

Differentialgleichungen

Jochen Merker

Sommersemester 2011

zuletzt aktualisiert am 14. September 2017

Inhaltsverzeichnis

1	Elementare Lösungsmethoden und allgemeine Existenzsätze	7
1.1	Lösungen gewöhnlicher Differentialgleichungen	7
1.2	Elementar lösbare Differentialgleichungen 1. Ordnung	11
1.3	Elementar lösbare Differentialgleichungen 2. Ordnung	20
1.4	Der Existenz- & Eindeutigkeitsatz v. Picard-Lindelöf	26
1.5	Allgemeinere Existenz- und Eindeutigkeitsresultate	31
1.6	Stetige Abhängigkeit	38
1.7	Differenzierbare Abhängigkeit	41
2	Lineare Differentialgleichungen	45
2.1	Lineare Systeme	45
2.2	Lineare Differentialgleichungen höherer Ordnung	48
2.3	Die Reduktionsmethode von d'Alembert	49
2.4	Lineare Differentialgleichungen mit konstanten Koeffizienten	51
3	Qualitative Theorie	59
3.1	Grundlegende Begriffe	59
3.2	Lineare Flüsse	61
3.3	Linearisierung nichtlinearer Flüsse	68
3.4	Lyapunov-Funktionen	70
4	Rand- und Eigenwertprobleme	79
4.1	Sturm-Liouvillesche Randwertprobleme	79
4.2	Sturm-Liouvillesche Eigenwertprobleme	83
5	Elementare Lösungsmethoden für partielle Differentialgleichungen	85
5.1	Die Laplace-Gleichung	85
5.2	Die Diffusionsgleichung	88
5.3	Die Wellengleichung	93
	Sporadisches Sachwortverzeichnis	99

Einleitung

In der Vorlesung „Differentialgleichungen“ werden wir uns mit Methoden beschäftigen, die es erlauben, gewöhnliche Differentialgleichungen explizit zu lösen oder zumindest die Existenz von Lösungen zu beweisen. Darüberhinaus werden wir Eigenschaften von Lösungen wie deren Eindeutigkeit, stetige Abhängigkeit oder qualitatives Verhalten diskutieren. Abgeschlossen wird die Vorlesung durch ein kurzes Kapitel über elementare Methoden zur Ermittlung von Lösungen spezieller partieller Differentialgleichungen.

Dieser Text ist dazu gedacht, Ihnen einen kurzen Überblick über die in der Vorlesung behandelten Themen zu geben. Darüberhinaus sind für interessierte und sehr begabte Studenten auch immer mal wieder weiterführende Literaturtipps angegeben. Lassen Sie sich aber nicht entmutigen, wenn Sie dort einen Blick hineinwerfen und nichts verstehen, das verlangt auch keiner von Ihnen.

Falls Sie Fragen, Anregungen oder Wünsche haben, sprechen Sie mich einfach an.

Viel Vergnügen und viel Erfolg beim Studium von Differentialgleichungen !

Kapitel 1

Elementare Lösungsmethoden und allgemeine Existenzsätze

In diesem Kapitel wird es darum gehen, einerseits spezielle Differentialgleichungen explizit zu lösen (was nicht ganz so hilfreich ist, wie es auf den ersten Blick erscheint, denn tatsächlich gibt es für die meisten interessanten Differentialgleichungen keine expliziten Lösungsmethoden) und andererseits Sätze zu beweisen, die die Lösbarkeit von allgemeinen Differentialgleichungen garantieren (was allein auch nicht sehr hilfreich ist, denn dadurch weiß man noch lange nicht, wie sich Lösungen verhalten). Darüberhinaus werden wir uns fragen, inwieweit die Lösungen einer Differentialgleichung eindeutig sind oder stetig/differenzierbar von den gegebenen Daten abhängen.

1.1 Lösungen gewöhnlicher Differentialgleichungen

Im Folgenden wird davon ausgegangen, dass Sie mit den grundlegenden Begriffen der Differentialrechnung für Funktionen einer reellen Variablen vertraut sind, wie sie z.B. in [Königsberger] entwickelt werden. Insbesondere sollte Ihnen bekannt sein, dass eine Abbildung $x : I \rightarrow \mathbb{R}^n$ auf einem Intervall $I \subset \mathbb{R}$ eine (parametrisierte) Kurve im \mathbb{R}^n genannt wird und differenzierbar in $t \in I$ heißt, falls der Grenzwert

$$x'(t) := \lim_{h \rightarrow 0} \frac{x(t+h) - x(t)}{h}$$

existiert, den man die (erste) Ableitung von x in t nennt (oder auch den Tangentialvektor an x zum Zeitpunkt t), siehe [Königsberger, 12.1]. Ist x in jedem Punkt $t \in I$ differenzierbar und die Ableitung $x' : I \rightarrow \mathbb{R}^n$ stetig, dann nennt man x stetig differenzierbar oder eine C^1 -Kurve. Rekursiv definiert man analog die m -te Ableitung $x^{(m)}$ und C^m -Kurven.

Gewöhnliche Differentialgleichungen sind nun Gleichungen, die nicht nur von den Werten sondern auch von den Ableitungen einer auf einem Intervall $I \subset \mathbb{R}$ definierten Funktion $x : I \rightarrow \mathbb{R}$ (im Fall $n = 1$) bzw. einer Kurve $x : I \rightarrow \mathbb{R}^n$ abhängen.

Definition 1.1. Ist $\Omega \subset \mathbb{R} \times (\mathbb{R}^n)^m$ und $f : \Omega \rightarrow \mathbb{R}^n$, so heißt die Gleichung

$$x^{(m)} = f(t, x, x', \dots, x^{(m-1)}) \tag{1.1}$$

eine (explizite¹) n -dimensionale Differentialgleichung m -ter Ordnung.

¹Noch allgemeiner als explizite Differentialgleichungen sind Differentialgleichungen von der Form $\tilde{f}(t, x, x', \dots, x^{(m-1)}, x^{(m)}) = 0$. Solche Differentialgleichungen werden implizit genannt und lassen sich nur dann in (1.1) umschreiben, wenn man nach $x^{(m)}$ auflösen kann. Wir werden uns in dieser Vorlesung nahezu ausschließlich mit expliziten Differentialgleichungen beschäftigen.

Man nennt (1.1) häufig auch ein n -dimensionales System von Differentialgleichungen m -ter Ordnung. Aufgrund des Vorkommens von Ableitung in (1.1) muss man natürlich präzise sagen, was man unter einer Lösung von (1.1) verstehen will.

Definition 1.2. Eine Kurve $x: I \rightarrow \mathbb{R}^n$ heißt eine (klassische) Lösung der Differentialgleichung (1.1), falls x eine m -mal differenzierbare Kurve mit $(t, x(t), x'(t), \dots, x^{(m-1)}(t)) \in \Omega$ und $x^{(m)}(t) = f(t, x(t), x'(t), \dots, x^{(m-1)}(t))$ für alle $t \in I$ ist.

Ist $f: \Omega \rightarrow \mathbb{R}^n$ stetig und x eine klassische Lösung der Differentialgleichung (1.1), dann ist offensichtlich x nicht nur m -mal differenzierbar, sondern sogar eine C^m -Kurve, da die rechte Seite von (1.1) stetig in t ist. Dies ist ein erster Hinweis darauf, dass klassische Lösungen von Differentialgleichungen oft viel regulärer sind als in Definition 1.2 gefordert wird.

Definition 1.1 ist recht allgemein, da an f überhaupt keine Anforderungen gestellt werden. Stellt man an f gewisse Bedingungen, dann benennt man auch den zugehörigen Typ von Differentialgleichungen entsprechend.

Definition 1.3. • Ist die Funktion f auf der rechten Seite von (1.1) unabhängig von t , so spricht man von einer **autonomen** Differentialgleichung.

- Ist $\Omega = I \times (\mathbb{R}^n)^m$ mit einem Intervall $I \subset \mathbb{R}$ und $(y_1, \dots, y_m) \mapsto f(t, y_1, \dots, y_m)$ für jedes $t \in I$ linear, dann nennt man (1.1) eine (homogene) **lineare** Differentialgleichung.

Als erstes Resultat wollen wir nun zeigen, dass man jede Differentialgleichung höherer Ordnung in eine äquivalente Differentialgleichung erster Ordnung umschreiben kann (die dafür aber eine höhere Dimension als die ursprüngliche Differentialgleichung hat).

Satz 1.4. Zu einer durch $f: \Omega \subset \mathbb{R} \times (\mathbb{R}^n)^m \rightarrow \mathbb{R}^n$ gegebenen Differentialgleichung m -ter Ordnung definiere man $\tilde{f}: \Omega \rightarrow (\mathbb{R}^n)^m$ durch

$$\tilde{f}(t, y_1, y_2, \dots, y_m) := (y_2, y_3, \dots, y_m, f(t, y_1, \dots, y_m)).$$

Dann ist x genau dann eine Lösung von $x^{(m)} = f(t, x, x', \dots, x^{(m-1)})$, wenn die durch $y(t) := (x(t), x'(t), \dots, x^{(m-1)}(t))$ definierte Kurve im $(\mathbb{R}^n)^m$ eine Lösung von $y' = \tilde{f}(t, y)$ ist.

Beweis: Ist x eine Lösung von $x^{(m)} = f(t, x, x', \dots, x^{(m-1)})$, dann ist y einmal differenzierbar und es gilt

$$\begin{aligned} y'_1 &= x' = y_2 \\ y'_2 &= x'' = y_3 \\ &\vdots \\ y'_m &= x^{(m)} = f(t, x, x', \dots, x^{(m-1)}) = f(t, y_1, y_2, \dots, y_m) \end{aligned}$$

also löst y die DGL $y' = \tilde{f}(t, y)$.

Ist umgekehrt y eine Lösung von $y' = \tilde{f}(t, y)$, dann hat die erste Komponente $x := y_1$ von y gerade y_k als $(k-1)$ -te Ableitung für $k = 2, \dots, m$, denn $y'_{k-1} = y_k$ gilt aufgrund der Definition der ersten $m-1$ Komponenten von \tilde{f} . Insbesondere ist $x^{(m-1)} = y_m$ differenzierbar und erfüllt aufgrund der Definition der letzten Komponente von \tilde{f} die Differentialgleichung

$$x^{(m)} = y'_m = f(t, y_1, y_2, \dots, y_m) = f(t, x, x', \dots, x^{(m-1)}).$$

□

Aufgrund von Satz (1.4) reicht es also aus, Systeme von Differentialgleichungen erster Ordnung zu betrachten, und diesen Spezialfall von Definition (1.1) wollen wir noch einmal festhalten.

Definition 1.5. Ist $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ und $f: \Omega \rightarrow \mathbb{R}^n$, so heißt die Gleichung

$$x' = f(t, x) \tag{1.2}$$

eine (explizite) n -dimensionale Differentialgleichung erster Ordnung oder auch ein n -dimensionales System von Differentialgleichungen erster Ordnung.

Die Abbildung f in (1.2) bezeichnet man auch als zeitabhängiges Vektorfeld auf \mathbb{R}^n . Tatsächlich gibt $f(t, x)$ gerade den Tangentialvektor an, den eine Lösungskurve haben soll, wenn sie zum Zeitpunkt t durch den Punkt x läuft.

Aufgabe 1.6. Zeigen Sie, dass autonome und nichtautonome Differentialgleichungen im folgenden Sinne äquivalent sind: Definiert man zu $f: \Omega \subset \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ das zeitunabhängige Vektorfeld $\tilde{f}(t, x) := (1, f(t, x))$ auf $\Omega \subset \mathbb{R}^{n+1}$, dann ist x genau dann eine Lösung von $x' = f(t, x)$ mit $x(t_0) = x_0$, wenn $y(s) := (s + t_0, x(s + t_0))$ eine Lösung von $y' = \tilde{f}(y)$ mit $y(0) = (t_0, x_0)$ ist. Es reicht also aus, autonome Systeme von Differentialgleichungen erster Ordnung zu betrachten.

Im Allgemeinen hat eine Differentialgleichung sehr viele verschiedene Lösungen, jedoch kann man hoffen, dass die Lösung einer Differentialgleichung erster Ordnung eindeutig ist, wenn man sich zusätzlich zur Differentialgleichung noch einen Anfangswert vorgibt.

Definition 1.7. Man sagt, eine Kurve x löst die Differentialgleichung (1.2) zum Anfangswert x_0 bei t_0 , wenn x eine Lösung von (1.2) ist und zusätzlich $x(t_0) = x_0$ gilt.

Beispiel 1.8. Zu vorgegebenem $\lambda \in \mathbb{R}$ hat die Differentialgleichung erster Ordnung $x' = \lambda x$ die Lösungen $x(t) = Ce^{\lambda t}$ mit einer beliebigen Konstanten $C \in \mathbb{R}$ (wie man durch Einsetzen in die Differentialgleichung leicht nachprüfen kann). Dies sind tatsächlich alle möglichen Lösungen der Differentialgleichung $x' = \lambda x$, denn löst x diese DGL, dann gilt auch $x'(t)e^{-\lambda t} = \lambda x(t)e^{-\lambda t}$ und daher

$$(xe^{-\lambda t})'(t) = x'(t)e^{-\lambda t} - \lambda x(t)e^{-\lambda t} = 0.$$

Also ist die Funktion $t \mapsto x(t)e^{-\lambda t}$ konstant, d.h. $x(t) = Ce^{\lambda t}$ gilt mit einer Konstanten $C \in \mathbb{R}$.

Nun wird die Konstante C aber durch die Anfangsbedingung $x(t_0) = x_0$ eindeutig festgelegt, denn aus dieser Bedingung ergibt sich $C = x_0 e^{-\lambda t_0}$. Somit gibt es genau eine Lösung der Differentialgleichung $x' = \lambda x$ zum Anfangswert x_0 bei t_0 , nämlich $x(t) = x_0 e^{\lambda(t-t_0)}$

Leider gibt es Differentialgleichungen mit stetiger rechter Seite f , deren Lösung durch eine Anfangsbedingung nicht eindeutig festgelegt wird.

Beispiel 1.9. Die autonome Differentialgleichung erster Ordnung $x' = \sqrt{|x|}$ hat die Lösungen $x(t) = \frac{(t+C)^2}{4}$ auf dem Intervall $[-C, \infty)$ (wie man wiederum durch Einsetzen in die Differentialgleichung leicht nachprüfen kann). Daher ist $x(t) := \frac{t^2}{4}$ auf dem Intervall $[0, \infty)$ eine Lösung zum Anfangswert $x_0 = 0$ bei $t = 0$, aber auch die konstante Lösung $\tilde{x}(t) := 0$ ist offensichtlich eine Lösung zum Anfangswert $x_0 = 0$ bei $t = 0$.

Ein anderes Problem ist, dass sich Lösungen von Differentialgleichungen selbst bei überall definierter und glatter rechter Seite $f: \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ im Allgemeinen nicht auf ganz \mathbb{R} fortsetzen lassen, d.h. zu einem Anfangswert muss es nicht unbedingt eine auf ganz \mathbb{R} definierte Lösung x geben. Man nennt eine Lösung x einer Differentialgleichung **global**, falls sie auf ganz \mathbb{R} definiert ist.

Beispiel 1.10. Die autonome Differentialgleichung $x' = x^2$ hat zum Anfangswert x_0 bei $t = 0$ die Lösung $x(t) = \frac{x_0}{1-tx_0}$. Diese Lösung ist sogar eindeutig, denn ist x eine Lösung mit $x(t) \neq 0$, dann gilt

$$\left(\frac{1+tx(t)}{x(t)}\right)' = \frac{(1+tx(t))x'(t) - x(t)(x(t)+tx'(t))}{(x(t))^2} = 0.$$

Jedoch lässt sich x bei $x_0 > 0$ nicht auf ein größeres Intervall als $(-\infty, \frac{1}{x_0})$ fortsetzen, da $\lim_{t \nearrow \frac{1}{x_0}} x(t) = +\infty$ gilt, und ähnlich für $x_0 < 0$. Daher gibt es zu keinem Anfangswert $x_0 \neq 0$ eine globale Lösung.

Treten bei einer autonomen Differentialgleichung diese beiden unschönen Situationen aber nicht auf, d.h. existieren für eine autonome Differentialgleichung erster Ordnung eindeutige globale Lösungen, dann ist der Lösungsoperator der Differentialgleichung ein Fluss, wie wir in Satz 1.13 zeigen werden.

Definition 1.11. Eine Abbildung $\Phi: \mathbb{R} \times \Omega \rightarrow \Omega$ heißt Fluss auf der Menge Ω , falls $\Phi(0, \cdot) = \text{Id}_\Omega$ und $\Phi(s, \Phi(t, x)) = \Phi(s+t, x)$ für alle $s, t \in \mathbb{R}$ und $x \in \Omega$ gilt.

Insbesondere gilt für einen Fluss Φ die Gleichung $\Phi(-t, \Phi(t, x)) = x$ für alle $t \in \mathbb{R}$ und $x \in \Omega$, d.h. $\Omega \ni x \mapsto \Phi(t, x) \in \Omega$ hat $y \mapsto \Phi(-t, y)$ als inverse Abbildung, und aufgrund der Gültigkeit von $\Phi(s, \Phi(t, x)) = \Phi(s+t, x)$ ist $t \mapsto \Phi(t, \cdot)$ ein Gruppenhomomorphismus von der abelschen Gruppe $(\mathbb{R}, +)$ in die Halbgruppe $(\text{Abb}(\Omega, \Omega), \circ)$ aller Abbildungen von Ω in sich mit der Komposition \circ als Verknüpfung.

Modellieren die Punkte von Ω den Zustand eines Systems (z.B. eines physikalischen oder biologischen Systems), so drücken die Eigenschaften, die ein Fluss erfüllt, gerade aus, dass die zeitliche Entwicklung des Systems deterministisch ist: Die gesamte zukünftige Entwicklung $t \mapsto \Phi(t, x)$ des Systems ist eindeutig durch den aktuellen Zustand $x \in \Omega$ bestimmt und unterliegt somit keinerlei zufälligem Einfluss.

Um zu beweisen, dass der Lösungsoperator einer autonomen Differentialgleichungen ein Fluss ist, ist die folgende Eigenschaft wesentlich, die im Allgemeinen nur für autonome Differentialgleichungen gilt.

Lemma 1.12. Ist x Lösung einer autonomen Differentialgleichung $x' = f(x)$, dann ist für jedes $t \in \mathbb{R}$ auch $s \mapsto x(s+t)$ eine Lösung der Differentialgleichung.

Beweis: Es gilt $\frac{d}{ds}(x(s+t)) = x'(s+t) = f(x(s+t))$. □

Satz 1.13. Besitzt die autonome Differentialgleichung erster Ordnung $x' = f(x)$ mit der auf der Teilmenge $\Omega \subset \mathbb{R}^n$ definierten rechten Seite $f: \Omega \rightarrow \mathbb{R}^n$ zu jedem Anfangswert $x_0 \in \Omega$ bei $t = 0$ eine eindeutige globale Lösung und bezeichnet man diese mit $t \mapsto \Phi(t, x_0)$, so ist Φ ein Fluss.

Beweis: Offensichtlich gilt aufgrund der Anfangsbedingung $\Phi(0, x) = x$ für alle $x \in \Omega$, und da zu gegebenem $t \in \mathbb{R}$ und $x \in \Omega$ sowohl $s \mapsto \Phi(s, \Phi(t, x))$ als auch (nach Lemma 1.12) $s \mapsto \Phi(t+s, x)$ Lösungen zum Anfangswert $\Phi(t, x)$ bei $s = 0$ sind, müssen diese aufgrund der Eindeutigkeit gleich sein, d.h. es gilt $\Phi(s, \Phi(t, x)) = \Phi(s+t, x)$ für alle $s, t \in \mathbb{R}$ und $x \in \Omega$. □

Umgekehrt kann man zu jedem Fluss Φ auf $\Omega \subset \mathbb{R}^n$, für den $t \mapsto \Phi(t, x)$ für jedes $x \in \Omega$ differenzierbar ist, durch $f(x) := \frac{\partial \Phi}{\partial t}(0, x)$ ein zeitunabhängiges Vektorfeld definieren, für das die

globale Lösung zum Anfangswert $x_0 \in \Omega$ bei $t = 0$ durch $t \mapsto \Phi(t, x_0)$ gegeben ist. In der Tat, aufgrund von $\Phi(s + t, x) = \Phi(s, \Phi(t, x))$ gilt

$$\frac{d}{dt}\Phi(t, x_0) = \frac{d}{ds}\Phi(s + t, x_0)|_{s=0} = \frac{d}{ds}\Phi(s, \Phi(t, x_0))|_{s=0} = f(\Phi(t, x_0)).$$

Bemerkung 1.14. *Diese bemerkenswerte Korrespondenz zwischen Flüssen und autonomen Differentialgleichungen wird genauer und allgemeiner in der Theorie (nicht)linearer Halbgruppen studiert, siehe [Engel, Nagel, Barbu].*

Im Folgenden wird es unser Anliegen sein, Bedingungen an die rechte Seite f einer Differentialgleichung aufzustellen, die die (globale) Existenz und Eindeutigkeit von Lösungen garantieren, denn nur dann kann man Satz (1.13) anwenden. Zunächst aber wollen wir für einige spezielle Typen von Differentialgleichungen Methoden vorstellen, mit denen man diese explizit lösen kann.

1.2 Elementar lösbare Differentialgleichungen 1. Ordnung

1.2.1 Lineare Differentialgleichungen 1. Ordnung

Eine (eindimensionale) lineare Differentialgleichung erster Ordnung hat die Form

$$x' = a(t)x + b(t) \tag{1.3}$$

mit vorgegebenen Funktionen $a, b : I \rightarrow \mathbb{R}$ auf einem Intervall I , d.h. die rechte Seite ist durch die affin-lineare Funktion $f : I \times \mathbb{R} \rightarrow \mathbb{R}$, $f(t, x) = a(t)x + b(t)$, gegeben. Im Fall $b = 0$ nennt man (1.3) homogen, und ist a von t unabhängig, dann spricht man von einer linearen Differentialgleichung mit konstanten Koeffizienten.

Beispiel 1.15. *Bezeichnet $x(t)$ die Menge (z.B. in Mol) eines radioaktiven Materials zum Zeitpunkt t , das mit der konstanten Rate a in andere Stoffe zerfällt, dann genügt x der homogenen linearen Differentialgleichung $x' = -ax$ mit konstanten Koeffizienten.*

Kommt aber ständig neues radioaktives Material mit der Rate $b(t)$ hinzu und kann man die Zerfallsrate beeinflussen, so dass sie durch eine zeitabhängige Funktion $a(t)$ gegeben ist, dann genügt x der allgemeinen linearen Differentialgleichung (1.3).

Betrachten wir zunächst den homogenen Fall $x' = a(t)x$ und setzen a als stetig voraus. Ist x auf einem $t_0 \in I$ enthaltenden offenen Intervall positiv, dann kann man dort auf beiden Seiten der Gleichung durch x teilen und erhält wegen $\frac{d}{dt} \ln(x(t)) = \frac{x'(t)}{x(t)}$ die Gleichung

$$\frac{d}{dt} \ln(x(t)) = a(t).$$

Integration über $[t_0, t]$ ergibt $\ln(x(t)) - \ln(x(t_0)) = \int_{t_0}^t a(s) ds$ und daher

$$x(t) = x_0 \exp\left(\int_{t_0}^t a(s) ds\right) \tag{1.4}$$

bei $x(t_0) = x_0$. Tatsächlich ist diese Funktion – wie man einfach durch Einsetzen überprüfen kann – für jede Wahl von (t_0, x_0) eine Lösung von $x' = a(t)x$ (auch wenn sie Null oder negativ wird).

Außerdem ist sie die einzige Lösung der Differentialgleichung zum Anfangswert $x(t_0) = x_0$ auf I , denn ist x eine Lösung von $x' = a(t)x$ auf I , so kann man ähnlich wie in Beispiel 1.8 mit $\exp\left(-\int_{t_0}^t a(s) ds\right)$ auf beiden Seiten multiplizieren und erhält mit der Produktregel

$$\left(x(t) \exp\left(-\int_{t_0}^t a(s) ds\right)\right)' = 0.$$

Demzufolge ist die Funktion $I \ni t \mapsto x(t) \exp\left(-\int_{t_0}^t a(s) ds\right)$ konstant, insbesondere gilt die Gleichung $x(t) = C \exp\left(-\int_{t_0}^t a(s) ds\right)$ mit einer Konstanten $C \in \mathbb{R}$ für alle $t \in I$, und durch Einsetzen von t_0 erhält man $C = x(t_0)$.

Um auch eine Lösung der inhomogenen linearen Differentialgleichung (1.3) zu erhalten, suchen wir mit Hilfe eines Tricks nach Lösungen, den man Variation der Konstanten nennt. Wir setzen nämlich einfach einmal den Ansatz $x(t) := C(t) \exp\left(\int_{t_0}^t a(s) ds\right)$ mit einer noch zu bestimmenden Funktion $C(\cdot)$ (statt einer Konstanten, deshalb auch der Name „Variation der Konstanten“) in (1.3) ein und erhalten mit Hilfe der Produktregel

$$C'(t) \exp\left(\int_{t_0}^t a(s) ds\right) + a(t)x(t) = a(t)x(t) + b(t).$$

Der Ansatz liefert also

$$C'(t) = \exp\left(-\int_{t_0}^t a(s) ds\right) b(t),$$

und durch Integration ergibt sich daher mit $C(t_0) = x_0$ sowie der Abkürzung $A(t) := \int_{t_0}^t a(s) ds$ für die Stammfunktion von a mit $A(t_0) = 0$ bei stetigem b

$$x(t) = \left(x_0 + \int_{t_0}^t \exp(-A(s))b(s) ds\right) \exp(A(t)) \tag{1.5}$$

als Lösung von (1.3). Wir halten dieses Ergebnis im folgenden Satz fest.

Satz 1.16 (allgemeine Lösung der linearen DGL erster Ordnung). *Sind $a, b: I \rightarrow \mathbb{R}$ stetige Funktionen auf einem Intervall I , ist $t_0 \in I$ und $x_0 \in \mathbb{R}$, so hat (1.3) unter der Anfangsbedingung $x(t_0) = x_0$ genau eine Lösung $x: I \rightarrow \mathbb{R}$, nämlich (1.5).*

Man bemerke, dass dieser Satz bei $I = \mathbb{R}$ sogar die Existenz und Eindeutigkeit globaler Lösungen garantiert.

Aufgabe 1.17. *Weisen Sie auch im inhomogenen Fall die Eindeutigkeit der Lösung von (1.3) nach.*

1.2.2 Trennung der Variablen

Hat eine (eindimensionale) lineare Differentialgleichung erster Ordnung die Form

$$x' = g(t)h(x) \tag{1.6}$$

mit vorgegebenen Funktionen $g: I \rightarrow \mathbb{R}$ und $h: J \rightarrow \mathbb{R}$ auf (offenen) Intervallen $I, J \subset \mathbb{R}$, so spricht man von einer Differentialgleichung mit getrennten Variablen. Beispielsweise ist jede autonome

Differentialgleichung $x' = f(x)$ oder auch jede homogene lineare Differentialgleichung $x' = a(t)x$ eine Differentialgleichung mit getrennten Variablen.

Für Differentialgleichungen mit getrennten Variablen gibt es eine relativ simple Methode (die Trennung der Variablen genannt wird), mit der man das zugehörige Anfangswertproblem zum Anfangswert $x(t_0) = x_0$ bei $(t_0, x_0) \in I \times J$ lokal lösen kann. Sind nämlich g und h stetig und ist $h(x_0) \neq 0$ (andernfalls ist offensichtlich die konstante Funktion $t \mapsto x_0$ auf I eine Lösung), so gibt es aufgrund der Stetigkeit von h zumindest ein offenes Intervall $\tilde{J} \subset J$ mit $x_0 \in \tilde{J}$ und $h(x) \neq 0$ für $x \in \tilde{J}$. Somit kann man die Stammfunktionen

$$G : I \rightarrow \mathbb{R} \quad , \quad G(t) := \int_{t_0}^t g(s) ds \quad \text{und}$$

$$H : \tilde{J} \rightarrow \mathbb{R} \quad , \quad H(x) := \int_{x_0}^x \frac{1}{h(y)} dy$$

mit $G(t_0) = 0$ und $H(x_0) = 0$ definieren. Nun ist aber $H' = \frac{1}{h}$ entweder auf ganz \tilde{J} positiv oder negativ, d.h. H ist streng monoton. Also besitzt H eine Umkehrfunktion $H^{-1} : H(\tilde{J}) \rightarrow \tilde{J}$, und $H(\tilde{J})$ ist ein offenes Intervall um $H(x_0) = 0$. Aufgrund der Stetigkeit von G und $G(t_0) = 0$ gibt es dann ein offenes Intervall $\tilde{I} \subset I$ um t_0 mit $G(\tilde{I}) \subset H(\tilde{J})$. Auf diesem Intervall \tilde{I} definieren wir nun $x(t) := H^{-1}(G(t))$, d.h. $x(t)$ erhält man durch Auflösen der Gleichung $H(x(t)) = G(t)$. Dann gilt $x(t_0) = H^{-1}(G(t_0)) = H^{-1}(0) = x_0$ und nach den Differentiationsregeln

$$x'(t) = (H^{-1})'(G(t)) \cdot G'(t) = \frac{1}{H'(H^{-1}(G(t)))} \cdot G'(t) = \frac{1}{H'(x(t))} \cdot G'(t) = h(x(t)) \cdot g(t)$$

aufgrund von $H' = \frac{1}{h}$ und $G' = g$. Also ist x auf \tilde{I} eine Lösung von (1.6) zum Anfangswert $x(t_0) = x_0$. Wir haben somit den folgenden Satz bewiesen.

Satz 1.18. *Seien $g : I \rightarrow \mathbb{R}$ und $h : J \rightarrow \mathbb{R}$ stetige Funktionen auf offenen Intervallen $I, J \subset \mathbb{R}$ und sei $(t_0, x_0) \in I \times J$ vorgegeben.*

- *Ist $h(x_0) = 0$, dann hat (1.6) zum Anfangswert $x(t_0) = x_0$ die konstante Lösung $x(t) = x_0$ auf ganz I .*
- *Ist $h(x_0) \neq 0$, dann besitzt (1.6) zum Anfangswert $x(t_0) = x_0$ auf einem hinreichend kleinen offenen Intervall $\tilde{I} \subset I$ um t_0 eine Lösung $x : \tilde{I} \rightarrow \mathbb{R}$, und diese erhält man durch Auflösen von*

$$\int_{x_0}^{x(t)} \frac{1}{h(y)} dy = \int_{t_0}^t g(s) ds \tag{1.7}$$

nach $x(t)$.

Bemerkung 1.19.

- *Satz 1.18 garantiert nur die Existenz lokaler Lösungen des Anfangswertproblems zu (1.6), aber nicht die Existenz von globalen Lösungen. Tatsächlich müssen i.a. keine globalen Lösungen von (1.6) existieren, wie Beispiel (1.10) zeigt.*
- *Aus dem Beweis von Satz 1.18 kann man desweiteren ablesen, dass Lösungen zu Anfangswerten x_0 mit $h(x_0) \neq 0$ zumindest auf dem gewonnenen hinreichend kleinen Intervall \tilde{I} eindeutig sind (schließlich existiert die Umkehrabbildung von H). Jedoch muss dies für größere Intervalle nicht mehr der Fall sein, wie auch Beispiel (1.9) zeigt.*

- Im Nachhinein erklärt Satz 1.18, warum wir im vorigen Abschnitt zur Lösung der linearen Differentialgleichung $x' = a(t)x$ durch x geteilt und dann integriert haben, dort haben wir nämlich gerade Trennung der Variablen benutzt.

Beispiel 1.20.

- (a) Betrachtet man die Population einer Spezies (z.B. Bakterien in einer Nährlösung) und sieht man die Anzahl $x(t)$ der Individuen zur Zeit $t \in I$ als kontinuierlich an (was nur bei sehr großen Populationen Sinn macht), so kann man die zeitliche Entwicklung der Population durch eine reelle Funktion $x: I \rightarrow \mathbb{R}$ modellieren. Den Quotienten $\frac{x'}{x}$ bezeichnet man als relative Änderungsrate. Diese relative Änderungsrate modelliert man häufig durch eine allein von der Zeit t und der Populationsgröße x abhängige Funktion $r(t, x)$ (die z.B. die Differenz der Geburten- und Sterberate zur Zeit t bei Populationsgröße x angibt). In diesem Modell erfüllt x also die Differentialgleichung

$$x' = r(t, x)x. \quad (1.8)$$

- (b) Das einfachste Modell (1.8) ist sicherlich durch die Wahl $r(t, x) = \lambda$ mit einer Konstanten $\lambda \in \mathbb{R}$ gegeben, und wie wir schon aus Beispiel 1.8 wissen, haben die Lösungen die Form $Ce^{\lambda t}$. Hier wird man nur nichtnegative Populationsgrößen zulassen, und tatsächlich folgt aus $C > 0$ auch $x(t) > 0$. In Abhängigkeit von λ wächst die Population $x(t)$ für $t \rightarrow \infty$ entweder ins Unendliche ($\lambda > 0$), bleibt konstant ($\lambda = 0$) oder geht gegen Null, d.h. die Population stirbt aus ($\lambda < 0$).

- (c) Ein etwas komplexeres Modell (1.8) ist durch eine lineare relative Änderungsrate $r(t, x) = \lambda(1 - \frac{x}{x_{\max}})$ mit der maximalen Wachstumsrate $\lambda > 0$ und der maximal tragbaren Populationsgröße $x_{\max} > 0$ gegeben. Überschreitet x die maximal tragbare Populationsgröße x_{\max} , so wird r negativ und die Population sinkt. Neben der konstanten Lösung 0 gibt es noch eine weitere konstante Lösung, nämlich $x = x_{\max}$. Diese Lösung ist eine sogenannte global asymptotisch stabile Ruhelage, d.h. sie hängt nicht von t ab und alle (positiven) Lösungen streben für $t \rightarrow \infty$ gegen x_{\max} . Tatsächlich, Trennung der Variablen liefert für $x' = \lambda(1 - \frac{x}{x_{\max}})x$ bei $x(t_0) = x_0$ die Gleichung

$$\int_{x_0}^{x(t)} \frac{1}{(y - x_{\max})y} dy = - \int_{t_0}^t \frac{\lambda}{x_{\max}} ds.$$

Mittels der Partialbruchzerlegung $\frac{1}{(y-x_{\max})y} = \frac{1}{x_{\max}} \left(\frac{1}{y-x_{\max}} - \frac{1}{y} \right)$ ergibt sich

$$\frac{1}{x_{\max}} \ln \left| \frac{y - x_{\max}}{y} \right| \Big|_{x_0}^{x(t)} = - \frac{\lambda}{x_{\max}} (t - t_0)$$

und somit $\frac{x_{\max} - x(t)}{x(t)} = \frac{x_{\max} - x_0}{x_0} e^{-\lambda(t-t_0)}$ bei $x_0 \in (0, x_{\max})$ und $\frac{x(t) - x_{\max}}{x(t)} = \frac{x_0 - x_{\max}}{x_0} e^{-\lambda(t-t_0)}$ bei $x_0 \in (x_{\max}, \infty)$. Die Lösung zum Anfangswert $x(t_0) = x_0 \geq 0$ ist also

$$x(t) = \begin{cases} 0 & x_0 = 0 \\ \frac{x_{\max}}{1 + \frac{x_{\max} - x_0}{x_0} e^{-\lambda(t-t_0)}} & 0 < x_0 < x_{\max} \\ x_{\max} & x_0 = x_{\max} \\ \frac{x_{\max}}{1 - \frac{x_0 - x_{\max}}{x_0} e^{-\lambda(t-t_0)}} & x_{\max} < x_0 \end{cases}$$

und strebt für $x_0 > 0$ wie behauptet bei $t \rightarrow \infty$ gegen x_{\max} .

1.2.3 Substitution

Manchmal kann man durch Substitution eine kompliziert aussehende Differentialgleichung in eine schon bekannte Differentialgleichung überführen. Aus einer Lösung dieser bekannten Gleichung erhält man dann durch Rücksubstitution eine Lösung der ursprünglichen Differentialgleichung. Wir wollen diese Methode zur Lösung von Differentialgleichungen anhand dreier Beispiele kennenlernen.

Lineare Substitution

Hat eine Differentialgleichung die Form

$$x' = f(ax + bt + c) \quad (1.9)$$

mit einer Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ und Konstanten $a, b, c \in \mathbb{R}$, $a \neq 0$, dann kann man zu einer Lösung x die Funktion $y(t) := ax(t) + bt + c$ betrachten. Diese erfüllt

$$y'(t) = ax'(t) + b = af(ax + bt + c) + b = af(y) + b,$$

löst also die autonome Differentialgleichung $y' = af(y) + b$. Umgekehrt erhält man aus einer Lösung y von $y' = af(y) + b$ durch $x(t) := \frac{1}{a}(y(t) - bt - c)$ aber auch eine Lösung der ursprünglichen Differentialgleichung (1.9). Da autonome Differentialgleichungen mittels Trennung der Variablen oft leicht lösbar sind, ist es uns also durch Substitution des linearen Terms $ax + bt + c$ gelungen, die kompliziert aussehende Differentialgleichung (1.9) in eine einfach zu lösende Differentialgleichung zu überführen.

Beispiel 1.21. *Substituiert man in der Differentialgleichung $x' = \frac{1}{(t-x)^2+1}$ den linearen Term $y := -x + t$, dann erhält man für y die Differentialgleichung*

$$y' = -\frac{1}{y^2+1} + 1.$$

Die rechte Seite ist gleich $\frac{y^2}{y^2+1}$, daher erhält man mittels Trennung der Variablen

$$\int_{y_0}^{y(t)} 1 + \frac{1}{z^2} dz = \int_{y_0}^{y(t)} \frac{z^2+1}{z^2} dz = \int_{t_0}^t 1 ds,$$

und daraus $y(t) - \frac{1}{y(t)} - y_0 + \frac{1}{y_0} = (t - t_0)$ oder

$$y(t)^2 - \left(y_0 - \frac{1}{y_0} + (t - t_0) \right) y(t) - 1 = 0.$$

Die Lösung von $y' = -\frac{1}{y^2+1} + 1$ zum Anfangswert $y(t_0) = y_0$ ist also

$$y(t) = \frac{1}{2} \left(y_0 - \frac{1}{y_0} + (t - t_0) \right) \pm \sqrt{\frac{1}{4} \left(y_0 - \frac{1}{y_0} + (t - t_0) \right)^2 + 1},$$

und die Lösung der ursprünglichen Gleichung zum Anfangswert $x(t_0) = x_0$ ist

$$x(t) = t - \frac{1}{2} \left(t - x_0 - \frac{1}{t_0 - x_0} \right) \mp \sqrt{\frac{1}{4} \left(t - x_0 - \frac{1}{t_0 - x_0} \right)^2 + 1},$$

wegen $y_0 = -x_0 + t_0$. Man bemerke, dass man aufgrund der quadratischen Abhängigkeit statt $-x + t$ ebensogut $x - t$ hätte substituieren können. Außerdem hat man bei $x_0 = t_0$ nur scheinbar ein Problem mit der Existenz der Lösung: Die Lösung der DGL für y zum Anfangswert $y(t_0) = 0$ ist die konstante Lösung $y = 0$, und zu dieser gehört die Lösung $x = t$ zum Anfangswert $x_0 = t_0$.

Ähnlichkeitsdifferentialgleichungen

Ähnlichkeitsdifferentialgleichungen haben die Form

$$x' = f\left(\frac{ax + bt + e_1}{cx + dt + e_2}\right) \quad (1.10)$$

mit einer invertierbaren Matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ und einem Vektor $e \in \mathbb{R}^2$. Im Spezialfall $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

und $e = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ nennt man die Differentialgleichung (1.10) eine homogene Differentialgleichung, sie hat dann die Form

$$x' = f\left(\frac{x}{t}\right). \quad (1.11)$$

Diese Bezeichnung macht einerseits Sinn, denn die Steigung x' einer Lösung ist bei einer homogenen Differentialgleichung auf allen Ursprungsgeraden $x = mt$ dieselbe. Insbesondere ist mit einer Lösung x auch $y(t) := \frac{x(st)}{s}$ für jedes $s \neq 0$ eine Lösung, denn

$$y'(t) = x'(st) = f\left(\frac{x(st)}{st}\right) = f\left(\frac{y(t)}{t}\right)$$

Andererseits ist die Bezeichnung „homogen“ ungünstig gewählt, da bereits lineare Differentialgleichungen (1.3) bei $b = 0$ als homogen bezeichnet werden.

Den allgemeinen Fall einer Ähnlichkeitsdifferentialgleichung kann man durch Koordinatenwechsel immer auf den einfacheren Fall einer homogenen Differentialgleichungen zurückführen. In der Tat, da $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ invertierbar ist, hat das lineare Gleichungssystem $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_0 \\ t_0 \end{pmatrix} = - \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}$ eine eindeutige Lösung (x_0, t_0) , und in den verschobenen Koordinaten $s := t - t_0$, $y(s) := x(s + t_0) - x_0$, lautet die Differentialgleichung

$$\begin{aligned} \frac{d}{ds}y(s) &= \frac{d}{ds}x(s + t_0) = x'(s + t_0) = f\left(\frac{ax(s + t_0) + b(s + t_0) + e_1}{cx(s + t_0) + d(s + t_0) + e_2}\right) \\ &= f\left(\frac{ay(s) + bs + ax_0 + bt_0 + e_1}{cy(s) + ds + cx_0 + dt_0 + e_2}\right) \stackrel{\text{LGS}}{=} f\left(\frac{ay(s) + bs}{cy(s) + ds}\right) = f\left(\frac{a\frac{y(s)}{s} + b}{c\frac{y(s)}{s} + d}\right), \end{aligned}$$

ist also eine homogene Differentialgleichung (wobei die neue, nur von $\frac{y}{s}$ abhängige rechte Seite von (1.11) durch $\tilde{f}(z) = f\left(\frac{az+b}{cz+d}\right)$ gegeben ist). Die homogene Differentialgleichung (1.11) kann man aber durch die Substitution $y := \frac{x}{t}$ lösen, denn unter dieser Substitution geht sie in die Differentialgleichung

$$y' = \frac{f(y) - y}{t}$$

über, welche man mittels Trennung der Variablen lösen kann.

Beispiel 1.22. *Homogene Differentialgleichungen treten insbesondere dann auf, wenn die Differentialgleichung nicht von der Wahl der Maßeinheiten abhängen soll. Modelliert beispielsweise t den Preis von Benzin in Euro und $x(t)$ die Nachfrage beim Preis t in Litern, so soll sich das durch eine Differentialgleichung gegebene Modell für die Preis-Nachfrage-Relation bei Umrechnung des Preises in Dollar und der Nachfrage in Gallonen nicht ändern. Genau dies ist bei der speziellen homogenen Differentialgleichung*

$$x' = \lambda \frac{x}{t}$$

der Fall, mit $x(t)$ ist für $a, b \neq 0$ auch $y(t) := ax(bt)$ eine Lösung, denn

$$y'(t) = \lambda abx'(bt) = \lambda ab \frac{x(bt)}{bt} = \lambda \frac{y(t)}{t}.$$

Die Differentialgleichung $x' = \lambda \frac{x}{t}$ ist aber sogar linear und kann leicht mittels Trennung der Variablen gelöst werden, aus

$$\ln(x(t)) - \ln(x_0) = \int_{x_0}^{x(t)} \frac{1}{y} dy = \lambda \int_{t_0}^t \frac{1}{s} ds = \lambda(\ln(t) - \ln(t_0))$$

ergibt sich $x(t) = x_0 \left(\frac{t}{t_0}\right)^\lambda$ für $\lambda, t_0 \neq 0$.

Bernoulli- und Riccati-Differentialgleichungen

Eine Differentialgleichung der Form

$$x' + a(t)x + b(t)x^p = 0 \tag{1.12}$$

heißt bei $p \neq 1$ Bernoulli-Differentialgleichung. Die Substitution $y := x^{1-p}$ führt wegen

$$y' = (1-p)x^{-p}x' = (p-1)x^{-p}(a(t)x + b(t)x^p) = (p-1)a(t)y + (p-1)b(t)$$

auf eine lineare Differentialgleichung, die i.a. inhomogen ist und nicht-konstante Koeffizienten besitzt. Formel (1.5) erlaubt einem prinzipiell, diese lineare Differentialgleichung zu lösen, und die Rücksubstitution $x = y^{1/(1-p)}$ liefert daraus Lösungen der Bernoulli-Differentialgleichung. Man beachte, dass der geschilderte Lösungsweg rigoros nur für positive Lösungen Sinn macht. Für negative Lösungen sollte man Potenzen x^p als $|x|^p$ und x^{1-p} bzw. $y^{1/(1-p)}$ als vorzeichenbehaftete Potenz $\text{sign}(x)|x|^{1-p}$ bzw. $\text{sign}(y)|y|^{1/(1-p)}$ interpretieren, sonst kommt es zu Uneindeutigkeiten.

Ist in einer Bernoulli-Differentialgleichung $p = 2$ und die rechte Seite nicht Null, sondern eine von t abhängige Funktion $c(t)$, d.h. hat die Differentialgleichung die Form

$$x' + a(t)x + b(t)x^2 = c(t) \tag{1.13}$$

dann spricht man von einer Riccati-Differentialgleichung. Solche Differentialgleichungen kann man elementar lösen, falls man schon eine partikuläre Lösung x_p kennt (z.B. geraten hat). Dann geht die Riccati-Differentialgleichung (1.13) durch die Substitution $y = x - x_p$ wegen

$$\begin{aligned} y' &= x' - x'_p = \left(c(t) - a(t)x - b(t)x^2\right) - \left(c(t) - a(t)x_p - b(t)x_p^2\right) \\ &= -a(t)(x - x_p) - b(t)(x^2 - x_p^2) = -a(t)(x - x_p) - b(t)(x - x_p)\left((x - x_p) + 2x_p\right) \\ &= -\left(a(t) + 2b(t)x_p(t)\right)(x - x_p) - b(t)(x - x_p)^2 = -\left(a(t) + 2b(t)x_p(t)\right)y - b(t)y^2 \end{aligned}$$

in die Bernoulli-Differentialgleichung $y' + (a(t) + 2b(t)x_p(t))y + b(t)y^2 = 0$ über. Diese kann man wie zuvor geschildert lösen und erhält durch anschließende Rücksubstitution $x = y + x_p$ Lösungen der Riccati-Differentialgleichung.

Beispiel 1.23. Die Riccati-Differentialgleichung $x' + (1 - 2t^2)x + tx^2 = t - t^3 + 1$ hat die spezielle Lösung $x_p(t) = t$. Die Differenz $y := x - x_p$ erfüllt die Bernoulli-Differentialgleichung

$$y' + ((1 - 2t^2) + 2t \cdot t)y + ty^2 = y' + y + ty^2 = 0 \quad .$$

Mit der Substitution $z = 1/y$ ergibt sich die lineare Differentialgleichung $z' = z + t$ mit allgemeiner Lösung $z(t) = Ce^t - (t + 1)$. Zweimalige Rücksubstitution ergibt insgesamt $x(t) = \frac{1}{Ce^t - (t+1)} + t$ als allgemeine Lösung der obigen Riccati-Differentialgleichung.

1.2.4 Exakte Differentialgleichungen

Ist $\Omega \subset \mathbb{R} \times \mathbb{R}$ offen und $f: \Omega \rightarrow \mathbb{R}$ stetig, so überdecken die Lösungen von $x' = f(t, x)$ ganz Ω ². Mit anderen Worten bilden die Lösungen also eine Schar von ebenen Kurven, die ganz Ω überdecken. Solch eine Schar von Kurven kann man häufig besser implizit durch eine Gleichung der Form $\Phi(t, x) = C$ mit einer differenzierbaren Funktion $\Phi: \Omega \rightarrow \mathbb{R}$ und einer frei wählbaren Konstanten C angeben statt explizit in der Form $x = x(t; C)$ einer allgemeinen Lösung.

