



# Gewöhnliche Differentialgleichungen

JOCHEN GLÜCK

Vorlesungsmanuskript  
Sommersemester 2020



# Inhaltsverzeichnis

<b>Stand des Manuskripts</b>	<b>v</b>
<b>Benutzungsanleitung: Wie verwende ich dieses Manuskript?</b>	<b>vii</b>
Aufbau . . . . .	vii
Dies ist ein Vorlesungsmanuskript . . . . .	viii
Achtung: Fehler! . . . . .	viii
<b>1 Differentialgleichungen – Was und wozu?</b>	<b>1</b>
1.1 Eine erste Begriffsklärung . . . . .	1
1.2 Fragestellungen . . . . .	5
1.3 Wodurch wird eine Differentialgleichung gewöhnlich? . . . . .	7
1.4 Ein erster Blick Richtung Eindeutigkeit – Anfangswerte und Anfangswertprobleme . . . . .	8
1.5 Ergänzungen . . . . .	10
<b>2 Modellierung und geometrische Intuition</b>	<b>15</b>
2.1 Modellierung mit Hilfe von Differentialgleichungen . . . . .	15
2.2 Welche Werte darf eine Größe annehmen? . . . . .	19
2.3 Zustandsraum und Vektorfelder . . . . .	22
2.4 Ergänzungen . . . . .	27
<b>3 Was sind Differentialgleichungen – noch einmal</b>	<b>29</b>
3.1 Begriffsklärung: Zweiter Versuch . . . . .	29
3.2 Differentialgleichungen erster und höherer Ordnung . . . . .	33
3.3 Differentialgleichungen zweiter Ordnung und Phasenraum . . . . .	33
3.4 Autonome und nicht-autonome Differentialgleichungen . . . . .	35
3.5 Ergänzungen . . . . .	36
<b>4 Existenz und Eindeutigkeit von Lösungen</b>	<b>37</b>
4.1 Lokale Existenz und globale Eindeutigkeit: Der Satz von Picard– Lindelöf . . . . .	37
4.2 Lokale Existenz: Der Satz von Peano . . . . .	48
4.3 Das maximale Existenzintervall . . . . .	50
4.4 Globale Existenz . . . . .	59
4.5 Lösungsfluss für autonome Differentialgleichungen . . . . .	66
4.6 Ergänzungen . . . . .	68
<b>5 Lineare Differentialgleichungen</b>	<b>71</b>
5.1 Der Begriff der linearen Differentialgleichung . . . . .	71

5.2	Die Struktur des Lösungsraums . . . . .	73
5.3	Homogene Systeme und Fundamentallösungen . . . . .	77
5.4	Homogene Systeme im autonomen Fall: Die Matrix-Exponentialfunktion . . . . .	84
5.5	Berechnungsverfahren für die Matrix-Exponentialfunktion . . . . .	90
5.6	Inhomogenitäten . . . . .	100
5.7	Differentialgleichungen höherer Ordnung: Ein Wiedersehen . . . . .	104
5.8	Ergänzungen . . . . .	110
<b>6</b>	<b>Berechnung von Lösungen im nichtlinearen Fall</b>	<b>113</b>
6.1	Eine schlechte (oder gute?) Nachricht vorab . . . . .	113
6.2	Trennung der Variablen . . . . .	114
6.3	Erhaltungsgrößen . . . . .	121
6.4	Exakte Differentialgleichungen und integrierende Faktoren . . . . .	127
6.5	Ergänzungen . . . . .	134
<b>7</b>	<b>Qualitative Eigenschaften von Lösungen</b>	<b>137</b>
7.1	Fehlerfortpflanzung und des Lemma von Gronwall . . . . .	137
7.2	Stetige Abhängigkeit und der Begriff der Wohlgestelltheit . . . . .	142
7.3	Invarianz von Mengen . . . . .	148
7.4	Differenzierbare Abhängigkeit vom Anfangswert . . . . .	155
7.5	Ergänzungen . . . . .	158
<b>8</b>	<b>Langzeitverhalten autonomer Systeme</b>	<b>161</b>
8.1	Langzeitverhalten im linearen Fall . . . . .	161
8.2	Stabilität und Attraktivität . . . . .	168
8.3	Linearisierte Stabilität . . . . .	172
8.4	Ljapunov-Funktionen . . . . .	179
8.5	Ergänzungen . . . . .	183
<b>9</b>	<b>Zwei Beispiele im Detail</b>	<b>185</b>
9.1	Ein Epidemie-Modell . . . . .	185
9.2	Endliche Approximation der Wärmeleitungsgleichung auf einem Intervall . . . . .	192
	<b>Appendices</b>	<b>197</b>
<b>A</b>	<b>Metrische und normierte Räume</b>	<b>199</b>
A.1	Metrische Räume: Grundbegriffe . . . . .	199
A.2	Konvergenz und Vollständigkeit . . . . .	201
A.3	Topologische Begriffe und Stetigkeit . . . . .	203
A.4	Normierte Räume . . . . .	204
A.5	Kompaktheit . . . . .	208

A.6 Zusammenhängende Mengen und Zusammenhangsargumente	211
<b>B Fixpunktsätze</b>	<b>217</b>
B.1 Aperitif: Wozu Fixpunktsätze? . . . . .	217
B.2 Der Banachsche Fixpunktsatz . . . . .	218
B.3 Die Fixpunktsätze von Brouwer und Schauder . . . . .	221
<b>C Matrix-Analysis</b>	<b>223</b>
C.1 Matrix-Normen . . . . .	223
C.2 Diagonalisierbarkeit und die Jordansche Normalform . . . . .	226
<b>Literaturverzeichnis</b>	<b>229</b>



## *Stand des Manuskripts*

- Es handelt sich um den vorläufig fertigen Stand des Manuskriptes.<sup>1</sup>
- Datum dieses vorläufig fertigen Standes: 26. August 2020

---

<sup>1</sup>Vorläufig deshalb, weil das Manuskript womöglich noch aktualisiert wird, wenn mir eine größere Zahl an Fehlern bekannt werden sollte.



# Benutzungsanleitung: Wie verwende ich dieses Manuskript?

## Aufbau

Der Aufbau des Manuskripts ist einfach:

- **Nummerierung von Abschnitten:** Es ist unterteilt in Kapitel (nummeriert als 1, 2, ...), und jedes Kapitel ist unterteilt in Abschnitte (nummeriert als 1.1, 1.2, ...).
- **Nummerierung von Umgebungen:** Definitionen, Sätze, Bemerkungen, usw. sind innerhalb der Abschnitte mit einer gemeinsamen Nummer durchlaufend nummeriert.
- **Einstiegsfragen:** Am Anfang jedes Kapitels finden Sie eine Liste mit *Fragen zum Einstieg*; sie dienen zur Motivation des Stoffes in diesem Kapitel.

Bitte lesen Sie diese Fragen jeweils in Ruhe durch und denken Sie ein wenig darüber nach. Dies wird Ihnen beim Verständnis und der Einordnung des Stoffes helfen.

- **Aufgaben:** Im Text sind immer wieder kleine Aufgaben eingestreut. Diese Aufgaben tauchen nicht auf den Übungsblätter auf; sie sind meist sehr kurz und sind als Verständnishilfe beim Nachbereiten der Vorlesung gedacht.

Nehmen Sie sich bitte beim Lesen des Manuskripts die Zeit um diese Aufgaben zu bearbeiten, denn dies wird Ihnen das Verständnis des Stoffes erleichtern.

- **Weitere Literatur:** Am Ende vieler Abschnitte finden Sie einen Absatz mit der Bezeichnung *Literaturstellen und verwandte Ergebnisse*; dort werden verschiedene Abschnitte in Büchern vorgeschlagen, in denen der Stoff des Abschnitts noch weiter oder tiefer behandelt wird.
- **Ergänzungen:** Am Ende jedes Kapitels steht ein Abschnitt mit dem Titel *Ergänzungen*. In diesen Abschnitten finden Sie Zusatzinformationen, die nicht Teil des Vorlesungsstoffes sind und auch nicht in der Vorlesung behandelt werden.

Diese Ergänzungen können Ihnen aber beim Verständnis helfen oder Ihre Neugier auf weitere verwandte Themen wecken. Bitte entscheiden Sie selbst, welche Ergänzungen Sie lesen möchten und welche nicht.

- **Anhang:** Am Ende des Manuskripts befindet sich ein Anhang aus mehreren Kapiteln (nummeriert als A, B, ...). Dort werden Konzepte und Resultate aus anderen Vorlesungen wiederholt, die für die Behandlung von *Gewöhnlichen Differentialgleichungen* wichtig sind.
- **Biografische Informationen:** Wenn im Text der Name eines Mathematikers oder einer Mathematikerin auftaucht, finden Sie eine kurze Information zu dieser Person in einer Fußnote. Der Name der Person ist in vielen dieser Fußnoten zu einem Eintrag im *MacTutor History of Mathematics Archive* verlinkt; wenn Sie auf den Namen klicken, können Sie biografische Informationen zur Person in diesem Online-Archiv nachlesen.

Direkter Link zum *MacTutor Archive*: [Bitter hier klicken](#).

## Dies ist ein Vorlesungsmanuskript

Die Betonung beim Begriff *Vorlesungsmanuskript* liegt auf *Vorlesung* (im Sommersemester 2020 aber etwas weniger als sonst). Dies bedeutet:

- Nehmen Sie an der Vorlesung teil.  
In der speziellen Epidemie-Situation im Sommersemester 2020 bedeutet dies konkreter: Sehen Sie sich die Vorlesungsvideos an – möglichst konzentriert und ohne Ablenkung.
- Mindestens genauso wichtig: Bearbeiten Sie die Übungsaufgaben und geben Sie Ihre Lösungen zur Korrektur ab.
- Wenn Sie Fragen haben: Kontaktieren Sie mich gerne und fragen Sie!

## Achtung: Fehler!

Dieses Manuskript enthält mit an Sicherheit grenzender Wahrscheinlichkeit Fehler. Mit hoher Wahrscheinlichkeit enthält es sogar viele Fehler.

- Wenn Sie glauben, einen Fehler entdeckt zu haben: Geben Sie mir bitte Bescheid (wirklich!), am einfachsten per E-Mail.
- E-Mail-Adresse: [jochen.glueck@uni-passau.de](mailto:jochen.glueck@uni-passau.de)

# Differentialgleichungen – Was und wozu?

## Fragen zum Einstieg.

- (a) Warum wächst eine Bakterienpopulation in einer Petrischale exponentiell, solange die Population noch klein ist?
- (b) Welche Form kann eine Satellitenbahn haben? Woran liegt das?
- (c) Das Kohlenstoff-Isotop C14 ist radioaktiv und hat eine Halbwertszeit von ca. 5700 Jahren. Wenn Sie ein Mikrogramm C14 betrachten, wie viel davon ist nach 1000 Jahren noch übrig?
- (d) Suchen Sie im Internet nach dem Begriff *C14-Methode*<sup>1</sup>.

## 1.1 Eine erste Begriffsklärung

In quantitativen Wissenschaften betrachtet man sehr häufig Größen, die sich mit der Zeit ändern – z.B. Ortsangaben in der Physik, Konzentrationen in der Chemie, Populationsgrößen in der Biologie, Aktienkurse in der Finanzmathematik, usw. Wenn wir solch eine Größe mit dem Symbol  $x$  bezeichnen und die Zeit mit  $t$ , dann ist  $x$  also eine Funktion, die von der Variablen  $t$  abhängt.

In vielen Modellen ist es unmöglich, die Funktion  $x$  ad hoc zu bestimmen. Betrachten Sie zum Beispiel die folgenden – sehr einfachen – Modelle, die den Einstiegsfragen zu diesem Kapitel entnommen sind:

- Beispiele 1.1.1.** (a) Sei  $b$  die Anzahl Bakterien in einer Petrischalen-Kultur. Welche Form hat die von der Zeit abhängige Funktion  $b$ ?
- (b) Es bezeichne  $x(t) \in \mathbb{R}^3$  die Position eines Satelliten zum Zeitpunkt  $t$  (in diesem Fall ist  $x(t)$  also nicht eine einzelne Größe, sondern ein Vektor, der aus drei Größen besteht). Wenn Sie den Ort  $x(0)$  sowie die Geschwindigkeit zum Zeitpunkt 0 kennen, woher wissen Sie dann, wo sich der Satellit zum Zeitpunkt  $t$  befinden wird?
- (c) Es bezeichne  $k$  die Konzentration des Isotops C14 in der gesamten Kohlenstoffmenge einer archäologischen Probe. Wie hängt  $k$  von der Zeit  $t$  ab?

---

<sup>1</sup>Auch *Radiokarbonmethode*, entwickelt vom US-amerikanischen Chemiker *Willard Libby* (1908 – 1980), der dafür 1960 mit dem Nobelpreis für Chemie ausgezeichnet wurde.

Während es in vielen – erst recht in komplizierteren – Situationen nicht möglich ist, ad hoc zu erkennen, von welcher Gestalt die gesuchte Funktion sein muss, ist es andererseits sehr oft möglich einen Zusammenhang zwischen der Größe und ihrer zeitlichen Änderung, d.h. ihrer Ableitung, anzugeben. Um dies zu veranschaulichen, betrachten wir wieder die drei Situationen aus den Beispielen 1.1.1:

**Beispiele 1.1.2.** (a) Bakterien vermehren sich durch Zellteilung, und solange sie genügend Ressourcen und Platz zur Verfügung haben und die Rahmenbedingungen sich nicht ändern, tun sich dies mit einer festen Rate. Deshalb ist es sinnvoll anzunehmen, dass die Zunahme der Bakterienanzahl, d.h. die Ableitung  $\dot{b}(t)$ , proportional zur Anzahl der bereits vorhandenen Bakterien ist.

Dies bedeutet, dass es eine reelle Zahl  $c > 0$  gibt mit der Eigenschaft

$$\dot{b}(t) = cb(t).$$

Diese Gleichung ist plausibel, solange die Anzahl der Bakterien nicht zu groß ist – denn wenn die Petrischale sich füllt, bleiben nicht mehr genügend Ressourcen um die Reproduktionsrate aufrecht zu erhalten.

(b) In der klassischen Mechanik gelten die sogenannten *Newtonschen Bewegungsgleichungen*: Sie besagen, dass die Beschleunigung  $a$ , die ein Objekt erfährt, multipliziert mit seiner Masse  $m$  gleich der Kraft  $F$  ist, die auf das Objekt wird, d.h. es gilt  $ma = F$ .

Weil  $x(t)$  die Position des Satelliten zum Zeitpunkt  $t$  ist, ist seine Geschwindigkeit zum Zeitpunkt  $t$  gleich der Ableitung  $\dot{x}(t)$  und seine Beschleunigung  $a$  gleich der zweiten Ableitung  $\ddot{x}(t)$ . Es gilt also  $m\ddot{x}(t) = F$ .

Die Gravitationskraft, die auf einen Satelliten wirkt hängt aber von einer aktuellen position ab – sie nimmt zum Beispiel ab, wenn der Satellit sich weiter von der Erde entfernt. Das heißt,  $F$  ist in Wirklichkeit eine Funktion, die von der aktuellen Position des Satelliten abhängt. Somit erhalten wir insgesamt die Gleichung

$$m\ddot{x}(t) = F(x(t))$$

für die Position des Satelliten. (In der Physik ist eine genaue Formel für  $F(x(t))$  bekannt – Sie können diese unter dem Stichwort *Newtonsches Gravitationsgesetz* nachlesen.)

(c) Aus der Theorie des radioaktiven Zerfalls weiß man, dass ein C14-Atom (ebenso wie jedes andere radioaktive Atom) in einem vorgegebenen Zeitraum mit einer festen Wahrscheinlichkeit zerfällt. Also ist die Rate

mit der die Konzentration  $k(t)$  abnimmt – d.h. die negative Ableitung  $-\dot{k}(t)$  – direkt proportional zur momentanen Konzentration  $k(t)$ . Es gibt also eine Zahl  $c > 0$  so, dass die Gleichung

$$-\dot{k}(t) = ck(t)$$

gilt.

Ein paar Kommentare zur Notation sind an dieser Stelle angebracht:

**Bemerkungen 1.1.3.** (a) Für die Ableitung einer Funktion  $x : \mathbb{R} \rightarrow \mathbb{R}$  (oder allgemeiner: einer Funktion  $x : \mathbb{R} \rightarrow \mathbb{R}^d$ ) gibt es verschiedene Notationen. In der Analysis 1 wird oft die Notation  $x'$  für die Ableitung der Funktion  $x$  eingeführt. In der Physik und in der Theorie der Differentialgleichungen ist auch die Notation  $\dot{x}$  sehr üblich – insbesondere, wenn die Variable, von der  $x$  abhängt anschaulich als Zeit interpretiert wird.

Sie können grundsätzlich beiden Notationen verwenden, und Sie sollten auch beide Notationen kennen.

(b) Lassen Sie sich nicht durch die Notation verwirren:

In der Analysis I werden Funktionen häufig als  $f, g, h, \dots$  bezeichnet, und die Variablen werden auf mit  $x, y, z, \dots$  oder  $x_1, x_2, x_3, \dots$  bezeichnet.

In der Theorie der gewöhnlichen Differentialgleichungen werden die gesuchten Funktionen oft (aber natürlich nicht immer) mit  $x$  bezeichnet, und die Variable mit  $t$ . Der Grund dafür ist, dass die Variable in vielen Anwendungen (wie in den Beispielen 1.1.2) die Zeit darstellt.

Wenn man über konkrete Anwendungen spricht (wie in obigen Beispielen) ist es oft sinnvoll die gesuchte Funktion statt mit  $x$  mit einem Buchstaben zu bezeichnen, der aufgrund der konkreten Anwendung naheliegt – z.B. mit eine  $b$  für die Bakterienanzahl oder einem  $k$  für die Konzentration.

An den Beispielen 1.1.2 erkennt man folgendes: Die Beschreibung einer zeitabhängigen Größe  $x$  passiert in vielen Modellen nicht, indem man direkt eine Formel für  $x$  angibt. Stattdessen findet man häufig eine Gleichung für den Zusammenhang zwischen  $x$  und ihrer Ableitung  $\dot{x}$  (oder auch ihren höheren Ableitungen wie z.B.  $\ddot{x}$ ).

Eine solche Gleichung bezeichnet man als *Differentialgleichung*. Lassen Sie uns dies noch einmal in einer (vorerst etwas vagen) Definition festhalten:

**Definition 1.1.4 (Differentialgleichung – erster Definitionsversuch).** Eine *Differentialgleichung* ist eine Gleichung für eine Funktion  $x$ , die einen Zusammenhang zwischen  $x$  und den Ableitungen von  $x$  herstellt.

Die Funktion  $x$  wird hierbei als Unbekannte – d.h. als das gesuchte Objekt – aufgefasst.

Aus den Beispielen 1.1.2 kann man noch weitere wichtige Erkenntnisse gewinnen:

- Wenn man  $x(t)$  zu einem Zeitpunkt (oder zu allen Zeitpunkten) berechnen will, muss man also die gegebene Differentialgleichung lösen.
- Situationen, die aus unterschiedlichen Wissenschaften stammen und die auf den ersten Blick sehr verschieden wirken, können manchmal mit dergleichen oder einer sehr ähnlichen Differentialgleichung beschrieben werden. Zum Beispiel unterschieden sich die Differentialgleichungen

$$\dot{b}(t) = cb(t) \quad \text{und} \quad -\dot{k}(t) = ck(t)$$

in den Beispielen 1.1.2(a) und (c) nur durch ein Vorzeichen.

- Um eine Differentialgleichung für eine Größe herzuleiten, benötigt man ein Verständnis des zugrundeliegenden wissenschaftlichen Fachgebiets: Für die Gleichung der Bakterienanzahl muss man zumindest grob wissen, wie Bakterien sich vermehren; um die Gleichung für die Satellitenposition herleiten zu können, muss man die Newtonschen Bewegungsgleichungen kennen (und damit man mit der Gleichung konkret rechnen kann, benötigt man außerdem eine Formel für  $F(x(t))$  – diese Formel, die oben nicht explizit angegeben wurde, stammt aus der Gravitationstheorie).
- Differentialgleichungen enthalten häufig sogenannte *Parameter*, die von der genauen Situation abhängen, die man betrachtet.

In Beispiel 1.1.2(a) wird der Parameter  $c$  unter anderem von der Art des Bakteriums abhängen, von der Nährstoffkonzentration in der Petrischale und von der Temperatur.

In Beispiel 1.1.2(c) hängt der Parameter  $c$  davon ab, welches radioaktive Isotop man betrachtet: Er hat für das Kohlenstoff-Isotop C14 einen völlig anderen Wert als zum Beispiel für das Plutonium-Isotop Pu239.

Die Parameter müssen in konkreten Anwendungen durch Messung und durch statistische Verfahren bestimmt werden.

- In konkreten Anwendungen sind Differentialgleichungen ein Teil von Modellen – damit unterliegen Sie den Modellannahmen! Wenn die Modellannahmen schlecht sind oder in der Realität nicht zutreffen, sagt die Differentialgleichung nichts über das wirkliche Verhalten der Größe aus. Zum Beispiel:

In Beispiel 1.1.2(a) haben wir angenommen, dass die Rahmenbedingungen konstant sind und dass alle Bakterien genügend Nährstoffe zur Verfügung haben um sich optimal zu vermehren. Die erste dieser Bedingungen stimmt zum Beispiel nicht mehr, wenn sich die Temperatur im Laufe der Zeit ändert; die zweite Bedingung stimmt nicht mehr, wenn die Petrischale voller wird. In diesen Fällen ist die angegebene Differentialgleichung nicht mehr geeignet um  $b(t)$  realistisch zu beschreiben.

In Beispiel 1.1.2(b) haben wir verwendet, dass die Newtonschen Bewegungsgleichungen gelten und dass nur Gravitationskräfte auf den Satelliten wirken. Die erste Annahme ist nur für Geschwindigkeiten (näherungsweise) richtig, die deutlich unter der Lichtgeschwindigkeit liegen (was für Satelliten tatsächlich der Fall ist). Die zweite Annahme ist nur dann vernünftig, wenn der Satellit sich weit genug von der Erde entfernt befindet, dass (fast) keine atmosphärische Reibung mehr auftritt. Selbst dann muss man aber noch beachten, dass die Erde nicht die einzige relevante Gravitationsquelle ist – zum Beispiel beeinflusst die Gravitation des Mondes die Bahn von Satelliten.

## 1.2 Fragestellungen

Wir wollen also Gleichungen studieren, die laut Definitionen 1.1.4 einen Zusammenhang zwischen einer Funktion  $x$  und den Ableitungen von  $x$  herstellen. Solche Gleichung zu „studieren“ bedeutet, dass man sich relevante Fragen über diese Gleichungen stellt und versucht sie zu beantworten. Welche Fragen über Differentialgleichungen könnten relevant sein? Hier sind einige Beispiele:

- (a) *Modellierung*: Wenn man das Verhalten einer gesuchten Größe in einer bestimmten Situation modellieren will, was muss man tun, um zu einer Differentialgleichung zu gelangen?

Dies wurde in Abschnitt 1.2 bereits an einigen Beispielen erläutert; in Kapitel 2 gehen wir noch einmal anhand weiterer Beispiele darauf ein.

- (b) *Geometrische Intuition*: Wie kann man sich eine Differentialgleichung und ihre Lösungen anschaulich vorstellen?

Dies besprechen wir ebenfalls in Kapitel 2.

- (c) *Existenz von Lösungen*: Wenn der Parameter  $c$  in Beispiel 1.1.2 konkret gewählt ist – sagen wir zum Beispiel  $c = 3$  – gibt es dann überhaupt eine Funktion  $x : \mathbb{R} \rightarrow \mathbb{R}$ , welche die Differentialgleichung

$$\dot{x}(t) = 3x(t)$$

für alle  $t \in \mathbb{R}$  erfüllt?

Auf die Existenz von Lösungen gehen wir in Kapitel 4 ein.

- (d) *Eindeutigkeit von Lösungen*: Angenommen, wir finden eine Funktion  $x : \mathbb{R} \rightarrow \mathbb{R}$ , die die Differentialgleichung

$$\dot{x}(t) = 3x(t)$$

erfüllt – ist dieses dann die einzige Funktion, die das tut? Oder gibt es mehrere solcher Funktionen? Falls es mehrere solcher Funktionen gibt, können wir vielleicht eine zusätzliche Bedingung zur Differentialgleichung hinzufügen, damit die Lösung eindeutig wird?

Die Eindeutigkeit von Lösungen wird ebenfalls in Kapitel 4 besprochen.

- (e) *Berechnung der Lösung*: Kann man die Lösung einer Differentialgleichung explizit berechnen? Für Anwendungen wäre dies sicherlich nützlich, weil man dann die gesuchte Funktion  $x$  konkret angeben kann.

Falls es möglich ist, die Lösung zu berechnen, mit welchen Rechenverfahren kann man dies tun?

Für einige Klassen von Differentialgleichungen beantworten wir diese Frage in den Kapiteln 5 und 6.

- (f) *Qualitative Eigenschaften der Lösung*: Falls es für eine konkrete Differentialgleichung nicht möglich sein sollte, die Lösung explizit zu berechnen (oder falls es uns zumindest nicht gelingt), kann man dann anhand der Differentialgleichung zumindest herleiten, dass die Lösung – wenn wir sie schon nicht kennen – bestimmte Eigenschaften hat, für die wir uns vielleicht interessieren?

Solche Eigenschaften könnten zum Beispiel das Verhalten der Lösung für große Zeiten sein: Konvergiert zum Beispiel  $x(t)$  für  $t \rightarrow \infty$  immer gegen einen Grenzwert?

Solche qualitativen Eigenschaften untersuchen wir in den Kapiteln 7 und 8.

- (g) *Numerische Berechnung der Lösung*: Falls es uns nicht gelingt, die Lösung einer Differentialgleichung (analytisch) auszurechnen, aber wir trotzdem konkrete Werte für  $x(t)$  benötigen (diese ist zum Beispiel in den Ingenieurwissenschaften andauernd der Fall), können wir die Lösung  $x$  dann mit Hilfe von numerischen Verfahren zumindest näherungsweise bestimmen?

Dies ist ein äußerst wichtiges und tiefgreifendes Thema; allerdings ist es nicht Teil der einführenden Veranstaltung.

### 1.3 Wodurch wird eine Differentialgleichung gewöhnlich?

Bisher war nur von *Differentialgleichungen* die Rede. Warum also heißt die Veranstaltung *Gewöhnliche Differentialgleichungen*?

Der Grund ist, dass es in vielen Modellen auch Größen gibt, die – im Gegensatz zu den oben diskutierten Beispielen 1.1.2 – nicht nur von einer Variablen, sondern von mehreren Variablen abhängen. Differentialgleichungen, die das Verhalten solcher Größen beschreiben, enthalten oft Ableitungen nach verschiedenen Variablen. In der Sprache der Analysis 2 nennt man solche Ableitungen *partielle Ableitungen* und die entsprechenden Differentialgleichungen heißen deshalb *partielle Differentialgleichungen*. Das folgende Beispiel zeigt eine sehr bekannte partielle Differentialgleichung.

**Beispiel 1.3.1 (Eine partielle Differentialgleichung: Wärmeleitungsgleichung).** Betrachten Sie einen sehr langen und dünnen Metallstab der Länge  $L$ , der entlang seiner Mantelfläche thermisch isoliert ist. Wir interessieren uns für jeden Zeitpunkt  $t$  für die Temperatur  $u$  an jeder Position im Metallstab. Da die Temperaturverteilung im Stab nicht unbedingt gleichmäßig sein muss (zum Beispiel, wenn der Stab an seinen Enden gekühlt wird), hängt  $u$  also nicht nur von  $t$  aber, sondern auch von der Position  $z$  im Stab, die man betrachtet.

Weil wir angenommen haben, dass der Stab sehr lang und dünn ist, unterscheiden wir nur zwischen Positionen entlang der Länge des Stabes; das bedeutet,  $z$  ist eine reelle Zahl aus dem Intervall  $[0, L]$ , und  $u$  ist, wenn wir zum Beispiel nur Zeiten  $t \geq 0$  betrachten, eine Funktion

$$u : [0, \infty) \times [0, L] \rightarrow \mathbb{R}.$$

Aus physikalischen Gesetzen kann man eine Differentialgleichung herleiten, denen die Temperatur  $u$  genügt. Sie lautet

$$\frac{\partial}{\partial t} u(t, z) = c \frac{\partial^2}{\partial z^2} u(t, z),$$

wobei der Parameter  $c$  eine Konstante ist, die vom Material abhängt, aus dem der Metallstab besteht (d.h.  $c$  ist eine sogenannte *Materialkonstante*). Diese Differentialgleichung bezeichnet man üblicherweise als *Wärmeleitungsgleichung*; es handelt sich um eine partielle Differentialgleichungen, weil sie eine Ableitung nach mindestens zwei verschiedenen Variablen enthält.

Wenn man die Lösung  $u$  dieser Differentialgleichung – genauer: die Lösungen, denn damit die Lösung eindeutig ist, muss man noch weitere Bedingungen festlegen – genügend gut versteht (oder sie sogar explizit berechnen kann), kann man also die Temperaturverteilung im Stab zu verschiedenen Zeiten angeben.

Wenn die gesuchte Funktion hingegen nur von einer reellen Variablen abhängt, können natürlich keine partiellen Ableitungen nach anderen Variablen auftreten. Man spricht in diesem Fall von einer *gewöhnlichen Differentialgleichung* statt von einer *partiellen Differentialgleichung*. Lassen Sie uns dies in der folgenden – ebenfalls etwas vagen – Definition festhalten.

**Definition 1.3.2 (Gewöhnliche Differentialgleichung).** Eine Differentialgleichung heißt *gewöhnlich*, wenn die gesuchte Funktion nur von einer einzigen ein-dimensionalen Variablen abhängt.

Sowohl gewöhnliche als auch partielle Differentialgleichungen sind enorm wichtig für viele Anwendungen. In der Veranstaltung *Gewöhnliche Differentialgleichungen* beschäftigen wir uns, was Sie nicht überraschen wird, nur mit gewöhnlichen Differentialgleichungen. Partielle Differentialgleichungen werden in einer eigenen Vorlesung behandelt.

## 1.4 Ein erster Blick Richtung Eindeutigkeit – Anfangswerte und Anfangswertprobleme

Lassen Sie uns an dieser Stelle vorab schon einmal einen Blick auf das Thema *Eindeutigkeit von Lösungen* werfen. Betrachten wir wieder die drei konkreten Situationen aus den Beispielen 1.1.1 und 1.1.2:

**Beispiele 1.4.1.** (a) Die Anzahl an Bakterien  $b(t)$  zum Zeitpunkt  $t$  haben wir durch die Differentialgleichung

$$\dot{b}(t) = cb(t)$$

für eine Konstante  $c > 0$  beschrieben. Können wir erwarten, dass es nur eine Lösung dieser Differentialgleichung gibt?

Stellen Sie sich vor, im Labor werden mehrere Bakterienkulturen in verschiedenen Petrischalen gezüchtet; nehmen wir an, diese Petrischalen sollen alle identisch sein, mit der Ausnahme dass sich die Anzahl der Bakterien beim Ansetzen der Kulturen zwischen den einzelnen Schalen wesentlich unterscheidet. Dann wäre es nicht sinnvoll anzunehmen, dass die Anzahl der Bakterien in jeder Schale zu jedem Zeitpunkt gleich ist – sie sind ja noch nicht einmal zum Zeitpunkt  $t = 0$  gleich!

(b) Sei wieder  $x(t) \in \mathbb{R}^3$  die Position unseres Satelliten zum Zeitpunkt  $t$ . Aus den Newtonschen Bewegungsgleichungen hatten wir die Differentialgleichung

$$m\ddot{x}(t) = F(x(t))$$

#### 1.4. Ein erster Blick Richtung Eindeutigkeit – Anfangswerte und Anfangswertprobleme

---

erhalten, wobei  $F(x(t))$  bezeichnet, die auf den Satelliten wirkt, wenn er sich an der Stelle  $x(t)$  befindet. Können wir erwarten, dass diese Differentialgleichung nur eine Lösung hat?

Das wäre mehr als verwegen, denn wo der Satellit sich zum Zeitpunkt  $t$  befindet, wird ja nicht nur vom Kraftfeld der Erdanziehung abhängen, sondern auch davon, wo der Satellit sich befunden hat, als er von einem Raumfahrzeug ausgesetzt wurde – also von seiner Position  $x(0)$  zum Zeitpunkt 0.

Selbst wenn aber die Position des Satelliten zum Zeitpunkt 0 bekannt ist, können wir nicht erwarten, dass dadurch seine Bahn bereits bestimmt ist: Stellen Sie sich zwei verschiedene Satelliten vor, die an der selben Position starten, aber zum Startzeitpunkt einen deutlich verschiedenen Geschwindigkeitsvektor besitzen. Dann werden die beiden Positionen dieser Satelliten sofort auseinander laufen. Also müssen wir nicht nur die Position  $x$  zum Zeitpunkt 0, sondern auch die Geschwindigkeit  $\dot{x}(0)$  zum Zeitpunkt 0 kennen, damit wir erwarten können, dass die Bahn des Satelliten (d.h. die Lösung  $x$  unserer Differentialgleichung) durch unsere Kenntnisse eindeutig bestimmt ist.

- (c) Mit der Konzentration  $k(t)$  des Kohlenstoff-Isotops C14 in unserer archäologischen Probe verhält es sich ähnlich wie mit der Bakterienkultur. Wir hatten diese durch die Differentialgleichung

$$-\dot{k}(t) = ck(t) \tag{1.1}$$

beschrieben (mit einer festen Konstante  $c > 0$ ). Anschaulich ist es aber nicht sinnvoll zu erwarten, dass diese Differentialgleichung nur eine Lösung hat – denn wenn wir verschiedene Proben mit verschiedenen Startkonzentrationen betrachten, so wird in jeder Probe der zeitliche Verlauf der C14-Konzentration die Differentialgleichung (1.1) lösen. Aber trotzdem wird der zeitliche Verlauf nicht in allen Proben gleich sein (d.h. es gibt mehrere Funktionen  $k$ , die (1.1) lösen) – den die C14-Konzentration kann sich ja vielleicht schon zum Zeitpunkt 0 unterscheiden (je nachdem, welche Proben wir vorliegen haben).

Dass also die Differentialgleichung (1.1) nur eine Lösung besitzt, können wir vernünftiger Weise nur dann erwarten, wenn wir denn Wert von  $k$  zum Zeitpunkt 0 fixieren.

**Bemerkung 1.4.2.** Man kann sich naheliegender Weise fragen, warum man für die Bakterienpopulation und die C14-Konzentration nur den Funktionswert  $b$  bzw.  $k$  zum Zeitpunkt 0 festlegen muss, um eine eindeutige Lösung zu erhalten, während man für den Satelliten sowohl die Position  $x$  also auch

die Geschwindigkeit  $\dot{x}$  zum Zeitpunkt 0 festlegen muss. Wir werden später noch sehen (zum Beispiel in Abschnitt 3.1), dass dies mit der sogenannten *Ordnung* der Differentialgleichung zusammenhängt, d.h. damit, welche Ordnung von Ableitung in der Differentialgleichung höchstens auftritt.

Wenn man den Wert der gesuchten Lösung einer Differentialgleichung zum Zeitpunkt 0 – also zum Anfang des zeitlichen Verlaufs – festlegt, dann hat man kurz gesprochen einen *Anfangswert* festgelegt; die Differentialgleichung zusammen mit diesem Anfangswert bezeichnet man dann als *Anfangswertproblem*. Man beachte dabei, dass man nicht unbedingt den Zeitpunkt 0 als Anfangszeitpunkt betrachten muss; dies haben wir in den Beispielen 1.4.1 der Einfachheit halber so gemacht, aber man kann den Anfangswert stattdessen auch zu irgendeinem anderen Zeitpunkt  $t_0$  festlegen.

Lassen Sie uns einen ersten Definitionsversuch des Begriffs *Anfangswertproblem* noch einmal explizit festhalten. In Abschnitt 3.1 werden wir den Begriff dann mathematische präzise fassen.

**Definition 1.4.3 (Anfangswertproblem – erster Definitionsversuch).** Ein *Anfangswertproblem* ist eine Differentialgleichung für eine Funktion  $x$ , bei der zusätzlich der Wert von  $x$  zu einem bestimmten Zeitpunkt  $t_0$  festgelegt wird.

Wenn in der Differentialgleichung nicht nur die erste Ableitung von  $x$ , sondern Ableitungen bis zur Ordnung  $n \geq 2$  vorkommen, muss man zum Zeitpunkt  $t_0$  nicht nur den Wert von  $x$ , sondern auch den Wert der Ableitungen  $x^{(1)}, \dots, x^{(n-1)}$  festlegen um ein Anfangswertproblem zu erhalten.

Unsere Diskussion in den Beispielen 1.4.1 legt nahe, dass man für ein Anfangswertproblem eine eindeutige Lösung erwarten kann (im Gegensatz zu einer Differentialgleichung ohne Anfangswert!). Man beachte, dass wir keinerlei mathematisches Argument für diese Eindeutigkeit gegeben haben. Wir haben uns lediglich überlegt, warum es aus Anwendungssicht zumindest sinnvoll ist, zu erwarten, dass man die Werte zu einem Zeitpunkt  $t_0$  festlegen muss, wenn man eine eindeutig bestimmte Lösung möchte. Einen präzisen mathematischen Satz über die Eindeutigkeit der Lösungen von Anfangswertproblemen – den *Satz von Picard–Lindelöf* – werden wir in Abschnitt 4.1 beweisen.

## 1.5 Ergänzungen

### Kontinuierliche und diskrete Zeit: Differentialgleichungen vs. Differenzgleichungen

Betrachten wir erneut eine Größe  $x$ , die von der Zeit abhängt. Da Differentialgleichungen Ableitungen enthalten, kann das Verhalten von  $x$  nur dann durch

eine Differentialgleichung beschrieben werden, wenn die Zeit als Kontinuum – also als ein Intervall in  $\mathbb{R}$  – modelliert wird. Dies ist häufig sinnvoll, zum Beispiel bei vielen Modellen aus der Physik.

Allerdings gibt es auch Anwendungen, in denen es sinnvoller erscheint, die Zeit als eine Folge von diskreten Zeitschritten zu modellieren. Wenn man die Zeitschritte mit natürlichen Zahlen durchnummeriert, ist  $x$  dann eine Abbildung von  $\mathbb{N}$  nach  $\mathbb{R}$ , also nichts anderes als eine Folge von reellen Zahlen. Wir geben einigen Beispiele an.

**Beispiel 1.5.1 (Jährliche Verzinsung).** Lassen Sie uns ein sehr einfaches Beispiel für ein Bankkonto betrachten:

Es bezeichne  $x$  den Saldo des Kontos. Sie eröffnen das Konto am Beginn eines Jahres, und der Saldo  $x(0) = 0$ . Nehmen wir – nur der Einfachheit halber – an, Sie zahlen nun am Anfang jedes Jahres – beginnend direkt nach der Kontoeröffnung – den Betrag  $b$  auf das Konto ein, und am Ende eines Jahres wird die Summe, die über das ganze Jahr auf dem Konto war mit einem Zinssatz  $p$  verzinst (üblicherweise wird  $p$  also im Moment, je nach Art das Kontos eine Zahl zwischen  $0\% = 0$  und  $1\% = 0.01$  sein.)

Wir nummerieren die Jahre mit natürlichen Zahlen durch, d.h. am Ende des ersten Jahres (nach Zinszahlung) bezeichnen wir den Saldo mit  $x(1)$ , am Ende des zweiten Jahres mit  $x(2)$ , usw. Können Sie direkt eine Formel für  $x(n)$  angeben, wenn  $n$  irgendeine natürlich Zahl ist?

Solch eine Formel direkt zu sehen ist schwierig, aber man kann sehr leicht eine *Differenzgleichung* für  $x$  angeben: Wenn am Ende eines Jahres nach Zinszahlung der Saldo  $x(n)$  lautet, wieviel kommt dann im nächsten Jahr durch Einzahlung am Jahresanfang und nächster Zinszahlung am Jahresende hinzu?

Aus den oben beschriebenen Umständen sehen Sie, dass der Betrag  $(1 + p)(x(n) + b)$  hinzukommt, d.h. die Differenz zwischen  $x(n + 1)$  und  $x(n)$  beträgt

$$x(n + 1) - x(n) = (1 + p)(x(n) + b) \quad (1.2)$$

für jedes  $n \in \mathbb{N}_0$ . Es sollte klar sein, warum man die Gleichung (1.2) als *Differenzgleichung* bezeichnet.

Das Beispiel 1.5.1 ist unrealistisch einfach; dennoch sehr gut eine wichtige Idee, die bei der Modellierung eines Sachverhalts auftritt (egal ob in kontinuierlicher Zeit oder in diskreter Zeit): Es ist häufig schwierig bis unmöglich direkt eine explizite Formel für die gesuchte Größe anzugeben. Andererseits ist es häufig viel einfacher eine Differenzgleichung oder Differentialgleichung zu finden, die das Verhalten der Größe beschreibt.

**Aufgaben 1.5.2.** (a) Können Sie die Differenzgleichung (1.2) *lösen*, also eine explizite Formel für  $x(n)$  in Abhängigkeit von  $n$  angeben?

Denken Sie daran, dass wir  $x(0) = 0$  vorausgesetzt haben.

- (b) Wenn  $x(0)$  eine andere Zahl als 0 ist (d.h. wenn das Konto mit einem von Null verschiedenen Saldo eröffnet wurde), können Sie dann ebenfalls die Lösung von (1.2) angeben?
- (c) Was lernen Sie aus (a) und (b) über die *Eindeutigkeit* der Lösung von (1.2)?

Wir geben noch ein weiteres, sehr bekanntes Beispiel für eine Differenzengleichung an.

**Beispiel 1.5.3 (Fibonacci-Folge).** Die *Fibonacci-Folge*<sup>2</sup> ist diejenige Folge reeller Zahlen  $(f_n)_{n \in \mathbb{N}_0}$ , die die folgenden Bedingungen erfüllt:

- (a) Es gilt  $f_0 = 1$  und  $f_1 = 1$ .
- (b) Für jede Zahl  $n \in \mathbb{N}_0$  gilt  $f_{n+2} = f_{n+1} + f_n$ .

Man kann eine explizite Formel für  $f_n$  finden, aber diese ist erstaunlich kompliziert (vielleicht haben Sie diese aber in der Lineare Algebra 1-Vorlesung als Anwendung für die Diagonalisierung von Matrizen besprochen). In den Ergänzungen am Ende von Kapitel 5 kommen wir nochmals auf die Fibonacci-Folge zurück.

**Aufgaben 1.5.4.** (a) Inwiefern halten Sie es für sinnvoll, die Gleichung  $f_{n+2} = f_{n+1} + f_n$  als *Differenzengleichung* zu bezeichnen?

- (b) Können wir eigentlich sicher sein, dass es nur eine Folge  $(f_n)_{n \in \mathbb{N}_0}$  gibt, welche die Eigenschaften erfüllt, die in Beispiel 1.5.3 aufgelistet sind?

Wir sprechen kurz noch eine weitere Situation an, bei der Differenzengleichungen auftauchen.

**Beispiel 1.5.5 (Approximation einer Differentialgleichung durch eine Differenzengleichung).** In vielen technischen Anwendungen (zum Beispiel in der Luftfahrt oder in der Automobilindustrie) müssen Mikrocontroller die Lösung einer Differentialgleichung in Echtzeit berechnen. Die Berechnung der Lösung wird dabei mit einer festen Taktung durchgeführt – d.h. für die Berechnung auf dem Mikrocontroller erscheint die Zeit nicht kontinuierlich, sondern diskret.

Betrachten Sie zum Beispiel die sehr einfache Differentialgleichung

$$\dot{x}(t) = -x(t) + 1. \tag{1.3}$$

---

<sup>2</sup>Benannt nach Leonardo Fibonacci (ca. 1170 – 1250), italienischer Mathematiker.

Es sei  $\tau$  ein kleine Zeit, und unser Mikrocontroller berechne immer im Abstand  $\tau$  den neuen Wert von  $x$ . Er bestimmt also die Werte  $y_0 := x(0)$ ,  $y_1 := x(\tau)$ ,  $y_2 := x(2\tau)$ ,  $y_3 := x(3\tau)$ , usw.

Können Sie eine Differenzgleichung für die Folge  $(y_n)_{n \in \mathbb{N}_0}$  angeben, die die Differentialgleichung (1.3) approximiert? Wenn Sie diese Differenzgleichung aufgeschrieben haben: Können Sie angeben, welche Berechnung der Mikrocontroller in jedem Schritt durchführen muss?



## Modellierung und geometrische Intuition

### Fragen zum Einstieg.

- (a) Denken Sie an die Anzahl der Bakterien  $b$  in Beispiel 1.1.1 zurück. Unabhängig von der tatsächlichen zeitlichen Entwicklung der Bakterienpopulation: Welche Werte kann die Größe  $b$  (zu irgendeinem Zeitpunkt) sinnvoller Weise überhaupt annehmen? Welche Werte würden Ihrer Ansicht nach überhaupt keinen Sinn ergeben?  
Selbe Frage für die Position  $x$  des Satelliten aus Beispiel 1.1.1(b).
- (b) Wie würden Sie eine Funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  graphisch veranschaulichen? Können Sie das an einem konkreten Beispiel zeigen?
- (c) Wie würden Sie eine Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}^2$  graphisch veranschaulichen? Können Sie das ebenfalls an einem konkreten Beispiel zeigen?
- (d) Wie würden Sie eine Funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  graphisch veranschaulichen? Und natürlich: Können Sie das an einem konkreten Beispiel zeigen?
- (e) Wenn  $x(t) \in \mathbb{R}^3$  die Position eines Objektes zum Zeitpunkt  $t$  bezeichnet, warum ist dann die Ableitung  $\dot{x}(t)$  gleich seiner Geschwindigkeit zum Zeitpunkt  $t$ ? Und warum ist die zweite Ableitung  $\ddot{x}(t)$  gleich seiner Beschleunigung zum Zeitpunkt  $t$ ?

## 2.1 Modellierung mit Hilfe von Differentialgleichungen

Wir gehen in diesem Abschnitt noch einmal etwas genauer darauf ein, wie man ein konkretes Problem mit Hilfe einer Differentialgleichung modellieren kann. Hierfür bietet sich noch einmal die Bakterienpopulation aus Beispiel 1.1.1(a) an, da man an diesem Beispiel sehr einfach verschiedene interessante Phänomene aufzeigen kann.

Lassen Sie uns zuerst noch einmal diskutieren, wie man zu einer Differentialgleichung zur Beschreibung der Bakterienpopulation gelangen kann: In Beispiel 1.1.2(a) hatten wir sehr knapp argumentiert, dass die Veränderungsrate  $\dot{b}(t)$  der Bakterienanzahl direkt proportional zur aktuellen Anzahl  $b(t)$  der Bakterien sein wird (zumindest solange  $b(t)$  klein ist), und daraus geschlossen, dass es eine Konstante  $c > 0$  gibt derart, dass die Differentialgleichung

$$\dot{b}(t) = cb(t)$$

gilt. Dies war ein sehr knappes Argument und überzeugt Sie vielleicht nur bedingt. Lassen Sie uns deshalb eine etwas detailliertere Hereustik besprechen, um zu dieser Differentialgleichung zu gelangen:

**Beispiel 2.1.1 (Herleitung der Populationsgleichung mit Hilfe von Differenzenquotienten).** Lassen Sie uns davon ausgehen, dass jedes Bakterium in einer Zeiteinheit im Mittel  $c$  neue Bakterien erzeugt; dabei ist  $c > 0$  eine reelle Zahl, die nicht ganzzahlig sein muss (wenn z.B. ein Bakterium im Schnitt länger als eine Zeiteinheit braucht um sich zu vermehren, wird  $c$  kleiner als 1 sein). Dann erhalten wir für die Gesamtanzahl  $b(t)$  der Bakterien die Gleichung

$$b(t+1) - b(t) = cb(t);$$

dies ist eine sogenannte *Differenzgleichung*. Damit könnte man z.B. eine Formel für  $b(n)$  für alle natürlichen Zahlen  $n$  berechnen.

Nun ist es aus verschiedenen Gründen oft vorteilhaft, nicht nur ganzzahlige Zeiten zu betrachten, sondern eine kontinuierliche Zeit. Zwei dieser Gründe lauten:

- Die Gesamtanzahl der Bakterien ist sehr groß, und es vermehren sich nicht alle Bakterien zur gleichen Zeit, sondern über die Zeit verteilt; deshalb ist es realistischer, einen kontinuierlichen Zeitverlauf anzunehmen.
- Modelle mit kontinuierlicher Zeit lassen sich mathematisch oft einfacher handhaben, weil dann die Mittel der Differential- und Integralrechnung bereitstehen. (Pointiert ausgedrückt: Es ist oft einfacher, mit Ableitungen und Integralen zu rechnen als mit Differenzen und Summen.)

Wenn wir zu einem Modell in kontinuierlicher Zeit gelangen möchten, können wir dies erreichen, indem wir die Zeitschritte immer kleiner wählen: In einer halben Zeiteinheit erzeugt ein Bakterium in etwa  $c/2$  neue Bakterien, und in einer (kurzen) Zeitspanne der Länge  $\tau$  erzeugt ein Bakterium in etwa  $\tau c$  neue Bakterien. Also erhalten wir für jede gegebene Zeit  $t$  und für einen kleinen Zeitschritt  $\tau$  die Gleichung

$$b(t+\tau) - b(t) = \tau cb(t).$$

Nun dividieren wir noch durch  $\tau$  und gelangen so zur Gleichung

$$\frac{b(t+\tau) - b(t)}{\tau} = cb(t).$$

Die linke Seite kommt Ihnen vermutlich vertraut vor: Es handelt sich um einen Differenzenquotienten. Nun betrachten wir immer kleinere Zeitschritte

$\tau$  – d.h. wir betrachten den Grenzwert für  $\tau \downarrow 0$  – und erhalten damit die Differentialgleichung

$$\dot{b}(t) = cb(t).$$

Wir sind also zur selben Gleichung gelangt, die Sie bereits aus Beispiel 1.1.2(a) kennen.

Lassen Sie uns nun eine Variation der Differentialgleichung  $\dot{b}(t) = cb(t)$  diskutieren: Wir wollen eine Bateriaenpopulation betrachten, wenn die Temperatur nicht konstant ist:

**Beispiel 2.1.2 (Eine nicht-autonome Populationsgleichung).** Die Reproduktionsrate eines Bakteriums wird in den meisten Fällen von der Umgebungstemperatur abhängen. Wenn diese sich mit der Zeit ändert, ist also auch die Anzahl  $c$  der neuen Bakterien, die ein Bakterium in einer Zeiteinheit erzeugt, zeitabhängig – d.h. wir müssen die Konstante  $c$  durch eine Zahl  $c(t)$  ersetzen, die von der Zeit abhängt.

Vom Zeitpunkt  $t$  aus betrachtet wird jedes Bakterium in einem nachfolgenden kleinen Zeitabschnitt der Länge  $\tau$  also im Mittel etwa  $\tau c(t)$  Bakterien erzeugen. Somit erhalten wir zum Zeitpunkt  $t$  die Differenzgleichung

$$b(t + \tau) - b(t) = \tau c(t)b(t).$$

Wie in Beispiel 2.1.1 teilen wir nun durch  $\tau$  und lassen  $\tau$  gegen 0 streben. Damit gelangen wir zur Differentialgleichung

$$\dot{b}(t) = c(t)b(t).$$

Hier sehen Sie zum ersten Mal in diesem Manuskript ein Beispiel für eine sogenannte *nicht-autonome Differentialgleichung* – so nennt man Differentialgleichungen, bei denen nicht nur die gesuchte Funktion (in diesem Fall also  $b$ ) von der Zeit abhängt, sondern auch manche Koeffizienten in der Differentialgleichung (in diesem Fall also  $c$ ).

Für unser letztes Beispiel in diesem Abschnitt kehren wir zum autonomen Fall zurück – wir nehmen also wieder an, dass die Rahmenbedingungen zeitlich konstant sind – aber wir berücksichtigen nun, dass die Reproduktionsrate der Bakterien abnehmen wird, wenn ihre Anzahl steigt, z.B. weil sie sich gegenseitig Nährstoffe wegnehmen.

**Beispiel 2.1.3 (Logistische Gleichung).** Betrachten wir wieder die Frage, wieviele neue Bakterien ein einzelnes Bakterium in einer Zeiteinheit erzeugt. Wie in den vorangehenden Beispielen gehen wir davon aus, dass dies  $c$  Stück wären, wenn es keine Konkurrenz um Nährstoffe durch andere Bakterien gäbe.

Da aber tatsächlich andere Bakterien vorhanden sind, wird die tatsächliche Reproduktion etwas niedriger sein.

Um das Modell einfach zu halten, treffen wir hier die – stark vereinfachende – Annahme, dass  $c$  um einen Wert gesenkt wird, der proportional zur bereits vorhandenen Anzahl an Bakterien ist – denn je mehr Bakterien vorhanden sind, desto stärker ist die Konkurrenz um Nährstoffe. Dann gibt es also eine Zahl  $d > 0$  derart, dass ein Bakterium zum Zeitpunkt  $t$  etwa  $c - db(t)$  neue Bakterien in einer Zeiteinheit erzeugt. (Interessant ist hierbei auch, dass  $c - db(t)$  negativ sein kann, wenn  $b(t)$  genügend groß ist! Dieses Modell beinhaltet also die Annahme, dass manche Bakterien absterben, wenn es zu viele von ihnen gibt).

In einem kleinen Zeitschritt  $\tau$  erzeugt ein einzelnes Bakterium also  $\tau(c - db(t))$  neue Bakterien, und wir erhalten somit zum Zeitpunkt  $t$  die Differenzgleichung

$$b(t + \tau) - b(t) = \tau(c - db(t))b(t).$$

Dividieren durch  $\tau$  und der Grenzübergang  $\tau \downarrow 0$  führen uns somit zur Differentialgleichung

$$\dot{b}(t) = (c - db(t))b(t).$$

Diese Differentialgleichung heißt *logistische Gleichung*.

Man beachte: Wenn  $b(t)$  zu Beginn des Experiments noch sehr klein ist, kann man den Term  $db(t)$  näherungsweise vernachlässigen. Dann erhält man die Differentialgleichung  $\dot{b}(t) = cb(t)$  zurück, die wir in den vorangehenden Beispielen diskutiert hatten.

**Aufgabe 2.1.4.** Wir würde die Differentialgleichung für  $b$  aussehen, wenn wir in Beispiel 2.1.3 auch noch eine zeitliche Veränderung der Temperatur zulassen würden?

Zum Abschluss ist es wichtig, noch einmal darauf hinzuweisen, dass alle Modelle, die wir in diesem Unterabschnitt besprochen haben, die Realität sehr stark simplifizieren. Zum Beispiel haben wir in keinem Modell berücksichtigt, dass die räumliche Verteilung der Bakterien in der Petrischale eine Rolle für ihre Reproduktion spielt (weil sich die Bakteriendichte an verschiedenen Stellen der Schale unterscheidet). Würden wir die räumliche Verteilung berücksichtigen, so müssten wir noch zwei Ortsvariablen einführen und würden somit eine partielle Differentialgleichung erhalten.

**Literaturstellen und verwandte Ergebnisse.** Zahlreiche weitere einführende Beispiele dazu, wie man verschiedene Vorgänge mit Hilfe von Differentialgleichungen modellieren kann, finden Sie unter anderem in [Arn92, Abschnitte 1.6 bis 1.18], in [Heu04, Abschnitt 1 in Kapitel I] und in [PW19, Abschnitt 1.1].

## 2.2 Welche Werte darf eine Größe annehmen?

In den bisherigen Beispielen haben wir uns auf den Standpunkt gestellt, dass wir eine Funktion  $x : \mathbb{R} \rightarrow \mathbb{R}$  suchen (oder im Falle des Satelliten eine Funktion  $x : \mathbb{R} \rightarrow \mathbb{R}^3$ ), die eine bestimmte Differentialgleichung erfüllt. Dabei haben wir zwei wichtige Fragen ignoriert:

1. Ist es in der konkreten Situation wirklich sinnvoll zuzulassen, dass  $x(t)$  beliebige reelle Zahlen (bzw. beliebige Punkte im  $\mathbb{R}^3$ ) als Werte annehmen darf? Oder sind hier gewisse Einschränkungen angebracht?
2. Suchen wir wirklich eine Funktion  $x$ , die als Definitionsbereich die komplette reelle Achse besitzt – d.h. benötigen wir unbedingt eine Lösung unserer Differentialgleichung, die zu allen Zeiten existiert? Oder kann es auch sinnvoll sein, sich mit einer Lösung  $x$  zufrieden zugeben, die vielleicht nur auf einem Teilintervall von  $\mathbb{R}$  definiert ist?

Wir gehen zuerst auf die erste Frage näher ein. Die folgende Liste von Beispielen zeigt, dass es sowohl aus Anwendungssicht als auch aus mathematischer Sicht sinnvoll sein kann, die Werte, die eine Lösung annehmen darf, einzuschränken:

**Beispiele 2.2.1.** (a) Betrachten wir wieder die Bakterienpopulation in der Petrischale. Aus biologischer Sicht kann man verschiedene Einwände dagegen bringen, für  $b(t)$  beliebige reelle Zahlen zuzulassen:

- Wir haben  $b(t)$  als Anzahl der Bakterien definiert. Eine Anzahl ist aber immer ganzzahlig.
- Es ergibt sicherlich keinen Sinn, von einer negativen Anzahl von Bakterien zu sprechen. Also sollte man zumindest nur Zahlen  $\geq 0$  für  $b(t)$  zulassen (oder sogar nur Zahlen  $> 0$ , wenn man sich nicht für die Situation interessiert, in der es gar keine Bakterien in der Schale gibt).

Der erstgenannte Einwand ist eher theoretischer Natur und spielt in Wirklichkeit keine große Rolle: Die Anzahl der Bakterien in einer Bakterienkultur ist derart hoch, dass es keinen großen Unterschied macht, ob man für ihre Beschreibung wirklich nur natürliche Zahlen oder auch reelle Zahlen verwendet. Zudem ist es aus mathematischer Sicht oft praktischer mit reellen Zahlen zu arbeiten.

Der zweitgenannte Einwand hingegen ist ganz wesentlich. Wir werden nachher noch besprechen, welche Möglichkeiten man hat um diese Einschränkung mathematisch abzubilden.

- (b) Sprechen wir wieder über die Position  $x(t)$  unseres Satelliten im Erdorbit. Damit wir seine Position überhaupt interpretieren können, müssen wir uns auf ein Koordinatensystem einigen; der Einfachheit halber wollen wir hier annehmen, dass der Ursprung dieses Koordinatensystem der Erdmittelpunkt ist (die Wahl der Koordinatenachsen spielt in diesem Beispiel keine große Rolle, nur der Ursprung ist im folgenden wichtig).

Nun ist es nicht besonders sinnvoll, beliebige Punkte aus  $\mathbb{R}^3$  für  $x(t)$  zuzulassen – denn wenn der Satellit näher als den Erdradius an den Mittelpunkt des Koordinatensystems gelangt, kollidiert er mit der Erde. Das wäre nicht nur äußerst ärgerlich (und gefährlich), sondern würde auch dafür sorgen, dass unser Modell für die Bewegung des Satelliten ab diesem Zeitpunkt nicht mehr realistisch wäre.<sup>1</sup>

Wenn wir die Erde als Kugel mit Radius  $R$  annähern<sup>2</sup>, dann ist es also sinnvoll zu fordern, dass die Euklidische Norm von  $x$  immer größer als  $R$  ist.

- (c) Auch bei der C14-Konzentration  $k(t)$ , die wir mehrmals angesprochen haben, ist es sinnvoll, zu fordern, dass  $k(t) \geq 0$  (oder sogar  $k(t) > 0$ ) gilt.
- (d) Zuletzt betrachten wir noch ein rein mathematisches Argument. Sehen wir uns (willkürlicher Weise) die Differentialgleichung

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

an, wobei die Funktion  $x$  Werte in  $\mathbb{R}$  annehmen soll. Hier kommen schon aus mathematischer Sicht nur Funktionen  $x$  als Lösung in Frage, die nicht den Wert 1 annehmen – denn wäre  $x(t)$  zu einem Zeitpunkt  $t$  gleich 1, so wäre die rechte Seite der Gleichung gar nicht definiert.

Um die Probleme, die in den vorangehenden Beispielen diskutiert wurden, zu lösen, gibt es – zumindest für die Beispiele (a)–(c) – zwei prinzipielle Möglichkeiten:

---

<sup>1</sup>Übrigens wird unser Modell für die Satellitenbewegung schon vor einer Kollision mit der Erde unrealistisch: Wenn der Satellit zu weit in die Atmosphäre eintritt, treten Reibungskräfte auf, die nicht nur von der Position des Satelliten, sondern auch von seiner Geschwindigkeit abhängen. Diese Reibungskräfte hatten wir in unserer Differentialgleichung nicht berücksichtigt, und sie führen zudem zu Hitzeentwicklung, die den Satelliten unter Umständen zerstören kann, bevor der die Erdoberfläche erreicht: Der Satellit “verglüht” in diesem Fall.

<sup>2</sup> $R$  beträgt etwas weniger als 6400 Kilometer. Etwas genauer wäre es übrigens, die Erdform als Rotationsellipsoiden anzunähern. Wie in der vorangehenden Fußnote erläutert, wäre es außerdem klug, nicht den Erdradius selbst, sondern die Summe aus Erdradius und der Höhe einiger Atmosphärenschichten zu betrachten.

- *Möglichkeit 1:* Man definiert sich eine feste Menge  $\Omega$  und definiert von Anfang an, dass eine Funktion  $x$  nur dann als Lösung der Differentialgleichung anerkannt wird, wenn all ihre Werte in  $\Omega$  liegen.

Im Fall der Bakterienpopulation könnte man  $\Omega = [0, \infty)$  wählen (oder  $\Omega = (0, \infty)$ ), im Fall des Satelliten könnte man  $\Omega = \{y \in \mathbb{R}^3 : \|y\| > R\}$  wählen (wobei  $\|\cdot\|$  die Euklidische Norm auf  $\mathbb{R}^3$  bezeichnet), und im Falle der Kohlenstoffkonzentration könnte man wiederum  $\Omega = [0, \infty)$  (oder  $\Omega = (0, \infty)$ ) wählen.

Auch das Problem aus Beispiel 2.2.1(d) kann man auf diese Weise lösen: Man definiert dann einfach  $\Omega = \mathbb{R} \setminus \{1\}$ .

- *Möglichkeit 2:* Man lässt zunächst – hypothetisch – auch Lösungen zu, die Werte annehmen, welche aus Anwendungssicht keinen Sinn ergeben. Anschließend untersucht man dann, ob die Lösungen, die man berechnet oder deren Existenz man bewiesen hat, tatsächlich nur Werte im gewünschten Bereich annimmt.

**Bemerkungen 2.2.2.** (a) Die Möglichkeit 1 wird durch die gesamte Theorie, die wir von Kapitel 3 an entwickeln werden, unterstützt. Allerdings gibt es eine wesentliche Einschränkung: Viele Teile der Theorie basieren darauf, dass die Menge  $\Omega$  offen ist, wodurch man zum Beispiel die Wahl  $\Omega = [0, \infty)$  in Möglichkeit 1 nicht einfach treffen kann.

- (b) Die Möglichkeit 2 hängt stark mit der sogenannten *Invarianz* von Mengen zusammen. Was dies bedeutet, werden wir in Abschnitt 7.3 besprechen.
- (c) In konkreten Situationen kombiniert man häufig die Möglichkeiten 1 und 2: Man lässt nur Lösungen zu, die Werte in einer Teilmenge  $\Omega$  von  $\mathbb{R}$  (oder  $\mathbb{R}^d$ ) annehmen; aber die Menge  $\Omega$ , die man zulässt, ist trotzdem zunächst oft größer als die Menge der Werte, die aus Anwendungssicht eigentlich sinnvoll wären. Nachdem man eine Lösung ausgerechnet hat (oder zumindest die Existenz einer Lösung gezeigt hat), untersucht man dann, ob die Werte der Lösung tatsächlich in dem (kleineren) Bereich liegen, für den man sich interessiert.

Nun kommen wir noch kurz auf die zweite Frage zurück, die wir zu Beginn des Abschnitts 2.2 angesprochen hatten: *Soll (und kann) man immer erwarten, dass die Lösungen einer Differentialgleichung für alle Zeiten definiert sind?*

Im Wesentlichen lautet die Antwort: *Nein*. Hier sind zwei unterschiedliche Gründe für diese Antwort:

- Es kann zum Beispiel passieren, dass eine Lösung zu einem bestimmten Zeitpunkt den Bereich  $\Omega$  verlässt, den wir als zulässigen Bereich für die Werte unserer Lösungen festgelegt hatten. Im konkreten Beispiel des Satelliten kann dies zum Beispiel bedeuten: Der Satellit könnte nach endlicher Zeit auf die Erde stürzen.<sup>3</sup> Zu diesem Zeitpunkt hätte seine Position  $x(t)$  den Bereich

$$\Omega = \{y \in \mathbb{R}^3 : \|y\| > R\}$$

verlassen, und wir hätten ab diesem Zeitpunkt keine Lösung unserer Differentialgleichung mehr (da wir von einer Lösung gefordert hatten, dass sie Werte in  $\Omega$  annimmt).

- Wir werden in einige Beispielen das Phänomen beobachten, dass die Lösungen mancher Differentialgleichungen in endlicher Zeit gegen  $\infty$  streben. Solche eine Lösung “explodiert” in gewissem Sinne, und ab diesem Zeitpunkt kann man nicht mehr erwarten, dass die Lösung weiter existiert.

Kurzum: Wir müssen damit rechnen, dass Lösungen von Differentialgleichungen nicht für alle Zeiten existieren, sondern nur für ein bestimmtes Zeitintervall. Dem werden wir in Kapitel 3 Rechnung tragen, wenn wir den Lösungsbegriff für eine Differentialgleichung genau definieren.

## 2.3 Zustandsraum und Vektorfelder

**Zustände** In den Beispielen in Kapitel 1 hatten wir Differentialgleichungen formuliert und dabei implizit angenommen, dass wir eine Lösung  $x : \mathbb{R} \rightarrow \mathbb{R}^d$  suchen (wobei  $d$  in den konkreten Beispielen gleich 1 oder 3 war). Soeben haben wir jedoch in Abschnitt 2.2 gesehen, dass wir tatsächlich etwas andere Anforderungen an eine Lösung haben:

Eine Lösung ist womöglich nicht auf ganz  $\mathbb{R}$  definiert, sondern nur auf einem Teilintervall von  $\mathbb{R}$  und sie bildet nicht unbedingt nach  $\mathbb{R}^d$  ab, sondern nur in eine Teilmenge  $\Omega \subseteq \mathbb{R}^d$ . Für diese Menge  $\Omega$  führen wir einen eigenen Begriff ein:

**Definition 2.3.1 (Zustandsraum und Zustände).** Die Menge  $\Omega \subseteq \mathbb{R}^d$  der Werte (oder Punkte), die die Lösung einer Differentialgleichung annehmen darf, nennen wir den *Zustandsraum*, und die Elemente von  $\Omega$  bezeichnen wir manchmal auch als *Zustände*.

---

<sup>3</sup>Wenn wir, wie in den vorausgehenden Fußnoten angedeutet, atmosphärische Effekte berücksichtigen, würden wir an dieser Stelle stattdessen sagen: Der Satellit kann in die Atmosphäre eintreten und dort verglühen, oder zumindest aufgrund von Reibungskräften unser Modell ungültig machen.

**Ein einheitliche Form für Differentialgleichungen erster Ordnung** Betrachten wir noch einmal die Differentialgleichungen

$$\dot{b}(t) = cb(t) \quad \text{und} \quad -\dot{k}(t) = ck(t)$$

aus den Beispielen 1.1.2(a) und (c). Wir sagen, diese sind von *erster Ordnung*, weil in ihnen neben der gesuchten Funktion nur deren erste Ableitung vorkommt (die Differentialgleichung für die Satellitenposition  $x(t)$  ist hingegen von *zweiter Ordnung*; über die Veranschaulichung von Differentialgleichungen zweiter Ordnung sprechen wir später im Laufe der Vorlesung noch).

Wir einigen uns vorerst für beide Gleichungen auf den Zustandsraum  $\Omega := (0, \infty)$  – d.h. wir wollen nur solche Funktionen  $b$  bzw.  $k$  als Lösung zulassen, die von einem reellen Intervall nach  $\Omega$  abbilden. (Bitte beachten Sie aber: Das bedeutet nicht automatisch, dass auch die Ableitungen  $\dot{b}$  oder  $\dot{k}$  aus  $\Omega$  sein müssen!) Um beide Gleichungen noch auf eine etwas ähnlichere Form zu bringen, multiplizieren wir die zweite Gleichung noch mit  $-1$ , denn dann erhalten wir die Gleichungen

$$\dot{b}(t) = cb(t) \quad \text{und} \quad \dot{k}(t) = -ck(t),$$

bei denen links jeweils nur die Ableitung der gesuchten Funktion steht.

Von einem abstrakten Standpunkt aus betrachtet haben beide Differentialgleichungen eine ähnliche Gestalt: Wir können sie schreiben als

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

wobei  $x$  die gesuchte Funktion ist, die von einem reellen Intervall nach  $\Omega$  abbilden soll (in der ersten Gleichung also  $x = b$  und in der zweiten Gleichung  $x = k$ ), und wobei  $f : \Omega \rightarrow \mathbb{R}$  eine gegebene Funktion ist (in der ersten Differentialgleichung ist  $f(y) = cy$  für  $y \in \Omega$  und in der zweiten Differentialgleichung ist  $f(y) = -cy$  für  $y \in \Omega$ ).

Wenn man eine Differentialgleichung in der Form (2.1) betrachtet, dann nennt man die Funktion  $f$  oft die *rechte Seite* der Differentialgleichung – was natürlich etwas unsauber gesprochen ist, denn genau genommen steht ja nicht  $f$ , sondern  $f(x(t))$  auf der rechten Seite.

**Vektorfelder** Sehr häufig werden wir Differentialgleichungen der Form (2.1) betrachten, bei denen die gesuchte Funktion aber nicht reelle Zahlen, sondern Vektoren im  $\mathbb{R}^d$  als Werte annimmt: In diesem Fall legt man eine Menge  $\Omega \subseteq \mathbb{R}^d$  fest, aus der die Werte der gesuchten Lösung kommen sollen (und wie in Abschnitt 2.2 bereits erwähnt betrachtet man meist nur offene Mengen  $\Omega$ ). Außerdem benötigt man eine Funktion  $f : \Omega \rightarrow \mathbb{R}^d$ , und dann ergibt die Gleichung (2.1) auch in diesem mehr-dimensionalen Rahmen Sinn. Solche Funktionen  $f$  besitzen einen speziellen Namen:

**Definition 2.3.2 (Vektorfeld).** Sei  $d \in \mathbb{N}$  und sei  $\Omega \subseteq \mathbb{R}^d$  nichtleer und offen. Eine Funktion  $f : \Omega \rightarrow \mathbb{R}^d$  bezeichnet man als *Vektorfeld* auf  $\Omega$ .

Ein Vektorfeld bildet also vom Zustandsraum in den  $\mathbb{R}^d$  ab. Bitte beachten Sie, dass die Differentialgleichung  $\dot{x}(t) = f(x(t))$  nur dann Sinn ergibt, wenn  $f$  tatsächlich Vektoren der Dimension  $d$  (aus  $\Omega$ ) wieder auf Vektoren der Dimension  $d$  abbildet.

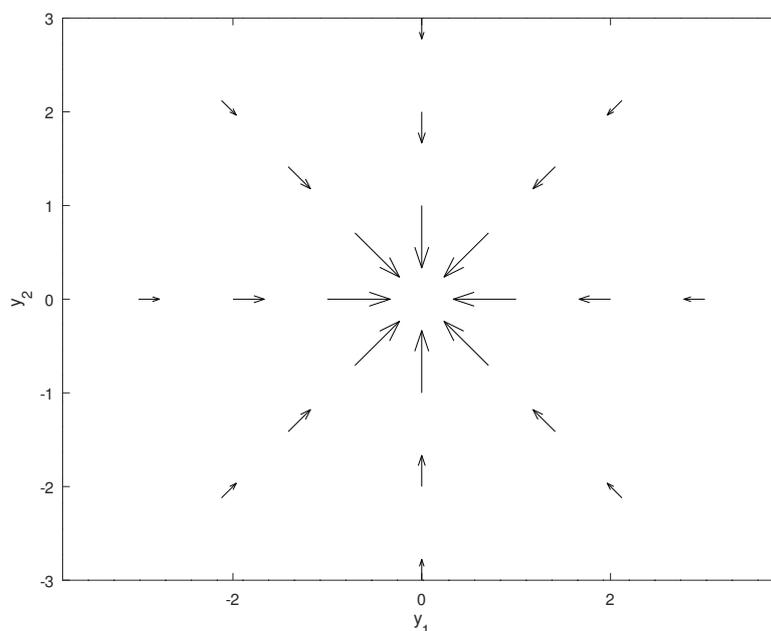
**Graphische Veranschaulichung von Vektorfeldern** Im Fall  $d = 2$  kann man ein Vektorfeld auf einem Gebiet  $\Omega \subseteq \mathbb{R}^2$  sehr gut veranschaulichen, indem man  $\Omega$  zeichnet und an einige Punkte  $y \in \Omega$  den Vektor  $f(y)$  anzeichnet. Wir geben ein konkretes Beispiel an:

**Beispiel 2.3.3.** Sei  $\Omega = \mathbb{R}^2 \setminus \{0\}$  und sei  $f : \Omega \rightarrow \mathbb{R}^2$  durch

$$f(y) = -\frac{2}{3} \frac{y}{\|y\|^2} \quad \text{für alle } y \in \Omega$$

gegeben (wobei  $\|\cdot\|$  die Euklidische Norm bezeichnet). Eine graphische Darstellung dieses Vektorfeldes finden Sie in Abbildung 2.3.3.

Abbildung 2.1: Graphische Darstellung des Vektorfeldes  $f$  aus Beispiel 2.3.3.



Natürlich kann man auch Vektorfelder im Falle  $d = 1$  so darstellen – in diesem Fall besteht das Bild natürlich nicht aus einer Fläche, sondern nur aus einer Linie, auf der Pfeile angebracht sind, die entlang dieser Linie zeigen.

**Aufgabe 2.3.4.** Betrachten Sie im Falle  $d = 1$  und  $\Omega = (0, \infty)$  das Vektorfeld  $f : \Omega \rightarrow \mathbb{R}$ , welches durch

$$f(y) = -\frac{1}{2}y \quad \text{für alle } y \in \Omega$$

gegeben ist. Stellen Sie  $f$  auf zwei verschiedene Weisen graphisch dar:

- So, wie Sie eine Funktion von  $(0, \infty)$  nach  $\mathbb{R}$  üblicherweise skizzieren würden (d.h. mit einer Rechtswert-Achse, an der Sie das Argument auftragen und einer Hochwertachse, an der Sie die Funktionswerte auftragen).
- Als Vektorfeld – d.h. so, wie Sie es in Beispiel 2.3.3 und Abbildung 2.1 gesehen haben, allerdings in einem eindimensionalen Bild.

**Graphische Darstellung der Lösung einer Differentialgleichung** Wenn wir auf einer offenen Menge  $\Omega \subseteq \mathbb{R}^2$  ein Vektorfeld  $f : \Omega \rightarrow \mathbb{R}^2$  gegeben haben und einen Anfangswert  $x_0 \in \Omega$  festhalten, dann gibt es eine sehr anschauliche Möglichkeit, Lösungen des Anfangswertproblems

$$\dot{x}(t) = f(x(t)), \quad x(0) = x_0 \tag{2.2}$$

zu verstehen: Eine Lösung  $x$ , die von einem Intervall nach  $\Omega$  abbildet, beschreibt, wie Sie aus der Analysis 2 wissen, eine Kurve in  $\Omega$ . Die Ableitung  $\dot{x}$  ist gerade die Geschwindigkeit (oder mathematischer ausgedrückt: der Tangentialvektor) der Kurve; sie gibt die Änderungsrate von  $x$  an.

Das Anfangswertproblem (2.2) kann man also folgendermaßen verstehen: Zum Zeitpunkt 0 startet die Kurve  $x$  an der Position  $x_0$ ; die Änderung der Kurve in diesem Moment ist gleich

$$\dot{x}(0) = f(x(0)) = f(x_0),$$

also gleich dem Vektor, den das Vektorfeld  $f$  an der Startposition  $x_0$  vorgibt. Die Ableitung  $\dot{x}(0)$  ist in etwa gleich dem Differenzenquotienten  $\frac{1}{\tau}(x(\tau) - x(0))$ , wenn  $\tau$  eine sehr kleine positive reelle Zahl ist. Anders ausgedrückt gilt also

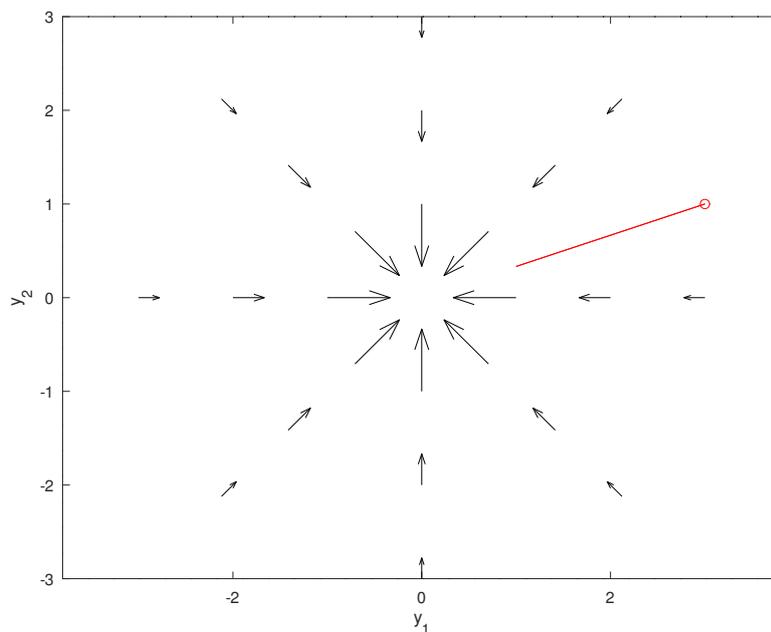
$$\frac{1}{\tau}(x(\tau) - x_0) = f(x_0),$$

und dies ist gleichbedeutend mit der Gleichung

$$x(\tau) = x_0 + \tau f(x_0).$$

## 2. MODELLIERUNG UND GEOMETRISCHE INTUITION

Abbildung 2.2: Eine Kurve (rot), die die Differentialgleichung  $\dot{x}(t) = f(x(t))$  löst; das Vektorfeld  $f$  stammt aus Beispiel 2.3.3. Der Anfangswert  $x_0 = (3, 1)^T$  ist als roter Kreis markiert und die Kurve verläuft Richtung Ursprung.



Um also die Position  $x(\tau)$  nach einer kleinen Zeit  $\tau$  zu erhalten, starten wir an der Stelle  $x_0$  und gehen dann  $\tau$  mal den Vektor  $f(x_0)$  entlang. Dann sind wir an der Stelle  $x(\tau)$ . Jetzt benutzen wir wieder die Differentialgleichung, laut der

$$\dot{x}(\tau) = f(x(\tau))$$

gilt, und wir können das vorangehende Prozedere wiederholen.

Auf diese Weise gibt uns das Vektorfeld  $f$  stets vor, in welche Richtung und wie schnell wir unsere momentane Position ändern müssen, und wir erhalten somit eine Kurve  $x$ , die auf einem von  $f$  vorgegebenen Weg durch  $\Omega$  führt. Für das Vektorfeld  $f$  aus Beispiel 2.3.3 und den Startpunkt  $x_0 = (3, 1)^T$  ist dies in Abbildung 2.2 dargestellt. In diesem einfachen Fall führt die Kurve auf einer Geraden in Richtung Ursprung; dies liegt daran, dass das Vektorfeld  $f$  in diesem Beispiel von jedem Punkt aus zum Ursprung zeigt und die Kurve somit immer weiter zum Ursprung hinbewegt.

Ähnlich können wir natürlich auch für ein ein-dimensionales Vektorfeld den Verlauf der Lösung der zugehörigen Differentialgleichung anhand einer Skizze des Vektorfeldes abschätzen:

**Aufgabe 2.3.5.** Sei  $\Omega = (0, \infty)$ . Betrachten Sie auf  $\Omega$  die logistische Gleichung

$$\dot{b}(t) = ((c - db(t))b(t))$$

aus Beispiel 2.1.3. Diese können wir abstrakt in der Form  $\dot{b}(t) = f(b(t))$  schreiben, wobei  $f : \Omega \rightarrow \mathbb{R}$  durch  $f(y) = (c - dy)y$  für alle  $y \in \Omega$  gegeben ist.

Um konkrete Zahlenwerte zur Verfügung zu haben, wählen wir in dieser Aufgabe  $c = \frac{1}{4}$  und  $d = \frac{1}{12}$ .

- Skizzieren Sie das Vektorfeld  $f$ .
- Geben Sie anhand Ihrer Skizze an, wohin die Lösung der Differentialgleichung verläuft, wenn der Anfangswert zum Zeitpunkt 0 gleich 1 ist, und wohin die Lösung verläuft, wenn der Anfangswert gleich 5 ist.

## 2.4 Ergänzungen

### Modellierung durch eine partiellen Differentialgleichung – ein Beispiel

Falls Sie Interesse haben einmal zu sehen, wie man auch eine partielle Differentialgleichung mit Hilfe von Differenzenquotienten herleiten kann, können Sie sich das folgende Beispiel näher ansehen. Darin geben wir eine “Herleitung” der Wärmeleitungsgleichung aus Beispiel 1.3.1 an.

**Beispiel 2.4.1 (Herleitung der Wärmeleitungsgleichung).** Wie bereits in Beispiel 1.3.1 betrachten wir die Temperatur  $u(t, z)$  eines dünnen Metallstabs der Länge  $L$  zur Zeit  $t \in [0, \infty)$  und an der Stelle  $z \in [0, L]$ ; also ist  $u$  eine Funktion von  $[0, \infty) \times [0, L]$  nach  $\mathbb{R}$ . Wir nehmen nachwievor an, dass der Stab entlang seiner Mantelfläche thermisch isoliert ist.

Wir betrachten nun einen festen Zeitpunkt  $t \in [0, \infty)$  and zwei Positionen  $z_1, z_2$ , die nahe aneinander liegen und die Bedingung  $0 < z_1 < z_2 < L$  erfüllen. Wir bezeichnen mit  $E(t)$  die thermische Energie, die im Stab zwischen den Positionen  $z_1$  und  $z_2$  zum Zeitpunkt  $t$  gespeichert ist. Weil  $z_1$  und  $z_2$  sehr nahe zusammenliegen, ist es sinnvoll davon auszugehen, dass die Temperatur im Bereich zwischen  $z_1$  und  $z_2$  ähnlich groß ist wie an der Stelle  $z_1$ .

Die thermische Energie  $E(t)$  ist proportional zum Produkt aus der Temperatur im Bereich zwischen  $z_1$  und  $z_2$  und der Länge des Intervalls. Also gibt es eine Konstante  $d > 0$  derart, dass

$$E(t) = du(t, z_1)(z_2 - z_1)$$

gilt. Nun betrachten wir die Frage, wie sich die Energie  $E(t)$  ändert, wenn wir einen kleinen Zeitschritt der Länge  $\tau$  weitergehen. Einerseits erhalten wir

dann natürlich

$$E(t + \tau) - E(t) = d(u(t + \tau, z_1) - u(t, z_1))(z_2 - z_1).$$

Andererseits kann eine Änderung der thermischen Energie im Intervall zwischen  $z_1$  und  $z_2$  – aufgrund der thermischen Isolation – nur dadurch zustande kommen, dass Energie durch die beiden Punkte  $z_1$  und  $z_2$  ins Intervall zufließt oder abfließt. In der Physik ist bekannt, dass der Durchfluss der thermischen Energie in einem Punkt direkt proportional zum räumlichen Temperaturabfall oder -anstieg in diesem Punkt ist. Es gibt also eine Konstante  $k > 0$  derart, dass im Zeitintervall der Länge  $\tau$  die thermische Energiedifferenz  $E(t + \tau) - E(t)$  (d.h. die thermische Energie, die in das Intervall  $[z_1, z_2]$  hineinfließt), gleich

$$k\tau \left( -\frac{\partial}{\partial z} u(t, z)|_{z=z_1} + \frac{\partial}{\partial z} u(t, z)|_{z=z_2} \right)$$

ist. Insgesamt ist also

$$d(u(t + \tau, z_1) - u(t, z_1))(z_2 - z_1) = k\tau \left( -\frac{\partial}{\partial z} u(t, z)|_{z=z_1} + \frac{\partial}{\partial z} u(t, z)|_{z=z_2} \right).$$

Wir teilen nun sowohl durch  $\tau$  also auch durch  $z_2 - z_1$ ; dann lassen wir  $\tau$  gegen 0 streben und  $z_2$  gegen  $z_1$ ; somit erhalten wir die partielle Differentialgleichung

$$d \frac{\partial}{\partial t} u(t, z_1) = k \frac{\partial^2}{\partial z^2} u(t, z)|_{z=z_1}.$$

Zuletzt teilen wir noch durch  $d$  und setzen  $c = k/d$ . Wenn wir berücksichtigen, dass obige Argumentation für alle  $z_1$  zwischen 0 und  $L$  funktioniert und wir die Ortsvariable deshalb wieder  $z$  statt  $z_1$  nennen, so gelangen wir also zur Gleichung

$$\frac{\partial}{\partial t} u(t, z) = c \frac{\partial^2}{\partial z^2} u(t, z),$$

und dies ist genau die Wärmeleitungsgleichung aus Beispiel 1.3.1.

Die „Herleitung“ in Beispiel 2.4.1 ist natürlich nur eine Heuristik. In der Literatur finden sich viele unterschiedliche Versionen solch einer Herleitung. Alle führen die Gleichung aber letztlich auf dieselben physikalischen Prinzipien zurück.

## Was sind Differentialgleichungen – noch einmal

### Fragen zum Einstieg.

- (a) Was genau war noch einmal der Unterschied zwischen einer Differentialgleichung und einem Anfangswertproblem? Warum haben wir Anfangswerte überhaupt eingeführt?
- (b) Können Sie sich erinnern, was man unter der Ordnung einer Differentialgleichung versteht?

### 3.1 Begriffsklärung: Zweiter Versuch

In diesem Abschnitt wollen wir nun eine präzise mathematische Definition der Begriffe *Differentialgleichung*, *Anfangswertproblem* und *Lösung* geben.

Wir beginnen zunächst mit dem allgemeinen Fall, in welchem die Differentialgleichung implizit definiert ist. Wir nennen ein Intervall  $J \subseteq \mathbb{R}$  *nicht-trivial*, wenn es mindestens zwei Punkte enthält (voraus dann natürlich folgt, dass  $J$  sogar überabzählbar viele Punkte enthält).

**Definition 3.1.1 (Differentialgleichung, Anfangswertproblem und Lösung: Implizite Form).** Seien  $d, n \in \mathbb{N}$ , sei  $\tilde{G} \subseteq \mathbb{R}^{1+d(n+1)}$  eine nichtleere Menge und sei  $\tilde{f} : \tilde{G} \rightarrow \mathbb{R}^d$  eine Abbildung.

- (i) Eine Gleichung der Form

$$\tilde{f}\left((t, x(t), \dot{x}(t), \ddot{x}(t), \dots, x^{(n)}(t))\right) = 0 \quad (3.1)$$

heißt eine *d-dimensionale gewöhnliche Differentialgleichung*. Falls die Funktionswerte von  $\tilde{f}$  nicht unabhängig von den letzten  $d$  Einträgen des Arguments von  $\tilde{f}$  sind, dann hat die Differentialgleichung die *Ordnung*  $n$ .

- (ii) Sei  $J \subseteq \mathbb{R}$  ein nicht-triviales Intervall. Eine Funktion  $x : J \rightarrow \mathbb{R}^d$  heißt *Lösung* der Differentialgleichung (3.1), wenn  $x$   $n$ -mal differenzierbar ist und die folgende Eigenschaft erfüllt:

Für alle  $t \in J$  gilt  $(t, x(t), \dot{x}(t), \ddot{x}(t), \dots, x^{(n)}(t)) \in \tilde{G}$  und es für alle  $t \in J$  gilt die Gleichung (3.1).

(iii) Sei  $t_0 \in \mathbb{R}$  und seien  $x_0, x_1, x_2, \dots, x_{n-1} \in \mathbb{R}^d$ . Dann heißt

$$\begin{cases} \tilde{f}(t, x(t), \dot{x}(t), \ddot{x}(t), \dots, x^{(n)}(t)) = 0, \\ x(t_0) = x_0, \quad \dot{x}(t_0) = x_1, \quad \dots, \quad x^{(n-1)}(t_0) = x_{n-1} \end{cases} \quad (3.2)$$

ein *Anfangswertproblem* zur Differentialgleichung (3.1).

(iv) Sei  $J \subseteq \mathbb{R}$  ein nicht-triviales Intervall. Eine Funktion  $x : J \rightarrow \mathbb{R}^d$  heißt *Lösung* des Anfangswertproblems (3.2), wenn  $x$  eine Lösung der Differentialgleichung in der ersten Zeile von (3.2) ist, wenn  $t_0 \in J$  gilt und wenn alle Gleichungen in der zweiten Zeile von (3.2) erfüllt sind.

Es ist lehrreich, die obige Definition mit den vorläufigen Definitionen 1.1.4 und 1.4.3 zu vergleichen.

Wir betrachten einige Beispiele um zu sehen, wie man konkrete Differentialgleichungen in Definition 3.1.1 einordnen kann. Hierzu sehen wir uns wieder einmal die Bakterienpopulation und den Satelliten an:

**Beispiele 3.1.2.** (a) Wir betrachten wieder die Bakterienpopulation  $b(t)$ . Wir wollen die Differentialgleichung

$$\dot{b}(t) = cb(t)$$

(mit der Konstanten  $c > 0$ ) in der allgemeinen Form aus Definition 3.1.1(i) darstellen.

Die Anzahl der Bakterien ist eine ein-dimensionale Größe, d.h. es ist  $d = 1$ . Außerdem kommen in der Differentialgleichung nur Ableitungen bis zur Ordnung 1 vor, also ist  $n = 1$ .

Als Zustandsraum  $\Omega$  wählen wir wie in Abschnitt 2.2 vorgeschlagen die Menge  $\Omega = (0, \infty)$ . Für die Zeit  $t$  wählen wir im voraus keine speziellen Einschränkungen. Also können wir

$$\tilde{G} = \mathbb{R} \times (0, \infty) \times \mathbb{R} = \{(t, y_0, y_1) \in \mathbb{R}^3 : y_0 > 0\}$$

wählen, und wir setzen

$$\tilde{f} : \tilde{G} \ni (t, y_0, y_1) \mapsto y_1 - cy_0 \in \mathbb{R}.$$

Dann ist die Differentialgleichung  $\dot{b}(t) = cb(t)$  äquivalent zu  $f(t, b(t), \dot{b}(t)) = 0$ .

(b) Lassen Sie uns noch einmal die Bakterienpopulation in der Petrischale betrachten – aber dieses mal gehen wir wie in Beispiel 2.1.2 davon aus,

dass die Temperatur zeitabhängig ist und die Zahl  $c$  somit von  $t$  abhängt. Wir wollen die Differentialgleichung

$$\dot{b}(t) = c(t)b(t)$$

aus Beispiel 2.1.2 in der Form aus Definition 3.1.1(i) schreiben.

Wie im vorangehenden Beispiel wählen wir dazu

$$\tilde{G} = \mathbb{R} \times (0, \infty) \times \mathbb{R} = \{(t, y_0, y_1) \in \mathbb{R}^3 : y_0 > 0\},$$

aber dieses mal definieren wir  $\tilde{f}$  durch

$$\tilde{f} : \tilde{G} \ni (t, y_0, y_1) \mapsto y_1 - c(t)y_0 \in \mathbb{R}.$$

Dann ist die Differentialgleichung  $\dot{b}(t) = c(t)b(t)$  wie gewünscht äquivalent zu  $f(t, b(t), \dot{b}(t)) = 0$ .

- (c) Sehen wir uns wieder die Position  $x(t)$  unseres Satelliten an, welche die Differentialgleichung  $m\ddot{x}(t) = F(x(t))$  erfüllt. Die Dimension ist nun  $d = 3$ , die Ordnung ist  $n = 2$ , und wir wollen nur Positionen in der Menge

$$\Omega = \{y \in \mathbb{R}^3 : \|y\| > R\}$$

zu lassen, wobei  $R$  den Erdradius bezeichnet. Wir wählen

$$\tilde{G} = \mathbb{R} \times \Omega \times \mathbb{R}^3 \times \mathbb{R}^3 = \{(t, y_0, y_1, y_2) \in \mathbb{R}^{10} : t \in \mathbb{R}, y_0 \in \Omega, y_1, y_2 \in \mathbb{R}^3\}$$

und

$$\tilde{f} : \tilde{G} \ni (t, y_0, y_1, y_2) \mapsto my_2 - F(y_0) \in \mathbb{R}^3.$$

Dann ist die Differentialgleichung  $m\ddot{x}(t) = F(x(t))$  äquivalent zur Differentialgleichung  $\tilde{f}(t, x(t), \dot{x}(t), \ddot{x}(t)) = 0$  aus Definition 3.1.1(i).

In Abschnitt 2.3 hatten wir besprochen, wie man Differentialgleichungen erster Ordnung für eine gesuchte Funktion  $x$  in manchen Fällen auf eine Form bringen kann, in der auf der linken Seite der Gleichung nur die erste Ableitung von  $x$  steht. Dasgleiche kann man auch versuchen, wenn man eine Differentialgleichung höherer Ordnung betrachtet; in diesem Fall versucht man, die Differentialgleichung so umzuschreiben, dass auf der linken Seite nur die höchste auftretende Ableitung von  $x$  steht.

Wenn dies gelingt, dann sagt man, dann nennen wir die so erhaltene Differentialgleichung *explizit*. Wir halten dies noch einmal in der folgenden Definition fest.

**Definition 3.1.3 (Differentialgleichung, Anfangswertproblem und Lösung: Explizite Form).** Seien  $d, n \in \mathbb{N}$  und sei  $G \subseteq \mathbb{R}^{1+dn}$  eine nichtleere Menge und sei  $f : G \rightarrow \mathbb{R}^d$  eine Abbildung.

- (i) Eine Gleichung der Form

$$x^{(n)}(t) = f\left((t, x(t), \dot{x}(t), \ddot{x}(t), \dots, x^{(n-1)}(t))\right) \quad (3.3)$$

heißt eine *explizite  $d$ -dimensionale gewöhnliche Differentialgleichung*; sie hat die *Ordnung  $n$* .

- (ii) Sei  $J \subseteq \mathbb{R}$  ein nicht-triviales Intervall. Eine Funktion  $x : J \rightarrow \mathbb{R}^d$  heißt *Lösung* der Differentialgleichung (3.3), wenn  $x$   $n$ -mal differenzierbar ist und die folgende Eigenschaft erfüllt:

Für alle  $t \in J$  ist  $(t, x(t), \dot{x}(t), \ddot{x}(t), \dots, x^{(n-1)}(t)) \in G$  und für alle  $t \in J$  gilt die Gleichung (3.3).

- (iii) Sei  $t_0 \in \mathbb{R}$  und seien  $x_0, x_1, x_2, \dots, x_{n-1} \in \mathbb{R}^d$ . Dann heißt

$$\begin{cases} x^{(n)}(t) = f\left((t, x(t), \dot{x}(t), \ddot{x}(t), \dots, x^{(n-1)}(t))\right), \\ x(t_0) = x_0, \quad \dot{x}(t_0) = x_1, \quad \dots, \quad x^{(n-1)}(t_0) = x_{n-1} \end{cases} \quad (3.4)$$

ein *Anfangswertproblem* zur Differentialgleichung (3.3).

- (iv) Sei  $J \subseteq \mathbb{R}$  ein nicht-triviales Intervall. Eine Funktion  $x : J \rightarrow \mathbb{R}^d$  heißt *Lösung* des Anfangswertproblems (3.4), wenn  $x$  eine Lösung der Differentialgleichung in der ersten Zeile von (3.4) ist, wenn  $t_0 \in J$  gilt und wenn alle Gleichungen in der zweiten Zeile von (3.4) erfüllt sind.

**Aufgabe 3.1.4.** Sehen Sie sich die Definitionen 3.1.1 und 3.1.3 noch einmal im Detail an.

- (a) Machen Sie sich dabei alle Unterschiede zwischen den beiden Definitionen klar.
- (b) Wie können Sie eine explizite Differentialgleichung in die implizite Form von Definition 3.1.1(i) bringen?

Alle Differentialgleichungen aus den Beispielen 3.1.2 lassen sich in explizite Form bringen. Es ist wichtig, dass Sie dies zur Übung konkret tun; dies ist Teil des zweiten Übungsblattes.

## 3.2 Differentialgleichungen erster und höherer Ordnung

Die folgende Proposition ist äußerst einfach, aber essentiell für viele Teile der Vorlesung. Sie zeigt, dass es in vielen Fällen genügt, Differentialgleichungen erster Ordnung zu betrachten. Der Einfachheit halber formulieren wir das Resultat nur für explizite Differentialgleichungen (bzw. Anfangswertprobleme), aber Sie können sich leicht überlegen, dass dasselbe auch für implizite Differentialgleichungen gilt.

**Proposition 3.2.1.** *Seien  $d, n \in \mathbb{N}$ , sei  $G \subseteq \mathbb{R}^{1+dn}$  eine nichtleere Menge und sei  $f : G \rightarrow \mathbb{R}^d$  eine Abbildung; sei  $t_0 \in \mathbb{R}$  und  $x_0, \dots, x_{n-1} \in \mathbb{R}^d$ . Wir definieren eine Funktion  $h : G \rightarrow \mathbb{R}^{dn}$  durch*

$$h(t, y_0, \dots, y_{n-1}) = (y_1, \dots, y_{n-1}, f(t, y_0, \dots, y_{n-1})) \quad \text{für } (t, y_0, \dots, y_{n-1}) \in G$$

(wobei  $t \in \mathbb{R}$  ist und  $y_0, \dots, y_{n-1} \in \mathbb{R}^d$  sind).

(a) Wenn  $x : J \rightarrow \mathbb{R}^d$  das Anfangswertproblem

$$\begin{cases} x^{(n)}(t) = f(t, x(t), \dot{x}(t), \ddot{x}(t), \dots, x^{(n-1)}(t)), \\ x(t_0) = x_0, \quad \dot{x}(t_0) = x_1, \quad \dots, \quad x^{(n-1)}(t_0) = x_{n-1} \end{cases} \quad (3.5)$$

löst, dann löst  $z := (x, \dot{x}, \dots, x^{(n-1)}) : J \rightarrow \mathbb{R}^{dn}$  das Anfangswertproblem

$$\begin{cases} \dot{z}(t) = h(t, z(t)), \\ z(t_0) = (x_0, \dots, x_{n-1}). \end{cases} \quad (3.6)$$

(b) Umgekehrt: Wenn  $z : J \rightarrow \mathbb{R}^{dn}$  das Anfangswertproblem (3.6) löst, dann löst die Funktion  $x : J \rightarrow \mathbb{R}^d$ , die aus den ersten  $d$  Komponenten von  $z$  besteht, das Anfangswertproblem (3.5).

*Beweis.* Der Beweis wird Teil des zweiten Übungsblattes sein. □

## 3.3 Differentialgleichungen zweiter Ordnung und Phasenraum

Speziell für Differentialgleichungen zweiter Ordnung gibt es den Begriff *Phasenraum*, der folgendermaßen definiert ist.

**Definition 3.3.1 (Phasenraum).** Sei  $d \in \mathbb{N}$ , sei  $G \subseteq \mathbb{R}^{1+2d}$  und  $f : G \rightarrow \mathbb{R}^d$ . Wenn man  $G$  in der Form  $G = I \times W$  für ein reelles Intervall  $I$  und eine Menge  $W \subseteq \mathbb{R}^{2d}$  schreiben kann, dann heißt die Menge  $W$  der *Phasenraum* der Differentialgleichung

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

Anders gesprochen: Wenn die Frage, welche Paare  $(x(t), \dot{x}(t))$  (bestehend aus Positionen und Geschwindigkeiten) für unsere Differentialgleichung zulässig sind, nicht von der Zeit  $t$  abhängt, dann bezeichnen wir die Menge dieser Paare als *Phasenraum* der Differentialgleichung.

**Beispiel 3.3.2 (Teilchen und Satellit).** (a) In der klassischen Mechanik betrachtet man oft ein Teilchen der Masse  $m$ , das sich in einer Menge  $\Omega \subseteq \mathbb{R}^d$  bewegt. Die Position  $x(t)$  des Teilchens zum Zeitpunkt  $t$  erfüllt die Differentialgleichung

$$\ddot{x}(t) = \frac{1}{m} \hat{F}(t, x(t), \dot{x}(t)), \quad (3.7)$$

wobei  $\hat{F} : \mathbb{R} \times \Omega \times \mathbb{R}^d$  die auf das Teilchen wirkende Kraft ist, welche im Allgemeinen von der Zeit, der Position und der Geschwindigkeit des Teilchens abhängen kann (wir nennen die Kraft in diesem Beispiel  $\hat{F}$  statt  $F$ , damit die Notation mit dem nachfolgenden Beispiel konsistent ist).

In der Notation von Beispiel 3.3.1 ist hier also  $G = \mathbb{R} \times \Omega \times \mathbb{R}^3$ , und  $\Omega \times \mathbb{R}^3$  ist der *Phasenraum*.

(b) Betrachten wir nun wieder einmal unseren Satelliten, der die Erde umkreist. Die möglichen Positionen des Satelliten werden, wie bereits mehrmals besprochen, durch die Menge

$$\Omega = \{y \in \mathbb{R}^3 : \|y\| > R\}$$

beschrieben, wobei  $R$  den Erdradius bezeichnet. Die Kraft  $F$ , die auf den Satelliten wirkt, hängt nur von seiner Position ab, nicht aber von seiner Zeit oder seiner Geschwindigkeit<sup>1</sup>, d.h.  $F$  ist eine Funktion von  $\Omega$  nach  $\mathbb{R}^3$ , und die Position des Satelliten erfüllt, wie Sie bereits wissen, die Differentialgleichung

$$\ddot{x}(t) = F(x(t)).$$

Wenn wir möchten, können wir diese Differentialgleichung in die Form aus Beispiel (a) bringen: Wir betrachten wieder den Phasenraum  $\Omega \times \mathbb{R}^3$ , setzen  $G = \mathbb{R} \times \Omega \times \mathbb{R}^3$  und definieren einfach

$$\hat{F} : G \ni (t, y_0, y_1) \mapsto F(y_0) \in \mathbb{R}^3.$$

<sup>1</sup>Sie wissen bereits, dass das nicht ganz richtig ist: Wenn der Satellit ein Stück weit in die oberen Atmosphärenschichten eindringt, wirkt auch Reibungskraft auf ihn, und diese hängt unter anderem von seiner Geschwindigkeit ab.

Außerdem wird die Gesamtkraft, die auf den Satelliten wirkt, im Allgemeinen sogar von der Zeit abhängen: Neben der Erdanziehung wirkt auf den Satelliten nämlich beispielsweise die Gravitation des Mondes – und diese ist aus Sicht des Satelliten zeitabhängig, da die Position des Mondes sich mit der Zeit ändert.

Dann können wir die Differentialgleichung  $\ddot{x}(t) = F(x(t))$  in der Form  $\ddot{x}(t) = \hat{F}(t, x(t), \dot{x}(t))$  schreiben.

**Aufgabe 3.3.3.** Benutzen Sie die Technik aus Abschnitt 3.2 um die Differentialgleichung (3.7) in eine Gleichung erster Ordnung umzuschreiben; diese Differentialgleichung erster Ordnung bezeichnen wir in dieser Aufgabe mit (\*).

Wenn  $x$  eine Lösung der Differentialgleichung (3.7) ist, dann beschreibt  $x(t)$  zum Zeitpunkt  $t$  den Ort des Teilchens. Sei nun  $z$  eine Lösung der Gleichung (\*). Was beschreibt  $z(t)$ ?

### 3.4 Autonome und nicht-autonome Differentialgleichungen

In Beispiel 2.1.2 hatten wir zum ersten Mal angesprochen, wann eine Differentialgleichung *autonom* oder *nicht-autonom* heißt. Mit Hilfe unserer formalen Definitionen in diesem Kapitel können wir dies nun präzisieren. Wir begnügen uns an dieser Stelle damit, den Fall expliziter Differentialgleichungen zu betrachten.

**Definition 3.4.1 (Autonome Differentialgleichung).** Die Differentialgleichung

$$x^{(n)}(t) = f\left(t, x(t), \dot{x}(t), \dots, x^{(n-1)}(t)\right)$$

in Definition 3.1.3 heißt *autonom*, wenn man  $G$  in der Form  $G = I \times H$  für eine Menge  $H \subseteq \mathbb{R}^{dn}$  und ein Intervall  $I \subseteq \mathbb{R}$  schreiben kann, und wenn  $f$  nicht von der Zeitkomponente seines Arguments abhängt – d.h. wenn es eine Funktion  $h : H \rightarrow \mathbb{R}^d$  gibt derart, dass

$$f(t, y_0, y_1, \dots, y_{n-1}) = h(y_0, y_1, \dots, y_{n-1})$$

für alle  $(t, y_0, y_1, \dots, y_{n-1}) \in G = I \times H$  gilt.<sup>2</sup>

Anders ausgedrückt: Eine explizite Differentialgleichung heißt autonom, wenn die folgenden beiden Objekte nicht von der Zeit abhängen:

- Die Menge, in der eine Lösung und ihre Ableitungen liegen dürfen.
- Die rechte Seite der Differentialgleichung.

<sup>2</sup>Man beachte, dass es für eine autonome Differentialgleichung eigentlich nicht viel Sinn ergibt, die zulässigen Zeiten von vorne herein auf ein Teilintervall  $I$  von  $\mathbb{R}$  einzuschränken; wenn man eine autonome Differentialgleichung vorliegen hat, spricht nichts dagegen,  $I$  durch die ganze reelle Achse  $\mathbb{R}$  zu ersetzen – wodurch die Lösungen dann natürlich eventuell auf einer größeren Menge von Zeiten definiert sein können.