Ist $(t, x(t))$ eine durch t parametrisierte Lösung von $\Phi(t, x) = C$, dann ergibt die Differentiation von $\Phi(t, x(t)) = C$ nach t unter Beachtung der Kettenregel die Differentialgleichung

$$\frac{\partial \Phi}{\partial t}(t, x) + \frac{\partial \Phi}{\partial x}(t, x)x' = 0$$

Daher bezeichnet man eine Differentialgleichung der Form

$$g(t, x) + h(t, x)x' = 0 \text{ oder symbolisch } g(t, x) dt + h(t, x) dx = 0 \quad (1.14)$$

mit zwei Funktionen $g, h: \Omega \rightarrow \mathbb{R}$ als exakte Differentialgleichung, wenn es eine Funktion $\Phi: \Omega \rightarrow \mathbb{R}$ mit $g = \frac{\partial \Phi}{\partial t}$ und $h = \frac{\partial \Phi}{\partial x}$ gibt. In diesem Fall sind die Lösungen der Differentialgleichung (1.14) implizit durch $\Phi(t, x) = C$ gegeben. Der folgende Satz stellt ein leicht zu überprüfendes Kriterium für die Exaktheit der Differentialgleichung (1.14) bereit.

Satz 1.24. *Ist $\Omega \subset \mathbb{R} \times \mathbb{R}$ ein einfach zusammenhängendes Gebiet (d.h. Ω ist offen, zusammenhängend und hat keine „Löcher“) und sind die Funktionen $g, h: \Omega \rightarrow \mathbb{R}$ stetig differenzierbar, dann ist die Differentialgleichung (1.14) genau dann exakt, wenn $\frac{\partial g}{\partial x} = \frac{\partial h}{\partial t}$ gilt.*

Beweis: Gibt es eine Funktion $\Phi: \Omega \rightarrow \mathbb{R}$ mit $g = \frac{\partial \Phi}{\partial t}$ und $h = \frac{\partial \Phi}{\partial x}$, dann ist aufgrund der stetigen Differenzierbarkeit von g und h die Funktion Φ zweimal stetig differenzierbar, also stimmen nach dem Satz von Schwarz die gemischten Ableitungen $\frac{\partial g}{\partial x} = \frac{\partial^2 \Phi}{\partial x \partial t}$ und $\frac{\partial h}{\partial t} = \frac{\partial^2 \Phi}{\partial t \partial x}$ überein und die eine Richtung des Satzes ist gezeigt. Für den Beweis der anderen Richtung verweisen wir auf [Königsberger, 11.3, 11.4]. \square

Für ein einfach zusammenhängendes Gebiet $\Omega \subset \mathbb{R} \times \mathbb{R}$ kann man also zur Lösung von (1.14) die folgenden Schritte durchführen:

- (a) Prüfe, ob $\frac{\partial g}{\partial x} = \frac{\partial h}{\partial t}$ gilt; wenn nicht, dann ist die Differentialgleichung nicht exakt und man muss nach einer anderen Lösungsmethode suchen (z.B. einen integrierenden Faktor finden).
- (b) Wenn die Differentialgleichung exakt ist, dann bestimme man durch Integration von g nach t (bei festgehaltenem x) eine Funktion $\Psi: \Omega \rightarrow \mathbb{R}$ mit $g = \frac{\partial \Psi}{\partial t}$.
- (c) Setzt man danach den Ansatz $\Phi(t, x) := \Psi(t, x) + \Gamma(x)$ in die Gleichung $h = \frac{\partial \Phi}{\partial x}$ ein, so erhält man $\Gamma'(x) = h(t, x) - \frac{\partial \Psi}{\partial x}(t, x)$. Die Funktion auf der rechten Seite hängt dann gar nicht von t ab (ansonsten hat man sich verrechnet), also liefert Integration nach x eine nur von x abhängige Funktion Γ , und $\Phi(t, x) := \Psi(t, x) + \Gamma(x)$ ist die gesuchte Funktion mit $g = \frac{\partial \Phi}{\partial t}$ und $h = \frac{\partial \Phi}{\partial x}$.
- (d) Zu einem Anfangswert $(t_0, x_0) \in \Omega$ ist die Lösung von (1.14) dann implizit durch $\Phi(t, x(t)) = C$ gegeben, wobei $C = \Phi(t_0, x_0)$ ist.

²Denn nach dem später formulierten Satz 1.43 von Peano existiert durch jeden Punkt $(t_0, x_0) \in \Omega$ eine nicht weiter fortsetzbare Lösung $x: I_{\max} \rightarrow \mathbb{R}$, die sich an den endlichen Randpunkten ihres offenen Existenzintervalls I_{\max} dem Rand von Ω beliebig nähert (oder gegen $\pm\infty$ geht).

Natürlich könnte man im zweiten und dritten Schritt auch erst h nach x (bei festgehaltenem t) integrieren, um ein Ψ mit $h = \frac{\partial \Psi}{\partial x}$ zu erhalten, und dann durch Einsetzen von $\Phi(t, x) = \Psi(t, x) + \Gamma(t)$ in $g = \frac{\partial \Phi}{\partial t}$ eine nur von t abhängige Funktion Γ zu erhalten.

Integrierender Faktor

Ist die Differentialgleichung (1.14) nicht exakt, so kann man manchmal durch Multiplikation mit einer Funktion $m : \Omega \rightarrow \mathbb{R}$, $m \neq 0$, aus (1.14) eine exakte Differentialgleichung machen. In diesem Fall nennt man m einen integrierenden Faktor (oder Eulerschen Multiplikator) der Differentialgleichung (1.14). Dazu beobachtet man, dass

$$m(t, x)g(t, x) + m(t, x)h(t, x)x' = 0$$

für ein einfach zusammenhängendes Gebiet $\Omega \subset \mathbb{R} \times \mathbb{R}$ genau dann exakt ist, wenn

$$m \frac{\partial g}{\partial x} + \frac{\partial m}{\partial x} g = m \frac{\partial h}{\partial t} + \frac{\partial m}{\partial t} h$$

gilt. Hängt insbesondere $\frac{\frac{\partial g}{\partial x} - \frac{\partial h}{\partial t}}{h}$ bei $h \neq 0$ nur von t ab, dann kann man aus

$$\frac{\frac{\partial g}{\partial x} - \frac{\partial h}{\partial t}}{h} = \frac{\frac{\partial m}{\partial t}}{m}$$

einen nur von t abhängigen integrierenden Faktor $m = m(t)$ bestimmen (hängt analog $\frac{\frac{\partial h}{\partial t} - \frac{\partial g}{\partial x}}{g}$ bei $g \neq 0$ nur von x ab, dann existiert ein nur von x abhängiger integrierender Faktor $m = m(x)$).

Beispiel 1.25. Die Differentialgleichung $(2t^2 + 2tx^2 + 1)x + (3x^2 + t)x' = 0$ auf $\Omega = \mathbb{R} \times \mathbb{R}$ ist wegen

$$\frac{\partial g}{\partial x} = 6tx^2 + 2t^2 + 1 \neq 1 = \frac{\partial h}{\partial t}$$

nicht exakt, aber

$$\frac{\frac{\partial g}{\partial x} - \frac{\partial h}{\partial t}}{h} = \frac{6tx^2 + 2t^2}{3x^2 + t} = 2t$$

hängt nur von t ab. Aus $2t = \frac{\frac{\partial m}{\partial t}}{m} = (\log |m|)'(t)$ ergibt sich der nur von t abhängige integrierende Faktor $m(t) = e^{t^2}$, und durch Multiplikation mit diesem geht die Differentialgleichung $(2t^2 + 2tx^2 + 1)x + (3x^2 + t)x' = 0$ in die exakte Differentialgleichung

$$e^{t^2}(2t^2 + 2tx^2 + 1)x + e^{t^2}(3x^2 + t)x' = 0$$

über. Integration von $e^{t^2}(3x^2 + t)$ nach x liefert $\Psi(t, x) = e^{t^2}(x^3 + tx)$, und aus $\Phi(t, x) = \Psi(t, x) + \Gamma(t)$ ergibt sich durch Einsetzen in $g = \frac{\partial \Phi}{\partial t}$ die Gleichung

$$e^{t^2}(2t^2 + 2tx^2 + 1)x = 2te^{t^2}x(x^2 + t) + e^{t^2}x + \Gamma'(t)$$

also $\Gamma'(t) = 0$ und daher z.B. $\Phi(t, x) = e^{t^2}x(x^2 + t)$. Die Lösungen von $(2t^2 + 2tx^2 + 1)x + (3x^2 + t)x' = 0$ sind also implizit durch $e^{t^2}x(x^2 + t) = C$ gegeben, eine Lösung zum Anfangswert $x(-1) = 1$ ist

also beispielsweise implizit durch $e^{t^2}x(x^2 + t) = 0$ oder explizit durch $x(t) = \begin{cases} \sqrt{|t|} & t \leq 0 \\ 0 & t \geq 0 \end{cases}$ gegeben.

1.3 Elementar lösbare Differentialgleichungen 2. Ordnung

1.3.1 Spezielle homogene lineare Differentialgleichungen 2. Ordnung

Lineare Schwingungsgleichungen mit Dämpfung

Eine homogene lineare Differentialgleichung zweiter Ordnung

$$x'' + 2\mu x' + \omega_0^2 x = 0 \quad (1.15)$$

mit konstanten Koeffizienten $\mu \geq 0$, $\omega_0 > 0$, nennt man auch eine lineare Schwingungsgleichung mit Dämpfung. Der Ansatz $x(t) = e^{\lambda t}$ führt auf

$$(\lambda^2 + 2\mu\lambda + \omega_0^2)e^{\lambda t} = 0$$

und somit auf $\lambda = -\mu \pm \sqrt{\mu^2 - \omega_0^2}$ (wobei der Term unter der Wurzel negativ und dann λ komplex sein kann). Im Fall $\mu \neq \omega_0$ sind also $e^{(-\mu + \sqrt{\mu^2 - \omega_0^2})t}$ und $e^{(-\mu - \sqrt{\mu^2 - \omega_0^2})t}$ (i.a. komplexe) linear unabhängige Lösungen, im Fall $\mu = \omega_0$ sind $e^{-\mu t}$ und $te^{-\mu t}$ linear unabhängige Lösungen, wie man durch Einsetzen in (1.15) und Berechnung der später eingeführten Wronski-Determinante leicht überprüfen kann.

Später werden wir solch ein System von linear unabhängigen Lösungen Fundamentalsystem nennen und zeigen, dass sich jede Lösung von (1.15) als Linearkombination der Lösungen eines Fundamentalsystems schreiben läßt. Alle Lösungen von (1.15) sind also von der Form

$$x(t) = c_1 e^{(-\mu + \sqrt{\mu^2 - \omega_0^2})t} + c_2 e^{(-\mu - \sqrt{\mu^2 - \omega_0^2})t} \text{ bei } \mu \neq \omega_0 \text{ bzw.} \\ x(t) = c_1 e^{-\mu t} + c_2 t e^{-\mu t} \text{ bei } \mu = \omega_0.$$

Wir wollen nun noch reelle Lösungen zu diesen i.a. komplexen Lösungen gewinnen und ihr langfristiges Verhalten diskutieren. Dazu unterscheiden wir vier Fälle.

- Ungedämpfte Schwingung ($\mu = 0$): In diesem Fall lautet die allgemeine reelle Lösung $x(t) = c_1 \cos(\omega_0 t) + c_2 \sin(\omega_0 t)$, jede Lösung ist also periodisch mit Periode $\frac{2\pi}{\omega_0}$.
- Schwache gedämpfte Schwingung ($0 < \mu < \omega_0$): In diesem Fall ist die allgemeine reelle Lösung durch $x(t) = c_1 e^{-\mu t} \cos(\omega t) + c_2 e^{-\mu t} \sin(\omega t)$ mit $\omega := \sqrt{\omega_0^2 - \mu^2} < \omega_0$ gegeben, die Schwingungen werden im Lauf der Zeit gedämpft und es gilt $x(t) \rightarrow 0$ für $t \rightarrow +\infty$.
- Aperiodischer Grenzfall ($\mu = \omega_0$): In diesem Fall wechselt die allgemeine reelle Lösung $x(t) = c_1 e^{-\mu t} + c_2 t e^{-\mu t}$ höchstens einmal ihr Vorzeichen und geht dann für $t \rightarrow +\infty$ gegen Null.
- Stark gedämpfte Schwingung ($\mu > \omega_0$): In diesem Fall ist mit $\kappa := \sqrt{\mu^2 - \omega_0^2} < \mu$ die allgemeine reelle Lösung durch $x(t) = c_1 e^{-(\mu - \kappa)t} + c_2 e^{-(\mu + \kappa)t}$ gegeben, sie wechselt niemals ihr Vorzeichen und konvergiert wegen $\mu \pm \kappa \geq 0$ für $t \rightarrow +\infty$ exponentiell schnell gegen Null.

Differentialgleichungen für orthogonale Polynome

Ist μ ein Maß auf einem Intervall (a, b) , so heißt eine Folge (P_n) von Polynomen n -ten Grades orthogonal in $L^2_\mu(a, b)$, falls $\int_a^b P_n(x) P_m(x) d\mu(x) = 0$ für $n \neq m$ gilt. Häufig erhält man orthogonale Polynome als Lösungen von parameterabhängigen homogenen linearen Differentialgleichung zweiter Ordnung.

Legendre-Polynome Die Legendre-Polynome $P_n(x) := \frac{1}{2^n n!} ((x^2 - 1)^n)^{(n)}$ sind orthogonal bzgl. des Lebesgue-Maßes auf $(-1, 1)$ und lösen auf $(-1, 1)$ die Legendre-Differentialgleichung

$$(1 - x^2)y'' - 2xy' + n(n + 1)y = 0, \quad (1.16)$$

die man auch als $-((x^2 - 1)y')' + n(n + 1)y = 0$ schreiben kann.

Tatsächlich ist $y(x) := ((x^2 - 1)^n)^{(n)}$ eine Lösung von (1.16), denn definiert man $z(x) := (x^2 - 1)((x^2 - 1)^n)' = 2nx(x^2 - 1)^n$, so gilt³ wegen $(xg)^{(n+1)} = xg^{(n+1)} + (n + 1)g^{(n)}$ einerseits

$$z^{(n+1)} = (2nx(x^2 - 1)^n)^{(n+1)} = 2nxy' + 2n(n + 1)y$$

und wegen $((x^2 - 1)g)^{(n+1)} = (x^2 - 1)g^{(n+1)} + 2(n + 1)xg^{(n)} + 2\binom{n+1}{2}g^{(n-1)}$ andererseits

$$\begin{aligned} z^{(n+1)} &= ((x^2 - 1)((x^2 - 1)^n)^{(n+1)}) \\ &= (x^2 - 1)((x^2 - 1)^n)^{(n+2)} + 2(n + 1)x((x^2 - 1)^n)^{(n+1)} + (n + 1)n((x^2 - 1)^n)^{(n)} \\ &= (x^2 - 1)y'' + 2(n + 1)xy' + n(n + 1)y. \end{aligned}$$

Ein Vergleich beider Ableitungen liefert daher $(1 - x^2)y'' - 2xy' + n(n + 1)y = 0$.

Außerdem sind die Polynome P_n und P_m bei $n \neq m$ orthogonal bzgl. des Lebesgue-Maßes auf $(-1, 1)$, denn es gilt $\int_{-1}^1 x^k P_n(x) dx = 0$ für alle $k < n$, da man die Gleichung $\int_{-1}^1 x^k ((x^2 - 1)^n)^{(m)} dx = 0$ für alle $k < m \leq n$ per Induktion nach m zeigen kann. Tatsächlich, es gilt $\int_{-1}^1 1((x^2 - 1)^n)' dx = 0$ im Fall $m = 1$, und aus der Gleichung $\int_{-1}^1 x^k ((x^2 - 1)^n)^{(m)} dx = 0$ für $k < m < n$ folgt

$$\begin{aligned} \int_{-1}^1 x^{k+1} ((x^2 - 1)^n)^{(m+1)} dx &= x^{k+1} ((x^2 - 1)^n)^{(m)} \Big|_{-1}^1 - (k + 1) \int_{-1}^1 x^k ((x^2 - 1)^n)^{(m)} dx = \\ &= x^{k+1} ((x^2 - 1)^n)^{(m)} \Big|_{-1}^1 = 0 \end{aligned}$$

wobei benutzt wurde, dass $(x^2 - 1)^n$ die n -fachen Nullstellen ± 1 hat, also $((x^2 - 1)^n)^{(m)}$ immer noch eine $(n - m)$ -fache Nullstelle in ± 1 hat.

Die Legendre-Differentialgleichung (1.16) bzw. ihre Umformung

$$y'' - \frac{2x}{1 - x^2}y' + \frac{n(n + 1)}{1 - x^2}y = 0$$

besitzt auf $(-1, 1)$ neben P_n noch eine weitere von P_n linear unabhängige Lösung Q_n , die man mittels des Ansatzes $Q_n(x) := z(x)P_n(x)$ erhalten kann. Setzt man diesen nämlich in die umgeformte Legendre-Differentialgleichung ein, so erhält man die Differentialgleichung

$$P_n z'' + (2P_n' - \frac{2x}{1 - x^2}P_n)z' = 0$$

und daraus für z' die homogene lineare Differentialgleichung erster Ordnung

$$z'' + (2\frac{P_n'}{P_n} - \frac{2x}{1 - x^2})z' = 0.$$

³vgl. Leibnizsche Produktregel $(fg)^{(n)} = \sum_{k=0}^n \binom{n}{k} f^{(n-k)} g^{(k)}$ für Produkte n -mal differenzierbarer Funktionen.

Beispielsweise gilt $P_1(x) = x$ und man muss die Differentialgleichung erster Ordnung $z'' + (\frac{2}{x} - \frac{2x}{1-x^2})z' = 0$ für z' lösen. Diese hat die Lösung

$$z' = e^{-\int(\frac{2}{x} - \frac{2x}{1-x^2})dx} = \frac{1}{x^2(1-x^2)},$$

und mit der Partialbruchzerlegung $z' = \frac{1}{x^2} + \frac{1}{2}(\frac{1}{x+1} - \frac{1}{x-1})$ ergibt sich $z = -\frac{1}{x} + \frac{1}{2} \ln(\frac{1+x}{1-x})$, also $Q_1(x) = \frac{x}{2} \ln(\frac{1+x}{1-x}) - 1$. Man bemerke, dass Q_1 im Gegensatz zu P_1 kein Polynom ist. Später werden wir noch genauer und allgemeiner diskutieren, wie man bei Kenntnis einer Lösung einer homogenen linearen Differentialgleichung die Ordnung zu reduzieren kann.

Hermite-Polynome Die Hermite-Polynome $H_n(x) := (-1)^n e^{x^2} (e^{-x^2})^{(n)}$ sind orthogonal bzgl. des Maßes $e^{-x^2} dx$ auf \mathbb{R} und lösen auf \mathbb{R} die Hermite-Differentialgleichung

$$y'' - 2xy' + 2ny = 0. \tag{1.17}$$

Tatsächlich ist H_n ein Polynom n -ten Grades, denn die n -fache Ableitung von e^{-x^2} hat wegen

$$(p_{n-1}(x)e^{-x^2})' = (p'_{n-1}(x) - 2xp_{n-1}(x))e^{-x^2}$$

die Form $p_n(x)e^{-x^2}$ mit einem Polynom n -ten Grades p_n . Mit dieser Bezeichnung ist $H_n(x) = (-1)^n p_n(x)$.

Um zu zeigen, dass $y(x) := e^{x^2} (e^{-x^2})^{(n)}$ eine Lösung von (1.17) ist, berechne man

$$\begin{aligned} (e^{-x^2} y)'' &= (e^{-x^2})^{(n+2)} = (-2xe^{-x^2})^{(n+1)} = -2x(e^{-x^2})^{(n+1)} - 2(n+1)(e^{-x^2})^{(n)} = \\ &= -2x(e^{-x^2} y)' - 2(n+1)e^{-x^2} y = e^{-x^2} (4x^2 y - 2xy' - 2(n+1)y) \end{aligned}$$

und

$$(e^{-x^2} y)'' = e^{-x^2} y'' + 2(e^{-x^2})' y' + (e^{-x^2})'' y = e^{-x^2} (y'' - 4xy' + (4x^2 - 2)y),$$

woraus durch Vergleich $y'' - 2xy' + 2ny = 0$ folgt. Für die Orthogonalität muss man nur $\int_{-\infty}^{\infty} x^k (e^{-x^2})^{(n)} dx = 0$ für alle $k < m \leq n$ zeigen. Tatsächlich gilt

$$\int_{-\infty}^{\infty} 1 \cdot (e^{-x^2})^{(n)} dx = \lim_{R \rightarrow \infty} (e^{-x^2})^{(n-1)}|_{-R}^R = 0$$

im Fall $m = 1$, und aus $\int_{-\infty}^{\infty} x^k (e^{-x^2})^{(n)} dx = 0$ für $k < m < n$ folgt

$$\int_{-\infty}^{\infty} x^{k+1} (e^{-x^2})^{(n)} dx = \lim_{R \rightarrow \infty} \left(x^{k+1} (e^{-x^2})^{(n-1)}|_{-R}^R \right) - (k+1) \int_{-\infty}^{\infty} x^k (e^{-x^2})^{(n-1)} dx = 0 - 0 = 0$$

nach Induktionsvoraussetzung.

Transformation auf konstante Koeffizienten

In diesem Abschnitt stellen wir uns die Frage, unter welchen Bedingungen wir die Zeit t so transformieren können, dass aus der allgemeinen homogenen linearen Differentialgleichung zweiter Ordnung

$$x'' + a_1(t)x' + a_0(t)x = 0$$

eine lineare Differentialgleichung mit konstanten Koeffizienten wird. Homogene lineare Differentialgleichungen mit konstanten Koeffizienten können wir nämlich durch den Ansatz $e^{\lambda t}$ ganz leicht lösen, wie wir schon am Beispiel der Gleichung (1.15) gesehen haben.

Solch eine Transformation auf konstante Koeffizienten ist genau dann möglich, wenn $\frac{a'_0+2a_1a_0}{2a_0^{\frac{3}{2}}}$ konstant ist. Tatsächlich, ist $s = s(t)$ und $y(s) = x(t)$, dann gilt

$$x'(t) = y'(s)s'(t) \text{ und } x''(t) = y''(s)|s'(t)|^2 + y'(s)s''(t),$$

also für y die Differentialgleichung

$$y'' + \frac{s''(t) + a_1(t)s'(t)}{|s'(t)|^2}y' + \frac{a_0(t)}{|s'(t)|^2}y(s) = 0$$

Um die Koeffizientenfunktion vor y zu einer Konstanten b_0 zu machen, muss man also $s(t) = b_0^{-\frac{1}{2}} \int |a_0(t)|^{\frac{1}{2}} dt$ wählen. Der Koeffizient vor y' lautet dann $b_0^{\frac{1}{2}} \frac{a'_0+2a_1a_0}{2a_0^{\frac{3}{2}}}$ und ist daher genau dann konstant, wenn wie behauptet $\frac{a'_0+2a_1a_0}{2a_0^{\frac{3}{2}}}$ konstant ist.

Beispiel 1.26. Für die Differentialgleichung $x'' + \frac{1}{t}x' + \frac{1}{(t \ln(t))^2}x = 0$ auf $(1, \infty)$ ist

$$\frac{\frac{-2}{(t \ln(t))^3}(\ln(t) + 1) + 2\frac{1}{t} \frac{1}{(t \ln(t))^2}}{2\frac{1}{(t \ln(t))^3}} = -1$$

konstant, also liefert die Transformation $s(t) := \int \frac{1}{t \ln(t)} dt = \ln(\ln(t))$, $y(s) := x(t)$, die lineare Differentialgleichung mit konstanten Koeffizienten $y'' - y' + y = 0$ für y . Diese Differentialgleichung hat die allgemeine reelle Lösung

$$y(s) = c_1 e^{\frac{s}{2}} \cos\left(\frac{\sqrt{3}}{2}s\right) + c_2 e^{\frac{s}{2}} \sin\left(\frac{\sqrt{3}}{2}s\right),$$

und daher lautet die allgemeine Lösung der ursprünglichen Differentialgleichung auf $(1, \infty)$

$$x(t) = c_1 \sqrt{\ln(t)} \cos\left(\frac{\sqrt{3}}{2} \ln(\ln(t))\right) + c_2 \sqrt{\ln(t)} \sin\left(\frac{\sqrt{3}}{2} \ln(\ln(t))\right).$$

1.3.2 Nichtlineare Schwingungen

Differentialgleichungen zweiter Ordnung der Form

$$x'' = -U'(x) \tag{1.18}$$

mit einer stetig differenzierbaren Funktion $U : \mathbb{R} \rightarrow \mathbb{R}$ treten in der Mechanik als Newtonsche Bewegungsgleichungen für ein eindimensionales Partikel in einem Potentialfeld auf. Offensichtlich kann man jede Differentialgleichung zweiter Ordnung der Form

$$x'' = f(x)$$

mit stetigem $f : \mathbb{R} \rightarrow \mathbb{R}$ durch Wahl einer Stammfunktion U von $-f$ in die Form (1.18) bringen.

Ist x eine Lösung von (1.18), dann ist $E := \frac{1}{2}(x'(t))^2 + U(x(t))$ konstant, denn es gilt

$$\frac{dE}{dt} = x'(t)x''(t) + U'(x(t))x'(t) = x'(t)(-U'(x(t))) + U'(x(t))x'(t) = 0.$$

Physikalisch interpretiert man U als potentielle Energie, $\frac{1}{2}(x'(t))^2$ als kinetische Energie und E als Gesamtenergie des Systems. Wegen $\frac{1}{2}(x'(t))^2 \geq 0$ verläuft eine Lösung x für alle Zeiten innerhalb der Menge $\{x \mid U(x) \leq E\}$.

Multipliziert man (1.18) mit x' und integriert auf beiden Seiten, so erhält man

$$\frac{1}{2}(x')^2 = E - U(x).$$

Also löst x eine Differentialgleichung erster Ordnung $x' = \pm\sqrt{2(E - U(x))}$, d.h. man hat die Differentialgleichung zweiter Ordnung (1.18) auf eine autonome Differentialgleichung erster Ordnung zurückgeführt, die man mittels Trennung der Variablen lösen kann. Das Integral $\int_{x_0}^x \frac{1}{\sqrt{2(E-U(\xi))}}$ gibt dabei die Zeit an, die das Teilchen benötigt hat, um von x_0 nach x zu gelangen. Besonders interessant ist der im folgenden Satz diskutierte Spezialfall von (1.18), da dort die Lösungen (möglicherweise nichtlineare) Schwingungen sind.

Satz 1.27. *Sei $U: \mathbb{R} \rightarrow \mathbb{R}$ stetig differenzierbar, sei $(t_0, x_0, v_0) \in \mathbb{R}^3$ und sei $E := \frac{1}{2}v_0^2 + U(x_0)$. Ist (x_A, x_B) ein x_0 enthaltendes Intervall mit $U(x_A) = U(x_B) = E$, $U(x) < E$ auf (x_A, x_B) und $U'(x_A), U'(x_B) \neq 0$, dann besitzt (1.18) zu den Anfangswerten $x(t_0) = x_0$, $x'(t_0) = v_0$, genau eine Lösung auf ganz \mathbb{R} . Diese hat Werte in $[x_A, x_B]$ und ist periodisch mit Periode $T := 2 \int_{x_A}^{x_B} \frac{1}{\sqrt{2(E-U(\xi))}} d\xi$.*

Beweis: Unter den angegebenen Voraussetzungen ist $x \mapsto \sqrt{2(E - U(x))}$ auf (x_A, x_B) stetig differenzierbar und nullstellenfrei, sowie stetig in den Punkten x_A, x_B . Darüberhinaus existieren die uneigentlichen Integrale $\int_{x_0}^{x_B} \frac{1}{\sqrt{2(E-U(\xi))}} d\xi$ und $\int_{x_A}^{x_0} \frac{1}{\sqrt{2(E-U(\xi))}} d\xi$. Das erste Integral kann man nämlich mit dem konvergenten uneigentlichen Integral $\int_{x_0}^{x_B} \frac{1}{\sqrt{x_B - \xi}} d\xi$ vergleichen, da wegen $U'(x_B) > 0$ und

$$\lim_{\xi \nearrow x_B} \frac{\sqrt{E - U(\xi)}}{\sqrt{x_B - \xi}} = \lim_{\xi \nearrow x_B} \sqrt{\frac{U(x_B) - U(\xi)}{x_B - \xi}} = \sqrt{U'(x_B)} \neq 0$$

gilt (analog für das zweite Integral). Also gibt es nach Satz (1.18) auf dem Bild $[t_A, t_B]$ von $[x_A, x_B]$ unter der Funktion

$$H : x \mapsto t_0 \pm \int_{x_0}^x \frac{1}{\sqrt{2(E - U(\xi))}} d\xi$$

eine Lösung x von $x' = \pm\sqrt{2(E - U(x))}$ (wobei als Vorzeichen $+$ bei $v_0 > 0$ bzw. $-$ bei $v_0 < 0$ zu wählen ist), die gerade durch die Umkehrfunktion $x(t) := H^{-1}(t)$ gegeben ist. Diese erfüllt $x(t_A) = x_A$, $x(t_B) = x_B$, $x(t) \in (x_A, x_B)$ für $t \in (t_A, t_B)$, $x'(t_A) = 0 = x'(t_B)$ und $x' > 0$ auf (t_A, t_B) bei $v_0 > 0$ (bzw. $x' < 0$ bei $v_0 < 0$). Darüberhinaus gilt sogar $x''(t_B) = -U'(x(t_B)) = -U'(x_B) < 0$ (und analog $x''(t_A) = -U'(x_A) > 0$), so dass man die gefundene Lösung durch $x(t) := x(2t_B - t)$ auf $[t_B, t_A + T]$ (man bemerke, dass $[t_A, t_B]$ nach Definition von T die Länge $\frac{T}{2}$ hat) und dann schließlich durch $x(t) := x(t - kT)$ mit dem $k \in \mathbb{N}$, für das $t - kT \in [t_A, t_A + T]$ gilt, zu einer T -periodischen Lösung auf ganz T fortsetzen kann. \square

Ist x_0 ein isoliertes Minimum von U , dann ist die konstante Funktion $x(t) := x_0$ die einzige Lösung von (1.18) zum Anfangswert $x(t_0) = x_0$, $x'(t_0) = 0$, denn einerseits gilt $x_0'' = 0 = -U'(x_0)$, und andererseits folgt für eine Lösung x des Anfangswertproblem aus $\frac{1}{2}(x')^2 + U(x) = U(x_0)$ aufgrund der Minimalität von $U(x_0)$ und $\frac{1}{2}(x')^2 \geq 0$ schon $U(x(t)) = U(x_0)$ und somit $x(t) = x_0$ aufgrund der Isoliertheit.

1.3.3 Autonome Differentialgleichungen zweiter Ordnung

Eine allgemeine autonome Differentialgleichung zweiter Ordnung hat die Form

$$x'' = f(x, x'). \quad (1.19)$$

Hängt f nur von x' ab, so liefert die Substitution $y := x'$ eine autonome Differentialgleichung erster Ordnung für y , die man mit Trennung der Variablen lösen kann, und anschließende Rücksubstitution ergibt $x = \int y(t) dt + \text{const.}$ Hängt f nur von x ab, so geht Gleichung (1.19) in Gleichung (1.18) über, und durch Multiplikation mit x' führt man die Gleichung auf eine autonome Differentialgleichung erster Ordnung zurück.

In der allgemeinen Situation kann man (1.19) immerhin noch auf eine nichtautonome Differentialgleichung erster Ordnung zurückführen, indem man $y(x) := x'(t(x))$ mit der Umkehrfunktion $t(x)$ von $x(t)$ betrachtet. Diese Funktion y erfüllt nämlich wegen

$$y'(x) = x''(t(x))t'(x) = f(x(t(x)), x'(t(x))) \frac{1}{x'(t(x))} = \frac{f(x, y)}{y}$$

die Differentialgleichung erster Ordnung $y' = \frac{f(x,y)}{y}$. Ist $y(x)$ eine Lösung dieser Differentialgleichung, so kann man dann $t(x)$ durch

$$t(x) = \int t'(x) dx = \int \frac{1}{x'(t(x))} dx = \int \frac{1}{y(x)} dx$$

zurückgewinnen und erhält die Lösung $x(t)$ von (1.19) als Umkehrfunktion von $t(x)$.

Beispiel 1.28. Die autonome Differentialgleichung zweiter Ordnung $x'' = \frac{(x')^2}{x}$ zu den Anfangswerten $x(0) = x_0 > 0$, $x'(0) = y_0 > 0$, geht bei $y(x) := x'(t(x))$ in die Differentialgleichung erster Ordnung $y' = \frac{y}{x}$ zum Anfangswert $y(x_0) = y_0$ über. Deren Lösung ist $y(x) = \frac{y_0}{x_0}x$, und aus

$$t(x) = \frac{x_0}{y_0} \int_{x_0}^x \frac{1}{\tilde{x}} d\tilde{x} = \frac{x_0}{y_0} \ln \left| \frac{x}{x_0} \right|$$

erhält man $x(t) = x_0 e^{\frac{y_0}{x_0} t}$.

Alternativ hätte man die Differentialgleichung $x'' = \frac{(x')^2}{x}$ auch durch x' teilen und zu $(\ln |x'|)' = (\ln |x|)'$ umschreiben können. Integration hätte $\ln \left| \frac{x'}{y_0} \right| = \ln \left| \frac{x}{x_0} \right|$ und daher die Differentialgleichung erster Ordnung $x' = \frac{y_0}{x_0}x$ zum Anfangswert $x(0) = x_0$ geliefert, aus der sich ebenso $x(t) = x_0 e^{\frac{y_0}{x_0} t}$ ergibt.

1.4 Der Existenz- & Eindeutigkeitsatz v. Picard-Lindelöf

Zwar haben wir in den vorigen Abschnitten einige wenige Differentialgleichungen explizit lösen können, für die meisten Differentialgleichungen gibt es aber keine elementaren Lösungsmethoden. Der in diesem Abschnitt präsentierte Satz von Picard-Lindelöf garantiert nun für eine recht allgemeine Klasse von Differentialgleichungen zumindest die lokale Existenz und Eindeutigkeit von Lösungen zu einem Anfangswert. Zusätzlich gibt sein Beweis ein Verfahren an, mit dessen Hilfe man die Lösung numerisch beliebig genau annähern kann.

Dazu schreibt man für stetiges $f : \Omega \subset \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ das Anfangswertproblem

$$x' = f(t, x) \quad , \quad x(t_0) = x_0 \quad , \quad (1.20)$$

durch Integration über $[t_0, t]$ zunächst in die Integralgleichung

$$x(t) = x_0 + \int_{t_0}^t f(s, x(s)) \, ds =: (Tx)(t) \quad (1.21)$$

um und beobachtet, dass eine Funktion $x : I \rightarrow \mathbb{R}^n$ genau dann eine Lösung von (1.20) ist, wenn x ein Fixpunkt des in (1.21) definierten Integraloperators T auf $C(I, \mathbb{R}^n)$ ist.

Aufgabe 1.29. *Weisen Sie nach, dass jede Lösung $x \in C(I, \mathbb{R}^n)$ von (1.21) auch eine Lösung von (1.20) ist.*

Um also die eindeutige Lösbarkeit von (1.20) zu beweisen, müssen wir nur nachweisen, dass der Operator T einen eindeutigen Fixpunkt besitzt, und dafür bietet sich natürlich der Banachsche Fixpunktsatz an.

Satz 1.30 (Banach). *Sei (X, d) ein vollständiger metrischer Raum und $T : X \rightarrow X$ eine Kontraktion, d.h. es gibt ein $L < 1$ mit $d(T(x), T(y)) \leq Ld(x, y)$. Dann gibt es einen eindeutigen Punkt $x^* \in X$ mit $T(x^*) = x^*$.*

Genauer konvergiert die durch $x_{k+1} := T(x_k)$ rekursiv definierte Folge für jeden Startwert $x_0 \in X$ gegen den Fixpunkt x^ von T .*

Beweis: Zunächst einmal kann es höchstens einen Fixpunkt x^* von T geben, denn ist \tilde{x}^* ein weiterer Fixpunkt, dann folgt aus $x^* = T(x^*)$ und $\tilde{x}^* = T(\tilde{x}^*)$ wegen

$$d(x^*, \tilde{x}^*) = d(T(x^*), T(\tilde{x}^*)) \leq Ld(x^*, \tilde{x}^*)$$

und $L < 1$ schon $d(x^*, \tilde{x}^*) = 0$ und somit $x^* = \tilde{x}^*$.

Nun wollen wir uns den gesuchten Fixpunkt konstruktiv verschaffen, indem wir beginnend mit einem beliebigen Startwert $x_0 \in X$ durch $x_{k+1} := T(x_k)$ rekursiv eine Folge definieren und zeigen, dass diese eine Cauchy-Folge ist. Dann konvergiert x_k aufgrund der Vollständigkeit von (X, d) nämlich gegen einen Punkt x^* , und daher auch $T(x_k)$ gegen $T(x^*)$ wegen $d(T(x_k), T(x^*)) \leq Ld(x_k, x^*) \rightarrow 0$. Aufgrund von

$$x^* = \lim_{k \rightarrow \infty} x_{k+1} = \lim_{k \rightarrow \infty} T(x_k) = T(x^*)$$

ist somit x^* ein Fixpunkt von T .

Zu zeigen bleibt also nur noch, dass die durch $x_{k+1} := T(x_k)$ rekursiv definierte Folge x_k eine Cauchy-Folge ist. Dazu bemerke man, dass wegen $d(x_{n+1}, x_n) = d(T(x_n), T(x_{n-1})) \leq Ld(x_n, x_{n-1})$ bei $l > k \geq 0$ die Ungleichung

$$d(x_l, x_k) \leq \sum_{n=k}^{l-1} d(x_{n+1}, x_n) \leq \sum_{n=k}^{l-1} L^{n-k} d(x_{k+1}, x_k) \leq \left(\sum_{n=0}^{l-k-1} L^n \right) L^k d(x_1, x_0)$$

gilt. Nutzt man nun noch aus, dass $\sum_{n=0}^{l-k-1} L^n \leq \sum_{n=0}^{\infty} L^n = \frac{1}{1-L}$ wegen $0 \leq L < 1$ gilt, so erhält man schließlich

$$d(x_l, x_k) \leq \frac{L^k}{1-L} d(x_1, x_0)$$

für alle $l > k \geq 0$. Also ist x_k wirklich eine Cauchy-Folge, denn zu $\varepsilon > 0$ findet man wegen $L^k \rightarrow 0$ ein $K \in \mathbb{N}$, ab dem $\frac{L^k}{1-L} d(x_1, x_0) \leq \varepsilon$ für alle $k \geq K$ gilt, und mit diesem K gilt dann auch $d(x_l, x_k) \leq \varepsilon$ für alle $k \geq K$. \square

Wir wollen noch den folgenden Spezialfall festhalten, der sich direkt aus dem Banachschen Fixpunktsatz und der Vollständigkeit abgeschlossener Teilmengen eines Banach-Raumes ergibt.

Korollar 1.31. *Ist $D \subset X$ abgeschlossene Teilmenge eines Banach-Raumes $(X, \|\cdot\|)$ und $T : D \subset X \rightarrow X$ eine kontrahierende Selbstabbildung von D , d.h. es gilt $T(D) \subset D$ und es gibt ein $L < 1$ mit $\|T(x) - T(y)\| \leq L\|x - y\|$, dann gibt es einen eindeutigen Fixpunkt $x^* \in D$ von T .*

Dieses Korollar wenden wir nun auf den in (1.21) definierten Integraloperator T auf einer abgeschlossenen Teilmenge des mit der Maximumsnorm $\|x\|_{\infty} := \max_{t \in I} |x(t)|$ versehenen Banach-Raumes $X := C(I, \mathbb{R}^n)$ für ein kompaktes Intervall I an, um den folgenden Satz von Picard-Lindelöf zu erhalten.

Satz 1.32. *Sei $f : \Omega \subset \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ stetig, sei $(t_0, x_0) \in \Omega$ ein innerer Punkt und genüge f nahe (t_0, x_0) einer Lipschitzbedingung in der Variablen x , d.h. es gibt eine Umgebung $U \subset \Omega$ von (t_0, x_0) und ein $L < \infty$ mit*

$$\|f(t, x) - f(t, \tilde{x})\| \leq L\|x - \tilde{x}\| \quad (1.22)$$

für alle $(t, x), (t, \tilde{x}) \in U$, dann existiert eine positive Zahl $\varepsilon > 0$ und eine eindeutige Lösung $x : [t_0 - \varepsilon, t_0 + \varepsilon] \rightarrow \mathbb{R}^n$ des Anfangswertproblems (1.20).

Beweis: Da (t_0, x_0) ein innerer Punkt von Ω ist und f nahe (t_0, x_0) einer lokalen Lipschitzbedingung genügt, gibt es ein abgeschlossenes Intervall I um t_0 und eine abgeschlossene Kugel $\overline{B_r(x_0)}$ um x_0 mit $I \times \overline{B_r(x_0)} \subset \Omega$, für die $\|f(t, x)\| \leq M$ für alle $(t, x) \in I \times \overline{B_r(x_0)}$ mit einem $M < \infty$ sowie (1.22) auf $U := I \times \overline{B_r(x_0)}$ gilt. Nun wähle man $\varepsilon > 0$ so klein, dass $[t_0 - \varepsilon, t_0 + \varepsilon] \subset I$, $M \cdot \varepsilon \leq r$ und $L \cdot \varepsilon \leq \frac{1}{2}$ gilt.

Ist T der Operator aus (1.21), dann gilt für beliebige stetige Kurven $x, \tilde{x} : [t_0 - \varepsilon, t_0 + \varepsilon] \rightarrow \overline{B_r(x_0)}$ sowohl

$$\|(Tx)(t) - x_0\| = \left\| \int_{t_0}^t f(s, x(s)) \, ds \right\| \leq \int_{\min(t_0, t)}^{\max(t_0, t)} \|f(s, x(s))\| \, ds \leq M \cdot |t - t_0| \leq r$$

für alle $t \in [t_0 - \varepsilon, t_0 + \varepsilon]$ als auch

$$\begin{aligned} \|(Tx)(t) - (T\tilde{x})(t)\| &\leq \int_{\min(t_0, t)}^{\max(t_0, t)} \underbrace{\|f(s, x(s)) - f(s, \tilde{x}(s))\|}_{\stackrel{(1.22)}{\leq} L\|x(s) - \tilde{x}(s)\|} ds \leq \int_{\min(t_0, t)}^{\max(t_0, t)} L\|x(s) - \tilde{x}(s)\| ds \\ &\leq L \cdot |t - t_0| \max_{|s-t_0| \leq \varepsilon} \|x(s) - \tilde{x}(s)\| \\ &\leq \frac{1}{2} \max_{|s-t_0| \leq \varepsilon} \|x(s) - \tilde{x}(s)\|. \end{aligned}$$

Also bildet T die abgeschlossene Teilmenge $C([t_0 - \varepsilon, t_0 + \varepsilon], \overline{B_r(x_0)})$ des mit der Maximumsnorm $\|x\|_\infty := \max_{|t-t_0| \leq \varepsilon} |x(t)|$ versehenen Banach-Raumes $C([t_0 - \varepsilon, t_0 + \varepsilon], \mathbb{R}^n)$ in sich ab und ist gleichzeitig eine Kontraktion. Nach dem Korollar 1.31 zum Banachschen Fixpunktsatz 1.30 gibt es daher einen eindeutigen Fixpunkt $x \in C([t_0 - \varepsilon, t_0 + \varepsilon], \overline{B_r(x_0)})$ von T und somit nach unserer Vorüberlegung eine eindeutige lokale Lösung des Anfangswertproblems (1.20). \square

Wie im Banachschen Fixpunktsatz kann man den Fixpunkt x von T und somit die eindeutige lokale Lösung auch konstruktiv durch Iteration annähern. Hier lautet die Iterationsvorschrift

$$x_{k+1}(t) = x_0 + \int_{t_0}^t f(s, x_k(s)) ds, \quad (1.23)$$

und x_k konvergiert für den konstanten Startwert x_0 auf einem kleinen Intervall um t_0 gegen die eindeutige Lösung des Anfangswertproblems (1.20).

Man beachte, dass der Satz von Picard-Lindelöf nur die lokale Existenz von Lösungen garantiert, denn $\varepsilon > 0$ kann sehr klein sein. Im folgenden Unterabschnitt diskutieren wir, unter welchen Umständen man die gewonnene Lösung auf ein größeres Intervall forsetzen kann. Vorher wollen wir aber noch ein nützliches Kriterium angeben, mit dem man die Gültigkeit der in Satz 1.32 geforderten lokalen Lipschitzbedingung leicht überprüfen kann.

Lemma 1.33. *Ist $f: \Omega \subset \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ stetig und im Inneren von Ω stetig partiell differenzierbar nach x , dann genügt f nahe jedes inneren Punktes $(t_0, x_0) \in \Omega$ einer Lipschitzbedingung in der Variablen x .*

Beweis: Nach dem Mittelwertsatz im \mathbb{R}^n gilt

$$\begin{aligned} |f(t, \tilde{x}) - f(t, x)| &= \left| \int_0^1 \partial_{\tilde{x}-x} f(t, x + s(\tilde{x} - x)) ds \right| \\ &\leq \int_0^1 |\partial_{\tilde{x}-x} f(t, x + s(\tilde{x} - x))| ds \leq \max_{s \in [0,1]} |\partial_{\tilde{x}-x} f(t, x + s(\tilde{x} - x))| \end{aligned}$$

Die Aussage folgt dann nach Definition der Operatornorm einfach aus

$$|\partial_{\tilde{x}-x} f(t, x + s(\tilde{x} - x))| \leq |\partial_{\mathbb{R}^n} f(t, x + s(\tilde{x} - x))| |\tilde{x} - x|$$

und der Tatsache, dass für eine kompakten Umgebung $I \times K \subset \Omega$ von (t_0, x_0) die stetige Funktion $(s, t, x, \tilde{x}) \mapsto |\partial_{\mathbb{R}^n} f(t, x + s(\tilde{x} - x))|$ auf $[0, 1] \times I \times K \times K$ ein Maximum besitzt. \square

Ist $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ offen und genügt f nahe jedes Punktes $(t_0, x_0) \in \Omega$ einer Lipschitzbedingung in der Variablen x , dann sagt man auch, dass f auf Ω einer lokalen Lipschitzbedingung genügt.

1.4.1 Fortsetzungssatz

Wir fragen uns nun, ob man die lokalen Lösungen zu einem Anfangswert, deren Existenz und Eindeutigkeit Satz 1.32 garantiert, zu einer „maximalen Lösung“ fortsetzen kann.

Definition 1.34. Eine Kurve $x: I_{\max} \rightarrow \mathbb{R}^n$ heißt *maximale Lösung des Anfangswertproblems* (1.20), falls x eine Lösung von (1.20) ist und jede andere Lösung von (1.20) nur die Einschränkung von x auf ein kleineres Intervall $I \subset I_{\max}$ ist.

Der folgende Satz garantiert die Existenz sowie die Eindeutigkeit von maximalen Lösungen und gibt darüberhinaus auch noch Auskunft darüber, was in den Randpunkten des maximalen Existenzintervalles I_{\max} passiert. Dazu sei daran erinnert, dass man für Kurven $x: (a, b) \rightarrow \mathbb{R}^n$ sagt, x strebe bei $t \nearrow a$ (bzw. $t \searrow b$) gegen unendlich, wenn die Norm von $x(t)$ für $t \nearrow a$ (bzw. $t \searrow b$) gegen unendlich strebt.

Satz 1.35. Sei $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ offen, sei $f: \Omega \rightarrow \mathbb{R}^n$ stetig, und genüge f auf Ω einer lokalen Lipschitzbedingung. Dann gibt es zu jedem Anfangswert $(t_0, x_0) \in \Omega$ eine maximale Lösung $x: I_{\max} \rightarrow \mathbb{R}^n$ des Anfangswertproblems (1.20), $I_{\max} = (a, b)$ ist offen, und falls $b < \infty$ (bzw. $a > -\infty$) gilt, dann strebt $x(t)$ für $t \nearrow b$ (bzw. $t \searrow a$) entweder gegen einen Randpunkt von Ω oder gegen unendlich.