Eine autonome explizite Differentialgleichung  $n$ -ter Ordnung ist also von der Form

$$x^{(n)}(t) = h(x(t), \dot{x}(t), \dots, x^{(n-1)}(t))$$

für eine Menge  $H \subseteq \mathbb{R}^{dn}$  und eine Abbildung  $h : H \rightarrow \mathbb{R}^d$ .

## 3.5 Ergänzungen

### Differential-algebraische Gleichungen

In dieser Vorlesung beschäftigen wir uns vor allem mit expliziten Differentialgleichungen. Eine Klasse von impliziten Differentialgleichungen sind die sogenannten *Differential-algebraischen Gleichungen*. Diese spielen zum Beispiel eine wichtige Rolle in der Modellierung von mechanischen oder elektrischen Systemen, die Nebenbedingungen unterliegen.

Wir werden in dieser Vorlesung aber nicht auf differential-algebraische Gleichungen eingehen.

# Existenz und Eindeutigkeit von Lösungen

## Fragen zum Einstieg.

- (a) Versuchen Sie zwei verschiedene Lösungen  $x : \mathbb{R} \rightarrow \mathbb{R}$  des folgenden Anfangswertproblems zu finden:

$$\begin{cases} \dot{x}(t) = |x(t)|^{1/2} & \text{für } t \in \mathbb{R}, \\ x(0) = 0. \end{cases}$$

- (b) Was bedeutet der Begriff *metrischer Raum*? Wann heißt ein metrischer Raum *vollständig*?
- (c) Suchen Sie im Internet nach den Begriffen *Laplacescher Dämon*<sup>1</sup> und *Determinismus*.

Wir benötigen in diesem Kapitel viele grundlegende Eigenschaften von metrischen und normierten Räumen, die Sie aus der Analysis 2 kennen. Der Vollständigkeit halber finden Sie einige wichtige dieser Eigenschaften in Anhang A in diesem Manuskript zusammengefasst. Außerdem spielen die Fixpunktsätze von Banach und von Schauder im Folgenden eine wichtige Rolle; einen Überblick über diese Fixpunktsätze sowie über die generelle Philosophie bei der Verwendung von Fixpunktsätzen finden Sie in Anhang B.

## 4.1 Lokale Existenz und globale Eindeutigkeit: Der Satz von Picard–Lindelöf

Wir wollen nun beweisen, dass ein Anfangswertproblem unter recht allgemeinen Bedingungen eine eindeutige Lösung besitzt. Aufgrund von Proposition 3.2.1 genügt es hierfür, Gleichungen erster Ordnung zu untersuchen. Die wichtigsten Aussagen in diesem Abschnitt sind ein globaler Eindeutigkeitssatz in Korollar 4.1.11 und ein lokaler Existenzsatz in Korollar 4.1.12; beide Resultate sind einfache Konsequenzen des lokalen Existenz- und Eindeutigkeitssatzes von Picard–Lindelöf, den wir in Satz 4.1.9 beweisen.

Der Satz von Picard–Lindelöf basiert auf der Idee, die Existenz und Eindeutigkeit der Lösung eines Anfangswertproblems mit Hilfe eines Fixpunktsatzes

<sup>1</sup>Benannt nach Pierre-Simon Laplace (1749 – 1827), französischer Mathematiker, Astronom, Physiker und Politiker.

aus der Funktionalanalysis beweisen. Hierfür ist es sehr nützlich, ein Anfangswertproblem zunächst in eine Integralgleichung umzuschreiben; diese Umformulierung ist der Inhalt der folgenden Proposition.

**Proposition 4.1.1.** Sei  $d \in \mathbb{N}$ , sei  $G \subseteq \mathbb{R}^{1+d}$  eine nichtleere Menge und sei  $f : G \rightarrow \mathbb{R}^d$  eine stetige Abbildung; sei zudem  $t_0 \in \mathbb{R}$  und  $x_0 \in \mathbb{R}^d$ .

Wenn  $J \subseteq \mathbb{R}$  ein nicht-triviales Intervall mit  $t_0 \in J$  ist und  $x : J \rightarrow \mathbb{R}^d$  eine Funktion mit der Eigenschaft  $(t, x(t)) \in G$  für alle  $t \in J$  ist, dann sind die folgenden Aussagen äquivalent:

(i) Die Funktion  $x$  löst das Anfangswertproblem

$$\begin{cases} \dot{x}(t) = f(t, x(t)), \\ x(t_0) = x_0. \end{cases}$$

(ii) Die Funktion  $x$  ist stetig und es gilt

$$x(t) = x_0 + \int_{t_0}^t f(s, x(s)) \, ds \quad (4.1)$$

für alle  $t \in J$ .

Man beachte, dass das Integral in (4.1) wohldefiniert ist (als Riemann-Integral), denn aus der Stetigkeit von  $x$  und  $f$  folgt, dass der Integrand stetig ist.

*Beweis von Proposition 4.1.1.* „(i)  $\Rightarrow$  (ii)“ Es gelte (i). Weil  $x$  eine Lösung ist, ist  $x$  nach Definition differenzierbar und somit stetig. Weil  $f$  ebenfalls stetig ist, folgt aus der Differentialgleichung, die  $x$  löst, dass  $\dot{x}$  ebenfalls stetig ist. Also ist  $x$  sogar stetig differenzierbar; somit können wir den Hauptsatz der Differential- und Integralrechnung anwenden. Mit dessen Hilfe erhalten für jedes  $t \in J$

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

dies zeigt (ii).

„(ii)  $\Rightarrow$  (i)“ Es gelte nun (ii). Dann ist  $x(t_0) = x_0 + 0 = x_0$ . Aus dem Hauptsatz der Differential- und Integralrechnung folgt, dass  $x$  stetig differenzierbar ist und dass  $\dot{x}(t) = f(t, x(t))$  für alle  $t \in J$  gilt. Dies zeigt (i).  $\square$

Wir wollen nun erläutern, wie man die Integralgleichung (4.1) als Fixpunktgleichung auffassen kann. Zu diesem Zweck nehmen wir (nur der besseren Anschaulichkeit halber) einmal kurz an, dass einfach  $G = \mathbb{R}^{1+d}$  gilt.

Betrachten wir ein kompaktes, nicht-triviales Intervall  $J \subseteq \mathbb{R}$ , das  $t_0$  enthält, und sehen wir uns den Funktionenraum

$$C(J; \mathbb{R}^d) := \{x : J \rightarrow \mathbb{R}^d : x \text{ ist stetig}\};$$

zusammen mit der Norm  $\|\cdot\|_\infty$ , die durch

$$\|x\|_\infty := \sup_{t \in J} \|x(t)\|$$

gegeben ist (wobei  $\|x(t)\|$  die Euklidische Norm des Vektors  $x(t) \in \mathbb{R}^d$  bezeichnet). Dann ist  $C(J; \mathbb{R}^d)$  zusammen mit der Norm  $\|\cdot\|_\infty$  ein vollständiger normierter Raum, also ein sogenannter Banachraum (siehe Beispiel A.4.7 im Anhang; allgemeine Aussagen über normierte Räume und Banachräume sowie weitere Beispiele für solche Räume finden Sie allgemein im Abschnitt A.4 im Anhang).

Nun schauen wir uns eine geeignete Abbildung  $T : C(J; \mathbb{R}^d) \rightarrow C(J; \mathbb{R}^d)$  an (d.h. eine Abbildung, die Funktionen auf Funktionen abbildet): Für jedes  $x \in C(J; \mathbb{R}^d)$  definieren wir  $T(x) \in C(J; \mathbb{R}^d)$  durch

$$(T(x))(t) = x(t_0) + \int_{t_0}^t f(s, x(s)) \, ds \quad \text{für alle } t \in J. \quad (4.2)$$

Die wesentliche Idee ist dann: Eine Funktion  $x \in C(J; \mathbb{R}^d)$  löst die Integralgleichung (4.1) (und somit das äquivalente Anfangswertproblem aus Proposition 4.1.1) genau dann, wenn  $x$  ein *Fixpunkt* von  $T$  ist, d.h. wenn  $T(x) = x$  ist. Die Existenz und Eindeutigkeit von solch einem Fixpunkt  $x$  wollen wir mit Hilfe eines geeigneten Fixpunktsatzes zeigen.

Wir müssen also folgende Aufgaben erledigen:

- Aufgaben-Liste 4.1.2.** (i) Wir müssen einen geeigneten Fixpunktsatz für diesen Plan finden.
- (ii) Wir müssen nachrechnen, dass die Abbildung  $T$  die Voraussetzungen des Fixpunktsatzes erfüllt.
- (iii) Wir müssen damit umgehen, dass im Allgemeinen eben nicht  $G = \mathbb{R}^{1+d}$  gilt; dies ist ein Problem bei der Definition der Abbildung  $T$ : Wenn wir in  $T$  einfach eine beliebige Funktion  $x \in C(J; \mathbb{R}^d)$  einsetzen, ist der Term  $f(s, x(s))$ , der im Integral in (4.2) auftritt, vielleicht gar nicht definiert, weil nicht sichergestellt ist, dass  $(s, x(s)) \in G$  gilt.

Der Sinn des verbleibenden Teils von Abschnitt 4.1 besteht darin, diese Liste abzuarbeiten und somit den Existenz- und Eindeutigkeitssatz von Picard–Lindelöf (Satz 4.1.9) zu beweisen. Unser Plan für die Abarbeitung von Liste 4.1.2 sieht hierbei folgendermaßen aus:

- Punkt (i) lösen wir, indem wir eine geeignete Version des Banachschen Fixpunktsatzes beweisen (Satz 4.1.4).
- Punkt (ii) lösen wir, indem wir eine geeignete Voraussetzung an die Funktion  $f$  auf der rechten Seite der Differentialgleichung stellen (eine sogenannte *lokale Lipschitz-Bedingung*).
- Punkt (iii) ist ein kleines bisschen technischer: Wir lösen dieses Problem, in dem wir annehmen, dass  $G$  offen ist, indem wir das Intervall  $J$  klein genug wählen (um den Startzeitpunkt  $t_0$  herum) und indem wir die Abbildung  $T$  nicht auf ganz  $C(J; \mathbb{R}^d)$ , sondern nur auf einer geeigneten Teilmenge  $M$  hiervon definieren.

Wir beginnen nun mit dem ersten Punkt, d.h. wir beweisen eine geeignete Version des Banachschen Fixpunktsatzes. Zur Formulierung des Banachschen Fixpunktes benötigt man das Konzept *Lipschitz-Stetigkeit*<sup>2</sup>, das in Definition B.2.1 im Anhang wiederholt wird. Die folgende einfache Aussage ist wichtig für das Verständnis des Banachschen Fixpunkt:

**Aufgabe 4.1.3.** Sei  $(M, d)$  ein metrischer Raum und sei  $f : M \rightarrow M$  Lipschitz-stetig mit Lipschitz-Konstante  $L \geq 0$ . Zeigen Sie: Für jedes  $n \in \mathbb{N}_0$  ist  $f^n$  (die  $n$ -fache Hintereinanderausführung von  $f$  mit sich selbst) Lipschitz-stetig mit Lipschitz-Konstante höchstens  $L^n$ .

Es ist wichtig darauf hinzuweisen, dass die Konstante  $L^n$  in obiger Aufgabe nicht immer optimal ist: Es kann vorkommen, dass die Lipschitz-Konstante von  $f^n$  deutlich kleiner ist als  $L^n$ :

Im Anhang finden Sie, inklusive Beweis, eine Variante des Banachschen Fixpunktsatzes, die häufig in Vorlesungen und in Lehrbüchern bewiesen wird (Satz B.2.3): Dort wird eine Lipschitz-stetige Funktion mit Lipschitz-Konstante  $L < 1$  betrachtet. Für unseren Zwecke verwenden wir aber die folgende etwas allgemeinere Version des Satzes:

**Satz 4.1.4 (Banachscher Fixpunktsatz<sup>3</sup>: Allgemeine Version).** Sei  $(M, d)$  ein nicht-leerer vollständiger metrischer Raum und sei  $f : M \rightarrow M$  eine Lipschitz-stetige Abbildung. Für jedes  $n \in \mathbb{N}_0$  bezeichne  $L_n$  die Lipschitzkonstante von  $f^n$  und es sei  $\sum_{n=0}^{\infty} L_n < \infty$ .

Dann besitzt  $f$  genau einen Fixpunkt  $\hat{x} \in M$ . Außerdem konvergiert für jedes  $x \in M$  die Folge der Iterierten  $(f^n(x))_{n \in \mathbb{N}}$  gegen  $\hat{x}$ .

---

<sup>2</sup>Benannt nach Rudolf Lipschitz (1832 – 1903), deutscher Mathematiker.

<sup>3</sup>Benannt nach Stefan Banach (1892 – 1945), polnischer Mathematiker; Mitbegründer der Funktionalanalysis und einer der bedeutendsten Mathematiker in der ersten Hälfte des 20. Jahrhunderts.

#### 4.1. Lokale Existenz und globale Eindeutigkeit: Der Satz von Picard–Lindelöf

*Beweis.* Wir setzen  $C := \sum_{n=0}^{\infty} L_n$ . Aus  $C < \infty$  folgt  $L_n \rightarrow 0$  für  $n \rightarrow \infty$ .

Wir zeigen nun zunächst, dass  $f$  höchstens einen Fixpunkt haben kann. Seien dazu  $\hat{x}_1$  und  $\hat{x}_2$  Fixpunkte von  $f$ , und sei  $n \in \mathbb{N}$  so groß, dass  $L_n < 1$  gilt. Dann ist

$$d(\hat{x}_1, \hat{x}_2) = d(f^n(\hat{x}_1), f^n(\hat{x}_2)) \leq L_n d(\hat{x}_1, \hat{x}_2);$$

wegen  $L_n < 1$  folgt hieraus  $d(\hat{x}_1, \hat{x}_2) = 0$ , also  $\hat{x}_1 = \hat{x}_2$ .

Sei nun  $x \in M$  fest. Wir wollen beweisen, dass die Folge  $(f^n(x))_{n \in \mathbb{N}}$  eine Cauchy-Folge ist. Hierzu definieren wir zunächst  $\delta := d(f(x), x)$  und bemerken, dass für alle  $n \in \mathbb{N}$  aus der Dreiecksungleichung die Abschätzung

$$d(f^n(x), x) \leq \sum_{k=0}^{n-1} d(f^{k+1}(x), f^k(x)) \leq \sum_{k=0}^{n-1} L_k d(f(x), x) \leq C\delta \quad (4.3)$$

folgt.

Wir wählen nun ein beliebiges, aber festes  $\varepsilon > 0$ . Wegen  $L_n \rightarrow 0$  für  $n \rightarrow \infty$  gibt es ein  $n_0 \in \mathbb{N}$  mit  $L_{n_0} C\delta < \varepsilon$ . Für alle  $n \geq n_0$  folgt nun aus (4.3)

$$d(f^n(x), f^{n_0}(x)) \leq L_{n_0} d(f^{n-n_0}(x), x) \leq L_{n_0} C\delta < \varepsilon;$$

also ist  $(f^n(x))_{n \in \mathbb{N}}$  tatsächlich eine Cauchy-Folge (siehe Aufgabe A.2.2(b) im Anhang).

Wegen der Vollständigkeit von  $(M, d)$  konvergiert also die Folge  $(f^n(x))_{n \in \mathbb{N}}$  gegen einen Punkt  $\hat{x} \in M$ . Wir müssen nur noch zeigen, dass  $\hat{x}$  ein Fixpunkt von  $f$  ist. Dazu berechnen wir

$$f(\hat{x}) = f\left(\lim_{n \rightarrow \infty} f^n(x)\right) = \lim_{n \rightarrow \infty} f(f^n(x)) = \lim_{n \rightarrow \infty} f^{n+1}(x) = \hat{x}.$$

Somit haben wir gezeigt, dass es genau einen Fixpunkt gibt, und dass die Folge  $(f^n(x))_{n \in \mathbb{N}}$  für jedes  $x \in M$  gegen diesen Fixpunkt konvergiert.  $\square$

**Aufgabe 4.1.5.** Zeigen Sie, dass die Variante des Banachschen Fixpunktsatzes in Satz B.2.3 im Anhang ein Spezialfall von Satz 4.1.4 ist.

Nun kommen wir zum zweiten Punkt aus der Liste 4.1.2: Um im Beweis von Satz 4.1.9 zu zeigen, dass der Banachsche Fixpunktsatz auf die Abbildung  $T$  angewendet werden kann, brauchen wir eine Voraussetzung an die Funktion  $f$  auf der rechten Seite der Differentialgleichung. Dieser Voraussetzung geben wir in der folgenden Definition einen Namen:

**Definition 4.1.6 (Lokale Lipschitz-Bedingung).** Sei  $d \in \mathbb{N}$ , sei  $G \subseteq \mathbb{R}^{1+d}$  nichtleer und offen, und sei  $f : G \rightarrow \mathbb{R}^d$  eine Abbildung. Wir sagen, die Abbildung  $f$  erfüllt die *lokale Lipschitz-Bedingung*, wenn es für jeden Punkt

$(t_0, x_0) \in G$  eine Umgebung  $U \subseteq G$  von  $(t_0, x_0)$  und eine Zahl  $L \geq 0$  gibt derart, dass die folgende Eigenschaft erfüllt ist:

Für alle  $t \in \mathbb{R}$  und alle  $x_1, x_2 \in \mathbb{R}^d$  mit  $(t, x_1) \in U$  und  $(t, x_2) \in U$  gilt

$$\|f(t, x_2) - f(t, x_1)\| \leq L\|x_2 - x_1\|;$$

hierbei bezeichnet  $\|\cdot\|$  die Euklidische Norm auf  $\mathbb{R}^d$ .

**Bemerkung 4.1.7.** Seien Sie bitte etwas vorsichtig bei der Interpretation des Begriffs *lokale Lipschitz-Bedingung* aus Definition 4.1.6:

- Die lokale Lipschitz-Bedingung impliziert im Allgemeinen *nicht*, dass  $f$  in einer Umgebung von  $(t_0, x_0)$  Lipschitz-stetig ist.
- Die lokale Lipschitz-Bedingung impliziert noch nicht einmal, dass  $f$  stetig bezüglich  $t$  ist (generell werden wir die Stetigkeit von  $f$  aber in allen nachfolgenden Sätzen zusätzlich voraussetzen).
- Die lokale Lipschitz-Bedingung bedeutet lediglich, dass  $f$  bezüglich der letzten  $d$  Komponenten in seinem Argument Lipschitz-stetig ist (solange man sich in  $U$  bewegt) und dass  $L$  unabhängig von  $t$  gewählt werden kann (solange man sich in  $U$  bewegt).

Die folgende Proposition enthält eine hinreichende Bedingung um die lokale Lipschitzbedingung nachzuprüfen.

**Proposition 4.1.8.** Sei  $d \in \mathbb{N}$ , sei  $G \subseteq \mathbb{R}^{1+d}$  nichtleer und offen, und sei  $f : G \rightarrow \mathbb{R}^d$  eine Abbildung. Falls  $f$  stetig differenzierbar ist, dann erfüllt  $f$  die lokale Lipschitzbedingung.

*Beweis.* Der Beweis ist eine Aufgabe auf dem zweiten Übungsblatt. □

Nun kommen wir wie angekündigt zum Eindeutigkeitssatz von Picard–Lindelöf. In dessen Beweis werden wir den Schritt (ii) aus der Liste 4.1.2 komplettieren und den Schritt (iii) durchführen.

**Satz 4.1.9 (Lokaler Existenz- und Eindeutigkeitssatz von Picard–Lindelöf<sup>4</sup>).** Sei  $d \in \mathbb{N}$ , sei  $G \subseteq \mathbb{R}^{1+d}$  offen und sei  $f : G \rightarrow \mathbb{R}^d$  eine stetige Funktion, die die lokale Lipschitz-Bedingung erfüllt. Zudem seien  $t_0 \in \mathbb{R}$  and  $x_0 \in \mathbb{R}^d$  mit  $(t_0, x_0) \in G$ . Dann gibt es ein  $\delta > 0$  und eine abgeschlossene Umgebung  $B$  von  $x_0$  in  $\mathbb{R}^d$  mit den folgenden Eigenschaften:

---

<sup>4</sup>Benannt nach Ernst Leonard Lindelöf (1870 – 1946), finnischer Mathematiker, und Charles Émile Picard (1856 – 1941), französischer Mathematiker.

#### 4.1. Lokale Existenz und globale Eindeutigkeit: Der Satz von Picard–Lindelöf

---

Es ist  $[t_0 - \delta, t_0 + \delta] \times B \subseteq G$ , und für jedes nicht-triviale Teilintervall  $I \subseteq [t_0 - \delta, t_0 + \delta]$ , das  $t_0$  enthält, gibt es genau eine Funktion  $x : I \rightarrow B$ , die das Anfangswertproblem

$$\begin{cases} \dot{x}(t) = f(t, x(t)), \\ x(t_0) = x_0 \end{cases} \quad (4.4)$$

löst.

*Beweis.* Wir versehen  $\mathbb{R}^d$  im ganzen Beweis mit der Euklidischen Norm  $\|\cdot\|$ . Durch geeignete Verschiebung des Problems können wir im gesamten Beweis annehmen, dass  $t_0 = 0$  ist.

Weil  $f$  die lokale Lipschitz-Bedingung erfüllt, gibt es eine Umgebung  $U \subseteq G$  von  $(0, x_0)$  und eine Zahl  $L \geq 0$  derart, dass

$$\|f(t, x_2) - f(t, x_1)\| \leq L\|x_2 - x_1\|$$

für alle  $t \in \mathbb{R}$  und alle  $x_1, x_2 \in \mathbb{R}^d$  mit  $(t, x_1) \in U$  und  $(t, x_2) \in U$  gilt. Weil  $(0, x_0)$  im Inneren von  $U$  liegt, gibt es ein Zahl  $\delta_0 > 0$  und eine abgeschlossene Kugel  $B$  im  $\mathbb{R}^d$  mit Radius  $r > 0$  und Mittelpunkt  $x_0$  derart, dass  $[-\delta_0, \delta_0] \times B \subseteq U \subseteq G$  gilt.

Im Folgenden werden wir feststellen, dass  $\delta_0$  vielleicht noch nicht klein genug ist (deshalb haben wir es  $\delta_0$  statt  $\delta$  genannt). Wir schrumpfen das Intervall  $[-\delta_0, \delta_0]$  noch ein wenig, indem wir folgendermaßen vorgehen:

Es bezeichne  $m \in [0, \infty)$  das Maximum von  $\|f(t, y)\|$  für  $(t, y)$  aus der kompakten Menge  $[-\delta_0, \delta_0] \times B$ . Jetzt wählen wir  $\delta > 0$  so klein, dass  $\delta \leq \delta_0$  und zugleich  $\delta m \leq r$  gilt. Mit diesem  $\delta$  können wir jetzt die Aussage des Satzes beweisen.

Natürlich gilt  $[-\delta, \delta] \times B \subseteq U \subseteq G$ . Sei nun also  $I \subseteq [-\delta, \delta]$  ein nicht-triviales Teilintervall, das 0 enthält. Wir betrachten zunächst den Fall, dass  $I$  abgeschlossen ist.

Wegen  $I \times B \subseteq U \subseteq G$  gilt für  $t \in I$  und  $x_1, x_2 \in B$  die Lipschitz-Abschätzung

$$\|f(t, x_2) - f(t, x_1)\| \leq L\|x_2 - x_1\|. \quad (4.5)$$

Laut Proposition 4.1.1 müssen wir nur zeigen, dass es genau eine stetige Abbildung  $x : I \rightarrow B$  gibt, welche die Integralgleichung

$$x(t) = x_0 + \int_0^t f(s, x(s)) \, ds \quad \text{für alle } t \in I \quad (4.6)$$

löst. Um dies zu zeigen, betrachten wir nun die Menge

$$M := \{x \in C(I; \mathbb{R}^d) : x(t) \in B \text{ für alle } t \in I\}.$$

Dies ist eine abgeschlossene Teilmenge des Banachraumes  $C(I; \mathbb{R}^d)$  und somit selbst ein vollständiger metrischer Raum (siehe Aufgabe A.3.2 im Anhang). Auf diesem  $M$  können wir nun unsere Abbildung  $T$  definieren: Für jedes  $x \in M$  definieren wir die Funktion  $T(x) : I \rightarrow \mathbb{R}^d$  durch

$$(T(x))(t) = x_0 + \int_0^t f(s, x(s)) \, ds \quad \text{für alle } t \in I;$$

dies ist wohldefiniert, denn für  $s \in I$  und  $x \in M$  ist  $x(s) \in B$ , also  $(s, x(s)) \in I \times B \subseteq G$ . Damit ist Punkt (iii) aus Liste 4.1.2 abgearbeitet. Wir müssen nun noch Punkt (ii) abarbeiten – d.h., wir müssen zeigen, dass  $T$  eine Lipschitz-stetige Abbildung von  $M$  nach  $M$  ist, die die Voraussetzungen des Banachschen Fixpunktsatzes 4.1.4 erfüllt.

Zunächst bemerken wir, dass die Funktion  $T(x)$  für jedes  $x \in M$  tatsächlich wieder ein Element von  $M$  ist; sei dazu  $x \in M$  fest. Wegen der Stetigkeit von  $f$  und  $x$  ist auch  $T(x)$  stetig, also ein Element von  $C(I; \mathbb{R}^d)$ . Nun fixieren wir  $t \in I$  und müssen noch  $(Tx)(t) \in B$  zeigen – d.h. wir müssen zeigen, dass  $(Tx)(t)$  von  $x_0$  höchstens den Abstand  $r$  hat. Dazu berechnen wir

$$\|(Tx)(t) - x_0\| = \left\| \int_0^t f(s, x(s)) \, ds \right\| \leq |t| m \leq \delta m \leq r,$$

denn so hatten wir  $\delta$  gewählt. Also ist tatsächlich  $T(x) \in M$ .

Zuletzt zeigen wir für jedes  $n \in \mathbb{N}_0$ , dass  $T^n$  Lipschitz-stetig mit Lipschitz-Konstante kleiner oder gleich  $\frac{\delta^n L^n}{n!}$  ist. Dazu muss man geschickt vorgehen und induktiv folgende etwas stärkere Aussage zeigen:<sup>5</sup>

Für alle  $x, \tilde{x} \in M$ , jedes  $n \in \mathbb{N}_0$  und jedes  $t \in I$  gilt

$$\|(T^n(x))(t) - (T^n(\tilde{x}))(t)\| \leq \frac{|t|^n L^n}{n!} \|x - \tilde{x}\|_\infty \quad (4.7)$$

Wir zeigen dies zunächst für alle  $t \geq 0$ . Für  $n = 0$  ist die Aussage offensichtlich; also sei die Aussage nun für ein festes  $n \in \mathbb{N}_0$  und alle  $0 \leq t \in I$  bereits bewiesen. Dann gilt

$$\begin{aligned} \|(T^{n+1}(x))(t) - (T^{n+1}(\tilde{x}))(t)\| &= \left\| \int_0^t f(s, (T^n(x))(s)) - f(s, (T^n(\tilde{x}))(s)) \, ds \right\| \\ &\stackrel{(4.5)}{\leq} \int_0^t L \|(T^n(x))(s) - (T^n(\tilde{x}))(s)\| \, ds \end{aligned}$$

<sup>5</sup>Die auf den ersten Blick widersprüchliche Bemerkung, dass man zum Beweis einer Aussage eine „etwas stärkere Aussage“ zeigen muss, ist ein Phänomen, das bei Induktionsbeweisen häufig vorkommt: Eine stärkere Aussage zu betrachten sorgt nämlich nicht nur dafür, dass man mehr zeigen muss, sondern auch dafür, dass man im Induktionsschritt mehr zur Verfügung hat.

$$\leq \int_0^t L \frac{s^n L^n}{n!} ds \|x - \tilde{x}\|_\infty = \frac{t^{n+1} L^{n+1}}{(n+1)!} \|x - \tilde{x}\|_\infty.$$

Für negative  $t$  zeigt man (4.7) auf dieselbe Weise (man muss nur etwas auf die Beträge und das Vorzeichen achten).

Indem man in (4.7) das Supremum über alle  $t \in I$  betrachtet, erhält man, dass  $T^n$  Lipschitz-stetig mit Lipschitz-Konstante höchstens  $\frac{\delta^n L^n}{n!}$  ist. Diese Zahlen sind über  $n \in \mathbb{N}$  summierbar – d.h.  $T$  erfüllt die Voraussetzungen des Banachschen Fixpunktsatzes 4.1.4. Also besitzt  $T$  genau einen Fixpunkt in  $M$  – und dies bedeutet gerade, dass es genau eine stetige Funktion  $x : I \rightarrow B$  gibt, welche die Integralgleichung (4.6) löst.

Zuletzt müssen wir noch den Fall betrachten, dass  $I$  nicht abgeschlossen ist. Die Existenzaussage folgt dann sofort aus dem bereits Gezeigten, da sie ja sogar auf dem Abschluss  $\bar{I}$  gilt. Die Eindeutigkeitsaussage folgt ebenfalls aus dem bereits Gezeigten, denn wir finden eine Folge von nicht-trivialen abgeschlossenen Intervallen  $I_n \subseteq I$ , die alle  $t_0$  enthalten, und deren Vereinigung gleich  $I$  ist. Wenn  $x, \tilde{x} : I \rightarrow B$  zwei Lösungen sind, dann stimmen sie (laut dem, was wir oben gezeigt haben) auf jedem der abgeschlossenen Intervalle  $I_n$  überein – und somit auch auf  $I$ .  $\square$

**Bemerkungen 4.1.10.** (a) Die etwas technisch anmutende Aussage im zweiten Absatz von Satz 4.1.9 wird ein wenig klarer, wenn man die Existenzaussage von der Eindeutigkeitsaussage trennt:

Für die Existenzaussage spielt das Intervall  $I$  im Grunde gar keine Rolle: Wir können einfach  $I = [t_0 - \delta, t_0 + \delta]$  wählen und erhalten damit, dass es eine Lösung  $x : [t_0 - \delta, t_0 + \delta] \rightarrow B$  des Anfangswertproblems (4.4) gibt.

Für die Eindeutigkeitsaussage hingegen spielt das Intervall  $I$  in folgendem Sinne eine Rolle: Wenn wir eine Lösung  $x : [t_0 - \delta, t_0 + \delta] \rightarrow B$  haben, und auf einem kleineren Intervall  $I$  nun noch eine weitere Lösung  $\tilde{x}$  betrachten, dann folgt aus der Eindeutigkeitsaussage, dass  $\tilde{x}$  schlichtweg die Einschränkung von  $x$  auf  $I$  ist.

In den beiden nachfolgenden Korollaren werden wir die Existenz- und Eindeutigkeitsaussage von Satz 4.1.9 ohnehin noch einmal getrennt betrachten, sodass die Situation dort noch einmal klarer werden sollte.

(b) Die lokale Lipschitz-Bedingung an  $f$  in Satz 4.1.9 kann man nicht einfach weglassen. Zum Beispiel ist diese Bedingung im Anfangswertproblem in der Einstiegsfrage (a) dieses Kapitels verletzt, und dieses Anfangswertproblem besitzt in der Tat zwei verschiedene Lösungen:

Zum einen die konstante 0-Funktion, und zum anderen die Funktion  $x : \mathbb{R} \rightarrow \mathbb{R}$ , die durch

$$x(t) = \begin{cases} 0 & \text{falls } t < 0, \\ \frac{t^2}{4} & \text{falls } t \geq 0 \end{cases}$$

gegeben ist.

- (c) Der Beweis des Satzes von Picard–Lindelöf liefert sogar ein Verfahren, um eine Lösung des gegebenen Anfangswertproblems zu bestimmen: Man wähle ein beliebiges  $\tilde{x} \in M$  (zum Beispiel die Funktion, die konstant gleich  $x_0$  ist). Der Banachsche Fixpunktsatz besagt dann, dass die Iterierten  $T^n(\tilde{x})$  für  $n \rightarrow \infty$  gegen den Fixpunkt von  $T$  – also gegen die Lösung des gegebenen Anfangswertproblems – konvergieren.

Man nennt die Funktionen  $T^n(\tilde{x})$  die *Picard-Iterierten* des Anfangswertproblems.

Die folgende Beobachtung ist vielleicht ein wenig überraschend, aber leicht zu beweisen: Die lokale Eindeutigkeitsaussage im Satz von Picard–Lindelöf lässt sich mit einem kleinen Trick – und ohne zusätzliche Voraussetzungen! – zu einer globalen Eindeutigkeitsaussage verallgemeinern:

**Korollar 4.1.11 (Globaler Eindeutigkeitsatz von Picard–Lindelöf).** Sei  $d \in \mathbb{N}$ , sei  $G \subseteq \mathbb{R}^{1+d}$  offen und sei  $f : G \rightarrow \mathbb{R}^d$  eine stetige Funktion, die die lokale Lipschitz-Bedingung erfüllt. Zudem seien  $t_0 \in \mathbb{R}$  und  $x_0 \in \mathbb{R}^d$  mit  $(t_0, x_0) \in G$ .

Sei  $J \subseteq \mathbb{R}$  ein nicht-triviales Intervall, das  $t_0$  enthält. Dann gibt es höchstens eine Funktion  $x : J \rightarrow \mathbb{R}^d$ , die das Anfangswertproblem

$$\begin{cases} \dot{x}(t) = f(t, x(t)), \\ x(t_0) = x_0, \end{cases}$$

löst.

*Beweis.* Seien  $x, \tilde{x} : J \rightarrow \mathbb{R}^d$  zwei Funktionen, die das gegebene Anfangswertproblem lösen; dann sind beide Funktionen stetig. Wir definieren

$$C := \{t \in J : x(t) = \tilde{x}(t)\}.$$

Die Menge  $C$  enthält die Zeit  $t_0$  und ist abgeschlossen in  $J$  (wegen der Stetigkeit von  $x$  und  $\tilde{x}$ ). Wir zeigen nun, dass  $C$  auch offen in  $J$  ist. Weil  $J$  zusammenhängend ist, folgt dann  $C = J$ , also  $\tilde{x} = x$ .

Sei  $t_1 \in C$  und sei  $x_1 := x(t_1) = \tilde{x}(t_1)$ . Dann lösen  $x$  und  $\tilde{x}$  beide das Anfangswertproblem

$$\begin{cases} \dot{x}(t) = f(t, x(t)), \\ x(t_1) = x_1, \end{cases} \quad (4.8)$$

Laut Satz 4.1.9 gibt es ein  $\delta > 0$  und eine Umgebung  $B$  von  $x_1$  derart, dass  $[t_1 - \delta, t_1 + \delta] \times B \subseteq G$  gilt und dass es auf jedem nicht-trivialen Teilintervall von  $[t_1 - \delta, t_1 + \delta]$ , das  $t_1$  enthält, nur eine Funktion gibt, die (4.8) löst. Auf dem Teilintervall  $J \cap [t_1 - \delta, t_1 + \delta]$  lösen aber  $x$  und  $\tilde{x}$  beide das Anfangswertproblem (4.8), also stimmen sie dort überein.

Damit haben wir  $J \cap [t_1 - \delta, t_1 + \delta] \subseteq C$  gezeigt. Insbesondere enthält  $C$  also die offene Umgebung  $J \cap (t_1 - \delta, t_1 + \delta)$  von  $t_1$  in  $J$ , d.h.  $C$  ist offen in  $J$ .  $\square$

Wir schließen diesen Unterabschnitt mit folgendem Korollar: Es enthält keine neuen Informationen im Vergleich zu Satz 4.1.9, aber wir führen in diesem Korollar nur die Existenzaussage aus dem Satz auf; dies führt dazu, dass die Aussage deutlich weniger technisch wird.

**Korollar 4.1.12 (Lokaler Existenzsatz von Picard–Lindelöf).** Sei  $d \in \mathbb{N}$ , sei  $G \subseteq \mathbb{R}^{1+d}$  offen und sei  $f : G \rightarrow \mathbb{R}^d$  eine stetige Funktion, die die lokale Lipschitz-Bedingung erfüllt. Zudem seien  $t_0 \in \mathbb{R}$  und  $x_0 \in \mathbb{R}^d$  mit  $(t_0, x_0) \in G$ .

Dann gibt es ein abgeschlossenes Intervall  $J \subseteq \mathbb{R}$ , das  $t_0$  im Inneren enthält, und eine Lösung  $x : J \rightarrow \mathbb{R}^d$  des Anfangswertproblems

$$\begin{cases} \dot{x}(t) = f(t, x(t)), \\ x(t_0) = x_0. \end{cases}$$

Zusammenfassend kann man sagen: Die wichtigste Beweisarbeit in diesem Abschnitt ist im Beweis von Satz 4.1.9 passiert; die wichtigsten Aussagen des Abschnitts finden sich aber in den Korollaren 4.1.11 und 4.1.12.

**Literaturstellen und verwandte Ergebnisse.** (a) Die Version des Banachschen Fixpunktsatzes, die wir in Satz 4.1.4 besprochen haben, finden Sie zum Beispiel auch in [Heu04, Satz 12.1 auf Seite 138]; diese Version wird dort als *Fixpunktsatz von Weissinger* bezeichnet.

(b) Einen etwas anderen Zugang als in diesem Manuskript finden Sie zum Beispiel im Buch [PW19]; dort wird zunächst in Satz 2.1.3 die Eindeutigkeit der Lösung mit Hilfe einer Integralungleichung gezeigt (ein ähnliches Argument wie dort werden wir auch in dieser Vorlesung später in Satz 7.1.4 sehen).

Interessant ist, dass man für dieses Argument die folgende Beobachtung benötigt: Wenn eine stetige Funktion  $f : \mathbb{R}^{1+d} \supseteq G \rightarrow \mathbb{R}^d$  die lokale Lipschitzbedingung erfüllt, dann erfüllt sie auf jeder kompakten Teilmenge von  $G$  sogar eine *globale* Lipschitzbedingung. Diese Aussage finden Sie z.B. in [PW19, Proposition 2.1.2].

Existenz von Lösungen wird in [PW19, Satz 2.2.2] ebenfalls mit Hilfe der Picard-Iteration bewiesen – allerdings wird dort eine andere Variante

des Banachschen Fixpunktsatzes verwendet, die Sie im vorliegenden Manuskript in Satz B.2.3 im Anhang nachlesen können; der Grund, warum diese schwächere Variante des Satzes ausreicht, ist, dass im Beweis von [PW19, Satz 2.2.2] eine geeignete Renormierung vorgenommen wird (welche wir nicht benötigen haben).

## 4.2 Lokale Existenz: Der Satz von Peano

In Bemerkung 4.1.10 hatten wir festgestellt, dass man die lokale Lipschitz-Bedingung im Satz von Picard–Lindelöf nicht einfach weglassen kann, da ansonsten die Eindeutigkeit der Lösung verloren gehen kann. Sie wundern sich vielleicht, ob auch die Existenz einer Lösung fehlschlagen kann, wenn man die lokale Lipschitz-Bedingung nicht voraussetzt.

Die Antwort lautet – wohl etwas überraschend – *Nein*: Um Existenz von Lösungen zu erhalten, genügt es tatsächlich, wenn die Funktion auf der rechten Seite der Differentialgleichung lediglich stetig ist. Dies ist der Inhalt des folgenden *Existenzsatzes von Peano*, der eine Verallgemeinerung von Korollar 4.1.12 darstellt:

**Satz 4.2.1 (Lokaler Existenzsatz von Peano<sup>6</sup>).** Sei  $d \in \mathbb{N}$ , sei  $G \subseteq \mathbb{R}^{1+d}$  offen und sei  $f : G \rightarrow \mathbb{R}^d$  eine stetige Funktion. Zudem seien  $t_0 \in \mathbb{R}$  und  $x_0 \in \mathbb{R}^d$  mit  $(t_0, x_0) \in G$ .

Dann gibt es ein abgeschlossenes Intervall  $J \subseteq \mathbb{R}$ , das  $t_0$  im Inneren enthält, und eine Lösung  $x : J \rightarrow \mathbb{R}^d$  des Anfangswertproblems

$$\begin{cases} \dot{x}(t) = f(t, x(t)), \\ x(t_0) = x_0. \end{cases}$$

Bitte beachten Sie: Der einzige Unterschied zwischen Satz 4.2.1 und Korollar 4.1.12 besteht darin, dass in Satz 4.1.12 die lokale Lipschitz-Bedingung weggelassen wurde. Der Satz von Peano ist also eine deutliche Verallgemeinerung der Existenzaussage von Picard–Lindelöf.

*Beweis von Satz 4.2.1.* Aufgrund eines Verschiebungsarguments können wir wieder ohne Einschränkung  $t_0 = 0$  annehmen.

Wir wählen eine abgeschlossene Kugel  $B \subseteq \mathbb{R}^d$  mit Mittelpunkt  $x_0$  und Radius  $r > 0$  und ein  $\delta > 0$ , wobei wir  $r$  und  $\delta$  so klein wie im Beweis des Satzes 4.1.9 von Picard–Lindelöf wählen (beachten Sie aber, dass die Menge  $U$  nun nicht mehr auftaucht, da  $f$  ja nicht die lokale Lipschitz-Bedingung

---

<sup>6</sup>Benannt nach Giuseppe Peano (1858 – 1932), italienischer Mathematiker.

erfüllt). Wir setzen nun einfach  $J := [-\delta, \delta]$ , definieren die abgeschlossene und konvexe Teilmenge

$$M := \{x \in C(J; \mathbb{R}^d) : x(t) \in B \text{ für alle } t \in J\}.$$

von  $C(J; \mathbb{R}^d)$  und die Abbildung  $T : M \rightarrow M$ , die durch

$$(Tx)(t) = x_0 + \int_0^t f(s, x(s)) \, ds \quad \text{für alle } t \in I;$$

für alle  $x \in C(J; \mathbb{R}^d)$  gegeben ist (dass  $T$  tatsächlich von  $M$  nach  $M$  abbildet, sieht man wie im Beweis des Satzes 4.1.9 von Picard–Lindelöf; dies funktioniert, weil wir  $\delta$  und  $r$  ja genauso gewählt haben, wie in diesem Beweis).

Laut Proposition 4.1.1 müssen wir nur zeigen, dass  $T$  einen Fixpunkt in  $M$  besitzt, und dies tun wir nun mithilfe des Schauderschen Fixpunktsatzes B.3.6.

Wir wissen bereits, dass  $M$  eine abgeschlossene und konvexe Teilmenge eines Banachraumes ist. Als nächstens müssen wir zeigen, dass  $T$  stetig ist (man beachte, dass wir keine Lipschitz-Stetigkeit von  $T$  erwarten können, da  $f$  nicht die lokale Lipschitz-Bedingung erfüllen muss – also müssen wir die Stetigkeit von  $T$  auf andere Weise zeigen). Sei also  $(x_n)$  eine Folge von Funktionen in  $M$ , die gegen eine Funktion  $x \in M$  konvergiert. Für jedes  $t \in J$  mit  $t \geq 0$  erhalten wir die Abschätzung

$$\|(Tx_n)(t) - (Tx)(t)\| \leq \int_0^t |f(s, x_n(s)) - f(s, x(s))| \, ds \leq \int_0^1 |f(s, x_n(s)) - f(s, x(s))| \, ds,$$

und eine analoge Abschätzung gilt für  $t \leq 0$ . Das Integral am Ende der Abschätzung hängt nicht von  $t$  ab und konvergiert aufgrund des Satzes von der dominerten Konvergenz gegen 0. Dies zeigt, dass

$$\|Tx_n - Tx\|_\infty \rightarrow 0 \quad \text{für } n \rightarrow \infty$$

gilt; also ist  $T$  tatsächlich stetig.

Zuletzt müssen wir, um den Schauderschen Fixpunktsatz anwenden zu können, noch zeigen, dass das Bild  $T(M)$  von  $T$  relativ kompakt ist. Dazu benutzen wir den Satz von Arzela–Ascoli (genauer: sein Korollar A.5.10 im Anhang):

Weil  $M$  beschränkt ist und  $T$  nach  $M$  abbildet, ist auch  $T(M)$  beschränkt. Außerdem können wir zeigen, dass alle Funktionen in  $T(M)$  Lipschitz-stetig sind und ihre Lipschitz-Konstante durch eine feste Zahl nach oben abgeschätzt werden kann: Sei nämlich  $L := \sup\{\|f(s, y)\| : s \in J, y \in B\}$ ; es ist  $L < \infty$ , weil  $f$  stetig und die Menge  $J \times B$  kompakt ist. Nun gilt für jedes  $x \in M$  und alle  $t_1, t_2 \in J$  mit  $t_1 \leq t_2$

$$\|(Tx)(t_2) - (Tx)(t_1)\| \leq \int_{t_1}^{t_2} \|f(s, x(s))\| \, ds \leq L(t_2 - t_1).$$

Also ist  $Tx$  Lipschitz-stetig mit Lipschitz-Konstante höchstens  $L$ . Somit sind die Voraussetzungen von Korollar A.5.10 für die Menge  $T(M)$  erfüllt, d.h.  $T(M)$  ist in der Tat relativ kompakt.

Also können wir den Fixpunktsatz B.3.6 von Schauder auf die Abbildung  $T$  anwenden; dieser besagt, dass  $T$  einen Fixpunkt besitzt, und dies zeigt die Behauptung.  $\square$

Man kann sich natürlich fragen, wie weit man das Spiel treiben kann – ist es beispielsweise möglich, sogar dann noch einen Existenzsatz zu zeigen, wenn man auch noch die Voraussetzung, dass  $f$  stetig ist, weglässt? Auf dem dritten Übungsblatt werden Sie sehen, dass die Antwort auf diese Frage im Allgemeinen *Nein* lautet.

**Literaturstellen und verwandte Ergebnisse.** (a) Einen Beweis des Satzes von Peano, der ohne den Schauderschen Fixpunktsatz auskommt, finden Sie zum Beispiel in [PW19, Satz 6.1.1]; anstelle des Schauderschen Fixpunktsatzes wird dort ein geeignetes Approximationsargument sowie der Satz von Picard–Lindelöf benutzt. Kompaktheit und der Satz von Arzelà–Ascoli spielen aber auch im dortigen Beweis eine wichtige Rolle.

(b) Ein nochmals anderer Beweis des Satzes von Peano – mit Hilfe sogenannter *Polygonzüge* – wird zum Beispiel in [Har02, Aufgabe 2.1 auf Seite 11] erläutert.

### 4.3 Das maximale Existenzintervall

Während man den globalen Eindeutigkeitsatz aus Korollar 4.1.11 als sehr zufriedenstellendes Ergebnis betrachten kann, haben die Existenzsätze aus Korollar 4.1.12 und Satz 4.2.1 den Nachteil, dass sie nur die lokale Existenz einer Lösung garantieren. In vielen Anwendungen möchte man aber gerne Kriterien dafür haben, dass eine Lösung für alle Zeiten, oder zumindest für Zeiten aus einem größeren Intervall, existiert.

Zu diesem Zweck führen wir in diesem Abschnitt das sogenannte *maximale Existenzintervall* eines Anfangswertproblems ein und studieren dessen Eigenschaften.

Wir beginnen zunächst mit der folgenden einfachen Proposition, die beschreibt, wie sich zwei nebeneinander liegende Lösungen einer Differentialgleichung *verkleben* lassen.

**Proposition 4.3.1.** Sei  $d \in \mathbb{N}$ , sei  $G \subseteq \mathbb{R}^{1+d}$  offen und sei  $f : G \rightarrow \mathbb{R}^d$  eine stetige Funktion.

Es seien  $I_1, I_2 \subseteq \mathbb{R}$  zwei nicht-triviale Intervalle, die sich in genau einem Punkt  $\hat{t}$  schneiden. Seien  $x_1 : I_1 \rightarrow \mathbb{R}^d$  und  $x_2 : I_2 \rightarrow \mathbb{R}^d$  Lösungen der Differentialgleichung  $\dot{x}(t) = f(t, x(t))$ , und es gelte  $x_1(\hat{t}) = x_2(\hat{t})$ .

Dann ist auch die Funktion  $x : I := I_1 \cup I_2 \rightarrow \mathbb{R}^d$ , die durch

$$x(t) = \begin{cases} x_1(t) & \text{falls } t \in I_1, \\ x_2(t) & \text{falls } t \in I_2 \end{cases}$$

gegeben ist, eine Lösung derselben Differentialgleichung.

Zum Verständnis des folgenden Beweises ist es sinnvoll, dass Sie sich noch einmal die genaue Definition des Lösungsbegriffs einer Differentialgleichung aus Definition 3.1.3 vergegenwärtigen.

*Beweis von Proposition 4.3.1.* Ohne Einschränkung können wir annehmen, dass  $I_1$  links von  $I_2$  liegt. Es kann nur im Punkt  $\hat{t}$  zu einem Problem kommen – denn dort ist  $x_1$  nur linksseitig differenzierbar und  $x_2$  nur rechtsseitig differenzierbar, also ist nicht klar, ob  $x$  dort differenzierbar ist ( $x$  ist in  $\hat{t}$  aber natürlich von beiden Seiten aus differenzierbar mit linksseitiger Ableitung  $\dot{x}_1(\hat{t})$  und rechtsseitiger Ableitung  $\dot{x}_2(\hat{t})$ ). Nun beobachten wir, dass

$$\dot{x}_1(\hat{t}) = f(\hat{t}, x_1(\hat{t})) = f(\hat{t}, x_2(\hat{t})) = \dot{x}_2(\hat{t})$$

gilt. Also stimmt die linksseitige Ableitung von  $x$  im Punkt  $\hat{t}$  mit der rechtsseitigen Ableitung überein, und somit ist  $x$  in  $\hat{t}$  differenzierbar. Für seine Ableitung in diesem Punkt gilt

$$\dot{x}(\hat{t}) = \dot{x}_1(\hat{t}) = f(\hat{t}, x_1(\hat{t})) = f(\hat{t}, x(\hat{t})).$$

Also erfüllt  $x$  auch im Punkt  $\hat{t}$  die Differentialgleichung. □

Die nächste Proposition und die nachfolgende Definition beschreiben, was mit dem Begriff *maximales Existenzintervall* gemeint ist.

**Proposition 4.3.2.** Sei  $d \in \mathbb{N}$ , sei  $G \subseteq \mathbb{R}^{1+d}$  offen und sei  $f : G \rightarrow \mathbb{R}^d$  eine stetige Funktion, die die lokale Lipschitz-Bedingung erfüllt. Zudem seien  $t_0 \in \mathbb{R}$  and  $x_0 \in \mathbb{R}^d$  mit  $(t_0, x_0) \in G$ .

Es bezeichne  $I_{\max} \subseteq \mathbb{R}$  die Vereinigung aller nicht-trivialen Intervalle in  $\mathbb{R}$ , auf denen eine Lösung des Anfangswertproblems

$$\begin{cases} \dot{x}(t) = f(t, x(t)), \\ x(t_0) = x_0 \end{cases} \quad (4.9)$$

existiert. Dann existiert auch eine Lösung  $x : I_{\max} \rightarrow \mathbb{R}^d$  des Anfangswertproblems (4.9).

*Beweis.* Wenn  $I_1$  und  $I_2$  zwei Intervalle sind, auf denen eine Lösung unseres Anfangswertproblems existiert, dann folgt aus dem globalen Eindeutigkeitsatz in Korollar 4.1.11, dass die beiden Lösungen des Anfangswertproblems, die auf  $I_1$  beziehungsweise  $I_2$  definiert sind, auf dem Durchschnitt  $I_1 \cap I_2$  übereinstimmen.

Deshalb können wir eine Funktion  $x : I_{\max} \rightarrow \mathbb{R}^d$  folgendermaßen definieren: Für  $t \in I_{\max}$  wählen wir ein Intervall  $I_1 \subseteq I_{\max}$ , das  $t$  enthält, und auf dem eine Lösung  $x_1 : I_1 \rightarrow \mathbb{R}^d$  von (4.9) existiert; dann setzen wir  $x(t) := x_1(t)$ . Wegen unserer Beobachtung aus dem vorangehenden Abschnitt hängt die Definition von  $x(t)$  nicht von der Wahl des Intervalls  $I_1$  ab.

Daraus folgt, dass  $x$  auf jedem Intervall  $I_1 \subseteq I_{\max}$ , auf dem eine Lösung  $x_1$  des Anfangswertproblems existiert, mit  $x_1$  übereinstimmt. Dies impliziert, dass  $x$  selbst die Differentialgleichung und die Anfangsbedingung erfüllt, also selbst eine Lösung des Anfangswertproblems ist.  $\square$

Es folgt sofort aus der Definition von  $I_{\max}$  in obiger Proposition, dass  $I_{\max}$  jedes andere Intervall, auf dem eine Lösung des Anfangswertproblems existiert, enthält. Somit ist  $I_{\max}$  das *größte* Intervall, auf dem eine Lösung von (4.9) existiert; dies begründet die folgende Begriffsbildung:

**Definition 4.3.3 (Maximales Existenzintervall).** Unter den Voraussetzungen von Proposition 4.3.2 nennen wir das Intervall  $I_{\max}$  das *maximale Existenzintervall* des Anfangswertproblems (4.9).

Man beachte, dass wir zur Konstruktion des maximalen Existenzintervalls die lokale Lipschitzbedingung an die Funktion  $f$  auf der rechten Seite der Differentialgleichung gefordert haben: Wenn diese Bedingung nicht erfüllt ist, haben wir im Allgemeinen keine Eindeutigkeit der Lösungen, und somit würde der Wert  $x(t)$  im Beweis von Proposition 4.3.2 von der Wahl des Intervalls  $I_1$  abhängen, dass zu seiner Definition verwendet wurde.

Aus der Tatsache, dass wir zwei Lösungen, deren Definitionsbereiche sich in einem Punkt schneiden, verkleben können (Proposition 4.3.1) folgt, dass das maximale Existenzintervall immer offen ist:

**Proposition 4.3.4.** *Unter den Voraussetzungen von Proposition 4.3.2 ist das maximale Existenzintervall  $I_{\max}$  des Anfangswertproblems (4.9) offen.*

*Beweis.* Sei  $x : I_{\max} \rightarrow \mathbb{R}^d$  die Lösung des Anfangswertproblems (4.9), die auf ganz  $I_{\max}$  definiert ist.

Wir betrachten zunächst die rechte Grenze  $t_+$  von  $I_{\max}$ : Wir nehmen widerspruchshalber an, dass diese endlich und in  $I_{\max}$  enthalten ist. Nun „starten wir die Differentialgleichung im Punkt  $t_+$  neu“, d.h. wir definieren

$x_+ := x(t_+)$  und betrachten das Anfangswertproblem

$$\begin{cases} \dot{\tilde{x}}(t) = f(t, \tilde{x}(t)), \\ \tilde{x}(t_+) = x_+. \end{cases} \quad (4.10)$$

Dieses Anfangswertproblem besitzt laut des lokalen Existenzsatzes in Korollar 4.1.12 eine Lösung  $\tilde{x}$ , die in einer Umgebung von  $t_+$  definiert ist. Wir schränken die Lösung  $\tilde{x}$  nun auf die rechtsseitige Umgebung von  $t_+$  ein (inklusive des Zeitpunktes  $t_+$  selbst). Dann sind  $x$  und  $\tilde{x}$  zwei Lösungen der zu  $f$  gehörenden Differentialgleichung und stimmen im Punkt  $t_+$  überein. Also können wir sie laut Proposition 4.3.1 zu einer neuen Lösung  $\hat{x}$  der Differentialgleichung verkleben.

Dann löst aber  $\hat{x}$  das Anfangswertproblem (4.9) und ist auf einem größeren Intervall als  $I_{\max}$  definiert. Widerspruch.

Am linken Rand des Intervalls  $I_{\max}$  kann man analog argumentieren.  $\square$

Für den linken und rechten Randpunkt des maximalen Existenzintervalls führen wir die folgende Notation ein:

**Notation 4.3.5 (Endpunkte des maximalen Existenzintervalls).** Unter den Voraussetzungen von Proposition 4.3.2 bezeichnen wir den linken Endpunkt von  $I_{\max}$  mit  $t_-$  und den rechten Endpunkt mit  $t_+$ .

Wenn wir die Abhängigkeit der Objekte  $I_{\max}$ ,  $t_-$  und  $t_+$  von den Anfangsdaten  $(t_0, x_0)$  in der Notation explizit zum Ausdruck bringen wollen, dann schreiben wir stattdessen auch  $I_{\max}(t_0, x_0)$ , sowie  $t_-(t_0, x_0)$  und  $t_+(t_0, x_0)$ .

Man beachte, dass stets  $-\infty \leq t_- < t_0 < t_+ \leq \infty$  gilt, und dass wir in Proposition 4.3.4 gezeigt haben, dass keiner der beiden Punkte  $t_-$  und  $t_+$  jemals in  $I_{\max}$  enthalten ist.

Nun kommen wir zu des Pudels Kern: Der nächste Satz gibt eine genaue Klassifikation des möglichen Verhaltens der Lösung an den Randpunkten des maximalen Existenzintervalls an: Bei  $t_-$  und  $t_+$  gibt es jeweils nur drei Möglichkeiten, wie die Lösung sich verhalten kann. In bestimmten Situationen kann man zwei der drei Möglichkeiten ausschließen, und somit zum Beispiel die Existenz von Lösungen auf ganz  $\mathbb{R}$  zeigen (siehe das Beispiel 4.3.10 und den nachfolgenden Abschnitt 4.4 für Details).

**Satz 4.3.6 (Charakterisierung des Randverhaltens von Lösungen).** *Es seien die Voraussetzungen der Proposition 4.3.2 erfüllt und es sei  $x : I_{\max} \rightarrow \mathbb{R}^d$  die Lösung des Anfangswertproblems (4.9), die auf ganz  $I_{\max}$  definiert ist. Dann gilt genau eine der folgenden zwei Aussagen für die rechte Grenze  $t_+$  von  $I_{\max}$ :*

- (a) *Es ist  $t_+ = \infty$  (d.h. die Lösung existiert unendlich lange).*

- (b) Es ist  $t_+ < \infty$  und für jedes kompakte Teilmenge  $K \subseteq G$  gibt es ein  $t_1 \in [t_0, t_+)$  mit  $(t, x(t)) \notin K$  für alle  $t \in [t_1, t_+)$  (d.h. die Lösung existiert nur endlich lange und das Tupel aus Zeit und Lösung verlässt jede kompakte Teilmenge von  $G$  schließlich).

Ein analoge Aussage gilt am linken Ende  $t_-$  von  $I_{\max}$ .

Bitte beachten Sie, dass die Bedingung  $t_+ < \infty$  in Aussage (b) nötig ist, damit die Aussagen (a) und (b) sich gegenseitig ausschließen (es kann nämlich durchaus vorkommen, dass  $t_+ = \infty$  ist und die Lösung trotzdem jede kompakte Teilmenge von  $G$  schließlich verlässt); würden wir die Bedingung  $t_+ < \infty$  in (b) weglassen, könnten wir also noch immer sagen, dass stets mindestens eine der beiden Aussagen (a) und (b) gilt, aber wir könnten nicht mehr sagen, dass stets *genau* eine der beiden Aussagen gilt.

Zum Beweis des Satzes benötigen wir das folgende einfache Resultat über Differenzierbarkeit:

**Lemma 4.3.7.** *Wir betrachten ein kompaktes Intervall  $[a, b] \subseteq \mathbb{R}$  mit  $a < b$ . Sei  $d \in \mathbb{N}$  und sei  $x : [a, b] \rightarrow \mathbb{R}^d$  stetig auf  $[a, b]$  und stetig differenzierbar auf  $[a, b)$ . Wenn der Limes  $d := \lim_{t \uparrow b} \dot{x}(t)$  in  $\mathbb{R}$  existiert, dann ist  $x$  auch in  $b$  differenzierbar mit (linksseitiger) Ableitung  $\dot{x}(b) = d$ .*

*Beweis.* Wir definieren eine Funktion  $y : [a, b] \rightarrow \mathbb{R}^d$  durch

$$y(t) = \begin{cases} \dot{x}(t) & \text{für } t \in [a, b), \\ d & \text{für } t = b. \end{cases}$$

Dann ist  $y$  nach Voraussetzung stetig. Also besitzt  $y$  laut Hauptsatz der Differential- und Integralrechnung eine stetig differenzierbare Stammfunktion  $\tilde{x}$  auf  $[a, b]$ . Wir können die Integrationskonstante der Stammfunktion so wählen, dass  $x(a) = \tilde{x}(a)$  gilt.

Die Funktion  $x - \tilde{x}$  ist stetig auf  $[a, b]$  und stetig differenzierbar mit Ableitung 0 auf  $[a, b)$ . Also ist  $x - \tilde{x}$  auf  $[a, b)$  – und aus Stetigkeitsgründen dann auch auf  $[a, b]$  – konstant. Die beiden Funktionen  $x$  und  $\tilde{x}$  stimmen aber im Punkt  $a$  überein, also sind sie gleich.

Somit ist  $x$  gleich der Funktion  $\tilde{x}$ , die auf ganz  $[a, b]$  stetig differenzierbar ist; insbesondere ist  $x$  im Punkt  $b$  differenzierbar. Für die Ableitung von  $x$  im Punkt  $b$  erhalten wir

$$\dot{x}(b) = \dot{\tilde{x}}(b) = y(b) = d,$$

womit die Behauptung bewiesen ist. □

Nun können wir Satz 4.3.6 beweisen.

*Beweis von Satz 4.3.6.* Offensichtlich schließen sich die zwei Aussagen (a) und (b) gegenseitig aus, also kann höchstens eine der Aussagen gelten. Um zu zeigen, dass mindestens eine der beiden Aussagen gilt, nehmen wir widerspruchshalber an, dass beide Aussagen nicht erfüllt sind.

Dann ist  $t_+ < \infty$ , da (a) nicht erfüllt ist und es gibt eine kompakte Menge  $K \subseteq G$  derart, dass  $x(t)$  die Menge  $K$  nicht schließlich verlässt – d.h. wir finden eine Folgen  $t_n \uparrow t_+$  derart, dass  $(t_n, x(t_n)) \in K$  für jeden Index  $n$  gilt. Indem wir  $K$ , falls nötig, etwas vergrößern, können wir sogar ein  $\varepsilon > 0$  finden derart, dass jeder Vektor  $(t_n, x(t_n))$  in  $\mathbb{R}^{1+d}$  mindestens den (Euklidischen) Abstand  $\varepsilon$  vom Rand  $\partial K$  von  $K$  hat.

Wir gehen nun in zwei Schritten vor um einen Widerspruch zu erhalten

*Schritt 1:* Wir zeigen, dass sogar  $(t, x(t)) \in K$  für alle  $t$  gilt, die genügend nahe an  $t_+$  liegen: Dazu setzen wir

$$L := \sup \left\{ \left\| (t, f(t, y)) \right\| : (t, y) \in K \right\},$$

wobei  $\|\cdot\|$  hier die Euklidische Norm im  $\mathbb{R}^{1+d}$  bezeichnet. Weil  $f$  stetig und  $K$  kompakt ist, gilt  $L < \infty$ .

Betrachten wir nun eine Zeit  $t_{n_0}$  für ein solch großes  $n_0$ , dass  $t_{n_0}$  höchstes Abstand  $\frac{\varepsilon}{2L}$  zu  $t_+$  hat. Würde  $(t, x(t))$  irgendwann nach der Zeit  $t_{n_0}$  die kompakte Menge  $K$  wieder verlassen, dann könnten wir die größte Zeit  $u \in (t_{n_0}, t_+)$  wählen, für welche  $(t, x(t))$  auf dem ganzen Zeitintervall  $[t_{n_0}, u]$  in  $K$  liegt. Dann muss  $(u, x(u))$  auf dem Rand von  $K$  liegen, also gilt

$$\left\| (u, x(u)) - (t_{n_0}, x(t_{n_0})) \right\| \geq \varepsilon.$$

Zugleich ist aber

$$(u, x(u)) - (t_{n_0}, x(t_{n_0})) = \int_{t_{n_0}}^u (1, \dot{x}(s)) \, ds = \int_{t_{n_0}}^u (1, f(s, x(s))) \, ds,$$

und somit

$$\begin{aligned} \left\| (u, x(u)) - (t_{n_0}, x(t_{n_0})) \right\| &\leq \int_{t_{n_0}}^u \left\| (1, f(s, x(s))) \right\| \, ds \\ &\leq \int_{t_{n_0}}^u L \, ds = (u - t_{n_0})L \leq (t_+ - t_{n_0})L \leq \frac{\varepsilon}{2}. \end{aligned}$$

Dies ist ein Widerspruch; also muss wie behauptet  $(t, x(t)) \in K$  für alle  $t \in [t_{n_0}, t_+)$  gelten.

*Schritt 2:* Nun zeigen wir, dass sich die Lösung  $x$  in den Punkt  $t_+$  fortsetzen lässt und auf  $I_{\max} \cup \{t_+\}$  noch immer eine Lösung des Anfangswertproblems ist;

dies widerspricht der Maximalität von  $I_{\max}$ , womit der Beweis dann komplett ist.

Wegen Schritt 1 gilt  $(t, x(t)) \in K$  für alle  $t \in [t_{n_0}, t_+)$ . Wir verwenden die Konstante  $L$  aus Schritt 1 weiter und beachten dabei, dass  $\|f(t, y)\| \leq L$  für alle  $(t, y) \in K$  gilt.

Für zwei Zeiten  $t_1, t_2 \in [t_{n_0}, t_+)$  ist stets

$$x(t_2) - x(t_1) = \int_{t_1}^{t_2} \dot{x}(s) \, ds = \int_{t_1}^{t_2} f(s, x(s)) \, ds,$$

und somit, falls  $t_1 < t_2$  ist,

$$\|x(t_2) - x(t_1)\| \leq \int_{t_1}^{t_2} \|f(s, x(s))\| \, ds \leq \int_{t_1}^{t_2} L \, ds = (t_2 - t_1)L.$$

(Das war eine sehr ähnliche – aber etwas einfachere – Rechnung, wie wir sie schon in Schritt 1 verwendet haben.) Dies impliziert, dass für jede Folge  $(s_n)$  in  $I_{\max}$  mit  $s_n \uparrow t_+$  die Folge  $(x(s_n))$  eine Cauchy-Folge ist und somit konvergiert. Daraus können wir schließen, dass  $x(t)$  für  $t \uparrow t_+$  gegen einen Vektor  $x_1 \in K \subseteq G$  konvergiert.

Wir können also  $x$  im Punkt  $t_+$  durch  $x_1$  fortsetzen und erhalten somit eine stetige Funktion, die auch im Punkt  $t_+$  definiert ist (und die wir der Einfachheit halber wieder mit  $x$  bezeichnen). Weil für  $t < t_+$  die Differentialgleichung  $\dot{x}(t) = f(t, x(t))$  gilt und  $f$  stetig ist, folgt  $\dot{x}(t) \rightarrow f(t_+, x(t_+))$  für  $t \uparrow t_+$ .

Die Ableitung von  $x$  ist also stetig auf  $I$  und konvergiert für  $t \uparrow t_+$  gegen den Wert  $f(t_+, x(t_+))$ . Somit sagt uns das Lemma 4.3.7, dass die (fortgesetzte) Funktion  $x$  sogar auf  $I_{\max} \cup \{t_+\}$  differenzierbar ist und die Differentialgleichung erfüllt. Dies widerspricht, wie am Beginn von Schritt 2 angekündigt, der Maximalität von  $I_{\max}$ .  $\square$

Zur Untersuchung konkreter Differentialgleichungen (siehe beispielsweise das unten folgende Beispiel 4.3.10) ist es praktisch, die Aussage von Satz 4.3.6 etwas abzuwandeln, derart, dass man drei Alternativen für das Verhalten der Lösung bei  $t_+$  erhält; dies ist der Zweck des folgenden Korollars.

**Korollar 4.3.8 (Charakterisierung des Randverhaltens von Lösungen).** *Es seien die Voraussetzungen der Proposition 4.3.2 erfüllt und es sei  $x : I_{\max} \rightarrow \mathbb{R}^d$  die Lösung des Anfangswertproblems (4.9), die auf ganz  $I_{\max}$  definiert ist. Dann gilt genau eine der drei folgenden Aussagen für die rechte Grenze  $t_+$  von  $I_{\max}$ :*

- (a) *Es ist  $t_+ = \infty$  (d.h. die Lösung existiert unendlich lange).*

- (b) Es ist  $t_+ < \infty$  und  $\liminf_{t \uparrow t_+} \text{dist}((t, x(t)), \partial G) = 0$  (d.h. die Lösung kommt dem Rand des Definitionsbereichs von  $f$  beliebig nahe).<sup>7</sup>
- (c) Es ist  $t_+ < \infty$ , sowie  $\liminf_{t \uparrow t_+} \text{dist}((t, x(t)), \partial G) > 0$  und  $\lim_{t \uparrow t_+} \|x(t)\| = \infty$  (d.h. die Lösung bleibt vom Rand von  $G$  entfernt, aber es gibt einen sogenannten Blow-up in endlicher Zeit).

Ein analoge Aussage gilt am linken Ende  $t_-$  von  $I_{\max}$ .

*Beweis.* Offensichtlich schließen sich die drei Aussagen (a), (b) und (c) gegenseitig aus, also kann höchstens eine der Aussagen gelten. Um zu zeigen, dass mindestens eine der drei Aussagen gilt, nehmen wir widerspruchshalber an, dass alle drei Aussagen nicht erfüllt sind.

Dann ist  $t_+ < \infty$ , da (a) nicht erfüllt ist, und es ist

$$\liminf_{t \uparrow t_+} \text{dist}((t, x(t)), \partial G) = 0,$$

da (b) nicht erfüllt ist. Weil zudem auch (c) nicht erfüllt ist, gilt

$$\liminf_{t \uparrow t_+} \|x(t)\| < \infty.$$

Somit finden wir eine monoton wachsende Folge  $(t_n) \subseteq I_{\max}$  mit  $t_n \uparrow t_+$  derart, dass die Folge  $((t_n, x(t_n))) \subseteq G$  beschränkt ist und vom Rand  $\partial G$  wegbleibt. Also gibt es eine kompakte Menge  $K \subseteq G$  und eine Zahl  $\varepsilon > 0$  derart, dass alle Vektoren  $x(t_n)$  in  $K$  liegen. Dies widerspricht aber Satz 4.3.6.  $\square$

**Aufgabe 4.3.9.** Leiten Sie aus Korollar 4.3.8 eine Version derselben Aussage für autonome Differentialgleichungen her.

Damit Sie das Korollar 4.3.8 richtig verstehen, ist es wichtig, dass Sie diese Aufgabe lösen können. Sollten Sie Schwierigkeiten mit der Aufgabe haben, fragen Sie bitte unbedingt in der Übung noch einmal nach.

Wir schließen den Abschnitt mit einem einfachen Anwendungsbeispiel für Korollar 4.3.8:

**Beispiel 4.3.10.** Wir betrachten die logistische Gleichung für die Bakterienpopulation aus Beispiel 2.1.3: Gegeben sei eine Anfangspopulation  $b_0 \in (0, \infty)$  und Konstanten  $c, d > 0$ . Wir sehen uns das Anfangswertproblem

$$\begin{cases} \dot{b}(t) = (c - db(t))b(t), \\ b(0) = b_0 \end{cases} \quad (4.11)$$

<sup>7</sup>Hierbei bezeichnet für Punkt  $y \in \mathbb{R}^{1+d}$  und  $A \subseteq \mathbb{R}^{1+d}$  die Notation  $\text{dist}(y, A)$  den Abstand zwischen dem Punkt  $y$  und der Menge  $A$ , d.h. es ist  $\text{dist}(y, A) = \inf\{\|y - a\| : a \in A\}$ . Bitte beachten Sie, dass dies insbesondere  $\text{dist}(y, \emptyset) = \infty$  impliziert (da man üblicherweise  $\inf \emptyset = \infty$  definiert); dies ist konsistent mit der Anschauung, dass man „unendlich lange braucht, um vom Punkt  $y$  zu einem Punkt der leeren Menge zu gelangen“.

an. Dabei wollen wir beliebige Zeiten  $t$  zulassen, aber nur Werte für  $b(t)$ , die strikt positiv sind. Also wählen wir  $G := \mathbb{R} \times (0, \infty) \subseteq \mathbb{R}^2$  und setzen  $f : G \rightarrow \mathbb{R}$  gleich

$$f(t, y) = (c - dy)y$$

für alle  $(t, y) \in G$ .

Wir behaupten: Die Lösung des Anfangswertproblems (4.11) existiert unendlich lange in die Zukunft, d.h. es gilt  $t_+ = \infty$ .

*Beweis.* Die Funktion  $f : G \rightarrow \mathbb{R}$  ist stetig differenzierbar und erfüllt somit laut Proposition 4.1.8 die lokale Lipschitzbedingung. Somit können wir die Theorie aus den Abschnitten 4.1 und 4.3 anwenden.

Wir wissen, dass genau eine der drei Aussagen aus Korollar 4.3.8 erfüllt ist, und wir wollen zeigen, dass die erste Aussage erfüllt ist. Also genügt es zu beweisen, dass die anderen beiden Aussagen nicht erfüllt sein können.

Der Rand von  $G$  ist gleich  $\mathbb{R} \times \{0\}$ ; and der Differentialgleichung in (4.11) sieht man: Sobald  $b(t)$  kleiner als  $c/d$  wird, wird die Ableitung  $\dot{b}(t)$  positiv – d.h.,  $b(t)$  stetig sofort wieder an. Also kann sich  $b(t)$  für  $t \uparrow t_+$  nicht beliebig nahe an die Null annähern, und somit kann sich  $(t, b(t))$  für  $t \uparrow t_+$  nicht beliebig nahe dem Rand von  $G$  annähern. Möglichkeit (b) in Korollar 4.3.8 scheidet also aus.

Aber auch die Möglichkeit (c) scheidet aus: Sobald nämlich  $b(t)$  größer als  $c/d$  wird, sieht man an der Differentialgleichung in (4.11), dass die Ableitung  $\dot{b}(t)$  negativ wird – d.h.  $b(t)$  nimmt sofort wieder ab. Also kann  $b(t)$  für  $t \uparrow t_+$  nicht gegen  $\infty$  gehen.

Somit bleibt nur Möglichkeit (a), d.h. es ist  $t_+ = \infty$ . □

Die Argumentation in obigem Beispiel ist der besseren Anschaulichkeit halber bewusst etwas vage gehalten; insbesondere sind Argumente wie „sobald  $b(t)$  kleiner als  $c/d$  wird, wird die Ableitung positiv, und  $b(t)$  nimmt somit wieder zu“ mathematisch nicht wirklich präzise. Alle Argumente in obigem Beispiel können aber mathematisch präzisiert werden.

Beispiel 4.3.10 vermittelt eine erste Vorstellung davon, wie Korollar 4.3.8 verwendet werden kann, um in manchen konkreten Beispielen zu beweisen, dass  $t_+ = \infty$  ist. Im nächsten Abschnitt stellen wir einige etwas systematischere Vorgehensweisen vor, mit denen man in einigen Situationen beweisen kann, dass die Lösung eines gegebenen Anfangswertproblems unendlich lange existiert (d.h. dass  $t_+ = \infty$  gilt).

**Literaturstellen und verwandte Ergebnisse.** Ein etwas anderer Beweis der Beschreibung des Randverhaltens von Lösungen wird zum Beispiel in [PW19, Abschnitt 2.3] präsentiert. Anstelle des Argumentes, das wir in Schritt 2 des

Beweises von Satz 4.3.6 verwendet haben, wird in dieser Referenz eine geeignete quantitative Version des Existenzsatzes von Picard–Lindelöf verwendet, die in [PW19, Bemerkung 2.2.3] besprochen wird.

## 4.4 Globale Existenz

In diesem Abschnitt geben wir mehrere hinreichende Kriterien für *globale Existenz* der Lösungen von Differentialgleichungen (bzw. der zugehörigen Anfangswertprobleme) an. Um technische Schwierigkeiten zu vermeiden beschränken wir uns hierbei auf den wichtigen Fall, in dem der Definitionsbereich der rechten Seite  $f$  von der Form  $I \times \Omega$  ist, wobei  $\Omega \subseteq \mathbb{R}^d$  nicht-leer und offen ist und  $I \subseteq \mathbb{R}$  ein nicht-leeres offenes (und somit nicht-triviales) Intervall ist.

Für den Begriff *globale Existenz* geben wir keine präzise Definition an, weil man damit – je nach Kontext – geringfügig verschiedene Dinge meinen kann. Im Grundsatz geht es aber immer um die Größe des maximalen Existenzintervalls  $I_{\max}$  der Anfangswertprobleme, die zur Differentialgleichung

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

gehören. Es gilt natürlich stets  $I_{\max} \subseteq I$ , und mit der Frage nach *globaler Existenz* meint man, ob  $I_{\max} = I$  gilt. Manchmal interessiert man sich auch nur für *globale Existenz nach rechts* (oder kann nur diese zeigen); damit meint man die Frage, ob  $t_+$  gleich dem rechten Rand von  $I$  ist.

Wir stellen zwei Kriterien vor, die manchmal nützlich sind um globale Existenz zu beweisen: *Ljapunov-Funktionen* und *lineare Beschränktheit*.

### Ljapunov-Funktionen

Anschaulich gesprochen ist eine *Ljapunov-Funktion* einer Differentialgleichung eine reell-wertige Funktion, die auf dem Zustandsraum einer Differentialgleichung definiert ist und welche die Eigenschaft hat, dass sie entlang der Lösungen der Differentialgleichung monoton fallend ist.

Diese Eigenschaft wird in der folgenden Definition und der anschließenden Proposition präzisiert.

**Definition 4.4.1 (Ljapunov<sup>8</sup>-Funktion).** Seien  $I \subseteq \mathbb{R}$  und  $\Omega \subseteq \mathbb{R}^d$  nichtleer und offen, wobei  $I$  ein Intervall sei. Sei  $f : I \times \Omega \rightarrow \mathbb{R}^d$  stetig und erfülle die lokale Lipschitz-Bedingung.

<sup>8</sup>Benannt nach Aleksandr Michajlovič Ljapunov (1857–1918), russischer Mathematiker und Physiker

Eine stetige Funktion  $V : \Omega \rightarrow \mathbb{R}$  heißt *Ljapunov-Funktion* für die Differentialgleichung

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

falls für jedes nicht-triviale Intervall  $J \subseteq I$  und jede Lösung  $x : J \rightarrow \Omega$  dieser Differentialgleichung gilt: Die Funktion

$$V \circ x : J \ni t \mapsto V(x(t)) \in \mathbb{R}$$

ist monoton fallend<sup>9</sup>.

Man beachte in Definition 4.4.1, dass  $V$  nicht etwa auf  $I \times \Omega$ , sondern nur auf  $\Omega$  definiert ist. Wir haben die Definition einer Ljapunov-Funktion also so gewählt, dass sie nicht von der Zeit abhängt<sup>10</sup>.

Mit der folgenden Proposition lässt sich herausfinden, ob eine gegebene Funktion  $V$  eine Ljapunov-Funktion ist (allerdings ist die Proposition nur anwendbar, wenn  $V$  stetig differenzierbar ist):

**Proposition 4.4.2.** *Seien  $I, \Omega$  und  $f$  wie in Definition 4.4.1. Es sei  $V : \Omega \rightarrow \mathbb{R}$  stetig differenzierbar. Dann sind äquivalent:*

- (i)  $V$  ist eine Ljapunov-Funktion für die Differentialgleichung  $\dot{x}(t) = f(t, x(t))$ .
- (ii) Für alle  $t \in I$  und alle  $y \in \Omega$  gilt  $\langle \nabla V(y), f(t, y) \rangle \leq 0$ .

Hierbei bezeichnet  $\langle \cdot, \cdot \rangle$  das übliche Skalarprodukt auf  $\mathbb{R}^d$ .

*Beweis.* „(ii)  $\Rightarrow$  (i)“ Sei  $J \subseteq I$  ein nicht-triviales Intervall und  $x : J \rightarrow \Omega$  eine Lösung der Differentialgleichung  $\dot{x}(t) = f(t, x(t))$ . Dann ist  $x$  stetig differenzierbar; weil  $V$  nach Voraussetzung stetig differenzierbar ist, ist auch die Abbildung  $\varphi := V \circ x : J \ni t \mapsto V(x(t)) \in \mathbb{R}$  stetig differenzierbar, und ihre Ableitung ist laut Kettenregel gleich

$$\dot{\varphi}(t) = \langle \nabla V(x(t)), \dot{x}(t) \rangle = \langle V(x(t)), f(t, x(t)) \rangle \leq 0$$

für  $t \in J$ . Also ist  $\varphi$  monoton fallend.

„(i)  $\Rightarrow$  (ii)“ Wir nehmen an, dass (ii) falsch ist, und wir zeigen nun, dass dann auch (i) falsch ist. Weil (ii) falsch ist, gibt es ein  $t_0 \in I$  und ein  $y_0 \in \Omega$  mit der Eigenschaft

$$\langle \nabla V(y_0), f(t_0, y_0) \rangle > 0.$$

---

<sup>9</sup>Hier verstehen wir *monoton fallend* in folgendem Sinne: Eine Funktion  $\varphi : J \rightarrow \mathbb{R}$  heißt *monoton fallend*, wenn für alle  $t_1, t_2 \in J$  mit  $t_1 \leq t_2$  die Ungleichung  $\varphi(t_1) \geq \varphi(t_2)$  gilt; d.h. es muss im Falle  $t_1 < t_2$  nicht unbedingt die strikte Ungleichung  $\varphi(t_1) > \varphi(t_2)$  gelten.

<sup>10</sup>In Abschnitt 8.4 werden wir Ljapunov-Funktionen verwenden, um das Langzeitverhalten der Lösungen von Differentialgleichungen zu studieren (insbesondere die Konvergenz gegen sogenannte *Gleichgewichtspunkte*); dort werden wir sogar noch etwas restriktiver sein und Ljapunov-Funktionen nur für autonome Differentialgleichungen betrachten.

Aufgrund des lokalen Existenzsatzes in Korollar 4.1.12 gibt es ein nicht-triviales Intervall  $J \subseteq I$ , das  $t_0$  enthält, und eine Lösung  $x : J \rightarrow \Omega$  des Anfangswertproblems

$$\begin{cases} \dot{x}(t) = f(t, x(t)), \\ x(t_0) = y_0. \end{cases}$$

Wir betrachten nun wieder die Abbildung  $\varphi := V \circ x : J \ni t \mapsto V(x(t)) \in \mathbb{R}$ . Dieses ist stetig differenzierbar, und ihre Ableitung im Punkt  $t_0$  ist gleich

$$\dot{\varphi}(t_0) = \langle \nabla V(x(t_0)), \dot{x}(t_0) \rangle = \langle V(x(t_0)), f(t_0, x(t_0)) \rangle = \langle V(y_0), f(t_0, y_0) \rangle > 0.$$

Also ist  $\varphi$  in einer Umgebung des Zeitpunktes  $t_0$  nicht monoton fallend, d.h.  $V$  ist keine Ljapunov-Funktion.  $\square$

Wenn man eine Ljapunov-Funktion findet<sup>11</sup>, die am Rand von  $\Omega$  sowie „in der Nähe von unendlich“ sehr groß wird, dann kann man daraus globale Existenz der Lösungen einer Differentialgleichung für positive Zeiten folgern. Wir präzisieren und beweisen diese Aussage in folgendem Satz.

**Satz 4.4.3 (Globale Existenz nach rechts via Ljapunov-Funktion).** *Seien  $I \subseteq \mathbb{R}$  und  $\Omega \subseteq \mathbb{R}^d$  nicht-leer und offen, wobei  $I$  ein Intervall sei. Sei  $f : I \times \Omega \rightarrow \mathbb{R}^d$  stetig und erfülle die lokale Lipschitz-Bedingung. Angenommen, es gibt eine Ljapunov-Funktion  $V : \Omega \rightarrow \mathbb{R}$  mit den folgenden beiden Eigenschaften:*

(a) *Es gilt  $\lim_{\|y\| \rightarrow \infty} V(y) = \infty$ .*<sup>12</sup>

(b) *Es gilt  $\lim_{y \rightarrow \partial\Omega} V(y) = \infty$ .*<sup>13</sup>

*Für jedes  $t_0 \in I$  und jedes  $x_0 \in \Omega$  ist dann der rechte Randpunkt  $t_+$  des maximalen Existenzintervalls von*

$$\begin{cases} \dot{x}(t) = f(t, x(t)), \\ x(t_0) = x_0 \end{cases}$$

*gleich dem rechten Randpunkt von  $I$  – anders ausgedrückt: alle Lösungen der Differentialgleichung  $\dot{x}(t) = f(t, x(t))$  existieren global nach rechts.*

<sup>11</sup>Hier sollte man zwei Punkte beachten: (1) Für eine Differentialgleichung kann es durchaus mehr als eine Ljapunov-Funktion geben; (2) Für eine konkrete Differentialgleichung kann es bisweilen schwer sein, eine Ljapunov-Funktion zu finden. Häufig versucht man eine Ljapunov-Funktion zu finden, indem man ein intuitives Wissen über das jeweilige Fachgebiet benutzt, aus dem die Differentialgleichung stammt – in der Physik eignet sich zum Beispiel die Energie oft als Ljapunov-Funktion.

<sup>12</sup>Beachten Sie, dass diese Bedingung leer – also automatisch erfüllt – ist, wenn  $\Omega$  beschränkt ist.

<sup>13</sup>Beachten Sie, dass diese Bedingung leer – also automatisch erfüllt – ist, wenn  $\Omega = \mathbb{R}^d$  ist (also leeren Rand hat).

*Beweis.* Wir fixieren Anfangsdaten  $t_0 \in I$  und  $x_0 \in \Omega$ ; sei  $x : I_{\max} \rightarrow \Omega$  die Lösung des zugehörigen Anfangswertproblems, die auf dem maximalen Existenzintervall definiert ist.

Es bezeichne  $b \in (-\infty, \infty]$  den rechten Randpunkt von  $I$ . Dann gilt natürlich  $t_+ \leq b$ . Nun verwenden wir die Beschreibung des Verhaltens von  $x$  in der Nähe von  $t_+$  aus Korollar 4.3.8. Es muss eine der drei Möglichkeiten aus diesem Korollar auftreten.

Wenn Möglichkeit (a) auftritt, gilt  $\infty = t_+ \leq b \leq \infty$ , also gleicht  $t_+ = b = \infty$ .

Als nächstes nehmen wir an, dass der Fall (b) auftritt. Es ist

$$\partial(I \times \Omega) = (\bar{I} \times \partial\Omega) \cup (\partial I \times \bar{\Omega}).$$

Weil die Abbildung  $I_{\max} \ni t \mapsto V(x(t))$  monoton fallend ist, aber  $V$  am Rand von  $\Omega$  gegen  $\infty$  strebt, muss  $x$  für  $t \uparrow t_+$  vom Rand von  $\Omega$  wegbleiben. Also kann  $\liminf_{t \uparrow t_+} \text{dist}((t, x(t)), \partial(I \times \Omega)) = 0$  nur gelten, wenn  $t$  sich für  $t \uparrow t_+$  dem rechten Randpunkt  $b$  von  $I$  annähert. Somit ist  $t_+ = b$ .

Nun betrachten wir noch den letzten Fall (c). Wie im vorangehenden Fall beobachten wir, dass die Ljapunov-Funktion  $V$  entlang der Lösung  $x$  monoton fallend ist, und dass sie aber gegen  $\infty$  streben würde, wenn die Norm ihres Argumentes gegen  $\infty$  geht. Somit kann Fall (c) nicht eintreten.  $\square$

Als Beispiel für die Anwendung dieses Satzes betrachten wir noch einmal die logistische Gleichung; die globale Existenz nach rechts für die Lösungen dieser Gleichung hatten wir in Beispiel 4.3.10 schon gezeigt, indem wir argumentiert hatten, weshalb die Möglichkeiten (b) und (c) aus Korollar 4.3.8 für diese Gleichung nicht auftreten können. Nun zeigen wir, wie man mit Hilfe einer Ljapunov-Funktion zum selben Ergebnis gelangen kann:

**Beispiel 4.4.4.** Wir betrachten erneut die logistische Gleichung für die Bakterienpopulation aus Beispiel 2.1.3: Es seien eine Anfangspopulation  $b_0 \in (0, \infty)$  und Konstanten  $c, d > 0$  gegeben, und wir sehen uns erneut das Anfangswertproblem

$$\begin{cases} \dot{b}(t) = (c - db(t))b(t), \\ b(0) = b_0 \end{cases} \quad (4.12)$$

an. Wie in Beispiel 4.3.10 wählen wir  $G := \mathbb{R} \times (0, \infty) \subseteq \mathbb{R}^2$  – d.h. in der Notation von Satz 4.4.3 ist  $I = \mathbb{R}$  und  $\Omega = (0, \infty)$  – und setzen  $f : G \rightarrow \mathbb{R}$  gleich

$$f(t, y) = (c - dy)y$$

für alle  $t \in I = \mathbb{R}$  und alle  $y \in \Omega = (0, \infty)$ .

In Beispiel 4.3.10 hatten wir bereits gezeigt: Die Lösung des Anfangswertproblems (4.12) existiert unendlich lange in die Zukunft, d.h. es gilt  $t_+ = \infty$ . Wir wollen dies nun noch einmal mit Hilfe von Satz 4.4.3 zeigen.

*Beweis. Erster Versuch eine passende Ljapunov-Funktion zu finden:* Wir beobachten zunächst: Für  $0 < y < \frac{c}{d}$  ist  $f(t, y) > 0$  und für  $y > \frac{c}{d}$  ist  $f(t, y) < 0$  (jeweils für alle  $t \in \mathbb{R}$ ). Anders ausgedrückt: Das ein-dimensionale Vektorfeld

$$\Omega \ni y \mapsto (c - dy)y \in \mathbb{R}$$

zeigt immer zum Punkt  $\frac{c}{d}$  hin – d.h. wir würden anschaulich erwarten, dass die Lösungen der Differentialgleichung zum Punkt  $\frac{c}{d}$  hin verlaufen. Wenn wir also eine Funktion  $V : \Omega \rightarrow \mathbb{R}$  wählen, die am Punkt  $\frac{c}{d}$  ihr Minimum hat und mit steigender Entfernung von  $\frac{c}{d}$  wächst, dann sollte dies eine Ljapunov-Funktion sein.

Deshalb liegt es nahe, konkret  $V : \Omega \rightarrow \mathbb{R}$  durch  $V(y) = (c - dy)^2$  für alle  $y \in \Omega$  zu definieren. Diese Funktion ist stetig differenzierbar, also können wir Proposition 4.4.2 anwenden um nachzuprüfen, ob  $V$  tatsächlich eine Ljapunov-Funktion ist: Für jedes  $t \in I = \mathbb{R}$  und jedes  $y \in \Omega = (0, \infty)$  gilt

$$\langle \nabla V(y), f(t, y) \rangle = -2d(c - dy) \cdot (c - dy)y = -2d(c - dy)^2 y \leq 0;$$

d.h.  $V$  ist tatsächlich eine Ljapunov-Funktion.

Leider können wir mit diesem  $V$  aber nicht Satz 4.4.3 anwenden, denn dieses  $V$  erfüllt zwar Voraussetzung (a), aber nicht Voraussetzung (b) des Satzes.

*Zweiter Versuch eine passende Ljapunov-Funktion zu finden:* Wir versuchen nun, die Funktion  $V$  aus dem ersten Versuch geeignet zu „manipulieren“ derart, dass wir immer noch eine Ljapunov-Funktion erhalten, diese aber auch am Rand von  $\Omega$  gegen  $\infty$  strebt.

Es liegt nahe, die Funktion  $V$  aus dem ersten Versuch so zu verändern, dass wir die Funktion

$$\tilde{V} : \Omega \rightarrow \mathbb{R}, \quad \tilde{V}(y) = \frac{(c - dy)^2}{y} \text{ für } y \in \Omega$$

erhalten. Diese Funktion hat die Ableitung  $\nabla \tilde{V}(y) = \frac{-2d(c - dy)}{y} - 2 \frac{(c - dy)^2}{y^2}$  für  $y \in \Omega$ , und somit folgt für alle  $y \in \Omega$  und alle Zeiten  $t$

$$\langle \nabla \tilde{V}(y), f(t, y) \rangle = -2d(c - dy)^2 - 2(c - dy)^2 \frac{c - dy}{y} = -2c(c - dy)^2 \frac{1}{y} \leq 0.$$

Also ist  $\tilde{V}$  aufgrund von Proposition 4.4.2 tatsächlich eine Ljapunov-Funktion für unsere Differentialgleichung.

Zugleich gilt aber  $\lim_{y \rightarrow \infty} \tilde{V}(y) = \infty$  und  $\lim_{y \rightarrow 0} \tilde{V}(y) = \infty$ . Also können wir Satz 4.4.3 anwenden: Dieser besagt, dass  $t_+$  gleich dem rechten Randpunkt des Intervalls  $I = \mathbb{R}$  ist, d.h.  $t_+ = \infty$ .  $\square$

### Lineare Beschränktheit

Nun stellen wir noch ein Kriterium für globale Existenz vor, das insbesondere für *lineare Differentialgleichungen*, die wir in Kapitel 5 studieren werden, sehr nützlich ist. Wir nehmen hierbei an, dass die rechte Seite  $f$  auf einer Menge der Form  $I \times \mathbb{R}^d$  definiert, wobei  $I$  ein reelles Intervall ist – das bedeutet, das Kriterium funktioniert nur, wenn der Zustandsraum  $\Omega$  der Differentialgleichung gleich ganz  $\mathbb{R}^d$  ist. Wir verwenden den folgenden Begriff:

**Definition 4.4.5 (Lineare Beschränktheit).** Sei  $I \subseteq \mathbb{R}$  ein nicht-leeres und offenes Intervall. Sei  $f : I \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ . Wir nennen  $f$  *linear beschränkt*, wenn es zwei stetige Funktionen  $\alpha, \beta : I \rightarrow [0, \infty)$  gibt derart, dass die Abschätzung

$$\|f(t, y)\| \leq \alpha(t) + \beta(t)\|y\|$$

für alle  $t \in I$  und alle  $y \in \mathbb{R}^d$  gilt.

Die Wahl des Begriffs *linear beschränkt* kann man dadurch motivieren, dass für jedes feste  $t$  die obere Schranke ein Polynom ersten Grades in der Norm  $\|y\|$  ist – der Begriff *linear beschränkt* bezieht sich also auf das Wachstumsverhalten von  $f$  bzgl. der Variablen  $y$ .

Aus linearer Beschränktheit von  $f$  folgern wir nun ein Kriterium für globale Existenz von Lösungen:

**Satz 4.4.6 (Globale Existenz via linearer Beschränktheit).** Sei  $I \subseteq \mathbb{R}$  ein nicht-leeres und offenes Intervall. Sei  $f : I \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  stetig und erfülle die lokale Lipschitz-Bedingung.

Wenn  $f$  linear beschränkt ist, dann ist für jedes  $t_0 \in I$  und jedes  $x_0 \in \mathbb{R}^d$  das maximale Existenzintervall von

$$\begin{cases} \dot{x}(t) = f(t, x(t)), \\ x(t_0) = x_0 \end{cases}$$

gleich  $I$  – anders ausgedrückt: alle Lösungen der Differentialgleichung  $\dot{x}(t) = f(t, x(t))$  existieren global.

*Beweis.* Es sei  $I = (a, b)$ , wobei  $-\infty \leq a < b \leq \infty$  ist, und seien  $\alpha$  und  $\beta$  wie in Definition 4.4.5. Wir fixieren Anfangsdaten  $t_0 \in I$  und  $x_0 \in \mathbb{R}^d$  und bezeichnen mit  $x : I_{\max} \rightarrow \mathbb{R}^d$  die Lösung des zugehörigen Anfangswertproblems, die auf dem maximalen Existenzintervall definiert ist. Es gilt natürlich  $a \leq t_- < t_0 < t_+ \leq b$ , und wir müssen  $t_- = a$  und  $t_+ = b$  zeigen.

Wir beginnen am rechten Rand von  $I_{\max}$ , das heißt bei  $t_+$ . Lassen Sie uns widerspruchshalber annehmen, dass  $t_+ < b$  ist. Da alle Voraussetzungen aus Korollar 4.3.8 erfüllt sind, muss eine der drei Alternativen aus diesem Korollar

eintreten. Die Alternative (a) scheidet aber aus, da wir  $t_+ < b$  angenommen haben.

Ebenso scheidet Alternative (b) aus: Es gilt nämlich  $\partial(I \times \mathbb{R}^d) = (\partial I) \times \mathbb{R}^d$ , und somit  $\text{dist}((t, x(t)), \partial(I \times \mathbb{R}^d)) = b - t \geq b - t_+ > 0$  für alle  $t \in I_{\max}$ . Also muss Alternative (c) eintreten, d.h. es gilt  $\|x(t)\| \rightarrow \infty$  für  $t \uparrow t_+$ .

Weil  $[t_0, t_+]$  ein kompaktes Teilintervall von  $I$  ist (hier verwenden wir erneut, dass wir  $t_+ < b$  angenommen haben) und weil  $\alpha$  und  $\beta$  stetig sind, gibt es eine reelle Zahl  $M \geq 0$  mit der Eigenschaft  $\alpha(t), \beta(t) \leq M$  für alle  $t \in [t_0, t_+]$ . Wegen  $\|x(t)\| \rightarrow \infty$  für  $t \uparrow t_+$  gibt es zudem eine Zeit  $t_1 \in [t_0, t_+]$  mit der Eigenschaft  $\|x(t)\| \geq 1$  für alle  $t \in [t_1, t_+]$ .

Nun verwenden wir, dass  $x$  die Differentialgleichung löst. Für die stetige differenzierbare Funktion  $\varphi : [t_1, t_+) \ni t \mapsto \|x(t)\|^2 \in [1, \infty)$  erhalten wir somit die Abschätzung<sup>14</sup>

$$\begin{aligned} \dot{\varphi}(t) &= 2\langle x(t), \dot{x}(t) \rangle = 2\langle x(t), f(t, x(t)) \rangle \leq 2|\langle x(t), f(t, x(t)) \rangle| \\ &\leq 2\|x(t)\| \|f(t, x(t))\| \leq 2\|x(t)\| (\alpha(t) + \beta(t)\|x(t)\|) \\ &\leq 2\|x(t)\| (M + M\|x(t)\|) \leq 4M\|x(t)\|^2 = 4M\varphi(t). \end{aligned}$$

für alle  $t \in [t_1, t_+)$ , wobei wir in der letzten Abschätzung verwendet haben, dass  $1 \leq \|x(t)\|$  gilt. Wegen  $\varphi(t) \geq 1 > 0$  können wir durch  $\varphi(t)$  teilen und erhalten somit die Ungleichung

$$\frac{\dot{\varphi}(t)}{\varphi(t)} \leq 4M$$

für alle  $t \in [t_1, t_+)$ . Für jedes  $\tau \in [t_1, t_+)$  können wir diese Ungleichung von  $t_1$  bis  $\tau$  integrieren; damit erhalten wir

$$\int_{t_1}^{\tau} \frac{\dot{\varphi}(t)}{\varphi(t)} dt \leq 4M \int_{t_1}^{\tau} dt = 4M(\tau - t_1).$$

Andererseits können wir den Ausdruck, der in obiger Formel ganz links steht, berechnen: Mit der Substitution  $q = \varphi(t)$  erhalten wir

$$\int_{t_1}^{\tau} \frac{\dot{\varphi}(t)}{\varphi(t)} dt = \int_{\varphi(t_1)}^{\varphi(\tau)} \frac{1}{q} dq = \ln \left| \frac{\varphi(\tau)}{\varphi(t_1)} \right| = \ln \frac{\varphi(\tau)}{\varphi(t_1)};$$

den Betrag konnten wir in der letzten Gleichheit weglassen, weil  $\varphi(t)$  für alle  $t$  positiv ist. Damit haben wir insgesamt die Ungleichung

$$\ln \frac{\varphi(\tau)}{\varphi(t_1)} \leq 4M(\tau - t_1)$$

<sup>14</sup>Die Abschätzung  $\dot{\varphi}(t) \leq 4M\varphi(t)$  ist ein einfaches Beispiel für eine sogenannte *Differentialungleichung*. Eng verwandt damit sind *Integralungleichungen*; auf diese werden wir später im Laufe der Vorlesung zurückkommen – insbesondere, in Abschnitt 7.1, wo wir das sogenannte *Gronwall-Lemma* (Lemma 7.1.1) besprechen.

für alle  $\tau \in [t_1, t_+)$  gezeigt, und weil die reelle Exponentialfunktion monoton steigend ist, folgt hieraus

$$\varphi(\tau) \leq \varphi(t_1)e^{4M(\tau-t_1)}$$

für alle  $\tau \in [t_1, t_+)$ . Es gilt aber  $\varphi(\tau) = \|x(\tau)\|^2 \rightarrow \infty$  für  $\tau \uparrow t_+$ . Widerspruch. Also war unsere Annahme falsch, d.h. es muss  $t_+ = b$  gelten.

Dass auch  $t_- = a$  gilt, kann man entweder mit einem ähnlichen Argument wie oben beweisen, oder via sogenannter *Zeitumkehr* aus dem bereits gezeigten folgern. Die Details besprechen wir auf Übungsblatt 4.  $\square$

## 4.5 Lösungsfluss für autonome Differentialgleichungen

In diesem Abschnitt wollen wir noch kurz das Konzept des *Lösungsflusses* von Differentialgleichungen ansprechen. Wir betrachten hierzu autonome Differentialgleichungen (siehe Definition 3.4.1) erster Ordnung. Wir lassen das Anfangswertproblem im Zeitpunkt 0 starten und nehmen an, dass die Lösung für jedes solche Anfangswertproblem global nach rechts existiert.<sup>15</sup>

**Definition 4.5.1 (Lösungsfluss).** Sei  $I \subseteq \mathbb{R}$  ein offenes Intervall, das den Zeitpunkt 0 enthält, und dessen rechte Grenze gleich  $\infty$  ist. Sei  $\Omega \subseteq \mathbb{R}^d$  nicht-leer und offen; sei  $h : \Omega \rightarrow \mathbb{R}^d$  stetig, und die Abbildung

$$f : I \times \Omega \ni (t, y) \mapsto h(y) \in \mathbb{R}^d$$

erfülle die lokale Lipschitzbedingung.

Wir nehmen an, dass für jedes  $x_0 \in \Omega$  die Lösung des Anfangswertproblems

$$\begin{cases} \dot{x}(t) = h(x(t)), \\ x(0) = x_0 \end{cases} \quad (4.13)$$

global nach rechts existiert, d.h. dass  $t_+(0, x_0) = \infty$  gilt.<sup>16</sup>

Für jedes  $t \in [0, \infty)$  und jedes  $x_0 \in \Omega$  bezeichne  $\varphi_t(x_0) \in \Omega$  den Vektor, den die Lösung des Anfangswertproblems (4.13) zum Zeitpunkt  $t$  annimmt. Dann heißt die Familie von Abbildungen<sup>17</sup>  $(\varphi_t)_{t \in [0, \infty)}$  der *Lösungsfluss* der Differentialgleichung  $\dot{x}(t) = h(x(t))$ .

<sup>15</sup>Man kann den Begriff *Lösungsfluss* auch unter etwas allgemeineren Voraussetzungen definieren; die hier beschriebene Situation ist aber besonders anschaulich, und erspart uns einige Technikalitäten.

<sup>16</sup>Zur Erinnerung: Hinreichende Kriterien hierfür finden Sie in den Sätzen 4.4.3 und 4.4.6.

<sup>17</sup>Man beachte: Für jedes  $t \in [0, \infty)$  ist  $\varphi_t$  eine Abbildung von  $\Omega$  nach  $\Omega$ .

Unter den Voraussetzungen der obigen Definition existiert die Lösung des Anfangswertproblems auch dann global nach rechts, wenn wir einen anderen Startzeitpunkt  $t_0$  wählen. Außerdem folgt aus der Autonomie der Differentialgleichung eine bestimmte algebraische Eigenschaft des Lösungsflusses. Diese Eigenschaften präzisieren wir in folgendem Satz:

**Satz 4.5.2.** *Unter den Voraussetzungen von Definition 4.5.1 gilt:*

- (a) Für jedes  $t_0 \in I$  und jedes  $x_0 \in \Omega$  existiert die Lösung des Anfangswertproblems

$$\begin{cases} \dot{x}(t) = h(x(t)), \\ x(t_0) = x_0 \end{cases}$$

global nach rechts, und dessen Lösung ist zu jedem Zeitpunkt  $t \geq t_0$  ist gegeben durch  $\varphi_{t-t_0}(x_0)$ .

- (b) Für alle  $t_1, t_2 \in I$  gilt

$$\varphi_{t_2}(\varphi_{t_1}(x_0)) = \varphi_{t_2+t_1}(x_0) \quad \text{für alle } x_0 \in \Omega,$$

oder in anderen Worten:  $\varphi_{t_2} \circ \varphi_{t_1} = \varphi_{t_2+t_1}$ .

*Beweis.* (a) Dieses Zeitverschiebungs-Resultat sieht man folgendermaßen: Man rechnet direkt nach, dass die Abbildung

$$x : [t_0, \infty) \ni t \mapsto \varphi_{t-t_0}(x_0)$$

das angegebene Anfangswertproblem löst; es gilt nämlich  $x(t_0) = \varphi_0(x_0) = x_0$  und

$$\dot{x}(t) = \frac{d}{dt} \varphi_{t-t_0}(x_0) = h(\varphi_{t-t_0}(x_0)) = h(x(t)).$$

für alle  $t \in [t_0, \infty)$ ,<sup>18</sup> insbesondere ist der rechte Randpunkt des maximalen Existenzintervalls des Anfangswertproblems gleich  $\infty$ .

<sup>18</sup>Bitte beachten Sie, dass wir für die zweite Gleichheit die Autonomie der Differentialgleichung verwendet haben: würden wir eine nicht-autonome-Differentialgleichung mit rechter Seite  $f : I \times \Omega \rightarrow \mathbb{R}^d$  betrachten, so würden wir stattdessen

$$\dot{x}(t) = \frac{d}{dt} \varphi_{t-t_0}(x_0) = f(t-t_0, \varphi_{t-t_0}(x_0)) = f(t-t_0, x(t))$$

erhalten, und würden dann im letzten Ausdruck die Zeitverschiebung um  $-t_0$  im ersten Argument von  $f$  nicht mehr los. Deshalb stimmt der Satz in dieser Form nicht für nicht-autonome Differentialgleichungen.

(b) Wir fixieren  $t_1 \in [0, \infty)$ . Die Abbildung  $\psi : [0, \infty) \ni t \mapsto \varphi_t(\varphi_{t_1}(x_0)) \in \mathbb{R}^d$  löst nach Definition des Lösungsflusses das Anfangswertproblem

$$\begin{cases} \dot{x}(t) = h(x(t)), \\ x(0) = \varphi_{t_1}(x_0). \end{cases}$$

Durch Nachrechnen sieht man genau wie im Beweis von (a)<sup>19</sup>, dass auch die Abbildung  $\tilde{\psi} : [0, \infty) \ni t \mapsto \varphi_{t+t_1}(x_0)$  dieses Anfangswertproblem löst. Aufgrund des globalen Eindeutigkeitsatzes von Picard–Lindelöf (Korollar 4.1.11) gilt also  $\psi = \tilde{\psi}$ . Dies zeigt (b).  $\square$

**Literaturstellen und verwandte Ergebnisse.** Weitere Informationen zum Lösungsfluss einer autonomen Differentialgleichung finden Sie zum Beispiel in [PW19, Abschnitt 4.4] und in [Tes12, Abschnitt 6.2].