Beweis: Wir zeigen zunächst: Sind $x: I \rightarrow \mathbb{R}^n$ und $\tilde{x}: \tilde{I} \rightarrow \mathbb{R}^n$ zwei Lösungen des Anfangswertproblems (1.20), dann gilt $x = \tilde{x}$ auf dem (nichtleeren) Intervall $I \cap \tilde{I}$. Denn angenommen, x und \tilde{x} wären (ohne Einschränkung) rechts von t_0 nicht identisch, dann gäbe es rechts von t_0 ein größtes $\tilde{t} \in I \cap \tilde{I}$ mit $x(t) = \tilde{x}(t)$ für alle $t \in [t_0, \tilde{t}]$, und \tilde{t} ist nicht rechter Randpunkt des Intervalls $I \cap \tilde{I}$. Zum neuen Anfangswert $(\tilde{t}, x(\tilde{t}) = \tilde{x}(\tilde{t}))$ existiert aber nach Satz 1.32 lokal eine eindeutige Lösung auf $[\tilde{t} - \varepsilon, \tilde{t} + \varepsilon]$ mit einem $\varepsilon > 0$, Aufgrund der Eindeutigkeit stimmt diese auf $[\tilde{t} - \varepsilon, \tilde{t} + \varepsilon]$ mit x und \tilde{x} überein, im Widerspruch dazu, dass $\tilde{t} \in I \cap \tilde{I}$ die größte Zahl mit $x(t) = \tilde{x}(t)$ für alle $t \in [t_0, \tilde{t}]$ war. Also war die Annahme falsch, x und \tilde{x} wären rechts von t_0 nicht identisch (links von t_0 zeigt man dieses analog).

Eine Konsequenz aus dieser weitreichenden Form der Eindeutigkeit ist, dass man zu jeder beliebigen Familie von Lösungen $x_i: I_i \rightarrow \mathbb{R}^n$ des Anfangswertproblems (1.20) eine Lösung x von (1.20) auf $\bigcup_i I_i$ erhält, indem man $x(t) := x_i(t)$ bei $t_i \in I_i$ setzt. Denn für Punkte $t \in I_i \cap I_j$ gilt ja $x_i(t) = x_j(t)$, so dass x wohldefiniert und natürlich auch eine Lösung des Anfangswertproblems (1.20) ist. Wendet man diese Ausdehnungseigenschaft auf die Familie aller Lösungen an, so erhält man also eine maximale Lösung $x: I_{\max} \rightarrow \mathbb{R}^n$, und I_{\max} muss offen sein, denn zu einem Randpunkt \tilde{t} von I_{\max} könnte man mittels Satz 1.32 wieder ein Lösung zum Anfangswert $(\tilde{t}, x(\tilde{t}))$ auf $[\tilde{t} - \varepsilon, \tilde{t} + \varepsilon]$ mit einem $\varepsilon > 0$ finden, im Widerspruch zur Definition von x .

Zu zeigen bleibt nur noch, dass bei $I_{\max} = (a, b)$ mit $b < \infty$ (bzw. $a > -\infty$) der Punkt $x(t)$ für $t \nearrow b$ (bzw. $t \searrow a$) gegen einen Randpunkt von Ω oder gegen unendlich strebt. Wäre dies nämlich nicht der Fall, dann liegt $x([t_0, b))$ (bzw. $x((a, t_0])$) in einer kompakten Teilmenge von Ω . Auf dieser wäre f als stetige Funktion aber beschränkt, d.h. $x'(t) = f(t, x(t))$ wäre auf $[t_0, b)$ (bzw. $(a, t_0])$ beschränkt, und somit wäre x Lipschitz-stetig auf $[t_0, b)$ (bzw. $(a, t_0])$. Eine Lipschitz-stetige Funktion überführt aber Cauchy-Folgen in Cauchy-Folgen, so dass der Grenzwert $\lim_{t \nearrow b} x(t)$ (bzw. $\lim_{t \searrow a} x(t)$) existiert. Setzt man x mit Hilfe dieses Grenzwertes auf $[t_0, b]$ (bzw. $[a, t_0]$) fort, so ist x und damit auch $t \mapsto f(t, x(t))$ stetig auf $[t_0, b]$ (bzw. $[a, t_0]$). Daher kann man in der für $t \in [t_0, b)$ (bzw. $t \in (a, t_0]$) gültigen Gleichung

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

den Grenzübergang $t \nearrow b$ (bzw. $t \searrow a$) vollziehen und erhält die Gültigkeit dieser Gleichung auch für $t = b$ (bzw. $t = a$). Dies bedeutet aber gerade, dass x in $t = b$ (bzw. $t = a$) die Ableitung $x'(b) = f(b, x(b))$ (bzw. $x'(a) = f(a, x(a))$) hat, so dass x eine Lösung des Anfangswertproblems (1.20) auf $[t_0, b]$ (bzw. $[a, t_0]$) wäre, im Widerspruch dazu, dass $I_{\max} = (a, b)$ das maximale Existenzintervall ist. \square

Eine Konsequenz dieses fundamentalen Satzes ist das folgende Korollar. Die Voraussetzung, dass $x : I_{\max} \rightarrow \mathbb{R}^n$ beschränkt zu sein hat, werden wir dabei mit Hilfe von Lyapunov-Funktionen zumindest in Spezialfällen a priori nachweisen können.

Korollar 1.36. *Genügt die stetige Funktion $f : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ einer lokalen Lipschitzbedingung und ist die maximale Lösung $x : I_{\max} \rightarrow \mathbb{R}^n$ zum Anfangswert (t_0, x_0) beschränkt, dann gilt $I_{\max} = (-\infty, \infty)$, d.h. x ist eine globale Lösung.*

Beweis: Wäre $I_{\max} = (a, b)$ mit $b < \infty$ (bzw. $a > -\infty$) dann würde $x(t)$ für $t \nearrow b$ (bzw. $t \searrow a$) entweder gegen einen Randpunkt von Ω oder gegen unendlich streben, beides kann aber nicht sein, denn $\Omega := \mathbb{R} \times \mathbb{R}^n$ hat keinen Randpunkt und x ist beschränkt. \square

Aufgabe 1.37. *Beweisen Sie, dass das Anfangswertproblem $x' = f(t, x)$, $x(t_0) = x_0$, eine eindeutige globale Lösung $x : I \rightarrow \mathbb{R}^n$ besitzt, falls $f : I \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine stetige Funktion ist, die mit einer Konstanten $L < \infty$ der globalen Lipschitzbedingung $|f(t, x) - f(t, \tilde{x})| \leq L|x - \tilde{x}|$ für alle $t \in I$ und $x, \tilde{x} \in \mathbb{R}^n$ genügt.*

Reparametrisierungen

Häufig kommt es beim Studium einer autonomen Differentialgleichung nicht darauf an, in welchen Zeiten die Lösungskurven durchlaufen werden, sondern allein die Gestalt der Lösungskurven ist wichtig. Die Gesamtheit aller (unparametrisierten) Lösungskurven nennt man das Phasenportrait der autonomen Differentialgleichung.

Definition 1.38. *Ist $x : I \rightarrow \mathbb{R}^n$ eine maximale Lösung von $x' = f(x)$, dann nennt man das Bild $x(I)$ die Spur der Lösung x . Die Menge aller Spuren von Lösungen der autonomen Differentialgleichung $x' = f(x)$ nennt man ihr Phasenportrait.*

Multipliziert man das Vektorfeld $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ einer autonomen Differentialgleichung $x' = f(x)$ mit einer positiven skalaren Funktion $\lambda : \mathbb{R}^n \rightarrow (0, \infty)$, so hat die Differentialgleichung $x' = (\lambda f)(x)$ dasselbe Phasenportrait wie die ursprüngliche Differentialgleichung, nur werden die Lösungskurven in anderen Zeitintervallen durchlaufen. Tatsächlich ist die Richtung der Tangentialvektoren $f(x)$ und $\lambda(x)f(x)$ dieselbe, nur ihre Länge ist unterschiedlich. Dies deutet schon darauf hin, dass die Phasenportraits beider Differentialgleichungen identisch sind (für einen formalen Beweis siehe [Chicone, Proposition 1.28]). Daher sagt man auch, die Differentialgleichung $x' = (\lambda f)(x)$ entsteht aus $x' = f(x)$ durch Reparametrisierung der Zeit.

Besitzt eine autonome Differentialgleichung $x' = f(x)$ zu jedem Anfangswert eine globale Lösung, so nennt man das Vektorfeld f **vollständig**. Da es autonome Differentialgleichungen $x' = f(x)$ gibt, für die nicht alle Lösungen global existieren, stellt sich die Frage, ob man nicht durch Multiplikation von $f(x)$ mit einer positiven skalaren Funktion $\lambda(x)$ ein vollständiges Vektorfeld λf erhalten kann, das dann nach der vorigen Beobachtung dasselbe Phasenportrait besitzt.

Proposition 1.39. *Ist das Vektorfeld $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ stetig differenzierbar, so ist $\frac{f(x)}{1+|f(x)|^2}$ ein vollständiges Vektorfeld.*

Beweis: Mit f ist auch $\frac{f(x)}{1+|f(x)|^2}$ stetig differenzierbar, zu jedem Anfangswert (t_0, x_0) existiert also eine maximale Lösung $x : I_{\max} \rightarrow \mathbb{R}^n$ von $x' = \frac{f(x)}{1+|f(x)|^2}$ mit $x(t_0) = x_0$. Aufgrund von

$$|x(t) - x(t_0)| = \left| \int_{t_0}^t x'(s) ds \right| = \left| \int_{t_0}^t \frac{f(x(s))}{1+|f(x(s))|^2} ds \right| \leq \int_{t_0}^t \frac{|f(x(s))|}{1+|f(x(s))|^2} ds \leq |t - t_0|$$

gilt $|x(t)| \leq |x(t_0)| + |t - t_0|$ für alle $t \in I_{\max}$. Wäre nun $I_{\max} = (a, b)$ mit $b < \infty$ (bzw. $a > -\infty$) dann würde $x(t)$ für $t \nearrow b$ (bzw. $t \searrow a$) nach dem Fortsetzungssatz 1.35 gegen unendlich streben, da $\Omega := \mathbb{R}^n$ als Teilmenge des \mathbb{R}^n keine Randpunkte hat, dies widerspricht aber der Ungleichung $|x(t)| \leq |x(t_0)| + |b - t_0|$ für $t_0 \leq t < b$ (bzw. $|x(t)| \leq |x(t_0)| + |a - t_0|$ für $a < t \leq t_0$). Also ist $I_{\max} = (-\infty, \infty)$, d.h. zu jedem Anfangswert existiert eine globale Lösung von $x' = \frac{f(x)}{1+|f(x)|^2}$. \square

1.5 Allgemeiner Existenz- und Eindeutigkeitsresultate

1.5.1 Der Satz von Peano

Ist die rechte Seite f der Differentialgleichung $x' = f(t, x)$ zwar stetig, genügt aber nahe (t_0, x_0) keiner lokalen Lipschitzbedingung in x , so kann man mittels eines Kompaktheitsarguments immerhin noch die Existenz einer Lösung von $x' = f(t, x)$ zum Anfangswert $x(t_0) = x_0$ zeigen. Eindeutigkeit kann jedoch nicht mehr garantiert werden, wie schon Beispiel 1.9 gezeigt hat. Bevor wir zu dieser als Satz von Peano bezeichneten Existenzaussage kommen, müssen wir aber zunächst klären, welche Funktionenfolgen in $C(I, \mathbb{R}^n)$ konvergente Teilfolgen besitzen, und dabei spielt der folgende Begriff eine wichtige Rolle.

Definition 1.40. Eine (abzählbare) Familie $x_k : I \rightarrow \mathbb{R}^n$ ($k \in \mathbb{N}$) stetiger Kurven heißt gleichgradig stetig, falls gilt:

$$\forall \varepsilon > 0 \exists \delta > 0 : \forall k \in \mathbb{N} \forall s, t \in I : (|s - t| \leq \delta \Rightarrow |x_k(s) - x_k(t)| \leq \varepsilon)$$

Gleichgradige Stetigkeit ist also mehr als Stetigkeit bzw. gleichmäßige Stetigkeit jedes einzelnen x_k . Denn während δ bei Stetigkeit von ε , k und t bzw. bei gleichmäßiger Stetigkeit von ε und k abhängen darf, wird bei gleichgradiger Stetigkeit zusätzlich gefordert, dass δ für alle x_k gleich gewählt werden kann, d.h. δ unabhängig von k ist und nur noch von ε abhängt.

Beispiel 1.41. Die (überabzählbare) Familie aller Kurven $x : I \rightarrow \mathbb{R}^n$, die Lipschitzstetig mit Lipschitzkonstante $\leq L$ für ein feste $0 < L < \infty$ sind, ist gleichgradig stetig, denn zu ε kann man $\delta := \frac{\varepsilon}{L}$ unabhängig von der Kurve x wählen.

Der Satz von Arzela-Ascoli charakterisiert für ein kompaktes Intervall I die kompakten Teilmengen des Banach-Raumes $C(I, \mathbb{R}^n)$ als die Familien von Kurven, die gleichgradig stetig und gleichmäßig beschränkt sind. Wir brauchen nur einen etwas schwächeren Spezialfall dieser Aussage, der aber auch noch als Satz von Arzela-Ascoli bezeichnet wird.

Satz 1.42 (Arzela-Ascoli). Ist die Folge $x_k : I \rightarrow \mathbb{R}^n$, $k \in \mathbb{N}$, von Kurven auf einem kompakten Intervall I gleichgradig stetig und gleichmäßig beschränkt⁴, dann besitzt sie eine auf I gleichmäßig konvergente Teilfolge.

⁴ x_k heißt gleichmäßig beschränkt, falls es ein $M < \infty$ mit $|x_k(t)| \leq M$ für alle $t \in I$ und $k \in \mathbb{N}$ gibt.

Beweis: siehe z.B. [Walter, I.7.IV]. □

Der Satz 1.42 von Arzela-Ascoli erlaubt also insbesondere, aus einer konstruierten Folge von Näherungslösungen des Anfangswertproblems $x' = f(t, x)$, $x(t_0) = x_0$, eine in $C(I, \mathbb{R}^n)$ konvergente Teilfolge auszuwählen, falls die Näherungslösungen gleichgeradig stetig und gleichmäßig beschränkt sind. Im folgenden Satz von Peano wird gezeigt, dass einerseits solche Näherungslösungen existieren und andererseits die Grenzkurve einer konvergenten Teilfolge von Näherungslösungen schon selbst eine lokale Lösung des Anfangswertproblems ist.

Satz 1.43 (Peano). *Ist $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ offen und $f : \Omega \rightarrow \mathbb{R}^n$ stetig, dann gibt es zu jedem Anfangswert (t_0, x_0) ein $\varepsilon > 0$ und eine Lösung $x : [t_0 - \varepsilon, t_0 + \varepsilon] \rightarrow \mathbb{R}^n$ von $x' = f(t, x)$, $x(t_0) = x_0$.*

Beweis: Da Ω offen ist, gibt es eine kompakte Umgebung $[t_0 - \tilde{\varepsilon}, t_0 + \tilde{\varepsilon}] \times \overline{B_r(x_0)} \subset \Omega$ von (t_0, x_0) , und da f stetig ist, nimmt $|f|$ auf $[t_0 - \tilde{\varepsilon}, t_0 + \tilde{\varepsilon}] \times \overline{B_r(x_0)}$ einen maximalen Wert $M < \infty$ an.

Sei nun $\varepsilon := \min(\tilde{\varepsilon}, \frac{r}{M})$. Wir konstruieren zunächst mittels des sogenannten Eulerschen Polygonzugverfahrens eine Folge von stetigen (stückweise linearen) Näherungslösungen $x_k : [t_0 - \varepsilon, t_0 + \varepsilon] \rightarrow \mathbb{R}^n$ des Anfangswertproblems mit der Eigenschaft, dass

$$|x'_k(t) - f(t, x_k(t))| \leq \frac{1}{k}$$

für alle $t \in [t_0 - \varepsilon, t_0 + \varepsilon]$ gilt (wobei $x'_k(t)$ die rechtsseitige oder linksseitige Ableitung bezeichnet).

Dazu beobachten wir, dass f auf der kompakten Teilmenge $[t_0 - \tilde{\varepsilon}, t_0 + \tilde{\varepsilon}] \times \overline{B_r(x_0)}$ als stetige Funktion sogar gleichmäßig stetig ist, es also zu jedem $k \in \mathbb{N}$ ein $\delta_k > 0$ gibt mit

$$|(\tilde{t}, \tilde{x}) - (t, x)| \leq \delta_k \Rightarrow |f(\tilde{t}, \tilde{x}) - f(t, x)| \leq \frac{1}{k}.$$

Um die Näherungslösung x_k rechts von t_0 zu definieren zerteilen wir $[t_0, t_0 + \varepsilon]$ in Teilintervalle $t_0 < t_1 < \dots < t_{N-1} < t_N = t_0 + \varepsilon$ der Maximallänge $\min(\frac{\delta_k}{2}, \frac{\delta_k}{2M})$ und setzen $x_k(t_0) := x_0$ sowie rekursiv

$$x_k(t) := x_k(t_i) + (t - t_i)f(t_i, x_k(t_i)) \text{ für } t \in [t_i, t_{i+1}].$$

Dies liefert stückweise lineare Kurven x_k , deren Graph einerseits wegen

$$\begin{aligned} |x_k(t) - x_0| &\leq |x_k(t) - x_k(t_i)| + \sum_{j=1}^i |x_k(t_j) - x_k(t_{j-1})| \\ &\leq M \left((t - t_i) + \sum_{j=1}^i (t_j - t_{j-1}) \right) \leq M\varepsilon \leq r \end{aligned}$$

für $t \in [t_i, t_{i+1}]$ komplett in $[t_0, t_0 + \tilde{\varepsilon}] \times \overline{B_r(x_0)}$ verläuft und die andererseits die gewünschte Näherungseigenschaft

$$|x'_k(t) - f(t, x_k(t))| = |f(t_i, x_k(t_i)) - f(t, x_k(t))| \leq \frac{1}{k}$$

für $t \in [t_i, t_{i+1}]$ erfüllen wegen $|t - t_i| \leq \frac{\delta_k}{2}$ und

$$|x_k(t) - x_k(t_i)| \leq M(t - t_i) \leq M \frac{\delta_k}{2} = \frac{\delta_k}{2}.$$

Die Näherungslösungen x_k sind gleichmäßig beschränkt, da

$$|x_k(t)| \leq |x_0| + M(t - t_0) \leq |x_0| + M\varepsilon \leq |x_0| + r$$

gilt, und gleichgeradig stetig, da wegen

$$\begin{aligned} |x_k(s) - x_k(t)| &\leq |x_k(s) - x_k(t_j)| + \cdots + |x_k(t_{i+1}) - x_k(t)| \\ &\leq M((s - t_j) + \cdots + (t_{i+1} - t)) = M(s - t) \end{aligned}$$

für $t_i \leq t \leq t_{i+1} \leq \cdots \leq t_j \leq s \leq t_{j+1}$ alle x_k dieselbe Lipschitzkonstante M besitzen. Analog kann man links von t_0 Näherungslösungen mit denselben Eigenschaften konstruieren, und zusammengesetzt erhält man eine gleichgeradig stetige und gleichmäßig beschränkte Folge von Näherungslösungen $x_k : [t_0 - \varepsilon, t_0 + \varepsilon] \rightarrow \mathbb{R}^n$.

Nach dem Satz von Arzela-Ascoli kann man nun eine Teilfolge der x_k finden (wir bezeichnen diese Teilfolge wiederum mit x_k), die gleichmäßig auf $[t_0 - \varepsilon, t_0 + \varepsilon]$ gegen eine stetige Kurve $x : [t_0 - \varepsilon, t_0 + \varepsilon] \rightarrow \mathbb{R}^n$ konvergiert. Insbesondere konvergiert auch $f(\cdot, x_k(\cdot))$ gleichmäßig gegen $f(\cdot, x(\cdot))$, und da $x'_k(\cdot) - f(\cdot, x_k(\cdot))$ aufgrund der Näherungseigenschaft gleichmäßig gegen Null konvergiert, führt der Grenzübergang $k \rightarrow \infty$ in

$$x_k(t) = x_0 + \int_{t_0}^t x'_k(s) ds = x_0 + \int_{t_0}^t f(s, x_k(s)) + (x'_k(s) - f(s, x_k(s))) ds$$

für alle $t \in [t_0 - \varepsilon, t_0 + \varepsilon]$ auf die Gleichung

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

Dies bedeutet aber gerade, dass x eine Lösung des Anfangswertproblems ist. \square

Ähnlich wie im Beweis von Satz 1.35 kann man darüberhinaus noch zeigen, dass sich eine mittels des Satzes von Peano gefundene lokale Lösung x des Anfangswertproblems $x' = f(t, x)$, $x(t_0) = x_0$, mit stetiger rechte Seite f zu einer nicht weiter fortsetzbaren Lösung auf einem maximalen Intervall I_{\max} erweitern läßt (siehe [Walter, I.7.VI], allerdings ist dies aufgrund der fehlenden Eindeutigkeit keine maximale Lösung im Sinne von Definition (1.34). Die Eigenschaft, dass $I_{\max} = (a, b)$ offen ist und bei $b < \infty$ sich $x(t)$ für $t \nearrow b$ dem Rand von Ω nähert oder unendlich wird, gilt aber auch für nur stetiges f .

1.5.2 Schwache Lösungen

Bisher haben wir die Existenz von Lösungen nur für stetige rechte Seiten bewiesen. Manchmal benötigt man aber auch Aussagen für Vektorfelder mit zeitabhängigen Sprüngen. Dies erfordert einen weitergefassten Lösungsbegriff als den in 1.2 definierten. Dazu führen wir zunächst absolut stetige Funktionen ein.

Definition 1.44. *Eine Kurve $x : [a, b] \rightarrow \mathbb{R}^n$ heißt absolut stetig, falls es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt mit $\sum_{i=1}^N |x(b_i) - x(a_i)| \leq \varepsilon$ für N paarweise disjunkte Intervalle $[a_i, b_i] \subset [a, b]$ der Gesamtlänge $\sum_{i=1}^N |b_i - a_i| \leq \delta$.*

Die Kurve x ist nach dem Hauptsatz der Differential- und Integralrechnung für das Lebesgue-Integral ([Elstrodt, VII.4.14]) genau dann absolut stetig auf jedem kompakten Teilintervall von I , wenn es ein lokal Lebesgue-integrierbares $\rho : I \rightarrow \mathbb{R}^n$ mit

$$x(t) = x(t_0) + \int_{t_0}^t \rho(s) ds$$

für alle $t, t_0 \in I$ gibt. Statt ρ schreibt man üblicherweise x' und nennt diese x' die schwache Ableitung von x . Tatsächlich ist jede stetig differenzierbare Kurve x absolut stetig und es gilt $x(t) = x(t_0) + \int_{t_0}^t x'(s) ds$.

Bemerkung 1.45. *Im Beweis des Satzes 1.43 von Peano haben wir schon spezielle schwache Ableitungen verwendet (ohne sie so zu nennen), denn dort wurden mittels des Eulerschen Polygonzugverfahrens stückweise lineare Kurven definiert. Deren schwache Ableitung ist stückweise konstant und existiert somit in den Teilungspunkten nicht klassisch, sondern höchstens einseitig.*

Eine Verallgemeinerung des klassischen Lösungsbegriffs ist durch die folgende Definition schwacher Lösungen gegeben. In der Literatur werden schwache Lösungen auch als Lösungen im Sinne von Carathéodory bezeichnet, siehe [Walter, II.10.XII].

Definition 1.46. *Eine Kurve $x: I \rightarrow \mathbb{R}^n$ heißt schwache Lösung der Differentialgleichung (1.2) zum Anfangswert x_0 bei t_0 , wenn x absolut stetig auf jedem kompakten Teilintervall von I ist, die Differentialgleichung (1.2) für fast alle $t \in I$ erfüllt ist und $x(t_0) = x_0$ gilt.*

Beispiel 1.47. *Die lineare Differentialgleichung $x' = \text{sign}(t)x$, hat zum Anfangswert x_0 bei t_0 die schwache Lösung $x(t) = x_0 \exp(|t| - |t_0|)$. Diese Funktion x ist zwar absolut stetig, aber in $t = 0$ nicht klassisch differenzierbar.*

Während man die Differentialgleichung $x' = \text{sign}(t)x$ möglicherweise als rein akademisches Beispiel ansehen kann, ist $x' = -\frac{1}{2}(1 + \text{sign}(t))x$ eine häufig in Anwendungen auftretende Differentialgleichung, die Schaltvorgänge modelliert. Tatsächlich haben ihre schwachen Lösungen die Form

$$x(t) = \begin{cases} x_0 & t \leq 0 \\ x_0 \exp(-t) & t \geq 0 \end{cases} \text{ und modellieren einen Abschaltvorgang.}$$

Um die Existenz von schwachen Lösungen zu beweisen, benötigen wir noch einige weitere Begriffe. Zunächst spezialisieren wir uns auf absolut stetige Kurven x , deren schwache Ableitung x' in $L^p_{loc}(I, \mathbb{R}^n)$ für ein $1 < p < \infty$ statt nur in $L^1_{loc}(I, \mathbb{R}^n)$ liegt. Den entsprechenden Raum von Kurven bezeichnet man mit $W^{1,p}_{loc}(I, \mathbb{R}^n)$, und den Raum von absolut stetigen Kurven x mit $x' \in L^p(I, \mathbb{R}^n)$ bezeichnet man mit $W^{1,p}(I, \mathbb{R}^n)$. Ausgestattet mit der Norm $\|x\|_{W^{1,p}} := \left(\int_I |x(t)|^p + |x'(t)|^p dt\right)^{1/p}$ ist $W^{1,p}(I, \mathbb{R}^n)$ sogar ein Banach-Raum für $1 < p < \infty$.

Als nächstes wollen wir Bedingungen an ein (der Einfachheit halber auf ganz $I \times \mathbb{R}^n$ definiertes) zeitabhängiges Vektorfeld formulieren, die garantieren, dass $t \mapsto f(t, x(t))$ für eine stetige Kurve $x \in C(I, \mathbb{R}^n)$ messbar oder sogar L^p -integrierbar ist.

Definition 1.48. *Eine Abbildung $f: I \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ heißt Carathéodory-Vektorfeld, falls*

- $t \mapsto f(t, x)$ messbar für jedes feste $x \in \mathbb{R}^n$ ist,
- $x \mapsto f(t, x)$ stetig für fast alle $t \in I$ ist.

Lemma 1.49. *Ist $f: I \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ ein Carathéodory-Vektorfeld und gilt mit einem $1 < p < \infty$ zusätzlich die Wachstumsbedingung, dass es*

- zu jedem kompakten Teilintervall $[a, b] \subset I$ und jedem $r < \infty$ eine Funktion $\gamma_r \in L^p([a, b])$ mit $|f(t, x)| \leq \gamma_r(t)$ für alle $x \in \mathbb{R}^n$ mit $|x| \leq r$ und fast alle $t \in [a, b]$ gibt,

dann liegt $t \mapsto f(t, x(t))$ für jede Kurve $x \in C(I, \mathbb{R}^n)$ in $L^p_{loc}(I, \mathbb{R}^n)$.

Beweis: Zunächst zeigen wir, dass das durch $y(t) := f(t, x(t))$ definierte $y : I \rightarrow \mathbb{R}^n$ meßbar ist. Dazu beobachte man, dass x stetig und daher meßbar ist, d.h. x ist fast überall auf I der punktweise Limes von Treppenfunktionen x_k . Da f ein Carathéodory-Vektorfeld ist, sind die Funktionen $t \mapsto f(t, x_k(t))$ meßbar und es gilt $y(t) = f(t, x(t)) = \lim_{k \rightarrow \infty} f(t, x_k(t))$ für fast alle $t \in I$ aufgrund der Stetigkeitsbedingung an f . Also ist y fast überall auf I der punktweise Limes von meßbaren Funktionen und somit selbst meßbar.

Desweiteren sei $[a, b]$ ein kompaktes Teilintervall von I und $r := \max_{t \in [a, b]} |x(t)|$, dann gibt es aufgrund der Wachstumsbedingung ein $\gamma_r \in L^p([a, b])$, mit dem die Abschätzung $|y(t)| \leq \gamma_r(t)$ für fast alle $t \in [a, b]$ gilt. Also ist y nach dem Satz von Lebesgue über majorisierte Konvergenz selbst L^p -integrierbar über $[a, b]$ und es gilt

$$\int_a^b |y(t)|^p dt \leq \int_a^b \gamma_r(t) dt.$$

□

Die in Lemma 1.49 genannte allgemeine Wachstumsbedingung kann man natürlich durch speziellere, leichter nachprüfbare Bedingungen ersetzen, z.B. dadurch dass

- $|f(t, x)| \leq C(|x|)\gamma(t)$ mit einem monoton wachsenden $C : [0, \infty) \rightarrow [0, \infty)$ und $\gamma \in L^p(I)$ gilt.

Wiederum liegt es nahe, Existenz von schwachen Lösungen durch Anwendung des Banachschen Fixpunktsatzes zu beweisen.

Satz 1.50. *Genügt das Carathéodory-Vektorfeld $f : I \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit einem $1 < p < \infty$ der in Lemma 1.49 genannten L^p -Wachstumsbedingung und zusätzlich der L^p -Lipschitzbedingung, dass*

- zu jedem kompakten Teilintervall $[a, b] \subset I$ und jedem $r < \infty$ eine Funktion $L_r \in L^p([a, b])$ mit $|f(t, x) - f(t, \tilde{x})| \leq L_r(t)|x - \tilde{x}|$ für alle $x, \tilde{x} \in \mathbb{R}^n$ mit $|x|, |\tilde{x}| \leq r$ und fast alle $t \in [a, b]$ existiert,

dann gibt es zu jedem Anfangswert $(t_0, x_0) \in I \times \mathbb{R}^n$ ein $\varepsilon > 0$ und eine eindeutige schwache Lösung $x \in W^{1,p}([t_0 - \varepsilon, t_0 + \varepsilon], \mathbb{R}^n)$ von (1.20).

Beweis: Wie im Beweis des Satzes 1.32 von Picard-Lindelöf ist $x \in W^{1,p}([t_0 - \varepsilon, t_0 + \varepsilon], \mathbb{R}^n)$ genau dann eine Lösung von (1.20), wenn x ein Fixpunkt des in (1.21) definierten Operators T auf $C([t_0 - \varepsilon, t_0 + \varepsilon], \mathbb{R}^n)$ ist.

Wähle nun zum Anfangswert (t_0, x_0) ein kompaktes Teilintervall $[a, b] \subset I$ mit $t_0 \in (a, b)$ und ein $r < \infty$, dann gibt es aufgrund der Wachstumsbedingung ein $\gamma_r \in L^p([a, b])$ mit $|f(t, x)| \leq \gamma_r(t)$ für alle $x \in \overline{B_r(x_0)}$ und fast alle $t \in [a, b]$, und aufgrund der Lipschitzbedingung ein $L_r \in L^p([a, b])$ mit $|f(t, x) - f(t, \tilde{x})| \leq L_r(t)|x - \tilde{x}|$ für alle $x, \tilde{x} \in B_r(x_0)$ und fast alle $t \in [a, b]$. Sei $p' := \frac{p}{p-1}$ der duale Exponenten zu p , dann wähle $\varepsilon > 0$ so klein, dass $[t_0 - \varepsilon, t_0 + \varepsilon] \subset [a, b]$, $\left(\int_a^b |\gamma_r(s)|^p ds\right) \cdot \varepsilon^{p'} \leq r$ und $\left(\int_a^b |L(s)|^p ds\right) \varepsilon^{p'} \leq \frac{1}{2}$ gilt.

Durch diese Wahl von ε gilt für beliebige stetige Kurven $x, \tilde{x} : [t_0 - \varepsilon, t_0 + \varepsilon] \rightarrow \overline{B_r(x_0)}$ sowohl

$$|(Tx)(t) - x_0| \leq \int_{t_0}^t |f(s, x(s))| ds \leq \left(\int_a^b |\gamma_r(s)|^p ds\right) \cdot |t - t_0|^{p'} \leq r$$

für alle $t \in [t_0 - \varepsilon, t_0 + \varepsilon]$ als auch

$$\begin{aligned}
|(Tx)(t) - (T\tilde{x})(t)| &\leq \int_{t_0}^t |f(s, x(s)) - f(s, \tilde{x}(s))| ds \\
&\leq \int_{t_0}^t L(s) |x(s) - \tilde{x}(s)| ds \\
&\leq \left(\int_a^b |L(s)|^p ds \right) |t - t_0|^{p'} \max_{|s-t_0| \leq \varepsilon} |x(s) - \tilde{x}(s)| \\
&\leq \frac{1}{2} \max_{|s-t_0| \leq \varepsilon} |x(s) - \tilde{x}(s)|.
\end{aligned}$$

Also ist T eine Selbstabbildung auf $C([t_0 - \varepsilon, t_0 + \varepsilon], \overline{B_r(x_0)})$ und eine Kontraktion, so dass der Banachsche Fixpunktsatz 1.30 einen eindeutigen Fixpunkt x von T und somit eine eindeutige lokale Lösung $x \in W^{1,p}([t_0 - \varepsilon, t_0 + \varepsilon], \mathbb{R}^n)$ des Anfangswertproblems (1.20) liefert. \square

Analog zum Satz von Peano kann auch hier auf die L^p -Lipschitzbedingung verzichtet werden, dann erhält man zwar immer noch Existenz, aber nicht unbedingt Eindeutigkeit schwacher Lösungen.

1.5.3 Eindeutigkeit bei monotonen Vektorfeldern

Um die Eindeutigkeit von Lösungen bei Differentialgleichungen mit stetigen (oder Carathéodory-) Vektorfeldern zu zeigen, die keiner der bisher von uns formulierten Lipschitzbedingungen genügen, ist die folgende Definition hilfreich, in der $\langle \cdot, \cdot \rangle$ das Euklidische Skalarprodukt auf dem \mathbb{R}^n bezeichnet.

Definition 1.51. *Man sagt, ein zeitabhängiges Vektorfeld $f : I \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ genügt einer (globalen) einseitigen Lipschitzbedingung, falls es ein $L < \infty$ mit*

$$\langle f(t, x) - f(t, \tilde{x}), x - \tilde{x} \rangle \leq L|x - \tilde{x}|^2 \quad (1.24)$$

für alle $x, \tilde{x} \in \mathbb{R}^n$ und fast alle $t \in I$ gibt.

Satz 1.52. *Genügt die rechte Seite f der Differentialgleichung $x' = f(t, x)$ einer einseitigen Lipschitzbedingung, dann gibt es zu jedem Anfangswert vorwärts in der Zeit höchstens eine schwache Lösung.*

Beweis: Sind x, y zwei klassische (oder schwache) Lösungen auf dem Intervall I mit $x(t_0) = y(t_0)$ für $t_0 \in I$, dann gilt $x'(t) = f(t, x(t))$ und $y'(t) = f(t, y(t))$, also

$$(x - y)'(t) = f(t, x(t)) - f(t, y(t)).$$

Multipliziert man diese Gleichung mit der Differenz $x - y$, so erhält man die Differentialungleichung

$$\left(\frac{1}{2}|x - y|^2\right)'(t) = \langle f(t, x(t)) - f(t, y(t)), x(t) - y(t) \rangle \leq 2L\left(\frac{1}{2}|x - y|^2\right)(t).$$

Mittels des Lemmas 1.53 von Gronwall folgt $\frac{1}{2}|x - y|^2(t) \leq \frac{1}{2}|x - y|^2(t_0)e^{2L(t-t_0)} \leq 0$ für (fast) alle $t \in I \cap [t_0, \infty)$ wegen $x(t_0) = y(t_0)$, also gilt $x(t) = y(t)$ für (fast) alle $t \in I \cap [t_0, \infty)$. \square

Es ist noch das Lemma von Gronwall zu zeigen. Dieses besagt insbesondere, dass für eine Funktion u aus der Gültigkeit der Differentialungleichung $u' \leq \beta u$ auf einem Intervall I mit linkem Randpunkt $t_0 \in I$ bei $u(t_0) = 0$ oder der äquivalenten Integralungleichung $u(t) \leq \beta \int_{t_0}^t u(s) ds$ mit $\beta \geq 0$ schon $u \leq 0$ auf I gefolgert werden kann.

Lemma 1.53. *Ist $u : I \rightarrow \mathbb{R}$ eine stetige Funktion auf einem Intervall I mit linkem Randpunkt $t_0 \in I$ und gilt für alle $t \in I$ die Abschätzung*

$$u(t) \leq \alpha + \beta \int_{t_0}^t u(s) ds$$

mit zwei Konstanten $\alpha, \beta \geq 0$, dann gilt für alle $t \in I$ die Ungleichung

$$u(t) \leq \alpha e^{\beta(t-t_0)}.$$

Beweis: Nur der Fall $\beta > 0$ ist nichttrivial. Mit $v(t) := \int_{t_0}^t u(s) ds$ gilt nach Voraussetzung $v'(t) = u(t) \leq \alpha + \beta v(t)$. Setzt man $w(t) := e^{-\beta(t-t_0)}v(t)$, so gilt $w(t_0) = 0$ sowie

$$w'(t) = e^{-\beta(t-t_0)}(v'(t) - \beta v(t)) \leq \alpha e^{-\beta(t-t_0)}.$$

Also ergibt sich durch Integration von t_0 bis t

$$w(t) \leq \frac{\alpha}{\beta}(1 - e^{-\beta(t-t_0)}).$$

Wegen $w(t) = e^{-\beta(t-t_0)}v(t)$ folgt daraus $v(t) \leq \frac{\alpha}{\beta}(e^{\beta(t-t_0)} - 1)$. Somit gilt die gewünschte Ungleichung

$$u(t) \leq \alpha + \beta v(t) \leq \alpha + \alpha(e^{\beta(t-t_0)} - 1) = \alpha e^{\beta(t-t_0)}.$$

□

Setzt man im Lemma von Gronwall stattdessen voraus, dass u auf einem Intervall I nichtnegativ ist (wobei nun t_0 ein innerer Punkt von I sein darf) und der Ungleichung $u(t) \leq \alpha + \beta \left| \int_{t_0}^t u(s) ds \right|$ genügt, dann kann man daraus sogar $u(t) \leq \alpha e^{\beta|t-t_0|}$ folgern.

Man könnte die Bedingung (1.24) auch als eine Monotonie-Bedingung bezeichnen, denn tatsächlich ist sie zur Monotonie von $L \text{Id} - f$ äquivalent.

Definition 1.54. *Ein zeitabhängiges Vektorfeld $g : I \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ heißt monoton, falls $\langle g(t, x) - g(t, \tilde{x}), x - \tilde{x} \rangle \geq 0$ für alle $x, \tilde{x} \in \mathbb{R}^n$ und fast alle $t \in I$ gilt.*

Monotone Vektorfelder spielen in den Anwendungen eine große Rolle. Beispielsweise sind Gradientenvektorfelder $g = \text{grad } V$ zu konvexen C^1 -Funktionen $V : \mathbb{R}^n \rightarrow \mathbb{R}$ monoton, denn die Bedingung

$$\langle (\text{grad } V)(x) - (\text{grad } V)(\tilde{x}), x - \tilde{x} \rangle \geq 0$$

für alle $x, \tilde{x} \in \mathbb{R}^n$ ist äquivalent zur Konvexität von V .

Korollar 1.55. *Ist g monoton, dann besitzt die Differentialgleichung $x' + g(t, x) = 0$ zu jedem Anfangswert vorwärts in der Zeit höchstens eine Lösung.*

Beweis: Das zeitabhängige Vektorfeld $f := -g$ genügt offensichtlich für jedes $L \geq 0$ der einseitigen Lipschitzbedingung

$$\langle f(t, x) - f(t, y), x - y \rangle \leq 0 \leq L|x - y|^2.$$

□

Beispiel 1.56. *Für jedes $1 < p < 2$ sind die Lösungen der Differentialgleichung $x' = -\text{sign}(x)|x|^{p-1}$ eindeutig, obwohl die rechte Seite nicht Lipschitz-stetig ist.*

1.6 Stetige Abhängigkeit

Bei (globaler) Existenz und Eindeutigkeit, also z.B. für global Lipschitz-stetiges $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$, generiert eine autonome Differentialgleichung $x' = f(x)$ nach Satz 1.13 einen Fluß Φ . Ob die Abbildungen $x \mapsto \Phi(t, x)$ aber stetig sind, ist nicht offensichtlich. Dies ist gerade dann der Fall, wenn die Lösungen der Differentialgleichung stetig von ihrem Anfangswert abhängen.

Wir wollen in diesem Abschnitt Aussagen beweisen, die die stetige Abhängigkeit sowohl vom Anfangswert als auch von der rechten Seite garantieren. Eine solche Aussage wird im folgenden Satz formuliert.

Satz 1.57. *Sei $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ offen, sei $f: \Omega \rightarrow \mathbb{R}^n$ und sei $x: [a, b] \rightarrow \mathbb{R}^n$ Lösung des Anfangswertproblems $x' = f(t, x)$, $x(t_0) = x_0$. Ist f auf einer kompakten Umgebung $K \subset [a, b] \times \mathbb{R}^n$ des Graphen von x stetig und genügt f auf K einer (globalen) Lipschitzbedingung, dann hängt die Lösung x in folgendem Sinne stetig vom Anfangswert und von der rechten Seite ab:*

Zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$ mit der Eigenschaft, dass bei

$$|t_0 - \tilde{t}_0| \leq \delta, |x_0 - \tilde{x}_0| \leq \delta \text{ und } |f(t, x) - \tilde{f}(t, x)| \leq \delta \text{ für alle } (t, x) \in K$$

mit einem stetigen $\tilde{f}: K \rightarrow \mathbb{R}^n$, das einer (globalen) Lipschitzbedingung genügt, die Lösung \tilde{x} des Anfangswertproblems $\tilde{x}' = \tilde{f}(t, \tilde{x})$, $\tilde{x}(\tilde{t}_0) = \tilde{x}_0$, auf ganz $[a, b]$ definiert ist und $|x(t) - \tilde{x}(t)| \leq \varepsilon$ für alle $t \in [a, b]$ gilt.

Beweis: Aufgrund der Stetigkeit von f und der Kompaktheit von K gibt es ein $M < \infty$ mit $|f(t, x)| \leq M$ für alle $(t, x) \in K$. Da \tilde{f} auf K einer globalen Lipschitzbedingung genügt, gibt es ein $L < \infty$ mit

$$|\tilde{f}(t, y) - \tilde{f}(t, \tilde{y})| \leq L|y - \tilde{y}|.$$

Außerdem ist K eine Umgebung von (t_0, x_0) , also kann man $\delta > 0$ so klein wählen, dass aus $|t_0 - \tilde{t}_0| < \delta$ und $|x_0 - \tilde{x}_0| < \delta$ schon $(\tilde{t}_0, \tilde{x}_0) \in K$ folgt. Dann verläuft \tilde{x} zumindest schon einmal für Zeiten nahe \tilde{t}_0 innerhalb von K .

Die Lösungen x bzw. \tilde{x} der Differentialgleichungen $x' = f(t, x)$ bzw. $\tilde{x}' = \tilde{f}(t, \tilde{x})$ erfüllen natürlich auch die zugehörigen Integralgleichungen

$$x(t) = x_0 + \int_{t_0}^t f(s, x(s)) ds \text{ bzw. } \tilde{x}(t) = \tilde{x}_0 + \int_{\tilde{t}_0}^t \tilde{f}(s, \tilde{x}(s)) ds.$$

Daraus ergibt sich zumindest für die Zeiten, innerhalb derer \tilde{x} noch in K verläuft, die Abschätzung

$$\begin{aligned} |x(t) - \tilde{x}(t)| &\leq |x_0 - \tilde{x}_0| + \left| \int_{t_0}^t f(s, x(s)) ds - \int_{\tilde{t}_0}^t \tilde{f}(s, \tilde{x}(s)) ds \right| \leq \\ &|x_0 - \tilde{x}_0| + \left| \int_{\tilde{t}_0}^{t_0} f(s, x(s)) ds \right| + \left| \int_{\tilde{t}_0}^t f(s, x(s)) - \tilde{f}(s, \tilde{x}(s)) ds \right| \leq \\ &|x_0 - \tilde{x}_0| + M|t_0 - \tilde{t}_0| + \left| \int_{\tilde{t}_0}^t f(s, x(s)) - \tilde{f}(s, x(s)) ds \right| + \left| \int_{\tilde{t}_0}^t \tilde{f}(s, x(s)) - \tilde{f}(s, \tilde{x}(s)) ds \right| \leq \\ &|x_0 - \tilde{x}_0| + M|t_0 - \tilde{t}_0| + \delta|t - \tilde{t}_0| + L \int_{\tilde{t}_0}^t |x(s) - \tilde{x}(s)| ds. \end{aligned}$$

Die ersten drei Summanden kann man durch $(1 + M + \max(b - \tilde{t}_0, \tilde{t}_0 - a))\delta$ abschätzen, und nach dem Lemma 1.53 von Gronwall nachfolgenden Bemerkung gilt daher

$$|x(t) - \tilde{x}(t)| \leq (1 + M + \max(b - \tilde{t}_0, \tilde{t}_0 - a))\delta e^{L|t - \tilde{t}_0|}$$

Wählt man $\delta > 0$ also zusätzlich so klein, dass $(1 + M + \max(b - \tilde{t}_0, \tilde{t}_0 - a))\delta e^{\delta \max(b - \tilde{t}_0, \tilde{t}_0 - a)}$ nicht nur kleiner als ε , sondern auch kleiner als $\text{dist}(\partial K, \text{Graph}(x))$ ist, dann gilt einerseits die gewünschte Abschätzung $|x(t) - \tilde{x}(t)| \leq \varepsilon$ und andererseits verläuft \tilde{x} für alle Zeiten $t \in [a, b]$ innerhalb von K , denn $|x(t) - \tilde{x}(t)| < \text{dist}(\partial K, \text{Graph}(x))$. \square

Bezeichnen wir die Lösung von $x' = f(t, x)$ zum Anfangswert $x(t_0) = x_0$ auf $[a, b]$ mit $x(t; t_0, x_0, f)$, dann besagt Satz 1.57 nichts anderes, als dass $(t_0, x_0, f) \mapsto (t \mapsto x(t; t_0, x_0, f))$ als Abbildung von $\Omega \times \text{Lip}(\Omega, \mathbb{R}^n) \subset \Omega \times C(\Omega, \mathbb{R}^n)$ nach $C([a, b], \mathbb{R}^n)$ stetig ist. Dabei bezeichnet $\text{Lip}(\Omega, \mathbb{R}^n)$ die einer Lipschitzbedingung genügenden $f \in C(\Omega, \mathbb{R}^n)$ und Konvergenz in $C(\Omega, \mathbb{R}^n)$ bedeutet gleichmäßige Konvergenz auf kompakten Teilmengen (während $C([a, b], \mathbb{R}^n)$ wie üblich mit der gleichmäßigen Konvergenz auf $[a, b]$ versehen ist).

Definition 1.58. *Man sagt, f_n konvergiert gleichmäßig auf kompakten Teilmengen von Ω gegen f , falls es für jedes $\varepsilon > 0$ und jede kompakte Teilmenge $K \subset \Omega$ ein $N \in \mathbb{N}$ gibt mit $|f - f_n| \leq \varepsilon$ auf K für alle $n \geq N$.*

Die zu diesem Konvergenzbegriff gehörige Topologie auf $C(\Omega, \mathbb{R}^n)$ ist die kompakt-offene Topologie, deren offene Teilmengen von $\{f \in C(\Omega, \mathbb{R}^n) \mid f(K) \subset U\}$ für kompakte $K \subset \Omega$ und offene $U \subset \mathbb{R}^n$ erzeugt werden.

Satz 1.57 ist in der Hinsicht unbefriedigend, dass die Gültigkeit einer Lipschitzbedingung essentiell in den Beweis eingeht und insbesondere $(t_0, x_0, f) \mapsto (t \mapsto x(t; t_0, x_0, f))$ nur als Abbildung auf $\Omega \times \text{Lip}(\Omega, \mathbb{R}^n)$ und nicht auf ganz $\Omega \times C(\Omega, \mathbb{R}^n)$ stetig ist. Der folgende Satz behebt dieses Problem und ist insbesondere dann anwendbar, wenn f nur stetig ist, zum Anfangswert (t_0, x_0) aber trotzdem eine eindeutige Lösung existiert.