## 4.6 Ergänzungen

### Anfangswertprobleme vs. Randwertprobleme

Lassen Sie uns eine Differentialgleichung zweiter Ordnung betrachten – zum Beispiel die sehr einfache autonome und ein-dimensionale Differentialgleichung

$$\ddot{x}(t) = -\lambda x(t). \tag{4.14}$$

für eine Konstante  $\lambda \in \mathbb{R}$ . In dem man diese Gleichung in eine Differentialgleichung erster Ordnung umschreibt (siehe Abschnitt 3.2), sieht man, dass man als Anfangswert zu einem Zeitpunkt  $t_0$  sowohl den Wert  $x(t_0)$  als auch den Wert  $\dot{x}(t_0)$  festlegen muss, um eine eindeutige Lösung zu erhalten.

Nun kann man sich fragen, ob man anstelle von  $\dot{x}(t_0)$  nicht vielleicht auch die Ableitung  $\dot{x}(t_1)$  an einem anderen Zeit  $t_1 \neq t_0$  vorschreiben kann – oder ob es vielleicht sogar möglich ist, anstelle der Ableitung von  $x$  einfach den Funktionswert von  $x$  noch an einem weiteren Zeitpunkt  $t_1 \neq t_0$  vorzuschreiben.

Prinzipiell ist dies möglich – man erhält dann ein sogenanntes *Randwertproblem* anstelle eines Anfangswertproblems; der Name „Randwertproblem“ rührt daher, dass man in Anwendungen häufig  $t_0$  und  $t_1$  als Randpunkte eines Intervalls betrachtet und nach einer Lösung der Differentialgleichung sucht, die auf dem Intervall zwischen  $t_0$  und  $t_1$  definiert ist. Bezüglich solcher Randwertproblemen sollte man beachten:

---

<sup>19</sup>Auch in dieser Rechnung verwendet man wie in (a) die Autonomie der Differentialgleichung.

- Obwohl es sich hierbei gemäß unserer Definition 1.3.2 eigentlich auch noch um gewöhnliche Differentialgleichungen handelt (eben mit Randbedingungen anstelle von Anfangsbedingungen), ist das Verhalten von Randwertproblemen der Theorie der partiellen Differentialgleichungen deutlich näher als der Theorie der gewöhnlichen Differentialgleichungen. Deshalb werden Randwertprobleme für gewöhnlich den partiellen Differentialgleichungen zugerechnet, selbst wenn die gesuchte Funktion nur von einer Variablen abhängt.
- Die Existenz- und Eindeigkeitstheorie von Randwertproblemen ist – wie es generell für partielle Differentialgleichungen gilt – komplexer als für Anfangswertprobleme von gewöhnlichen Differentialgleichungen. Gibt man zum Beispiel die Werte  $x(0)$  und  $x(1)$  vor, so erhält man für die Differentialgleichung (4.14) tatsächlich eine eindeutig bestimmte Lösung – es sei denn,  $\lambda$  ist von der Form  $k^2\pi^2$  für eine natürliche Zahl  $k$ ; für diese speziellen Werte von  $\lambda$  erhält man keine Eindeigkeit<sup>20</sup>.

---

<sup>20</sup>Dieses auf den ersten Blick äußerst merkwürdige Verhalten lässt sich funktionalanalytisch mit der Spektraltheorie unbeschränkter linearer Operatoren zufriedenstellend erklären.



# Lineare Differentialgleichungen

## Fragen zum Einstieg.

- Denken Sie zurück an die *linearen Gleichungssysteme*, die Sie in der Lineare Algebra 1 kennengelernt haben. Welche allgemeinen Aussagen kann man über die Lösungsmenge solcher Gleichungen treffen?
- Sei  $V$  ein Vektorraum über  $\mathbb{R}$  und  $U \subseteq V$ . Wann nennt man  $U$  einen Untervektorraum von  $V$ ?
- Was haben Ebenen im  $\mathbb{R}^3$  mit Untervektorräumen des  $\mathbb{R}^3$  zu tun?
- Betrachten Sie zwei Proben – nennen wir sie  $A$  und  $B$  – derselben radioaktiven Substanz (beispielsweise C14, das bereits in Beispiel 1.1.1(c) angesprochen wurde). Wenn die Probe  $A$  zum Zeitpunkt 0 dreimal so viele Teilchen enthält wie die Probe  $B$ , enthält Probe  $A$  dann auch zu späteren Zeitpunkten immer drei mal so viele Teilchen wie Probe  $B$ ?

## 5.1 Der Begriff der linearen Differentialgleichung

Wir beginnen direkt mit der Definition des Begriffs der *linearen Differentialgleichung*. Wir definieren und behandeln in den Abschnitte 5.1 bis 5.6 lineare Differentialgleichungen erster Ordnung; in Abschnitt 5.7 befassen wir uns schließlich mit linearen Differentialgleichungen höherer Ordnung.

**Definition 5.1.1 (Lineare Differentialgleichung).** Sei  $\emptyset \neq G \subseteq \mathbb{R}^{1+d}$  offen und sei  $f : G \rightarrow \mathbb{R}^d$ . Wir nennen die Differentialgleichung erster Ordnung

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

*linear*, wenn  $G$  von der Form  $G = I \times \mathbb{R}^d$  für ein (nicht-leeres und offenes) Intervall  $I \subseteq \mathbb{R}$  ist und wenn es Abbildungen  $A : I \rightarrow \mathbb{R}^{d \times d}$  und  $b : I \rightarrow \mathbb{R}^d$  gibt derart, dass

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

für alle  $t \in I$  und alle  $y \in \mathbb{R}^d$  gilt.

In anderen Worten: Eine lineare Differentialgleichung erster Ordnung ist eine Differentialgleichung auf  $I \times \mathbb{R}^d$  (für ein nicht-leeres offenes Intervall  $I \subseteq \mathbb{R}$ ), die von der Form

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

für eine Matrix-wertige Funktion  $A : I \rightarrow \mathbb{R}^{d \times d}$  und eine Vektor-wertige Funktion  $b : I \rightarrow \mathbb{R}^d$  ist.

**Bemerkung 5.1.2.** Angenommen, wir sind in der Situation von Definition 5.1.1 und die gegebene Differentialgleichung ist linear, wie in dieser Definition beschreiben. Dann kann man sich leicht überlegen:

- (a) Die Funktionen  $A : I \rightarrow \mathbb{R}^{d \times d}$  und  $b : I \rightarrow \mathbb{R}^d$  sind durch  $f$  eindeutig bestimmt.
- (b) Die Funktionen  $A$  und  $b$  sind genau dann stetig, wenn  $f$  stetig ist.
- (c) Wenn  $A$  und  $b$  stetig sind, dann erfüllt  $f$  die lokale Lipschitzbedingung.

**Aufgabe 5.1.3.** Beweisen Sie die drei Aussagen in Bemerkung 5.1.2.

Aus Aussage (c) in Bemerkung 5.1.2 folgt, dass die Lösung eines Anfangswertproblems, das zu einer linearen Differentialgleichung gehört, stets global eindeutig ist, falls die Funktionen  $A$  und  $b$  stetig sind (vgl. Korollar 4.1.11). Die Aussage (b) in Bemerkung 5.1.2 rechtfertigt außerdem die folgende weitere Begriffsbildung:

**Definition 5.1.4 (Homogene und inhomogene lineare Differentialgleichungen<sup>1</sup>).** Sei in der Situation von Definition 5.1.1 die gegebene Differentialgleichung linear und seien alle Bezeichnungen so wie in dieser Definition.

- (i) Die Funktion  $b : I \rightarrow \mathbb{R}^d$  heißt die *Inhomogenität* der Differentialgleichung.
- (ii) Die Differentialgleichung heißt *homogen*, wenn  $b$  konstant gleich 0 ist (d.h. wenn  $b(t) = 0$  für alle  $t \in I$  gilt). Andernfalls heißt sie *inhomogen*.
- (iii) Die Differentialgleichung

$$\dot{x}(t) = Ax(t)$$

auf  $I \times \mathbb{R}^d$  heißt die zur Differentialgleichung

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

gehörende *homogene Differentialgleichung*.

Wir schließen diesen einführenden Abschnitt mit der folgenden Existenzaussage: Die Lösung jedes Anfangswertproblems, das zu einer linearen Differentialgleichung mit stetiger rechter Seite gehört, existiert global. Genauer:

---

<sup>1</sup>Manchmal benutzt man in den hier definierten Begriffen anstelle des Ausdrucks *lineare Differentialgleichung* auch den Ausdruck *lineares Differentialgleichungssystem*, oder kurz *lineares System*.

**Proposition 5.1.5 (Globale Existenz für lineare Differentialgleichungen).**

Sei  $I \subseteq \mathbb{R}$  ein nicht-leeres und offenes Intervall und seien  $A : I \rightarrow \mathbb{R}^{d \times d}$  und  $b : I \rightarrow \mathbb{R}^d$  stetig. Sei  $t_0 \in I$  und  $x_0 \in \mathbb{R}^d$ . Dann ist das maximale Existenzintervall des auf  $I \times \mathbb{R}^d$  definierten Anfangswertproblems

$$\begin{cases} \dot{x}(t) = A(t)x(t) + b(t), \\ x(t_0) = x_0 \end{cases}$$

gleich  $I$ .

*Beweis.* Wir bemerken zunächst: Die rechte Seite der Differentialgleichung erfüllt laut Bemerkung 5.1.2(c) die lokale Lipschitzbedingung; somit ist das maximale Existenzintervall des Anfangswertproblems überhaupt definiert (siehe Proposition 4.3.2 and Definition 4.3.3).

Wir definieren nun  $\alpha, \beta : I \rightarrow [0, \infty)$  durch

$$\alpha(t) = \|A(t)\| \quad \text{und} \quad \beta(t) = \|b(t)\|$$

für alle  $t \in I$ , wobei  $\|A(t)\|$  die Matrixnorm von  $A(t)$  bezeichnet, die durch die Euklidische Norm auf  $\mathbb{R}^d$  induziert wird (vgl. Abschnitt C.1 im Anhang für Details). Dann sind  $\alpha$  und  $\beta$  stetig, da  $A$  und  $b$  stetig sind. Zugleich gilt

$$\|A(t)y + b(t)\| \leq \beta(t) + \alpha(t)\|y\|$$

für alle  $t \in I$  und alle  $y \in \mathbb{R}^d$ ; d.h. die Funktion, die die rechte Seite unserer Differentialgleichung beschreibt, ist linear beschränkt<sup>2</sup>. Also folgt die globale Existenz der Lösungen aus Satz 4.4.6.  $\square$

## 5.2 Die Struktur des Lösungsraums

In diesem Abschnitt besprechen wir die Struktur der Menge aller Lösungen – d.h. des sogenannten *Lösungsraumes* – einer linearen Differentialgleichung. In der folgenden Definition 5.2.1 wird hierzu zunächst die Differentialgleichung ohne Anfangsdaten betrachtet (weshalb es nicht nur eine Lösung, sondern eben einen ganzen Raum von Lösungen gibt). Der Zusammenhang mit Anfangswerten wird aber gleich darauf in Satz 5.2.2 hergestellt.

Für ein nicht-triviales Intervall  $I \subseteq \mathbb{R}$  bezeichnen wir mit

$$C^1(I; \mathbb{R}^d)$$

<sup>2</sup>Wobei hier die Bezeichnungen  $\alpha$  und  $\beta$  im Vergleich zu Definition 4.4.5 vertauscht sind.

den Vektorraum aller stetig differenzierbaren Funktionen von  $I$  nach  $\mathbb{R}^d$ ; wir statteten diesen im Moment nicht mit einer Norm aus<sup>3</sup>.

**Definition 5.2.1 (Lösungsraum einer linearen Differentialgleichung).** Sei  $I \subseteq \mathbb{R}$  ein nicht-leeres und offenes Intervall, und seien  $A : I \rightarrow \mathbb{R}^{d \times d}$  und  $b : I \rightarrow \mathbb{R}^d$  stetig. Die Menge aller Funktionen  $x : I \rightarrow \mathbb{R}^d$ , die die Differentialgleichung

$$\dot{x}(t) = A(t)x(t) + b(t) \quad (5.1)$$

lösen, bezeichnen wir als den *Lösungsraum* der Differentialgleichung (5.1).

Der Lösungsraum von (5.1) ist stets eine Teilmenge von  $C^1(I; \mathbb{R}^d)$ ; außerdem folgt aus Proposition 5.1.5, dass der Lösungsraum nie leer ist. Der folgende Satz beschreibt die genaue Struktur des Lösungsraums für homogene Differentialgleichungen. In der anschließenden Proposition beschreiben wir auch die Struktur des Lösungsraumes für inhomogene Differentialgleichungen.

**Satz 5.2.2.** Sei  $I \subseteq \mathbb{R}$  ein nicht-leeres und offenes Intervall, und sei  $A : I \rightarrow \mathbb{R}^{d \times d}$  stetig. Wir bezeichnen den Lösungsraum der homogenen linearen Differentialgleichung

$$\dot{x}(t) = A(t)x(t)$$

mit  $\mathcal{L}_{\text{hom}}$ . Sei außerdem  $t_0 \in I$  fest gewählt. Für jedes  $x_0 \in \mathbb{R}^d$  bezeichnen wir mit  $Sx_0 \in C^1(I; \mathbb{R}^d)$  die auf ganz  $I$  definierte Lösung des Anfangswertproblems<sup>4</sup>

$$\begin{cases} \dot{x}(t) = A(t)x(t), \\ x(t_0) = x_0. \end{cases}$$

Dann ist  $S : \mathbb{R}^d \rightarrow C^1(I; \mathbb{R}^d)$ ,  $x_0 \mapsto Sx_0$  eine injektive lineare Abbildung, deren Bild gleich  $\mathcal{L}_{\text{hom}}$  ist. Insbesondere ist  $\mathcal{L}_{\text{hom}}$  ein  $d$ -dimensionaler Untervektorraum von  $C^1(I; \mathbb{R}^d)$ .

In anderen Worten:  $S$  ist eine lineare Bijektion zwischen  $\mathbb{R}^d$  und dem  $d$ -dimensionalen Untervektorraum  $\mathcal{L}_{\text{hom}}$  von  $C^1(I; \mathbb{R}^d)$ .

<sup>3</sup>Wenn  $I$  kompakt ist, kann man diesen Raum, wenn man möchte, zum Beispiel mit der Norm  $\|\cdot\|_{C^1}$  ausstatten, die durch  $\|x\|_{C^1} = \max\{\|x\|_{\infty}, \|\dot{x}\|_{\infty}\}$  für alle  $x \in C^1(I; \mathbb{R}^d)$  gegeben ist – durch diese Norm wird  $C^1(I; \mathbb{R}^d)$  zu einem Banachraum. Allerdings betrachten wir im folgenden ohnehin meist offene Intervalle  $I$  – und in diesem Fall kann man diese Norm nicht verwenden, da stetige Funktionen auf einem offenen Intervall nicht beschränkt sein müssen.

<sup>4</sup>Solch eine Lösung gibt es aufgrund von Proposition 5.1.5.

*Beweis von Satz 5.2.2.* Wir zeigen zunächst, dass  $S$  linear ist: Seien  $x_0, \tilde{x}_0 \in \mathbb{R}^d$  und  $\lambda, \tilde{\lambda} \in \mathbb{R}$ . Dann rechnet man sofort nach, dass  $\lambda Sx_0 + \tilde{\lambda} S\tilde{x}_0$  das Anfangswertproblem

$$\begin{cases} \dot{x}(t) = A(t)x(t), \\ x(t_0) = \lambda x_0 + \tilde{\lambda} \tilde{x}_0 \end{cases}$$

löst; somit ist diese Funktion gleich  $S(\lambda x_0 + \tilde{\lambda} \tilde{x}_0)$ . Also ist  $S$  in der Tat linear.

Die Injektivität von  $S$  sieht man folgendermaßen: Sind  $x_0, \tilde{x}_0 \in \mathbb{R}^d$  und ist  $Sx_0 = S\tilde{x}_0$ , so sind die beiden letztgenannten Funktionen insbesondere am Punkt  $t_0$  gleich, also ist

$$x_0 = (Sx_0)(t_0) = (S\tilde{x}_0)(t_0) = \tilde{x}_0.$$

Offensichtlich ist das Bild von  $S$  in  $\mathcal{L}_{\text{hom}}$  enthalten. Ist umgekehrt  $x \in \mathcal{L}_{\text{hom}}$ , so gilt  $x = Sx(t_0)$ , also ist  $x$  im Bild von  $S$  enthalten. Somit ist gezeigt, dass das Bild von  $S$  gleich  $\mathcal{L}_{\text{hom}}$  ist.

Damit folgt aber auch, dass  $\mathcal{L}_{\text{hom}}$  ein Untervektorraum von  $C^1(I; \mathbb{R}^d)$  ist, denn das Bild einer linearen Abbildung ist immer ein Untervektorraum des Bildraumes. Weil die lineare Abbildung  $S$  von  $\mathbb{R}^d$  nach  $\mathcal{L}_{\text{hom}}$  bijektiv ist, hat  $\mathcal{L}_{\text{hom}}$  die gleiche Dimension wie  $\mathbb{R}^d$ , namentlich  $d$ .  $\square$

**Aufgabe 5.2.3.** Führen Sie den Schritt zu Beginn des Beweises von Satz 5.2.2, von dem behauptet wurde, man „rechnet [...] sofort nach“, explizit aus.

Nun beweisen wir eine Proposition, die den Zusammenhang zwischen dem Lösungsraum einer inhomogenen linearen Differentialgleichung und dem Lösungsraum der zugehörigen homogenen Differentialgleichung beschreibt.

**Proposition 5.2.4.** Sei  $I \subseteq \mathbb{R}$  eine nicht-leeres und offenes Intervall, und seien  $A : I \rightarrow \mathbb{R}^{d \times d}$  und  $b : I \rightarrow \mathbb{R}^d$  stetig. Es bezeichne  $\mathcal{L}_{\text{hom}}$  den Lösungsraum der homogenen Differentialgleichung

$$\dot{x}(t) = A(t)x(t)$$

und  $\mathcal{L}_{\text{inh}}$  den Lösungsraum der inhomogenen Differentialgleichung

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

Dann gilt für jedes  $x_{\text{inh}} \in \mathcal{L}_{\text{inh}}$  die Gleichheit  $\mathcal{L}_{\text{inh}} = x_{\text{inh}} + \mathcal{L}_{\text{hom}}$ .

Diese Proposition zeigt, dass der Lösungsraum der inhomogenen Differentialgleichung ein sogenannter *affiner Unterraum* von  $C^1(I; \mathbb{R}^d)$  ist – d.h. eine Teilmenge, die durch Verschiebung eines Untervektorraumes um einen Vektor entsteht. Dies ist völlig analog zur Situation für lineare Gleichungssysteme, die Sie aus der Linearen Algebra kennen.

Proposition 5.2.4 hat folgende Konsequenz: Wollen wir alle Lösungen einer inhomogenen linearen Differentialgleichung bestimmen, so genügt es alle Lösungen der zugehörigen homogenen Differentialgleichung sowie eine einzelne Lösung der inhomogenen Gleichung zu bestimmen.

*Beweis von Proposition 5.2.4.* Sei  $x_{\text{inh}} \in \mathcal{L}_{\text{inh}}$  fest. Wir müssen die Mengengleichheit  $\mathcal{L}_{\text{inh}} = x_{\text{inh}} + \mathcal{L}_{\text{hom}}$  zeigen.

„ $\subseteq$ “ Sei  $\tilde{x}_{\text{inh}}$  eine Funktion aus  $\mathcal{L}_{\text{inh}}$ . Dann kann man sofort nachrechnen, dass die Funktion  $x_{\text{hom}} := \tilde{x}_{\text{inh}} - x_{\text{inh}}$  die homogene Differentialgleichung  $\dot{x}(t) = Ax(t)$  löst, d.h. dass  $x_{\text{hom}} \in \mathcal{L}_{\text{hom}}$  ist. Hieraus erhalten wir  $\tilde{x}_{\text{inh}} = x_{\text{inh}} + x_{\text{hom}} \in x_{\text{inh}} + \mathcal{L}_{\text{hom}}$ .

„ $\supseteq$ “ Sei  $\tilde{x} \in x_{\text{inh}} + \mathcal{L}_{\text{hom}}$ ; dann können wir  $\tilde{x}$  schreiben als  $\tilde{x} = x_{\text{inh}} + x_{\text{hom}}$  für eine geeignete Funktion  $x_{\text{hom}} \in \mathcal{L}_{\text{hom}}$ . Mit Hilfe dieser Gleichheit kann man sofort nachrechnen, dass  $\tilde{x}$  die inhomogene Differentialgleichung löst, d.h. dass  $\tilde{x} \in \mathcal{L}_{\text{inh}}$  gilt.  $\square$

**Aufgabe 5.2.5.** Führen Sie diejenigen Schritte im Beweis von Proposition 5.2.4, von denen behauptet wurde, man könne sie „sofort nachrechnen“, explizit aus.

Wir schließen diesen Abschnitt mit einem einfachen Beispiel, das die Anwendung der vorangehenden Resultate illustriert.

**Beispiel 5.2.6.** Wir betrachten die ein-dimensionale autonome Differentialgleichung

$$\dot{x}(t) = 2x(t) + 1,$$

und als Zeiten lassen wir alle reellen Zahlen zu. Aus Proposition 5.1.5 wissen wir, dass wir für jede Anfangsbedingung das zugehörige maximale Existenzintervall gleich ganz  $\mathbb{R}$  ist. Den Lösungsraum  $\mathcal{L}_{\text{hom}}$  der zugehörigen homogenen Gleichung

$$\dot{x}(t) = 2x(t)$$

können wir mit Hilfe von Satz 5.2.2 bestimmen: Dazu fixieren wir zum Beispiel den Startzeitpunkt  $t_0 = 0$ . Für jedes  $x_0 \in \mathbb{R}$  ist die Lösung des Anfangswertproblems

$$\begin{cases} \dot{x}(t) = 2x(t), \\ x(0) = x_0 \end{cases}$$

gegeben durch  $\mathbb{R} \ni t \mapsto e^{2t}x_0 \in \mathbb{R}$ ; dies wissen Sie aus der Übung (Aufgabe 1 auf Blatt 1). Somit gilt

$$\mathcal{L}_{\text{hom}} = \{\mathbb{R} \ni t \mapsto e^{2t}x_0 \in \mathbb{R} : x_0 \in \mathbb{R}\};$$

dies ist ein ein-dimensionaler Untervektorraum von  $C^1(\mathbb{R}; \mathbb{R})$ . Ein Lösung  $x_{\text{inh}}$  der inhomogenen Differentialgleichung kann man mit etwas Geschick erraten: Zum Beispiel ist

$$x_{\text{inh}} : \mathbb{R} \ni t \mapsto -\frac{1}{2} \in \mathbb{R}$$

eine Lösung der inhomogenen Gleichung. Aus Proposition 5.2.4 erhalten wir somit, dass der Lösungsraum  $\mathcal{L}_{\text{inh}}$  der inhomogenen Gleichung durch

$$\mathcal{L}_{\text{inh}} = \{\mathbb{R} \ni t \mapsto e^{2t} x_0 - \frac{1}{2} \in \mathbb{R} : x_0 \in \mathbb{R}\}$$

gegeben ist.

Das obige Beispiel illustriert, wie unsere bisherigen Ergebnisse über lineare Differentialgleichungen in einer sehr einfachen Situation genutzt werden können um den Lösungsraum der Gleichung zu bestimmen. Wir werden im Laufe des Kapitels aber noch ausgefeiltere Techniken kennenlernen, mit denen man viele lineare Differentialgleichungen sehr detailliert untersuchen kann.

### 5.3 Homogene Systeme und Fundamentallösungen

Proposition 5.2.4 legt nahe, dass es beim Studium linearer Differentialgleichungen sinnvoll ist, zunächst die zugehörige homogene Gleichung zu untersuchen, und sich um die inhomogene Gleichung separat zu kümmern. Die homogene Gleichung werden wir in diesem und im nächsten Abschnitt betrachten; inhomogene Gleichungen studieren wir dann in Abschnitt 5.6.

Im kompletten Abschnitt 5.3 betrachten wir die folgende Situation:

*Sei  $I \subseteq \mathbb{R}$  ein nicht-leeres offenes Intervall und sei  $A : I \rightarrow \mathbb{R}^{d \times d}$  stetig. Wir betrachten die homogene lineare Differentialgleichung*

$$\dot{x}(t) = A(t)x(t). \tag{5.2}$$

Aus Satz 5.2.2 wissen wir, dass der Lösungsraum von (5.2) – nennen wir ihn wieder  $\mathcal{L}_{\text{hom}}$  – ein  $d$ -dimensionaler Untervektorraum von  $C^1(I; \mathbb{R}^d)$  ist. In der folgenden Proposition beschreiben wir, wie man erkennen kann, ob ein System aus  $d$  Funktionen in  $\mathcal{L}_{\text{hom}}$  linear unabhängig ist (und somit eine Basis von  $\mathcal{L}_{\text{hom}}$  bildet).

**Proposition 5.3.1.** *Seien  $x_1, \dots, x_d : I \rightarrow \mathbb{R}^d$  Lösungen von (5.2). Dann sind die folgenden Aussagen äquivalent:*

- (i) *Das System, das aus den Funktionen  $x_1, \dots, x_d$  besteht, ist linear unabhängig (und somit eine Basis von  $\mathcal{L}_{\text{hom}}$ ).*

- (ii) Für jedes  $t_0 \in I$  ist das System aus den Vektoren  $x_1(t_0), \dots, x_d(t_0)$  in  $\mathbb{R}^d$  linear unabhängig.
- (iii) Es gibt ein  $t_0 \in I$  derart, dass das System aus den Vektoren  $x_1(t_0), \dots, x_d(t_0)$  in  $\mathbb{R}^d$  linear unabhängig ist.

*Beweis.* „(ii)  $\Rightarrow$  (iii)“ Diese Implikation ist offensichtlich.

„(iii)  $\Rightarrow$  (i)“ Diese Implikation folgt sofort aus der Definition von linearer Unabhängigkeit.

„(i)  $\Rightarrow$  (ii)“ Dies ist die einzige nicht-triviale Implikation, und hier geht tatsächlich ein, dass wir es mit Lösungen einer Differentialgleichung zu tun haben:

Es gelte (i) und es sei  $t_0 \in I$  beliebig, aber fest. Seien  $\lambda_1, \dots, \lambda_d \in \mathbb{R}$  mit

$$\lambda_1 x_1(t_0) + \dots + \lambda_d x_d(t_0) = 0.$$

Dann löst die Funktion  $x := \lambda_1 x_1 + \dots + \lambda_d x_d : I \rightarrow \mathbb{R}^d$  das Anfangswertproblem

$$\begin{cases} \dot{x}(t) = A(t)x(t), \\ x(t_0) = 0. \end{cases}$$

Aber die konstante Null-Funktion ist offensichtlich ebenfalls eine Lösung dieses Anfangswertproblems. Wegen des globalen Eindeigkeitsatzes von Picard–Lindelöf (Korollar 4.1.11) folgt somit  $x = 0$ . Aber die Funktionen  $x_1, \dots, x_d$  sind laut (i) linear unabhängig; also folgt  $\lambda_1 = \dots = \lambda_d = 0$ .  $\square$

Wir führen nun einige Bezeichnungen ein, die uns die Formulierung der nachfolgenden Resultate in diesem Kapitel vereinfachen.

**Definition 5.3.2 (Fundamentalsysteme und Fundamentalmatrizen).** Es seien  $x_1, \dots, x_d : I \rightarrow \mathbb{R}^d$  Lösungen von (5.2).

- (i) Man bezeichnet die Abbildung

$$X : I \ni t \mapsto X(t) := (x_1(t), \dots, x_d(t)) \in \mathbb{R}^{d \times d}$$

als eine *Lösungsmatrix* von (5.2).

- (ii) Das System, das aus den Lösungen  $x_1, \dots, x_d$  besteht, heißt ein *Fundamentalsystem* von (5.2), falls es linear unabhängig ist.

In diesem Fall heißt die Lösungsmatrix  $X$  aus (a) eine *Fundamentalmatrix* von (5.2).

- (iii) Sei  $t_0 \in I$ . Man nennt das System, das aus den Lösungen  $x_1, \dots, x_d$  besteht, das *Hauptfundamentalsystem* von (5.2), das zum Startzeitpunkt  $t_0$  gehört, wenn

$$x_1(t_0) = e_1, \quad \dots, \quad x_d(t_0) = e_d$$

gilt.<sup>5,6</sup>

In diesem Fall nennt man  $X$  die *Hauptfundamentalmatrix* von (5.2), die zum Startzeitpunkt  $t_0$  gehört.

**Aufgabe 5.3.3.** Weshalb existiert für jedes fest gewählte  $t_0 \in I$  immer ein Hauptfundamentalsystem (bzw. eine Hauptfundamentalmatrix) zum Startzeitpunkt  $t_0$ ?

**Bemerkungen 5.3.4.** (a) Zu Teil (i) von Definition 5.3.2:

Dass  $X : I \rightarrow \mathbb{R}^{d \times d}$  eine Lösungsmatrix von (5.2) ist, kann man in Matrix-Schreibweise auch folgendermaßen ausdrücken: Es bedeutet, dass die Abbildung  $X$  differenzierbar ist und dass

$$\dot{X}(t) = A(t)X(t)$$

für alle  $t \in I$  gilt.

(b) Zu Teil (ii) von Definition 5.3.2:

Ein Fundamentalsystem von (5.2) ist also, anders gesprochen, schlichtweg eine Basis des Lösungsraumes von (5.2).

Laut Proposition 5.3.1 gilt: Eine Lösungsmatrix  $X$  ist eine Fundamentalmatrix genau dann, wenn die Matrix  $X(t_0)$  für jedes  $t_0 \in I$  invertierbar ist, genau dann, wenn die Matrix  $X(t_0)$  für mindestens ein  $t_0 \in I$  invertierbar ist.

(c) Zu Teil (iii) von Definition 5.3.2:

Die Bedingung  $x_1(t_0) = e_1, \dots, x_d(t_0) = e_d$  bedeutet anders ausgedrückt, dass  $X(t_0)$  die Einheitsmatrix ist.

Es ist sinnvoll von *dem* Hauptfundamentalsystem und *der* Hauptfundamentalmatrix zu sprechen, welche zum Zeitpunkt  $t_0$  gehören, denn diese sind aufgrund des globalen Eindeutigkeitsatzes von Picard–Lindelöf eindeutig bestimmt.

---

<sup>5</sup>Hierbei bezeichnen  $e_1, \dots, e_d \in \mathbb{R}^d$  die kanonischen Einheitsvektoren.

<sup>6</sup>Man beachte, dass ein Hauptfundamentalsystem aufgrund von Proposition 5.3.1 immer ein Fundamentalsystem ist.

Die folgenden Eigenschaften von Lösungsmatrizen (bzw. Fundamentalmatrizen) sind einfach, aber wichtig:

**Proposition 5.3.5.** Sei  $X : I \rightarrow \mathbb{R}^{d \times d}$  eine Lösungsmatrix von (5.2).

(a) Ist  $C \in \mathbb{R}^{d \times d}$ , so ist auch die Abbildung

$$XC : I \ni t \mapsto X(t)C \in \mathbb{R}^{d \times d}$$

eine Lösungsmatrix von (5.2).

(b) Sei  $C \in \mathbb{R}^{d \times d}$ . Die Abbildung  $XC$  ist eine Fundamentalmatrix genau dann, wenn  $X$  eine Fundamentalmatrix und  $C$  invertierbar ist.

(c) Sei  $t_0 \in I$  und sei  $X$  eine Fundamentalmatrix. Dann ist  $XX(t_0)^{-1}$  die Hauptfundamentalmatrix, die zum Startzeitpunkt  $t_0$  gehört.

(d) Sei  $t_0 \in I$  und sei  $X$  die Hauptfundamentalmatrix, die zum Startzeitpunkt  $t_0$  gehört. Für  $x_0 \in \mathbb{R}^d$  ist dann

$$x := Xx_0 : I \ni t \mapsto X(t)x_0 \in \mathbb{R}^d$$

die (eindeutig bestimmte) Lösung des Anfangswertproblems

$$\begin{cases} \dot{x}(t) = A(t)x(t), \\ x(t_0) = x_0. \end{cases} \quad (5.3)$$

*Beweis.* (a) Jede Spalte von  $XC$  ist eine Linearkombination der Spalten von  $X$  und löst somit die Differentialgleichung.

(b) Wir fixieren ein beliebiges  $t_0 \in I$ ; dann folgt aus Bemerkung 5.3.4(b):  $XC$  ist eine Fundamentalmatrix genau dann, wenn  $(XC)(t_0)$  invertierbar ist, genau dann, wenn  $X(t_0)C$  invertierbar ist, genau dann, wenn  $X(t_0)$  und  $C$  invertierbar sind, genau dann, wenn  $X$  eine Fundamentalmatrix ist und  $C$  invertierbar ist.

(c) Weil  $X(t_0)$  laut Bemerkung 5.3.4(b) invertierbar ist, ist die inverse Matrix  $X(t_0)^{-1}$  wohldefiniert. Wegen (b) ist  $XX(t_0)^{-1}$  eine Fundamentalmatrix, und offensichtlich ist  $XX(t_0)^{-1}$  zum Zeitpunkt  $t_0$  gleich der Einheitsmatrix.

(d) Weil  $X$  die Hauptfundamentalmatrix ist, die zum Startzeitpunkt  $t_0$  gehört, erfüllt  $x$  die Anfangsbedingung. Außerdem ist  $x$  eine Linearkombination der Spalten von  $X$  und somit eine Lösung der Differentialgleichung.  $\square$

**Beispiel 5.3.6 (Ein zwei-dimensionale autonome Differentialgleichung).** Lassen Sie uns die zwei-dimensionale autonome Differentialgleichung

$$\dot{x}(t) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} x(t) \quad (5.4)$$

betrachten, die auf dem Intervall  $I = \mathbb{R}$  definiert ist. Dann kann man nachrechnen, dass die Abbildung  $X : \mathbb{R} \rightarrow \mathbb{R}^{2 \times 2}$ , die durch

$$X(t) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$$

für alle  $t \in \mathbb{R}$  gegeben ist, eine Lösungsmatrix ist. Außerdem ist  $X(0)$  die Einheitsmatrix, und somit insbesondere invertierbar; also ist  $X$  die Hauptfundamentalmatrix, die zum Startzeitpunkt 0 gehört.

*Bemerkung:* In Aufgabe 16 auf Übungsblatt 4 haben Sie bereits die Familie von Abbildungen  $(X(t))_{t \in [0, \infty)}$  betrachtet<sup>7</sup> und nachgerechnet, dass diese der Lösungsfluss der Differentialgleichung (5.4) ist.

Die Aussage „ $X$  ist die Hauptfundamentalmatrix von (5.4), die zum Startzeitpunkt 0 gehört“, sagt im Grund nichts anderes aus – allerdings betrachten wir nun auch negative Zeiten.

**Bemerkungen 5.3.7.** Aussage (d) in Proposition 5.3.5 verdeutlicht die Bedeutung von Hauptfundamentalmatrizen:

- (a) Wenn man die Hauptfundamentalmatrix kennt, die zum Startzeitpunkt  $t_0$  gehört, dann kann man für jedes  $t_0 \in I$  sofort die Lösung des Anfangswertproblems (5.3) bestimmen.
- (b) Es gibt einen engen Zusammenhang zwischen Proposition 5.3.5(d) und Satz 5.2.2:

Ist  $X$  die Hauptfundamentalmatrix, die zum Startzeitpunkt  $t_0$  gehört, so ist die Abbildung

$$\mathbb{R}^d \ni x_0 \mapsto Xx_0 \in C^1(I; \mathbb{R}^d)$$

gerade die Abbildung  $S$  aus Satz 5.2.2.

- (c) Es gibt auch einen Zusammenhang zwischen Proposition 5.3.5(d) und dem Lösungsfluss für autonome Systeme, den wir in Abschnitt 4.5 besprochen haben:

Enthalte  $I$  das Intervall  $[0, \infty)$ , sei  $t_0 = 0$  und sei  $A(t)$  unabhängig von  $t$  (letztere Aussage bedeutet, dass die Differentialgleichung (5.2) autonom ist).

Sei  $X$  die Hauptfundamentalmatrix, die zum Startzeitpunkt  $t_0 = 0$  gehört; für jede Zeit  $t$  identifizieren wir die Matrix  $X(t) \in \mathbb{R}^{d \times d}$  in kanonischer Weise mit einer linearen Abbildung von  $\mathbb{R}^d$  nach  $\mathbb{R}^d$ . Dann ist  $(X(t))_{t \in [0, \infty)}$  der Lösungsfluss der Differentialgleichung (5.2).

<sup>7</sup>Wobei wir hier wieder jede Matrix in  $\mathbb{R}^{d \times d}$  in kanonischer Weise mit einer linearen Abbildung  $\mathbb{R}^d \rightarrow \mathbb{R}^d$  identifizieren.

- (d) Man kann leicht sehen, dass auch die umgekehrte Aussage in Proposition 5.3.5 gilt. Das heißt genauer: Ist  $X : I \rightarrow \mathbb{R}^{d \times d}$  eine Funktion derart, dass für jedes  $x_0 \in \mathbb{R}^d$  die Abbildung

$$x := Xx_0 : I \ni t \mapsto X(t)x_0 \in \mathbb{R}^d$$

das Anfangswertproblem (5.3) löst, dann ist  $X$  eine Lösungsmatrix der Differentialgleichung (5.2) und sogar eine Hauptfundamentalmatrix, die zum Startzeitpunkt  $t_0$  gehört.<sup>8</sup>

In Abschnitt 5.4 werden wir systematisch untersuchen, wie man für jede autonome lineare Differentialgleichung und für jeden Startzeitpunkt die Hauptfundamentalmatrix angeben kann.

Im aktuellen Abschnitt beschäftigen wir uns aber noch mit der – einfacheren – Frage, wie man (auch im nicht-autonomen Fall) die Hauptfundamentalmatrix für eine ein-dimensionale Differentialgleichung angeben kann.

**Proposition 5.3.8 (Hauptfundamentalmatrix im ein-dimensionalen Fall).** *Es sei  $d = 1$ .<sup>9</sup> Wir fixieren  $t_0 \in I$ . Dann ist die Hauptfundamentalmatrix  $X$  von (5.2), die zum Zeitpunkt  $t_0$  gehört, gegeben durch*

$$X(t) = \exp\left(\int_{t_0}^t A(s) \, ds\right) \in \mathbb{R}$$

für alle  $t \in I$ . Anders ausgedrückt ist die Lösung der Anfangswertproblems

$$\begin{cases} \dot{x}(t) = A(t)x(t), \\ x(t_0) = x_0 \end{cases}$$

für  $x_0 \in \mathbb{R}$  gegeben durch

$$I \ni t \mapsto x(t) := \exp\left(\int_{t_0}^t A(s) \, ds\right)x_0 \in \mathbb{R}.$$

**Aufgabe 5.3.9.** Beweisen Sie Proposition 5.3.8.

**Beispiele 5.3.10 (Einige Beispiele im 1D-Fall).** (a) Sie haben den Spezialfall von Proposition 5.3.8, in dem  $I = \mathbb{R}$  und  $t_0 = 0$  ist, im Grunde schon in Aufgabe 17(b) auf Übungsblatt 4 behandelt.

---

<sup>8</sup>Um dies zu sehen setze man für  $x_0$  einfach der Reihe nach die kanonischen Einheitsvektoren  $e_1, \dots, e_d$  ein.

<sup>9</sup>Dann ist die Abbildung  $A$  also einfach eine Abbildung von  $I$  nach  $\mathbb{R} = \mathbb{R}^{1 \times 1}$ .

- (b) Im autonomen Fall, d.h. wenn  $A(t)$  für alle Zeiten  $t$  gleich einer festen Zahl ist – welche wir ebenfalls mit  $A$  bezeichnen<sup>10</sup> –, erhält man aus Proposition 5.3.8 den einfachen, aber wichtigen Spezialfall

$$X(t) = \exp((t - t_0)A)$$

für alle  $t \in I$ .

- (c) Wir betrachten noch das – willkürlich gewählte – Beispiel  $I = (0, \infty)$  und  $A(t) = \frac{1}{t}$  für alle  $t \in I$ , d.h. wir betrachten die Differentialgleichung

$$\dot{x}(t) = \frac{x(t)}{t}.$$

Für  $t_0 \in (0, \infty)$  erhalten wir aus Proposition 5.3.8, dass die Hauptfundamental„matrix“  $X$ , die zu diesem Startzeitpunkt gehört: Es ist

$$X(t) = \exp\left(\int_{t_0}^t \frac{1}{s} ds\right) = \exp(\ln t - \ln t_0) = \frac{t}{t_0}$$

für alle  $t \in (0, \infty)$ . Für jedes  $x_0 \in \mathbb{R}$  ist also die Lösung des Anfangswertproblems

$$\begin{cases} \dot{x}(t) = \frac{x(t)}{t}, \\ x(t_0) = x_0 \end{cases}$$

gegeben durch  $x(t) = \frac{t}{t_0}x_0$  für alle  $t \in (0, \infty)$ .

Wir schließen den Abschnitt mit einer Bemerkung zur sogenannten *Wronski-Determinante* einer Lösungsmatrix.

**Bemerkung 5.3.11 (Wronski<sup>11</sup>-Determinante).** Sei  $X : I \rightarrow \mathbb{R}^{d \times d}$  eine Lösungsmatrix von (5.2). Dann ist die Abbildung

$$\varphi : I \ni t \mapsto \det X(t) \in \mathbb{R}$$

eine stetig differenzierbare Abbildung; man nennt sie die *Wronski-Determinante* der Lösungsmatrix  $X$ .

Aufgrund von Bemerkung 5.3.4(b) ist  $\varphi(t_0) \neq 0$  für mindestens ein  $t_0 \in I$  genau dann, wenn  $\varphi(t) \neq 0$  für alle  $t \in I$  gilt, genau dann, wenn  $X$  eine Fundamentalmatrix von (5.2) ist.

<sup>10</sup>Das ist eigentlich etwas unpräzise, denn  $A$  ist ja eine Abbildung von  $I$  nach  $\mathbb{R}$ , und selbst wenn diese nur einen konstanten Wert annimmt, dürfte man diesen genau genommen nicht mit demselben Symbol bezeichnen wie die Abbildung selbst.

<sup>11</sup>Benannt nach Józef Maria Hoene-Wronski (1778 – 1853), polnischer Mathematiker und Philosoph.

Man kann zeigen, dass die Wronski-Determinante  $\varphi$  die Differentialgleichung

$$\dot{\varphi}(t) = \text{spur}(A(t)) \varphi(t)$$

für alle  $t \in I$  erfüllt<sup>12</sup>, siehe zum Beispiel [PW19, Lemma 3.1.2]. Eine Anwendung dieser Differentialgleichung für die Wronski-Determinante werden wir später in der Vorlesung im Beweis von Satz 7.4.3 sehen. Aufgrund von Proposition 5.3.8 können wir die Wronski-Determinante  $\varphi$  aufgrund der obenstehenden Differentialgleichung sogar berechnen, wenn wir sie zu einem Zeitpunkt  $t_0$  kennen: Es ist

$$\varphi(t) = \exp\left(\int_{t_0}^t \text{spur}(A(s)) \, ds\right) \varphi(t_0)$$

für alle  $t \in I$ .

## 5.4 Homogene Systeme im autonomen Fall: Die Matrix-Exponentialfunktion

Im vorangehenden Abschnitt hatten wir homogene lineare Differentialgleichungen untersucht und für sie den Begriff *Fundamentalmatrix* definiert. Nun wollen wir hiervon den Spezialfall betrachten, in dem die Differentialgleichung zusätzlich noch autonom ist, d.h. wir studieren in diesem Abschnitt die  $d$ -dimensionale Differentialgleichung

$$\dot{x}(t) = Ax(t) \tag{5.5}$$

für eine zeitunabhängige Matrix  $A \in \mathbb{R}^{d \times d}$ . Da die Matrix  $A$  nicht von der Zeit abhängt, gibt es hier auch keinen Grund sich auf ein Zeitintervall  $I$  einzuschränken, welches kleiner als  $\mathbb{R}$  ist<sup>13</sup> – d.h. wir wählen als Zeitintervall in diesem Abschnitt immer  $I = \mathbb{R}$ . Somit ist die Differentialgleichung (5.5) insgesamt auf dem Gebiet  $G = \mathbb{R}^{1+d}$  definiert.

Ziel des Abschnitts ist es, für gegebenes  $t_0 \in \mathbb{R}$  die Hauptfundamentalmatrix von (5.5) zu finden, die zum Startzeitpunkt  $t_0$  gehört.

Die Idee ist nun folgendermaßen: Im Fall  $d = 1$  kennen wir die Hauptfundamentalmatrix, die zum Startzeitpunkt  $t_0$  gehört, bereits: Sie ist laut Beispiel 5.3.10(b) gleich der Abbildung

$$\mathbb{R} \ni t \mapsto \exp(A(t - t_0)) \in \mathbb{R}$$

<sup>12</sup>Hierbei verwenden für jede Matrix  $B \in \mathbb{R}^{d \times d}$  die Bezeichnung  $\text{spur}(B)$  für die *Spur* von  $B$  – d.h. es ist  $\text{spur}(B) \in \mathbb{R}$  die Summe der Diagonaleinträge von  $B$ .

<sup>13</sup>Im nicht-autonomen Fall hingegen ist diese Einschränkung manchmal nötig, weil  $A(t)$  dann vielleicht nur für Zeiten  $t$  aus einem bestimmten Intervall  $I$  definiert ist; dies ist beispielsweise in Beispiel 5.3.10(c) der Fall.

## 5.4. Homogene Systeme im autonomen Fall: Die Matrix-Exponentialfunktion

(wobei  $A$  wegen  $d = 1$  ein Element von  $\mathbb{R}^{1 \times 1} = \mathbb{R}$  – also schlichtweg eine reelle Zahl – ist). Nun sind wir etwas frech und versuchen, im höher-dimensionalen Fall einfach dieselbe Formel zu verwenden, um eine Fundamentalmatrix zu erhalten; die Frage ist nur, wie man die Exponentialfunktion zu verstehen hat, wenn man eine Matrix anstatt einer Zahl einsetzt.

Ein mögliche – und gute – Antwort lautet: Man definiert die Exponentialfunktion einfach genauso wie für Zahlen, nämlich über die Exponentialreihe. Aus Gründen, die wir später noch sehen werden, ist es sinnvoll, dies nicht nur für reelle, sondern sogar für komplexe Matrizen zu tun. Dies führt uns zu folgender Proposition und der darauffolgenden Definition.

Um die folgenden Aussagen und ihre Beweise richtig zu verstehen, ist es wichtig, dass Sie mit dem Begriff der *Matrix-Norm* vertraut sind. Wir verstehen  $\mathbb{C}^{d \times d}$  bzw.  $\mathbb{R}^{d \times d}$  im Folgenden immer mit der *induzierten 2-Norm*, die in Anhang C.1 erläutert wird.

**Proposition 5.4.1.** *Sei  $A \in \mathbb{C}^{d \times d}$ . Dann konvergiert die Reihe*

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

*in  $\mathbb{C}^{d \times d}$ . Genauer bedeutet dies, dass die Folge der Partialsummen*

$$\left( \sum_{k=0}^n \frac{A^k}{k!} \right)_{n \in \mathbb{N}_0}$$

*gegen eine Matrix in  $\mathbb{C}^{d \times d}$  konvergiert<sup>14</sup>. Hat die Matrix  $A$  nur reelle Einträge, so hat auch der Grenzwert der reelle Einträge.*

*Beweis.* Es gilt

$$\sum_{k=0}^{\infty} \left\| \frac{A^k}{k!} \right\| \leq \sum_{k=0}^{\infty} \frac{\|A\|^k}{k!} = e^{\|A\|} < \infty,$$

also ist die in der Proposition angegebene Reihe *absolut konvergent* (siehe Definition A.4.9 im Anhang). Weil  $\mathbb{C}^{d \times d}$  zusammen mit der induzierten 2-Norm ein Banachraum ist, folgt aus Proposition A.4.10 im Anhang, dass die Reihe konvergiert.

Wenn alle Einträge der Matrix  $A$  reell sind, dann sind auch alle Einträge der  $n$ -ten Partialsummen  $\sum_{k=0}^n \frac{A^k}{k!}$  reell – und laut Proposition C.1.3(b) konvergiert die Folge der Partialsummen nicht nur in Norm, sondern auch eintragsweise gegen ihren Grenzwert, also sind auch die Einträge des Grenzwertes reell.<sup>15</sup> □

<sup>14</sup>Die Konvergenz von Reihen wird in Definition A.4.9 und in Proposition A.4.10 im Anhang etwas genauer besprochen.

<sup>15</sup>Hier haben wir verwendet, dass  $\mathbb{R}$  in  $\mathbb{C}$  abgeschlossen ist.

**Definition 5.4.2 (Matrix-Exponentialfunktion).** Sei  $A \in \mathbb{C}^{d \times d}$ . Dann definieren wir

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

Die Abbildung

$$\exp : \mathbb{C}^{d \times d} \ni A \mapsto \exp(A) \in \mathbb{C}^{d \times d}$$

nennen wir die *Matrix-Exponentialfunktion*.

Als erstes bemerken wir, dass die Matrix-Exponentialfunktion stetig ist:

**Proposition 5.4.3.** Die Abbildung  $\exp : \mathbb{C}^{d \times d} \rightarrow \mathbb{C}^{d \times d}$  ist stetig.

*Beweis.* Das ist Teil des fünften Übungsblattes.  $\square$

Für komplexe Zahlen  $a, b$  wissen Sie, dass die sogenannte *Funktionalgleichung* der Exponentialfunktion gilt, d.h. es ist  $e^{a+b} = e^a e^b$ . Dasselbe gilt auch für Matrizen  $A$  und  $B$  – aber im Allgemeinen nur, wenn  $A$  und  $B$  kommutieren! Dies ist der Inhalt der nächsten Proposition.

**Proposition 5.4.4.** Seien  $A, B \in \mathbb{C}^{d \times d}$  und sei  $AB = BA$ . Dann gilt

$$e^{A+B} = e^A e^B.$$

Insbesondere ist  $e^A$  invertierbar mit inverser Matrix  $(e^A)^{-1} = e^{-A}$ .

*Beweis.* Per Induktion kann man unter Verwendung von  $AB = BA$  zeigen, dass der binomische Lehrsatz

$$(A + B)^n = \sum_{k=0}^n \binom{n}{k} A^k B^{n-k}$$

für die Matrizen  $A$  und  $B$  gilt.<sup>16</sup>

Mit Hilfe des binomischen Lehrsatzes kann man die Funktionalgleichung nun genauso zeigen, wie Sie es für den skalaren Fall aus der Analysis 1 kennen.

Nun zeigen wir noch die letzte Aussage. Weil  $A$  und  $-A$  kommutieren, gilt laut der soeben gezeigten Funktionalgleichung  $e^A e^{-A} = e^{A+(-A)} = e^0 = \text{id}_{\mathbb{C}^d}$ .  $\square$

Bevor wir erläutern, wie genau die Matrix-Exponentialfunktion mit Fundamentalsystemen von linearen Differentialgleichungen zusammenhängt, diskutieren wir einige Beispiele.

<sup>16</sup>Achtung: Hier geht wesentlich ein, dass  $A$  und  $B$  kommutieren – ohne diese Annahme ist der binomische Lehrsatz für Matrizen im Allgemeinen falsch!

## 5.4. Homogene Systeme im autonomen Fall: Die Matrix-Exponentialfunktion

**Beispiele 5.4.5.** (a) Für die Nullmatrix  $0 \in \mathbb{C}^{d \times d}$  gilt

$$e^0 = \sum_{k=0}^{\infty} \frac{0^k}{k!} = \text{id}_{\mathbb{C}^d}.$$

(Dies hatten wir am Ende des Beweises von Proposition 5.4.4 bereits benutzt.)

(b) Für die Matrix

$$A = \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix} \in \mathbb{C}^{2 \times 2}$$

gilt

$$\begin{aligned} e^A &= \sum_{k=0}^{\infty} \frac{1}{k!} \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix}^k = \sum_{k=0}^{\infty} \begin{pmatrix} \frac{(-1)^k}{k!} & 0 \\ 0 & \frac{2^k}{k!} \end{pmatrix} \\ &= \begin{pmatrix} \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} & 0 \\ 0 & \sum_{k=0}^{\infty} \frac{2^k}{k!} \end{pmatrix} = \begin{pmatrix} e^{-1} & 0 \\ 0 & e^2 \end{pmatrix}. \end{aligned}$$

Allgemeiner kann man für eine Diagonalmatrix  $A \in \mathbb{C}^{d \times d}$  mit Diagonaleinträgen  $\lambda_1, \dots, \lambda_d$  mit Hilfe derselben Rechnung wie oben sehen, dass

$$\exp \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_d \end{pmatrix} = \begin{pmatrix} e^{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_d} \end{pmatrix}$$

gilt.

(c) Sei  $a \in \mathbb{C}$ . Für die Matrix

$$A = \begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix}$$

gilt  $A^2 = 0$ . Somit ist

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!} = \frac{A^0}{0!} + \frac{A^1}{1!} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}.$$

Weil die Matrix  $A$  mit jedem Vielfachen der Einheitsmatrix  $\text{id}_{\mathbb{C}^2}$  kommutiert, erhalten wir aus Proposition 5.4.4 für jede Zahl  $b \in \mathbb{C}$

$$\begin{aligned} \exp \begin{pmatrix} b & a \\ 0 & b \end{pmatrix} &= \exp \left( \begin{pmatrix} b & 0 \\ 0 & b \end{pmatrix} + \begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix} \right) = \exp \begin{pmatrix} b & 0 \\ 0 & b \end{pmatrix} \cdot \exp \begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} e^b & 0 \\ 0 & e^b \end{pmatrix} \cdot \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} e^b & e^b a \\ 0 & e^b \end{pmatrix}. \end{aligned}$$

- (d) Sei  $P \in \mathbb{C}^{d \times d}$  eine *Projektion*, d.h. eine Matrix, die die Gleichung  $P^2 = P$  erfüllt. Dann gilt für alle natürlichen Zahlen  $k \geq 1$  die Gleichung  $P^k = P$ , für die Zahl  $k = 0$  hingegen gilt  $P^k = \text{id}_{\mathbb{C}^d}$ . Somit gilt

$$\begin{aligned} e^P &= \sum_{k=0}^{\infty} \frac{P^k}{k!} = \text{id}_{\mathbb{C}^d} + \sum_{k=1}^{\infty} \frac{1}{k!} P = \text{id}_{\mathbb{C}^d} - P + \sum_{k=0}^{\infty} \frac{1}{k!} P \\ &= \text{id}_{\mathbb{C}^d} - P + eP = \text{id}_{\mathbb{C}^d} + (e-1)P. \end{aligned}$$

Ein konkretes Beispiel für eine Projektion im 2-dimensionalen ist die Matrix

$$P = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \in \mathbb{C}^{2 \times 2}.$$

Für diese Matrix gilt also

$$e^P = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{e-1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{e+1}{2} & \frac{e-1}{2} \\ \frac{e-1}{2} & \frac{e+1}{2} \end{pmatrix}.$$

Wenn man will, kann man diese Matrix übrigens noch etwas umschreiben: Es gilt

$$e^P = e^{1/2} \begin{pmatrix} \frac{e^{1/2} + e^{-1/2}}{2} & \frac{e^{1/2} - e^{-1/2}}{2} \\ \frac{e^{1/2} - e^{-1/2}}{2} & \frac{e^{1/2} + e^{-1/2}}{2} \end{pmatrix} = e^{1/2} \begin{pmatrix} \cosh \frac{1}{2} & \sinh \frac{1}{2} \\ \sinh \frac{1}{2} & \cosh \frac{1}{2} \end{pmatrix}.$$

Auf dem fünften Übungsblatt werden Sie sehen, dass hinter dieser Formel ein etwas allgemeineres Prinzip steckt.

Im nächsten Abschnitt werden wir ein systematisches Verfahren besprechen um die Matrix-Exponentialfunktion für eine konkret gegebene Matrix zu berechnen. Nun wollen wir aber zum Ziel des aktuellen Abschnitts zurückkehren und erläutern, was die Matrix-Exponentialfunktion mit Differentialgleichungen und mit Fundamentalmatrizen zu tun hat. Dies ist der Inhalt des folgenden Satzes und seines Korollars.

**Satz 5.4.6.** Sei  $A \in \mathbb{C}^{d \times d}$ . Dann ist die Abbildung

$$\mathbb{R} \ni t \mapsto e^{tA} \in \mathbb{C}^{d \times d}$$

stetig differenzierbar, und ihre Ableitung ist gleich

$$\frac{d}{dt} e^{tA} = A e^{tA} = e^{tA} A.$$

für alle  $t \in \mathbb{R}$ .

#### 5.4. Homogene Systeme im autonomen Fall: Die Matrix-Exponentialfunktion

*Beweis.* Wir beweisen zuerst, dass die gegebene Abbildung im Punkt 0 differenzierbar ist mit Ableitung  $A$ . Hierfür müssen nur zeigen, dass

$$\frac{e^{sA} - \text{id}_{\mathbb{C}^d}}{s} - A \rightarrow 0 \quad \text{für } s \rightarrow 0$$

gilt. Für jedes  $s \neq 0$  berechnen wir dazu

$$\begin{aligned} \left\| \frac{e^{sA} - \text{id}_{\mathbb{C}^d}}{s} - A \right\| &= \left\| \frac{\sum_{k=1}^{\infty} \frac{s^k A^k}{k!}}{s} - A \right\| = \left\| \sum_{k=2}^{\infty} \frac{s^{k-1} A^k}{k!} \right\| \leq |s| \sum_{k=2}^{\infty} \frac{s^{k-2} \|A\|^k}{k!} \\ &= |s| \|A\|^2 \sum_{k=0}^{\infty} \frac{s^k \|A\|^k}{(k+2)!} \leq |s| \|A\|^2 \sum_{k=0}^{\infty} \frac{s^k \|A\|^k}{k!} = |s| \|A\|^2 e^{s\|A\|}. \end{aligned}$$

Für  $s \rightarrow 0$  geht der ganz rechts stehende Ausdruck gegen 0; also haben wir die Differenzierbarkeit im Zeitpunkt 0 und die behauptete Form der Ableitung in diesem Punkt gezeigt.

Sei nun  $t \in \mathbb{R}$  beliebig. Für alle  $s \neq t$  erhalten wir, weil  $tA$  und  $sA$  kommutieren, aus Proposition 5.4.4

$$\begin{aligned} \frac{e^{tA} - e^{sA}}{t-s} - Ae^{tA} &= \frac{e^{tA} - e^{sA}}{t-s} - Ae^{sA} + A(e^{sA} - e^{tA}) \\ &= e^{sA} \left( \frac{e^{(t-s)A} - \text{id}_{\mathbb{C}^d}}{t-s} - A \right) + A(e^{sA} - e^{tA}) \end{aligned}$$

Nun betrachten wir das Verhalten der einzelnen Terme für  $s \rightarrow t$ : Wegen Proposition 5.4.3 gilt dann  $e^{sA} \rightarrow e^{tA}$ ; außerdem haben wir bereits gezeigt, dass  $\frac{e^{(t-s)A} - \text{id}_{\mathbb{C}^d}}{t-s}$  gegen  $A$  konvergiert. Insgesamt erhalten wir also

$$\frac{e^{tA} - e^{sA}}{t-s} - Ae^{tA} \rightarrow e^{tA}(A - A) + A(e^{tA} - e^{tA}) = 0$$

für  $s \rightarrow t$ ; dies zeigt, dass die in der Proposition gegebene Abbildung im Zeitpunkt  $t$  tatsächlich differenzierbar ist und dort die Ableitung  $Ae^{tA}$  besitzt.

Weil die Funktion  $t \mapsto e^{tA}$  differenzierbar ist und die in der Proposition angegebene Differentialgleichung erfüllt, ist sie sogar stetig differenzierbar. Zuletzt bemerken wir, dass aus der Reihendarstellung von  $e^{tA}$  folgt, dass diese Matrix mit  $A$  kommutiert.  $\square$

**Korollar 5.4.7.** Sei  $A \in \mathbb{R}^{d \times d}$  und sei  $t_0 \in \mathbb{R}$ . Dann ist die Hauptfundamental-Matrix der Differentialgleichung

$$\dot{x}(t) = Ax(t),$$

die zum Startzeitpunkt  $t_0$  gehört, gegeben durch

$$\mathbb{R} \ni t \mapsto e^{(t-t_0)A} \in \mathbb{R}^{d \times d}.$$

Es ist lehrreich, die Aussage des Korollars mit der analogen eindimensionalen Aussage zu vergleichen, die wir bereits in Beispiel 5.3.10(b) betrachtet haben.

*Beweis von Korollar 5.4.7.* Sei  $x_0 \in \mathbb{R}^d$  und sei  $x : \mathbb{R} \ni t \mapsto e^{(t-t_0)A}x_0 \in \mathbb{R}^d$ ; wir müssen lediglich zeigen, dass diese Funktion  $x$  das Anfangswertproblem

$$\begin{cases} \dot{x}(t) = Ax(t), \\ x(t_0) = x_0 \end{cases}$$

löst.

*Anfangsbedingung:* Es gilt

$$x(t_0) = e^{0 \cdot A}x_0 = \text{id}_{\mathbb{C}^d}x_0 = x_0.$$

*Differentialgleichung:* Wegen Satz 5.4.6 ist die Funktion  $x$  differenzierbar, und ihre Ableitung ist gegeben durch

$$\dot{x}(t) = \frac{d}{dt}e^{(t-t_0)A}x_0 = e^{-t_0A} \frac{d}{dt}e^{tA}x_0 = e^{-t_0A}Ae^{tA}x_0 = Ae^{(t-t_0)A}x_0 = Ax(t)$$

für alle  $t \in \mathbb{R}$ ; also erfüllt  $x$  tatsächlich die Differentialgleichung.  $\square$

Sie wundern sich vielleicht, weshalb wir die Matrix-Exponentialfunktion für komplexe Matrizen eingeführt haben, obwohl wir sie doch in Korollar 5.4.7 nur für reelle Matrizen benötigen. Ein Grund dafür ist, dass bei der allgemeinen Berechnungsmethode für die Matrix-Exponentialfunktion, die wir im nächsten Abschnitt besprechen, manchmal komplexe Matrizen auftauchen – selbst dann, wenn man  $e^A$  nur für reelle Matrizen  $A$  berechnen will.

**Literaturstellen und verwandte Ergebnisse.** Die Lösung von homogenen linearen Differentialgleichungen mit nicht-konstanten Koeffizienten (d.h. der nicht-autonome Fall) ist deutlich schwieriger als der in diesem Abschnitt besprochene autonome Fall; es gibt für den nicht-autonomen Fall kein allgemeines Lösungsverfahren.<sup>17</sup>

Eine Methode, die im nicht-autonomen Fall hilfreich sein kann, wird zum Beispiel in [PW19, Abschnitt 3.3.1] vorgestellt.

## 5.5 Berechnungsverfahren für die Matrix-Exponentialfunktion

Einige Möglichkeiten um in bestimmten Situationen die Matrix  $\exp(A)$  für eine vorgegebene Matrix  $A \in \mathbb{C}^{d \times d}$  zu bestimmen haben Sie bereits in den Beispielen 5.4.5 sowie in Aufgabe 16 auf Übungsblatt 5 kennengelernt.

<sup>17</sup>Mit Ausnahme des ein-dimensionalen Falls – diesen haben wir ja bereits in Proposition 5.3.8 besprochen.

In diesem Abschnitt lernen Sie ein allgemeines Verfahren kennen, mit dem man für konkret vorgegebene Matrizen  $A$  die Exponentialfunktion  $\exp(A)$  ausrechnen kann. Die Grundidee beruht darauf, die Matrix  $A$  zu *diagonalisieren*. Allerdings gibt es natürlich auch nicht diagonalisierbare Matrizen; in diesem Fall verwendet man die *Jordan-Normalform* einer Matrix, die eine Verallgemeinerung der Diagonalisierung darstellt. Beide Konzepte sind Teil der Linearen Algebra. Eine kurze Zusammenfassung hierzu finden Sie auch in Anhang C.2.

Wir unterteilen den Abschnitt in drei (nicht nummerierte) Unterabschnitte: Im ersten betrachten wir den (wichtigen und einfacheren) Spezialfall diagonalisierbarer Matrizen; im zweiten betrachten wir sogenannte *Block-Diagonalmatrizen* bzw. – systemtheoretisch ausgedrückt – Systeme, die in mehrere Teilsysteme entkoppeln – und im dritten Unterabschnitt sehen wir uns allgemeine Matrizen und ihre Jordan-Normalform an.

### Diagonalisierbare Matrizen

Wir beginnen mit dem wichtigen Spezialfall, in dem die Matrix  $A$  diagonalisierbar ist.

**Satz 5.5.1.** Sei  $A \in \mathbb{C}^{d \times d}$  diagonalisierbar, d.h. es gebe eine invertierbare Matrix  $T \in \mathbb{C}^{d \times d}$  und eine Diagonalmatrix  $D \in \mathbb{C}^{d \times d}$  mit Diagonaleinträgen  $\lambda_1, \dots, \lambda_d$  derart, dass  $A = TDT^{-1}$  ist. Dann gilt

$$e^A = Te^DT^{-1} = T \begin{pmatrix} e^{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_d} \end{pmatrix} T^{-1}.$$

Es ist fast ein wenig übertrieben, dass wir dieses Resultat als „Satz“ bezeichnet haben, denn der Beweis, den wir gleich führen werden, ist beinahe eine Trivialität; die wesentliche Erkenntnis hinter dem Satzes steckt darin, überhaupt erst zu bemerken, wie sehr sich die Berechnung der Matrix-Exponentialfunktion durch Diagonalisierung vereinfacht.

*Beweis von Satz 5.5.1.* Die Definition von  $e^A$  und  $e^D$  als Exponentialreihe liefert

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!} = \sum_{k=0}^{\infty} \frac{(TDT^{-1})^k}{k!} = \sum_{k=0}^{\infty} \frac{TD^kT^{-1}}{k!} = T \sum_{k=0}^{\infty} \frac{D^k}{k!} T^{-1} = Te^DT^{-1};$$

für die vorletzte Gleichheit haben wir verwendet, dass die Matrixmultiplikation stetig ist, damit wir die Faktoren  $T$  und  $T^{-1}$  aus der unendlichen Reihe ziehen dürfen. Der wesentliche Schritt in oben stehender Rechnung ist die

dritte Gleichheit: hier haben wir verwendet, dass  $(TDT^{-1})^k = TD^kT^{-1}$  für alle  $k \in \mathbb{N}_0$ .

Dass  $e^D$  die behauptete Form hat, haben wir bereits in Beispiel 5.4.5(b) gesehen.  $\square$

Bitte beachten Sie, dass man mit Satz 5.5.1 auch  $e^{tA}$  für jedes  $t \in \mathbb{R}$  ausrechnen kann, wenn  $A$  diagonalisierbar ist; denn aus  $A = TDT^{-1}$  folgt ja  $tA = T(tD)T^{-1}$ , und  $tD$  ist ebenfalls eine Diagonalmatrix.

Wir demonstrieren die Anwendung von Satz 5.5.1 anhand von zwei Matrizen, deren Exponentialfunktion Sie in Aufgabe 16 auf Blatt 5 mit einer anderen Methode berechnen.

**Beispiele 5.5.2.** (a) Es sei

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in \mathbb{C}^{2 \times 2}.$$

Man kann leicht das charakteristische Polynom  $p_A$  dieser Matrix berechnen: für alle  $\lambda \in \mathbb{C}$  gilt

$$p_A(\lambda) = \det(\lambda \text{id} - A) = \lambda^2 - 1 = (\lambda - 1)(\lambda + 1).$$

Also besitzt  $A$  die beiden Eigenwert  $-1$  und  $1$ . Mit Hilfe des Gauß-Algorithmus kann man je einen zugehörigen Eigenvektor berechnen<sup>18</sup>, zum Beispiel die beiden Vektoren  $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$  und  $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ . Wenn wir also

$$D := \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{und} \quad T := \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

setzen, dann gilt  $A = TDT^{-1}$ , und somit erhalten wir aus Satz 5.5.1 für alle  $t \in \mathbb{R}$

$$\begin{aligned} e^{tA} &= T e^{tD} T^{-1} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} e^{-t} & 0 \\ 0 & e^t \end{pmatrix} \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} e^{-t} & -e^{-t} \\ e^t & e^t \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} e^{-t} + e^t & -e^{-t} + e^t \\ -e^{-t} + e^t & e^{-t} + e^t \end{pmatrix} = \begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix}; \end{aligned}$$

hierbei haben wir die Inverse von  $T^{-1}$  benötigt; man kann sie zum Beispiel mit Hilfe des Gauß-Algorithmus ausrechnen.

(b) Betrachten wir nun die Matrix

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathbb{C}^{2 \times 2}.$$

<sup>18</sup>In diesem einfachen Fall kann man die Eigenvektoren auch schlicht erraten.

Mit einer kurzen Rechnung erhält man die beiden Eigenwerte  $-i$  und  $i$ , sowie z.B.  $\begin{pmatrix} 1 \\ i \end{pmatrix}$  und  $\begin{pmatrix} 1 \\ -i \end{pmatrix}$  als zugehörige Eigenvektoren. Wir können also

$$D := \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} \quad \text{und} \quad T := \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}$$

definieren. Somit gilt für alle  $t \in \mathbb{R}$

$$\begin{aligned} e^{tA} &= T e^{tD} T^{-1} = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \begin{pmatrix} e^{-it} & 0 \\ 0 & e^{it} \end{pmatrix} \frac{1}{2} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \begin{pmatrix} e^{-it} & -ie^{-it} \\ e^{it} & ie^{it} \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} e^{-it} + e^{it} & i(-e^{-it} + e^{it}) \\ i(e^{-it} - e^{it}) & e^{-it} + e^{it} \end{pmatrix} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}. \end{aligned}$$

**Bemerkung 5.5.3.** Das obige Beispiel (b) zeigt, weshalb es notwendig war, die Matrix-Exponentialfunktion auch für komplexe Matrizen einzuführen: Obwohl wir lediglich die Exponentialfunktion der reellen Matrix

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

berechnen wollten, tauchte in der Berechnung die komplexe Matrix

$$D = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}$$

auf, und wir mussten für sie die Exponentialfunktion berechnen. Abstrakt gesprochen liegt dies daran, dass Matrizen mit reellen Einträgen manchmal komplexe Eigenwerte haben.

### Block-Diagonalmatrizen und entkoppelte Differentialgleichungen

In Beispiel 5.4.5(b) hatten Sie gesehen, dass es sehr leicht ist, die Exponentialfunktion einer Diagonalmatrix zu berechnen, und im vorangehenden Unterabschnitt haben wir uns das zu Nutze gemacht um allgemeiner die Exponentialfunktion von diagonalisierbaren Matrizen zu bestimmen.

Anstelle von Diagonalmatrizen kann man auch etwas allgemeiner sogenannte *Block-Diagonalmatrizen* betrachten. Eine Matrix  $A \in \mathbb{C}^{d \times d}$  heißt *Block-Diagonalmatrix*, wenn man sie in der Form

$$A = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{pmatrix} \quad (5.6)$$

schreiben kann, wobei  $A_1 \in \mathbb{C}^{d_1 \times d_1}, \dots, A_m \in \mathbb{C}^{d_m \times d_m}$  quadratische Matrizen sind.<sup>19,20</sup>

Wenn man die Exponentialfunktion einer Block-Diagonalmatrix ausrechnen möchte, dann kann man dies einfach blockweise tun:

**Proposition 5.5.4.** *Für eine Block-Diagonalmatrix  $A$  wie in (5.6) gilt*

$$e^A = \begin{pmatrix} e^{A_1} & & 0 \\ & \ddots & \\ 0 & & e^{A_m} \end{pmatrix},$$

wobei die Ausdrücke  $e^{A_1}, \dots, e^{A_m}$  ebenfalls im Sinne der Matrix-Exponentialfunktion zu verstehen sind.

*Beweis.* Man kann leicht nachrechnen, dass

$$\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{pmatrix}^k = \begin{pmatrix} (A_1)^k & & 0 \\ & \ddots & \\ 0 & & (A_m)^k \end{pmatrix}$$

für alle  $k \in \mathbb{N}_0$  gilt. Somit folgt die Behauptung aus der Exponentialreihe, über welche die Matrix-Exponentialfunktion definiert ist.  $\square$

Wenn die Blöcke  $A_1, \dots, A_m$  alle die Größe  $d_1 = \dots = d_m = 1$  haben – also schlichtweg Zahlen sind – dann ist  $A$  einfach eine Diagonalmatrix, und die Aussage von Proposition 5.5.4 ist dann die Formel für die Exponentialfunktion von Diagonalmatrizen, die wir bereits in Beispiel 5.4.5(b) berechnet haben.

**Beispiel 5.5.5.** Sei

$$A = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 4 \end{pmatrix} \in \mathbb{C}^{3 \times 3}.$$

Dann können wir  $tA$  für jedes  $t \in \mathbb{R}$  schreiben als

$$tA = \begin{pmatrix} tA_1 & 0 \\ 0 & tA_2 \end{pmatrix}, \quad \text{mit } A_1 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathbb{C}^{2 \times 2} \quad \text{und } A_2 = 4 \in \mathbb{C}^{1 \times 1}.$$

Deswegen erhalten wir aus Proposition 5.5.4 die Formel

$$e^{tA} = \begin{pmatrix} e^{tA_1} & 0 \\ 0 & e^{tA_2} \end{pmatrix} = \begin{pmatrix} \cos t & -\sin t & 0 \\ \sin t & \cos t & 0 \\ 0 & 0 & e^{4t} \end{pmatrix}$$

<sup>19</sup>Hierbei gilt natürlich  $d_1 + \dots + d_m = d$ .

<sup>20</sup>Rein formal betrachtet kann man jede Matrix als Block-Diagonalmatrix auffassen, die aus nur einem Block besteht (mit  $d_1 = d$ ). Das ist aber natürlich nicht besonders hilfreich.

für alle  $t \in \mathbb{R}$ .<sup>21</sup>

Die beiden folgenden Bemerkungen 5.5.6 und 5.5.8 sind wichtig um die Berechnung der Matrix-Exponentialfunktion mit Hilfe von Blockmatrizen richtig zu verstehen und einzuordnen.

**Bemerkung 5.5.6 (Entkoppelte Differentialgleichungen).** Für eine Matrix  $A \in \mathbb{R}^{d \times d}$  wollen wir  $e^{tA}$  für Zeiten  $t \in \mathbb{R}$  ja in erster Linie deshalb ausrechnen, weil wir eine Fundamentalmatrix für die autonome und homogene lineare Differentialgleichung

$$\dot{x}(t) = Ax(t) \tag{5.7}$$

suchen.<sup>22</sup> Deswegen ist es naheliegend sich zu fragen, was es eigentlich für die Differentialgleichung (5.7) bedeutet, wenn man  $A$  als Block-Diagonalmatrix schreiben kann.

Lassen Sie uns dies im Fall von zwei Blöcken besprechen<sup>23</sup>, d.h. wir betrachten den Fall

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} \in \mathbb{C}^{d \times d}$$

mit zwei quadratischen Matrizen  $A_1 \in \mathbb{R}^{d_1 \times d_1}$  und  $A_2 \in \mathbb{R}^{d_2 \times d_2}$ . Lassen Sie uns auch den Vektor  $x(t)$  schreiben als  $x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}$  mit  $x_1(t) \in \mathbb{R}^{d_1}$  und  $x_2(t) \in \mathbb{R}^{d_2}$  – d.h.  $x_1(t)$  enthält die ersten  $d_1$  Komponenten von  $x(t)$  und  $x_2(t)$  enthält die verbleibenden  $d_2$  Komponenten. Die Differentialgleichung (5.7) kann man mit dieser Notation schreiben als

$$\begin{pmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{pmatrix} = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} A_1 x_1(t) \\ A_2 x_2(t) \end{pmatrix}.$$

D.h. wir erhalten für  $x_1(t)$  und  $x_2(t)$  die beiden Differentialgleichungen

$$\dot{x}_1(t) = A_1 x_1(t) \quad \text{und} \quad \dot{x}_2(t) = A_2 x_2(t),$$

die völlig unabhängig voneinander sind – in anderen Worten: Die Differentialgleichung (5.7) *entkoppelt* in zwei Differentialgleichungen, die nichts miteinander zu tun haben!

<sup>21</sup>Hierbei haben wir verwendet, dass wir  $e^{tA_1}$  in Beispiel 5.5.2(b) bereits berechnet haben. Ansonsten müsste man  $e^{tA_1}$  an dieser Stelle eben noch ausrechnen um eine explizite Formel für  $e^{tA}$  zu erhalten.

<sup>22</sup>Denn laut Korollar 5.4.7 ist die Abbildung  $\mathbb{R} \ni \rightarrow e^{(t-t_0)A} \in \mathbb{R}^{d \times d}$  die Hauptfundamentalmatrix zum Startzeitpunkt  $t_0$  für die Differentialgleichung (5.7).

<sup>23</sup>Wir schränken uns hier nur der einfacheren Notation halber auf zwei Blöcke ein; für mehrere Blöcke verhält sich alles analog.

Wir können also beide Differentialgleichungen unabhängig voneinander lösen, und die Lösungen dann wieder zu einer Lösung der ursprünglichen Differentialgleichung (5.7) zusammensetzen. Dies ist deswegen relevant, weil es zum Beispiel oft viel einfacher ist zwei zwei-dimensionale Differentialgleichungen zu lösen (oder zu untersuchen) als eine vier-dimensionale.

**Aufgabe 5.5.7.** (a) Gehen Sie Bemerkung 5.5.6 noch einmal durch und sehen Sie sich an, wie hergeleitet wurde, dass die Differentialgleichung (5.7) in zwei niedriger-dimensionale Differentialgleichungen entkoppelt. Dass  $A$  eine Block-Diagonalmatrix ist, wurde dort nur an einer einzigen Stelle benutzt.

Wo genau im Argument befindet sich diese Stelle?

(b) Dass die Differentialgleichung (5.7) in die beiden Differentialgleichungen

$$\dot{x}_1(t) = A_1 x_1(t) \quad \text{und} \quad \dot{x}_2(t) = A_2 x_2(t)$$

entkoppelt, kann man ebenfalls benutzen, um die Aussage

$$e^{tA} = \begin{pmatrix} e^{tA_1} & 0 \\ 0 & e^{tA_2} \end{pmatrix}$$

herzuleiten (die wir in Proposition 5.5.4 bereits mit Hilfe der Exponentialreihe begründet haben).

Können Sie dieses Argument genau ausführen?

**Bemerkung 5.5.8 (Entkopplung und Ummummerierung von Indizes).** Schauen wir uns einmal die Matrix

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \in \mathbb{C}^3$$

an. Auf den ersten Blick ist es nicht möglich, diese Matrix in mehrere Diagonalblöcke aufzuteilen.

Aber andererseits: Die Matrix  $A$  kann man ja als lineare Abbildung  $\mathbb{C}^3 \rightarrow \mathbb{C}^3$  auffassen, und man sieht leicht, dass ihre Wirkung auf der zweiten Komponente eines Vektors im  $\mathbb{C}^3$  nichts mit ihrer Wirkung auf der ersten und dritten Komponente zu tun hat.

Wir können also die Einträge in der zweiten Zeile bzw. Spalte von  $A$  zu einer  $1 \times 1$ -Matrix  $A_1 = -1$  zusammenfassen, und die Wirkung der ersten und dritten Zeile bzw. Spalte von  $A$  zu einer  $2 \times 2$ -Matrix  $A_2 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ . Für alle

$t \in \mathbb{R}$  ist

$$e^{tA_1} = e^{-t} \quad \text{und} \quad e^{tA_2} = \begin{pmatrix} e^t \cosh t & e^t \sinh t \\ e^t \sinh t & e^t \cosh t \end{pmatrix},$$

wobei die Formel für  $e^{tA_2}$  zum Beispiel aus Aufgabe 16(c) auf Übungsblatt 5 folgt. Indem wir die jeweiligen Einträge wieder auf ihre ursprünglichen Positionen zurückschreiben, die zu den Zeilen- und Spaltenindizes der Matrix  $A$  gehörten, erhält man somit

$$e^{tA} = \begin{pmatrix} e^t \cosh t & 0 & e^t \sinh t \\ 0 & e^{-t} & 0 \\ e^t \sinh t & 0 & e^t \cosh t \end{pmatrix}.$$

Die Matrix  $A$  war hier also nicht direkt in Block-Diagonalgestalt, sondern die „Blöcke“ waren über verschiedene, teils nicht benachbarte Indizes verstreut.<sup>24,25</sup>

### Allgemeine Matrizen und Jordan-Normalform

Weil manche Matrizen  $A \in \mathbb{R}^{d \times d}$  nicht diagonalisierbar sind, kann man den Satz 5.5.1 auf diese Matrizen nicht anwenden. Allerdings besitzt, wie Sie aus der Linearen Algebra 2 wissen, jede quadratische Matrix  $A$  eine *Jordan-Normalform*, d.h. man kann  $A$  schreiben als

$$A = TJT^{-1}, \tag{5.8}$$

wobei  $J$  nun keine Diagonalmatrix ist, sondern eine Block-Diagonalmatrix, deren Diagonalblöcke sogenannte *Jordan-Blöcke* sind. Präziser gesagt ist  $J$  also von der Form

$$J = \begin{pmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_m \end{pmatrix} \tag{5.9}$$

mit Matrizen  $J_1 \in \mathbb{C}^{d_1 \times d_1}, \dots, J_m \in \mathbb{C}^{d_m \times d_m}$ , und die  $k$ -te Block-Matrix  $J_k$  hat die spezielle Gestalt

$$J_k = \begin{pmatrix} \lambda_k & 1 & 0 & \cdots & 0 \\ 0 & \lambda_k & 1 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & & & \lambda_k & 1 \\ 0 & \cdots & \cdots & 0 & \lambda_k \end{pmatrix}. \tag{5.10}$$

<sup>24</sup>Man könnte sagen, dass wir die Matrix  $A$  durch Umnummerieren der Indizes zu einer Block-Diagonalmatrix gemacht haben.

<sup>25</sup>Mit Hilfe von Permutationsmatrizen lässt sich das gesamte Argument in dieser Bemerkung etwas präziser fassen.

Dabei ist jede Zahl  $\lambda_k$  ( $k = 1, \dots, m$ ) ein Eigenwert von  $A$ , und es können auch mehrere der Zahlen  $\lambda_1, \dots, \lambda_m$  gleich sein (d.h. in der Jordan-Normalform von  $A$  können mehrere Jordan-Blöcke mit demselben Eigenwert vorkommen).

Mit Hilfe der Jordan-Normalform können wir ein Verfahren angeben um die Matrixexponentialfunktion einer beliebigen Matrix  $A$  zu bestimmen.<sup>26</sup>

**Satz 5.5.9.** Sei  $A \in \mathbb{C}^{d \times d}$ ; die Jordan-Normalform von  $A$  bezeichnen wir mit derselben Notation wie in den Formeln (5.8), (5.9) und (5.10). Dann gilt für jedes  $t \in \mathbb{R}$

$$e^{tA} = T \begin{pmatrix} e^{tJ_1} & & 0 \\ & \ddots & \\ 0 & & e^{tJ_m} \end{pmatrix} T^{-1},$$

und für jedes  $k \in \{1, \dots, m\}$  ist

$$e^{tJ_k} = e^{t\lambda_k} \begin{pmatrix} \frac{t^0}{0!} & \frac{t^1}{1!} & \frac{t^2}{2!} & \cdots & \frac{t^{d_k-1}}{(d_k-1)!} \\ 0 & \frac{t^0}{0!} & \frac{t^1}{1!} & \cdots & \vdots \\ \vdots & & \ddots & \ddots & \frac{t^2}{2!} \\ \vdots & & & \frac{t^0}{0!} & \frac{t^1}{1!} \\ 0 & \cdots & \cdots & 0 & \frac{t^0}{0!} \end{pmatrix}.$$

*Beweis.* Sei  $t \in \mathbb{R}$  fest. Aufgrund der Reihenformel für die Matrix-Exponentialfunktion gilt

$$e^{tA} = \sum_{n=0}^{\infty} \frac{t^n (TJT^{-1})^n}{n!} = \sum_{n=0}^{\infty} T \frac{t^n J^n}{n!} T^{-1} = T e^{tJ} T^{-1}.$$

Die behauptete Gestalt von  $e^{tJ}$  folgt aus Proposition 5.5.4. Also bleibt noch die Formel für  $e^{tJ_k}$  zu zeigen; sei hierzu  $k \in \{1, \dots, m\}$  fest. Aus Aufgabe 16(b) auf Übungsblatt 5 folgt

$$e^{tJ_k} = e^{t\lambda_k \text{id} + t(J_k - \lambda_k \text{id})} = e^{t\lambda_k} e^{t(J_k - \lambda_k \text{id})}.$$

Deshalb genügt es, die Formel für den Fall  $\lambda_k = 0$  zu zeigen. Sei nun also  $\lambda_k = 0$ .

Für diesen Fall besteht  $J_k$  aber nur aus Nullen, mit Ausnahme der ersten Super-Diagonalen, auf der sich nur Einsen befinden. Durch jedes Potenzieren von  $J_k$  wandert die Superdiagonale mit Einsen um eine Position weiter nach rechts oben, d.h. genauer:

<sup>26</sup>Allerdings nur, wenn man die Jordan-Normalform von  $A$  tatsächlich kennt bzw. ausrechnen kann, was keineswegs selbstverständlich ist!

Die Matrix  $(J_k)^n$  ist gleich 0, falls  $n \geq d_k$  ist, und im Fall  $n \in \{0, \dots, d_k - 1\}$  besteht die Matrix  $(J_k)^n$  nur aus Nullen, mit Ausnahme der  $n$ -ten Superdiagonalen, die nur aus Einsen besteht. Mit Hilfe der Exponentialreihe

$$e^{tJ_k} = \sum_{n=0}^{\infty} \frac{(tJ_k)^n}{n!} = \sum_{n=0}^{d_k-1} \frac{t^n (J_k)^n}{n!}$$

erhalten wir deshalb die behauptete Formel im Fall  $\lambda_k = 0$ .  $\square$

Man beachte: Wenn die Matrix  $A$  diagonalisierbar ist, dann haben alle Jordanblöcke die Größe  $d_k = 1$ , und somit ist die Aussage von Satz 5.5.9 in diesem Fall einfach dieselbe Aussage wie in Satz 5.5.1.

Lassen Sie uns zum Schluss dieses Abschnitts noch ein Beispiel für die Anwendung von Satz 5.5.9 diskutieren:

**Beispiel 5.5.10.** Betrachten wir die Matrix

$$A = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \in \mathbb{C}^{3 \times 3}.$$

Mit Methoden der Linearen Algebra kann man die Jordan-Normalform von  $A$  ausrechnen: Man erhält  $A = TJT^{-1}$  mit

$$J = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{und} \quad T = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Somit erhalten wir für  $t \in \mathbb{R}$  aus Satz 5.5.9<sup>27</sup>

$$e^{tA} = Te^{tJ}T^{-1} = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} e^t \begin{pmatrix} 1 & t & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} = e^t \begin{pmatrix} 1 & t & -t \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

**Literaturstellen und verwandte Ergebnisse.** Der enge Zusammenhang zwischen Linearer Algebra, Matrixexponentialfunktion und linearen Differentialgleichungen wird unter anderem in dem Buch [HS74] sehr detailliert besprochen.

<sup>27</sup>In diesem einfachen Beispiel hätte man  $e^{tA}$  aber auch ohne Satz 5.5.9 berechnen können: Nämlich, indem man verwendet, dass das Quadrat der Matrix

$$\begin{pmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

gleich 0 ist.

## 5.6 Inhomogenitäten

In diesem Abschnitt kommen wir zurück zu allgemeinen linearen Differentialgleichungen, die Inhomogenitäten enthalten können. Wir betrachten diesselbe Situation wie in Abschnitt 5.3 – wo wir unter anderem Fundamentalmatrizen eingeführt hatten –, allerdings gehen wir nun einen Schritt weiter und lassen auch eine Inhomogenität in der Differentialgleichung zu:

Sei  $I \subseteq \mathbb{R}$  ein nicht-leeres offenes Intervall und seien  $A : I \rightarrow \mathbb{R}^{d \times d}$  und  $b : I \rightarrow \mathbb{R}^d$  stetige Abbildungen. Wir betrachten die inhomogene lineare Differentialgleichung

$$\dot{x}(t) = A(t)x(t) + b(t). \quad (5.11)$$

Lassen Sie uns ohne Umschweife zum Hauptresultat des Abschnitts kommen:

**Satz 5.6.1 (Variation der Konstanten-Formel).** Sei  $t_0 \in I$  und sei  $X : I \ni t \mapsto X(t) \in \mathbb{R}^{d \times d}$  eine Fundamentalmatrix der homogenen Differentialgleichung  $\dot{x}(t) = A(t)x(t)$ . Für jedes  $x_0 \in \mathbb{R}^d$  ist dann die (globale) Lösung des Anfangswertproblems

$$\begin{cases} \dot{x}(t) = A(t)x(t) + b(t), \\ x(t_0) = x_0 \end{cases}$$

die Funktion  $x : I \rightarrow \mathbb{R}^d$  mit

$$x(t) = X(t)X(t_0)^{-1}x_0 + X(t) \int_{t_0}^t X(s)^{-1}b(s) \, ds$$

für alle  $t \in I$ .

Man beachte: Wenn  $X$  sogar eine Hauptfundamentalmatrix zum Startzeitpunkt  $t_0$  ist, dann vereinfacht sich der erste Summand  $X(t)X(t_0)^{-1}x_0$  in der Lösungsformel zu  $X(t)x_0$ . Die Bezeichnung *Variation der Konstanten-Formel* erklären wir nach dem Beweis in Bemerkung 5.6.2.

*Beweis von Satz 5.6.1.* Sei  $x : I \rightarrow \mathbb{R}^d$  die globale Lösung des Anfangswertproblems<sup>28</sup>. Wir definieren eine Abbildung  $c : I \rightarrow \mathbb{R}^d$  durch

$$c(t) = X(t)^{-1}x(t) \quad \text{für alle } t \in I.$$

Wir verwenden die Notation<sup>29</sup>  $GL(\mathbb{R}^d) := \{A \in \mathbb{R}^{d \times d} : A \text{ ist invertierbar}\}$ ; weil die Abbildung

$$GL(\mathbb{R}^d) \ni A \mapsto A^{-1} \in GL(\mathbb{R}^d)$$

<sup>28</sup>Zur Erinnerung: Diese existiert laut Proposition 5.1.5

<sup>29</sup>Wie Sie vielleicht aus der Linearen Algebra wissen, wird diese Menge als *Allgemeine lineare Gruppe auf  $\mathbb{R}^d$*  bezeichnet; die Abkürzung GL stammt aus dem Englischen und steht für *general linear group*.

stetig differenzierbar ist, ist auch  $c$  stetig differenzierbar.

Aus der Gleichung  $x(t) = X(t)c(t)$  und der Differentialgleichung, die  $x$  erfüllt, erhalten wir für alle  $t \in I$

$$\begin{aligned} Ax(t) + b(t) = \dot{x}(t) &= \dot{X}(t)c(t) + X(t)\dot{c}(t) \\ &= A(t)X(t)c(t) + X(t)\dot{c}(t) = Ax(t) + X(t)\dot{c}(t), \end{aligned}$$

also  $b(t) = X(t)\dot{c}(t)$ , und somit  $\dot{c}(t) = X(t)^{-1}b(t)$ . Wenn wir die Variable  $t$  durch  $s$  ersetzen und von  $t_0$  bis  $t$  nach  $s$  integrieren, erhalten wir hieraus für alle  $t \in I$

$$c(t) - c(t_0) = \int_{t_0}^t \dot{c}(s) \, ds = \int_{t_0}^t X(s)^{-1}b(s) \, ds,$$

also

$$X(t)^{-1}x(t) = c(t) = c(t_0) + \int_{t_0}^t X(s)^{-1}b(s) \, ds.$$

Den Vektor  $c(t_0)$  können wir leicht ausrechnen: Es gilt  $c(t_0) = X(t_0)^{-1}x(t_0) = X(t_0)^{-1}x_0$ . Indem wir die vorangehende Gleichung für  $X(t)^{-1}x(t)$  noch mit  $X(t)$  multiplizieren, erhalten wir also schließlich für alle  $t \in I$  die Formel

$$x(t) = X(t)X(t_0)^{-1}x_0 + X(t) \int_{t_0}^t X(s)^{-1}b(s) \, ds;$$

das ist gerade die Behauptung.  $\square$

**Bemerkung 5.6.2.** Der Name *Variation der Konstanten-Formel* hat folgenden Hintergrund: Weil  $X$  ein Fundamentalsystem für die homogene Differentialgleichung

$$\dot{x}(t) = A(t)x(t)$$

ist, ist jede Lösung zu dieser Differentialgleichung von der Form  $I \ni t \mapsto X(t)c \in \mathbb{R}^d$  für einen festen Vektor  $c \in \mathbb{R}^d$ <sup>30</sup>. Man kann den Vektor  $c$  als eine Art „Integrationskonstante“ auffassen<sup>31</sup>.

Wenn man nun die Inhomogenität  $b(t)$  in der Differentialgleichung hinzunimmt, so kann man auf die Idee kommen, eine Lösung der inhomogenen Differentialgleichung

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

<sup>30</sup>Der Vektor  $c$  hängt natürlich mit dem Startwert  $x_0$ , den man sich zu einem Zeitpunkt  $t_0 \in I$  vorgibt, zusammen, nämlich über die Formel  $c = X(t_0)^{-1}x_0$ .

<sup>31</sup>wobei es sich natürlich genau genommen um  $d$  Konstanten handelt, da  $c$  ja ein Vektor mit  $d$  Einträgen ist.

zu suchen, indem man „die Konstanten  $c$  variiert“ – sie also durch eine zeitabhängige Funktion ersetzt. Wenn man diese Idee verfolgen möchte, will man also die Lösung  $x$  der inhomogenen Gleichung nun in der Form

$$x(t) = X(t)c(t)$$

schreiben.

Dass dies tatsächlich zum Ziel führt, haben Sie im Beweis von Satz 5.6.1 gesehen, denn dort hatten wir ja gerade  $c(t) = X(t)^{-1}x(t)$  definiert, dann die Funktion  $c(t)$  berechnet, und daraus am Ende  $x(t)$  bestimmt.

Für den Fall, dass die Koeffizientenmatrix  $A$  nicht von der Zeit abhängt, können wir die Lösungsformel aus Satz 5.6.1 folgendermaßen umschreiben:

**Korollar 5.6.3 (Faltungsformel für inhomogene Differentialgleichungen).**

Es sei die Abbildung  $A$  konstant, d.h. es gelte  $A(t) = A$  für alle  $t \in \mathbb{R}^{d \times d}$  und eine feste Matrix  $A \in \mathbb{R}^{d \times d}$ .<sup>32</sup> Sei  $t_0 \in I$  und  $x_0 \in \mathbb{R}^d$ . Dann ist die (globale) Lösung des Anfangswertproblems

$$\begin{cases} \dot{x}(t) = Ax(t) + b(t), \\ x(t_0) = x_0 \end{cases}$$

die Funktion  $x : I \rightarrow \mathbb{R}^d$  mit

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

für alle  $t \in I$ .

*Beweis.* Laut Korollar 5.4.7 ist

$$X : \mathbb{R} \ni t \mapsto e^{(t-t_0)A} \in \mathbb{R}^{d \times d}$$

die zum Startzeitpunkt  $t_0$  gehörende Hauptfundamentalmatrix der homogenen Differentialgleichung  $\dot{x}(t) = Ax(t)$ , also insbesondere eine Fundamentalmatrix. Also folgt die Formel sofort aus der Lösungsformel in Satz 5.6.1.  $\square$

Lassen Sie uns kurz besprechen, weshalb man in Korollar 5.6.3 von einer *Faltungsformel* spricht:

<sup>32</sup>Hier sind wir wieder etwas unsauber mit der Notation: Eigentlich dürfe man mit dem Symbol  $A$  nicht zugleich eine konstante Abbildung  $A : I \rightarrow \mathbb{R}^{d \times d}$  und eine feste Matrix aus  $\mathbb{R}^{d \times d}$  bezeichnen.

**Bemerkung 5.6.4.** Wenn  $f, g : [0, \infty) \rightarrow \mathbb{R}$  zwei (stetige, oder allgemeiner: messbare) Funktionen mit der Eigenschaft  $\int_0^\infty |f(s)| ds < \infty$  und  $\int_0^\infty |g(s)| ds < \infty$  sind, dann kann man eine weitere Funktion von  $[0, \infty)$  nach  $\mathbb{R}$  definieren, die man als  $f \star g$  notiert und durch die Formel

$$(f \star g)(t) = \int_0^t f(t-s)g(s) ds$$

für alle  $t \in [0, \infty)$  festlegt. Die Funktion  $f \star g$  bezeichnet man als *Faltung* von  $f$  und  $g$ . Die Faltung, und verschiedene Varianten davon, spielen in zahlreichen Gebieten der Mathematik eine Rolle.<sup>33</sup>

Eine Variante der Faltung kommt auch in der Lösungsformel in Korollar 5.6.3 vor: Wenn man dort den Startzeitpunkt  $t_0$  gleich 0 wählt, dann ist das Integral, welches die Inhomogenität enthält, gleich

$$\int_0^t e^{(t-s)A} b(s) ds;$$

es handelt sich also in gewissem Sinne um eine Faltung der (matrixwertigen) Fundamentalmatrix  $t \mapsto e^{tA}$  mit der (vektorwertigen) Inhomogenität  $t \mapsto b(t)$ .

Wir diskutieren nun zwei Beispiele für die Anwendung von Korollar 5.6.3:

**Beispiele 5.6.5.** (a) Wir betrachten das zwei-dimensionale Anfangswertproblem

$$\begin{cases} \dot{x}(t) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} x(t) + \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \\ x(0) = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \end{cases}$$

auf dem Zeitintervall  $t = I$ . Laut Beispiel 5.5.2 gilt

$$\exp \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} t = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$$

für alle  $t \in \mathbb{R}$ . Also ist die Lösung  $x : \mathbb{R} \rightarrow \mathbb{R}^2$  unseres Anfangswertproblems laut der Faltungsformel aus Korollar 5.6.3 gleich

$$\begin{aligned} x(t) &= \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \begin{pmatrix} 2 \\ 0 \end{pmatrix} + \int_0^t \begin{pmatrix} \cos(t-s) & -\sin(t-s) \\ \sin(t-s) & \cos(t-s) \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} ds \\ &= \begin{pmatrix} 2 \cos t \\ 2 \sin t \end{pmatrix} + \int_0^t \begin{pmatrix} \cos(t-s) \\ \sin(t-s) \end{pmatrix} ds = \begin{pmatrix} 2 \cos t \\ 2 \sin t \end{pmatrix} + \begin{pmatrix} \int_0^t \cos(t-s) ds \\ \int_0^t \sin(t-s) ds \end{pmatrix} \end{aligned}$$

<sup>33</sup>Beispielsweise in der Fourieranalysis, in der Theorie der partiellen Differentialgleichungen und in der Stochastik.

$$= \begin{pmatrix} 2 \cos t \\ 2 \sin t \end{pmatrix} + \begin{pmatrix} -\sin(t-s)|_{s=0}^{s=t} \\ \cos(t-s)|_{s=0}^{s=t} \end{pmatrix} = \begin{pmatrix} 2 \cos t + \sin t \\ 2 \sin t + 1 - \cos t \end{pmatrix}$$

für alle  $t \in \mathbb{R}$ .

- (b) Lassen Sie uns nun noch ein ein-dimensionales Beispiel eines Anfangswertproblems betrachten, nämlich

$$\begin{cases} \dot{x}(t) = x(t) + t, \\ x(1) = 2, \end{cases}$$

auf dem Zeitintervall  $I = \mathbb{R}$ . Für die Lösung  $x$  gilt laut Korollar 5.6.3

$$\begin{aligned} x(t) &= e^{t-1} 2 + \int_1^t e^{t-s} s \, ds = 2e^{t-1} + e^t \int_1^t e^{-s} s \, ds \\ &= 2e^{t-1} + e^t \left( -e^{-s} s \Big|_{s=1}^{s=t} + \int_1^t e^{-s} \, ds \right) \\ &= 2e^{t-1} + e^t \left( -e^{-t} t + e^{-1} - e^{-t} + e^{-1} \right) = 4e^{t-1} - (t+1). \end{aligned}$$

für alle  $t \in I$ .

## 5.7 Differentialgleichungen höherer Ordnung: Ein Wiedersehen

In diesem Abschnitt wollen wir lineare Differentialgleichungen höherer Ordnung besprechen. Wir schränken uns hier auf ein-dimensionale Gleichungen ein und nehmen zudem an, dass die zugehörige homogene Gleichung autonom ist. Für solche Gleichungen kann man die Lösungen sehr präzise beschreiben.

Seien also  $a_0, \dots, a_{d-1} \in \mathbb{R}$ . Wir betrachten in diesem Abschnitt die ein-dimensionale Gleichung  $d$ -ter Ordnung

$$x^{(d)}(t) = a_0 x(t) + a_1 \dot{x}(t) + a_2 \ddot{x}(t) + \dots + a_{d-1} x^{(d-1)}(t) \quad (5.12)$$

auf dem Zeitintervall  $\mathbb{R}$ . Wenn noch zusätzlich eine Inhomogenität

$$b : I \rightarrow \mathbb{R}$$

gegeben ist, die auf einem nicht-leeren, offenen Intervall  $I \subseteq \mathbb{R}$  definiert ist, können wir auch die inhomogene Gleichung

$$x^{(d)}(t) = a_0 x(t) + a_1 \dot{x}(t) + a_2 \ddot{x}(t) + \dots + a_{d-1} x^{(d-1)}(t) + b(t) \quad (5.13)$$

auf dem Zeitintervall  $I$  betrachten. Mit der Technik aus Abschnitt 3.2 können wir diese Differentialgleichungen in Gleichungen erster Ordnung umschreiben: Wenn wir

$$z := \begin{pmatrix} x \\ \dot{x} \\ \ddot{x} \\ \vdots \\ x^{(d-1)} \end{pmatrix} : I \rightarrow \mathbb{R}^d$$

setzen, dann ist die Differentialgleichung (5.13) äquivalent zur Gleichung

$$\dot{z}(t) = \underbrace{\begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & 0 & 1 \\ a_0 & a_1 & a_1 & \dots & a_{d-1} \end{pmatrix}}_{=:A} z(t) + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ b(t) \end{pmatrix}, \quad (5.14)$$

also zu einer  $d$ -dimensionalen linearen Differentialgleichung erster Ordnung. Wenn die Inhomogenität  $b$  gleich 0 ist, fällt der rechts stehende Vektor weg, d.h. wir erhalten dann die homogene lineare Differentialgleichung

$$\dot{z}(t) = Az(t)$$

(welche wiederum sogar auf dem Zeitintervall  $\mathbb{R}$  Sinn ergibt). Man beachte, dass diese Differentialgleichung autonom ist, weil wir die Koeffizienten  $a_0, \dots, a_{d-1}$  zeit-unabhängig gewählt haben. Würden wir zeitabhängige Koeffizienten betrachten, dann wäre natürlich auch die Matrix  $A$  von der Zeit abhängig, und somit wäre die Differentialgleichung  $\dot{z}(t) = A(t)z(t)$  nicht autonom.

Lassen Sie uns der Vollständigkeit halber noch einmal präzise festhalten, wie der Lösungsraum der homogenen Differentialgleichung (5.12) mit dem Lösungsraum der homogenen Differentialgleichung  $\dot{z}(t) = Az(t)$  zusammenhängt:

**Proposition 5.7.1.** Sei  $\mathcal{L}_{\text{hom}}^{a_0, \dots, a_{d-1}} \subseteq C^d(\mathbb{R}; \mathbb{R})$  die Menge aller Lösungen der homogenen Differentialgleichung  $d$ -ter Ordnung (5.12), und sei  $\mathcal{L}_{\text{hom}}^A \subseteq C^1(\mathbb{R}; \mathbb{R}^d)$  der Lösungsraum der homogenen Differentialgleichung 1-ter Ordnung  $\dot{z}(t) = Az(t)$ .

Dann ist  $\mathcal{L}_{\text{hom}}^{a_0, \dots, a_{d-1}}$  ein Untervektorraum von  $C^d(\mathbb{R}; \mathbb{R})$ , und die Abbildung

$$x \mapsto \begin{pmatrix} x \\ \dot{x} \\ \ddot{x} \\ \vdots \\ x^{(d-1)} \end{pmatrix}$$

ist ein Isomorphismus zwischen den Vektorräumen  $\mathcal{L}_{\text{hom}}^{a_0, \dots, a_{d-1}}$  und  $\mathcal{L}_{\text{hom}}^A$ . Insbesondere besitzt  $\mathcal{L}_{\text{hom}}^{a_0, \dots, a_{d-1}}$  die Dimension  $d$ .

*Beweis.* Dass  $\mathcal{L}_{\text{hom}}^{a_0, \dots, a_{d-1}}$  ein Untervektorraum von  $C^d(\mathbb{R}; \mathbb{R})$  ist, und dass die gegebene Abbildung linear ist, rechnet man sofort nach. Dass die gegebene Abbildung bijektiv ist, folgt aus Proposition 3.2.1.  $\square$

Das Hauptresultat dieses Abschnitts ist der folgende Satz, der eine Basis des Lösungsraumes der homogenen Gleichung  $d$ -ter Ordnung (5.12) beschreibt.