Satz 1.59. *Sei $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ offen, konvergiere $f_n \in C(\Omega, \mathbb{R}^n)$ gleichmäßig auf kompakten Teilmengen von Ω gegen $f \in C(\Omega, \mathbb{R}^n)$ und gelte $(t_{0n}, x_{0n}) \rightarrow (t_0, x_0)$ in Ω für $n \rightarrow \infty$. Bezeichne desweiteren x_n für jedes $n \in \mathbb{N}$ eine nicht weiter fortsetzbare Lösung von $x' = f_n(t, x)$, $x_n(t_{0n}) = x_{0n}$. Dann gilt:*

Existiert eine eindeutige Lösung x des Anfangswertproblems $x' = f(t, x)$, $x(t_0) = x_0$, auf dem kompakten Intervall $[a, b]$, dann existiert für genügend große n auch x_n auf $[a, b]$ und x_n konvergiert gleichmäßig auf $[a, b]$ gegen x .

Beweis: Wir zeigen o.B.d.A. nur die gleichmäßige Konvergenz von x_n gegen x auf dem Intervall $[t_0, b]$.

Da der Graph von x über $[a, b]$ kompakt ist, gibt es eine kompakte Teilmenge K von Ω , in deren Inneren der Graph von x über $[a, b]$ verläuft. Aufgrund der Stetigkeit von f und der Kompaktheit von K gibt es ein $M < \infty$ mit $|f(t, x)| < M$ für alle $(t, x) \in K$. Darüberhinaus gibt es aufgrund der gleichmäßigen Konvergenz von f_n gegen f auf K auch ein $\tilde{N} \in \mathbb{N}$ mit $|f_n(t, x)| < M$ für alle $n \geq \tilde{N}$. Wähle nun $\varepsilon > 0$ so klein, dass $[t_0, t_0 + \varepsilon] \subset [a, b]$ sowie $[t_0 - \varepsilon, t_0 + \varepsilon] \times B_{3M\varepsilon}(x_0) \subset K$ gilt, und $N \geq \tilde{N}$ so groß, dass $|t_{0n} - t_0| < \varepsilon$ sowie $|x_{0n} - x_0| < M\varepsilon$ für alle $n \geq N$ gilt.

Dann gilt für $n \geq N$ und $t \in [t_0, t_0 + \varepsilon]$ zumindest solange, wie x_n innerhalb von $[t_0 - \varepsilon, t_0 + \varepsilon] \times B_{3M\varepsilon}(x_0)$ verläuft, die Abschätzung

$$|x_n(t) - x_{0n}| \leq \left| \int_{t_{0n}}^t f_n(s, x_n(s)) \, ds \right| \leq M|t - t_{0n}|$$

und daher

$$|x_n(t) - x_0| \leq |x_n(t) - x_{0n}| + |x_{0n} - x_0| \leq M(|t - t_{0n}| + \varepsilon) \leq M(|t - t_0| + |t_0 - t_{0n}| + \varepsilon) \leq 3M\varepsilon.$$

Somit ist x_n für $n \geq N$ tatsächlich auf ganz $[t_0, t_0 + \varepsilon]$ definiert und verläuft innerhalb von $[t_0 - \varepsilon, t_0 + \varepsilon] \times B_{3M\varepsilon}(x_0)$. Insbesondere ist x_n gleichmäßig beschränkt auf $[t_0, t_0 + \varepsilon]$ und wegen

$$|x_n(t) - x_n(\tilde{t})| \leq \left| \int_{\tilde{t}}^t f_n(s, x_n(s)) \, ds \right| \leq M|t - \tilde{t}|$$

auch gleichgradig stetig. Nach dem Satz 1.42 von Arzela-Ascoli gibt es also eine Teilfolge x_{n_k} von x_n , die gleichmäßig auf $[t_0, t_0 + \varepsilon]$ gegen eine stetige Kurve \tilde{x} konvergiert, und Limes-Bildung $k \rightarrow \infty$ in

$$x_{n_k}(t) = x_{0n} + \int_{t_{0n}}^t f_{n_k}(s, x_{n_k}(s)) \, ds$$

liefert $\tilde{x}(t) = x_0 + \int_{t_0}^t f(s, \tilde{x}(s)) \, ds$. Daher löst \tilde{x} das Anfangswertproblem $x' = f(t, x)$, $x(t_0) = x_0$, und stimmt also aufgrund der Eindeutigkeit auf $[t_0, t_0 + \varepsilon]$ mit x überein. Dies gilt auch für den Grenzwert jeder anderen auf $[t_0, t_0 + \varepsilon]$ gleichmäßig konvergenten Teilfolge von x_n . Würde nun nicht die gesamte Folge x_n gleichmäßig auf $[t_0, t_0 + \varepsilon]$ gegen x konvergieren, dann gäbe es ein $\tilde{\varepsilon} > 0$ und eine Teilfolge x_{n_k} von x_n mit $\max_{[t_0, t_0 + \varepsilon]} |x_{n_k} - x| \geq \tilde{\varepsilon}$, im Widerspruch dazu würde aber eine Teilfolge dieser Teilfolge wieder gleichmäßig auf $[t_0, t_0 + \varepsilon]$ gegen x konvergieren. Also konvergiert schon die gesamte Folge x_n gleichmäßig auf $[t_0, t_0 + \varepsilon]$ gegen x .

Nun müssen wir dieses Resultat nur noch auf das gesamte Intervall $[t_0, b]$ ausdehnen. Sei dazu

$$\tilde{t} := \sup\{t \geq t_0 \mid x_n \rightarrow x \text{ gleichmäßig auf } [t_0, t]\},$$

dann müssen wir $\tilde{t} = b$ zeigen. Angenommen, es wäre $\tilde{t} < b$, dann können wir immer noch ein $\tilde{\varepsilon}$ finden mit $[\tilde{t}, \tilde{t} + \tilde{\varepsilon}] \subset [a, b]$ sowie $[\tilde{t} - 2\tilde{\varepsilon}, \tilde{t} + 2\tilde{\varepsilon}] \times B_{4M\tilde{\varepsilon}}(x(\tilde{t})) \subset K$. Außerdem können wir ein \tilde{t}_0 mit $\tilde{t} - \tilde{\varepsilon} < \tilde{t}_0 < \tilde{t}$ finden. Setze $\tilde{x}_0 := x(\tilde{t}_0)$, dann gilt

$$|x(\tilde{t}) - \tilde{x}_0| \leq \left| \int_{\tilde{t}_0}^{\tilde{t}} f(s, x(s)) \, ds \right| \leq M\tilde{\varepsilon}$$

und daher

$$[\tilde{t}_0 - \tilde{\varepsilon}, \tilde{t}_0 + \tilde{\varepsilon}] \times B_{3M\tilde{\varepsilon}}(\tilde{x}_0) \subset [t_0 - 2\tilde{\varepsilon}, t_0 + 2\tilde{\varepsilon}] \times B_{4M\tilde{\varepsilon}}(x(\tilde{t})).$$

Ersetze nun im ersten Teil dieses Beweises t_0 und t_{0n} durch \tilde{t}_0 , x_{0n} durch $x_n(\tilde{t}_0)$ und ε durch $\tilde{\varepsilon}$, dann erhält man die gleichmäßige Konvergenz von x_n gegen x auf $[\tilde{t}_0, \tilde{t}_0 + \tilde{\varepsilon}]$, aber nach Definition von \tilde{t}_0 gilt $\tilde{t}_0 + \tilde{\varepsilon} > \tilde{t}$, im Widerspruch zur Definition von \tilde{t} . Somit ist $\tilde{t} = b$ und x_n konvergiert auf ganz $[t_0, b]$ gleichmäßig gegen x . \square

Als Konsequenz aus Satz 1.59 ist $(t_0, x_0, f) \mapsto (t \mapsto x(t; t_0, x_0, f))$ als Abbildung von $\Omega \times C(\Omega, \mathbb{R}^n)$ nach $C([a, b], \mathbb{R}^n)$ in den Punkten (t_0, x_0, f) stetig, für die auf $[a, b]$ eine eindeutige Lösung des Anfangswertproblems $x' = f(t, x)$, $x(t_0) = x_0$, existiert. Man bemerke, dass auf die Eindeutigkeit nicht verzichtet werden kann, denn existieren verschiedene Lösungen zum selben Anfangswert, dann liegt selbstverständlich keine stetige Abhängigkeit vor. Das folgende Korollar garantiert abschließend die stetige Abhängigkeit von Parametern.

Korollar 1.60. *Sei $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ offen, Λ ein metrischer Raum und $f : \Omega \times \Lambda \rightarrow \mathbb{R}^n$ ein stetiges zeit- und parameterabhängiges Vektorfeld. Existiert auf $[a, b]$ eine eindeutige Lösung $x(t; t_0, x_0, \lambda_0)$ von $x' = f(t, x, \lambda_0)$ zum Anfangswert $x(t_0) = x_0$, dann existieren für jedes (s, ξ, λ) nahe (t_0, x_0, λ) Lösungen $x(t; s, \xi, \lambda)$ auf $[a, b]$ von $x' = f(t, x, \lambda)$ zum Anfangswert $x(s) = \xi$, und $(s, \xi, \lambda) \mapsto (t \mapsto x(t; s, \xi, \lambda))$ ist als Abbildung von $\Omega \times \Lambda$ nach $C([a, b], \mathbb{R}^n)$ stetig in (t_0, x_0, λ_0) .*

Beweis: Wir haben nur zu zeigen: Ist die Abbildung $\Omega \times \Lambda \ni (t, x, \lambda) \mapsto f(t, x, \lambda) \in \mathbb{R}^n$ stetig, dann ist auch

$$\Lambda \ni \lambda \mapsto ((t, x) \mapsto f(t, x, \lambda)) \in C(\Omega, \mathbb{R}^n)$$

stetig bzgl. der Topologie der gleichmäßigen Konvergenz auf kompakten Teilmengen von Ω .

Diese Aussage ist wahr, denn ist $\lambda_0 \in \Lambda$ und K eine kompakte Teilmenge von Ω , dann gibt es aufgrund der Stetigkeit von f zu $\varepsilon > 0$ und jedem $(t_0, x_0) \in K$ Umgebungen $U_{(t_0, x_0)} \subset \Omega$ von (t_0, x_0) und $V_{(t_0, x_0)} \subset \Lambda$ von λ_0 mit $|f(t, x, \lambda) - f(t, x, \lambda_0)| < \varepsilon$ für alle $(t, x) \in U_{(t_0, x_0)}$ und $\lambda \in V_{(t_0, x_0)}$. Nun ist K kompakt, also reichen endlich viele Umgebungen $U_{(t_i, x_i)}$, $i = 1, \dots, N$, aus, um K zu überdecken. Mit dem Durchschnitt $V := \bigcap_{i=1}^N V_{(t_i, x_i)}$ der zugehörigen Umgebungen $V_{(t_i, x_i)} \subset \Lambda$ von λ_0 gilt dann aber $|f(t, x, \lambda) - f(t, x, \lambda_0)| < \varepsilon$ für alle $(t, x) \in K$ und $\lambda \in V$, was zu zeigen war. \square

1.7 Differenzierbare Abhängigkeit

Nachdem wir die stetige Abhängigkeit vom Anfangswert und von Parametern zeigen konnten, stellt sich als nächstes die Frage, ob die Abbildung $(t_0, x_0, \lambda_0) \mapsto (t \mapsto x(t; t_0, x_0, \lambda_0))$ nach t_0 , x_0 und λ_0 differenzierbar ist, wobei $x(t; t_0, x_0, \lambda_0)$ die Lösung der parameterabhängigen Differentialgleichung $x' = f(t, x, \lambda_0)$ zum Anfangswert $x(t_0) = x_0$ bezeichnet, $\Lambda \subset \mathbb{R}^k$ eine offene Teilmenge ist und angenommen wird, dass $f : \Omega \times \Lambda \rightarrow \mathbb{R}^n$ stetig differenzierbar ist.

Zunächst wollen wir den speziellen Fall der differenzierbaren Abhängigkeit vom Anfangswert x_0 klären. Sei dazu $f = f(x)$ ein zeit- und parameterunabhängiges Vektorfeld und bezeichne $x(t; x_0)$ die Lösung der autonomen Differentialgleichung $x' = f(x)$ zum Anfangswert $x(0) = x_0$ auf dem Intervall $[a, b]$. Wüssten wir schon, dass sowohl $x(t; x_0)$ als auch $\frac{\partial x}{\partial t}(t; x_0)$ nach x_0 differenzierbar ist, dann könnte man beide Seiten der Gleichung $\frac{\partial x}{\partial t}(t; x_0) = f(t, x(t; x_0))$ nach x_0 ableiten und auf der linken Seite die Differentiationsreihenfolge vertauschen. Daraus würde sich die Differentialgleichung

$$Y' = \frac{\partial f}{\partial x}(t, x(t; x_0))Y \tag{1.25}$$

für die Matrixwertige Kurve $Y(t) := \frac{\partial x}{\partial x_0}(t; x_0)$ ergeben. Diese Differentialgleichung wird die Variationsgleichung zu $x' = f(x)$ genannt und ist ein n -dimensionales homogenes lineares System. Über homogene lineare Systeme von Differentialgleichungen $y' = A(t)y$ weiß man durch Satz 2.3 (auf den wir hier schon einmal vorgreifen) aber sehr gut Bescheid. Insbesondere gibt es ein ausgezeichnetes Fundamentalsystem Y auf ganz I , das $Y' = A(t)Y$ zum Anfangswert $Y(0) = \text{Id}$ löst, und hängt A zusätzlich stetig von Parametern ab (wie im Fall der Variationsgleichung (1.25), wo $A(t; x_0) := \frac{\partial f}{\partial x}(t, x(t; x_0))$ von x_0 stetig abhängt), dann hängt nach unseren Erkenntnissen über stetige Abhängigkeit auch Y von diesen Parametern stetig ab. Dies deutet schon darauf hin, dass die Variationsgleichung (1.25) ein nützliches Hilfsmittel beim Beweis der differenzierbaren Abhängigkeit der Lösung $x(t; x_0)$ vom Anfangswert x_0 sein wird.

Satz 1.61. *Sei $\Omega \subset \mathbb{R}^n$ offen, sei $x_0 \in \Omega$, sei $f : \Omega \rightarrow \mathbb{R}^n$ stetig differenzierbar und existiere die Lösung $x(\cdot; x_0)$ von $x' = f(x)$ zum Anfangswert $x(0) = x_0$ auf dem Intervall $[a, b]$. Dann existiert die Ableitung $\frac{\partial x}{\partial x_0}(t; x_0)$ für alle $t \in [a, b]$ und löst die Variationsgleichung (1.25) zum Anfangswert $Y(0) = \text{Id}$.*

Beweis: Zunächst einmal existieren nach unseren Erkenntnissen über stetige Abhängigkeit eindeutige Lösungen $x(t; \tilde{x}_0)$ von $x' = f(x)$ zum Anfangswert $x(0) = \tilde{x}_0$ auf ganz $[a, b]$, falls \tilde{x}_0 nahe genug an x_0 liegt, und $\tilde{x}_0 \mapsto x(\cdot; \tilde{x}_0)$ ist stetig.

Um für jedes $t \in [a, b]$ die Differenzierbarkeit von $\tilde{x}_0 \mapsto x(t; \tilde{x}_0)$ in $\tilde{x}_0 = x_0$ zu zeigen, betrachte die Kurve

$$y(t; x_0, h) := x(t; x_0 + h) - x(t; x_0)$$

und die zeit- und parameterabhängige Matrix

$$A(t; x_0, h) := \int_0^1 \frac{\partial f}{\partial x}(t, x(t; x_0) + sy(t; x_0, h)) ds.$$

Dann gilt nach dem Mittelwertsatz

$$\begin{aligned} \frac{\partial y}{\partial t}(t; x_0, h) &= \frac{\partial x}{\partial t}(t; x_0 + h) - \frac{\partial x}{\partial t}(t; x_0) = \\ f(t, x(t; x_0 + h)) - f(t, x(t; x_0)) &= A(t; x_0, h)y(t; x_0, h). \end{aligned}$$

Also löst $y(t; x_0, h)$ das homogene lineare System $y' = A(t; x_0, h)y$ zum Anfangswert $y(0; x_0, h) = x_0 + h - x_0 = h$ bei $t = 0$. Daher kann man y mit dem ausgezeichneten und von all seinen Parametern stetig abhängendem Fundamentalsystem $Y(t; x_0, h)$ von $Y' = A(t; x_0, h)Y$, $Y(0) = \text{Id}$, als $y(t; x_0, h) = Y(t; x_0, h)h$ schreiben. Daraus ergibt sich

$$\begin{aligned} |x(t; x_0 + h) - x(t; x_0) - Y(t; x_0, 0)h| &= |y(t; x_0, h) - Y(t; x_0, 0)h| = \\ |Y(t; x_0, h)h - Y(t; x_0, 0)h| &\leq |Y(t; x_0, h) - Y(t; x_0, 0)||h|, \end{aligned}$$

also wegen $Y(t; x_0, h) \rightarrow Y(t; x_0, 0)$ für $h \rightarrow 0$ die Differenzierbarkeit von $x(t; x_0)$ nach x_0 mit Ableitung $Y(t; x_0, 0)$. Diese Ableitung erfüllt aber wegen $A(t; x_0, 0) = \frac{\partial f}{\partial x}(t, x(t; x_0))$ die Variationsgleichung (1.25). \square

Insbesondere ist die Ableitung von $x(t; x_0)$ nach x_0 sogar stetig und auch die Zeitableitung x' nach dem Anfangswert differenzierbar, denn es gilt $(x')(t; x_0) = f(t, x(t; x_0))$ und die rechte Seite ist nach x_0 differenzierbar.

Den allgemeinen Fall kann man auf den in Satz 1.61 behandelten Fall der differenzierbaren Abhängigkeit vom Anfangswert bei einer autonomen Differentialgleichung zurückführen. Dazu reicht es zu beobachten, dass analog zu 1.6 das zeit- und parameterabhängige Anfangswertproblem $x' = f(t, x, \lambda_0)$, $x(t_0) = x_0$, mit $y(t-t_0) := (t, x(t), \lambda)^T$, $s := t-t_0$, zum autonomen Anfangswertproblem

$$\frac{dy}{ds} = \begin{pmatrix} t \\ x \\ \lambda \end{pmatrix}' = \begin{pmatrix} 1 \\ f(t, x, \lambda) \\ 0 \end{pmatrix} =: \tilde{f}(y), \quad y(0) = \begin{pmatrix} t_0 \\ x_0 \\ \lambda_0 \end{pmatrix}$$

äquivalent ist. Die Variationsgleichung (1.25) lautet hier wegen $t = s+t_0$, $\lambda = \lambda_0$ und $x(t; t_0, x_0, \lambda_0) = y_2(t-t_0; (t_0, x_0, \lambda_0))$ gerade

$$\frac{d}{dt} \begin{pmatrix} 1 & 0 & 0 \\ \frac{\partial x}{\partial t_0} + f & \frac{\partial x}{\partial x_0} & \frac{\partial x}{\partial \lambda_0} \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ \frac{\partial f}{\partial t} & \frac{\partial f}{\partial x} & \frac{\partial f}{\partial \lambda} \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ \frac{\partial x}{\partial t_0} + f & \frac{\partial x}{\partial x_0} & \frac{\partial x}{\partial \lambda_0} \\ 0 & 0 & 1 \end{pmatrix}.$$

Also ergibt sich aus Satz 1.61 das folgende Korollar.

Korollar 1.62. *Seien $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ und $\Lambda \subset \mathbb{R}^k$ offen, sei $(t_0, x_0) \in \Omega$ und $\lambda_0 \in \Lambda$, sei $f : \Omega \times \Lambda \rightarrow \mathbb{R}^n$ ein stetig differenzierbares zeit- und parameterabhängiges Vektorfeld, und existiere die Lösung $x(\cdot; t_0, x_0, \lambda_0)$ von $x' = f(t, x, \lambda_0)$ zum Anfangswert $x(t_0) = x_0$ auf dem Intervall $[a, b]$.*

Dann existieren für alle $t \in [a, b]$ die partiellen Ableitungen von $x(t; t_0, x_0, \lambda_0)$ nach t_0 , x_0 , λ_0 . Diese sind stetig und

- $\frac{\partial x}{\partial t_0}(\cdot; t_0, x_0, \lambda_0)$ löst das homogene lineare System $w' = \frac{\partial f}{\partial x}(t, x(t; t_0, x_0, \lambda_0), \lambda_0)w$ zum Anfangswert $w(t_0) = -f(t_0, x_0, \lambda_0)$
- $\frac{\partial x}{\partial x_0}(\cdot; t_0, x_0, \lambda_0)$ löst $Y' = \frac{\partial f}{\partial x}(t, x(t; t_0, x_0, \lambda_0), \lambda_0)Y$ zum Anfangswert $Y(t_0) = \text{Id}$
- $\frac{\partial x}{\partial \lambda_0}(\cdot; t_0, x_0, \lambda_0)$ löst $Z' = \frac{\partial f}{\partial x}(t, x(t; t_0, x_0, \lambda_0), \lambda_0)Z + \frac{\partial f}{\partial \lambda}(t, x(t; t_0, x_0, \lambda_0), \lambda_0)$ zum Anfangswert $Z(t_0) = 0$.

Man bemerke, dass wir die stetige Differenzierbarkeit von $x(t; t_0, x_0, \lambda_0)$ in t_0 , x_0 und λ_0 bisher nur für feste $t \in [a, b]$ formuliert haben. Tatsächlich ist aber sogar die Abbildung $(t_0, x_0, \lambda_0) \mapsto (t \mapsto x(t; t_0, x_0, \lambda_0))$ in den Banachraum $C([a, b], \mathbb{R}^n)$ stetig differenzierbar, denn die im Beweis von Satz 1.61 verwendete Konvergenz $Y(t; x_0, h) \rightarrow Y(t; x_0, 0)$ für $h \rightarrow 0$ gilt aufgrund unserer Erkenntnisse über stetige Abhängigkeit nicht nur punktweise für $t \in [a, b]$, sondern sogar gleichmäßig auf $[a, b]$.

Kapitel 2

Lineare Differentialgleichungen

Lineare Systeme von Differentialgleichungen und lineare Differentialgleichungen höherer Ordnung spielen nicht nur eine wichtige Rolle in vielen Anwendungen, sondern auch innerhalb der Mathematik, da sie bei der Linearisierung von nichtlinearen Differentialgleichungen entlang von Lösungen auftreten und qualitative Aussagen über deren Verhalten ermöglichen.

2.1 Lineare Systeme

In Verallgemeinerung von (1.3) hat eine (explizite) lineare n -dimensionale Differentialgleichung erster Ordnung die Form

$$x' = A(t)x + b(t) \quad (2.1)$$

mit einer zeitabhängigen Matrix $A : I \rightarrow \mathbb{R}^{n \times n}$ und einem zeitabhängigen Vektor $b : I \rightarrow \mathbb{R}^n$ auf einem Intervall $I \subset \mathbb{R}$. Man nennt (2.1) auch ein n -dimensionales lineares System von Differentialgleichungen erster Ordnung. Im Fall $b = 0$ nennt man (2.1) homogen, und ist A von t unabhängig, dann spricht man von einem linearen System mit konstanten Koeffizienten.

Zunächst wollen wir die globale Existenz und Eindeutigkeit einer Lösung des zu (2.1) gehörigen Anfangswertproblems zeigen.

Satz 2.1. *Sei $I \subset \mathbb{R}$ ein Intervall und seien $A : I \rightarrow \mathbb{R}^{n \times n}$ sowie $b : I \rightarrow \mathbb{R}^n$ stetig, dann gibt es zu jedem $(t_0, x_0) \in I \times \mathbb{R}^n$ eine eindeutige Lösung $x : I \rightarrow \mathbb{R}^n$ des Anfangswertproblems $x' = A(t)x + b(t)$, $x(t_0) = x_0$.*

Beweis: Die rechte Seite der Differentialgleichung (2.1) ist durch die affin-lineare Abbildung $f : I \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, $f(t, x) = A(t)x + b(t)$, gegeben. Diese genügt auf jedem kompakten Teilintervall $[a, b] \subset I$ mit dem Maximum $L := \max_{t \in [a, b]} |A(t)|$ (wobei $|A|$ die von der gewählten Norm $|\cdot|$ auf \mathbb{R}^n induzierte Matrixnorm (=Operatornorm) von A bezeichnet) der globalen Lipschitzbedingung

$$|f(t, x) - f(t, \tilde{x})| \leq L|x - \tilde{x}|.$$

Daher besitzt die Differentialgleichung nach 1.37 auf jedem kompakten Teilintervall $[a, b] \subset I$ und somit auch auf ganz I eine eindeutige Lösung x zum Anfangswert $x(t_0) = x_0$. \square

Bemerkung 2.2. *Ebenso existieren eindeutige globale schwache Lösungen $x \in W^{1,p}(I, \mathbb{R}^n)$ des Anfangswertproblems $x' = A(t)x + b(t)$, $x(t_0) = x_0$, wenn nur $A \in L^p(I, \mathbb{R}^{n \times n})$ und $b \in L^p(I, \mathbb{R}^{n \times n})$*

gilt. Dies folgt leicht aus unserer Diskussion schwacher Lösungen, denn $f(t, x) = A(t)x + b(t)$ ist in diesem Fall ein Carathéodory-Vektorfeld, das einer L^p -Wachstumsbedingung und der globalen L^p -Lipschitzbedingung

$$|f(t, x) - f(t, \tilde{x})| \leq L(t)|x - \tilde{x}|$$

mit $L(t) := |A(t)| \in L^p(I, \mathbb{R})$ genügt.

Im Gegensatz zu allgemeinen nichtlinearen Differentialgleichungssystemen besitzen lineare Systeme also unter sehr schwachen Voraussetzungen an die Koeffizienten eindeutige globale Lösungen zu Anfangswerten. Wir interessieren uns im Folgenden für die genaue Struktur der Lösungen von Systemen von linearen Differentialgleichungen. Dabei ist es wie bei der Diskussion von linearen Gleichungen in der linearen Algebra nützlich, zuerst den homogenen und erst danach den inhomogenen Fall zu diskutieren.

2.1.1 Homogene lineare Systeme

Satz 2.3. Sei $I \subset \mathbb{R}$ ein Intervall und $A : I \rightarrow \mathbb{R}^{n \times n}$ stetig. Dann bilden die Lösungen des homogenen linearen Systems $x' = Ax$ einen n -dimensionalen Vektorraum, und für Lösungen x_1, \dots, x_k von $x' = Ax$ sind die folgenden Aussagen äquivalent:

- (a) x_1, \dots, x_k sind linear unabhängige Kurven.
- (b) Es gibt ein $t_0 \in I$, so dass die Vektoren $x_1(t_0), \dots, x_k(t_0)$ linear unabhängig sind.
- (c) Für jedes $t \in I$ sind die Vektoren $x_1(t), \dots, x_k(t)$ linear unabhängig.

Beweis: Sind x und \tilde{x} Lösungen von $x' = A(t)x$ auf I , dann ist auch $\lambda x + \tilde{\lambda} \tilde{x}$ für beliebige $\lambda, \tilde{\lambda}$ wegen

$$\begin{aligned} (\lambda x + \tilde{\lambda} \tilde{x})' &= \lambda x' + \tilde{\lambda} \tilde{x}' = \\ \lambda A(t)x + \tilde{\lambda} A(t)\tilde{x} &= A(t) (\lambda x + \tilde{\lambda} \tilde{x}) \end{aligned}$$

eine Lösung, also bilden die Lösungen einen Untervektorraum von $C(I, \mathbb{R}^n)$.

Offensichtlich gilt desweiteren (iii) \Rightarrow (ii) \Rightarrow (i), so dass nur noch (i) \Rightarrow (iii) gezeigt werden muss. Seien dazu die Kurven x_1, \dots, x_k linear unabhängig. Angenommen, es gäbe ein $t \in I$, so dass die Vektoren $x_1(t), \dots, x_k(t)$ linear abhängig wären, dann gäbe es $(\lambda_1, \dots, \lambda_k) \neq 0$ mit $\lambda_1 x_1(t) + \dots + \lambda_k x_k(t) = 0$. Nun ist die Linearkombination $x := \lambda_1 x_1 + \dots + \lambda_k x_k$ aber auch eine Lösung und hat in t den Anfangswert 0. Aufgrund der Eindeutigkeit muss also $x = 0$ auf ganz I gelten (denn die konstante Kurve 0 ist eine Lösung zum Anfangswert $x(t) = 0$), dies würde aber $\lambda_1 x_1 + \dots + \lambda_k x_k = 0$ mit $(\lambda_1, \dots, \lambda_k) \neq 0$ bedeuten und somit der linearen Unabhängigkeit von x_1, \dots, x_k widersprechen.

Aus der Äquivalenz von (i),(ii) und (iii) folgt auch sofort, dass der Lösungsraum die Dimension n hat. Denn wegen (ii) \Rightarrow (i) folgt aus der linearen Unabhängigkeit der Anfangswerte zu einem Zeitpunkt $t_0 \in I$ auch schon die lineare Unabhängigkeit der Lösungen. Nach Satz 2.1 existieren aber eindeutige Lösungen x_i auf I zu den linear unabhängigen Anfangswerten $x(t_0) = e_i$, $i = 1, \dots, n$ (wobei e_i den i -ten Einheitsvektor bezeichnet), also hat der Lösungsraum mindestens die Dimension n . Andererseits müssten für $n + 1$ linear unabhängige Lösungen x_1, \dots, x_{n+1} und beliebiges $t \in I$ wegen (i) \Rightarrow (iii) auch schon die $(n + 1)$ Vektoren $x_1(t), \dots, x_{n+1}(t)$ im \mathbb{R}^n linear unabhängig sein, was unmöglich ist. \square

Definition 2.4. Eine Basis x_1, \dots, x_n des Lösungsraumes des n -dimensionalen homogenen linearen Systems $x' = Ax$ nennt man ein *Fundamentalsystem*.

Häufig fasst man solch eine Basis als Matrix-wertige Kurve $X(t) := (x_1(t) \ \dots \ x_n(t))$ auf, in deren Spalten gerade die einzelnen Lösungen x_1, \dots, x_n stehen, und nennt dann auch X ein Fundamentalsystem. Tatsächlich gilt die Matrix-Differentialgleichung $X' = A(t)X$ (mit dem Produkt von Matrizen auf der rechten Seite), und jede Lösung von $x' = Ax$ hat die Form $x(t) = X(t)c$ mit einem konstanten Vektor $c \in \mathbb{R}^n$. Ein ausgezeichnetes Fundamentalsystem erhält man durch Lösen der Matrix-Differentialgleichung $X' = A(t)X$ zum Anfangswert $X(t_0) = \text{Id}$. Mit diesem Fundamentalsystem ist nämlich $x(t) := X(t)x_0$ die Lösung von $x' = A(t)x$ zum Anfangswert $x(t_0) = x_0$.

Bevor wir uns inhomogenen linearen Systemen zuwenden, wollen wir noch ein Kriterium angeben, mit dem man die in Satz 2.3 formulierten äquivalenten Aussagen über die lineare Unabhängigkeit von Lösungen rechnerisch überprüfen kann. Dazu nennen wir eine Matrix X ein Lösungssystem, falls $X' = A(t)X$ gilt (verlangen aber im Gegensatz zu einem Fundamentalsystem nicht die Regularität der Matrizen $X(t)$).

Satz 2.5. Ist $X: I \rightarrow \mathbb{R}^{n \times n}$ ein Lösungssystem der homogenen linearen Differentialgleichung $x' = A(t)x$, dann genügt die Wronski-Determinante genannte Funktion $y := \det(X)$ der homogenen linearen Differentialgleichung $y' = \text{Spur}(A(t))y$, wobei $\text{Spur}(A(t))$ die Spur der Matrix $A(t)$ bezeichnet.

Beweis: Sei $X(t) = (x_1(t) \ \dots \ x_n(t))$, dann gilt

$$(\det(X))' = \sum_{i=1}^n \det(x_1, \dots, x_{i-1}, x_i', x_{i+1}, \dots, x_n)$$

aufgrund der Linearität der Determinante in den Spalten. Ist nun \tilde{X} das ausgezeichnete Fundamentalsystem mit $\tilde{X}(t_0) = \text{Id}$, dann gilt $\tilde{x}_i(t_0) = e_i$ (mit dem i -ten Einheitsvektor e_i) und daher wegen $\tilde{x}_i'(t_0) = A(t_0)\tilde{x}_i(t_0) = A(t_0)e_i$ auch

$$(\det(\tilde{X}))'(t_0) = \sum_{i=1}^n \det(e_1, \dots, e_{i-1}, A(t_0)e_i, e_{i+1}, \dots, e_n) = \sum_{i=1}^n a_{ii}(t_0) = \text{Spur}(A(t_0)).$$

Jedes allgemeine Lösungssystem X erfüllt aber $X(t) = \tilde{X}(t)X(t_0)$ mit dem ausgezeichneten Fundamentalsystem \tilde{X} zum Anfangswert $\tilde{X}(t_0) = \text{Id}$. Also gilt aufgrund der Multiplikativität der Determinanten $\det(X(t)) = \det(\tilde{X}(t))\det(X(t_0))$ und insbesondere

$$y'(t) = (\det(X(t)))' = (\det(\tilde{X}(t)))' \det(X(t_0))$$

was für $t = t_0$ gerade $y'(t_0) = \text{Spur}(A(t_0))y(t_0)$ liefert. Nun war aber t_0 beliebig, so dass man $y'(t) = \text{Spur}(A(t))y(t)$ für alle $t \in I$ erhält. \square

Da die Lösung der Differentialgleichung $y' = \text{Spur}(A(t))y$ zum Anfangswert $y(t_0) = y_0$ durch $y(t) = \exp\left(\int_{t_0}^t \text{Spur}(A(s)) ds\right) y_0$ gegeben ist, ergibt sich als Konsequenz aus diesem Satz, dass für ein Lösungssystem die Wronski-Determinante entweder für alle Zeiten verschwindet (dies ist bei $y_0 = 0$ der Fall) oder für alle Zeiten ungleich Null ist (falls $y_0 \neq 0$ gilt). Man bemerke, dass diese Aussage gleichbedeutend mit der Äquivalenz von (b) und (c) in Satz 2.3 ist.

2.1.2 Inhomogene lineare Systeme

Satz 2.6. Sei $I \subset \mathbb{R}$ ein Intervall und seien $A: I \rightarrow \mathbb{R}^{n \times n}$ sowie $b: I \rightarrow \mathbb{R}^n$ stetig. Ist x_p eine Lösung des inhomogenen linearen Systems (2.1), dann hat jede weitere Lösung x von (2.1) die Form $x = x_p + x_h$ mit einer Lösung x_h des homogenen linearen Systems $x' = A(t)x$.

Beweis: Sind x_p und x Lösungen von $x' = A(t)x + b(t)$, dann ist $x - x_p$ wegen

$$(x - x_p)' = (A(t)x + b(t)) - (A(t)x_p + b(t)) = A(t)(x - x_p)$$

eine Lösung x_h des homogenen linearen Systems $x' = A(t)x$ und es gilt $x = x_p + x_h$. \square

Abstrakt besagt der vorige Satz, dass die Lösungen des inhomogenen linearen Systems (2.1) einen affinen Unterraum von $C(I, \mathbb{R}^n)$ bilden, dessen linearer Anteil der Lösungsraum des zugehörigen homogenen linearen Systems ist. Neben der allgemeinen Lösung des homogenen linearen Systems muss man also nur eine einzige partikuläre Lösung des inhomogenen linearen Systems kennen, dann hat man schon eine vollständige Kenntnis von der allgemeinen Lösung des inhomogenen linearen Systems. Wie im eindimensionalen Fall kann man eine partikuläre Lösung durch Variation der Konstanten gewinnen.

Satz 2.7. Sei $I \subset \mathbb{R}$ ein Intervall und seien $A: I \rightarrow \mathbb{R}^{n \times n}$ sowie $b: I \rightarrow \mathbb{R}^n$ stetig. Ist X ein Fundamentalsystem des homogenen linearen Systems $x' = A(t)x$, dann erhält man eine Lösung x_p des inhomogenen linearen Systems $x' = A(t)x + b(t)$ durch den Ansatz $x_p(t) = X(t)c(t)$, aus dem sich bis auf einen konstanten Vektor $c(t) = \int_{t_0}^t X(s)^{-1}b(s) ds$ ergibt.

Beweis: Setzt man den Ansatz $x_p(t) = X(t)c(t)$ in $x' = A(t)x + b(t)$ ein, so ergibt sich $X'c + Xc' = A(t)Xc + b(t)$ und daher wegen $X' = A(t)X$ sofort $c'(t) = X(t)^{-1}b(t)$, d.h. bis auf einen konstanten Vektor $c(t) = \int_{t_0}^t X(s)^{-1}b(s) ds$. \square

Zur Anwendung des Satzes sei allerdings angemerkt, dass es im Gegensatz zum Eindimensionalen oft nicht ganz leicht ist, das Inverse des Fundamentalsystems $X(t)$ bzw. eine Lösung c der Gleichung $X(t)c'(t) = b(t)$ konkret auszurechnen.

2.2 Lineare Differentialgleichungen höherer Ordnung

In Verallgemeinerung von (1.3) hat eine (explizite) lineare (eindimensionale) Differentialgleichung n -ter Ordnung die Form

$$x^{(n)} + a_{n-1}(t)x^{(n-1)} + \dots + a_1(t)x' + a_0(t)x = b(t) \quad (2.2)$$

mit zeitabhängigen Koeffizienten $a_0, a_1, \dots, a_{n-1} : I \rightarrow \mathbb{R}$ und einer Inhomogenität $b : I \rightarrow \mathbb{R}$ auf einem Intervall $I \subset \mathbb{R}$. Nach Satz 1.4 ist die lineare Differentialgleichung n -ter Ordnung (2.2) über $y(t) = (x(t), x'(t), \dots, x^{(n-1)}(t))$ zum linearen n -dimensionalen System

$$y' = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -a_0(t) & -a_1(t) & -a_2(t) & \dots & -a_{n-1}(t) \end{pmatrix} y + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ b(t) \end{pmatrix} \quad (2.3)$$

äquivalent. Insbesondere kann man die Sätze 2.1, 2.3 und 2.6 auf lineare Differentialgleichungen n -ter Ordnung übertragen und erhält das folgende Resultat.

Satz 2.8. Sei $I \subset \mathbb{R}$ ein Intervall und seien $a_0, a_1, \dots, a_n: I \rightarrow \mathbb{R}$ sowie $b: I \rightarrow \mathbb{R}$ stetig. Dann existiert zu jedem $(t_0, x_0, \dots, x_{n-1}) \in I \times \mathbb{R}^n$ genau eine Lösung $x: I \rightarrow \mathbb{R}$ von (2.2) zu den Anfangswerten $x(t_0) = x_0, x'(t_0) = x_1, \dots, x^{(n-1)}(t_0) = x_{n-1}$.

Die Lösungen der homogenen Gleichung $x^{(n)} + a_{n-1}(t)x^{(n-1)} + \dots + a_1(t)x' + a_0(t)x = 0$ bilden einen n -dimensionalen Vektorraum, und kennt man eine Lösung x_p der inhomogenen Gleichung (2.2), dann hat jede weitere Lösung x der inhomogenen Gleichung (2.2) die Form $x = x_p + x_h$ mit einer Lösung x_h der homogenen Gleichung.

Wieder bezeichnet man eine Basis des Lösungsraumes der homogenen Gleichung als Fundamentalsystem. Auch die Bestimmung einer partikulären Lösung x_p der inhomogenen Gleichung (2.2) mittels Variation der Konstanten überträgt sich von linearen Systemen auf lineare Differentialgleichungen höherer Ordnung. Die Berechnung wird dabei sogar einfacher, denn für n linear unabhängige Lösungen x_1, \dots, x_n der homogenen linearen Differentialgleichung n -ter Ordnung

$$x^{(n)} + a_{n-1}(t)x^{(n-1)} + \dots + a_1(t)x' + a_0(t)x = 0$$

ist das Fundamentalsystem des zugehörigen Systems (2.3) durch

$$Y = \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ x'_1 & x'_2 & \dots & x'_n \\ \vdots & \vdots & \vdots & \vdots \\ x_1^{(n-1)} & x_2^{(n-1)} & \dots & x_n^{(n-1)} \end{pmatrix}$$

gegeben, und die Inhomogenität des Systems hat nur in der letzten Zeile einen nicht verschwindenden Eintrag. Während also eine partikuläre Lösung des Systems allgemein durch $y_p(t) = Y(t) \int_{t_0}^t Y(s)^{-1} (0, \dots, 0, b(s))^T ds$ gegeben ist, hat hier die Lösung z von $Y(s)z = (0, \dots, 0, b(s))^T$ nach der Cramerschen Regel die Form $z_i = (-1)^{n+i} \frac{W_i(s)}{W(s)} b(s)$ mit den Wronski-Determinanten $W(s) := \det(Y(s))$ und $W_i(s) := \det(Y_i(s))$, wobei Y_i die $((n-1) \times (n-1))$ -Matrix bezeichnet, bei der man die i -te Spalte und die letzte Zeile gestrichen hat ($W_i(s)$ ist also die Wronski-Determinante zu $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$ bis zur $(n-2)$ -ten Ableitung). Da uns außerdem nur die erste Zeile interessiert, ist eine partikuläre Lösung x_p der linearen Differentialgleichung n -ter Ordnung gegeben durch

$$x_p(t) = \sum_{i=1}^n (-1)^{n+i} x_i(t) \int_{t_0}^t \frac{W_i(s)}{W(s)} b(s) ds. \quad (2.4)$$

2.3 Die Reduktionsmethode von d'Alembert

Kennt man eine nichttriviale Lösung eines homogenen linearen Systems der Dimension n bzw. einer homogenen linearen Differentialgleichung n -ter Ordnung, so kann man mittels eines von d'Alembert entdeckten Verfahrens die Dimension bzw. die Ordnung um 1 reduzieren und hat nur noch ein homogenes lineares System der Dimension $n-1$ bzw. eine homogene lineare Differentialgleichung $(n-1)$ -ter Ordnung zu lösen. Diese Reduktionsmethode ist insbesondere im Fall $n=2$ nützlich, denn eine eindimensionale lineare Differentialgleichungen erster Ordnung (1.3) besitzt die explizite Lösung (1.4), also braucht man im Fall $n=2$ nur eine Lösung zu erraten und erhält daraus schon die allgemeine Lösung.

2.3.1 Reduktion der Dimension

Ist $\tilde{x} : I \rightarrow \mathbb{R}^n$ eine nichttriviale Lösung des n -dimensionalen homogenen linearen Systems $x' = A(t)x$ und gilt o.B.d.A. $\tilde{x}_1 \neq 0$ auf dem Intervall I (ansonsten verkleinere man I oder benenne Komponenten um), dann erhält man durch den Ansatz

$$x(t) := \lambda(t)\tilde{x}(t) + \begin{pmatrix} 0 \\ y_1(t) \\ \vdots \\ y_{n-1}(t) \end{pmatrix} \quad (2.5)$$

mit einer skalaren Funktion $\lambda : I \rightarrow \mathbb{R}$ ein $(n-1)$ -dimensionales homogenes lineares System für $y : I \rightarrow \mathbb{R}^{n-1}$. In der Tat, die durch den Ansatz definierte Kurve $x : I \rightarrow \mathbb{R}^n$ ist genau dann eine Lösung von $x' = Ax$, wenn

$$\lambda' \tilde{x} + \lambda \tilde{x}' + \begin{pmatrix} 0 \\ y_1' \\ \vdots \\ y_{n-1}' \end{pmatrix} = \lambda A \tilde{x} + A \begin{pmatrix} 0 \\ y_1 \\ \vdots \\ y_{n-1} \end{pmatrix}$$

gilt. Wegen $\tilde{x}' = A\tilde{x}$ ist dieses System zu den Gleichungen

$$\begin{aligned} \lambda' \tilde{x}_1 &= \sum_{j=2}^n a_{1j} y_{j-1} \quad \text{und} \\ \lambda' \tilde{x}_{i+1} + y_i' &= \sum_{j=2}^n a_{(i+1)j} y_{j-1} \quad (i = 1, \dots, n-1) \end{aligned}$$

äquivalent. Aus der ersten Gleichung erhält man wegen $\tilde{x}_1 \neq 0$ die Gleichung

$$\lambda' = \sum_{j=2}^n \frac{1}{\tilde{x}_1} a_{1j} y_{j-1} \quad (2.6)$$

und setzt man diese in die verbleibenden Gleichungen ein, so folgt

$$y_i' = \sum_{j=2}^n \left(a_{(i+1)j} - \frac{\tilde{x}_{i+1}}{\tilde{x}_1} a_{1j} \right) y_{j-1}.$$

Dies ist das gewünschte $(n-1)$ -dimensionale homogene lineare System für y .

Hat man nun ein Fundamentalsystem y_1, \dots, y_{n-1} dieses $(n-1)$ -dimensionalen Systems gewonnen (z.B. durch Raten einer weiteren Lösung und wiederholte Anwendung der Reduktionsmethode, oder da das System zufällig konstante Koeffizienten hatte), dann erhält man aus jeder dieser Lösungen y_i zunächst durch Integration von (2.6) eine skalare Funktion λ_i und danach aus (2.5) eine Lösung x_i des ursprünglichen n -dimensionalen Systems. Zu zeigen bleibt noch, dass diese $(n-1)$ Lösungen x_1, \dots, x_{n-1} zusammen mit $x_n := \tilde{x}$ ein Fundamentalsystem bilden, also linear unabhängig sind. Dazu sei $x_i = \lambda_i \tilde{x} + \begin{pmatrix} 0 \\ y_i \end{pmatrix}$ und $\mu_1 x_1 + \dots + \mu_n x_n = 0$. Teilt man die erste Zeile dieser Gleichung durch $\tilde{x}_1 \neq 0$, so erhält man

$$\mu_1 \lambda_1 + \dots + \mu_{n-1} \lambda_{n-1} + \mu_n = 0.$$

Multipliziert man diese Gleichung mit \tilde{x} und zieht sie von $\mu_1 x_1 + \dots + \mu_n x_n = 0$ ab, dann erhält man $\mu_1 y_1 + \dots + \mu_{n-1} y_{n-1} = 0$. Aufgrund der linearen Unabhängigkeit von y_1, \dots, y_{n-1} folgt daraus $\mu_1 = \dots = \mu_{n-1} = 0$, und damit dann schließlich auch $\mu_n = 0$, was die lineare Unabhängigkeit von x_1, \dots, x_n beweist.