**Satz 5.7.2.** Wir betrachten das Polynom  $p(\lambda) = \lambda^d - \sum_{\ell=0}^{d-1} a_\ell \lambda^\ell$ . Seien  $\lambda_1, \dots, \lambda_m \in \mathbb{R}$  die reellen Nullstellen dieses Polynoms mit Vielfachheiten  $\alpha_1, \dots, \alpha_m$  und seien  $\mu_1, \dots, \mu_n$  und  $\bar{\mu}_1, \dots, \bar{\mu}_n$  die nicht-reellen Nullstellen dieses Polynoms<sup>34</sup> mit Vielfachheiten  $\beta_1, \dots, \beta_n$ .<sup>35</sup>

Für  $k \in \{1, \dots, m\}$  und  $j \in \{0, \dots, \alpha_k - 1\}$  sei

$$e_{k,j} : \mathbb{R} \ni t \mapsto t^j e^{\lambda_k t} \in \mathbb{R},$$

und für  $k \in \{1, \dots, n\}$  und  $j \in \{0, \dots, \beta_k - 1\}$  seien

$$c_{k,j} : \mathbb{R} \ni t \mapsto t^j e^{\operatorname{Re}(\mu_k)t} \cos(\operatorname{Im}(\mu_k)t) \in \mathbb{R},$$

$$s_{k,j} : \mathbb{R} \ni t \mapsto t^j e^{\operatorname{Re}(\mu_k)t} \sin(\operatorname{Im}(\mu_k)t) \in \mathbb{R}.$$

Dann bilden diese Funktionen zusammen eine Basis des Lösungsraumes von (5.12).

Zum Beweis benötigen wir das folgende Hilfsmittel:

**Lemma 5.7.3.** Seien  $\lambda_1, \dots, \lambda_m \in \mathbb{C}$  paarweise verschiedene Zahlen und sei  $\alpha \in \mathbb{N}$ . Für jedes  $k \in \{1, \dots, m\}$  und jedes  $j \in \{0, \dots, \alpha - 1\}$  betrachten wir die Funktion

$$e_{k,j} : \mathbb{R} \ni t \mapsto t^j e^{\lambda_k t} \in \mathbb{C}$$

in  $C(\mathbb{R}; \mathbb{C})$ . Dann ist das System, das aus diesen  $m\alpha$  Funktionen besteht, linear unabhängig.

<sup>34</sup>Man beachte: Da alle Koeffizienten des Polynoms reell sind, treten alle nicht-reellen Nullstellen als konjugiert-komplexe Paare auf.

<sup>35</sup>Somit gilt also  $d = (\alpha_1 + \dots + \alpha_m) + 2(\beta_1 + \dots + \beta_n)$ .

## 5.7. Differentialgleichungen höherer Ordnung: Ein Wiedersehen

*Beweis.* Den Beweis lagern wir in eine Übungsaufgabe aus. □

Nun können wir Satz 5.7.2 beweisen:

*Beweis von Satz 5.7.2.* Wir betrachten zunächst eine beliebige Nullstelle  $\nu \in \mathbb{C}$  des Polynomes  $p$  mit Vielfachheit  $\gamma \in \mathbb{N}$ . Für  $j \in \{0, \dots, \gamma - 1\}$  betrachten wir die – eventuell komplex-wertige – Funktion

$$x : \mathbb{R} \ni t \mapsto t^j e^{\nu t} \in \mathbb{C}.$$

Wir zeigen, dass diese Funktion die Differentialgleichung (5.12) erfüllt. Dazu benutzen wir folgenden Trick: Wir schauen uns zunächst eine Funktion in zwei komplexen Variablen an, nämlich

$$\mathbb{C} \times \mathbb{C} \ni (t, \lambda) \mapsto t^j e^{\lambda t} \in \mathbb{C};$$

der Vorteil dieser Sichtweise ist, dass wir den Funktionsterm  $t^j e^{\lambda t}$  nun schreiben können als  $t^j e^{\lambda t} = \frac{\partial^j}{\partial \lambda^j} e^{\lambda t}$ , und somit gilt für jedes  $\ell \in \{0, \dots, d\}$

$$\frac{\partial^\ell}{\partial t^\ell} t^j e^{\lambda t} = \frac{\partial^\ell}{\partial t^\ell} \frac{\partial^j}{\partial \lambda^j} e^{\lambda t} = \frac{\partial^j}{\partial \lambda^j} \frac{\partial^\ell}{\partial t^\ell} e^{\lambda t} = \frac{\partial^j}{\partial \lambda^j} \lambda^\ell e^{\lambda t}.$$

Somit erhalten wir

$$\begin{aligned} \frac{\partial^d}{\partial t^d} t^j e^{\lambda t} - \sum_{\ell=0}^{d-1} a_\ell \frac{\partial^\ell}{\partial t^\ell} t^j e^{\lambda t} &= \frac{\partial^j}{\partial \lambda^j} \left( \lambda^d e^{\lambda t} - \sum_{\ell=0}^{d-1} \lambda^\ell e^{\lambda t} \right) = \frac{\partial^j}{\partial \lambda^j} (p(\lambda) e^{\lambda t}) \\ &= \sum_{\ell=0}^j \binom{j}{\ell} p^{(\ell)}(\lambda) t^{j-\ell} e^{\lambda t}. \end{aligned}$$

Weil  $\nu$  eine Nullstelle von  $p$  von Ordnung  $\gamma$  ist und  $j \leq \gamma - 1$  gilt, verschwinden die Ableitungen  $p^{(\ell)}$  für  $\ell \in \{0, \dots, j\}$  alle an der Stelle  $\nu$ . Also folgt

$$x^{(d)}(t) - \sum_{\ell=0}^{d-1} a_\ell x^{(\ell)}(t) = \sum_{\ell=0}^j \binom{j}{\ell} p^{(\ell)}(\nu) t^{j-\ell} e^{\nu t} = 0,$$

d.h.  $x$  löst tatsächlich die Differentialgleichung.

Weil alle Koeffizienten der Differentialgleichung reell sind, lösen auch der Realteil und der Imaginärteil von  $x$  die Differentialgleichung; also sind alle Funktionen, die im Satz angegeben sind, Lösungen der Differentialgleichung.

Zuletzt bemerken wird, dass aus Lemma 5.7.3 folgt, dass die im Satz angegebenen Funktionen linear unabhängig sind. Weil es sich um genau  $d$  Funktionen handelt, bilden sie also eine Basis des Lösungsraumes (denn dieser hat laut Proposition 5.7.1 die Dimension  $d$ ). □

**Bemerkung 5.7.4.** Alternativ kann man Satz 5.7.2 auch beweisen, indem man die Jordan-Normalform der Matrix  $A$  bestimmt, die am Anfang dieses Abschnitts angegeben war, und anschließend die Matrix-Exponentialfunktion berechnet.

Lassen Sie uns nun zwei Beispiele diskutieren, in denen die vorangehenden Resultate angewendet werden:

**Beispiele 5.7.5.** (a) Sei  $a \in (0, \infty)$  fest. Betrachten wir noch einmal die Differentialgleichung

$$(*) \quad \ddot{x}(t) = -x(t) + \sin(at),$$

auf dem Zeitintervall  $\mathbb{R}$ , die Sie bereits aus Aufgabe 26 b) auf Blatt 6 kennen.<sup>36</sup> Im folgenden besprechen wir, wie man diese Differentialgleichung auch mit Hilfe von Satz 5.7.2 behandeln kann.

Um eine Basis des Lösungsraums der zugehörigen homogenen Gleichung

$$\ddot{x}(t) = -x(t)$$

zu finden, können wir laut Satz 5.7.2 das Polynom  $p(\lambda) = \lambda^2 - (-1)\lambda^0 = \lambda^2 + 1$  betrachten; seine Nullstellen sind  $-i$  und  $i$  (jeweils mit Vielfachheit 1), und somit bilden die beiden Funktionen  $\cos$  und  $\sin$  eine Basis des Lösungsraumes von  $\ddot{x}(t) = -x(t)$ .

Nun verwenden wir Proposition 5.7.1: Laut dieser bilden die beiden Funktionen

$$\begin{pmatrix} \cos \\ \dot{\cos} \end{pmatrix} = \begin{pmatrix} \cos \\ -\sin \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} \sin \\ \dot{\sin} \end{pmatrix} = \begin{pmatrix} \sin \\ \cos \end{pmatrix}$$

eine Basis des Lösungsraumes der zugehörigen homogenen Gleichung erster Ordnung, die durch

$$\dot{z}(t) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} z(t)$$

gegeben ist. Es ist also

$$\mathbb{R} \ni t \mapsto X(t) := \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix} \in \mathbb{R}^{2 \times 2}$$

---

<sup>36</sup>Wir verwenden für den Streckungsfaktor im Sinus hier die Bezeichnung  $a$  statt  $\lambda$  (welche wir in Aufgabe 26 b) verwendet hatten), weil wir in Satz 5.7.2 das Symbol  $\lambda$  zur Bezeichnung der Variable des Polynoms  $p$  verwendet haben, und dies auch in diesem Beispiel tun möchten.

eine Fundamentalmatrix des homogenen Systems erster Ordnung. Nun können wir ähnlich wie in der Lösung von Aufgabe 26 b) fortfahren: Wir schreiben die ursprüngliche Gleichung (\*) als inhomogenes System erster Ordnung, nämlich als

$$\dot{z}(t) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} z(t) + \begin{pmatrix} 0 \\ \sin(at) \end{pmatrix}$$

und verwenden z.B. die Variation der Konstanten-Formel aus Satz 5.6.1 um – für gegebene Anfangsdaten – die Lösung dieser Gleichung zu bestimmen.

(b) Lassen Sie uns die Differentialgleichung

$$(*) \quad \ddot{x}(t) = -x(t) + 2\dot{x}(t) + t$$

auf dem Zeitintervall  $I$  betrachten.

Um eine Basis des Lösungsraumes der zugehörigen homogenen Differentialgleichung  $\ddot{x}(t) = -x(t) + 2\dot{x}(t)$  zu erhalten, können wir laut Satz 5.7.2 das Polynom  $P(\lambda) = \lambda^2 - 2\lambda + 1 = (\lambda - 1)^2$  betrachten; es besitzt die zweifache reelle Nullstelle 1; somit ist eine Basis des Lösungsraums von  $\ddot{x}(t) = -x(t) + 2\dot{x}(t)$  laut Satz 5.7.2 gegeben durch die beiden Funktionen

$$\mathbb{R} \ni t \mapsto e^t \in \mathbb{R},$$

$$\mathbb{R} \ni t \mapsto te^t \in \mathbb{R}.$$

Eine Fundamentalmatrix des zugehörigen homogenen Systems erster Ordnung

$$z(t) = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix} z(t)$$

ist somit aufgrund von Proposition 5.7.1 gegeben durch

$$\mathbb{R} \ni t \mapsto X(t) := \begin{pmatrix} e^t & te^t \\ \frac{d}{dt}e^t & \frac{d}{dt}te^t \end{pmatrix} = \begin{pmatrix} e^t & te^t \\ e^t & (1+t)e^t \end{pmatrix} \in \mathbb{R}^{2 \times 2}.$$

Die ursprüngliche Gleichung (\*\*) ist äquivalent zu dem System erster Ordnung

$$\dot{z}(t) = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix} z(t) + \begin{pmatrix} 0 \\ t \end{pmatrix}.$$

Da wir oben eine Fundamentalmatrix des zugehörigen homogenen Systems berechnet haben, können wir nun – falls entsprechende Anfangsdaten gegeben sind – das inhomogene System mit Hilfe der Variation der Konstanten-Formel aus Satz 5.6.1 lösen.

**Literaturstellen und verwandte Ergebnisse.** Das Thema dieses Abschnitts wird zum Beispiel in [PW19, Abschnitt 3.4.3] besprochen; allerdings werden im dortigen Satz 3.4.2 komplexwertige Lösungen zugelassen und die Formulierung des Satzes für reellwertige Lösungen nur implizit angegeben; deshalb sieht dieser Satz etwas einfacher aus als seine komplett ausformulierte reelle Variante, die wir Satz 5.7.2 besprochen haben.

## 5.8 Ergänzungen

### Die Fibonacci-Folge – eine diskrete Verwandte

In diesem Unterabschnitt kommen wir noch einmal auf die sogenannte *Fibonacci-Folge* zurück, die bereits in den Ergänzungen zu Kapitel 1 in Beispiel 1.5.3 kurz angesprochen wurde; dies ist die Folge  $(f_n)_{n \in \mathbb{N}_0}$  von natürlichen Zahlen, die durch die folgende Rekursionsvorschrift definiert ist:

$$f_0 = 1, f_1 = 1, \quad f_{n+2} = f_{n+1} + f_n \quad \text{für alle } n \in \mathbb{N}_0.$$

Während wir in Abschnitt 5.7 Differentialgleichungen höherer Ordnung besprochen hatten, handelt es sich bei der Rekursion-Vorschrift für die Fibonacci-Folge um eine *Differenzgleichung* zweiter Ordnung. Ähnlich wie für Differentialgleichungen kann man auch Differenzgleichungen höherer Ordnung in Differenzgleichungen erster Ordnung umschreiben, indem man die Dimension erhöht: Man definiert dazu

$$z_n := \begin{pmatrix} f_n \\ f_{n+1} \end{pmatrix} \quad \text{für alle } n \in \mathbb{N}_0$$

und erhält somit, dass die Folge  $(z_n)$  der Rekursionsvorschrift

$$z_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad z_{n+1} = \underbrace{\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}}_{=:A} z_n \quad \text{für alle } n \in \mathbb{N}_0$$

erfüllt. Hieraus folgt

$$z_n = A^n z_0 \quad \text{für alle } n \in \mathbb{N}_0.$$

Diese Lösungsformel ist sehr ähnlich zur Lösungsformel  $z(t) = e^{tA} z(0)$  für die Differentialgleichung  $\dot{z}(t) = Az(t)$ . Und ebenso, wie man  $e^{tA}$  für konkretes  $A$  durch Diagonalisieren von  $A$  (oder mit Hilfe der Jordan-Normalform von  $A$ , falls  $A$  nicht diagonalisierbar wäre) explizit berechnen kann, ist dies auch mit  $A^n$  möglich:

Wenn man  $A = TDT^{-1}$  für eine Diagonalmatrix  $D$  und eine invertierbare Matrix  $A$  hat, dann folgt  $A^n = TD^nT^{-1}$  für alle  $n \in \mathbb{N}_0$ . Indem man also die oben definierte Matrix

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$

diagonalisiert, kann man eine explizite Darstellung für  $z_n$  und somit für alle Folgenglieder  $f_n$  der Fibonacci-Folge berechnen. Dies ist eine Möglichkeit, um die bekannte Formel

$$f_n = \frac{1}{\sqrt{5}} \left( \left( \frac{1 + \sqrt{5}}{2} \right)^n - \left( \frac{1 - \sqrt{5}}{2} \right)^n \right) \quad \text{für alle } n \in \mathbb{N}$$

herzuleiten (beachten Sie, dass die Formel in dieser Darstellung für  $n = 0$  hingegen nicht stimmt).



## Berechnung von Lösungen im nichtlinearen Fall

### Fragen zum Einstieg.

- (a) Geben Sie eine Differentialgleichung an, die nicht linear ist!
- (b) Können Sie eine Lösung des Anfangswertproblems

$$\begin{cases} \dot{x}(t) = -x(t)^2, \\ x(0) = 0. \end{cases}$$

angeben? Falls ja: Ist die Lösung, die Sie gefunden haben, die einzige, die es gibt?

- (c) Können Sie eine Lösung des Anfangswertproblems

$$\begin{cases} \dot{x}(t) = -x(t)^2, \\ x(0) = 1. \end{cases}$$

angeben? Falls ja: Ist die Lösung, die Sie gefunden haben, die einzige, die es gibt?

### 6.1 Eine schlechte (oder gute?) Nachricht vorab

In diesem Abschnitt wollen wir nur die folgenden Bemerkung festhalten:

Für die meisten nicht-linearen Differentialgleichungen kann man keine explizite Lösungsformel angeben; oder anders ausgedrückt: in vielen Fällen hat man nicht die Möglichkeit, die Lösung eines Anfangswertproblems „analytisch“ zu berechnen. Bereits für solche fundamentalen Beispiele wie das  $N$ -Körper-Problem in der Physik<sup>1</sup> hat man keinerlei Möglichkeit die Lösungen explizit auszurechnen.

Daraus folgt, dass man folgendes tun sollte:

- Für einige (wenige) nicht-lineare Differentialgleichungen / Anfangswertprobleme kann man tatsächlich die Lösungen explizit berechnen; man sollte Methoden entwickeln um dies zu tun.

Dies ist der Inhalt dieses Kapitels 6.

<sup>1</sup>Welches die Bewegung von  $N$  Punktmassen beschreibt, die nur durch ihre Gravitation wechselwirken.

- Man sollte eine *qualitative Theorie* von Differentialgleichungen entwickeln – d.h., wenn man die Lösungen schon nicht explizit berechnen kann, sollte man zumindest Techniken entwickeln, um ihre qualitativen Eigenschaften zu untersuchen. Dies wird der Inhalt der Kapitel 7 und 8 sein.
- Man sollte eine Theorie für die *Numerik von Differentialgleichungen* entwickeln, damit man die Lösungen von Differentialgleichungen zumindest näherungsweise am Computer berechnen kann. Dies ist in der Tat ein großes und wichtiges Teilgebiet der Numerik; in dieser Vorlesung werden wir es aber nicht behandeln.

Warum ist obige Nachricht nun schlecht? Natürlich, weil es schöner wäre, möglichst explizite Lösungsformeln für möglichst viele Differentialgleichungen berechnen zu können.

Warum ist die Nachricht aber auch gut? Weil die Unmöglichkeit Objekte, die uns interessieren, explizit zu berechnen oft die Grundlage für die Entwicklung einer sehr reichhaltigen und eleganten Theorie ist mit dem Ziel diese Objekte besser verstehen zu können.<sup>2</sup>

## 6.2 Trennung der Variablen

Eine Klasse von Differentialgleichungen, die man – mehr oder weniger – explizit lösen kann, sind ein-dimensionale Differentialgleichungen erster Ordnung, deren *Variablen sich trennen lassen* (und diese beinhalten insbesondere alle autonomen ein-dimensionalen Differentialgleichungen erster Ordnung!). Hierbei ist das Wort „Variable“ folgendermaßen zu verstehen: Hat man eine Differentialgleichung der Form

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

wobei die gesuchte Funktion  $x$   $\mathbb{R}$ -wertig ist, dann kann man die Sichtweise einnehmen, dass sowohl  $t$  eine Variable ist – die *Zeitvariable* – als auch, dass  $x$  eine Variable ist; diese Sichtweise wirkt aus rein mathematischer Sicht vielleicht etwas merkwürdig, da wir ja auf dem Standpunkt stehen, dass wir eine *Funktion*  $x$  suchen. Zum Beispiel aus Sicht der Physik beschreiben  $t$  und  $x$  aber nur Größen, die in einem physikalischen Modell auftreten, und die eben mit den *Variablen*  $t$  and  $x$  bezeichnet werden.<sup>3</sup>

Mit *Trennung* der Variablen ist hier gemeint, dass man  $f(t, x(t))$  manchmal in der Form  $f(t, x(t)) = f_1(t)f_2(x(t))$  für zwei geeignete Funktionen  $f_1$  und  $f_2$

---

<sup>2</sup>Und dies ist auch für die Gewöhnlichen Differentialgleichungen der Fall.

<sup>3</sup>Das man  $x$  als explizit abhängig von der Zeit auffasst, führt dann zur der mathematischen Sichtweise, dass  $x$  eine Funktion ist.

schreiben kann. Dies ist natürlich bei weitem nicht immer möglich – aber wenn es möglich ist, dann sagt man, dass man „die Variablen trennen kann“ – und in diesem Fall ist es tatsächlich manchmal möglich die Lösungen der Differentialgleichung konkret auszurechnen.

Bevor wir die Details hierzu in einem allgemeinen Satz formulieren, illustrieren wir dieses Vorgehen zunächst in etwas heuristischer Weise an einigen Beispielen.

**Beispiele 6.2.1.** (a) Betrachten wir das Anfangswertproblem

$$\begin{cases} \dot{x}(t) = -x(t)^2, \\ x(0) = 1, \end{cases}$$

das in der Einstiegsfrage (c) dieses Kapitel auftaucht. Hier können wir tatsächlich „die Variablen trennen“, denn  $-x(t)^2$  hängt ja gar nicht explizit von  $t$  ab (sondern  $t$  taucht nur innerhalb der gesuchten Funktion  $x$  auf).

Nun kann man folgende Idee benutzen: Aus dem lokalen Existenzsatz von Picard–Lindelöf folgt, dass die Differentialgleichung zumindest lokal um 0 eine Lösung besitzt. In einem (womöglich kleinen) Intervall  $U$  um  $t = 0$  ist die Lösung zudem nicht 0, weil  $x(0) = 1$  gilt und  $x$  stetig ist. Also können wir für  $t \in U$  durch  $-x(t)^2$  teilen und erhalten somit, dass die Gleichung

$$\frac{\dot{x}(t)}{-x(t)^2} = 1 \quad \text{für alle } t \in U.$$

Wenn wir die Variable  $t$  innerhalb dieser Gleichung in  $s$  umbenennen und anschließend von 0 bis  $t$  nach  $s$  integrieren, dann folgt somit

$$\int_0^t \frac{\dot{x}(s)}{-x(s)^2} = t$$

für alle  $t \in U$ . Die linke Seite dieser Gleichung kann man mit Hilfe der Substitution  $y := x(t)$  berechnen; sie ist gleich

$$\int_{x(0)}^{x(t)} \frac{1}{-y^2} dy = \frac{1}{x(t)} - \frac{1}{x(0)} = \frac{1}{x(t)} - 1.$$

Weil dieser Term gleich  $t$  ist, erhalten durch Auflösen nach  $x(t)$  die Gleichung

$$x(t) = \frac{1}{1+t}$$

für alle  $t \in U$ .

Nun, da wir die Lösung für alle  $t$  nahe bei 0 berechnet haben, kann man umgekehrt nachrechnen, dass diese Formel sogar eine Lösung des Anfangswertproblems für alle  $t \in (-1, \infty)$  definiert.

- (b) Betrachten wir das Anfangswertproblem

$$\begin{cases} \dot{x}(t) = -x(t)^2, \\ x(0) = 0, \end{cases}$$

aus Einstiegsfrage (b) dieses Kapitels. Die Differentialgleichung ist hier diesselbe wie im vorangehenden Beispiel, aber der Anfangswert zum Startzeitpunkt  $t_0 = 0$  ist 0 anstelle von 1.

Wir können aber nicht mehr so argumentieren wie im vorangehenden Beispiel, denn die Lösung  $x$  – die aufgrund des lokalen Existenzsatzes von Picard–Lindelöf in einer Umgebung um  $t = 0$  existiert – ist zum Zeitpunkt 0 gleich 0; wir können also nicht einfach in einer Umgebung des Zeitpunktes 0 durch  $x(t)$  teilen.

Andererseits macht die Bedingung  $x(0) = 0$  die Situation hier sogar einfacher – denn man sieht ganz leicht, dass die konstante Nullfunktion  $\mathbb{R} \ni t \mapsto 0 \in \mathbb{R}$  das Anfangswertproblem löst.

- (c) Betrachten wir nun ein Beispiel einer nicht-autonomen Differentialgleichungen: Wir sehen uns das Anfangswertproblem

$$\begin{cases} \dot{x}(t) = tx(t), \\ x(0) = 1 \end{cases}$$

an. Die Differentialgleichung ist linear und homogen, d.h. wir können – wenn wir möchten – einfach die Theorie aus Kapitel 5 verwenden um das Anfangswertproblem zu lösen: Laut Proposition 5.3.8 ist die Lösung  $x : \mathbb{R} \rightarrow \mathbb{R}$  des Anfangswertproblems durch

$$x(t) = \exp\left(\int_0^t s \, ds\right) \cdot 1 = e^{\frac{1}{2}t^2}$$

für alle  $t \in \mathbb{R}$  gegeben.

Es ist aber lehrreich zu sehen, dass wir die Differentialgleichung stattdessen auch durch *Trennung der Variablen* lösen können. Wir gehen dazu vor wie in Beispiel (a) oben:

Es bezeichne  $x$  die Lösung des Anfangswertproblems. In einem Intervall  $U$  um  $t = 0$  ist  $x(t) \neq 0$ , also können wir durch  $x(t)$  teilen und von 0 bis  $t$

integrieren. Somit erhalten wir

$$\int_0^t \frac{\dot{x}(s)}{x(s)} ds = \int_0^t s ds,$$

also  $\ln \frac{|x(t)|}{|x(0)|} = \frac{1}{2}t^2$  für alle  $t \in U$  (wobei wir wieder die Substitution  $y := x(t)$  verwendet haben). Weil  $x(0)$  positiv ist und  $x$  in  $U$  keine Nullstelle hat, ist  $x(t) > 0$  für alle  $t \in U$ ; also können wir den Betrag um  $x(t)$  weglassen und erhalten somit durch Auflösen nach  $x(t)$  die Formel

$$x(t) = \exp\left(\frac{1}{2}t^2\right)x(0) = \exp\left(\frac{1}{2}t^2\right)$$

für alle  $t \in U$ . Nun kann man wiederum a posteriori nachrechnen, dass diese Formel sogar eine Funktion auf ganz  $\mathbb{R}$  definiert, welche das Anfangswertproblem löst.

Also haben wir auch auf diese Weise die Lösung des Anfangswertproblems gefunden.

Die *Trennung der Variablen*-Methode, die in den vorangehenden Beispielen illustriert wurde, kann man also folgendermaßen zusammenfassen: Wenn sich in einer ein-dimensionalen Differentialgleichung erster Ordnung die Variablen trennen lassen – d.h. wenn die Differentialgleichung sich in der Form  $\dot{x}(t) = f_1(t)f_2(x(t))$  schreiben lässt –, dann kann zur Berechnung der Lösung für die Anfangsbedingung  $x(t_0) = x_0$  folgendermaßen vorgehen:

- Man unterscheidet, ob  $f_2(x_0)$  gleich 0 oder ungleich 0 ist: Wenn  $f_2(x_0) = 0$  ist, ist die konstante Funktion  $t \mapsto x_0$  die Lösung des Anfangswertproblems. Wenn hingegen  $f_2(x_0) \neq 0$  ist, kann man folgendermaßen weitermachen:
- Man schreibt die Differentialgleichung um in die Gleichung

$$\frac{\dot{x}(t)}{f_2(x(t))} = f_1(t);$$

dies ist in einem offenen Intervall um  $t_0$  möglich, da  $f_2(x(t))$  dort nicht 0 ist. Dann benennt man die Variable um – sagen wir in  $s$  – und integriert von  $t_0$  bis  $t$  nach  $s$ .

Das Integral links kann man mit Hilfe der Substitution  $y := x(s)$  vereinfachen (wobei man ausnutzt, dass  $\dot{x}(s)$  im Integranden als Faktor vorkommt).

- Nun löst man die so erhaltene Gleichung nach  $x(t)$  auf (was rein rechnerisch durchaus ein Problem sein kann, wenn man zum Beispiel die

Stammfunktionen der auftretenden Integrale nicht explizit bestimmen kann, oder wenn die Stammfunktion so kompliziert ist, dass das Auflösen der Gleichung Schwierigkeiten macht).

- Nun hat man die Lösung für  $t$  nahe an  $t_0$  berechnet. Zuletzt überlegt man sich, auf einem Intervall welcher Größe diese Funktion sich fortsetzen lässt, und prüft dann nach, dass die fortgesetzte Funktion tatsächlich noch die Differentialgleichung löst.

Diese Beschreibung ist natürlich nachwievor sehr heuristisch, und das Fortsetzungsargument, dass am Ende nötig ist, ist – zusammen mit der notwendigen Überprüfung, ob die fortgesetzte Funktion tatsächlich noch die Differentialgleichung löst – etwas lästig.

Deshalb ist es eine gute Idee, das oben beschriebene Vorgehen mathematisch präzise in einen Satz zu fassen. Um den Satz möglichst übersichtlich zu formulieren, brauchen wir ein paar kleine Vorbereitungen. Wir beginnen mit dem Begriff *Zusammenhangskomponente*, den wir im Folgenden für Teilmengen von  $\mathbb{R}$  definieren<sup>4</sup>.

**Definition 6.2.2 (Zusammenhangskomponenten von Teilmengen von  $\mathbb{R}$ ).** Sei  $S \subseteq \mathbb{R}$  und  $s_0 \in S$ . Dann nennt man das größte Intervall in  $\mathbb{R}$ , das in  $S$  enthalten ist und das  $s_0$  enthält,<sup>5</sup> die *Zusammenhangskomponente von  $S$ , die  $s_0$  enthält*.

Wenn  $S$  offen ist, dann sieht man leicht, dass auch alle Zusammenhangskomponenten von  $S$  offen sind.

**Proposition 6.2.3 (Lokale Lipschitz-Bedingung für getrennte Variablen).** Seien  $I, J \subseteq \mathbb{R}$  zwei nichtleere offene Intervalle und seien  $f_1 : I \rightarrow \mathbb{R}$  und  $f_2 : J \rightarrow \mathbb{R}$  zwei Funktionen; es sei  $f_1$  stetig, und  $f_2$  sei lokal Lipschitz-stetig in dem Sinne, dass es für jeden Punkt  $y \in J$  eine Umgebung  $V \subseteq J$  gebe, auf der die Einschränkung  $f_2|_V$  Lipschitz-stetig ist. Dann erfüllt

$$I \times J \ni (t, y) \mapsto f_1(t)f_2(y) \in \mathbb{R}$$

die lokale Lipschitzbedingung<sup>6</sup>.

**Aufgabe 6.2.4.** Beweisen Sie Proposition 6.2.3.

Nun können wir den versprochenen Satz formulieren.

<sup>4</sup>Dieser Begriff lässt sich aber auch allgemeiner für Teilmengen eines metrischen Raumes oder sogar eines topologischen Raumes definieren.

<sup>5</sup>Man kann sich leicht überlegen, dass es solch ein größtes Intervall gibt.

<sup>6</sup>Zur Erinnerung: Den Begriff *lokale Lipschitz-Bedingung* hatten wir in Definition 4.1.6 eingeführt.

**Satz 6.2.5 (Trennung der Variablen).** Seien  $I, J$  und  $f_1, f_2$  wie in Proposition 6.2.3. Sei  $t_0 \in J$  und  $x_0 \in I$  und sei  $x : I_{\max} \rightarrow \mathbb{R}$  die Lösung des Anfangswertproblems

$$\begin{cases} \dot{x}(t) = f_1(t)f_2(x(t)), \\ x(t_0) = x_0, \end{cases}$$

die auf dem maximalen Existenzintervall  $I_{\max}$  definiert ist. Dann gilt:

- (a) Wenn  $f_2(x_0) = 0$  ist, dann ist  $I_{\max} = I$  und  $x(t) = x_0$  für alle  $t \in I_{\max}$ .
- (b) Sei nun  $f_2(x_0) \neq 0$  und sei  $\hat{J}$  die Zusammenhangskomponente von  $\{y \in J : f_2(y) \neq 0\}$ , die  $x_0$  enthält. Wir setzen

$$F_1 : I \ni t \mapsto \int_{t_0}^t f_1(s) \, ds \in \mathbb{R}, \quad F_2 : \hat{J} \ni y \mapsto \int_{x_0}^y \frac{1}{f_2(z)} \, dz \in \mathbb{R}.$$

Dann ist  $F_2$  streng monoton<sup>7</sup>, das maximale Existenzintervall  $I_{\max}$  ist gleich der Zusammenhangskomponente von

$$\{t \in I : F_1(t) \in F_2(\hat{J})\},$$

die  $t_0$  enthält, und für alle  $t \in I_{\max}$  gilt

$$x(t) = F_2^{-1}(F_1(t)).$$

Insbesondere gilt  $x(t) \in \hat{J}$  und somit  $f_2(x(t)) \neq 0$  für alle  $t \in I_{\max}$ .

Die Aussage des Satzes wird durch das oben beschriebene Verfahren zur Trennung der Variablen motiviert; allerdings enthält der Satz, wie Sie sehen, einige zusätzliche Informationen – insbesondere die genaue Beschreibung des maximalen Existenzintervalls der Lösung.

Obwohl der Beweis des Satzes nicht unbedingt schwierig ist, ist er an einigen Stellen recht technisch. Wir verzichten deshalb darauf, den Beweis innerhalb der Vorlesung zu geben. Wenn Sie möchten, können Sie den Beweis aber in den Ergänzungen zum diesem Kapitel in Abschnitt 6.5 nachlesen.

Es ist wichtig zu bemerken, dass sich bei ein-dimensionalen autonomen Differentialgleichungen immer die Variablen trennen lassen – denn für solche Gleichungen hängt die rechte Seite ja gar nicht explizit von der Zeit ab<sup>8,9</sup>.

Für solche Differentialgleichungen formulieren wir nun die Aussage von Satz 6.2.5 noch einmal explizit in einem Korollar:

<sup>7</sup>D.h.  $F_2$  ist entweder auf ganz  $\hat{J}$  streng monoton steigend oder auf ganz  $\hat{J}$  streng monoton fallend.

<sup>8</sup>D.h. wir können in diesem Fall die Funktion  $f_1$  als konstant 1 wählen.

<sup>9</sup>Sehen Sie sich zur Illustration am besten noch einmal die Beispiele 6.2.1(a) und (b) an.

**Korollar 6.2.6 (Ein-dimensionale autonome Differentialgleichungen).** Sei  $J \subseteq \mathbb{R}$  ein nichtleeres offenes Intervall und sei  $f_2 : J \rightarrow \mathbb{R}$  lokal Lipschitz-stetig<sup>10,11</sup>. Es sei  $t_0 \in \mathbb{R}$  und  $x_0 \in J$ . Für das (auf dem Zeitintervall  $\mathbb{R}$ ) definierte Anfangswertproblem

$$\begin{cases} \dot{x}(t) = f_2(x(t)), \\ x(t_0) = x_0, \end{cases}$$

sei  $x : I_{\max} \rightarrow \mathbb{R}$  diejenige Lösung, die auf dem maximalen Existenzintervall  $I_{\max}$  definiert ist. Dann gilt:

- (a) Wenn  $f_2(x_0) = 0$  ist, dann ist  $I_{\max} = \mathbb{R}$  und  $x(t) = x_0$  für alle  $t \in \mathbb{R}$ .
- (b) Sei nun  $f_2(x_0) \neq 0$  und sei  $\hat{J}$  die Zusammenhangskomponente von  $\{y \in J : f_2(y) \neq 0\}$ , die  $x_0$  enthält. Wir setzen

$$F_2 : \hat{J} \ni y \mapsto \int_{x_0}^y \frac{1}{f_2(z)} dz \in \mathbb{R}.$$

Dann ist  $F_2$  streng monoton, es ist  $I_{\max}$  gleich dem Intervall  $F_2(\hat{J}) + t_0$ , und für alle  $t \in I_{\max}$  gilt

$$x(t) = F_2^{-1}(t - t_0).$$

Insbesondere gilt  $x(t) \in \hat{J}$  und somit  $f_2(x(t)) \neq 0$  für alle  $t \in I_{\max}$ .

*Beweis.* Das Korollar folgt aus Satz 6.2.5, wenn wir dort  $I = \mathbb{R}$  setzen und  $f_1$  als die konstante 1-Funktion wählen. Man muss nur beachten, dass

$$\begin{aligned} \{t \in I : F_1(t) \in F_2(\hat{J})\} &= \{t \in I : t - t_0 \in F_2(\hat{J})\} \\ &= F_2(\hat{J}) + t_0 \end{aligned}$$

gilt – d.h. diese Menge ist selbst bereits ein Intervall.  $\square$

Zur Illustration zeigen wir noch einmal, wie man das Anfangswertproblem in Beispiel 6.2.1(a) direkt mit Hilfe von Korollar 6.2.6 lösen kann.

**Beispiel 6.2.7.** Wir betrachten nochmals das ein-dimensionale Anfangswertproblem

$$\begin{cases} \dot{x}(t) = -x(t)^2, \\ x(0) = 1 \end{cases}$$

<sup>10</sup>Den Begriff *lokal Lipschitz-stetig* hatten wir in Proposition 6.2.3 definiert.

<sup>11</sup>In Definition 3.4.1, wo wir den Begriff der *autonomen Differentialgleichung* eingeführt hatten, haben wir die rechte Seite einer autonomen Differentialgleichung mit  $h$  bezeichnet; um mit Satz 6.2.5 konsistent zu sein, nennen wir die Funktion hier lieber  $f_2$ .

auf dem Zeitintervall  $\mathbb{R}$ . Wir befinden uns somit in der Situation von Korollar 6.2.6, wobei  $f_2 : \mathbb{R} \ni y \mapsto -y^2 \in \mathbb{R}$ ,  $t_0 = 0$  und  $x_0 = 1$  ist.

In der Notation von Korollar 6.2.6 ist somit  $\hat{f} = (0, \infty)$ , und

$$F_2(y) = \frac{1}{y} - 1$$

für alle  $y \in (0, \infty)$ . Also folgt aus dem Korollar, dass

$$I_{\max} = F_2((0, \infty)) + t_0 = (-1, \infty)$$

ist, und dass die Lösung  $x : (-1, \infty) \rightarrow \mathbb{R}$  durch

$$x(t) = F_2^{-1}(t - t_0) = \frac{1}{t + 1}$$

für alle  $t \in (-1, \infty)$  gilt. Dies ist (natürlich) dasselbe Ergebnis, das wir in Beispiel 6.2.1(a) bereits berechnet hatten.

Wir schließen unseren Abschnitt zur Trennung der Variablen mit der folgenden Bemerkung.

**Bemerkung 6.2.8.** Man darf Satz 6.2.5 und Korollar 6.2.7 nicht in dem Sinne missverstehen, dass man die Lösung einer Differentialgleichung, die die Voraussetzungen des Satzes oder des Korollars erfüllt, immer explizit berechnen könnte.

Hat man zum Beispiel eine ein-dimensionale autonome Differentialgleichung erster Ordnung vorliegen und will somit Korollar 6.2.7 zur Berechnung der Lösung anwenden, so muss man zunächst das Integral berechnen, das in der Definition der Funktion  $F_2$  auftaucht, und anschließend muss man  $F_2$  invertieren. Beide Schritte lassen sich in vielen Fällen nicht unbedingt explizit durchrechnen.

**Literaturstellen und verwandte Ergebnisse.** Informationen zur Lösung von ein-dimensionalen Differentialgleichungen durch Trennung der Variablen finden Sie zum Beispiel auch in [Heu04, Abschnitt 8 in Kapitel II] und [PW19, Abschnitt 1.5].

## 6.3 Erhaltungsgrößen

Das Verfahren zur *Trennung der Variablen* aus Abschnitt 6.2 lässt sich, wie Sie gesehen haben, unter anderem zur Behandlung von ein-dimensionalen autonomen Differentialgleichungen erster Ordnung verwenden. Das Verfahren ist aber auf den ein-dimensionalen Fall eingeschränkt.

Im aktuellen Abschnitt wollen wir eine Möglichkeit besprechen, wie man auch für mehrdimensionale autonome Systeme gewisse explizite Informationen berechnen kann: wir besprechen sogenannte *Erhaltungsgrößen* einer Differentialgleichung. Man kann darüber diskutieren, ob dieser Unterabschnitt in Kapitel 6 richtig eingeordnet ist – denn Erhaltungsgrößen liefern uns nicht unbedingt eine Möglichkeit zur expliziten Berechnung einer Lösung; aber wie Sie gleich sehen werden, liefern Erhaltungsgrößen uns zumindest die Möglichkeit, gewisse Informationen über die Lösung explizit zu berechnen.

Wir beginnen direkt mit der Definition des Begriffs *Erhaltungsgröße*. Es handelt sich hierbei um eine enge Verwandte der *Ljapunov-Funktion*, die Sie bereits aus Definition 4.4.1 kennen.<sup>12</sup> Wir führen den Begriff der *Erhaltungsgröße* im selben technischen Rahmen ein wie den Begriff der *Ljapunov-Funktion* – also für im Allgemeinen nicht-autonome Differentialgleichungen erster Ordnung, wobei die rechte Seite aber auf einem Produkt aus einem Zeitintervall und einem Zustandsraum definiert ist.

An dieser Stelle sei aber bereits betont, dass Erhaltungsgrößen besonders für autonome Differentialgleichungen von Bedeutung sind.

**Definition 6.3.1 (Erhaltungsgröße).** Seien  $I \subseteq \mathbb{R}$  und  $\Omega \subseteq \mathbb{R}^d$  nichtleer und offen, wobei  $I$  ein Intervall sei. Sei  $f : I \times \Omega \rightarrow \mathbb{R}^d$  stetig und erfülle die lokale Lipschitz-Bedingung.

Eine stetige Funktion  $V : \Omega \rightarrow \mathbb{R}$  heißt *Erhaltungsgröße* oder *erstes Integral* der Differentialgleichung

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

falls für jedes nicht-triviale Intervall  $J \subseteq I$  und jede Lösung  $x : J \rightarrow \mathbb{R}^d$  dieser Differentialgleichung gilt: Die Funktion

$$V \circ x : J \ni t \mapsto V(x(t)) \in \mathbb{R}$$

ist konstant.

Wenn wir die Definition einer Erhaltungsgröße mit der Definition einer Ljapunov-Funktion vergleichen, erkennen wir sofort:

**Bemerkung 6.3.2.** Seien  $I$ ,  $\Omega$  und  $f$  wie in Definition 6.3.1; dann gilt: Eine stetige Funktion  $V : \Omega \rightarrow \mathbb{R}$  ist genau dann eine Erhaltungsgröße der Differentialgleichung

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

wenn sowohl  $V$  als auch  $-V$  eine Ljapunov-Funktion für diese Differentialgleichung ist.

<sup>12</sup>Es ist deshalb ratsam, dass Sie sich diese Definition noch einmal ansehen.

Daraus erhalten wir sofort eine Charakterisierung dafür, ob eine gegebene Funktion  $V$  ein Ljapunov-Funktion ist (unter der Zusatzvoraussetzung, dass  $V$  stetig differenzierbar ist).

**Proposition 6.3.3.** *Seien  $I$ ,  $\Omega$  und  $f$  wie in Definition 6.3.1 und sei  $V : \Omega \rightarrow \mathbb{R}^d$  stetig differenzierbar. Dann sind äquivalent:*

(i)  $V$  ist eine Erhaltungsgröße für die Differentialgleichung  $\dot{x}(t) = f(t, x(t))$ .

(ii) Für alle  $t \in I$  und alle  $y \in \Omega$  gilt  $\langle \nabla V(y), f(t, y) \rangle = 0$ .

Hierbei bezeichnet  $\langle \cdot, \cdot \rangle$  das übliche Skalarprodukt auf  $\mathbb{R}^d$ .

*Beweis.* Dies folgt sofort aus der Beschreibung von Erhaltungsgrößen mittels Ljapunov-Funktionen in Bemerkung 6.3.2 und aus der Charakterisierung von Ljapunov-Funktionen in Proposition 4.4.2.  $\square$

Erhaltungsgrößen liefern uns zwar keine Lösungsformel für ein Anfangswertproblem – aber zumindest eine konkret berechenbare Information über die Lösung des Anfangswertproblems: Wenn nämlich  $V$  eine Erhaltungsgröße ist und wir uns zusätzlich die Anfangsbedingung  $x(t_0) = x_0$  vorgeben, dann folgt natürlich für alle  $t$  im maximalen Existenzintervall, dass  $V(x(t)) = V(x_0)$  gilt – und somit verläuft die Lösung komplett in der Menge

$$\{y \in \Omega : V(y) = V(x_0)\}.$$

Wenn zum Beispiel  $d = 2$  ist und  $V$  nicht zu trivial ist, handelt es sich bei dieser Menge oft um eine Linie im  $\mathbb{R}^2$  – wir kennen dann zwar keine Formel für  $x(t)$ , aber wir wissen zumindest eine Linie, innerhalb derer die Lösung komplett verläuft.

Es ist lehrreich, dies zunächst an einem Beispiel zu betrachten, für das man die Lösung auch explizit berechnen kann:

**Beispiel 6.3.4.** Betrachten wir die zwei-dimensionale autonome Differentialgleichung

$$(*) \quad \dot{x}(t) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} x(t).$$

Weil diese Differentialgleichung linear ist, kennen Sie bereits eine Methode um sie zu lösen: Für die Anfangsbedingung  $x(0) = x_0$  mit einem Vektor  $x_0 \in \mathbb{R}^2$  erhalten wir mit Hilfe der Matrix-Exponentialfunktion die Lösung

$$x(t) = \underbrace{\begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}}_{=:A} x_0$$

for  $t \in \mathbb{R}$ ; dies bedeutet, der Startvektor  $x_0$  wird bis zum Zeitpunkt  $t$  um den Winkel  $t$  rotiert – die Lösung verläuft als im Kreis.

Vergessen wir nun für einen Augenblick, dass wir die Lösung in diesem Beispiel explizit berechnen können. Lassen Sie uns für einen Moment die Funktion  $V : \mathbb{R}^2 \ni y \mapsto \|y\|^2 \in \mathbb{R}$  betrachten<sup>13</sup>.

Die Funktion  $V$  ist eine Erhaltungsgröße der Differentialgleichung (\*); dies folgt aus Proposition 6.3.3, denn für alle  $y \in \mathbb{R}^2$  gilt

$$\langle \nabla V(y), Ay \rangle = \langle 2y, Ay \rangle = 0.$$

Das heißt, wenn  $x : \mathbb{R} \rightarrow \mathbb{R}^2$  die Lösung von (\*) mit Anfangsbedingung  $x(0) = x_0$  bezeichnet, dann gilt für alle  $t \in \mathbb{R}$

$$\|x(t)\|^2 = V(x(t)) = V(x_0) = \|x_0\|^2.$$

Dies bedeutet also, dass die Lösung in der Menge aller Punkte verläuft, die den Abstand  $\|x_0\|$  zum Nullpunkt haben – d.h. die Lösung verläuft innerhalb einer Kreislinie.

An diesem einfachen Beispiel können Sie erkennen, dass eine Erhaltungsgröße uns bereits recht explizite geometrische Informationen über die Lösungen einer Differentialgleichung liefern kann, selbst wenn wir die Lösung nicht konkret ausrechnen (oder, in anderen Beispielen, nicht konkret ausrechnen können). Andererseits muss man beachten, dass die geometrische Information, die wir mit Hilfe der Erhaltungsgröße  $V$  gewonnen haben, natürlich weniger aussagt als es die explizite Lösung der Differentialgleichung tut: Mit Hilfe der Erhaltungsgröße haben wir bisher nur erkannt, dass die Lösung innerhalb einer Kreislinie verläuft, aber nicht, in welcher Geschwindigkeit und in welcher Richtung sie das tut<sup>14</sup>.

In diesem Beispiel kann man aber tatsächlich sehr einfach auch eine Information über die Geschwindigkeit der Lösung  $x$  gewinnen: Die Matrix  $A$  ist nämlich orthogonal – d.h. sie erhält die Euklidische Norm. Wenn wir in der Differentialgleichung auf beiden Seiten die Norm nehmen, sehen wir somit, dass

$$\|\dot{x}(t)\| = \|x(t)\| = \|x_0\|$$

für alle  $t$  gilt, wobei wir im letzten Schritt verwenden haben, dass  $V$  eine Erhaltungsgröße ist. Jetzt wissen wir also nicht nur, dass die Lösung  $x$  innerhalb der Kreislinie

$$\{y \in \mathbb{R}^2 : \|y\| = \|x_0\|\}$$

<sup>13</sup>Wobei wir mit  $\|\cdot\|$  wieder die Euklidische Norm auf  $\mathbb{R}^d$  bezeichnen.

<sup>14</sup>Und das bisherige Argument, das die Erhaltungsgröße verwendet, zeigt noch nicht einmal, dass die Lösung den Kreis überhaupt einmal komplett durchläuft.

verläuft, sondern auch, dass sie dies mit konstanter Winkelgeschwindigkeit tut; wir wissen sogar, wie groß diese Geschwindigkeit ist: ihr Betrag ist gleich  $\|x_0\|$ . Wenn wir auch noch die Richtung der Bewegung von  $x$  sehen möchten, genügt es in diesem Beispiel, wenn wir uns einfach die Geometrie des Vektorfeldes ansehen, dass zur Differentialgleichung (\*) gehört.<sup>15</sup>

Wir weisen zum Schluss dieses Beispiels noch explizit darauf hin, dass sich Informationen über die Geschwindigkeit der Lösung nicht in allen Beispielen so einfach gewinnen lassen wie für die Differentialgleichung (\*).

Das wohl klassischste Beispiel für eine Erhaltungsgröße schlechthin ist die Energie in der Physik. Wir wollen sie im folgenden anhand der Newtonschen Bewegungsgleichung aus der klassischen Mechanik diskutieren. Diese Bewegungsgleichung ist zunächst eine Differentialgleichung zweiter Ordnung; aber indem wir sie in eine Differentialgleichung erster Ordnung umschreiben, können wir die Definition 6.3.1 verwenden.

**Beispiel 6.3.5 (Konservative Systeme in der klassischen Mechanik).** Die Newtonsche Bewegungsgleichung in der klassischen Mechanik ist folgendermaßen gegeben:

Es sei  $\Omega \subseteq \mathbb{R}^d$  nichtleer und offen, und es sei  $F : \Omega \rightarrow \mathbb{R}^d$  ein Vektorfeld, dass lokal Lipschitz-stetig ist<sup>16</sup>. Hierbei fasst man  $F$  aus physikalischer Sicht als ein *Kraftfeld* auf. Außerdem sei  $M \in \mathbb{R}^{d \times d}$  eine symmetrische und positiv-definite Matrix, die sogenannte *Massenmatrix* des Systems.<sup>17</sup> Die Newtonsche Bewegungsgleichung ist die  $d$ -dimensionale Gleichung zweiter Ordnung

$$(*) \quad \ddot{x}(t) = M^{-1}F(x(t)),$$

wobei wir für die Geschwindigkeit  $\dot{x}(t)$  (die in der Gleichung gar nicht explizit auftaucht) den ganzen Raum  $\mathbb{R}^d$  zulassen.

Die Bewegungsgleichung (\*) heißt *konservativ*, wenn das Kraftfeld  $F$  ein sogenanntes *Gradientenfeld* ist, d.h. wenn es eine (stetig) differenzierbare Funktion  $U : \Omega \rightarrow \mathbb{R}$  gibt mit der Eigenschaft  $F = -\nabla U$ . Dann nennt man  $U$  häufig das zugehörige *Potential* oder die zugehörige *potentielle Energie*. Für den Rest dieses Beispiels nehmen wir an, dass (\*) konservativ ist mit zugehörigem Potential  $U$ .

<sup>15</sup>Zur Erinnerung: Den Begriffs des *Vektorfeldes* und seinen Zusammenhang mit Differentialgleichungen hatten wir in Abschnitt 2.3 besprochen.

<sup>16</sup>D.h. für jedes  $y \in \Omega$  gebe es eine Umgebung  $W$  von  $y$  in  $\Omega$  derart, dass die Einschränkung von  $F$  auf  $W$  Lipschitz-stetig ist

<sup>17</sup>Aus mathematischer Sicht könnte man, wie Sie an der Gleichung (\*) sehen, die Massenmatrix auch in das Kraftfeld  $F$  mit aufnehmen; aber für die physikalische Intuition ist es vermutlich besser, wenn wir sie gesondert aufführen.

Nun schreiben wir (\*) in eine Gleichung erster Ordnung um: Mit der Notation  $v := \dot{x}$  ist die Gleichung (\*) äquivalent zu Differentialgleichung

$$(**) \quad \begin{pmatrix} \dot{x}(t) \\ \dot{v}(t) \end{pmatrix} = \begin{pmatrix} v(t) \\ M^{-1}F(x(t)) \end{pmatrix},$$

die auf  $\Omega \times \mathbb{R}^d \subseteq \mathbb{R}^d \times \mathbb{R}^d$  (und dem Zeitintervall  $\mathbb{R}$ ) definiert ist.

Für die Differentialgleichung (\*\*) kann man eine Erhaltungsgröße angeben: Es sei

$$E : \Omega \times \mathbb{R}^d \ni (y_1, y_2) \mapsto U(y_1) + \frac{1}{2} \|M^{1/2}y_2\|^2 = U(y_1) + \frac{1}{2} \langle My_2, y_2 \rangle.$$

Dann ist  $E$  eine Erhaltungsgröße für (\*\*), denn für alle  $(y_1, y_2) \in \Omega \times \mathbb{R}^d$  gilt

$$\left\langle \nabla E(y_1, y_2), \begin{pmatrix} y_2 \\ M^{-1}F(y_1) \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} U(y_1) \\ My_2 \end{pmatrix}, \begin{pmatrix} y_2 \\ -M^{-1}\nabla U(y_1) \end{pmatrix} \right\rangle = 0.$$

Wenn also  $x$  eine Lösung der Bewegungsgleichung (\*) bezeichnet, dann ist  $(x, \dot{x})$  eine Lösung der Gleichung erster Ordnung (\*\*), und somit ist die Größe

$$E(x(t), \dot{x}(t)) = U(x(t)) + \frac{1}{2} \langle M\dot{x}(t), \dot{x}(t) \rangle$$

– die sogenannte *Energie*<sup>18</sup> des Systems – unabhängig von der Zeit  $t$ .

Der Grund, warum man konservative System *konservativ* nennt, ist, dass in ihnen die Energie erhalten bleibt.

Um das recht abstrakt formulierte Beispiel 6.3.5 mit etwas Leben zu füllen, werden wir auf Übungsblatt 9 einige konkretere Beispiele für konservative Systeme in der klassischen Mechanik besprechen.