2.3.2 Reduktion der Ordnung

Ist \tilde{x} eine nichttriviale Lösung der homogenen linearen Differentialgleichung n -ter Ordnung

$$Lx := x^{(n)} + a_{n-1}(t)x^{(n-1)} + \dots + a_1(t)x' + a_0(t)x = 0$$

mit $\tilde{x} \neq 0$ auf I und würde man das zuvor geschilderte Reduktionsverfahren auf das zugehörige System (2.3) und dessen Lösung $(\tilde{x}, \tilde{x}', \dots, \tilde{x}^{(n-1)})$ anwenden, so würde sich als reduziertes $(n-1)$ -dimensionales System im Allgemeinen nicht wieder ein homogenes lineares System ergeben, das zu einer homogenen linearen Differentialgleichung der Ordnung $(n-1)$ gehört. Daher ist es nützlich, den Ansatz zu $x(t) = y(t)\tilde{x}(t)$ abzuändern und y so zu bestimmen, dass x ebenfalls $Lx = 0$ erfüllt. Die allgemeine Produktregel ergibt mit $a_n := 1$

$$Lx := \sum_{i=0}^n a_i(t) \sum_{j=0}^i \binom{i}{j} y^{(j)} \tilde{x}^{(i-j)} = \sum_{j=0}^n \left(\sum_{i=j}^n \binom{i}{j} a_i(t) \tilde{x}^{(i-j)} \right) y^{(j)}$$

Der Koeffizient vor y (d.h. bei $j = 0$) ist gerade $\sum_{i=0}^n a_i(t) \tilde{x}^{(i)} = L\tilde{x} = 0$, also gilt $Lx = 0$ genau dann, wenn y' die homogene lineare Differentialgleichung

$$\sum_{j=0}^{n-1} \left(\sum_{i=j+1}^n \binom{i}{j+1} a_i(t) \tilde{x}^{(i-j-1)} \right) (y')^{(j)} = 0$$

der Ordnung $(n-1)$ löst. Hat man nun $(n-1)$ linear unabhängige Lösungen y'_1, \dots, y'_{n-1} dieser Gleichung bestimmt und sind y_1, \dots, y_{n-1} zugehörige Stammfunktionen, so ist $y_1 \tilde{x}, \dots, y_{n-1} \tilde{x}, \tilde{x}$ tatsächlich ein Fundamentalsystem der ursprünglichen homogenen linearen Differentialgleichung n -ter Ordnung, denn aus

$$\mu_1 y_1 \tilde{x} + \dots + \mu_{n-1} y_{n-1} \tilde{x} + \mu_n \tilde{x} = 0$$

folgt nach Division durch $\tilde{x} \neq 0$ und anschließende Differentiation

$$\mu_1 y'_1 + \dots + \mu_{n-1} y'_{n-1} = 0.$$

Aufgrund der linearen Unabhängigkeit von y'_1, \dots, y'_{n-1} ergibt sich daraus $\mu_1 = \dots = \mu_{n-1} = 0$ und damit auch $\mu_n = 0$, was die lineare Unabhängigkeit von $y_1 \tilde{x}, \dots, y_{n-1} \tilde{x}, \tilde{x}$ beweist.

2.4 Lineare Differentialgleichungen mit konstanten Koeffizienten

In diesem Abschnitt wollen wir lineare Differentialgleichungen mit konstanten Koeffizienten explizit durch Angabe eines Fundamentalsystems lösen. Für Systeme kann man ein Fundamentalsystem mittels Transformation auf Jordansche Normalform bestimmen, für Gleichungen höherer Ordnung muss man nur die Nullstellen des charakteristischen Polynoms berechnen.

2.4.1 Lineare Systeme mit konstanten Koeffizienten

Ein n -dimensionales lineares System von Differentialgleichungen mit konstanten Koeffizienten hat die Form

$$x' = Ax + b(t) \quad (2.7)$$

mit einer konstanten Matrix $A \in \mathbb{R}^{n \times n}$. Zunächst interessiert uns wieder der homogene Fall $b = 0$, der später insbesondere beim Studium des Verhaltens eines nichtlinearen Systems nahe einer Ruhelage eine große Rolle spielen wird.

Das ausgezeichnete Fundamentalsystem X von $x' = Ax$ mit $X(0) = \text{Id}$ kann man mit Hilfe der Exponentialabbildung auf Matrizen sofort hinschreiben, es ist nämlich durch $X(t) = \exp(tA)$ gegeben. Dabei ist die Exponentialabbildung auf Matrizen ganz analog zum Eindimensionalen durch die Reihe

$$\exp(A) := \sum_{k=0}^{\infty} \frac{A^k}{k!}$$

definiert, die einen unendlichen Konvergenzradius besitzt und deshalb auch für jede Matrix $A \in \mathbb{R}^{n \times n}$ konvergiert. Wie in der Analysis I kann man desweiteren für $A, B \in \mathbb{R}^{n \times n}$ mit $AB = BA$ die Funktionalgleichung $\exp(A + B) = \exp(A)\exp(B)$ mittels des Cauchy-Produktes von absolut konvergenten Reihen unter Ausnutzung der Kommutativität $AB = BA$ beweisen, siehe [Königsberger, 1.6].

Satz 2.9. *Ist $A \in \mathbb{R}^{n \times n}$, so ist die eindeutige Lösung von $x' = Ax$ zum Anfangswert $x(0) = x_0$ durch $x(t) = \exp(tA)x_0$ gegeben.*

Beweis: Die Kurve $X: t \mapsto \exp(tA)$ in $\mathbb{R}^{n \times n}$ erfüllt die Gleichung $\exp((s+t)A) = \exp(sA)\exp(tA)$ sowie $\exp(0A) = \text{Id}$. Da für $|h| \leq 1$ die Abschätzung

$$|\exp(hA) - \text{Id} - hA| \leq |h|^2 \sum_{k=2}^{\infty} \frac{|A|^k}{k!}$$

gilt, folgt desweiteren für die Kurve $x(t) := \exp(tA)x_0$ nicht nur $x(0) = x_0$, sondern auch

$$\begin{aligned} x'(t) &= \lim_{h \rightarrow 0} \frac{x(t+h) - x(t)}{h} = \lim_{h \rightarrow 0} \frac{\exp((t+h)A)x_0 - \exp(tA)x_0}{h} = \\ &= \exp(tA) \left(\lim_{h \rightarrow 0} \frac{\exp(hA) - \text{Id}}{h} x_0 \right) = \exp(tA)Ax_0 = A \exp(tA)x_0 = Ax(t) \quad . \end{aligned}$$

Dies zeigt, dass $x(t) = \exp(tA)x_0$ wirklich die Differentialgleichung $x' = Ax$ zum Anfangswert $x(0) = x_0$ löst. \square

Die allgemeine Formel $x(t) = \exp(tA)x_0$ aus Satz 2.9 für die Lösung von $x' = Ax$, $x(0) = x_0$, ist noch nicht vollkommen befriedigend, denn bisher wissen wir nicht, wie man den Grenzwert der Reihe $\exp(tA)$ für eine konkret gegebene Matrix A explizit ausrechnen kann. Offensichtlich kann man aber alles, was wir in diesem Abschnitt bisher für reelle Matrizen gemacht haben, auch für komplexe Matrizen machen. Desweiteren gilt $\exp(S^{-1}AS) = S^{-1}\exp(A)S$ für jede reguläre Matrix $S \in \mathbb{C}^{n \times n}$, denn

$$\exp(S^{-1}AS) = \sum_{k=0}^{\infty} \frac{(S^{-1}AS)^k}{k!} = S^{-1} \left(\sum_{k=0}^{\infty} \frac{A^k}{k!} \right) S = S^{-1} \exp(A)S.$$

Daher bietet es sich zur Berechnung von $\exp(tA)$ an, A durch Transformation mit S auf Jordansche Normalform zu bringen.

Im einfachsten Fall ist A diagonalisierbar, so dass man A in eine Diagonalmatrix $B = S^{-1}AS = \text{diag}(\lambda_1, \dots, \lambda_n)$ transformieren kann, wobei die (i.a. komplexen) Zahlen λ_j die Eigenwerte von A sind und S als Spalten linear unabhängige Eigenvektoren v_j zu den Eigenwerten λ_j enthält. In diesem Fall ergibt sich somit $\exp(tA) = S \text{diag}(e^{t\lambda_1}, \dots, e^{t\lambda_n})S^{-1}$. Setzt man $c := S^{-1}x_0$, so haben die Lösungen von $x' = Ax$ also die Form

$$x(t) = S \text{diag}(\exp(t\lambda_1), \dots, \exp(t\lambda_n))c = \sum_{j=1}^n c_j e^{t\lambda_j} v_j.$$

Im komplizierteren Fall ist A nicht diagonalisierbar, dann kann man A aber immerhin noch auf Jordansche Normalform $B = S^{-1}AS$ transformieren, wobei

$$B = \text{diag}(\lambda_1 \text{Id} + N_1, \dots, \lambda_r \text{Id} + N_r) \quad \text{mit} \quad N_j = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

gilt und S als Spalten linear unabhängige Eigen- und Hauptvektoren zu den Eigenwerten λ_j enthält. Dabei sind Hauptvektoren 0-ter Stufe zum Eigenwert λ_j nichts anderes als Eigenvektoren, und w_i heißt Hauptvektor i -ter Stufe zum Eigenwert λ_j , wenn $\lambda_j w_i + w_{i-1} = Aw_i$ mit einem Hauptvektor w_{i-1} der Stufe $(i-1)$ zu λ_j gilt. Man bemerke, dass sich die die Hauptvektoren zu λ_j definierende Gleichung auch als

$$(w_{k_j}, \dots, w_0)(\lambda_j \text{Id} + N_j) = A(w_{k_j}, \dots, w_0)$$

mit der maximalen Stufe k_j eines Hauptvektors (die gerade der Größe des Jordan-Blocks entspricht) schreiben läßt und somit direkt aus einem Block der Gleichung $B = S^{-1}AS$ entspringt. Aus der Blockgestalt von B ergibt sich

$$\exp(tB) = \text{diag}(\exp(t(\lambda_1 \text{Id} + N_1)), \dots, \exp(t(\lambda_r \text{Id} + N_r))),$$

und man kann leicht ausrechnen, wie die Exponentialabbildung auf einen einzelnen Jordan-Block wirkt.

Lemma 2.10. *Das Fundamentalsystem X des m -dimensionalen linearen Systems $x' = (\lambda \text{Id} + N)x$ mit $X(0) = \text{Id}$ ist*

$$X(t) = e^{\lambda t} \begin{pmatrix} 1 & t & \frac{t^2}{2} & \dots & \frac{t^{m-1}}{(m-1)!} \\ 0 & 1 & t & \dots & \frac{t^{m-2}}{(m-2)!} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & t \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

Beweis: Mit dem Ansatz $X(t) := e^{\lambda t} Y(t)$ hat man wegen $X' = \lambda X(t) + e^{\lambda t} Y'(t)$ nur noch $Y' = NY$, $Y(0) = \text{Id}$, zu lösen. Die Matrix N^k hat aber nur noch auf der k -ten rechten oberen

Nebendiagonalen Einsen und ist für $k \geq m$ die Nullmatrix, also ist

$$Y(t) = \exp(tN) = \begin{pmatrix} 1 & t & \frac{t^2}{2} & \cdots & \frac{t^{m-1}}{(m-1)!} \\ 0 & 1 & t & \cdots & \frac{t^{m-2}}{(m-2)!} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & t \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

□

Somit ist $\exp(tA) = S \operatorname{diag}(\exp(t(\lambda_1 \operatorname{Id} + N_1)), \dots, \exp(t(\lambda_r \operatorname{Id} + N_r)))S^{-1}$, und setzt man $c := S^{-1}x_0$, so sind die Lösungen von $x' = Ax$ Linearkombinationen von

$$e^{t\lambda_j} w_{0,j}, \dots, e^{t\lambda_j} \left(w_{k_j,j} + t w_{k_j-1,j} + \cdots + \frac{t^{k_j-1}}{(k_j-1)!} w_{0,j} \right) \quad (j = 1, \dots, r),$$

mit den Koeffizienten aus c , wobei $w_{i,j}$ einen Hauptvektor der Stufe i zu λ_j bezeichnet. Die Bestimmung des Fundamentalsystems $\exp(tA)$ eines linearen Systems mit konstanter Koeffizientenmatrix $A \in \mathbb{R}^{n \times n}$ ist also nicht viel schwieriger als die Berechnung der Jordanschen Normalform von A .

Man beachte jedoch, dass die Eigenwerte und Eigenvektoren auch für eine reelle Matrix A komplex sein können. Allerdings treten die Eigenwerte und Eigenvektoren für eine reelle Matrix A immer in konjugiert komplexen Paaren auf. Aus einem komplexen Fundamentalsystem kann man daher ein reelles Fundamentalsystem gewinnen, indem man jedes konjugiert komplexes Paar durch seinen Real- bzw. Imaginärteil ersetzt.

Kurz erwähnt werden soll abschließend noch der inhomogene Fall, bei dem $b \neq 0$ in (2.7) gilt. Natürlich kann man diesen Fall ganz allgemein mittels Variation der Konstanten handhaben, rechenstechnisch ist es für spezielle rechte Seiten b wie z.B. polynomiale, exponentielle oder trigonometrische vektorwertige Funktionen aber häufig besser, mittels eines geschickten Ansatzes eine partikuläre Lösung zu suchen. Solche Ansätze haben dabei meist eine ähnliche Form wie die Inhomogenität b . Genauer werden wir die Wahl eines speziellen Ansatzes für lineare Differentialgleichungen höherer Ordnung mit konstanten Koeffizienten diskutieren.

2.4.2 Lineare Differentialgleichungen höherer Ordnung mit konstanten Koeffizienten

Eine lineare Differentialgleichung n -ter Ordnung mit konstanten Koeffizienten hat die Form

$$x^{(n)} + a_{n-1}x^{(n-1)} + \cdots + a_1x' + a_0x = b(t) \quad (2.8)$$

mit Konstanten a_0, \dots, a_{n-1} . Natürlich könnte man ein explizites Fundamentalsystem von (2.8) dadurch bestimmen, dass man zum zugehörigen System übergeht. Die zugehörige Matrix hat das charakteristische Polynom $(-1)^n(\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0)$. Die Nullstellen dieses Polynoms sind gerade die Eigenwerte λ_j der zugehörigen Matrix, und aus der ersten Zeile von Lemma 2.10 sieht man, dass jede Lösung eine Linearkombination der Funktionen $t^{q-1}e^{\lambda_j t}$ ist, wobei q von 1 bis zur algebraischen Vielfachheit von λ_j läuft.

Wir wollen aber auch noch einen direkten elementaren Beweis dieser expliziten Form des Fundamentalsystems von (2.8) geben. Dazu bemerke man, dass man in jedes abstrakte Polynom $P \in \mathbb{C}[X]$ mit möglicherweise komplexen Koeffizienten, wie z.B. das Polynom

$$P(X) := X^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0$$

mit Höchstkoeffizient $a_n := 1$, für die Variable X nicht nur reelle oder komplexe Zahlen einsetzen kann, sondern auch Matrizen und sogar Differentialoperatoren wie z.B. den Differentialoperator erster Ordnung $D := (\cdot)' = \frac{d}{dt}$. Potenzen D^k interpretiert man dabei als k -malige Anwendung von D . Dadurch kann man die lineare Differentialgleichung n -ter Ordnung (2.8) einfach als $P(D)x = b$ mit dem Differentialoperator n -ter Ordnung $P(D)$ schreiben. Um nun ein Fundamentalsystem von (2.8) zu erhalten, berechnen wir, wie der Differentialoperator $P(D)$ auf den Ansatz $e^{\lambda t}$ wirkt.

Lemma 2.11. *Für jedes $P \in \mathbb{C}[X]$ und jedes $\lambda \in \mathbb{C}$ gilt $P(D)e^{\lambda t} = P(\lambda)e^{\lambda t}$.*

Beweis: Aus $De^{\lambda t} = \lambda e^{\lambda t}$ folgt $D^k e^{\lambda t} = \lambda^k e^{\lambda t}$ und damit

$$P(D)e^{\lambda t} = \sum_{k=0}^n D^k e^{\lambda t} = \sum_{k=0}^n a_k \lambda^k e^{\lambda t} = P(\lambda)e^{\lambda t}.$$

□

Insbesondere ist also für jede Nullstelle $\lambda \in \mathbb{C}$ von $P(\lambda)$ die Funktion $x(t) := e^{\lambda t}$ eine Lösung der homogenen Differentialgleichung n -ter Ordnung $P(D)x = 0$.

Satz 2.12. *Sei $P \in \mathbb{C}[X]$ ein Polynom vom Grad n . Hat P genau n paarweise verschiedene Nullstellen $\lambda_j \in \mathbb{C}$, $j = 1, \dots, n$, so bilden die Funktionen $e^{\lambda_j t}$ ein Fundamentalsystem von $P(D)x = 0$.*

Beweis: Wir müssen nur noch die lineare Unabhängigkeit der Funktionen $e^{\lambda_j t}$, $j = 1, \dots, n$, zeigen. Die zugehörige Wronski-Determinante lautet an der Stelle $t = 0$

$$\det \left(\begin{pmatrix} 1 & 1 & \dots & 1 \\ \lambda_1 & \lambda_2 & \dots & \lambda_n \\ \vdots & \vdots & & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \dots & \lambda_n^{n-1} \end{pmatrix} \right) = \prod_{i>j} (\lambda_i - \lambda_j),$$

da die Matrix gerade die Vandermondsche Matrix ist, deren Determinante man schon aus der linearen Algebra kennt. Diese Determinante ist aber ungleich Null, da $\lambda_i \neq \lambda_j$ für $i \neq j$ vorausgesetzt wurde. □

Ist $\lambda \in \mathbb{C}$ eine mehrfache Nullstelle der Vielfachheit k von P , dann ist $t^{q-1}e^{\lambda t}$ für jedes $q = 1, \dots, k$ eine Lösung der homogenen Differentialgleichung n -ter Ordnung $P(D)x = 0$. Tatsächlich, wegen $\frac{d^{q-1}}{d\lambda^{q-1}}e^{\lambda t} = t^{q-1}e^{\lambda t}$ gilt aufgrund der Vertauschbarkeit der Ableitungen nach t und λ

$$P(D)(t^{q-1}e^{\lambda t}) = P\left(\frac{d}{dt}\right)\left(\frac{d^{q-1}}{d\lambda^{q-1}}e^{\lambda t}\right) = \frac{d^{q-1}}{d\lambda^{q-1}}(P(D)e^{\lambda t}) = \frac{d^{q-1}}{d\lambda^{q-1}}(P(\lambda)e^{\lambda t}) = 0,$$

wobei im letzten Schritt benutzt wurde, dass λ (für festes t) auch eine k -fache Nullstelle von $P(\lambda)e^{\lambda t}$ ist wegen $\frac{d^{m-1}P}{d\lambda^{m-1}}(\lambda) = 0$ für alle $1 \leq m \leq k$ und der verallgemeinerten Produktregel

$$\frac{d^{q-1}}{d\lambda^{q-1}}(P(\lambda)e^{\lambda t}) = \sum_{m=1}^q \binom{q-1}{m-1} \frac{d^{m-1}P}{d\lambda^{m-1}}(\lambda) t^{q-m} e^{\lambda t}.$$

Satz 2.13. *Sei $P \in \mathbb{C}[X]$ ein Polynom vom Grad n . Hat P die paarweise verschiedenen Nullstellen $\lambda_j \in \mathbb{C}$, $j = 1, \dots, r$, der Vielfachheit k_j , so bilden die Funktionen $t^{q-1}e^{\lambda_j t}$, $q = 1, \dots, k_j$, $j = 1, \dots, r$, ein Fundamentalsystem von $P(D)x = 0$.*

Beweis: Wir müssen nur noch die lineare Unabhängigkeit der Funktionen $t^{q-1}e^{\lambda_j t}$, $q = 1, \dots, k_j$, $j = 1, \dots, r$, zeigen. Eine beliebige Linearkombination dieser Funktionen hat offenbar die Form $\sum_{j=1}^r p_j(t)e^{\lambda_j t}$, wobei p_j ein komplexes Polynom ist und λ_j paarweise verschiedene komplexe Zahlen sind. Wir müssen daher zeigen, dass $\sum_{j=1}^r p_j(t)e^{\lambda_j t} = 0$ nur dann auf einem Intervall gelten kann, wenn alle Polynome p_j gleich Null sind. Für $r = 1$ ist dies klar (Induktionsanfang). Angenommen, diese Aussage sei schon für r Summanden bewiesen und es gelte

$$\sum_{j=1}^{r+1} p_j(t)e^{\lambda_j t} = 0$$

auf einem Intervall I . Multiplikation dieser Gleichung mit $e^{-\lambda_{r+1}t}$ liefert

$$\sum_{j=1}^r p_j(t)e^{(\lambda_j - \lambda_{r+1})t} + p_{r+1}(t) = 0$$

auf I , und leitet man dies Gleichung $(\deg(p_{r+1})+1)$ -mal ab, so erhält man wegen $D^{\deg(p_{r+1})+1}p_{r+1} = 0$ eine Gleichung der Form

$$\sum_{j=1}^r q_j(t)e^{(\lambda_j - \lambda_{r+1})t} = 0$$

mit Polynomen q_j . Nach Induktionsvoraussetzung folgt daraus $q_j = 0$ für alle $j = 1, \dots, r$, dies impliziert aber auch $p_j = 0$, da die Ableitung von $r(t)e^{\mu t}$ ($\mu \neq 0$) gerade $(r'(t) + \mu r(t))e^{\mu t}$ ist, wobei $(r' + \mu r)$ denselben Grad wie r hat und insbesondere bei $r \neq 0$ auch ungleich Null ist. Schließlich folgt aus $p_j = 0$ für alle $j = 1, \dots, r$ aber wegen $\sum_{j=1}^{r+1} p_j(t)e^{\lambda_j t} = 0$ auch $p_{r+1} = 0$. \square

Man beachte, dass das gefundene Fundamentalsystem selbst dann aus komplexwertigen Funktionen bestehen kann, wenn P ein reelles Polynom ist, da reelle Polynome i.a. nur über \mathbb{C} in Linearfaktoren zerfallen. Jedoch treten in diesem Fall komplexe Nullstellen λ_j nur in konjugierten Paaren auf, d.h. mit λ_j ist auch $\bar{\lambda}_j$ eine Nullstelle. Ein reelles Fundamentalsystem kann man gewinnen, indem man die komplexen Lösungen $t^{q-1}e^{\lambda_j t}$, $t^{q-1}e^{\bar{\lambda}_j t}$ durch $t^{q-1} \operatorname{Re}(e^{\lambda_j t})$, $t^{q-1} \operatorname{Im}(e^{\lambda_j t})$ ersetzt.

Im Fall einer inhomogenen Differentialgleichung höherer Ordnung macht es für polynomiale, exponentielle oder trigonometrische Inhomogenitäten $b \neq 0$ in (2.8) häufig Sinn, statt mittels Variation der Konstanten durch einen speziellen Ansatz eine partikuläre Lösung zu suchen. Dieser Ansatz sollte dieselbe Form wie b haben, also auch polynomial, exponentiell oder trigonometrisch sein. Bevor wir ein genaueres Resultat dazu formulieren, welche Ansätze garantiert zu einer partikulären Lösung führen, sei noch kurz bemerkt, dass es ausreicht, den Fall $b(t) = p(t)e^{\mu t}$ mit einem Polynom p und einem Exponenten $\mu \in \mathbb{C}$ zu diskutieren. Denn eine allgemeine polynomiale, exponentielle oder trigonometrische Inhomogenität b ist eine Summe $b = \sum_i b_i$ von solchen Funktionen b_i , mit partikulären Lösungen x_i zu den einzelnen b_i ist dann aber auch die Summe $\sum_i x_i$ eine partikuläre Lösung zur Inhomogenität b .

Satz 2.14. *Sei $P \in \mathbb{C}[X]$ ein Polynom vom Grad n , und sei $b(t) = p(t)e^{\mu t}$ mit einem Polynom p und einem Exponenten $\mu \in \mathbb{C}$.*

- *Ist μ keine Nullstelle von P , dann existiert eine partikuläre Lösung x_p von $P(D)x = b$ die die Form $x_p(t) = q(t)e^{\mu t}$ mit einem Polynom q vom gleichen Grad wie p hat.*
- *Ist μ eine Nullstelle von P der Vielfachheit k , dann existiert eine partikuläre Lösung x_p von $P(D)x = b$ die die Form $x_p(t) = t^k q(t)e^{\mu t}$ mit einem Polynom q vom gleichen Grad wie p hat.*

Beweis:

- Hat p den Grad 0, d.h. $p(t) = p_0$ ist konstant, dann gilt nach Lemma 2.11 die Gleichung $P(D)(ce^{\mu t}) = cP(\mu)e^{\mu t}$, also gilt mit dem konstanten Polynom $q(t) := \frac{p_0}{P(\mu)}$ wie behauptet $P(D)(q(t)e^{\mu t}) = p_0e^{\mu t}$.

Den allgemeinen Fall beweisen wir nun per Induktion: Angenommen, die Aussage stimmt für Polynome p vom Grad m . Ist nun p ein Polynom vom Grad $m+1$ und ist \tilde{p} das durch $P(D)(x^{m+1}e^{\mu t}) = \tilde{p}(t)e^{\mu t}$ eindeutig bestimmte Polynom vom Grad $m+1$, dann gibt es eine Konstante c , so dass $p - c\tilde{p}$ ein Polynom ist, dessen Grad höchstens m beträgt. Nach Induktionsvoraussetzung gibt es dann aber ein Polynom \tilde{q} vom Höchstgrad m mit

$$P(D)(\tilde{q}(t)e^{\mu t}) = (p(t) - c\tilde{p}(t))e^{\mu t}.$$

Daraus ergibt sich aber mit dem Polynom $q(x) := cx^{m+1} + \tilde{q}(x)$ vom Höchstgrad $(m+1)$ die Gleichung

$$P(D)(q(t)e^{\mu t}) = p(t)e^{\mu t}.$$

- Da μ eine Nullstelle von P der Vielfachheit k ist, gibt es ein Polynom Q mit $P(X) = Q(X)(X - \mu)^k$ und $Q(\mu) \neq 0$. Nach dem ersten Teil gibt es ein Polynom \tilde{q} vom Höchstgrad m mit $Q(D)(\tilde{q}(t)e^{\mu t}) = p(t)e^{\mu t}$. Zu diesem \tilde{q} gibt es wiederum ein Polynom, dessen k -te Ableitung gerade \tilde{q} ist, und ohne Einschränkung hat dieses Polynom die Form $t^k q(t)$, denn beim Ableiten fallen die Terme bis zum Grad $(k-1)$ weg. Nun gilt

$$(D - \mu)^k(t^k q(t)e^{\mu t}) = (t^k q(t))^{(k)} e^{\mu t} = \tilde{q}(t)e^{\mu t},$$

denn die Ableitung von $r(t)e^{\mu t}$ ist gerade $(r'(t) + \mu r(t))e^{\mu t}$, also ist $(D - \mu)(r(t)e^{\mu t}) = r'(t)e^{\mu t}$. Daher gilt wie gewünscht

$$P(D)(t^k q(t)e^{\mu t}) = Q(D)(D - \mu)^k(t^k q(t)e^{\mu t}) = Q(D)(\tilde{q}(t)e^{\mu t}) = p(t)e^{\mu t}.$$

□

Eulersche Differentialgleichungen

Eine homogene lineare Differentialgleichung n -ter Ordnung der Form

$$t^n x^{(n)} + a_{n-1} t^{n-1} x^{(n-1)} + \dots + a_1 t x' + a_0 x = 0 \quad (2.9)$$

auf $(0, \infty)$ mit Konstanten a_0, \dots, a_{n-1} heißt Eulersche Differentialgleichung. Durch die Substitution $t = e^s$, $x(t) = y(\ln(t))$, oder äquivalenterweise $s = \ln(t)$, $y(s) = x(e^s)$, kann man eine Eulersche Differentialgleichung auf eine homogene lineare Differentialgleichung n -ter Ordnung mit konstanten Koeffizienten zurückführen, denn es gilt

$$\begin{aligned} x(t) &= y(s) \\ tx'(t) &= y'(s) \text{ wegen } y'(s) = x'(t)t \\ t^2 x''(t) &= y''(s) - y'(s) \text{ wegen } y''(s) = x''(t)t^2 + x'(t)t \\ t^3 x'''(t) &= y'''(s) - 3y''(s) + 2y'(s) \text{ wegen } y'''(s) = x'''(t)t^3 + 3x''(t)t^2 + x'(t)t \text{ usw.} \end{aligned}$$

und daher

$$y^{(n)} + b_{n-1} y^{(n-1)} + \dots + b_1 y' + b_0 y = 0$$

mit konstanten Koeffizienten b_0, b_1, \dots, b_{n-1} , die sich aus der Substitution und a_0, \dots, a_{n-1} berechnen lassen.

Beispiel 2.15. Die Eulersche Differentialgleichung $t^2x'' - 3tx' + 7x = 0$ auf dem Intervall $(0, \infty)$ führt durch die Substitution $t = e^s$, $x(t) = y(\ln(t))$, auf die Differentialgleichung $y'' - 4y' + 7y = 0$ mit konstanten Koeffizienten. Diese Gleichung hat das charakteristische Polynom $\lambda^2 - 4\lambda + 7 = 0$ mit den Nullstellen $\lambda_{1,2} = 2 \pm i\sqrt{3}$ und somit die allgemeine reelle Lösung

$$y(s) = c_1 e^{2s} \cos(\sqrt{3}s) + c_2 e^{2s} \sin(\sqrt{3}s).$$

Die allgemeine Lösung der betrachteten Eulerschen Differentialgleichung auf dem Intervall $(0, \infty)$ ist also

$$x(t) = c_1 t^2 \cos(\sqrt{3} \ln(t)) + c_2 t^2 \sin(\sqrt{3} \ln(t)).$$

Kapitel 3

Qualitative Theorie

In der qualitativen Theorie von Differentialgleichungen diskutiert man autonome Systeme von Differentialgleichungen nicht mehr im Hinblick auf Existenz, Eindeutigkeit oder lokale Abhängigkeit der Lösungen von den gegebenen Daten, sondern man versucht, einen Eindruck vom langfristigen Verhalten aller Lösungen zu bekommen. Dazu untersucht man insbesondere spezielle Lösungen wie Ruhelagen (=Gleichgewichte, Equilibria), periodische Orbits oder allgemeinere invariante Teilmengen des Phasenraumes auf Stabilität, denn stabile Lösungen bestimmen das langfristige Verhalten von in ihrer Nähe gelegenen Lösungen.

3.1 Grundlegende Begriffe

Da wir uns beim qualitativen Studium eines autonomen Systems von Differentialgleichungen

$$x' = f(x) \tag{3.1}$$

nicht so sehr für einzelne Lösungen, sondern für die Gesamtheit aller Lösungen und somit für das Phasenportrait interessieren (siehe Definition 1.38), spielt der von (3.1) generierte Lösungsoperator eine entscheidende Rolle. Wir werden üblicherweise voraussetzen, dass $f : \Omega \rightarrow \mathbb{R}^n$ auf einer offenen Teilmenge $\Omega \subset \mathbb{R}^n$ stetig ist und dort einer lokalen Lipschitzbedingung genügt, dann existiert nach Satz 1.32 zu jedem Anfangswert in Ω lokal eine eindeutige Lösung. Würde man desweiteren voraussetzen, dass alle Lösungen sogar auf ganz \mathbb{R} und damit global existieren, so wäre der Lösungsoperator nach Satz 1.13 bereits ein Fluss (siehe Definition 3.1). Diese Voraussetzung ist für manche Anwendungen aber zu stark, daher führen wir hier den etwas schwächeren Begriff eines lokalen Flusses ein.

Definition 3.1. Sei $Q \subset \mathbb{R} \times \Omega$ offen. Eine stetige Abbildung $\Phi: Q \rightarrow \Omega$ heißt lokaler Fluss (=kontinuierliches lokales dynamisches System) auf der Menge Q , falls $\{0\} \times \Omega \subset Q$ und $\Phi(0, \cdot) = \text{Id}_\Omega$ gilt, $\{t \in \mathbb{R} \mid (t, x) \in Q\}$ für jedes $x \in \Omega$ ein Intervall ist, sowie für jedes $(t, x) \in Q$ mit $(s, \Phi(t, x)) \in Q$ auch schon $(s + t, x) \in Q$ und $\Phi(s + t, x) = \Phi(s, \Phi(t, x))$ gilt.

Satz 3.2. Ist $\Omega \subset \mathbb{R}^n$ offen, genügt das stetige Vektorfeld $f: \Omega \rightarrow \mathbb{R}^n$ einer lokalen Lipschitzbedingung, und bezeichnet man die maximale Lösung von $x' = f(x)$ zum Anfangswert $x(0) = x_0$ mit $I_{x_0} \ni t \mapsto \Phi(t, x_0)$, so ist $Q := \{(t, x) \in \mathbb{R} \times \Omega \mid t \in I_x\}$ eine offene Umgebung von $\{0\} \times \Omega$ sowie eine Vereinigung $Q = \bigcup_{x \in \Omega} I_x \times \{x\}$ von Intervallen, und $\Phi: Q \rightarrow \Omega$ ist ein lokaler Fluss. Ist darüberhinaus f stetig differenzierbar, so ist auch Φ (in beiden Variablen) stetig differenzierbar.

Beweis: Der Satz fasst nur den lokalen Existenz- und Eindeutigkeitsatz 1.32, den Fortsetzungssatz 1.35 und unsere Resultate 1.57 bzw. 1.61 über stetige bzw. differenzierbare Abhängigkeit vom Anfangswert für den Fall einer autonomen Differentialgleichung zusammen. \square

Natürlich gelten die in Satz 3.2 formulierten Stetigkeits- und Differenzierbarkeitseigenschaften im Falle der globalen Existenz von Lösungen auch für einen Fluss. Betrachtet man nur nichtnegative Zeiten $t \geq 0$, so spricht man von einem (lokalen) Halbfluss statt von einem (lokalen) Fluss.

Die Menge Ω bezeichnet man als den Phasenraum (=Zustandsraum), auf dem der (lokale) Fluss definiert ist. Die Spur $x(I_{x_0}) = \Phi(I_{x_0}, x_0)$ einer maximalen Lösung x von (3.1) zu einem Anfangswert $x_0 \in \Omega$ bezeichnet man als Orbit (=Zeitorbit, Trajektorie, Phasenkurve) durch x_0 . Betrachtet man nur positive Zeiten, so spricht man vom (positiven) Halborbit durch x_0 . Wie wir aufgrund unserer Existenz- und Eindeutigkeitsresultate wissen, geht durch jeden Punkt des Phasenraums genau ein Orbit.

Spezielle Orbits haben oft einen großen Einfluss auf das langfristige Verhalten der Orbits für $t \rightarrow +\infty$ in einem größeren Teilbereich des Phasenraums. Solche Orbits sind oft sehr einfach, nämlich Ruhelagen oder periodische Orbits.

Definition 3.3. Ein Punkt $x_0 \in \Omega$ heißt *Ruhelage* (=Gleichgewicht, *Equilibrium*) der autonomen Differentialgleichung $x' = f(x)$, falls $f(x_0) = 0$ oder äquivalenterweise $\Phi(t, x_0) = x_0$ für alle $t \in \mathbb{R}$ gilt.

Definition 3.4. Ist x eine T -periodische Lösung der autonomen Differentialgleichung $x' = f(x)$ zum Anfangswert x_0 oder gilt äquivalenterweise $\Phi(T, x_0) = x_0$, so heißt die Spur von x bzw. der Orbit durch x_0 ein *periodischer Orbit*.

Lemma 3.5. Auf jeden Orbit $x(I_{x_0}) = \Phi(I_{x_0}, x_0)$ des von (3.1) generierten lokalen Flusses Φ trifft genau eine der folgenden Aussage zu:

- Der Orbit ist ein Punkt und somit eine Ruhelage.
- Der Orbit ist eine geschlossene Kurve und somit ein periodischer Orbit.
- Der Orbit ist nicht geschlossen, und dann ist $t \mapsto \Phi(t, x_0)$ eine Bijektion von I_{x_0} auf den Orbit.

Beweis: Die Abbildung $t \mapsto \Phi(t, x_0)$ auf dem maximalen Existenzintervall I_{x_0} der Lösung durch x_0 ist immer eine Surjektion auf den Orbit durch x_0 . Tritt also nicht der letzte Fall ein, dann muss $t \mapsto \Phi(t, x_0)$ nicht injektiv sein, d.h. es gibt ein $t \in I_{x_0}$ und ein $T > 0$ mit $t + T \in I_{x_0}$ und $\Phi(t + T, x_0) = \Phi(t, x_0)$. Dann gilt aufgrund der Eindeutigkeit aber schon $I_{x_0} = \mathbb{R}$ und $\Phi(t + T, x_0) = \Phi(t, x_0)$ für alle $t \in \mathbb{R}$. Kann man T beliebig klein wählen, dann ist der Orbit eine Ruhelage, ansonsten gibt es ein minimales $T > 0$ mit dieser Eigenschaft und der Orbit gehört zu einer periodischen Lösung mit Periode T . \square

Um das langfristige Verhalten eines (kontinuierlichen) dynamischen Systems erfassen zu können, d.h. das Verhalten der Lösungen für $t \rightarrow +\infty$, erweisen sich verschiedene Stabilitätsbegriffe als hilfreich. Vor deren Formulierung sei daran erinnert, dass $U \subset \mathbb{R}^n$ eine Umgebung der Teilmenge $A \subset \mathbb{R}^n$ heißt, falls es eine offene Menge $V \subset \mathbb{R}^n$ mit $A \subset V \subset U$ gibt. In den folgenden Definitionen beziehen wir uns immer auf ein fest gewähltes stetiges Vektorfeld $f : \Omega \rightarrow \mathbb{R}^n$, das einer lokalen Lipschitzbedingung genügt, und den von f generierten lokalen Fluss Φ .

Definition 3.6. Eine Menge $A \subset \Omega$ heißt (Lyapunov-)stabil unter dem von (3.1) generierten lokalen Fluss Φ , wenn es für jede Umgebung V von A eine Umgebung U von A gibt, für die jede Lösung zu einem Anfangswert $x \in U$ auf $[0, +\infty)$ definiert ist und $\Phi(t, x) \in V$ für alle $t \geq 0$ und $x \in U$ erfüllt.

Stabilität ist in gewisser Weise eine Verschärfung der auf kompakten Zeitintervallen vorhandenen stetigen Abhängigkeit. Tatsächlich, ist x_0 eine Ruhelage, so besagt die stetige Abhängigkeit vom Anfangswert gerade, dass $\lim_{x \rightarrow x_0} \Phi(t, x) = \Phi(t, x_0) = x_0$ gleichmäßig für alle $t \in [0, T]$ gilt. Dies bedeutet aber nichts anderes, als dass es für jede Umgebung V von $A := \{x_0\}$ eine Umgebung U von A mit $\Phi(t, x) \in V$ für alle $t \in [0, T]$ und $x \in U$ gibt. Die Ruhelage $\{x_0\}$ ist somit genau dann stabil, wenn diese Bedingung nicht nur auf $[0, T]$ sondern sogar auf $[0, +\infty)$ gilt.

Stabilität einer Ruhelage x_0 bedeutet jedoch noch nicht, dass in der Nähe von x_0 gestartete Lösungen langfristig auch gegen x_0 konvergieren. Beispielsweise kreisen bei den unter Punkt (d) von Abschnitt 3.2.2 diskutierten linearen Flüssen die Orbits um die Ruhelage 0, ohne ihr beliebig nahe zu kommen. Eine Menge, gegen die alle in der Nähe gestartete Lösungen konvergieren, nennt man einen Attraktor.

Definition 3.7. Eine Menge $A \subset \Omega$ heißt ein Attraktor des von (3.1) generierten lokalen Fluss Φ , falls es eine Umgebung U von A gibt, für die jede Lösung zu einem Anfangswert $x \in U$ auf $[0, +\infty)$ definiert ist und $\lim_{t \rightarrow +\infty} \text{dist}(\Phi(t, x), A) = 0$ erfüllt.

Andererseits muss ein Attraktor A nicht stabil sein, es könnte nämlich sein, dass sich beliebig nah bei A gestartete Lösungen zunächst einmal bis zu einer festen Distanz von A entfernen und erst danach die Distanz zu A wieder gegen Null konvergiert, siehe [Amann, (15.1.d)]. Ist dies nicht der Fall, so nennen wir A asymptotisch stabil.

Definition 3.8. Ist eine Menge $A \subset \Omega$ sowohl (Lyapunov-)stabil als auch ein Attraktor, so heißt sie asymptotisch stabil.

Zu guter letzt wollen wir noch sagen, wann wir zwei lokale Flüsse Φ auf Ω und $\tilde{\Phi}$ auf $\tilde{\Omega}$ als äquivalent betrachten wollen. Dies soll der Fall sein, wenn es eine topologische Konjugation zwischen Ihnen gibt.

Definition 3.9. Zwei lokale Flüsse Φ auf Ω (definiert auf Q) und $\tilde{\Phi}$ auf $\tilde{\Omega}$ (definiert auf \tilde{Q}) heißen topologisch äquivalent, wenn es eine topologische Konjugation zwischen ihnen gibt, d.h. eine Zahl $a > 0$ und einen Homöomorphismus $h : \Omega \rightarrow \tilde{\Omega}$ mit der Eigenschaft

$$\forall (t, x) \in Q : h(\Phi(t, x)) = \tilde{\Phi}(at, h(x)).$$

Sind die Flüsse Φ und $\tilde{\Phi}$ topologisch äquivalent, dann sieht das Phasenportrait von Φ samt zeitlicher Orientierung der Orbits in den neuen durch h gegebenen Koordinaten nämlich genauso aus wie das Phasenportrait von $\tilde{\Phi}$.

3.2 Lineare Flüsse

In diesem Abschnitt wollen wir das qualitative Verhalten von homogenen linearen Systemen mit konstanten Koeffizienten diskutieren, d.h. von Differentialgleichungen der Form

$$x' = Ax$$

mit einer konstanten Matrix $A \in \mathbb{R}^{n \times n}$. Wie wir schon aus Abschnitt 2.1 wissen, erzeugen solche Differentialgleichungen den (globalen) Fluss $\exp(tA)$. Da jede einzelne Abbildung $\exp(tA)$ linear ist, nennt man diesen Fluss linear.

Bezeichnet $B = S^{-1}AS$ die Jordansche Normalform von A , dann ist der Fluss $\exp(tA)$ zum Fluss $\exp(tB)$ topologisch äquivalent, indem man die Zeit unverändert läßt ($a = 1$) und als topologische Konjugation die lineare Abbildung $h(x) := S^{-1}x$ wählt, denn es gilt $S^{-1}\exp(tA)x = \exp(tB)S^{-1}x$. Tatsächlich sind die Flüsse $\exp(tA)$ und $\exp(tB)$ linear äquivalent, d.h. h ist nicht nur ein Homöomorphismus sondern sogar eine bijektive lineare Abbildung. Daher reicht es bei einer qualitativen Diskussion linearer Flüsse aus, $\exp(tA)$ für eine Matrix A in Jordanscher Normalformen zu untersuchen.

3.2.1 Stabilität

Zunächst wollen wir klären, wann die Ruhelage 0 eines homogenen linearen Systems $x' = Ax$ mit konstanten Koeffizienten stabil bzw. asymptotisch stabil ist.

Satz 3.10. *Der Nullpunkt von $x' = Ax$ ist genau dann eine stabile Ruhelage, wenn für alle Eigenwerte λ von A die Ungleichung $\operatorname{Re}(\lambda) \leq 0$ gilt und jeder Eigenwert λ mit $\operatorname{Re}(\lambda) = 0$ halbeinfach ist, d.h. seine algebraische und geometrische Vielfachheit stimmen überein.*

Beweis: Jede Lösung $x(t) = \exp(tA)x_0$ von $x' = Ax$ ist eine Linearkombination von Kurven der Form $t^k \exp(\lambda t)w$ mit Eigenwerten λ von A und Eigen- bzw. Hauptvektoren w von A . Sind also die angegebenen Voraussetzungen an die Eigenwerte erfüllt, dann gilt $\sup_{t \geq 0} |\exp(tA)| < \infty$, denn entweder gilt $\operatorname{Re}(\lambda) < 0$ und damit $t^k \exp(\lambda t) \rightarrow 0$ für $t \rightarrow \infty$, oder es gilt $\operatorname{Re}(\lambda) = 0$ und $k = 0$, so dass $t^k \exp(\lambda t)$ auf $[0, \infty)$ beschränkt bleibt. Wählt man also zu $\varepsilon > 0$ den Radius r so klein, dass $\left(\sup_{t \geq 0} |\exp(tA)|\right) r < \varepsilon$ gilt, dann folgt aus $x_0 \in B_r(0)$ schon $|\exp(tA)x_0| \leq |\exp(tA)| |x_0| < \varepsilon$. Dies bedeutet aber gerade, dass 0 eine stabile Ruhelage ist.

Ist umgekehrt eine der angegebenen Voraussetzungen verletzt, so gibt es ein $x_0 \neq 0$ mit $|\exp(tA)x_0| \rightarrow \infty$. Da dann auch $|\exp(tA)(\varepsilon x_0)| \rightarrow \infty$ für jedes $\varepsilon > 0$ gilt, ist 0 eine instabile Ruhelage. \square

Korollar 3.11. *Gibt es einen Eigenwert λ von A mit $\operatorname{Re}(\lambda) > 0$, so ist der Nullpunkt von $x' = Ax$ eine instabile Ruhelage.*

Satz 3.12. *Der Nullpunkt von $x' = Ax$ ist genau dann eine asymptotisch stabile Ruhelage, wenn für alle Eigenwerte λ von A die Ungleichung $\operatorname{Re}(\lambda) < 0$ gilt.*

Beweis: Wie Beweis von Satz 3.10 ist jede Lösung $x(t) = \exp(tA)x_0$ von $x' = Ax$ eine Linearkombination von Kurven der Form $t^k \exp(\lambda t)w$ mit Eigenwerten λ von A und Eigen- bzw. Hauptvektoren w von A . Ist also die angegebene Voraussetzung an die Eigenwerte erfüllt, dann gilt $\lim_{t \rightarrow \infty} \exp(tA)x_0 = 0$ wegen $t^k \exp(\lambda t) \rightarrow 0$ für $t \rightarrow \infty$ für λ mit $\operatorname{Re}(\lambda) < 0$. Also gilt auch $\exp(tA) \rightarrow 0$ in $L(\mathbb{R}^n)$ und somit ist 0 eine asymptotisch stabile Ruhelage.

Ist umgekehrt die angegebene Voraussetzung verletzt, so gibt es ein $x_0 \neq 0$ mit $|\exp(tA)x_0| \rightarrow \infty$ (im Fall dass einer Term $t^k \exp(\lambda t)w$ mit $\operatorname{Re}(\lambda) > 0$ oder mit $\operatorname{Re}(\lambda) = 0$ sowie $k > 0$ vorkommt) oder ein $x_0 \neq 0$ mit $|\exp(tA)x_0| = |x_0|$ (im Fall dass einer Term $\exp(\lambda t)w$ mit $\operatorname{Re}(\lambda) = 0$ vorkommt). \square

3.2.2 Ebene lineare Flüsse

Nachdem wir die Stabilität der Ruhelage 0 eines homogenen linearen Systems geklärt haben, wollen wir nun genauer wissen, wie sich die Lösungskurven eines homogenen linearen Systems qualitativ verhalten. Dazu betrachten wir zunächst ebene Systeme $x' = Ax$ mit $A \in \mathbb{R}^{2 \times 2}$. Um die verschiedenen Möglichkeiten der Jordanschen Normalform zu diskutieren, unterscheiden wir die folgenden Fälle.

- (a) A hat reelle Eigenwerte $\neq 0$ unterschiedlichen Vorzeichens: Dann ist die Ruhelage 0 instabil und A hat die Jordansche Normalform

$$B = S^{-1}AS = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

Nach der Koordinatentransformation mittels $y = h(x) = S^{-1}x$ hat die Lösung also die Form $\exp(tB)y = (e^{t\lambda_1}y_1, e^{t\lambda_2}y_2)$. Ist $\lambda_1 < 0 < \lambda_2$, dann streben alle Lösungen zu Anfangswerten mit $y_2 = 0$ gegen 0, während bei $y_2 \neq 0$ die Lösungen gegen Unendlich laufen. Die Ruhelage 0 nennt man in diesem Fall einen Sattel.

- (b) A hat nur Eigenwerte mit negativem Realteil: Dann ist die Ruhelage 0 asymptotisch stabil und man bezeichnet sie auch als Senke. Es treten die folgenden unterschiedlichen Fälle auf:

- Die Eigenwerte sind reell und halbeinfach, dann hat A wiederum die Jordansche Normalform

$$B = S^{-1}AS = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix},$$

nur diesmal mit $\lambda_1 < \lambda_2 < 0$ bzw. $\lambda_1 = \lambda_2 < 0$. In den neuen Koordinaten lautet die Lösung $\exp(tB)y = (e^{t\lambda_1}y_1, e^{t\lambda_2}y_2)$ und man bezeichnet die Ruhelage 0 als stabilen Knoten bzw. stabilen Fokus.

- Es gibt einen nicht halbeinfachen reellen Eigenwert, dann hat A die Jordansche Normalform

$$B = S^{-1}AS = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$$

mit $\lambda < 0$. In den neuen Koordinaten lautet die Lösung $\exp(tB)y = (e^{t\lambda}y_1 + te^{t\lambda}y_2, e^{t\lambda}y_2)$ und man bezeichnet die Ruhelage 0 als uneigentlichen stabilen Knoten.

- Die Eigenwerte sind komplex und haben somit die Form $\lambda = a + i\omega$ und $\bar{\lambda} = a - i\omega$ mit $a < 0$ und $\omega > 0$. In diesem Fall hat A die reelle Jordansche Normalform

$$B = S^{-1}AS = \begin{pmatrix} a & -\omega \\ \omega & a \end{pmatrix}.$$

Nach der Koordinatentransformation mittels $y = h(x) = S^{-1}x$ hat die Lösung also die Form $\exp(tB)y = e^{ta}(\cos(\omega t)y_1 - \sin(\omega t)y_2, \sin(\omega t)y_1 + \cos(\omega t)y_2)$ und man bezeichnet die Ruhelage 0 als stabilen Strudel.

- (c) A hat nur Eigenwerte mit positivem Realteil: Dann ist die Ruhelage 0 instabil und man bezeichnet sie auch als Quelle. Es treten analog zum vorigen Fall unterschiedliche Fälle auf, in denen man die Ruhelage 0 als instabilen (uneigentlichen) Knoten, Fokus und Strudel bezeichnet.

- (d) A hat rein imaginäre Eigenwerte: Dann haben die Eigenwerte die Form $\lambda = i\omega$ und $\bar{\lambda} = -i\omega$ mit $\omega > 0$ und A hat die reelle Jordansche Normalform

$$B = S^{-1}AS = \begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix}.$$

Alle Lösungen sind periodisch mit Periode $\frac{2\pi}{\omega}$, und man nennt die Ruhelage 0 ein Zentrum oder einen Wirbel. Die Ruhelage ist stabil, aber nicht asymptotisch stabil.

- (e) A hat einen Nulleigenwert: Dann gibt es eine Gerade von Ruhelagen und einen Eigenwert $\lambda \neq 0$, jede Lösung hat in den neuen Koordinaten also die Form $\exp(tB)y = (y_1, e^{t\lambda}y_2)$, oder A ist die Nullmatrix und alle Lösungen sind Ruhelagen.

3.2.3 Klassifikation hyperbolischer linearer Flüsse

Im vorigen Abschnitt haben wir ebene lineare Flüsse bis auf lineare Äquivalenz klassifiziert, wir haben nämlich das Verhalten aller Lösungen in Koordinaten angegeben, in denen A Jordansche Normalform hat. In höheren Dimensionen bereiten einem bei einer solchen Klassifikation besonders die Analoga der letzten beiden Fälle große Probleme, bei denen es einen Eigenwert mit Realteil 0 gibt, denn in diesem Fall muß man sehr viele Fallunterscheidungen machen. Daher betrachtet man in höheren Dimensionen üblicherweise zunächst einmal hyperbolische lineare Flüsse.

Definition 3.13. *Ein linearer Fluss $\exp(tA)$ heißt hyperbolisch, falls kein Eigenwert von A einen verschwindenden Realteil besitzt.*

Für einen hyperbolischen Fluss ist nur 0 eine Ruhelage, denn $Ax_0 = 0$ für ein $x_0 \neq 0$ würde bedeuten, dass 0 ein Eigenwert von A ist. Der Nullpunkt ist nach Satz 3.12 genau dann asymptotisch stabil, d.h. eine Senke, wenn $\operatorname{Re}(\lambda) < 0$ für alle Eigenwerte λ von A gilt, und er ist eine Quelle, d.h. $|\exp(tA)x_0| \rightarrow \infty$ gilt bei $t \rightarrow \infty$ für alle $x_0 \neq 0$, wenn $\operatorname{Re}(\lambda) > 0$ für alle Eigenwerte λ von A gilt.

Mit Hilfe des folgenden Lemmas wollen wir noch einmal quantitativ festhalten, dass die Lösungen sogar exponentiell schnell gegen 0 konvergieren, falls 0 eine asymptotisch stabile Ruhelage ist.

Lemma 3.14. *Gilt mit einer reellen Zahl $\alpha \in \mathbb{R}$ die Ungleichung $\operatorname{Re}(\lambda) < \alpha$ für jeden Eigenwert λ von $A \in \mathbb{R}^{n \times n}$, dann gibt es eine von einem Skalarprodukt induzierte Norm $\|\cdot\|$ auf \mathbb{R}^n mit $\|\exp(tA)\| \leq e^{t\alpha}$.*

Beweis: Fasst man A als komplexe Matrix auf, dann hat A bzgl. einer geeigneten Basis die Form $A = D + N$ mit einer Diagonalmatrix $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, wobei die Eigenwerte λ_i von A möglicherweise mehrfach vorkommen, und einer nilpotenten Matrix N , d.h. es gilt $N^n = 0$. Dabei gilt $DN = ND$ und man kann die Basis e_1, \dots, e_n so wählen, dass $Ne_i = e_{i-1}$ oder $Ne_i = 0$ gilt. Ersetzt man e_i durch $\tilde{e}_i := \delta^i e_i$ mit einem $\delta > 0$, so bleibt in der Darstellung von A bzgl. dieser neuen Basis D unverändert, während $N\tilde{e}_i = \delta\tilde{e}_i$ oder $N\tilde{e}_i = 0$ gilt. Also hat N höchstens in der rechten oberen Nebendiagonalen die von Null verschiedenen Einträge δ . Indem man die vom Euklidischen Skalarprodukt zu dieser Basis \tilde{e}_i induzierte Norm verwendet und δ genügen klein wählt, erhält man also zu jedem $\varepsilon > 0$ eine von einem Skalarprodukt induzierte Norm $\|\cdot\|$ mit $\|N\| \leq \varepsilon$. Die Norm $\|D\|$ von D ist offensichtlich das Maximum der Beträge aller Eigenwerte, daher erhält man für $\varepsilon > 0$ so klein, dass $\operatorname{Re}(\lambda) + \varepsilon \leq \alpha$ gilt, die Ungleichung

$$\|\exp(tD)\| = \max_{i=1, \dots, n} e^{t \operatorname{Re}(\lambda_i)} \leq e^{t(\alpha - \varepsilon)}.$$

Zusammen mit der Abschätzung $\|N\| \leq \varepsilon$ gilt also die gewünschte Ungleichung

$$\|\exp(tA)\| = \|\exp(tD)\exp(tN)\| \leq \|\exp(tD)\| \|\exp(tN)\| \leq e^{t(\alpha-\varepsilon)} e^{t\|N\|} \leq e^{t\alpha}.$$

□

Satz 3.15. Für $A \in \mathbb{R}^{n \times n}$ sind äquivalent:

- (a) Der Nullpunkt ist eine asymptotisch stabile Ruhelage (oder Senke) von $x' = Ax$.
- (b) Es gibt Konstanten $\alpha > 0$ und $\beta \geq 1$ mit $|\exp(tA)x| \leq \beta e^{-t\alpha}|x|$ für alle $t \geq 0$ und $x \in \mathbb{R}^n$.
- (c) Es gibt eine Konstante $\alpha > 0$ und eine von einem Skalarprodukt induzierte Norm $\|\cdot\|$ auf \mathbb{R}^n mit $\|\exp(tA)x\| \leq e^{-t\alpha}\|x\|$ für alle $t \geq 0$ und $x \in \mathbb{R}^n$.

Beweis: Wendet man Satz 3.12 und Lemma 3.14 mit einem $\alpha > 0$ an, für das $\operatorname{Re}(\lambda) < -\alpha$ für alle Eigenwerte λ von A gilt, so ergibt sich die Äquivalenz von (i) und (iii). Der Punkt (ii) ist aber zu (iii) äquivalent, da es aufgrund der Äquivalenz aller Normen zu der Norm $|\cdot|$ Konstanten $c, C > 0$ mit $c|x| \leq \|x\| \leq C|x|$ für alle $x \in \mathbb{R}^n$ gibt, man kann also $\beta = \frac{C}{c}$ wählen. □

Analog ist der Nullpunkt genau dann eine Quelle von $x' = Ax$, wenn $|\exp(tA)x| \geq \beta e^{t\alpha}|x|$ mit Konstanten $\alpha > 0$, $\beta \geq 1$, für alle $t \geq 0$ und $x \in \mathbb{R}^n$ gilt. Die vorigen beiden Resultate wollen wir nutzen, um zu zeigen, dass sich bei einem hyperbolischen linearen Fluß der Phasenraum in stabile und instabile Richtungen zerlegen läßt, die unter dem Fluss invariant sind und auf denen der Fluss kontrahierend bzw. expandierend ist. Dazu nennen wir einen linearen Fluss kontrahierend, wenn der Nullpunkt eine Senke ist, und expandierend, wenn der Nullpunkt eine Quelle ist.

Satz 3.16. Ist $\exp(tA)$ ein hyperbolischer linearer Fluss, dann kann man den Phasenraum eindeutig in eine direkte Summe $X_s \oplus X_u$ von Unterräumen zerlegen, so dass A eine Zerlegung $A = A_s \oplus A_u$ mit $A_s : X_s \rightarrow X_s$, $A_u : X_u \rightarrow X_u$, besitzt und $\exp(tA_s)$ auf X_s kontrahierend ist, während $\exp(tA_u)$ auf X_u expandierend ist.

Beweis: Wir fassen wiederum A als komplexe $(n \times n)$ -Matrix auf und definieren X_s als die Summe aller verallgemeinerten Eigenräume zu Eigenwerten λ mit $\operatorname{Re}(\lambda) < 0$, d.h.

$$X_s := \bigoplus_{\{\lambda \in \mathbb{C} \mid \lambda \text{ ist EW von } A, \operatorname{Re}(\lambda) < 0\}} \operatorname{Ker}((A - \lambda \operatorname{Id})^{m(\lambda)})$$

mit der algebraischen Vielfachheit $m(\lambda)$ von λ (die man auch durch $k + 1$ ersetzen könnte, wenn es höchstens einen Hauptvektor der Stufe k zu λ gibt). Analog definieren wir X_u als die Summe aller verallgemeinerten Eigenräume zu Eigenwerten λ mit $\operatorname{Re}(\lambda) > 0$ und erhalten damit aufgrund der der Hyperbolizität von A die Zerlegungen

$$\mathbb{C}^n = X_s \oplus X_u \text{ sowie } A = A_s \oplus A_u$$

mit den Einschränkungen A_s bzw. A_u von A auf die invarianten Unterräume X_s bzw. X_u . Da A_s auf X_s nur Eigenwerte λ mit $\operatorname{Re}(\lambda) < 0$ besitzt, ist $\exp(tA_s)$ nach Satz 3.15 kontrahierend, während $\exp(tA_u)$ expandierend ist.

Wäre $\mathbb{C}^n = X_1 \oplus X_2$ eine andere Zerlegung mit dieser Eigenschaft, dann gäbe es zu $x \in X_1$ Vektoren $y \in X_s$, $z \in X_u$, mit $x = y + z$. Aufgrund von $\exp(tA)x \rightarrow 0$ für $t \rightarrow \infty$ würde daraus auch $\exp(tA)y \rightarrow 0$ folgen. Nun liegt z aber im instabilen Unterraum X_u , so dass aufgrund

von $|\exp(tA)z| \geq \beta e^{t\alpha}|z|$ mit $\alpha > 0$, $\beta \geq 1$, nur noch $z = 0$ gelten kann. Daher ist $X_1 \subset X_s$. Umgekehrt erhält man auch $X_s \subset X_1$ und daher $X_1 = X_s$, und analog $X_2 = X_u$, d.h. die Zerlegung ist eindeutig. \square

Bei hyperbolischen linearen Flüssen mit $X_s, X_u \neq 0$ verhält sich die Ruhelage 0 also ganz ähnlich wie ein Sattel bei einem ebenen linearen Fluss, nur dass innerhalb der Räume X_s bzw. X_u das Verhalten viel komplexer sein kann als im ebenen Fall, wo diese Räume eindimensional sind. Durch dieses Verhalten unterscheiden sich lineare Flüsse bei einer Klassifikation bis auf lineare Äquivalenz recht stark, wie der folgende Satz besagt, bei dessen Formulierung $\sigma(A)$ die Menge aller Eigenwerte der Matrix A bezeichnet, die das Spektrum von A genannt wird.

Satz 3.17. *Zwei lineare Flüsse $\exp(tA)$ und $\exp(tB)$ sind genau dann linear äquivalent, wenn ein $a > 0$ existiert mit $\sigma(A) = \sigma(aB)$ und die geometrischen und algebraischen Vielfachheiten der korrespondierenden Eigenwerte übereinstimmen.*

Beweis: Sind $\exp(tA)$ und $\exp(tB)$ linear äquivalent, dann gibt es eine Zahl $a > 0$ und eine lineare Bijektion $h : \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit $h \circ \exp(tA) = \exp(atB) \circ h$ für alle $t \in \mathbb{R}$. Also gilt mit einer h repräsentierenden Matrix S die Gleichung

$$\exp(tA) = S^{-1} \exp(atB) S = \exp(atS^{-1}BS)$$

für alle $t \in \mathbb{R}$ und daher $A = S^{-1}(aB)S$. Daraus ergibt sich wegen

$$\det(A - \lambda \text{Id}) = \det(S^{-1}(aB)S - \lambda \text{Id}) = \det(aB - \lambda \text{Id})$$

die Gleichung $\sigma(A) = \sigma(aB)$ und wegen $A = S^{-1}(aB)S$ auch die Übereinstimmung der geometrischen und algebraischen Vielfachheiten der korrespondierenden Eigenwerte.

Gilt umgekehrt $\sigma(A) = \sigma(aB)$ und stimmen die geometrischen und algebraischen Vielfachheiten der korrespondierenden Eigenwerte überein, dann haben A und aB die gleiche Jordansche Normalform. Somit gibt es eine invertierbare Matrix S mit $A = S^{-1}(aB)S$, und mit $h(x) := Sx$ folgt die lineare Äquivalenz aus

$$\exp(tA) = \exp(atS^{-1}BS) = S^{-1} \exp(atB) S.$$

\square

Lässt man in Satz 3.17 C^1 -Diffeomorphismen h statt nur linearer Abbildungen bei der Äquivalenz zu, so ändert sich nichts an der Aussage des Satzes. Erlaubt man aber Homöomorphismen, dann kann man beispielsweise einen stabilen Strudel in einen stabilen Knoten transformieren. Dadurch erscheinen bei einer Klassifikation hyperbolischer linearer Flüsse bis auf topologische Äquivalenz viele Flüsse als gleich, die bei einer Klassifikation bis auf lineare Äquivalenz noch als verschieden angesehen wurden. Genauer besagt der folgende Satz, bei dem X_s^A bzw. X_s^B den in Satz 3.16 gefundenen stabilen Unterraum zu A bzw. B bezeichnet.

Satz 3.18. *Zwei hyperbolische lineare Flüsse $\exp(tA)$ und $\exp(tB)$ auf dem \mathbb{R}^n sind genau dann topologisch äquivalent, wenn $\dim(X_s^A) = \dim(X_s^B)$ gilt.*

Beweis: Ist $\mathbb{R}^n = X_s^A \oplus X_u^A$ eine Zerlegung für $A = A_s \oplus A_u$ wie in Satz 3.16, dann sind nach Lemma 3.19 die Flüsse $\exp(tA_s)$ und $\exp(-t \text{Id}_{X_s^A})$ topologisch äquivalent, und analog sind $\exp(tA_u)$ und $\exp(t \text{Id}_{X_u^A})$ topologisch äquivalent. Sind h_s^A und h_u^A die entsprechenden Homöomorphismen, dann ist auch $h^A(x_s + x_u) := h_s^A(x_s) + h_u^A(x_u)$ ein Homöomorphismus, der eine topologische Äquivalenz zwischen $\exp(tA)$ und $\exp(-t \text{Id}_{X_s^A}) \oplus \exp(t \text{Id}_{X_u^A})$ herstellt. Ebenso ist $\exp(tB)$

zu $\exp(-t \text{Id}_{X_s^B}) \oplus \exp(t \text{Id}_{X_u^B})$ topologisch äquivalent. Aufgrund der Gleichheit der Dimensionen $\dim(X_s^A) = \dim(X_s^B)$ und $\dim(X_u^A) = \dim(X_u^B)$ sind aber auch $\exp(-t \text{Id}_{X_s^A}) \oplus \exp(t \text{Id}_{X_u^A})$ und $\exp(-t \text{Id}_{X_s^B}) \oplus \exp(t \text{Id}_{X_u^B})$ topologisch äquivalent, so dass man insgesamt die topologische Äquivalenz von $\exp(tA)$ und $\exp(tB)$ erhält.

Ist umgekehrt $a > 0$ und h ein Homöomorphismus auf \mathbb{R}^n mit $h \circ \exp(tA) = \exp(atB) \circ h$, dann folgt $h(X_s^A) \subset X_s^B$, da h die Konvergenz gegen 0 für $t \rightarrow \infty$ erhält. Ebenso folgt $h^{-1}(X_s^B) \subset X_s^A$, also gilt $h(X_s^A) = X_s^B$, d.h. h bildet den Vektorraum X_s^A homöomorph auf X_s^B ab. Daher müssen aber die Dimensionen von X_s^A und X_s^B übereinstimmen. \square

Lemma 3.19. *Gelte für jeden Eigenwert λ von $A \in \mathbb{R}^{n \times n}$ die Ungleichung $\text{Re}(\lambda) < 0$. Dann ist $\exp(tA)$ topologisch äquivalent zu $\exp(-t \text{Id})$, wobei man die Zeit unverändert lassen kann.*

Beweis: Um einen Homöomorphismus $h: \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit $h \circ \exp(tA) = \exp(-t \text{Id}) \circ h$ zu definieren, setze man $h(0) := 0$ und $h(x) := e^t \frac{\exp(tA)x}{|\exp(tA)x|}$ für $x \neq 0$, wobei t die reelle Zahl mit $\|\exp(tA)x\| = 1$ für eine (fest gewählte) von einem Skalarprodukt induzierte Norm $\|\cdot\|$ mit $\|\exp(tA)\| \leq e^{-t\alpha}$ ($\text{Re}(\lambda) < -\alpha < 0$) sei, siehe Satz 3.15.

Tatsächlich gibt es genau ein solches t , denn $\|\exp(tA)x\| \leq e^{-t\alpha}\|x\|$ und

$$\|x\| = \|\exp(tA) \exp(-tA)x\| \leq e^{-t\alpha} \|\exp(-tA)x\| \implies \|\exp(-tA)x\| \geq e^{t\alpha} \|x\|$$

gilt für $t \geq 0$, also schneidet jeder Orbit $\exp(tA)x$ die Einheitssphäre $\mathbb{S} := \{x \mid \|x\| = 1\}$ bzgl. der Norm $\|\cdot\|$ in genau einem Punkt. Aus dieser Tatsache folgt zusätzlich, dass

$$\mathbb{R} \times \mathbb{S} \ni (t, x) \mapsto \exp(tA)x \in \mathbb{R}^n \setminus \{0\}$$

eine stetige Bijektion und aufgrund der Kompaktheit von \mathbb{S} sogar ein Homöomorphismus ist. Da mit der Einheitssphäre $S := \{x \mid |x| = 1\}$ bzgl. der ursprünglichen Norm $|\cdot|$ außerdem $\mathbb{S} \ni x \mapsto \frac{x}{|x|} \in S$ ein Homöomorphismus ist, ist auch $h: x \mapsto e^t \frac{\exp(tA)x}{|\exp(tA)x|}$ ein Homöomorphismus von $\mathbb{R}^n \setminus \{0\}$ auf sich.

Tatsächlich ist h sogar in 0 stetig und somit ein Homöomorphismus von \mathbb{R}^n auf sich, denn zu jeder Umgebung V von 0 innerhalb der Einheitskugel bzgl. der Norm $\|\cdot\|$ gibt es ein $t_0 > 0$ mit $e^{-t}S \subset V$ für alle $t \geq t_0$. Mit der offenen Umgebung

$$U := \left\{ x \mid \|x\| < \frac{1}{\|\exp(-t_0A)\|} \right\}$$

von 0 gilt dann $h(U) \subset V$. Tatsächlich, $\|\exp(-t_0A)x\| < 1$ gilt für $x \in U$ und daher

$$\|\exp((t - t_0)A)x\| \leq e^{-\alpha t} \|\exp(-t_0A)x\| < e^{-\alpha t} \leq 1,$$

also gilt für das eindeutige $t = t(x)$ mit $\|\exp(tA)x\| = 1$ die Ungleichung $t < -t_0$, aus der $h(x) = e^t \frac{\exp(tA)x}{|\exp(tA)x|} \in e^t S \subset V$ folgt.

Zuletzt muss noch gezeigt werden, dass h eine topologische Äquivalenz der Flüsse $\exp(tA)$ und $\exp(-t \text{Id})$ konstituiert. Gerade so war aber h definiert, denn ist $0 \neq x = \exp(sA)y$ mit $(s, y) \in \mathbb{R} \times \mathbb{S}$, dann gilt

$$h(\exp(tA)x) = h(\exp((t + s)A)y)) = e^{-(t+s)} \frac{y}{|y|} = e^{-t} \left(e^{-s} \frac{\exp(-sA)x}{|\exp(-sA)x|} \right) = e^{-t} h(x).$$

\square

3.3 Linearisierung nichtlinearer Flüsse

Nachdem wir im vorigen Abschnitt lineare Flüsse diskutiert haben, wenden wir uns in diesem Abschnitt nichtlinearen lokalen Flüssen zu und beweisen Kriterien, unter denen eine Ruhelage asymptotisch stabil bzw. instabil ist. Danach klassifizieren wir nichtlineare lokale Flüsse in der Nähe einer hyperbolischen Ruhelage bis auf topologische Äquivalenz. In beiden Fällen werden wir den nichtlinearen lokalen Fluss als Störung des linearen Flusses interpretieren, der durch Linearisierung der rechten Seite in der Ruhelage entsteht.

3.3.1 Linearisierte Stabilität von Ruhelagen

Zur Untersuchung der Stabilität von Ruhelagen bei nichtlinearen lokalen Flüssen bemerke man, dass man jede nichtlineare autonome Differentialgleichung $x' = f(x)$ mit einer stetig differenzierbaren rechten Seite f in der Nähe einer Ruhelage x_0 mit $A := \frac{\partial f}{\partial x}(x_0)$ als Störung $x' = A(x - x_0) + g(x - x_0)$ der linearen Differentialgleichung $x' = A(x - x_0)$ auffassen kann, wobei $g(x - x_0) = o(|x - x_0|)$ für $x \rightarrow x_0$ gilt. Insbesondere ist $x' = f(x)$ nahe der Ruhelage x_0 durch die Transformation $y := x - x_0$ zur gestörten linearen Differentialgleichung $y' = Ay + g(y)$ mit $g(y) = o(|y|)$ für $y \rightarrow 0$ nahe der Ruhelage 0 äquivalent. Nun ist aber y genau dann eine Lösung dieser Differentialgleichung zum Anfangswert $y(t_0) = y_0$, wenn y der Integralgleichung

$$y(t) = \exp((t - t_0)A)y_0 + \int_{t_0}^t \exp((t - s)A)g(y(s)) ds \quad (3.2)$$

genügt. Aus dieser Beobachtung ergibt sich das folgende Prinzip der linearisierten Stabilität gewinnen.

Satz 3.20. *Sei $\Omega \subset \mathbb{R}^n$ offen, sei $f : \Omega \rightarrow \mathbb{R}^n$ stetig differenzierbar und sei $x_0 \in \Omega$ ein Punkt mit $f(x_0) = 0$. Gilt $\operatorname{Re}(\lambda) < 0$ für alle Eigenwerte λ der Matrix $\frac{\partial f}{\partial x}(x_0) \in \mathbb{R}^{n \times n}$, dann ist die Ruhelage x_0 asymptotisch stabil.*

Beweis: Wie zuvor begründet, reicht es aus, die Stabilität der Ruhelage 0 für die gestörte lineare Gleichung $y' = Ay + g(y)$ mit $g(y) = o(|y|)$ für $y \rightarrow 0$ zu diskutieren. Nach Satz 3.15 gibt es Konstanten $\alpha > 0$ und $\beta \geq 1$ mit $|\exp(tA)| \leq \beta e^{-t\alpha}$ für alle $t \geq 0$. Aus der Integralgleichung (3.2) erhält man damit für die Lösung y zum Anfangswert $y(t_0) = y_0$ die Ungleichung

$$|y(t)| \leq \beta e^{-(t-t_0)\alpha} |y_0| + \beta \int_{t_0}^t e^{-(t-s)\alpha} |g(y(s))| ds$$

für $t_0 \leq t \in I_{y_0}$. Sei nun $\varepsilon > 0$ beliebig, dann existiert ein $0 < \delta < \varepsilon$ mit $|g(y)| \leq \frac{\varepsilon}{\beta} |y|$.

Wählt man nun den Anfangswert y_0 so klein, dass $|y_0| < \frac{\delta}{\beta}$ gilt, dann folgt $|y(t)| < \delta$ für alle $t_0 \leq t \in I_{y_0}$. Tatsächlich, andernfalls würde es nämlich ein y_0 mit $|y_0| < \frac{\delta}{\beta}$ und einen ersten Zeitpunkt $t_0 \leq \tilde{t} \in I_{y_0}$ geben mit $|y(\tilde{t})| = \delta$. Für die Zeiten $t_0 \leq t \leq \tilde{t}$ gilt aber

$$e^{t\alpha} |y(t)| \leq \delta e^{t_0\alpha} + \varepsilon \int_{t_0}^t e^{s\alpha} |y(s)| ds$$

und daher nach dem Lemma 1.53 von Gronwall

$$e^{t\alpha} |y(t)| \leq \delta e^{t_0\alpha} e^{\varepsilon(t-t_0)},$$

d.h. $|y(t)| \leq \delta e^{-(t-t_0)(\alpha-\varepsilon)} < \delta$, im Widerspruch zu $|y(\tilde{t})| = \delta$.

Somit entfernen sich Lösungen zu Anfangswerten $|y_0| < \frac{\delta}{\beta}$ nicht weiter als um δ von der Ruhelage 0, d.h. die Ruhelage 0 ist stabil. Darüberhinaus gilt die Ungleichung $|y(t)| \leq \delta e^{-(t-t_0)(\alpha-\varepsilon)}$, und aus dieser folgt, dass die Ruhelage 0 sogar asymptotisch stabil ist. \square

Ähnlich kann man das folgende Prinzip der linearisierten Instabilität herleiten.

Satz 3.21. Sei $\Omega \subset \mathbb{R}^n$ offen, sei $f: \Omega \rightarrow \mathbb{R}^n$ stetig differenzierbar und sei $x_0 \in \Omega$ ein Punkt mit $f(x_0) = 0$. Gilt $\operatorname{Re}(\lambda) > 0$ für einen Eigenwert λ der Matrix $\frac{\partial f}{\partial x}(x_0) \in \mathbb{R}^{n \times n}$, dann ist die Ruhelage x_0 instabil.

Beweis: Siehe [Amann, (15.5)]. \square

Beispiel 3.22. Ein allgemeines Räuber-Beute-Modell mit beschränktem Wachstum hat die Form

$$\begin{aligned} x' &= (\alpha - \beta y - \lambda x)x \\ y' &= (\delta x - \gamma - \mu y)y \end{aligned} \quad (3.3)$$

mit positiven Parametern $\alpha, \beta, \gamma, \delta, \lambda, \mu > 0$, wobei x die Population der Beute und y die Population der Räuber modelliert. Da Populationen nicht negativ werden können, ist man vorwiegend am Fluss im positiven Quadranten $\Omega := (0, \infty)^2$ oder auf dessen Rand $\partial\Omega$ interessiert.

Das System (3.3) hat die Ruhelagen $(0, 0)$, $(0, -\frac{\gamma}{\mu})$, $(\frac{\alpha}{\lambda}, 0)$ und $(\xi, \eta) := (\frac{\alpha\mu + \beta\gamma}{\lambda\mu + \beta\delta}, \frac{\alpha\delta - \lambda\gamma}{\lambda\mu + \beta\delta})$ wobei (ξ, η) der Schnittpunkt der Geraden $\alpha - \beta y - \lambda x = 0$ und $\delta x - \gamma - \mu y = 0$ ist. Bezeichnet $f(x, y)$ die rechte Seite von (3.3), so gilt für die Ableitung

$$Df(x, y) = \begin{pmatrix} \alpha - \beta y - 2\lambda x & -\beta x \\ \delta y & \delta x - \gamma - 2\mu y \end{pmatrix}.$$

Daraus ergibt sich für die ersten drei Ruhelagen

$$Df(0, 0) = \begin{pmatrix} \alpha & 0 \\ 0 & -\gamma \end{pmatrix}, \quad Df(0, -\frac{\gamma}{\mu}) = \begin{pmatrix} \alpha + \beta\frac{\gamma}{\mu} & 0 \\ * & \gamma \end{pmatrix} \quad \text{und} \quad Df(\frac{\alpha}{\lambda}, 0) = \begin{pmatrix} -\alpha & * \\ 0 & \delta\frac{\alpha}{\lambda} - \gamma \end{pmatrix}.$$

Da die Eigenwerte auf der Diagonalen stehen, sind $(0, 0)$ und $(0, -\frac{\gamma}{\mu})$ instabile Ruhelagen, während $(\frac{\alpha}{\lambda}, 0)$ asymptotisch stabil bei $\frac{\alpha}{\lambda} < \frac{\gamma}{\delta}$ und instabil bei $\frac{\alpha}{\lambda} > \frac{\gamma}{\delta}$ ist. Im Grenzfall $\frac{\alpha}{\lambda} = \frac{\gamma}{\delta}$ kann man mittels des Prinzips der linearisierten Stabilität/Instabilität keine Aussage treffen. Für die letzte Ruhelage (ξ, η) ergibt sich wegen $\alpha - \beta\eta - \lambda\xi = 0$ und $\delta\xi - \gamma - \mu\eta = 0$ als Ableitung

$$Df(\xi, \eta) = \begin{pmatrix} -\lambda\xi & -\beta\xi \\ \delta\eta & -\mu\eta \end{pmatrix}$$

mit den Eigenwerten $-\frac{\lambda\xi + \mu\eta}{2} \pm \frac{1}{2}\sqrt{(\lambda\xi + \mu\eta)^2 - 4\xi\eta(\lambda\mu + \beta\delta)}$. Für die Anwendung ist nur der Fall $\xi, \eta > 0$ interessant, d.h. $\frac{\alpha}{\lambda} > \frac{\gamma}{\delta}$, und in diesem Fall folgt aus

$$(\lambda\xi + \mu\eta)^2 - 4\xi\eta(\lambda\mu + \beta\delta) < (\lambda\xi + \mu\eta)^2 - 4\xi\eta\lambda\mu = (\lambda\xi - \mu\eta)^2 < (\lambda\xi + \mu\eta)^2$$

sofort, dass die Eigenwerte einen Realteil kleiner als Null haben, d.h. die Ruhelage (ξ, η) ist asymptotisch stabil, falls sie im positiven Quadranten liegt.

Für spezielle Parameter jedoch kann man mittels des Prinzips der linearisierten Stabilität/Instabilität häufig gerade nichts aussagen. Ist beispielsweise $\lambda = 0 = \mu$, so geht (3.3) in das Lotka-Volterra-System

$$\begin{aligned} x' &= (\alpha - \beta y)x \\ y' &= (\delta x - \gamma)y \end{aligned}$$

über, das die Ruhelagen $(0, 0)$ und $\left(\frac{\gamma}{\delta}, \frac{\alpha}{\beta}\right)$ besitzt. Die Ableitung in der zuletzt genannten Ruhelage ist

$$Df\left(\frac{\gamma}{\delta}, \frac{\alpha}{\beta}\right) = \begin{pmatrix} 0 & -\beta\frac{\gamma}{\delta} \\ \delta\frac{\alpha}{\beta} & 0 \end{pmatrix},$$

so dass die Eigenwerte rein imaginär sind, weswegen sich mittels des Prinzips der linearisierten Stabilität/Instabilität nichts aussagen lässt.

3.3.2 Der Satz von Hartman-Grobman

Sei f stetig differenzierbar, dann ist in der Nähe jedes Punktes x_0 mit $f(x_0) \neq 0$ der lokale Fluss zu $x' = f(x)$ zum vom konstanten Vektorfeld $x' = (1, 0, \dots, 0)$ nahe 0 erzeugten Fluss C^1 -äquivalent.

Aufgabe 3.23. *Beweisen Sie die vorige Behauptung.*

Ist jedoch x_0 eine Ruhelage, dann ist eine lokale Klassifikation wesentlich schwieriger. Der folgende Satz von Hartman-Grobman garantiert für hyperbolische Ruhelagen x_0 von $x' = f(x)$, dass der lokale Fluss nahe x_0 zum linearen Fluss $\exp(tA)$ nahe 0 mit der Matrix $A := \frac{\partial f}{\partial x}(x_0)$ topologisch äquivalent ist. Dabei heißt eine Ruhelage x_0 von $x' = f(x)$ wie zuvor hyperbolisch, falls die Matrix $A := \frac{\partial f}{\partial x}(x_0)$ keine Eigenwerte λ mit $\operatorname{Re}(\lambda) = 0$ besitzt.

Satz 3.24. *Sei $\Omega \subset \mathbb{R}^n$ offen, sei $f: \Omega \rightarrow \mathbb{R}^n$ stetig differenzierbar und sei x_0 eine hyperbolische Ruhelage von $x' = f(x)$. Dann ist der von f generierte lokale Fluss nahe x_0 topologisch äquivalent zum linearen Fluss $\exp(tA)$ nahe 0, wobei $A := \frac{\partial f}{\partial x}(x_0)$.*

Beweis: Für einen Beweis, der den Hartmanschen Linearisierungssatz für Abbildungen benutzt, siehe [Amann, (19.9)]. Einen Beweis, der direkt mit der Differentialgleichung arbeitet, findet man in [Chicone, 4.3.3] oder in [Pöschel, 3-D]. \square

Korollar 3.25. *In einem n -dimensionalen Raum gibt es genau $n + 1$ topologisch verschiedene Typen von hyperbolischen Ruhelagen einer autonomen Differentialgleichung $x' = f(x)$ mit stetig differenzierbarer rechter Seite f .*

Beweis: Nach dem Satz 3.24 von Hartman-Grobman ist der von f generierte lokale Fluss nahe einer hyperbolischen Ruhelagen zum Fluss seiner Linearisierung A nahe 0 topologisch äquivalent, deren topologische Äquivalenzklasse ist aber nach Satz 3.18 durch die Zahl $\dim(X_s^A) \in \{0, 1, \dots, n\}$ festgelegt. \square

Tatsächlich ist der Homöomorphismus h , der die topologische Äquivalenz im Satz 3.24 von Hartman-Grobman konstituiert, i.a. nicht Lipschitz-stetig oder gar ein C^k -Diffeomorphismus für $k \geq 1$. Unter bestimmten Umständen ist der von f generierte lokale Fluss nahe einer hyperbolischen Ruhelage aber doch C^k -äquivalent zum von seiner Linearisierung in x_0 generierten Fluss nahe 0. Kriterien dafür findet man in [Pöschel, 3-E]. Insbesondere kann man in diesem Fall mittels der Linearisierung den Typ der Ruhelage genauer bestimmen (im \mathbb{R}^2 : Strudel, Knoten, o.ä.).

3.4 Lyapunov-Funktionen

Lyapunov-Funktionen dienen häufig dazu, die Stabilität einer Ruhelage in den Fällen zu überprüfen, in denen man das Prinzip der linearisierten Stabilität/Instabilität nicht anwenden kann.

Tatsächlich ist die Existenz einer Lyapunov-Funktion in vielen Fällen äquivalent zur Stabilität der Ruhelage oder allgemeiner zur Stabilität einer Teilmenge des Phasenraumes. Darüberhinaus kann man Lyapunov-Funktionen auch benutzen, um die Invarianz gewisser Teilmengen des Phasenraumes zu zeigen oder ω -Limesmengen zu diskutieren. Bevor wir den Begriff der Lyapunov-Funktion definieren, wollen wir aber noch kurz positiv invariante Mengen und den Begriff der exponentiellen Stabilität einführen.

Definition 3.26. *Ein Menge $A \subset \Omega$ heißt positiv invariant unter dem von $x' = f(x)$ generierten lokalen Fluss Φ , falls $x \in A$ schon $\Phi(t, x) \in A$ für alle $t \geq 0$ impliziert.*

Definition 3.27. *Ein Menge $A \subset \Omega$ heißt exponentiell stabil unter dem von $x' = f(x)$ generierten lokalen Fluss Φ , wenn es eine Umgebung U von A und Konstanten $\alpha > 0$, $\beta \geq 1$, gibt, für die jede Lösung zum Anfangswert $x \in U$ auf $[0, +\infty)$ definiert ist und $\text{dist}(\Phi(t, x), A) \leq \beta e^{-\alpha t} \text{dist}(x, A)$ für alle $t \geq 0$ erfüllt.*

Lemma 3.28. *Ist $A \subset \Omega$ kompakt und exponentiell stabil, dann ist A auch asymptotisch stabil.*

Beweis: A ist ein Attraktor, denn für alle $x \in U$ gilt

$$\lim_{t \rightarrow +\infty} \text{dist}(\Phi(t, x), A) \leq \lim_{t \rightarrow +\infty} \beta e^{-\alpha t} \text{dist}(x, A) = 0.$$

Außerdem ist A stabil, denn sei V eine Umgebung von A , dann gibt es zur Umgebung $V \cap U$ von A ein $r > 0$ mit $\text{dist}(x, A) < r \Rightarrow x \in V \cap U$ (hier geht die Kompaktheit von A entscheidend ein), und wegen $\text{dist}(\Phi(t, x), A) \leq \beta e^{-\alpha t} \text{dist}(x, A)$ gibt es ein $T > 0$ mit $\text{dist}(x, A) < r \Rightarrow \text{dist}(\Phi(t, x), A) < r$ für $t \geq T$. Liegt also x in der Umgebung $B_r(A) := \{x \in \mathbb{R}^n \mid \text{dist}(x, A) < r\}$ von A , dann gilt $\Phi(t, x) \in V$ für $t \geq T$, und dies reicht für die Stabilität von A aus. \square

Nun wollen wir formulieren, welche Bedingungen wir von einer Lyapunov-Funktion verlangen. Wir behandeln hier nur stetig differenzierbare Lyapunov-Funktionen, man kann aber auch schwächere Differenzierbarkeitsbedingungen verlangen, siehe [Amann, Abschnitt 18].

Definition 3.29. *Eine stetig differenzierbare Funktion $V : \Omega \rightarrow \mathbb{R}$ heißt Lyapunov-Funktion für den von $x' = f(x)$ generierten lokalen Fluß Φ auf Ω , falls*

$$\dot{V}(x) := \lim_{t \rightarrow 0} \frac{V(\Phi(t, x)) - V(x)}{t} = \langle (\text{grad } V)(x), f(x) \rangle \leq 0$$

für alle $x \in \Omega$ gilt, d.h. V ist entlang von Lösungen der Differentialgleichung monoton fallend.

Die letzte Bemerkung der Definition folgt dabei aus der Gültigkeit von

$$\frac{d}{dt} V(x(t)) = \langle (\text{grad } V)(x(t)), x'(t) \rangle = \langle (\text{grad } V)(x(t)), f(x(t)) \rangle \leq 0$$

für eine Lösung x von $x' = f(x)$. Mittels des folgenden Satzes kann man nun die Stabilität von Teilmengen des Phasenraumes mit Hilfe einer Lyapunov-Funktion nachweisen.

Satz 3.30. *Sei $\Omega \subset \mathbb{R}^n$ offen, sei $f : \Omega \rightarrow \mathbb{R}^n$ lokal Lipschitz-stetig, und sei $V : \Omega \rightarrow \mathbb{R}$ eine Lyapunov-Funktion für $x' = f(x)$.*

Dann gilt: Ist $A := V^{-1}((-\infty, 0]) \subset \Omega$ kompakt und nicht leer, so ist A positiv invariant und stabil.

Gilt darüberhinaus $\dot{V}(x) < 0$ für alle $x \in \Omega \setminus A$, dann ist A asymptotisch stabil.

Gilt sogar $\dot{V}(x) \leq -p\alpha V(x)$ sowie $c \text{dist}(x, A)^p \leq V(x) \leq \tilde{c} \text{dist}(x, A)^p$ für alle $x \in \Omega$ mit Konstanten $\alpha, \tilde{c}, c, p > 0$, dann ist A exponentiell stabil.

Beweis: Die positive Invarianz der Teilmenge A folgt direkt aus

$$x_0 \in A \Rightarrow \forall 0 \leq t \in I_{x_0} : V(x(t)) \leq V(x_0) \leq 0 \Rightarrow \forall 0 \leq t \in I_{x_0} : x(t) \in A$$

für eine Lösung x von $x' = f(x)$ mit Anfangswert x_0 , sowie aus der Tatsache, dass sich aufgrund der Kompaktheit von $A \subset \Omega$ eine Lösung in A nicht dem Rand von Ω nähern kann, also für alle Zeiten existiert.

Bezeichne $B_r(A) := \{x \in \mathbb{R}^n \mid \text{dist}(x, A) < r\}$ für $r > 0$ die Menge aller Punkte, deren Abstand zu A kleiner als r ist. Um die Stabilität von A nachzuweisen, reicht es, zu jedem genügend kleinen $\varepsilon > 0$ ein $\delta > 0$ zu finden, für das $x \in B_\delta(A)$ schon $\Phi(t, x) \in B_\varepsilon(A)$ impliziert. Zunächst bemerke man, dass es aufgrund der Kompaktheit von $A := V^{-1}((-\infty, 0])$ ein $r > 0$ mit $\overline{B_r(A)} \subset \Omega$ gibt. Sei nun $0 < \varepsilon < r$ vorgegeben und $\gamma > 0$ das Minimum der stetigen Funktion V auf dem Rand $\partial B_\varepsilon(A)$. Zu diesem $\gamma > 0$ gibt es ein $0 < \delta < \varepsilon$ mit $V(x) < \gamma$ für $x \in B_\delta(A)$. Nun ist $V(x(t))$ entlang jeder Lösung von $x' = f(x)$ zu einem Anfangswert $x(0) \in B_\delta(A)$ monoton fallend. Insbesondere gilt $V(x(t)) \leq V(x(0)) < \gamma$ für alle $t \geq 0$, also erreicht $V(x(t))$ niemals den Wert γ , und somit erreicht $x(t)$ niemals den Rand $\partial B_\varepsilon(A)$. Dies bedeutet aber, dass $x(0) \in B_\delta(A)$ schon $x(t) \in B_\varepsilon(A)$ für alle $t \geq 0$ impliziert, was zu zeigen war.

Darüberhinaus existiert sogar der Grenzwert $V_\infty := \lim_{t \rightarrow \infty} V(x(t))$ für jede Lösung x von $x' = f(x)$ mit $V(x(0)) \leq \gamma$ für ein $\gamma > 0$, für das $V^{-1}((-\infty, \gamma]) \subset \Omega$ kompakt ist, denn V ist entlang von Lösungen monoton fallend und auf der kompakten Menge $V^{-1}((-\infty, \gamma])$ nach unten beschränkt. Gilt nun zusätzlich $\dot{V}(x) < 0$ für alle $x \in \overline{B_r(A)} \setminus A$, dann folgt sogar $V_\infty \leq 0$. Denn andernfalls würde \dot{V} auf der kompakten Teilmenge $\{x \in \overline{B_r(A)} \mid V_\infty \leq V(x) \leq \gamma\}$ von $\Omega \setminus A$ ein Maximum $-\alpha < 0$ besitzen, und mit diesem α würde $\frac{d}{dt}V(x(t)) \leq -\alpha$ gelten, im Widerspruch dazu, dass $V(x(t))$ nicht unterhalb V_∞ sinkt. Da außerdem für jede Umgebung U von A ein $\gamma > 0$ mit $V(x) < \gamma \Rightarrow x \in U$ existiert, dieses γ wegen $\lim_{t \rightarrow \infty} V(x(t)) = V_\infty \leq 0$ aber im Laufe der Zeit unterschritten wird, landet jede Lösung x zu einem Anfangswert in $\overline{B_r(A)}$ nach einiger Zeit in U , d.h. A ist attraktiv und somit auch asymptotisch stabil.

Gilt sogar $\dot{V}(x) \leq -p\alpha V(x)$ sowie $c \text{dist}(x, A)^p \leq V(x) \leq \tilde{c} \text{dist}(x, A)^p$ für alle $x \in \Omega$, dann folgt aus $\frac{d}{dt}V(x(t)) = \dot{V}(x(t)) \leq -p\alpha V(x(t))$ schon

$$c \text{dist}(x(t), A)^p \leq V(x(t)) \leq V(x(0))e^{-p\alpha t} \leq \tilde{c}e^{-p\alpha t} \text{dist}(x(0), A)^p,$$

also $\text{dist}(x(t), A) \leq \beta e^{-\alpha t} \text{dist}(x(0), A)$ für alle Startwerte in U und $t \geq 0$, wobei $\beta := (\frac{\tilde{c}}{c})^{\frac{1}{p}}$. Daher ist A sogar exponentiell stabil. \square

Hauptsächlich wendet man Satz 3.30 für den Fall an, dass $A = \{x^*\}$ eine Ruhelage ist, d.h. die Lyapunov-Funktion V in x^* ein striktes Minimum annimmt. Man bemerke, dass in der Formulierung von Satz 3.30 dieses Minimum Null sein muss, aber man kann in Satz 3.30 natürlich A durch $A := V^{-1}((-\infty, c])$ mit jeder anderen Zahl $c \in \mathbb{R}$ ersetzen. Außerdem reicht es aus, auf einer offenen Teilmenge $\Omega' \subset \Omega$ statt auf dem gesamten Phasenraum zu arbeiten. Diesen Spezialfall formulieren wir im folgenden Korollar.

Korollar 3.31. *Sei $\Omega \subset \mathbb{R}^n$ offen und sei $f : \Omega \rightarrow \mathbb{R}^n$ lokal Lipschitz-stetig.*

Ist $V : \Omega' \rightarrow \mathbb{R}$ eine Lyapunov-Funktion für $x' = f(x)$ auf einer Umgebung $\Omega' \subset \Omega$ des Punktes $x^ \in \Omega$ und hat V in x^* ein striktes lokales Minimum, dann ist x^* eine stabile Ruhelage von $x' = f(x)$.*

Gilt darüberhinaus $\dot{V}(x) < 0$ für alle $x_0 \neq x \in \Omega'$, dann ist x^ asymptotisch stabil, und gilt sogar $\dot{V}(x) \leq -p\alpha V(x)$ sowie $c|x - x^*|^p \leq V(x) \leq \tilde{c}|x - x^*|^p$ für alle $x \in \Omega'$ mit Konstanten $\alpha, \tilde{c}, c, p > 0$, dann ist x^* exponentiell stabil.*

Beispiel 3.32. In Beispiel 3.22 hatten wir gesehen, dass man mittels des Prinzips der linearisierten Stabilität keine Aussage über die Stabilität der Ruhelage $\left(\frac{\gamma}{\delta}, \frac{\alpha}{\beta}\right)$ des Lotka-Volterra-Systems

$$\begin{aligned}x' &= (\alpha - \beta y)x \\y' &= (\delta x - \gamma)y\end{aligned}$$

mit Parametern $\alpha, \beta, \gamma, \delta > 0$ treffen kann. Wir suchen nun mit Hilfe des Ansatzes $V(x, y) := F(x) + G(y)$ nach einer Lyapunov-Funktion. Mit $f := F'$ und $g := G'$ hat man wegen $(\text{grad } V)(x, y) = (f(x), g(y))$ die von einer Lyapunov-Funktion verlangte Bedingung

$$\begin{aligned}\dot{V}(x, y) &= \langle (\text{grad } V)(x, y), \begin{pmatrix} (\alpha - \beta y)x \\ (\delta x - \gamma)y \end{pmatrix} \rangle \\ &= (\alpha - \beta y)xf(x) + (\delta x - \gamma)yg(y) \leq 0\end{aligned}$$

zu erfüllen. Würde man $f(x) = \delta$ und $g(y) = \beta$ wählen, so würden schon einmal die gemischten Terme wegfallen, allerdings würde dann nicht $V(x, y) > V\left(\frac{\gamma}{\delta}, \frac{\alpha}{\beta}\right)$ für $(x, y) \neq \left(\frac{\gamma}{\delta}, \frac{\alpha}{\beta}\right)$ gelten. Daher wählen wir $f(x) = \delta - \gamma\frac{1}{x}$ und $g(y) = \beta - \alpha\frac{1}{y}$, dann gilt

$$(\alpha - \beta y)xf(x) + (\delta x - \gamma)yg(y) = \alpha(\delta x - \gamma) + \beta\gamma y - \alpha\delta x - \gamma(\beta y - \alpha) = 0,$$

d.h. V ist sogar konstant entlang von Lösungen und nicht nur monoton fallend. Im positiven Quadranten sind $F(x) = \delta x - \gamma \log(x)$ bzw. $G(y) = \beta y - \alpha \log(y)$ Stammfunktionen zu f bzw. g , und diese nehmen in $x = \frac{\gamma}{\delta}$ bzw. $y = \frac{\alpha}{\beta}$ ihr striktes Minimum an. Also ist in der Tat

$$V(x, y) = \delta x - \gamma \log(x) + \beta y - \alpha \log(y)$$

im positiven Quadranten eine Lyapunov-Funktion mit $V(x, y) > V\left(\frac{\gamma}{\delta}, \frac{\alpha}{\beta}\right)$ für $(x, y) \neq \left(\frac{\gamma}{\delta}, \frac{\alpha}{\beta}\right)$, so dass $\left(\frac{\gamma}{\delta}, \frac{\alpha}{\beta}\right)$ eine stabile, jedoch nicht asymptotisch stabile Ruhelage ist. Man erfährt hier sogar mehr, nämlich dass die Höhenlinien von V gerade die Orbits des dynamischen Systems im positiven Quadranten sind, und diese sind alle periodisch.

Dem Beweis von Satz 3.30 kann man darüberhinaus sogar entnehmen, dass im Fall der asymptotischen oder exponentiellen Stabilität der Einzugsbereich von A die Menge aller Punkte $x \in \Omega$ enthält, für die $V(x) \leq \gamma$ mit einem $\gamma > 0$ gilt, für das $V^{-1}((-\infty, \gamma]) \subset \Omega$ kompakt ist. Dabei ist der Einzugsbereich eines Attraktors A folgendermaßen definiert.

Definition 3.33. Ist $A \subset \Omega$ unter dem lokalen Fluss Φ auf Ω attraktiv, so heißt die Menge $\{x \in \Omega \mid \lim_{t \rightarrow \infty} \text{dist}(\Phi(t, x), A) = 0\}$ der Einzugsbereich von A .

Die Bestimmung des Einzugsbereichs eines Attraktors spielt in vielen Anwendungen eine wichtige Rolle. Das folgende Beispiel zeigt, dass die mittels einer Lyapunov-Funktion gefundene Teilmenge des Einzugsbereichs üblicherweise echt kleiner ist als der Einzugsbereich.

Beispiel 3.34. Für das ebene System

$$\begin{aligned}x' &= -x + x^2y \\y' &= -y\end{aligned}\tag{3.4}$$

sieht man bereits mit dem Prinzip der linearisierten Stabilität, dass die einzige Ruhelage $(0,0)$ asymptotisch stabil ist. Wir wollen aber genauer den Einzugsbereich von $(0,0)$ wissen und suchen dazu mit Hilfe des Ansatzes $V(x,y) := ax^2 + by^2$ nach einer Lyapunov-Funktion. Aus

$$\dot{V}(x,y) = -2ax^2(1-xy) - 2by^2$$

ergibt sich, dass V für $a, b > 0$ im von der Hyperbel $xy = 1$ berandeten Gebiet $\Omega' := \{(x,y) \mid xy < 1\}$ eine strikte Lyapunov-Funktion ist. Das Einzugsgebiet von $(0,0)$ enthält daher alle Punkte, die in einer kompakt in Ω' enthaltenen Ellipse $V^{-1}((-\infty, \gamma])$ liegen. Dies sind alle Punkte in $\{(x,y) \mid |xy| < 1\}$, daher enthält der Einzugsbereich von $(0,0)$ jeden Punkt (x,y) mit $|xy| < 1$.

Das System (3.4) ist glücklicherweise sogar so einfach, dass wir es explizit lösen: Die zweite Gleichung liefert $y(t) = y_0 e^{-t}$ und die erste Gleichung wird damit zu einer Bernoulli-Gleichung $x' + x - y_0 e^{-t} x^2 = 0$ mit Lösung $x(t) = \frac{2x_0}{(2-x_0 y_0) e^t + 2x_0 y_0 e^{-t}}$. Daher konvergiert $(x(t), y(t))$ bei $t \rightarrow \infty$ für Anfangswerte mit $|x_0 y_0| < 2$ gegen die Ruhelage $(0,0)$, während die Lösungen $(x(t), y(t)) = (\frac{x_0}{2} e^t, y_0 e^{-t})$ zu Anfangswerten mit $|x_0 y_0| = 2$ nicht gegen $(0,0)$ konvergieren und die Lösungen zu Anfangswerten mit $|x_0 y_0| > 2$ gar nur ein endliches positives Existenzintervall besitzen. Die von uns mit Hilfe von ellipsoiden Lyapunov-Funktionen gefundene Teilmenge $\{(x,y) \mid |xy| < 1\}$ ist also eine echte Teilmenge des Einzugsbereichs $\{(x,y) \mid |xy| < 2\}$ der Ruhelage $(0,0)$.

Analog zum Stabilitätssatz 3.30 kann man mit Hilfe von Funktionen, die $\dot{V} > 0$ erfüllen, auch einen Instabilitätssatz beweisen, siehe [Walter, 30.III/IX]. Darüberhinaus kann man auch Umkehrungen des Stabilitätssatzes 3.30 gewinnen, die aus der (asymptotischen bzw. exponentiellen) Stabilität die Existenz einer zugehörigen Lyapunov-Funktion folgern und somit beweisen, dass die Stabilität einer Teilmenge äquivalent zur Existenz einer Lyapunov-Funktion für diese Teilmenge ist, siehe die Referenzen in [Amann, (18.9)]. Statt mit solch allgemeinen Resultaten wollen wir uns zum Abschluss unserer Diskussion von Lyapunov-Funktionen mit zwei speziellen Typen von Systemen nichtlinearer Differentialgleichungen beschäftigen, nämlich mit Gradientensystemen und Hamiltonschen Systemen.

3.4.1 Gradientensysteme

Negative Gradientenvektorfelder zu konvexen Funktionen haben wir schon in Abschnitt 1.5.3 kennengelernt. Hier wollen wir uns mit allgemeinen Gradientensystemen beschäftigen, die die Form

$$x' = -(\text{grad } V)(x) \tag{3.5}$$

mit dem Gradientenvektorfeld $\text{grad } V : \Omega \rightarrow \mathbb{R}^n$ einer (nicht notwendig konvexen) C^2 -Funktion $V : \Omega \rightarrow \mathbb{R}$ auf einer offenen Menge $\Omega \subset \mathbb{R}^n$ besitzen. Wie wir schon in Satz 1.24 mit Verweis auf [Königsberger, 11.3,11.4] bemerkt haben, ist ein System $x' = f(x)$ für ein stetig differenzierbares Vektorfeld f auf einer einfach zusammenhängenden Menge $\Omega \subset \mathbb{R}^n$ genau dann ein Gradientensystem, wenn $\frac{\partial f_i}{\partial x_j} = \frac{\partial f_j}{\partial x_i}$ für alle $i, j = 1, \dots, n$ gilt.

Geometrisch bewegt sich jede Lösung eines Gradientensystems (3.5) im Punkt $x(t)$ in Richtung des steilsten Abstiegs $-(\text{grad } V)(x(t))$ der Funktion V in $x(t)$. Deswegen schneidet der Orbit durch einen regulären Punkt x von V , d.h. einen Punkt x mit $(\text{grad } V)(x) \neq 0$, die Niveaumenge $V^{-1}(\{x\})$ senkrecht, und man wird erwarten, dass jede Lösung von (3.5) für $t \rightarrow \infty$ gegen ein lokales Minimum oder zumindest gegen einen kritischen Punkt von V strebt. Bevor wir dies aber beweisen, wollen wir den Begriff der ω -Limesmenge eines Orbits einführen.

Definition 3.35. Für einen lokalen Fluss Φ auf Ω ist die ω -Limesmenge $\omega(x)$ des Orbits durch $x \in \Omega$ die Menge aller Punkte $y \in \Omega$, für die es eine Folge $t_k \rightarrow \infty$ mit $\Phi(t_k, x) \rightarrow y$ gibt.

Mit anderen Worten ist $\omega(x)$ die Menge aller Häufungspunkte von $\Phi(t, x)$ für $t \rightarrow \infty$, und man kann $\omega(x)$ alternativ durch

$$\omega(x) := \bigcap_{t>0} \overline{\Phi([t, \infty), x)} \quad (3.6)$$

definieren. Entsprechend kann man die α -Limesmenge $\alpha(x)$ des Orbits durch $x \in \Omega$ als die Menge aller Häufungspunkte von $\Phi(t, x)$ für $t \rightarrow -\infty$ definieren. Wir konzentrieren uns im folgenden Satz auf ω -Limesmengen, für α -Limesmengen existieren aber ganz analoge Resultate.

Satz 3.36. Sei $\Omega \subset \mathbb{R}^n$ offen und Φ ein lokaler Fluss auf Ω . Ist der positive Halborbit durch x relativ kompakt in Ω , so ist die ω -Limesmenge $\omega(x)$ nicht leer, kompakt, zusammenhängend und (positiv) invariant.

Ist darüberhinaus V eine Lyapunov-Funktion auf einer Umgebung Ω' des positiven Halborbits durch x , dann gibt es einen Wert γ von V mit $\omega(x) \subset V^{-1}(\{\gamma\})$, und insbesondere gilt $\omega(x) \subset \{x \in \Omega' \mid \dot{V}(x) = 0\}$.

Beweis: Aufgrund der relativen Kompaktheit des positiven Halborbits durch x existiert die Lösung durch x für alle Zeiten $t \geq 0$. Nach (3.6) ist $\omega(x)$ der Durchschnitt der ineinander geschachtelten nicht leeren und kompakten Mengen $\Phi([t, \infty), x)$, also ist $\omega(x)$ selbst nicht leer und kompakt.

Die positive Invarianz der Halborbite $\Phi([t, \infty), x)$ impliziert die positive Invarianz des Abschlusses $\overline{\Phi([t, \infty), x)}$, und $\omega(x)$ ist als Durchschnitt dieser positiv invarianten Mengen selbst positiv invariant. Man kann sogar die Invarianz von $\omega(x)$ zeigen, darauf verzichten wir aber hier.

Schließlich ist $\omega(x)$ zusammenhängend, denn wäre $\omega(x) = \omega_1 \cup \omega_2$ die Vereinigung zweier nicht leerer, abgeschlossener und disjunkter Mengen, so wären mit $\omega(x)$ auch ω_1, ω_2 kompakt. Daher gäbe es disjunkte offene Umgebungen U_i von ω_i , $i = 1, 2$, und weil $\omega(x)$ die Menge der Häufungspunkte des positiven Halborbits durch x ist, würde $\Phi([t, \infty)) \subset U_1 \cup U_2$ und $\Phi([t, \infty)) \cap U_i \neq \emptyset$ für $i = 1, 2$ mit einem genügend großen $t > 0$ gelten, im Widerspruch dazu, dass $\Phi([t, \infty))$ zusammenhängend ist.

Ist zu guter Letzt V eine Lyapunov-Funktion auf einer Umgebung Ω' von $\Phi([0, \infty), x)$, dann ist $t \mapsto V(\Phi(t, x))$ monoton fallend und konvergiert für $t \rightarrow \infty$ gegen $V_\infty := \inf\{V(\Phi(t, x)) \mid t \geq 0\}$. Aufgrund der Stetigkeit von V und der Definition von $\omega(x)$ folgt $V_\infty = V(y)$ für alle $y \in \omega(x)$, und da $\omega(x) \neq \emptyset$ gilt, ist V_∞ insbesondere endlich. Mit $\gamma := V_\infty$ folgt dann die Behauptung. \square

Nach diesen allgemeinen Aussagen über ω -Limesmengen wollen wir uns nun speziell Gradientensystemen (3.5) zuwenden. Für ein Gradientensystem $x' = -(\text{grad } V)(x)$ ist offensichtlich V selbst eine Lyapunov-Funktion, denn $\dot{V}(x) = -|(\text{grad } V)(x)|^2 \leq 0$ gilt für alle x . Desweiteren gilt $\dot{V}(x) = 0$ nur in den Punkten x mit $(\text{grad } V)(x) = 0$, d.h. nur in den kritischen Punkten von V . Aus der letzten Aussage $\omega(x) \subset \{x \in \Omega' \mid \dot{V}(x) = 0\}$ von Satz 3.36 ergibt sich daher die folgende Aussage.

Korollar 3.37. Sei $\Omega \subset \mathbb{R}^n$ offen und $V : \Omega \rightarrow \mathbb{R}$ zweimal stetig differenzierbar. Dann ist jeder ω -Limespunkt eines Orbits des Gradientensystems $x' = -(\text{grad } V)(x)$ ein kritischer Punkt von V .

Jedoch muss nicht jeder ω -Limespunkt eines Orbits auch ein lokales Minimum sein, z.B. könnte der Orbit aus einer stabilen Richtung in einen Sattel hineinlaufen. Liegt aber in x^* ein striktes lokales Minimum von V , dann ist x^* wegen $\dot{V}(x) = -|(\text{grad } V)(x)|^2 < 0$ für $x \neq x^*$ in einer Umgebung von x^* sogar asymptotisch stabil.

Korollar 3.38. Sei $\Omega \subset \mathbb{R}^n$ offen und $V : \Omega \rightarrow \mathbb{R}$ zweimal stetig differenzierbar. Dann sind strikte lokale Minima von V asymptotisch stabile Ruhelagen des Gradientensystems $x' = -(\text{grad } V)(x)$.

Eine weitere Konsequenz aus diesen Beobachtungen ist, dass es in Gradientensystemen keine periodischen Orbits geben kann. Solche hätten nämlich sich selbst als ω -Limesmengen, aber wie wir zuvor gesehen haben, besteht die ω -Limesmenge jedes Orbits nur aus Ruhelagen. Daher ist das Langzeitverhalten eines Gradientensystems nicht so komplex wie das Langzeitverhalten eines allgemeinen Systems. Es könnte jedoch beispielsweise passieren, dass die ω -Limesmenge eines Orbits ein Kreis von Ruhelagen ist und nicht nur eine einzelne Ruhelage.

3.4.2 Hamiltonsche Systeme

Hamiltonsche Systeme sind Systeme von $2n$ Differentialgleichungen der Form

$$\begin{aligned} x' &= \frac{\partial H}{\partial y} \\ y' &= -\frac{\partial H}{\partial x} \end{aligned} \tag{3.7}$$

mit einer C^2 -Funktion $H : \Omega \rightarrow \mathbb{R}$ auf einer offenen Teilmenge $\Omega \subset \mathbb{R}^{2n}$, wobei x und y jeweils Kurven in \mathbb{R}^n sind. In der Physik interpretiert man $x(t)$ als Konfiguration, $y(t)$ als Impuls und $H(x(t), y(t))$ als Gesamtenergie eines mechanischen Systems zum Zeitpunkt t . Die grundlegendste Eigenschaft eines Hamiltonschen Systems ist die Erhaltung der Funktion H , die die Gesamtenergie modelliert, denn tatsächlich gilt

$$\frac{d}{dt}H(x(t), y(t)) = \frac{\partial H}{\partial x}x' + \frac{\partial H}{\partial y}y' = \frac{\partial H}{\partial x} \frac{\partial H}{\partial y} - \frac{\partial H}{\partial y} \frac{\partial H}{\partial x} = 0.$$

Insbesondere ist H eine Lyapunov-Funktion mit $\dot{H}(x, y) = 0$, d.h. jeder Orbit verläuft innerhalb einer Niveaumenge von H . Für strikte lokale Minima von H ergibt sich aus unseren Erkenntnissen über Lyapunov-Funktionen das folgende Resultat.

Korollar 3.39. Sei $\Omega \subset \mathbb{R}^{2n}$ offen und $H : \Omega \rightarrow \mathbb{R}$ zweimal stetig differenzierbar. Dann sind strikte lokale Minima von H stabile Ruhelagen des Hamiltonschen Systems (3.7).

Jedoch gibt es in Hamiltonschen Systemen keine asymptotisch stabilen Ruhelagen oder periodischen Orbits. Denn wie wir im folgenden zeigen werden, ist der Fluss eines Hamiltonschen Systems volumenerhaltend. Dies aber widerspricht der Existenz asymptotisch stabiler Ruhelagen oder periodischer Orbits, denn aufgrund der asymptotischen Stabilität würde es eine (kompakte) Umgebung $U \subset \Omega$ der Ruhelage bzw. des periodischen Orbits geben, die sich unter dem Fluss für $t \rightarrow 0$ auf die 0-dimensionale Ruhelage bzw. den 1-dimensionalen periodischen Orbit zusammenzieht, d.h. im Widerspruch zur Volumenerhaltung würde $\text{Vol}(U) > 0$ und $\lim_{t \rightarrow \infty} \text{Vol}(\Phi(t, U)) = 0$ gelten.

Um nun zu beweisen, dass der Fluss eines Hamiltonschen Systems das Volumen jeder (kompakten) Teilmenge $U \subset \Omega$ erhält, bemerke man zunächst, dass das Vektorfeld $f(x, y) := (\frac{\partial H}{\partial y}(x, y), -\frac{\partial H}{\partial x}(x, y))$ auf der rechten Seite von (3.7) divergenzfrei ist, denn es gilt

$$\text{div}(f) = \sum_{i=1}^n \frac{\partial^2 H}{\partial x_i \partial y_i} - \sum_{i=1}^n \frac{\partial^2 H}{\partial y_i \partial x_i} = 0.$$

Damit folgt die behauptete Volumenerhaltung Hamiltonscher Flüsse

Satz 3.40. *Ist $\Omega \subset \mathbb{R}^n$ offen und ist $f : \Omega \rightarrow \mathbb{R}^n$ ein stetig differenzierbares und divergenzfreies Vektorfeld, dann ist der von $x' = f(x)$ erzeugte lokale Fluss Φ volumenerhaltend, d.h. für jede kompakte Teilmenge $U \subset \Omega$ und jedes t (im Durchschnitt aller Existenzintervalle zu Anfangswerten in U) gilt $\text{Vol}(\Phi(t, U)) = \text{Vol}(U)$.*

Beweis: Nach Satz 1.61 über die differenzierbare Abhängigkeit vom Anfangswert löst $Y(t, x) := \frac{\partial \Phi}{\partial x}(t, x)$ für festes $x \in \Omega$ die homogene lineare Variationsgleichung $Y' = \frac{\partial f}{\partial x}(\Phi(t, x))Y$. Insbesondere gilt nach Satz 2.5 für die Wronski-Determinante

$$\det(Y)' = \text{Spur}\left(\frac{\partial f}{\partial x}(\Phi(t, x))\right) \det(Y) = \text{div}(f)(\Phi(t, x)) \det(Y)$$

Also ergibt sich mittels des Transformationssatzes

$$\begin{aligned} \frac{d}{dt} \text{Vol}(\Phi(t, U)) &= \frac{d}{dt} \int_{\Phi(t, U)} 1 \, d\text{Vol}(y) = \frac{d}{dt} \int_U \det(Y(t; x)) \, d\text{Vol}(x) \\ &= \int_U \text{div}(f)(\Phi(t, x)) \det(Y(t, x)) \, d\text{Vol}(x) = \int_{\Phi(t, U)} \text{div}(f)(y) \, d\mu(y). \end{aligned}$$

Für das divergenzfreie Vektorfeld f gilt also $\frac{d}{dt} \text{Vol}(\Phi(t, U)) = 0$, d.h. die Funktion $t \mapsto \text{Vol}(\Phi(t, U))$ ist konstant. \square

Kapitel 4

Rand- und Eigenwertprobleme

Bisher haben wir hauptsächlich Anfangswertprobleme betrachtet, d.h. wir haben nach Lösungen einer Differentialgleichung erster Ordnung zu einem vorgegebenen Anfangswert an einer Stelle gesucht (siehe Definition 1.7). Für eine Differentialgleichung n -ter Ordnung musste man sich dabei Anfangswerte für die ersten $(n - 1)$ Ableitungen an einer Stelle vorgeben. In diesem Abschnitt wollen wir nun speziell für Differentialgleichungen zweiter Ordnung (d.h. $n = 2$) diskutieren, was passiert, wenn man sich statt Werten für die Lösung und ihre Ableitung an einer Stelle beispielsweise den Wert der Lösung an zwei verschiedenen Stellen vorgibt. Sucht man eine solche Lösung, so spricht man von einem Randwertproblem. Das folgende Beispiel zeigt, dass selbst im einfachsten Fall einer linearen Differentialgleichung zweiter Ordnung mit konstanten Koeffizienten die Lösbarkeit von Randwertproblemen nicht allgemein garantiert werden kann.

Beispiel 4.1. *Es gibt keine Lösung der homogenen linearen Differentialgleichung mit konstanten Koeffizienten $y'' + y = 0$, die den Randbedingungen $y(0) = 0$ und $y(\pi) = 1$ genügt. Denn jede Lösung y von $y'' + y = 0$ hat die Form $y(x) = c_1 \sin(x) + c_2 \cos(x)$ mit $c_1, c_2 \in \mathbb{R}$, aus $y(0) = 0$ folgt aber $c_2 = 0$, und es gibt kein $c_1 \in \mathbb{R}$ mit $c_1 \sin(\pi) = 1$ wegen $\sin(\pi) = 0$.*

Ziel dieses Kapitels wird es daher sein, notwendige und hinreichende Kriterien zu formulieren, die für lineare Differentialgleichungen zweiter Ordnung die eindeutige Lösbarkeit von Randwertproblemen garantieren, und Methoden anzugeben, mit denen man die Lösung explizit ausrechnen kann. Danach werden wir noch kurz Eigenwertprobleme diskutieren, die mit Randwertproblemen eng verwandt sind.

4.1 Sturm-Liouvillesche Randwertprobleme

Jede lineare Differentialgleichung zweiter Ordnung

$$y'' + a_1(x)y' + a_2(x)y = b(x)$$

mit stetigen Koeffizientenfunktionen a_1, a_2 und einer stetigen Inhomogenität b auf einem Intervall $[a, b]$ kann man in die selbstadjungierte Form

$$Ly := (p(x)y')' + q(x)y = g \tag{4.1}$$

mit stetigen Funktionen p, q, g auf $[a, b]$ umschreiben, wobei p stetig differenzierbar ist und $p > 0$ auf $[a, b]$ erfüllt.

Aufgabe 4.2. Zeigen Sie, dass $y'' + a_1(x)y' + a_2(x)y = b(x)$ durch Multiplikation mit $p(x) := e^{\int a_1(x) dx} > 0$ in die Form (4.1) übergeht.

Wir wollen nun (4.1) unter Randbedingungen diskutieren. Statt nur den Wert der Lösung an den beiden Stellen a, b vorzugeben, betrachten wir dabei gleich allgemeinere Randbedingungen der Form

$$\begin{aligned} R_1 y &:= \alpha_1 y(a) + \alpha_2 p(a) y'(a) = \eta_1 \\ R_2 y &:= \beta_1 y(b) + \beta_2 p(b) y'(b) = \eta_2 \end{aligned} \quad (4.2)$$

mit $0 \neq \alpha, \beta \in \mathbb{R}^2$ und beliebigem $\eta \in \mathbb{R}^2$, die auch die Werte der Ableitung an den Stellen a, b einbeziehen können. Sucht man unter den zuvor genannten Bedingungen an die Koeffizienten nach einer Lösung y von (4.1), die den Randbedingungen (4.2) genügt, so spricht man von einem Sturm-Liouvilleschen Randwertproblem.

Bemerkung 4.3. Gelten die zuvor genannten Bedingungen an die Funktionen p, q, g nur auf dem offenen Intervall (a, b) , so spricht man bei $p(x) \rightarrow +\infty$ für $x \rightarrow a, b$ von einem singulären bzw. bei $p(x) \rightarrow 0$ für $x \rightarrow a, b$ von einem degenerierten Problem. Solche Probleme sind i.a. viel schwieriger zu behandeln als der von uns diskutierte reguläre Fall.

Außerdem sei angemerkt, dass es neben Randbedingungen der Form (4.2) auch noch andere sinnvolle Randbedingungen gibt, wie z.B. die periodischen Randbedingungen $y(a) = y(b)$, $y'(a) = y'(b)$.

Der folgende fundamentale Satz formuliert das wichtigste Kriterium für die Lösbarkeit Sturm-Liouvillescher Randwertprobleme.

Satz 4.4. Für $p \in C^1([a, b])$ mit $p > 0$ auf $[a, b]$, $q \in C([a, b])$, $0 \neq \alpha, \beta \in \mathbb{R}^2$, ist das inhomogene Randwertproblem (4.1), (4.2), genau dann für jedes $g \in C([a, b])$ und $\eta \in \mathbb{R}^2$ eindeutig lösbar, falls das homogene Randwertproblem $Ly = 0$, $Ry = 0$, nur die triviale Lösung $y = 0$ besitzt, oder äquivalenterweise $\det \begin{pmatrix} R_1 y_1 & R_1 y_2 \\ R_2 y_1 & R_2 y_2 \end{pmatrix} \neq 0$ für ein Fundamentalsystem y_1, y_2 von $Ly = 0$ gilt.

Beweis: Aufgrund der Voraussetzungen an p, q ist (4.1) zu einer linearen Differentialgleichung zweiter Ordnung in der üblichen Form äquivalent, daher besitzt $Ly = g$ für jede Wahl von $g \in C([a, b])$ einen zweidimensionalen affinen Lösungsraum, und jede Lösung y von $Ly = g$ hat die Form

$$y = c_1 y_1 + c_2 y_2 + y_p$$

mit einem Fundamentalsystem y_1, y_2 von $Ly = 0$, einer partikulären Lösung y_p von $Ly = g$ und Konstanten $c_1, c_2 \in \mathbb{R}$. Die Randbedingung $Ry = \eta$ liefert das lineare Gleichungssystem

$$\begin{pmatrix} R_1 y_1 & R_1 y_2 \\ R_2 y_1 & R_2 y_2 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} \eta_1 - R_1 y_p \\ \eta_2 - R_2 y_p \end{pmatrix}$$

für die Konstanten c_1, c_2 , und dieses lineare Gleichungssystem ist genau dann für jede Wahl von $\eta \in \mathbb{R}^2$ eindeutig lösbar, falls die Matrix auf der linken Seite eine nichtverschwindende Determinante besitzt. Dies ist aber offenbar auch äquivalent dazu, dass bei $y_p = 0$ und $\eta = 0$ nur $(c_1, c_2) = 0$ eine Lösung ist, d.h. dass das homogene Randwertproblem $Ly = 0$, $Ry = 0$, nur die triviale Lösung $y = 0$ besitzt. \square

Der Beweis von Satz 4.4 besagt insbesondere, dass man zur Lösung des inhomogenen Randwertproblems (4.1), (4.2), nur ein lineares (2×2) -System lösen muss, falls man schon ein Fundamentalsystem von $Ly = 0$ und eine partikuläre Lösung von $Ly = g$ gefunden hat.

4.1.1 Greensche Funktion

Die ganze Schwierigkeit des vollen Randwertproblems (4.1), (4.2), ist bereits im (halb) homogenen Randwertproblem

$$Ly = g, \quad Ry = 0, \quad (4.3)$$

versteckt. Tatsächlich, man findet leicht eine Funktion $y_p \in C^2([a, b])$ mit $Ry_p = \eta$, und ist y_h eine Lösung von $Ly_h = g - Ly_p$, $Ry_h = 0$, dann ist $y := y_h + y_p$ die Lösung von $Ly = g$, $Ry = \eta$.

Unter Voraussetzung der Gültigkeit des in Satz 4.4 genannten Kriteriums für die eindeutige Lösbarkeit wollen wir nun das (halb)homogene Randwertproblem (4.3) nicht nur für eine rechte Seite $g \in C([a, b])$ lösen (dies könnte man leicht wie im Beweis von Satz 4.4), sondern wir suchen eine allgemeine Formel, die für beliebige rechte Seiten $g \in C([a, b])$ die Lösung von (4.3) angibt.

Dazu bestimmen wir zwei nichttriviale Lösungen y_1, y_2 der homogenen Gleichung $Ly = 0$, wobei y_1 zusätzlich die Gleichung $R_1y_1 = 0$ zu erfüllen hat, während y_2 der Gleichung $R_2y_2 = 0$ genügen soll. Dies ist nicht schwierig, denn gilt für $(\lambda, \mu) \neq (0, 0)$ etwa $\alpha_1\lambda + \alpha_2\mu = 0$, so kann man y_1 als Lösung von $Ly = 0$ zu den Anfangswerten $y_1(a) = \lambda$, $p(a)y_1'(a) = \mu$, wählen, und ganz analog für y_2 im Randpunkt b . Die Funktionen y_1, y_2 sind dann linear unabhängig, denn ansonsten würde y_2 ein Vielfaches von y_1 sein und somit auch $R_2y_1 = 0$ gelten, d.h. y_1 wäre eine nichttriviale Lösung von $Ly = 0$, $R_1y = 0 = R_2y$, im Widerspruch zu unserer Voraussetzung. Daher bilden die Lösungen y_1, y_2 ein Fundamentalsystem.

Um die inhomogene Gleichung $Ly = g$ unter den Randbedingungen $R_1y = 0 = R_2y$ zu lösen, beobachte man zunächst, dass $Ly = g$ äquivalent ist zu $y'' + \frac{p'(x)}{p(x)}y' + \frac{q(x)}{p(x)}y = \frac{g(x)}{p(x)}$. Nach (2.4) ist eine partikuläre Lösung dieser Gleichung durch

$$y_p(x) := -y_1(x) \int_a^x \frac{y_2(\xi) g(\xi)}{W(\xi) p(\xi)} d\xi + y_2(x) \int_a^x \frac{y_1(\xi) g(\xi)}{W(\xi) p(\xi)} d\xi$$

mit der Wronskideterminanten $W := y_1y_2' - y_2y_1'$ gegeben, und nach Satz 2.5 gilt $W' = -\frac{p'}{p}W$, d.h. $(pW)' = 0$, so dass pW konstant ist. Also gilt

$$y_p(x) = -\frac{y_1(x)}{p(a)W(a)} \int_a^x y_2(\xi)g(\xi) d\xi + \frac{y_2(x)}{p(a)W(a)} \int_a^x y_1(\xi)g(\xi) d\xi,$$

und man berechnet leicht $R_1y_p = 0$, $R_2y_p = -\frac{R_2y_1}{p(a)W(a)} \int_a^b y_2(\xi)g(\xi) d\xi$. Die allgemeine Lösung von $Ly = g$ hat nun die Form $y = y_p + c_1y_1 + c_2y_2$. Damit y auch noch den Randbedingungen $R_1y = 0 = R_2y$ genügt, muss wegen $R_1y_2 \neq 0$ offensichtlich $c_2 = 0$ gelten und daher $c_1 = -\frac{1}{p(a)W(a)} \int_a^b y_2(\xi)g(\xi) d\xi$. Somit ist die Lösung y von $Ly = g$, $R_1y = 0 = R_2y$, durch

$$y(x) = \frac{y_2(x)}{p(a)W(a)} \int_a^x y_1(\xi)g(\xi) d\xi + \frac{y_1(x)}{p(a)W(a)} \int_x^b y_2(\xi)g(\xi) d\xi$$

gegeben. Definiert man abschließend die Greensche Funktion zum Operator L unter den Randbedingungen $R_1y = 0 = R_2y$ durch

$$\Gamma(x, \xi) := \begin{cases} \frac{y_2(x)y_1(\xi)}{p(a)W(a)} & a \leq \xi \leq x \leq b \\ \frac{y_1(x)y_2(\xi)}{p(a)W(a)} & a \leq x \leq \xi \leq b \end{cases},$$

dann erhält man den folgenden Satz.

Satz 4.5. *Ist das (halb)homogene Randwertproblem (4.3) eindeutig lösbar, dann ist die Lösung y von (4.3) mittels der Greenschen Funktion Γ zum Operator L unter den Randbedingungen $R_1y = 0 = R_2y$ gegeben durch*

$$y(x) = \int_a^b \Gamma(x, \xi)g(\xi) d\xi.$$

Das wichtige an diesem Satz ist, dass es gelungen ist, den Lösungsoperator von (4.3) als Integraloperator $g \mapsto \int_a^b \Gamma(\cdot, \xi)g(\xi) d\xi$ mit einer stetigen Funktion Γ darzustellen. Über solche Integraloperatoren kann man nun weitreichende funktionalanalytische Aussagen machen und insbesondere die Kompaktheit des Lösungsoperators folgern, was wiederum weitreichende Konsequenzen für die Eigenwerte des Operators $-L$ unter den Randbedingungen $R_1y = 0 = R_2y$ hat. Wir wollen dies hier aber nicht vertiefen, sondern beschäftigen uns im Folgenden mit dem Maximumprinzip.

4.1.2 Maximum- und Minimumprinzipien

Das starke Minimumprinzip besagt, dass bei $g \leq 0$ eine Lösung y von $Ly = g$ auf (a, b) unter den homogenen Randbedingungen $y(a) = 0 = y(b)$ entweder konstant gleich Null oder auf ganz (a, b) positiv ist. Bevor wir aber diese Form des Minimumprinzips beweisen, zeigen wir zunächst eine schwächere Version.

Satz 4.6. *Gelte $p > 0$ und $q \leq 0$ in (a, b) . Liegen y und py' in $C^1((a, b))$, gilt $Ly \leq 0$ und besitzt y auf (a, b) ein negatives lokales Minimum, dann ist y konstant.*

Beweis: Vorbemerkung: Bekanntermaßen gelten in einem lokalen Minimum x^* von y auf einem offenen Intervall I die Gleichungen $y'(x^*) = 0$ und $y''(x^*) \geq 0$, nicht notwendig aber $y''(x^*) > 0$ und somit auch nicht unbedingt $(py')'(x^*) > 0$. Nehmen wir aber an, dass y links von x^* nicht konstant ist, dann gibt es zumindest einen Punkt $x \in I$ mit $(py')'(x) > 0$. Denn wähle $\tilde{x} \in I$ mit $\tilde{x} < x^*$ so, dass $y(\tilde{x}) > y(x^*)$ gilt, dann gibt es zwischen \tilde{x} und x^* zumindest einen Punkt \hat{x} mit $y'(\hat{x}) < 0$. Aus $(py')'(\hat{x}) < 0 = (py')'(x^*)$ folgt dann aber, dass es zwischen \hat{x} und x^* zumindest einen Punkt $x \in I$ mit $(py')'(x) > 0$ gibt. Analog behandelt man den Fall, dass y rechts von x^* nicht konstant ist.

Nun kommen wir zum Beweis des Satzes: Sei $x^* \in (a, b)$ eine lokale Minimalstelle mit $y(x^*) < 0$. Nehmen wir an, dass y auf (a, b) nicht konstant ist, dann kann man ein offenes Intervall I mit $x^* \in I \subset (a, b)$ finden, auf dem y negativ und nicht konstant ist. Dann würde wegen $qy \geq 0$ auf I aber auch $(py')' \leq Ly \leq 0$ auf I gelten, d.h. im Widerspruch zur Vorbemerkung würde kein Punkt $x \in I$ mit $(py')'(x) > 0$ existieren. Daher war die Annahme, y sei nicht konstant auf (a, b) , falsch. \square

Analog kann man natürlich auch ein schwaches Maximumprinzip beweisen.

Korollar 4.7. *Gelte $p > 0$ und $q \leq 0$ in (a, b) . Liegen y und py' in $C^1((a, b))$, gilt $Ly \geq 0$ und besitzt y auf (a, b) ein positives lokales Maximum, dann ist y konstant.*

Nun kommen wir zum angekündigten starken Minimumprinzip.

Satz 4.8. *Gelte $p \in C([a, b])$, $p > 0$ auf $[a, b]$, und es gebe eine Konstante $K > 0$ mit $-K \leq q \leq 0$ in $[a, b]$. Liegen y und py' in $C^1([a, b])$, gilt $Ly \leq 0$ und ist $y(a), y(b) \geq 0$, dann gilt entweder $y = 0$ auf ganz $[a, b]$, oder es gilt $y > 0$ auf (a, b) und bei $y(a) = 0$ bzw. $y(b) = 0$ folgt $y'(a) > 0$ bzw. $y'(b) < 0$.*

Beweis: Nach Satz 4.6 besitzt y kein negatives lokales Minimum in (a, b) . Aber y besitzt natürlich ein Minimum auf $[a, b]$, d.h. $y \geq 0$ muss gelten.

Angenommen, es gilt weder $y = 0$ auf ganz (a, b) noch $y > 0$ auf (a, b) , dann gibt es ein Intervall $(\alpha, \beta) \subset (a, b)$ mit $y > 0$ auf (α, β) und $y(\alpha) = 0$ oder $y(\beta) = 0$. Ohne Einschränkung betrachten wir nur den ersten Fall und nehmen an, es gelte $p > \delta > 0$ sowie $q > -K$ auf $[\alpha, \beta]$. Aus $y(\alpha) = 0$ folgt $y(x) = \int_{\alpha}^x y'(\xi) d\xi$, und da wegen $y \geq 0$ auch $y'(\alpha) = 0$ gilt, folgt aus $(py) < Ky$ auf $[\alpha, \beta]$ die Ungleichung

$$py'(x) \leq \int_{\alpha}^x Ky(\xi) d\xi \leq K(x - \alpha)\|y\|_{L^{\infty}([\alpha, x])}.$$

Kombiniert mit $y(x) = \int_{\alpha}^x y'(\xi) d\xi$ erhält man wegen $p > \delta$ die Ungleichung

$$y(x) \leq \frac{K}{\delta}(x - \alpha)^2\|y\|_{L^{\infty}([\alpha, x])}.$$

Es gibt aber beliebig nahe bei α Punkte $x > \alpha$ mit $y(x) = \|y\|_{L^{\infty}([\alpha, x])}$, und dies widerspricht der vorigen Ungleichung. Daher gilt entweder $y = 0$ auf ganz (a, b) oder $y > 0$ auf (a, b) . Im Sonderfall $\alpha = a$ führt der Beweis zudem die Annahme $y(a) = 0 = y'(a)$ und $y > 0$ auf (a, b) zum Widerspruch, so dass wegen $y \geq 0$ schon $y'(a) > 0$ gelten muss. \square

4.2 Sturm-Liouvillesche Eigenwertprobleme

Wie wir in Satz 4.4 gesehen haben, ist für eine Konstante λ das inhomogene Randwertproblem

$$Ly + \lambda y = g \quad \text{unter} \quad R_1 y = \eta_1, \quad R_2 y = \eta_2,$$

genau dann für jedes $g \in C([a, b])$ und $\eta \in \mathbb{R}^2$ eindeutig lösbar, falls das homogene Randwertproblem $Ly + \lambda y = 0$ unter den Randbedingungen $R_1 y = 0, R_2 y = 0$, nur die triviale Lösung $y = 0$ besitzt. In diesem Abschnitt interessieren wir uns genau für den gegenteiligen Fall, d.h. wir suchen nach Konstanten λ , für die dieses homogene Randwertproblem nicht nur die triviale Lösung besitzt. Solch ein Problem nennt man ein Eigenwertproblem.

Definition 4.9. Bei einem Eigenwertproblem für das Negative des in (4.1) definierten Operators L sucht man Werte λ , für die $Ly + \lambda y = 0$ unter den Randbedingungen $R_1 y = 0, R_2 y = 0$, eine nichttriviale Lösung $y \neq 0$ besitzt. Ist dies der Fall, so nennt man λ einen Eigenwert und y eine Eigenfunktion von $-L$.

Man lasse sich dabei nicht durch die Vorzeichenwahl verwirren, diese macht erst nach einer intensiveren Beschäftigung mit Eigenwertproblemen Sinn. Der folgende Satz besagt, dass es unter den üblichen Voraussetzungen eine ganze Folge von reellen Eigenwerten λ_k des Operators $-L$ mit $\lim_{k \rightarrow \infty} \lambda_k = \infty$ gibt.

Satz 4.10. Für $p \in C^1([a, b])$ mit $p > 0$ auf $[a, b]$, $q \in C([a, b])$, $0 \neq \alpha, \beta \in \mathbb{R}^2$, gibt es mit den in (4.1), (4.2), definierten Operatoren L, R_1, R_2 zum Eigenwertproblem $Ly + \lambda y = 0$ unter den Randbedingungen $R_1 y = 0, R_2 y = 0$, eine (unendliche) Folge $\lambda_1 < \lambda_2 < \dots < \lambda_k < \dots$ von reellen Eigenwerten mit $\lambda_k \rightarrow \infty$ für $k \rightarrow \infty$.

Für diese Aussage kann man einen eleganten funktionalanalytischen Beweis angeben, der auf der Kompaktheit des in Satz 4.5 erwähnten Integraloperators zur Greenschen Funktion Γ beruht (siehe [Walter, Paragraph 28]), oder einen elementaren Beweis, der auf der Prüfer-Transformation basiert (siehe [Walter, Paragraph 27]). Wir verzichten hier aber auf einen Beweis und diskutieren stattdessen ein Beispiel.

Beispiel 4.11. Wir suchen die Eigenwerte λ von $-L$ für den Differentialoperator $Ly := y'' - 2y'$ unter den Randbedingungen $y'(0) = 0 = y'(\pi)$, d.h. die Werte λ , für die $Ly + \lambda y = 0$ unter den Randbedingungen $y'(0) = 0 = y'(\pi)$ eine nichttriviale Lösung besitzt. Dabei arbeiten wir direkt mit L , obwohl L nicht die Form (4.1) hat.

Die Nullstellen des charakteristischen Polynoms zu $Ly + \lambda y = y'' - 2y' + \lambda y$ sind $\mu_{1,2} = 1 \pm \sqrt{1 - \lambda}$, also hat $Ly + \lambda y = 0$ die allgemeine Lösung $y(x) = c_1 e^{\mu_1 x} + c_2 e^{\mu_2 x}$ (bzw. $y(x) = c_1 e^x + c_2 x e^x$ bei $\lambda = 1$). Unter den Randbedingungen $y'(0) = 0 = y'(\pi)$ gibt es genau dann eine nichttriviale Lösung $y'' - 2y' + \lambda y = 0$, wenn

$$\begin{pmatrix} \mu_1 & \mu_2 \\ \mu_1 e^{\mu_1 \pi} & \mu_2 e^{\mu_2 \pi} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

eine nichttriviale Lösung besitzt (bzw. $\begin{pmatrix} 1 & 1 \\ e^\pi & (\pi + 1)e^\pi \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$). Dies ist genau dann der Fall, wenn die Determinante der Matrix verschwindet, d.h. $\mu_1 \mu_2 (e^{\mu_2 \pi} - e^{\mu_1 \pi}) = 0$ gilt oder äquivalenterweise

$$\lambda e^{(1 - \sqrt{1 - \lambda})\pi} (1 - e^{2\pi\sqrt{1 - \lambda}}) = 0,$$

also $\lambda = 0$ gilt oder $\sqrt{1 - \lambda}$ eine ganzzahlige Vielfache von i ist, d.h. $\lambda = k^2 + 1$ mit einem $k \in \mathbb{N}$ gilt (im Fall $\lambda = 1$ ist die Matrix regulär, also liegt kein Eigenwert vor). In diesem Fall hat das obige Gleichungssystem wegen $\mu_{1,2} = 1 \pm ik$ die Lösung $c_1 = 1 - ik$, $c_2 = -(1 + ik)$, und somit sind die zum Eigenwert $\lambda_k := k^2 + 1$ gehörigen Eigenfunktionen gerade Vielfache von

$$(1 - ik)e^{\mu_1 x} - c_2(1 + ik)e^{\mu_2 x} = e^x (e^{ikx} - e^{-ikx} - ik(e^{ikx} + e^{-ikx})) = 2ie^x (\sin(kx) - k \cos(kx)).$$

Die Eigenfunktionen zum Eigenwert $k^2 + 1$ von $-L$ sind also Vielfache der reellen Funktion $\exp(x)(\sin(kx) - k \cos(kx))$, während die zum Eigenwert 0 gehörigen Eigenfunktionen konstant sind.

Abschließend sei noch bemerkt, dass man Eigenwerte von $-L$ natürlich auch für Randbedingungen untersuchen kann, die nicht die Form (4.2) haben, z.B. periodische Randbedingungen.

Beispiel 4.12. Das Eigenwertproblem $y'' + \lambda y = 0$ hat unter den periodischen Randbedingungen $y(-\pi) = y(\pi)$, $y'(-\pi) = y'(\pi)$, die reellen Eigenwerte $\lambda_n = n^2$, $n \in \mathbb{N}_0$. Tatsächlich, ist $\lambda < 0$, dann ist die allgemeine Lösung von $y'' + \lambda y = 0$ durch

$$y(x) = c_1 e^{\sqrt{-\lambda}x} + c_2 e^{-\sqrt{-\lambda}x}$$

gegeben, aber y ist für keine Wahl $(c_1, c_2) \neq (0, 0)$ eine 2π -periodische Funktion. Für $\lambda = 0$ ist die allgemeine Lösung $y(x) = c_1 + c_2 x$, und bei $c_2 = 0$ erhält man Konstanten als 2π -periodische Eigenfunktionen. Ist dagegen $\lambda > 0$, so lautet die allgemeine Lösung

$$y(x) = c_1 \sin(\sqrt{\lambda}x) + c_2 \cos(\sqrt{\lambda}x),$$

und diese Funktionen sind 2π -periodisch genau dann, wenn $\lambda = n^2$ mit einem $n \in \mathbb{N}$ gilt, egal wie $c_1, c_2 \in \mathbb{R}$ gewählt wurden.

Kapitel 5

Elementare Lösungsmethoden für partielle Differentialgleichungen

Partielle Differentialgleichungen sind Gleichungen, in die nicht nur die Werte sondern auch die partiellen Ableitungen einer von mehreren Variablen abhängigen Funktion eingehen. Der Unterschied zu gewöhnlichen Differentialgleichungen ist also, dass Ableitungen nach verschiedenen Variablen vorkommen statt nur Ableitungen nach einer Variablen. Genauer hat eine vollständig nichtlineare m -dimensionale partielle Differentialgleichung k -ter Ordnung für eine Funktion $u: \Omega \rightarrow \mathbb{R}^m$ auf einem Gebiet $\Omega \subset \mathbb{R}^n$ die Form

$$f(x, u, Du, \dots, D^k u) = 0$$

mit einer Funktion $f: \Omega \times \mathbb{R}^m \times \mathbb{R}^{nm} \times \dots \times \mathbb{R}^{n^k m} \rightarrow \mathbb{R}^m$. Allerdings kann man über vollständig nichtlineare partielle Differentialgleichungen so gut wie keine allgemein gültigen Aussagen treffen. Daher konzentriert man sich üblicherweise auf spezielle Typen von (häufig linearen) partiellen Differentialgleichungen. Aber auch deren Theorie ist so weitläufig und komplex, dass es anmaßend und hoffnungslos wäre, darüber in diesem Kapitel einen Überblick geben zu wollen. Deshalb sprechen wir in diesem Kapitel nahezu keine Theorie an, sondern wir ermitteln stattdessen spezielle Lösungen von gewissen partiellen Differentialgleichungen, indem wir in die partielle Differentialgleichung Ansätze einsetzen, die auf gewöhnliche Differentialgleichungen führen.

5.1 Die Laplace-Gleichung

Eine partielle Differentialgleichung zweiter Ordnung ist die Laplace-Gleichung

$$\Delta u = 0 \tag{5.1}$$

für eine Funktion $u: \Omega \rightarrow \mathbb{R}$ auf einem Gebiet $\Omega \subset \mathbb{R}^n$, wobei der Laplace-Operator Δ für zweimal stetig differenzierbares u durch $\Delta u := \operatorname{div}(\operatorname{grad} u) = \sum_{i=1}^n \frac{\partial^2 u}{\partial x_i^2}$ definiert ist.

Definition 5.1. Eine Lösung u von (5.1) in Ω nennt man eine harmonische Funktion auf Ω .

Beispiel 5.2. Die Funktionen $u_1(x, y) := xy$ und $u_2(x, y) := x^2 - y^2$ sind harmonisch in jedem Gebiet $\Omega \subset \mathbb{R}^2$. Auch $u_3(x, y) := \frac{1}{2} \ln(x^2 + y^2)$ ist harmonisch, aber nur in Gebieten $\Omega \subset \mathbb{R}^2$, die nicht den Ursprung enthalten, denn u kann nicht stetig in den Punkt $(0, 0)$ fortgesetzt werden und ist somit insbesondere keine Lösung von (5.1) in einem Gebiet, das den Ursprung enthält.

Da der Laplace-Operator linear ist, bilden die harmonischen Funktionen auf Ω einen Vektorraum. Insbesondere ist eine Funktion u auf Ω nicht allein durch die Gültigkeit von (5.1) eindeutig festgelegt, sondern wie schon bei gewöhnlichen Differentialgleichungen muss man zusätzliche Bedingungen verlangen, um eine eindeutige Lösung von (5.1) zu bekommen. Während bei gewöhnlichen Differentialgleichungen üblicherweise die Anfangswerte eine Lösung eindeutig festgelegt haben, ist eine Lösung von (5.1) schon eindeutig durch ihre Werte auf dem Rand $\partial\Omega$ bestimmt. Dies ist nach unserer Diskussion von Randwertproblemen in 4.1 auch nicht weiter verwunderlich, denn ist $\Omega := (a, b)$ ein Intervall ($n = 1$), so ist (5.1) gerade die gewöhnliche Differentialgleichung zweiter Ordnung $u'' = 0$, und deren Lösung ist eindeutig durch die Werte $u(a)$ und $u(b)$ festgelegt. Als weiteres Beispiel dafür, dass harmonische Funktionen in Ω durch ihre Werte auf dem Rand $\partial\Omega$ eindeutig festgelegt sind, werden wir harmonische Funktionen auf dem Kreis $\Omega := B_1(0) \subset \mathbb{R}^2$ diskutieren. Dazu ist es nützlich, in Polarkoordinaten zu arbeiten.

5.1.1 Der Laplace-Operator in Polarkoordinaten

Wir werden häufig nach Lösungen von (5.1) (oder von anderen partiellen Differentialgleichungen, die den Laplace-Operator enthalten) suchen, die rotationssymmetrisch sind.

Definition 5.3. Eine Funktion $u: \Omega \rightarrow \mathbb{R}$ auf einem Gebiet $\Omega \subset \mathbb{R}^n$ heißt rotationssymmetrisch, falls es eine Funktion $y:]0, \infty[\rightarrow \mathbb{R}$ mit $u(x) = y(\|x\|)$ für alle $0 \neq x \in \Omega$ gibt.

Für rotationssymmetrische Funktionen u kann man Δu besonders leicht ausrechnen.

Lemma 5.4. Ist die Funktion u auf $\Omega \setminus \{0\}$ zweimal stetig differenzierbar und rotationssymmetrisch mit $u(x) = y(|x|)$ für alle $0 \neq x \in \Omega$, dann gilt

$$(\Delta u)(x) = y''(|x|) + \frac{n-1}{|x|} y'(|x|).$$

Beweis: Da u auf $\Omega \setminus \{0\}$ zweimal stetig differenzierbar ist, ist auch y auf $(0, \infty)$ zweimal stetig differenzierbar. Wegen $(\text{grad } u)(x) = y'(|x|) \frac{x}{|x|}$ gilt

$$(\Delta u)(x) = \text{div}\left(y'(|x|) \frac{x}{|x|}\right) = y''(|x|) \left(\sum_{i=1}^n \frac{x_i^2}{|x|^2}\right) + y'(|x|) \left(\sum_{i=1}^n \frac{|x| + x_i \frac{x_i}{|x|}}{|x|^2}\right)$$

also wegen $|x|^2 = \sum_{i=1}^n x_i^2$ und $\sum_{i=1}^n \frac{1}{|x|} = \frac{n}{|x|}$ die zu beweisende Formel. \square

Ist die Funktion u nicht rotationssymmetrisch, so vereinfachen sich manche Rechnungen trotzdem noch, wenn man von kartesischen Koordinaten (x_1, \dots, x_n) zu Polarkoordinaten $(r, \varphi_1, \dots, \varphi_{n-1})$ übergeht. In zwei Dimensionen (d.h. $n = 2$) erhält man mit $x = r \cos(\varphi)$, $y = r \sin(\varphi)$, für eine Funktion $u = u(r, \varphi)$ wegen $r = |(x, y)|$, $\varphi = \arctan(\frac{y}{x})$ für $x > 0$, die Gleichung

$$(\text{grad } u)(x, y) = \left(\frac{\partial u}{\partial r} \frac{x}{r} + \frac{\partial u}{\partial \varphi} \frac{(-y)}{r^2}, \frac{\partial u}{\partial r} \frac{y}{r} + \frac{\partial u}{\partial \varphi} \frac{x}{r^2}\right)$$

und damit

$$\begin{aligned} (\Delta u)(x, y) &= \frac{\partial^2 u}{\partial r^2} \frac{x^2}{r^2} + \frac{\partial u}{\partial r} \frac{y^2}{r^3} + \frac{\partial^2 u}{\partial \varphi^2} \frac{y^2}{r^4} + \frac{\partial u}{\partial \varphi} \frac{2xy}{r^4} + \frac{\partial^2 u}{\partial r^2} \frac{y^2}{r^2} + \frac{\partial u}{\partial r} \frac{x^2}{r^3} + \frac{\partial^2 u}{\partial \varphi^2} \frac{x^2}{r^4} + \frac{\partial u}{\partial \varphi} \frac{(-2xy)}{r^4} \\ &= \frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r} + \frac{1}{r^2} \frac{\partial^2 u}{\partial \varphi^2}. \end{aligned}$$

5.1.2 Harmonische Funktionen auf dem Kreis

Wir wollen in diesem Abschnitt mittels Trennung der Variablen nach harmonischen Funktionen auf dem Kreis $B_1(0) \subset \mathbb{R}^2$ suchen. Trennung der Variablen bedeutet dabei, dass wir für die Lösung u von $\Delta u = 0$ in Polarkoordinaten (r, φ) den Ansatz $u(r, \varphi) = R(r)\Phi(\varphi)$ machen und nach einer Funktion $R : [0, 1) \rightarrow \mathbb{R}$ sowie einer 2π -periodischen Funktion $\Phi : \mathbb{R} \rightarrow \mathbb{R}$ suchen, für die u eine Lösung der Laplace-Gleichung $\Delta u = 0$ in $B_1(0)$ ist. Setzt man diesen Ansatz in die zuvor ermittelte Formel

$$\Delta u = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r} + \frac{1}{r^2} \frac{\partial^2 u}{\partial \varphi^2} = 0$$

ein, so erhält man

$$R''\Phi + \frac{1}{r}R'\Phi + \frac{1}{r^2}R\Phi'' = 0.$$

Bringt man alle von r bzw. φ abhängigen Terme auf eine Seite, so ergibt sich daraus

$$\frac{r^2 R'' + rR'}{R} = -\frac{\Phi''}{\Phi}.$$

Da die Gleichung für beliebige $r \in [0, r_0)$ und $\varphi \in \mathbb{R}$ gelten soll, jedoch auf der linken Seite nur r und auf der rechten Seite nur φ vorkommt, müssen beide Seiten gleich einer Konstanten sein, d.h. es muss ein $\lambda \in \mathbb{R}$ geben mit

$$\Phi'' + \lambda\Phi = 0 \quad \text{und} \quad r^2 R'' + rR' - \lambda R = 0.$$

Weil Φ periodisch sein soll, ist die erste Gleichung $\Phi'' + \lambda\Phi = 0$ nichts anderes als das in 4.12 behandelte Eigenwertproblem unter periodischen Randbedingungen. Somit muß $\lambda = n^2$ für ein $n \in \mathbb{N}_0$ gelten, und als zugehörige Lösungen erhält man $\Phi_0 := \frac{a_0}{2}$ bei $\lambda = 0$ bzw. $\Phi_n(\varphi) := a_n \cos(n\varphi) + b_n \sin(n\varphi)$ bei $\lambda = n^2 > 0$. Setzt man den gefundenen Eigenwert in die zweite Gleichung ein, so ergibt sich

$$r^2 R'' + rR' - n^2 R = 0.$$

Für $n = 0$ hat diese Gleichung die allgemeine Lösung $R_0(r) = c_0 + d_0 \ln(r)$, die aber nur bei $d_0 = 0$ auch in $r = 0$ definiert ist. Für $n > 0$ ergibt sich aus dem Ansatz $R(r) = r^k$ die Gleichung $k(k-1) + k - n^2 = 0$ und somit als allgemeine Lösung $R_n(r) = c_n r^n + d_n r^{-n}$, wobei wieder $d_n = 0$ gelten muss, damit R_n auch in $r = 0$ definiert ist. Versteckt man nun noch im Produkt $R_n \Phi_n$ die Konstante c_n in den Konstanten a_n, b_n , so erhält man in Polarkoordinaten $u_0(r, \varphi) = \frac{a_0}{2}$ bzw.

$$u_n(r, \varphi) = (a_n \cos(n\varphi) + b_n \sin(n\varphi))r^n \quad \text{für} \quad n \in \mathbb{N}$$

als Lösungen von $\Delta u = 0$ in $B_1(0)$. Da der Operator Δ aber linear ist, sind auch beliebige Summen von u_n Lösungen, d.h. allgemeiner ist

$$u(r, \varphi) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(n\varphi) + b_n \sin(n\varphi))r^n$$

eine harmonische Funktion auf dem Kreis $B_1(0)$, falls die Koeffizienten $a_n, b_n \in \mathbb{R}$ so gewählt sind, dass die Reihe in $B_1(0)$ in einem geeigneten Sinne konvergiert.

Man bemerke, dass die gefundene Lösung durch ihre Werte auf dem Rand $\partial B_1(0)$ eindeutig festgelegt ist. Denn $u = g$ auf $\partial\Omega$ ist in Polarkoordinaten gleichbedeutend mit

$$g(\varphi) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(n\varphi) + b_n \sin(n\varphi)),$$

also muss die rechte Seite gerade die Fourier-Reihe von g sein, und diese ist eindeutig durch g bestimmt. Das Problem, Lösungen u einer partiellen Differentialgleichung in Ω zu suchen, die $u = g$ auf dem Rand $\partial\Omega$ erfüllen, bezeichnet man üblicherweise als Dirichlet-Problem. Macht man sich noch genauere Gedanken über die Konvergenz der Reihen, so erhält man den folgenden Satz.

Satz 5.5. *Das Dirichlet-Problem $\Delta u = 0$ in $B_1(0) \subset \mathbb{R}^2$, $u = g$ auf $\partial B_1(0)$, besitzt in Polarkoordinaten die eindeutige Lösung*

$$u(r, \varphi) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(n\varphi) + b_n \sin(n\varphi)) r^n,$$

falls $g = g(\varphi)$ eine quadratintegrierbare 2π -periodische Funktion ist und a_n, b_n ihre Fourier-Koeffizienten sind.

Ähnlich kann man harmonische Funktionen auf dem Kreis finden, die einer Neumannschen Randbedingung $\partial_\nu u = g$ auf $\partial\Omega$ genügen, wobei ν das äußere Einheitsnormalenvektorfeld an $\partial\Omega$ bezeichnet und im Falle $\Omega = B_1(0)$ durch $\nu(x) = \frac{x}{|x|}$ gegeben ist.

5.2 Die Diffusionsgleichung

Die lineare partielle Differentialgleichung

$$\frac{\partial u}{\partial t} = \Delta u \tag{5.2}$$

für eine zeitabhängige Funktion $u : (0, \infty) \times \Omega \rightarrow \mathbb{R}$ auf einem Gebiet $\Omega \subset \mathbb{R}^n$ heißt Diffusionsgleichung, da die zeitliche Entwicklung der Dichte bei einem Diffusionsprozess mit Hilfe von (5.2) modelliert werden kann. Andere Anwendungen interpretieren u als Temperatur, und in diesem Kontext wird (5.2) auch Wärmeleitungsgleichung genannt.

Wir suchen zunächst nach selbstähnlichen Lösungen von (5.2) auf $\Omega := \mathbb{R}^n$, d.h. nach Lösungen der Form $u(t, x) = t^\alpha y(t^\beta x)$ mit einer Funktion $y : \mathbb{R}^n \rightarrow \mathbb{R}$. Dazu bemerke man, dass mit $u : (0, \infty) \times \mathbb{R}^n \rightarrow \mathbb{R}$ auch $u(a^2 t, ax)$ für $a > 0$ eine Lösung von (5.2) ist, denn es gilt

$$\frac{\partial}{\partial t}(u(a^2 t, ax)) = a^2 \frac{\partial u}{\partial t}(a^2 t, ax) = a^2 (\Delta u)(a^2 t, ax) = \Delta(u(a^2 t, ax)).$$

Diese Invarianz unter Skalierungen deutet schon darauf hin, dass es möglicherweise eine selbstähnliche Lösung von (5.2) gibt, denn könnte man $a = \frac{1}{\sqrt{t}}$ setzen (was man aber nicht darf, da a in der obigen Rechnung als konstant angenommen wurde), so würde man nach der Skalierung die nur vom Quotienten $\frac{x}{\sqrt{t}}$ abhängige Funktion $u(1, \frac{x}{\sqrt{t}})$ erhalten.

Lemma 5.6. *Jede rotationssymmetrische selbstähnliche Lösung u der Diffusionsgleichung (5.2) auf \mathbb{R}^n ist ein Vielfaches von*

$$\Phi(t, x) := \frac{1}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{|x|^2}{4t}}.$$

Beweis: Da wir nach einer rotationssymmetrischen und selbstähnlichen Lösung von (5.2) suchen, betrachten wir den Ansatz $u(t, x) = t^\alpha y(t^\beta |x|)$. Setzt man diesen in (5.2) ein, so erhält man mit Hilfe von 5.4 die Gleichung

$$t^\alpha y' \beta t^{\beta-1} |x| + \alpha t^{\alpha-1} y = t^\alpha y'' t^{2\beta} + t^\alpha \frac{n-1}{|x|} y' t^\beta.$$

Um t aus dieser Gleichung zu eliminieren, muss man $\alpha - 1 = \alpha + 2\beta$ verlangen, dann folgt mit $\xi = t^\beta|x|$ und $y = y(\xi)$ die Gleichung

$$\beta\xi y' + \alpha y = y'' + \frac{n-1}{\xi}y'.$$

Damit man beide Seiten als Ableitung darstellen kann, muss man $\alpha = n\beta$ verlangen und mit ξ^{n-1} multiplizieren, dann erhält man

$$\beta(\xi^n y)' = (\xi y)''.$$

Nimmt man an, dass $\xi^n y(\xi)$ sowie $\xi^{n-1}y'(\xi)$ für $\xi \rightarrow \infty$ gegen Null konvergieren, so erhält man daraus $y' = \beta\xi y$ und somit $y(\xi) = Ce^{\beta\frac{\xi^2}{2}}$ mit einer Konstanten $C \in \mathbb{R}$. Aus den Gleichungen $\alpha - 1 = \alpha + 2\beta$ und $\alpha = n\beta$ folgt zudem $\alpha = -\frac{n}{2}$ und $\beta = -\frac{1}{2}$, so dass $\xi^2 = \frac{|x|}{t}$ und

$$u(t, x) := \frac{C}{t^{\frac{n}{2}}} e^{-\frac{|x|^2}{4t}}.$$

□

Die Funktion Φ wird häufig Fundamentallösung von (5.2) genannt. Der folgende Satz rechtfertigt diese Bezeichnung und auch die Wahl der Konstanten.

Aufgabe 5.7. Zeigen Sie, dass $\int_{\mathbb{R}^n} \Phi(t, x) dx = 1$ für alle $t > 0$ gilt.

Satz 5.8. Ist $u_0: \mathbb{R}^n \rightarrow \mathbb{R}$ integrierbar, dann löst $u(t, x) := \int_{\mathbb{R}^n} \Phi(t, x - y)u_0(y) dy$ die Diffusionsgleichung (5.2) in $(0, \infty) \times \mathbb{R}^n$.

Ist u_0 darüberhinaus stetig, so gilt $\lim_{t \searrow 0} u(t, x) = u_0(x)$ für alle $x \in \mathbb{R}^n$.

Beweis: Da $x \mapsto \Phi(t, x)$ für jedes $t > 0$ integrierbar ist, existiert das Faltungsintegral $u(t, x) := \int_{\mathbb{R}^n} \Phi(t, x - y)u_0(y) dy$. Außerdem ist die Funktion Φ auf $(0, \infty) \times \mathbb{R}^n$ beliebig oft differenzierbar und alle ihre Ableitungen sind für jedes $\varepsilon > 0$ auf $[\varepsilon, \infty) \times \mathbb{R}^n$ gleichmäßig beschränkt, also ist auch u auf $(0, \infty) \times \mathbb{R}^n$ beliebig oft differenzierbar und man darf bei der Berechnung der Ableitungen von u nach t bzw. x_i Differentiation und Integration vertauschen. Daher erhält man

$$\frac{\partial u}{\partial t}(t, x) = \int_{\mathbb{R}^n} \frac{\partial \Phi}{\partial t}(t, x - y)u_0(y) dy = \int_{\mathbb{R}^n} (\Delta \Phi)(t, x - y)u_0(y) dy = (\Delta u)(t, x),$$

d.h. u löst die Diffusionsgleichung auf \mathbb{R}^n .

Ist nun u_0 zusätzlich noch stetig (und daher als integrierbare und stetige Funktion insbesondere beschränkt), dann gibt es zu $x_0 \in \mathbb{R}^n$ und $\varepsilon > 0$ ein $\delta > 0$, für das $|y - x_0| < \delta$ die Ungleichung $|u_0(y) - u_0(x_0)| < \varepsilon$ impliziert. Ist nun $|x - x_0| < \frac{\delta}{2}$, dann gilt wegen $\int_{\mathbb{R}^n} \Phi(t, x) dx = 1$ die Ungleichung

$$\begin{aligned} |u(t, x) - u_0(x_0)| &= \left| \int_{\mathbb{R}^n} \Phi(t, x - y)(u_0(y) - u_0(x_0)) dy \right| \\ &\leq \int_{B_\delta(x_0)} \Phi(t, x - y)|u_0(y) - u_0(x_0)| dy + \int_{\mathbb{R}^n \setminus B_\delta(x_0)} \Phi(t, x - y)|u_0(y) - u_0(x_0)| dy. \end{aligned}$$

Das erste Integral ist kleiner als $\int_{B_\delta(x_0)} \Phi(t, x - y)\varepsilon dy \leq \varepsilon$. Für das zweite Integral ergibt sich

wegen $|y - x| \geq \frac{1}{2}|y - x_0|$ für x, y mit $|x - x_0| < \frac{\delta}{2}$, $|y - x_0| \geq \delta$, die Konvergenz

$$\begin{aligned} \int_{\mathbb{R}^n \setminus B_\delta(x_0)} \Phi(t, x - y) |u_0(y) - u_0(x_0)| dy &\leq 2 \|u_0\|_\infty \int_{\mathbb{R}^n \setminus B_\delta(x_0)} \Phi(t, x - y) dy \\ &= \frac{2 \|u_0\|_\infty}{(4\pi t)^{\frac{n}{2}}} \int_{\mathbb{R}^n \setminus B_\delta(x_0)} e^{-\frac{|x-y|^2}{4t}} dy \\ &\leq \frac{2 \|u_0\|_\infty}{(4\pi t)^{\frac{n}{2}}} \int_{\mathbb{R}^n \setminus B_\delta(x_0)} e^{-\frac{|x_0-y|^2}{16t}} dy \\ &= \frac{2 \|u_0\|_\infty}{(4\pi t)^{\frac{n}{2}}} \int_\delta^\infty e^{-\frac{r^2}{16t}} r^{n-1} dr \rightarrow 0 \end{aligned}$$

für $t \searrow 0$. Insbesondere unterschreitet der Wert des zweiten Integrals die Zahl $\varepsilon > 0$ für genügend kleines $t > 0$, also gilt für $|x - x_0| < \frac{\delta}{2}$ und genügend kleines $t > 0$ die Ungleichung $|u(t, x) - u_0(x_0)| < 2\varepsilon$, die zu zeigen war. \square

Aufgrund dieses Satzes bezeichnet man $u(t, x) := \int_{\mathbb{R}^n} \Phi(t, x - y) u_0(y) dy$ als Lösung der Diffusionsgleichung (5.2) auf \mathbb{R}^n unter der Anfangsbedingung $u(0, x) = u_0(x)$, und dies aufgrund des in der folgenden Aufgabe formulierten Resultats sogar auch dann noch, falls u_0 nur integrierbar und nicht stetig ist.

Aufgabe 5.9. Zeigen Sie, dass $u(t, \cdot) \rightarrow u_0$ in $L^1(\mathbb{R}^n)$ für $t \searrow 0$ gilt.

Bevor wir im folgenden Abschnitt nach wandernden Wellen in einer nichtlinearen Reaktions-Diffusionsgleichung suchen, wollen wir noch die Eindeutigkeit von Lösungen von (5.2) sowie die stetige Abhängigkeit für Anfangswerte aus $L^2(\mathbb{R}^n)$ zeigen.

Satz 5.10. Sind $u, \tilde{u}: (0, \infty) \times \mathbb{R}^n \rightarrow \mathbb{R}$ Lösungen der Diffusionsgleichung (5.2) zu den Anfangswerten $u_0, \tilde{u}_0 \in L^2(\mathbb{R}^n)$, dann gilt $\|u(t) - \tilde{u}(t)\|_{L^2} \leq \|u_0 - \tilde{u}_0\|_{L^2}$.

Inbesondere gibt es zu jedem Anfangswert aus $L^2(\mathbb{R}^n)$ höchstens eine Lösung von (5.2).

Beweis: Die Differenz $v := u - \tilde{u}$ löst aufgrund der Linearität auch die Diffusionsgleichung (5.2). Daher gilt nach partieller Integration

$$\frac{d}{dt} \|v\|_{L^2}^2 = \langle v, \frac{\partial v}{\partial t} \rangle_{L^2} = \langle v, \Delta v \rangle_{L^2} = -\langle \text{grad } v, \text{grad } v \rangle_{L^2} \leq 0.$$

Also gilt $\|v(t)\|_{L^2}^2 \leq \|v(0)\|_{L^2}^2$ für fast alle $t > 0$ und somit die Behauptung. \square

Abschließend wollen wir noch kurz anmerken, dass man für spezielle Gebiete $\Omega \neq \mathbb{R}^n$ natürlich auch mittels Trennung der Variablen nach speziellen Lösungen suchen kann, ähnlich wie wir dies in Abschnitt 5.1.2 für die Laplace-Gleichung getan haben und in Abschnitt 5.3.1 für die Wellengleichung tun werden.

5.2.1 Reaktions-Diffusionsgleichungen

Eine partielle Differentialgleichung der Form

$$\frac{\partial u}{\partial t} = \Delta u + f(u) \tag{5.3}$$

für eine zeitabhängige Funktion $u : (0, T) \times \Omega \rightarrow \mathbb{R}$ auf einem Gebiet $\Omega \subset \mathbb{R}^n$ heißt Reaktions-Diffusionsgleichung, wobei die vorgegebene (nichtlineare) Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ der Reaktionsterm genannt wird. Räumlich konstante Lösungen von (5.3) entsprechen offensichtlich gerade den Lösungen der gewöhnlichen Differentialgleichung $u' = f(u)$.

Beispiel 5.11. Gleichungen von der Form (5.3) treten beispielsweise in Populationsmodellen auf, wenn man zusätzlich annimmt, dass sich die Population räumlich im Gebiet Ω durch Diffusion verteilt. Geht man wie in Beispiel 1.20 von einer maximal tragbaren Populationsgröße u_{\max} aus und nimmt man $u_{\max} = 1$ an, so ergibt sich $f(u) = \lambda(1 - u)u$ als Reaktionsterm und daher als Reaktions-Diffusionsgleichung die sogenannte Fishersche Gleichung

$$\frac{\partial u}{\partial t} = \Delta u + \lambda(1 - u)u.$$

Wir suchen hier nach speziellen Lösungen von (5.3), die die Form

$$u(t, x) = y(x + ct)$$

mit einer Funktion $y : \mathbb{R}^n \rightarrow \mathbb{R}$ und $0 \neq c \in \mathbb{R}^n$ besitzen. Solche Lösungen werden wandernde Wellen (engl.: traveling waves) genannt, und c ist die Geschwindigkeit der Welle. Setzt man diesen Ansatz im Fall $n = 1$, $\Omega = \mathbb{R}$, in die Gleichung (5.3) ein, so erhält man für y die gewöhnliche Differentialgleichung

$$cy' = y'' + f(y)$$

oder äquivalenterweise das ebene System

$$\begin{aligned} y' &= z \\ z' &= cz - f(y). \end{aligned} \tag{5.4}$$

Setzt man voraus, dass $f(0) = 0 = f(1)$ gilt, so sind $(0, 0)$ und $(1, 0)$ Ruhelagen dieses Systems. In diesem Fall sind besonders Wellenfronten interessant, von denen man zusätzlich $\lim_{\xi \rightarrow -\infty} y(\xi) = 0$, $\lim_{\xi \rightarrow \infty} y(\xi) = 1$ und $0 \leq y \leq 1$ verlangt. Wellenfronten entsprechen daher gerade heteroklinen Orbits des ebenen Systems von $(0, 0)$ nach $(1, 0)$ in $[0, 1] \times \mathbb{R}$.

Definition 5.12. Sind x^* und \tilde{x}^* Ruhelagen eines lokalen Flusses, so nennt man den Orbit durch einen Punkt x einen heteroklinen Orbit von x^* nach \tilde{x}^* , wenn $\alpha(x) = \{x^*\}$ und $\omega(x) = \{\tilde{x}^*\}$ gilt.

Bezeichnet F eine Stammfunktion von f , so ist das ebene System (5.4) im Spezialfall $c = 0$ ein Hamiltonsches System mit $H(y, z) := \frac{1}{2}z^2 + F(y)$. Da H in diesem Fall konstant entlang von Orbits ist, verlaufen die Orbits innerhalb der Höhenlinien von H und man hat einen genauen Überblick über das Phasenportrait. Ist $c < 0$, dann ist H immerhin noch eine Lyapunov-Funktion für (5.4), denn

$$\dot{H}(y, z) = f(y)z + z(cz - f(y)) = c|z|^2 \leq 0.$$

Zwar ist $H^{-1}((-\infty, 0])$ i.a. nicht kompakt, jedoch ist bei $f(0) = 0$ und $f'(0) > 0$ der Punkt 0 ein striktes lokales Minimum von F und daher $H^{-1}((-\infty, 0]) \cap \Omega' = \{0\}$ für eine Umgebung Ω' von $(0, 0)$. Somit ist $(0, 0)$ stabil für $c \leq 0$, d.h. es kann im Fall $c \leq 0$ keine der von uns gesuchten Wellenfronten geben. Das folgende Lemma gibt Auskunft über den Fall $c > 0$.

Lemma 5.13. Gilt $f(0) = 0 = f(1)$ und ist f' in $[0, 1]$ streng monoton fallend mit $f'(0) > 0$, $f'(1) < 0$, dann sind $(0, 0)$ und $(1, 0)$ die einzigen Ruhelagen von (5.4) in $[0, 1] \times \mathbb{R}$. Die Ruhelage $(0, 0)$ ist ein instabiler Strudel für $0 < c < 2\sqrt{f'(0)}$ und ein instabiler Knoten für $c \geq 2\sqrt{f'(0)}$, während $(1, 0)$ für alle $c > 0$ ein Sattel ist.

Beweis: Da f' streng monoton fallend ist, ist f streng konkav auf $[0, 1]$, also sind 0 und 1 die einzigen Nullstellen von f in $[0, 1]$ und somit $(0, 0)$ und $(1, 0)$ die einzigen Ruhelagen von (5.4). Die Linearisierung der rechten Seite von (5.4) im Punkt (y, z) ist $\begin{pmatrix} 0 & 1 \\ -f'(y) & c \end{pmatrix}$. und hat die Eigenwerte $\lambda = \frac{c}{2} \pm \frac{1}{2}\sqrt{c^2 - 4f'(y)}$. Wegen $f'(0) > 0$ ist $(0, 0)$ für $0 < c < 2\sqrt{f'(0)}$ ein instabiler Strudel des linearisierten Systems, für $c = 2\sqrt{f'(0)}$ ein uneigentlicher instabiler Knoten und für $c > 2\sqrt{f'(0)}$ ein instabiler Knoten, während $(1, 0)$ wegen $f'(1) < 0$ für alle $c > 0$ ein Sattel ist. Abschließend muss man noch zeigen, dass nahe der hyperbolischen Ruhelagen $(0, 0)$ und $(1, 0)$ nicht nur eine topologische Äquivalenz, sondern sogar eine C^1 -Äquivalenz zur Linearisierung vorliegt, auf diesen Teil des Beweises gehen wir hier aber nicht ein. \square

Aufgrund von Lemma 5.13 kann auch im Fall $0 < c < 2\sqrt{f'(0)}$ keine der von uns gesuchten Wellenfronten existieren, denn wenn $(0, 0)$ ein instabiler Strudel ist, dann verläßt jede Lösung nahe $(0, 0)$ den Bereich $[0, 1] \times \mathbb{R}$ und wird zwischendurch auch einmal negativ in der ersten Komponente. Um die Existenz von Wellenfronten für $c > 2\sqrt{f'(0)}$ zu zeigen, müssen wir nachweisen, dass die stabile Mannigfaltigkeit von $(1, 0)$ eine negativ invariante Teilmenge des Abstoßungsbereiches von $(0, 0)$ schneidet.

Definition 5.14. Ist x^* Ruhelage eines lokalen Flusses Φ , so heißt die Teilmenge $W_s(x^*) := \{x \mid \lim_{t \rightarrow \infty} \Phi(t, x) = x^*\}$ die stabile und $W_u(x^*) := \{x \mid \lim_{t \rightarrow -\infty} \Phi(t, x) = x^*\}$ die instabile Mannigfaltigkeit von x^* .

Zur Erinnerung: Für hyperbolische lineare Flüsse haben wir in Satz 3.16 bewiesen, dass die stabile und die instabile Mannigfaltigkeit des Ursprungs lineare Unterräume sind und sie mit X_s und X_u bezeichnet. Für von einer nichtlinearen Differentialgleichung mit einem C^k -Vektorfeld als rechter Seite erzeugte lokale Flüsse kann man allgemeiner zeigen, dass für hyperbolische Ruhelagen x^* die Mengen $W_s(x^*)$ und $W_u(x^*)$ immersierte C^k -Untermannigfaltigkeiten sind, deren Tangentialraum in x^* der stabile bzw. instabile Unterraum der Linearisierung ist, siehe [Amann, (19.11), (19.12c)].

Dieses Resultat wenden wir nun unter den Voraussetzungen von Lemma 5.13 an. Der stabile Unterraum der Linearisierung in $(1, 0)$ wird durch den Eigenvektor $(-1, -\lambda_-)$ von $\begin{pmatrix} 0 & 1 \\ -f'(1) & c \end{pmatrix}$ zum Eigenwert $\lambda_- = \frac{c}{2} - \frac{1}{2}\sqrt{c^2 - 4f'(1)} < 0$ aufgespannt, und wir bezeichnen mit $\Gamma_1 \subset W_s((1, 0))$ den Orbit, der für $t \rightarrow \infty$ aus der Richtung $(-1, -\lambda_-)$ in $(1, 0)$ hineinläuft. Für $c > 2\sqrt{f'(0)}$ wird der am stärksten abstoßende Unterraum der Linearisierung in $(0, 0)$ durch den Eigenvektor $(1, \lambda_+)$ von $\begin{pmatrix} 0 & 1 \\ -f'(0) & c \end{pmatrix}$ zum Eigenwert $\lambda_+ = \frac{c}{2} + \frac{1}{2}\sqrt{c^2 - 4f'(0)} > 0$ aufgespannt, und wir bezeichnen mit $\Gamma_2 \subset W_u((0, 0))$ den Orbit, der für $t \rightarrow -\infty$ aus der Richtung $(1, \lambda_+)$ in $(0, 0)$ hineinläuft.

Aufgrund der Voraussetzungen von Lemma 5.13 hat f in $[0, 1]$ eine eindeutige Maximalstelle $a \in (0, 1)$. Desweiteren gilt $z' = 0$ auf dem Graphen $\{(y, z) \mid z = \frac{f(y)}{c}\}$ von $\frac{1}{c}f$ gilt. Daher hat die zweite Komponente jeder Lösung von (5.4) eine Extremalstelle, wenn Sie den Graphen von $\frac{1}{c}f$ schneidet, und zwar wegen $z'' = cz' - f'(y)y' = -f'(y)z$ eine Minimalstelle, falls der Schnitt in (y, z) mit $y \in (0, a)$ stattfindet, bzw. eine Maximalstelle, wenn der Schnitt in (y, z) mit $y \in (a, 1)$ erfolgt.

Nun betrachten wir den Orbit Γ_2 , der von $(0, 0)$ aus in den positiven Quadranten hineinläuft. Der Orbit Γ_2 kann den positiven Quadranten nicht durch die positive z -Achse verlassen, denn $\langle (-1, 0), (z, cz - f(y)) \rangle = -z < 0$ gilt für $z > 0$. Da Γ_2 nahe $(0, 0)$ oberhalb des Graphen von $\frac{1}{c}f$ liegt, kann Γ_2 aber auch den Graphen von $\frac{1}{c}f$ nicht schneiden, solange Γ_2 in $(0, a) \times (0, \infty)$ verläuft, denn ansonsten hätte die zweite Komponente des Orbits im Schnittpunkt als von der Zeit t abhängige Funktion ein lokales Minimum. Darüberhinaus gilt $y' = z > 0$, solange Γ_2 den

positiven Quadranten nicht verläßt, und daher verläßt Γ_2 den Bereich $(0, a) \times (0, \infty)$ durch die Gerade $y = a$ in einem Punkt (a, b) mit $b > \frac{f(a)}{c}$.

Für den Orbit Γ_1 liegt rückwärts in der Zeit gerade die umgekehrte Situation vor: Γ_1 läuft aus $(0, 1) \times (0, \infty)$ kommend nach $(1, 0)$ hinein, und Γ_1 kann wegen $\langle (0, -1), (z, cz - f(y)) \rangle = f(y) > 0$ für $z = 0, y \in (0, 1)$, nicht durch die positive y -Achse in das Gebiet $(0, 1) \times (0, \infty)$ hineingekommen sein. Da Γ_1 nahe $(1, 0)$ unterhalb des Graphen von $\frac{1}{c}f$ verläuft, kann Γ_1 aber auch den Graphen von $\frac{1}{c}f$ nicht schneiden, solange Γ_1 in $(a, 1) \times (0, \infty)$ verläuft.

Das von $[0, 1] \times \{0\}$, $\{(y, z) \mid a \leq y \leq 1, z = \frac{f(y)}{c}\}$, $\{(y, z) \mid y = a, \frac{f(a)}{c} \leq z \leq b\}$ und Γ_2 berandete Gebiet ist nun negativ invariant, denn es gilt $\langle (-f'(y), c), (z, cz - f(y)) \rangle = -f'(y)z > 0$ für $a < y \leq 1, z = \frac{f(y)}{c} > 0$, sowie $\langle (1, 0), (z, cz - f(y)) \rangle = z > 0$ für $y = a, \frac{f(a)}{c} \leq z \leq b$, und Γ_2 kann als Orbit nicht von einem anderen Orbit geschnitten werden. Außerdem liegt dieses Gebiet wegen $b \leq af(a)/c$ und

$$H(a, b) = \frac{1}{2}b^2 + F(a) \leq \frac{(af(a))^2}{2c^2} + F(a)$$

unter der zusätzlichen Voraussetzung $\frac{af(a)}{\sqrt{2(F(1)-F(a))}} < c$ innerhalb von $\{(y, z) \mid y < 1, z > 0, H(y, z) < F(1)\}$. und gehört somit wegen $\dot{H}(y, z) > 0$ für $z > 0$ zum Abstoßungsbereich von $(0, 0)$. Daher muss Γ_1 für $t \rightarrow -\infty$ von $(0, 0)$ her gekommen sein, d.h. Γ_1 ist der gesuchte heterokline Orbit. Da die stabile Untermannigfaltigkeit von $(1, 0)$ außerdem eindimensional ist, kann es zu einem festen c nicht noch weitere heterokline Orbits von $(0, 0)$ nach $(1, 0)$ geben, und dies beweist den folgenden Satz.

Satz 5.15. *Gilt $f(0) = 0 = f(1)$, ist f' in $[0, 1]$ streng monoton fallend mit $f'(0) > 0, f'(a) = 0$ für $a \in (0, 1)$ und $f'(1) < 0$, und bezeichnet F eine Stammfunktion von f , so existiert für jedes $c > \max(2\sqrt{f'(0)}, \frac{af(a)}{\sqrt{2(F(1)-F(a))}})$ genau eine Wellenfront von (5.3).*

Abschließend sei darauf hingewiesen, dass die oben benutzte Beweis-Methode, mit der wir den heteroklinen Orbit Γ_1 gefunden haben, allgemein unter dem Namen Wazewski-Prinzip in der Literatur bekannt ist.

5.3 Die Wellengleichung

Im letzten Abschnitt haben wir gesehen, dass in Reaktions-Diffusionsgleichungen wandernde Wellen als Lösungen auftreten können. In diesem Abschnitt wollen wir uns mit einer weiteren partiellen Differentialgleichung zweiter Ordnung beschäftigen, deren Lösungen sogar naturgemäß Wellen sind, nämlich mit der Wellengleichung

$$\frac{\partial^2 u}{\partial t^2} = c^2 \Delta u \tag{5.5}$$

für eine zeitabhängige Funktion $u: (0, \infty) \times \Omega \rightarrow \mathbb{R}$ auf einem Gebiet $\Omega \subset \mathbb{R}^n$, wobei die Zahl $c \in (0, \infty)$ die Geschwindigkeit der Wellen angibt.

Im Fall $\Omega = \mathbb{R}$ kann man leicht sehen, dass jede Lösung eine Überlagerung von Wellen ist. Denn der Differentialoperator aus (5.5) besitzt in diesem Fall die Zerlegung

$$\frac{\partial^2}{\partial t^2} - c^2 \frac{\partial^2}{\partial x^2} = \left(\frac{\partial}{\partial t} + c \frac{\partial}{\partial x} \right) \left(\frac{\partial}{\partial t} - c \frac{\partial}{\partial x} \right) = \left(\frac{\partial}{\partial t} - c \frac{\partial}{\partial x} \right) \left(\frac{\partial}{\partial t} + c \frac{\partial}{\partial x} \right).$$

Nun kann man aber leicht nachrechnen, dass auf \mathbb{R} die lineare partielle Differentialgleichung erster Ordnung $\frac{\partial u}{\partial t} \pm c \frac{\partial u}{\partial x} = 0$, die auch lineare Transportgleichung genannt wird, Lösungen von der Form $u(t, x) = y(x \mp ct)$ mit einer beliebigen stetig differenzierbaren Funktion $y: \mathbb{R} \rightarrow \mathbb{R}$ besitzt. Somit ist u auch eine Lösung der Wellengleichung (5.5), und aufgrund der Linearität ist daher

$$u(t, x) = y_1(x + ct) + y_2(x - ct)$$

für beliebige zweimal stetig differenzierbare Funktionen $y_1, y_2: \mathbb{R} \rightarrow \mathbb{R}$ eine Lösung der Wellengleichung (5.5) auf $\Omega = \mathbb{R}$. Die Lösung besteht also aus der Überlagerung einer nach links und einer nach rechts wandernden Welle.

Gibt man sich eine Anfangsauslenkung $u(0, x) = u_0(x)$ und eine Anfangsgeschwindigkeit $\frac{\partial u}{\partial t}(0, x) = u_1(x)$ der Welle vor, so kann man die Funktionen y_1, y_2 eindeutig aus

$$y_1(x) + y_2(x) = u_0(x) \quad \text{und} \quad cy_1'(x) - cy_2'(x) = u_1(x)$$

bestimmen. Tatsächlich ist die Lösung zu vorgegebenen Anfangsdaten auch im Mehrdimensionalen eindeutig, wie der folgende Satz besagt.

Satz 5.16. *Ist $\Omega \subset \mathbb{R}^n$ ein beschränktes Gebiet, so gibt es zu $u_0, u_1: \Omega \rightarrow \mathbb{R}$ höchstens eine Lösung u von (5.5) mit $u(0, x) = u_0(x)$ und $\frac{\partial u}{\partial t}(0, x) = u_1(x)$.*

Beweis: Sind u, \tilde{u} zwei Lösungen zu denselben Anfangsdaten, so löst die Differenz $v := u - \tilde{u}$ aufgrund der Linearität die Wellengleichung (5.2) zu den Nullfunktionen als Anfangsdaten. Definiere die Energie

$$E(t) := \frac{1}{2} \int_{\Omega} \left| \frac{\partial v}{\partial t}(t, x) \right|^2 + |(\text{grad } v)(t, x)|^2 dx,$$

dann gilt mittels partieller Integration

$$\frac{d}{dt} E(t) = \int_{\Omega} \frac{\partial v}{\partial t} \frac{\partial^2 v}{\partial t^2} + \left\langle \text{grad } v, \text{grad } \frac{\partial v}{\partial t} \right\rangle dx = \int_{\Omega} \frac{\partial v}{\partial t} \left(\frac{\partial^2 v}{\partial t^2} - \Delta v \right) dx = 0,$$

d.h. die Energie ist konstant entlang von Lösungen. Aufgrund von $v(0, \cdot) = 0 \Rightarrow (\text{grad } v)(0, \cdot) = 0$ und $\frac{\partial v}{\partial t}(0, \cdot) = 0$ gilt $E(0) = 0$, also auch $E(t) = 0$ für fast alle $t > 0$. Dies impliziert $\frac{\partial v}{\partial t} = 0$ und $\text{grad } v = 0$ in $(0, \infty) \times \Omega$, wegen $v(0, \cdot) = 0$ also auch $v = 0$ in $(0, \infty) \times \Omega$, was zu beweisen war. \square

5.3.1 Eine Wellengleichung mit Transportterm

In den Anwendungen treten in partiellen Differentialgleichungen mit einem Differentialoperator zweiter Ordnung in den Raumvariablen wie Δ häufig zusätzlich noch Terme mit Ableitungen erster Ordnung in den Raumvariablen auf, beispielsweise Transportterme der Form $\text{div}(f(u))$ mit einer Funktion $f: \mathbb{R} \rightarrow \mathbb{R}^n$. Wir wollen anhand der speziellen Wellengleichung mit linearem Transportterm

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} - 2 \frac{\partial u}{\partial x}$$

auf dem Intervall $\Omega := (0, \pi)$ unter den Neumannschen Randbedingungen $\frac{\partial u}{\partial x}(0) = 0 = \frac{\partial u}{\partial x}(\pi)$ zeigen, dass auch in diesem Fall Trennung der Variablen zum Ziel führen kann.

Dazu setzen wir den Ansatz $u(t, x) = T(t)X(x)$ in die Gleichung ein und erhalten $T''X = X''T - 2X'T$, d.h.

$$X'' - 2X' + \lambda X = 0 \quad \text{und} \quad T'' + \lambda T = 0$$

mit einer Konstanten λ . Die erste Gleichung ist gerade das in 4.11 behandelte Eigenwertproblem. Daher kommt für λ nur der Wert $\lambda_0 = 0$ oder $\lambda_k = k^2 + 1$ mit $k \in \mathbb{N}$ infrage, und für die zugehörige Eigenfunktion X_k gilt $X_0 = c_0$ bzw.

$$X_k := c_k e^x (\sin(kx) - k \cos(kx))$$

mit einer Konstanten c_k . Die zweite Gleichung hat mit $\lambda = k^2 + 1$ die allgemeine Lösung

$$T_k(t) = a_k \cos(\sqrt{k^2 + 1}t) + b_k \sin(\sqrt{k^2 + 1}t)$$

bzw. bei $\lambda = 0$ die Lösung $T(t) = a_0 + b_0 t$. Versteckt man nun noch im Produkt $T_k X_k$ die Konstante c_k in den Konstanten a_k, b_k , so ergibt sich

$$u_k(t, x) = e^x (\sin(kx) - k \cos(kx)) \left(a_k \cos(\sqrt{k^2 + 1}t) + b_k \sin(\sqrt{k^2 + 1}t) \right)$$

für jedes $k \in \mathbb{N}$ als Lösung bzw. $u_0(t, x) = a_0 + b_0 t$. Aufgrund der Linearität der Gleichung ist also auch

$$u(t, x) = a_0 + b_0 t + e^x \sum_{k=1}^{\infty} (\sin(kx) - k \cos(kx)) \left(a_k \cos(\sqrt{k^2 + 1}t) + b_k \sin(\sqrt{k^2 + 1}t) \right)$$

eine Lösung, falls die Koeffizienten $a_k, b_k \in \mathbb{R}$ so gewählt sind, dass die Reihe in $\mathbb{R} \times (0, \pi)$ in einem geeigneten Sinne konvergiert. Sucht man beispielsweise nach einer Lösung, bei der die Anfangsauslenkung durch $u(0, x) = e^x (2 \sin(3x) - 6 \cos(3x))$ gegeben und keine Anfangsgeschwindigkeit vorhanden ist, d.h. $\frac{\partial u}{\partial t}(0, x) = 0$ gilt, dann ergibt sich durch Koeffizientenvergleich $b_k = 0$, $a_3 = 2$ sowie $a_k = 0$ für $k \neq 3$, also als Lösung

$$u(t, x) = \cos(\sqrt{10}t) \exp(x) (2 \sin(3x) - 6 \cos(3x)) .$$

Literaturverzeichnis

- [Amann] H. AMANN, *Gewöhnliche Differentialgleichungen*, de Gruyter, 1983.
- [Barbu] V. BARBU, *Nonlinear semigroups and differential equations in Banach spaces*, Editura Academiei, 1976.
- [Chicone] C.C. CHICONE, *Ordinary Differential Equations with Applications*, Texts in applied Mathematics 34, Springer, 2006.
- [Coddington,Levinson] E. CODDINGTON, N. LEVINSON, *Theory of Ordinary Differential Equations*, McGraw-Hill, 1955.
- [Elstrodt] J. ELSTRODT, *Maß- und Integrationstheorie*, Springer, 1999.
- [Engel,Nagel] K.-J. ENGEL, R. NAGEL, *One-Parameter Semigroups for Linear Evolution Equations*, Graduate Texts in Mathematics, 2000.
- [Evans] L.C. EVANS, *Partial Differential Equations*, Graduate Studies in Mathematics 19, AMS, 1998.
- [Forster] O. FORSTER, *Analysis II*, Vieweg, 2008.
- [Königsberger] K. KÖNIGSBERGER, *Analysis 1*, Springer, 1990.
- [Königsberger] K. KÖNIGSBERGER, *Analysis 2*, Springer, 1993.
- [Pöschel] J. PÖSCHEL, *Dynamische Systeme*, Skript, <http://www.poeschle.de/ds-09/pdf/ds-skript.pdf>, 2009.
- [Strauss] W.A. STRAUSS, *Partielle Differentialgleichungen*, Vieweg, 1992.
- [Walter] W. WALTER, *Gewöhnliche Differentialgleichungen*, Springer, 2000.

Sporadisches Sachwortverzeichnis

dynamisches System

kontinuierliches lokales, 59

Fluss

lokaler, 59

Fundamentalsystem, 47

Gleichung

Laplace-, 85

Lösung

globale, 9

Methode

Trennung der Variablen, 12

Variation der Konstanten, 12

rotationssymmetrisch, 86

Variation

der Konstanten, 48

Vektorfeld

zeitabhängiges, 9