Wir schließen diesen Abschnitt mit zwei Bemerkungen:

**Bemerkung 6.3.6.** Wir betonen noch einmal, dass die Existenz einer Erhaltungsgröße nicht direkt eine Methode liefert um die Lösung eines Anfangswertproblems zu berechnen. Sie liefert aber zumindest gewisse quantitative Informationen, die konkret berechnet werden können – und die zum Beispiel benutzt werden können um geometrische Informationen über den Verlauf der Lösung zu erhalten (vgl. Beispiel 6.3.4).

**Bemerkung 6.3.7 (Noether-Theorem<sup>19</sup>).** Erhaltungsgrößen spielen in der mathematischen Physik eine zentrale Rolle (vgl. beispielsweise das einfache,

<sup>18</sup>Die Energie ist, wie Sie sehen, die Summe der beiden Zahlen  $U(x(t))$  und  $\frac{1}{2} \langle M\dot{x}(t), \dot{x}(t) \rangle$ ; die erstgenannte Größe  $U(x(t))$  bezeichnet man als die *potentielle Energie*, die zweitgenannte Größe  $\frac{1}{2} \langle M\dot{x}(t), \dot{x}(t) \rangle$  bezeichnet man als die *kinetische Energie*.

<sup>19</sup>Benannt nach Emmy Amalie Noether (1882 – 1935), deutsche Mathematikerin, die enormen Einfluss sowohl auf die Entwicklung der theoretischen Physik als auch der abstrakten Algebra hatte.

aber grundlegende Beispiel 6.3.5). Im Jahr 1918 bewies Emmy Noether, dass es eine Korrespondenz zwischen den Erhaltungsgrößen eines physikalischen Systems auf der einen Seite und Invarianzen des Systems unter Symmetrietransformationen auf der anderen Seite gibt.

Dieses Theorem wurde zu einer der Grundlagen der modernen Physik.

**Literaturstellen und verwandte Ergebnisse.** Die Energieerhaltung von Newtonschen Bewegungsgleichungen in einer Dimension wird im Falle einer Dimension sehr detailliert in [Arn92, Abschnitt 12 in Kapitel 2] besprochen. Ausführliche Informationen über Erhaltungsgrößen in der Mechanik finden Sie zudem in praktisch allen Lehrbüchern über klassische Mechanik.

## 6.4 Exakte Differentialgleichungen und integrierende Faktoren

In diesem Abschnitt diskutieren wir eine weitere Möglichkeit zur Behandlung von ein-dimensionalen nicht-autonomen Differentialgleichungen. Sei  $G \subseteq \mathbb{R}^2$  nichtleer und offen und sei  $f : G \rightarrow \mathbb{R}$ ; wir betrachten die ein-dimensionale Differentialgleichung

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

Angenommen wir können  $f$  schreiben als  $f = -\frac{M}{N}$  für zwei stetig differenzierbare Funktionen  $M, N : G \rightarrow \mathbb{R}$ , wobei  $N$  nirgends auf  $\Omega$  verschwindet<sup>20,21</sup>; dann lässt sich die Differentialgleichung  $\dot{x}(t) = f(t, x(t))$  schreiben als

$$\dot{x}(t) = -\frac{M(t, x(t))}{N(t, x(t))}, \quad (6.1)$$

und diese Differentialgleichung wiederum ist äquivalent zur implizit formulierten Differentialgleichung

$$N(t, x(t))\dot{x}(t) + M(t, x(t)) = 0. \quad (6.2)$$

Eine besondere Möglichkeit zur Lösung dieser Differentialgleichung ergibt sich in dem Spezialfall, in dem  $N$  und  $M$  die Komponenten eines Gradientenfeldes sind; in diesem Fall nennen wir die Differentialgleichung (6.2) *exakt*. Genauer:

<sup>20</sup>Das Minus zeichen schreiben wir vor den Bruch, weil dann im Folgenden die Vorzeichen etwas besser aufgehen; man könnte es natürlich auch in eine der Funktionen  $M$  oder  $N$  absorbieren.

<sup>21</sup>Für weite Teile der folgenden Argumentation würde es genügen, wenn  $M$  und  $N$  lediglich lokal Lipschitz-stetig sind.

**Definition 6.4.1 (Exakte Differentialgleichung).** Sei  $G \subseteq \mathbb{R}^2$  nichtleer und offen, seien  $M, N : G \rightarrow \mathbb{R}$  stetig differenzierbar, und  $N$  sei überall auf  $G$  ungleich 0. Die implizit formulierte ein-dimensionale Differentialgleichung (6.2) heißt *exakt*, wenn es eine (zweimal stetig) differenzierbare Funktion  $U : G \rightarrow \mathbb{R}$  gibt mit der Eigenschaft  $\partial_1 U = M$  und  $\partial_2 U = N$ .<sup>22</sup>

Sie wundern sich vielleicht, worin der Sinn dieser Definition bestehen soll. Es ist so: Die Differentialgleichung  $\dot{x}(t) = f(t, x(t))$  kann man mit einem einfachen Trick umschreiben in eine andere Differentialgleichung, die zwei-dimensional, aber dafür autonom ist; und wenn man ein Potential  $U$  wie in Definition 6.4.1 hat, dann ist  $U$  eine Erhaltungsgröße für diese zugehörige autonome zwei-dimensionale Gleichung; dies kann man manchmal benutzen um die ursprüngliche ein-dimensionale (und nicht-autonome) Differentialgleichung zu lösen.

Wir führen im folgenden die Details aus, beginnend mit folgender Aufgabe.

**Aufgabe 6.4.2.** Sei  $G \subseteq \mathbb{R}^2$  nichtleer und offen, und sei  $f : G \rightarrow \mathbb{R}$  eine Funktion. Zeigen Sie, dass die ein-dimensionale nicht-autonome Differentialgleichung

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

und die zwei-dimensionale autonome Differentialgleichung

$$\dot{z}(t) = \begin{pmatrix} 1 \\ f(z(t)) \end{pmatrix}$$

in folgendem Sinne äquivalent sind: Sei  $I \subseteq \mathbb{R}$  ein nicht-triviales Intervall.

- (a) Wenn  $x : I \rightarrow \mathbb{R}$  die ein-dimensionale nicht-autonome Differentialgleichung löst, dann löst

$$z : I \ni t \mapsto \begin{pmatrix} t \\ x(t) \end{pmatrix} \in \mathbb{R}^2$$

die zwei-dimensionale autonome Differentialgleichung.

- (b) Wenn  $z : I \rightarrow \mathbb{R}^2$  die zwei-dimensionale autonome Differentialgleichung löst, dann löst  $x := z_2$  (also die zweite Komponenten von  $z$ ) die ein-dimensionale nicht-autonome Differentialgleichung.

<sup>22</sup>Man beachte, dass – obwohl hier ebenfalls ein Gradientenfeld auftritt – diese Situation sehr verschieden ist von der Situation in Beispiel 6.3.5 (wo ebenfalls ein Gradientenfeld vorkam).

Mit Hilfe der Beobachtung in obiger Aufgabe kann man leicht sehen, was es nützt, wenn man ein Potential  $U$  wie in Definition 6.4.1 zur Verfügung hat; dies ist Inhalt der folgenden Proposition.

**Proposition 6.4.3.** *Seien  $G$ ,  $M$  und  $N$  wie in Definition 6.4.1 und sei die ein-dimensionale implizite Differentialgleichung (6.2) exakt; entsprechend sei  $U$  ebenfalls wie in Definition 6.4.1.*

*Wenn  $I \subseteq \mathbb{R}$  ein nicht-triviales Intervall und  $x : I \rightarrow \mathbb{R}$  eine Lösung von (6.2) ist, dann hängt  $U(t, x(t))$  nicht von  $t$  ab.*

*Beweis.* Wir benutzen unsere Beobachtung aus Aufgabe 6.4.2: Die Abbildung

$$z : I \ni t \mapsto \begin{pmatrix} t \\ x(t) \end{pmatrix} \in \mathbb{R}^2$$

löst die zwei-dimensionale autonome Differentialgleichung

$$\dot{z}(t) = \begin{pmatrix} 1 \\ -\frac{M(z(t))}{N(z(t))} \end{pmatrix},$$

welche man – und hier kommt nun  $U$  ins Spiel – auch in der Form

$$\dot{z}(t) = \frac{1}{\partial_2 U(z(t))} \begin{pmatrix} \partial_2 U(z(t)) \\ -\partial_1 U(z(t)) \end{pmatrix},$$

oder – noch etwas prägnanter – in der Form

$$\dot{z}(t) = \underbrace{\frac{1}{\partial_2 U(z(t))} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}}_{=: H(z(t))} \nabla U(z(t))$$

schreiben kann. Für diese Differentialgleichung ist  $U$  wegen Proposition 6.3.3 eine Erhaltungsgröße, denn es gilt

$$\langle \nabla U(y), H(y) \rangle = \frac{1}{\partial_2 U(y)} \langle \nabla U(y), \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \nabla U(y) \rangle = 0$$

für alle  $y \in G$ .

Also hängt  $U(z(t)) = U(t, x(t))$ , wie behauptet, nicht von  $t$  ab. □

Die Tatsache, dass  $U(t, x(t))$  nicht von  $t$  abhängt, kann man manchmal benutzen um  $x(t)$  zu berechnen: Wenn man die Anfangsbedingung  $x(t_0) = x_0$  (mit  $(t_0, x_0) \in G$ ) vorgibt, dann gilt für alle  $t \in I$  die Gleichheit

$$U(t, x(t)) = U(t_0, x_0),$$

und diese Gleichung kann man versuchen nach  $t$  aufzulösen. Wir demonstrieren dies in folgendem Beispiel. Wie bereits mehrmals in dieser Vorlesung schauen wir uns dazu wieder ein Beispiel an, dass auch mit einer anderen Methode, die Sie bereits kennen, gelöst werden kann – denn somit können Sie beide Methoden anhand des Beispiels vergleichen.

**Beispiel 6.4.4.** Wir betrachten das ein-dimensionale Anfangswertproblem

$$(*) \quad \begin{cases} \dot{x}(t) = -\frac{1}{t}x(t), \\ x(1) = 1 \end{cases}$$

auf dem Zeitintervall  $(0, \infty)$ . Weil die Differentialgleichung linear, homogen und ein-dimensional ist, können wir das Anfangswertproblem mit der Lösungsformel aus Proposition 5.3.8 lösen: Die Lösung  $x : (0, \infty) \rightarrow \mathbb{R}$  ist gegeben durch

$$x(t) = \exp\left(\int_1^t -\frac{1}{s} ds\right) \cdot 1 = \exp(-\ln|t|) = \frac{1}{|t|} = \frac{1}{t}$$

für alle  $t \in (0, \infty)$  (dass die Lösung auf dem Intervall  $(0, \infty)$  definiert ist, wissen wir a priori bereits aus Proposition 5.1.5).

Nun zeigen wir, wie man das Anfangswertproblem auch mit Hilfe des Konzepts der exakten Differentialgleichung lösen kann: Die Differentialgleichung können wir schreiben als  $\dot{x}(t) = f(t, x(t))$ , wobei

$$f : G := (0, \infty) \times \mathbb{R} \ni (y_1, y_2) \mapsto -\frac{1}{y_1}y_2 \in \mathbb{R}$$

ist. Wenn wir also  $M, N : G \rightarrow \mathbb{R}$  durch

$$M(y_1, y_2) = y_2 \quad \text{und} \quad N(y_1, y_2) = y_1$$

für alle  $(t, y) \in G$  definieren, dann ist  $f = -\frac{M}{N}$ . Das Vektorfeld

$$\begin{pmatrix} M \\ N \end{pmatrix} : G \ni (y_1, y_2) \mapsto \begin{pmatrix} M(y_1, y_2) \\ N(y_1, y_2) \end{pmatrix} = \begin{pmatrix} y_2 \\ y_1 \end{pmatrix}$$

ist der Gradient des Potentials  $U : G \ni (y_1, y_2) \mapsto y_1 y_2 \in \mathbb{R}$ , also ist die Differentialgleichung

$$t\dot{x}(t) + x(t) = 0,$$

die zur Differentialgleichung in (\*) äquivalent ist, exakt. Für die Lösung  $x : I_{\max} \rightarrow \mathbb{R}$  von (\*) gilt somit laut Proposition 6.4.3  $U(t, x(t)) = U(1, 1)$  für alle  $t \in I_{\max}$ ; für jedes solche  $t$  gilt also

$$tx(t) = 1,$$

und somit  $x(t) = \frac{1}{t}$ . Nun kann man direkt nachprüfen, dass die auf ganz  $(0, \infty)$  definierte Funktion  $t \mapsto \frac{1}{t}$  tatsächlich die Differentialgleichung (und natürlich auch die Anfangsbedingung) erfüllt. Also ist  $I_{\max} = (0, \infty)$  und  $x(t) = \frac{1}{t}$  für alle  $t \in (0, \infty)$ .

Anstatt die Theorie linearer Differentialgleichungen zu verwenden, kann man das Anfangswertproblem in diesem Beispiel also auch mit der Methode der exakten Differentialgleichungen lösen.

In der folgenden Bemerkung erinnern wir kurz daran, dass Sie bereits aus der Analysis II eine Methode kennen, mit der man erkennen kann, ob eine Differentialgleichung exakt ist:

**Bemerkung 6.4.5.** Seien  $G$ ,  $M$  und  $N$  wie in Definition 6.4.1. Betrachten Sie die folgenden beiden Aussagen:

- (i) Die Differentialgleichung

$$N(t, x(t))\dot{x}(t) + M(t, x(t)) = 0$$

ist exakt (d.h. anders ausgedrückt: das Vektorfeld  $(M, N)$  ist ein Gradientenfeld).

- (ii) Es gilt die sogenannte *Integrabilitätsbedingung*  $\partial_2 M = \partial_1 N$  auf ganz  $G$ .

Dann gilt stets die Implikation „(i)  $\Rightarrow$  (ii)“; dies folgt aus dem Satz von Schwarz. Wenn  $G$  *einfach zusammenhängend*<sup>23,24</sup> ist, dann gilt auch die umgekehrte Implikation „(i)  $\Leftarrow$  (ii)“.

Nun erwähnen wir noch die folgende Methode, mit der man eine nicht-exakte Differentialgleichung manchmal in eine exakte Differentialgleichung umformulieren kann:

**Bemerkung 6.4.6 (Integrierende Faktoren).** Die Grundidee dieses Abschnitts war, die rechte Seite  $f$  der ein-dimensionalen Differentialgleichung  $\dot{x}(t) = f(t, x(t))$  in der Form  $f = -\frac{M}{N}$  zu schreiben; wenn das Vektorfeld  $(M, N)$  ein Gradientenfeld ist, kann man das zugehörige Potential  $U$  berechnen und erhält dann für jede Lösung  $x$  der Differentialgleichung, dass  $U(t, x(t))$  nicht

<sup>23</sup>Wir erinnern daran, dass  $G$  *einfach zusammenhängend* heißt, wenn man jede geschlossene Kurve in  $G$  stetig zu einem Punkt zusammenziehen kann. Z.B. ist jedes Rechteck, jede Kreisfläche, und allgemeiner jede konvexe Menge einfach zusammenhängend. Es gibt natürlich auch nicht-konvexe einfach zusammenhängende Mengen (zum Beispiel die Fläche eines Viertelmondes, oder eine sternförmige Fläche).

<sup>24</sup>Beachten Sie bitte außerdem, dass die Begriffe *einfach zusammenhängend* und *zusammenhängend* verschiedene Dinge bedeuten; keine dieser beiden Eigenschaften impliziert die jeweils andere.

von  $t$  abhängt – somit kann man versuchen die Gleichung  $U(t, x(t)) = U(t_0, x_0)$  nach  $x(t)$  aufzulösen.

Wenn das Vektorfeld  $(M, N)$  kein Gradientenfeld ist, funktioniert diese Methode natürlich nicht – in anderen Worten: Die Differentialgleichung

$$N(t, x(t))\dot{x}(t) + M(t, x(t)) = 0$$

ist dann nicht exakt.

Nun können wir uns aber Folgendes überlegen: Wir haben ja nicht mit den Funktionen  $M$  und  $N$  begonnen, sondern mit der Funktion  $f$  – und die Funktionen  $M$  und  $N$  müssen dann lediglich so gewählt werden, dass  $f = -\frac{M}{N}$  gilt; diese Gleichheit ist aber natürlich bei Weitem nicht nur für eine Wahl von  $M$  und  $N$  erfüllt. Ganz im Gegenteil: Wir können beide Funktionen  $M$  und  $N$  jeweils mit dergleichen Funktion  $q : G \rightarrow \mathbb{R}$  multiplizieren; sofern  $q$  nirgends Null ist, gilt dann natürlich auch  $f = -\frac{qM}{qN}$ .

Wenn das Vektorfeld  $(M, N)$  kein Gradientenfeld ist, kann man durch geschickte Wahl von  $q$  manchmal erreichen, dass zumindest das Vektorfeld  $(qM, qN)$  ein Gradientenfeld ist; in anderen Worten: Die Differentialgleichung

$$(qN)(t, x(t))\dot{x}(t) + (qM)(t, x(t)) = 0$$

(die natürlich immer noch äquivalent zur ursprünglichen Differentialgleichung  $\dot{x}(t) = f(t, x(t))$  ist) ist nun exakt, und man kann somit die in diesem Abschnitt beschriebene Methode zur Lösung von  $\dot{x}(t) = f(t, x(t))$  anwenden.

Man nennt  $q$  in diesem Fall einen *integrierenden Faktor*<sup>25</sup>.

Wir geben im Folgenden ein einfaches Beispiel für die Verwendung solch eines integrierenden Faktors an.

**Beispiel 6.4.7.** Wir betrachten das ein-dimensionale Anfangswertproblem

$$(*) \quad \begin{cases} \dot{x}(t) = \frac{1}{t}x(t), \\ x(1) = 1, \end{cases}$$

auf dem Zeitintervall  $(0, \infty)$ , welches sich durch das Anfangswertproblem in Beispiel 6.4.4 nur durch das Vorzeichen in der Differentialgleichung unterscheidet.

Die Differentialgleichung ist wieder linear, und wir hatten diese bereits in Beispiel 5.3.10(c) besprochen: Mit Hilfe der Lösungsformel aus Proposition 5.3.8 erhalten wir die Lösung

$$x : I_{\max} = (0, \infty) \ni t \mapsto \exp\left(\int_1^t \frac{1}{s} ds\right) = t \in \mathbb{R}$$

<sup>25</sup>Die Wahl der Bezeichnung *integrierender Faktor* kann folgendermaßen interpretiert werden: Indem wir den Faktor  $q$  an  $M$  und  $N$  multiplizieren, erhalten wir das Vektorfeld  $(qM, qN)$ , das ein Gradientenfeld ist – der Faktor  $q$  sorgt also dafür, dass das Vektorfeld eine Stammfunktion besitzt.

(dass  $I_{\max} = (0, \infty)$  ist, folgt aus Proposition 5.1.5). Im Folgenden besprechen wir, wie man die Lösung stattdessen auch mit Hilfe eines integrierenden Faktors und der Methode der exakten Differentialgleichungen herleiten kann:

Die Differentialgleichung in unserem Anfangswertproblem (\*) ist  $\dot{x}(t) = f(t, x(t))$  mit

$$f : G := (0, \infty) \times \mathbb{R} \ni (y_1, y_2) \mapsto \frac{y_2}{y_1} \in \mathbb{R}.$$

Wenn wir  $M, N : G \rightarrow \mathbb{R}$  durch

$$M(y_1, y_2) = -y_2 \quad \text{und} \quad N(y_1, y_2) = y_1$$

für  $(y_1, y_2) \in G$  definieren, gilt also  $f = -\frac{M}{N}$ . Mit diesen Abbildungen  $M$  und  $N$  kommen wir aber nicht direkt weiter: Es gilt  $\partial_2 M \neq \partial_1 N$ , also ist  $(M, N)$  kein Gradientenfeld.

Wir können aber einen integrierenden Faktor verwenden: Sei  $q : G \rightarrow \mathbb{R}$ ,

$$q(y_1, y_2) = \frac{1}{y_1^2}$$

für  $(y_1, y_2) \in G$ . Dann ist das Vektorfeld

$$\begin{pmatrix} qM \\ qN \end{pmatrix} : G \ni (y_1, y_2) \mapsto \begin{pmatrix} -\frac{y_2}{y_1^2} \\ \frac{1}{y_1} \end{pmatrix} \in \mathbb{R}^2$$

ein Gradientenfeld – nämlich der Gradient des Potentials

$$U : G \ni \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \mapsto \frac{y_2}{y_1} \in \mathbb{R}.$$

Somit ist die zur Differentialgleichung in (\*) äquivalente Differentialgleichung

$$(qN)(t, x(t))\dot{x}(t) + (qM)(t, x(t)) = 0$$

exakt. Ist also  $x : I_{\max} \rightarrow \mathbb{R}$  die Lösung des Anfangswertproblems (\*), so gilt  $U(t, x(t)) = U(1, 1)$  für alle  $t \in I_{\max}$ ; dies impliziert für all solche  $t$ , dass

$$\frac{x(t)}{t} = 1, \quad \text{und somit} \quad x(t) = t$$

gilt. Nun wiederum kann man leicht nachrechnen, dass  $t \mapsto t$  die Differentialgleichung sogar auf ganz  $(0, \infty)$  löst, und erhält hieraus  $I_{\max} = (0, \infty)$ .

**Literaturstellen und verwandte Ergebnisse.** (a) Weitere Informationen zu exakten Differentialgleichungen finden Sie beispielsweise in [Heu04, Abschnitt 6 in Kapitel II].

(b) Integrierende Faktoren werden zum Beispiel in [Heu04, Abschnitt 7 in Kapitel II] näher behandelt.

## 6.5 Ergänzungen

### Beweis von Satz 6.2.5

In diesem Unterabschnitt geben wir einen Beweis von Satz 6.2.5 an, welchen wir in der Vorlesung ausgelassen hatten.

*Beweis von Satz 6.2.5.* (a) Klar ist, dass  $I_{\max}$  nicht größer als  $I$  sein kann; andererseits ist

$$I \ni t \mapsto x_0 \in \mathbb{R}$$

wegen  $f_2(x_0) = 0$  offensichtlich eine Lösung des gegebenen Anfangswertproblems.

(b) Weil  $\frac{1}{f_2}$  auf  $\hat{J}$  ungleich 0 ist und das Vorzeichen nicht wechselt, ist  $F_2$  streng monoton. Weil  $F_2$  streng monoton ist und  $\hat{J}$  offen ist, ist  $F_2(\hat{J})$  ebenfalls offen; somit ist auch die Menge

$$\{t \in I : F_1(t) \in F_2(\hat{J})\}$$

offen. Außerdem enthält diese Menge die Zeit  $t_0$  (wegen  $F_1(t_0) = 0 = F_2(x_0) \in F_2(\hat{J})$ ). Die Zusammenhangskomponente dieser Menge, die  $t_0$  enthält – bezeichnen wir sie für den Rest dieses Beweis mit  $I_0$  – ist ebenfalls offen.

Weil  $F_2$  streng monoton steigend und differenzierbar ist und die Ableitung von  $F_2$  – also die Funktion  $\frac{1}{f_2}$  – nirgends auf  $\hat{J}$  verschwindet, ist auch die Umkehrfunktion

$$F_2^{-1} : F_2(\hat{J}) \rightarrow \hat{J}$$

differenzierbar mit Ableitung

$$(F_2^{-1})'(y) = \frac{1}{F_2'(F_2^{-1}(y))} = f_2(F_2^{-1}(y))$$

für alle  $y \in F_2(\hat{J})$ . Wenn wir die Funktion

$$\hat{x} : I_0 \ni t \mapsto F_2^{-1}(F_1(t)) \in \mathbb{R}$$

definieren, dann ist diese Funktion also laut Kettenregel differenzierbar mit Ableitung

$$\frac{d}{dt} \hat{x}(t) = f_2(F_2^{-1}(F_1(t))) \cdot \frac{d}{dt} F_1(t) = f_2(\hat{x}(t)) f_1(t)$$

für alle  $t \in I_0$ . Also erfüllt  $\hat{x}$  die gegebene Differentialgleichung. Außerdem erfüllt  $\hat{x}$  die Anfangsbedingung, denn wegen  $F_1(t_0) = 0$  ist

$$\hat{x}(t_0) = F_2^{-1}(0) = x_0.$$

Also löst  $\tilde{x}$  unser Anfangswertproblem; insbesondere ist somit  $I_0 \subseteq I_{\max}$ , und  $\tilde{x}$  stimmt auf  $I_0$  mit  $x$  überein. Wir müssen nun noch zeigen, dass  $I_{\max}$  nicht größer als  $I_0$  sein kann. Dazu verwenden wir die folgenden Bezeichnungen:

- Es seien  $a, b$  der linke und der rechte Randpunkt von  $I$  (also  $-\infty \leq a < b \leq \infty$ ).
- Es seien  $a_0, b_0$  der linke und der rechte Randpunkt von  $I_0$ .
- Wie gewohnt seien  $t_-$  und  $t_+$  der linke und der rechte Randpunkt von  $I_{\max}$ .

Wir wissen bereits, dass  $a \leq t_- \leq a_0 < b_0 \leq t_+ \leq b$  gilt, und wir müssen die beiden Gleichheiten  $t_- = a_0$  und  $b_0 = t_+$  zeigen.

Lassen Sie uns zunächst den rechten Rand betrachten: Nehmen wir widerspruchshalber an, dass  $b_0 < t_+ \leq b$  gilt. Dann sind  $F_1$  und  $x$  beide am Punkt  $b_0$  definiert. Weil  $I_0$  offen ist, ist  $b_0$  kein Element von  $I_0$ . Nach Definition von  $I_0$  gilt somit  $F_1(b_0) \notin F_2(\hat{J})$ . Hieraus folgt

$$J \ni x(b_0) = \lim_{t \uparrow b_0} x(t) = \lim_{t \uparrow b_0} \tilde{x}(t) = \lim_{t \uparrow b_0} F_2^{-1}(F_1(t))$$

Wegen  $F_1(b_0) \notin F_2(\hat{J})$  folgt hieraus  $x(b_0) \notin \hat{J}$ . Also ist  $x(b_0)$  ein Punkt in  $J$ , der zugleich ein Randpunkt von  $\hat{J}$  ist, aber nicht in  $\hat{J}$  liegt – dies impliziert, dass  $f_2(x(b_0)) = 0$  gilt.

Damit erfüllt aber  $x$  das Anfangswertproblem

$$\begin{cases} \dot{x}(t) = f_1(t)f_2(x(t)), \\ x(b_0) = x(b_0); \end{cases}$$

dieses Anfangswertproblem wird aber auch von der konstanten Funktion  $I \ni t \mapsto x(b_0) \in \mathbb{R}$  erfüllt. Also folgt aus dem globalen Eindeutigkeitssatz von Picard–Lindelöf, dass  $x(t) = x(b_0)$  für alle  $t \in I_{\max}$  gilt. Insbesondere ist dann  $x_0 = x(t_0) = x(b_0)$ , was im Widerspruch zu  $f_2(x_0) \neq 0$  steht.

Somit ist  $b_0 = t_+$  gezeigt. Am linken Rand argumentiert man genauso.  $\square$

### Bernoulli-Differentialgleichung

Eine weitere Möglichkeit, mit der man einige konkrete Differentialgleichungen lösen kann, ist *Substitution*. Wir illustrieren diese Möglichkeit hier am Beispiel der *Bernoullischen Differentialgleichung*<sup>26</sup>:

<sup>26</sup>Benannt nach Jacob Bernoulli (1655 – 1705), Schweizer Mathematiker.

Sei  $I \subseteq \mathbb{R}$  ein nichtleeres offenes Intervall und seien  $\alpha, \beta : I \rightarrow \mathbb{R}$  stetig und  $\gamma \in \mathbb{R}$ . Seien  $t_0 \in I$  und  $x_0 \in (0, \infty)$ . Auf der Menge  $I \times (0, \infty)$  betrachten wir das Anfangswertproblem

$$(*) \quad \begin{cases} \dot{x}(t) = \alpha(t)x(t) + \beta(t)x(t)^\gamma, \\ x(t_0) = x_0. \end{cases}$$

Falls  $\gamma = 0$  oder  $\gamma = 1$  ist, ist die Differentialgleichung linear (homogen im Fall  $\gamma = 1$ , und inhomogen im Fall  $\gamma = 0$ , sofern  $\beta \neq 0$  ist); wir können das Anfangswertproblem in diesem Fall also mit den Methoden aus Kapitel 5 behandeln.

Also sei von nun an  $\gamma \notin \{0, 1\}$ .<sup>27</sup> In diesen Fällen kann man die Lösung folgendermaßen bestimmen:

Sei  $x : I_{\max} \rightarrow \mathbb{R}$  die Lösung des Anfangswertproblems. Da das Anfangswertproblem auf der Menge  $I \times (0, \infty)$  definiert ist, ist  $x(t) > 0$  für alle  $t \in I_{\max}$ . Also können wir eine Funktion  $z := x^{1-\gamma}$  definieren. Nun kann man leicht nachrechnen, dass  $z$  das Anfangswertproblem

$$\begin{cases} \dot{z}(t) = (1-\gamma)\alpha(t)z(t) + (1-\gamma)\beta(t), \\ z(t_0) = x_0^{1-\gamma} \end{cases}$$

löst. Die Differentialgleichung in diesem Anfangswertproblem ist linear, also kann man die Methoden aus Kapitel 5 verwenden um  $z$  zu berechnen. Hieraus kann man dann wegen  $x = z^{\frac{1}{1-\gamma}}$  die Lösung  $x$  des ursprünglichen Anfangswertproblems (\*) bestimmen.

Für weitere Informationen zur Bernoullischen Differentialgleichung sowie zu einer weiteren speziellen Klasse von Differentialgleichungen verweisen wir auf [Heu04, Abschnitt 4 in Kapitel II].

---

<sup>27</sup>Im Falle  $\gamma = 2$  ergibt die Differentialgleichung übrigens für geeignete Wahl von  $\alpha$  und  $\beta$  die logistische Gleichung, die Sie unter anderem aus Beispiel 2.1.3 kennen.

## Qualitative Eigenschaften von Lösungen

### Fragen zum Einstieg.

- (a) Betrachten Sie wieder einmal die Position  $x(t)$  des Satelliten aus den Beispielen 1.1.1(b) und 1.1.2(b).

Welche Informationen zum Zeitpunkt 0 braucht man nochmal um die Bahn des Satelliten vorherzusagen?

- (b) Wie genau kann man diese Informationen überhaupt kennen? Ist es möglich, sie völlig exakt zu bestimmen?

Was passiert mit unserer Bahnvorhersage, wenn wir bei der Messung der nötigen Informationen zum Zeitpunkt 0 einen kleinen Fehler machen?

- (c) Betrachten wir drei chemische Substanzen  $A$ ,  $B$  und  $C$ , die in einer Lösung miteinander reagieren. Die Konzentrationen der drei Substanzen zum Zeitpunkt  $t$  bezeichnen wir mit  $c_A(t)$ ,  $c_B(t)$  und  $c_C(t)$ . Die zeitliche Entwicklung der Konzentrationen kann man mit einer Differentialgleichung im  $\mathbb{R}^3$  beschreiben.

Angenommen, wir haben zum Zeitpunkt  $t = 0$  Konzentrationen mit den Werten  $c_A(0) = 2$ ,  $c_B(0) = 3.7$  und  $c_C(0) = 0.5$  vorliegen (in einer geeignet gewählten Einheit), und die Lösung der Differentialgleichung ergibt, dass wir dann zum Zeitpunkt  $t = 10$  die Konzentrationen  $c_A(10) = 0.4$ ,  $c_B(10) = -0.8$  und  $c_C(10) = 6.5$  haben.

Wie finden Sie das?

### 7.1 Fehlerfortpflanzung und des Lemma von Gronwall

In diesem Abschnitt besprechen wir, wie weit zwei Lösungen  $x$  und  $\tilde{x}$  einer Differentialgleichung höchstens auseinanderliegen können, wenn man weiß, wie weit sie zum einem Zeitpunkt  $t_0$  auseinander liegen. Wenn man  $x(t_0)$  als den „richtigen“ Startwert und  $\tilde{x}(t_0)$  als einen „gestörten“ Startwert ansieht – und  $\|\tilde{x}(t_0) - x(t_0)\|$  somit den „Fehler“ zum Zeitpunkt  $t_0$  beschreibt – dann bedeutet das eben Gesagte, dass wir untersuchen, wie sich der Fehler der Anfangsdaten im Laufe der Zeit fortpflanzt. Dies erklärt den Begriff *Fehlerfortpflanzung* in der Überschrift dieses Abschnitts.

Um die Fehlerfortpflanzung untersuchen zu können, benötigen wir als Hilfsmittel eine Version des sogenannten *Lemmas von Gronwall*, das eine Aussage über *Integralungleichungen* macht:

**Lemma 7.1.1 (Gronwall-Lemma<sup>1</sup>).** *Seien  $t_0 \in \mathbb{R}$  und sei  $I \subseteq \mathbb{R}$  ein nicht-triviales Intervall, das  $t_0$  als linken Randpunkt enthält. Zudem sei  $x_0 \in \mathbb{R}$ . Es seien  $x, A : I \rightarrow \mathbb{R}$  stetige Funktion mit  $A(t) \geq 0$  für alle  $t \in I$  und es gelte*

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

für alle  $t \in I$ . Dann gilt die Abschätzung

$$x(t) \leq \exp\left(\int_{t_0}^t A(s) \, ds\right)x_0$$

für alle  $t \in I$ .

Bevor wir das Lemma beweisen, wollen wir kurz diskutieren, wie man seine Aussage intuitiv verstehen kann. Betrachten wir dazu die Ungleichung, die im Lemma vorausgesetzt wird; wenn wir anstelle der Ungleichung eine Gleichung hätten, dann würde die Voraussetzung

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

für alle  $t \in I$  lauten. Diese Integralgleichung ist, wie wir bereits in Proposition 4.1.1 gezeigt haben, äquivalent zum Anfangswertproblem

$$\begin{cases} \dot{x}(t) = A(t)x(t), \\ x(t_0) = x_0; \end{cases}$$

und weil die Differentialgleichung in diesem Anfangswertproblem linear und homogen ist, können wir das Anfangswertproblem mit der Lösungsformel aus Proposition 5.3.8 lösen: seine Lösung lautet

$$x(t) = \exp\left(\int_{t_0}^t A(s) \, ds\right)x_0$$

für  $t \in I$ . Diese Lösungsformel ist gerade Aussage des obigen Lemmas – mit dem Unterschied, dass wir das Gleichheitszeichen wieder durch eine Ungleichung ersetzen.

---

<sup>1</sup>Manchmal auch *Grönwall-Lemma*, benannt nach Thomas Hakon Grönwall (1877 – 1932), schwedischer Mathematiker.

Im Grunde ist Lemma 7.1.1 also eine Variante von Proposition 5.3.8 für Ungleichungen anstelle von Gleichungen.<sup>2</sup>

*Beweis von Lemma 7.1.1.* Wir können ohne Einschränkung annehmen, dass  $I$  kompakt ist – denn wenn  $I$  nicht kompakt ist, betrachten wir für jedes  $b \in I$  mit  $b > t_0$  einfach das kleinere und kompakte Intervall  $[t_0, b] \subseteq I$  und zeigen die Behauptung für dieses Intervall.

Sei  $L := \sup_{t \in I} |A(t)|$ ; weil  $I$  kompakt und  $A$  stetig ist, folgt  $L < \infty$ . Wir verwenden nun die Picard-Iteration, die Sie bereits aus dem Beweis von Satz 4.1.9 und aus Aufgabe 7 auf Übungsblatt 2 kennen:

Sei  $T : C(I; \mathbb{R}) \rightarrow C(I; \mathbb{R})$  gegeben durch

$$(Ty)(t) = x_0 + \int_{t_0}^t A(s)y(s) \, ds \quad \text{für } t \in I$$

für alle  $y \in C(I; \mathbb{R})$ . Nun machen wir die folgenden drei Beobachtungen über die Abbildung  $T$ :

- (a) Für jedes  $y \in C(I; \mathbb{R})$  konvergiert die Folge der Picard-Iterierten  $(T^n y)_{n \in \mathbb{N}}$  gleichmäßig gegen die Lösung des Anfangswertproblems

$$\begin{cases} \dot{z}(t) = A(t)z(t), \\ z(t_0) = x_0, \end{cases}$$

d.h. gegen die Funktion

$$z : I \ni t \mapsto \exp\left(\int_{t_0}^t A(s) \, ds\right) x_0 \in \mathbb{R}.$$

Dies folgt wegen  $L < \infty$  aus Aufgabe 7 auf Übungsblatt 2.

- (b) Wenn  $y_1, y_2 \in C(I; \mathbb{R})$  zwei Funktionen sind, für die  $y_1 \leq y_2$  gilt (in dem Sinne, dass  $y_1(t) \leq y_2(t)$  für alle  $t \in I$  ist), dann folgt  $Ty_1 \leq Ty_2$ . Dies ist eine Konsequenz der Annahme  $A(t) \geq 0$  für alle  $t \in I$ .
- (c) Es gilt  $x \leq Tx$ ; dies ist die Ungleichung, die wir im Lemma als Voraussetzung annehmen.

<sup>2</sup>Beachten Sie bitte, dass im Gronwall-Lemma zwei Voraussetzungen gemacht werden, die in Proposition 5.3.8 nicht benötigt werden: nämlich, dass  $t_0$  der linke Randpunkt von  $I$  ist und dass  $A(t) \geq 0$  für alle  $t \in I$  gilt (vgl. hierzu auch Beispiel 7.1.2). Dass wir die Voraussetzung  $A(t) \geq 0$  für alle  $t \in I$  benötigen, liegt aber daran, dass wir die vorausgesetzte Ungleichung als Integralungleichung formuliert haben; wenn wir stattdessen eine Differentialungleichung voraussetzen, kann man die Voraussetzung  $A(t) \geq 0$  für alle  $t \in I$  weglassen.

Wegen Eigenschaft (b) können wir die Ungleichung in (c) iterieren: Aus  $x \leq Tx$  folgt  $Tx \leq T^2x$ , hieraus wiederum folgt  $T^2x \leq T^3x$ , usw. Induktiv erhalten wir somit  $x \leq Tx \leq T^2x \leq T^3x \leq \dots \leq T^n x$  für alle  $n \in \mathbb{N}$ . Weil  $T^n x$  gleichmäßig – und somit insbesondere punktweise – gegen  $z$  konvergiert, ist also  $x \leq z$ . Dies ist gerade die Behauptung des Lemmas.  $\square$

Das folgende einfache Beispiel zeigt, dass man die Voraussetzung  $A(t) \geq 0$  für alle  $t \in I$  in Lemma 7.1.1 nicht einfach weglassen kann:

**Beispiel 7.1.2.** Sei  $t_0 = 0$ ,  $I = [0, 1]$  und  $x_0 = -1$ . Außerdem sei  $A(t) = A := -1$  für alle  $t \in I$ . Wir betrachten die Funktion

$$x : I \ni t \mapsto 2(t-1) \in \mathbb{R}.$$

Dann gilt für alle  $t \in I$  die Ungleichung

$$x(t) \leq x_0 + \int_0^t Ax(s) \, ds$$

gilt, denn die rechte Seite der Ungleichung ist gleich

$$-1 - 2 \int_0^t s - 1 \, ds = -1 - t^2 + 2t = -(t-1)^2,$$

und diese Zahl ist für  $t \in [0, 1]$  größer oder gleich  $2(t-1)$ . Also ist die Ungleichung, die in Lemma 7.1.1 vorausgesetzt wird, erfüllt (die Annahme  $A(t) \geq 0$  für alle  $t \in I$  ist hingegen nicht erfüllt).

Die Ungleichung, die im Lemma gefolgert wird, stimmt jedoch nicht, denn es ist  $x(1) = 0$ , aber

$$\exp\left(\int_{t_0}^1 A(s) \, ds\right)x_0 = -\frac{1}{e}.$$

Wir können die Voraussetzung  $A(t) \geq 0$  für alle  $t \in I$  in Lemma 7.1.1 also nicht einfach weglassen.

Dass wir in Lemma 7.1.1 vorausgesetzt haben, dass  $t_0$  ein linker Randpunkt von  $I$  ist, ist lediglich der Tatsache geschuldet, dass die im Lemma auftretenden Integrale von  $t_0$  bis  $t$  das falsche Vorzeichen hätten, wenn wir  $t < t_0$  zulassen. Durch Zeitumkehr können wir jedoch leicht eine Version des Lemmas erhalten, in der wir Zeiten betrachten, die kleiner als  $t_0$  sind.

**Korollar 7.1.3 (Gronwall-Lemma nach links).** Sei  $t_0 \in \mathbb{R}$  und sei  $I \subseteq \mathbb{R}$  ein nicht-triviales Intervall, das  $t_0$  als rechten Randpunkt enthält. Es seien  $x, A : I \rightarrow \mathbb{R}$  stetige Funktionen mit  $A(t) \geq 0$  für alle  $t \in I$  und es gelte

$$x(t) \leq x_0 + \int_t^{t_0} A(s)x(s) \, ds$$

für alle  $t \in I$ . Dann gilt die Abschätzung

$$x(t) \leq \exp\left(\int_t^{t_0} A(s) ds\right) x_0$$

für alle  $t \in I$ .

*Beweis.* Dies folgt aus Lemma 7.1.1 durch Zeitumkehr. Details hierzu besprechen wir in der Übung.  $\square$

Als Konsequenz des Gronwall-Lemmas können wir nun leicht den folgenden Satz über die Fehlerfortpflanzung für Differentialgleichungen zeigen.

**Satz 7.1.4 (Fehlerfortpflanzung).** Sei  $G \subseteq \mathbb{R}^{1+d}$  offen und sei  $f : G \rightarrow \mathbb{R}^d$  eine stetige Funktion derart, dass für eine Zahl  $L \geq 0$  und für alle  $t \in \mathbb{R}$  und alle  $y_1, y_2 \in \mathbb{R}^d$  mit  $(t, y_1) \in G$  und  $(t, y_2) \in G$  die Lipschitz-Abschätzung

$$\|f(t, y_2) - f(t, y_1)\| \leq L \|y_2 - y_1\|$$

gilt. Sei  $I \subseteq \mathbb{R}$  ein nicht-triviales Intervall, sei  $t_0 \in I$  und seien  $x, \tilde{x} : I \rightarrow \mathbb{R}^d$  zwei Lösungen der Differentialgleichung

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

Dann gilt die Abschätzung

$$\|x(t) - \tilde{x}(t)\| \leq e^{L|t-t_0|} \|x(t_0) - \tilde{x}(t_0)\|.$$

für alle  $t \in I$ .

*Beweis.* Die beiden Funktionen  $x$  und  $\tilde{x}$  erfüllen die Integralgleichungen

$$x(t) = x(t_0) + \int_{t_0}^t f(s, x(s)) ds \quad \text{und} \quad \tilde{x}(t) = \tilde{x}(t_0) + \int_{t_0}^t f(s, \tilde{x}(s)) ds$$

für alle  $t \in I$ ; dies folgt aus dem Hauptsatz der Differential- und Integralrechnung (siehe auch Proposition 4.1.1). Wir betrachten nun die Abbildung

$$\varphi : I \ni t \mapsto \|x(t) - \tilde{x}(t)\| \in \mathbb{R}.$$

Wegen der Lipschitz-Bedingung an  $f$  erfüllt  $\varphi$  die Abschätzung

$$\begin{aligned} \varphi(t) &\leq \varphi(t_0) + \int_{t_0}^t L\varphi(s) ds \quad \text{für } t \geq t_0, \\ \text{und} \quad \varphi(t) &\leq \varphi(t_0) + \int_t^{t_0} L\varphi(s) ds \quad \text{für } t \leq t_0. \end{aligned}$$

Also folgt die behauptete Ungleichung

$$\varphi(t) \leq e^{L|t-t_0|} \varphi(t_0)$$

für  $t \geq t_0$  aus Lemma 7.1.1 und für  $t \leq t_0$  aus Korollar 7.1.3.  $\square$

Wir schließen diesen Abschnitt mit folgender Bemerkung.

**Bemerkung 7.1.5.** Satz 7.1.4 zeigt uns: Wenn wir den Anfangswert  $x(t_0)$  in einem Anfangswertproblem um einen Fehler der Größe  $\varepsilon$  stören, dann verändert dies die Lösung nach  $s$  Zeiteinheiten um einen Differenzvektor von Norm höchstens  $e^{L|s|}\varepsilon$  (falls die rechte Seite der Differentialgleichung die globale Lipschitzbedingung erfüllt, die wir in Satz 7.1.4 vorausgesetzt hatten).

Aus theoretischer Sicht ist es natürlich gut, solch eine Fehlerschranke überhaupt zur Verfügung zu haben. Bitte beachten Sie jedoch, dass die Fehlerschranke aus praktischer Sicht nicht unbedingt nützlich ist: Laut dieser Fehlerschranke ist es möglich, dass der Fehler im Laufe der Zeit exponentiell zunimmt! Das bedeutet zum Beispiel: Selbst wenn wir den Anfangswert  $x(t_0)$  zum Startzeitpunkt  $t_0$  nur ein klein wenig stören, kann dies noch einiger Zeit zu einer völlig anderen Lösung führen.

**Literaturstellen und verwandte Ergebnisse.** (a) Es gibt auch allgemeinere Varianten des Lemma von Gronwall als in Lemma 7.1.1; siehe zum Beispiel [PW19, Lemma 2.4.1].

(b) Resultate über *Differentialungleichungen* finden Sie zum Beispiel in [PW19, Abschnitt 2.4 und Abschnitt 4.2.1].

## 7.2 Stetige Abhängigkeit und der Begriff der Wohlgestellttheit

Von mathematischen Gleichungen, die Phänomene in der realen Welt modellieren, erwartet man häufig eine Reihe von Eigenschaften, die man unter dem Begriff *Wohlgestellttheit* zusammenfasst:

**Definition 7.2.1 (Wohlgestellttheit nach Hadamard<sup>3</sup>).** Wir nennen eine Gleichung *wohlgestellt*, wenn sie die drei folgenden Eigenschaften besitzt:

- (a) Es existiert eine Lösung.
- (b) Die Lösung ist eindeutig.
- (c) Die Lösung hängt stetig von den Daten der Gleichung ab.

Obige „Definition“ ist bewusst recht vage gehalten – denn wie genau man die drei aufgelisteten Eigenschaften versteht, hängt jeweils vom Kontext ab. Existenz- und Eindeutig von Lösungen in der Theorie der gewöhnlichen Differentialgleichungen haben wir bereits in Kapitel 4 ausführlich diskutiert.

---

<sup>3</sup>Benannt nach Jacques Hadamard (1865 – 1963), französischer Mathematiker.

In diesem Abschnitt 7.2 wollen wir nun genauer auf die dritte genannte Eigenschaften, die *stetige Abhängigkeit* der Lösungen, eingehen.

Hierbei muss man natürlich spezifizieren, über die Abhängigkeit von welchen Informationen man überhaupt sprechen möchte. Betrachten Sie zum Beispiel das  $d$ -dimensionale Anfangswertproblem

$$\begin{cases} \dot{x}(t) = f(t, x(t)), \\ x(t_0) = x_0. \end{cases}$$

Hier sind drei Daten gegeben und bestimmen die Lösung des Anfangswertproblems: Die Funktion  $f$ , der Startzeitpunkt  $t_0 \in \mathbb{R}$  und der Startwert  $x_0 \in \mathbb{R}^d$ . Im Folgenden wollen wir genauer besprechen, inwiefern die Lösung dieses Anfangswertproblems *stetig* von den Daten  $(t_0, x_0, f)$  abhängt. Dazu präzisieren wir den Begriff *stetige Abhängigkeit* wie folgt:

**Definition 7.2.2 (Stetige Abhängigkeit).** Sei  $G \subseteq \mathbb{R}^{1+d}$  offen und nichtleer und sei  $f : G \rightarrow \mathbb{R}^d$  stetig und erfülle die lokale Lipschitzbedingung. Seien  $t_0 \in \mathbb{R}$  und  $x_0 \in \mathbb{R}^d$  mit  $(t_0, x_0) \in G$  und sei  $x : I_{\max} \rightarrow \mathbb{R}^d$  die auf dem maximalen Existenzintervall definierte Lösung des Anfangswertproblems

$$\begin{cases} \dot{x}(t) = f(t, x(t)), \\ x(t_0) = x_0. \end{cases}$$

Wir sagen, dass  $x$  *stetig von den Daten*  $(t_0, x_0, f)$  *abhängt*, falls es für jedes nicht-triviale kompakte Intervall  $J \subseteq I_{\max}$  eine kompakte Umgebung  $K \subseteq G$  des Graphen

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

gibt, derart, dass für jedes  $\varepsilon > 0$  ein  $\delta > 0$  mit folgender Eigenschaft existiert:

Wenn  $\tilde{f} : G \rightarrow \mathbb{R}^d$  stetig ist und die lokale Lipschitzbedingung erfüllt, wenn  $(\tilde{t}_0, \tilde{x}_0) \in G$  ist und wenn die drei Abschätzungen

$$|\tilde{t}_0 - t_0| \leq \delta, \quad \|\tilde{x}_0 - x_0\| \leq \delta, \quad \sup_{(t,y) \in K} \|\tilde{f}(t,y) - f(t,y)\| \leq \delta$$

gelten, dann ist die Lösung  $\tilde{x} : \tilde{I}_{\max} \rightarrow \mathbb{R}^d$  des Anfangswertproblems

$$\begin{cases} \dot{\tilde{x}}(t) = \tilde{f}(t, \tilde{x}(t)), \\ \tilde{x}(\tilde{t}_0) = \tilde{x}_0. \end{cases}$$

auf ganz  $J$  definiert – d.h. es ist  $\tilde{I}_{\max} \supseteq J$  – und es gilt

$$\|\tilde{x}(t) - x(t)\| \leq \varepsilon$$

für alle  $t \in J$ .

**Bemerkung 7.2.3.** Definition 7.2.2 wirkt etwas technisch, aber im Grunde genommen ist sie nur die Formalisierung der folgenden vier Gedanken:

- (a) Man möchte gerne das ursprüngliche Anfangswertproblem etwas „stören“, indem man die Daten  $(t_0, x_0, f)$  zu  $(\tilde{t}_0, \tilde{x}_0, \tilde{f})$  abändert. Wenn die Änderung aller drei Daten klein genug ist, möchte man, dass sich auch die Lösung nicht zu sehr ändert.
- (b) Weil beim Stellen eines Anfangswertproblems a priori nicht immer klar ist, wie groß das maximale Existenzintervall ist, muss man bei der Präzisierung der Aussage „die Lösung ändert sich nicht zu sehr“ auch mit fordern, dass die Lösung  $\tilde{x}$  des gestörten Anfangswertproblems auf demselben Intervall  $J$  existiert, wie die Lösung  $x$  des ursprünglichen Anfangswertproblems.
- (c) Man kann nicht erwarten, dass die Lösung  $\tilde{x}$  des gestörten Anfangswertproblems genau das gleiche (oder ein größeres) maximale Existenzintervall hat wie  $x$ ; denn selbst wenn man das Anfangswertproblem nur ein wenig ändert, kann es natürlich passieren, dass das maximale Existenzintervall ein wenig schrumpft.

Deshalb wählt man in Definition 7.2.2 eine lokalere Betrachtungsweise, d.h. anstatt die Lösung  $x$  auf ganz  $I_{\max}$  zu betrachten, schränkt man sich auf ein kompaktes Teilintervall  $J \subseteq I_{\max}$  ein. Weil  $I_{\max}$  offen ist, sind die Randpunkte von  $J$  somit ein Stück von den Randpunkten von  $I_{\max}$  entfernt. Deshalb ist es sinnvoll zu hoffen, dass eine kleine Störung des Anfangswertproblems dazu führt, dass  $J$  noch immer im maximalen Existenzintervall der Lösung des gestörten Problems enthalten ist.

- (d) Dass man eine kompakte Umgebung  $K$  des Graphen der Einschränkung von  $x$  auf  $J$  betrachtet, liegt an folgender intuitiven Idee: Damit die Lösung  $\tilde{x}$  des gestörten Anfangswertproblems nahe an der Lösung  $x$  des ursprünglichen Anfangswertproblems liegt, sollte es reichen, wenn  $\tilde{f}$  in einer Umgebung von  $x$  nahe an  $f$  liegt.

Das zentrale Resultate über die stetige Abhängigkeit der Lösung von gewöhnlichen Differentialgleichungen ist der folgende Satz:

**Satz 7.2.4 (Stetige Abhängigkeit von den Daten).** Sei  $G \subseteq \mathbb{R}^{1+d}$  offen und nichtleer und sei  $f : G \rightarrow \mathbb{R}^d$  stetig und erfülle die lokale Lipschitzbedingung. Seien  $t_0 \in \mathbb{R}$  und  $x_0 \in \mathbb{R}^d$  mit  $(t_0, x_0) \in G$  und sei  $x : I_{\max} \rightarrow \mathbb{R}^d$  die auf dem maximalen Existenzintervall definierte Lösung des Anfangswertproblems

$$\begin{cases} \dot{x}(t) = f(t, x(t)), \\ x(t_0) = x_0. \end{cases}$$

Dann hängt  $x$  stetig von den Daten  $(t_0, x_0, f)$  ab.

Etwas prägnanter formuliert besagt der Satz also: Unter den Voraussetzungen, unter den wir die stetige Abhängigkeit in Definition 7.2.2 überhaupt definiert hatten, ist die stetige Abhängigkeit auch immer erfüllt.

Der Beweis von Satz 7.2.4 ist ein wenig technisch; er beruht unter anderem auf den folgenden Ideen:

- Man schreibt die auftretenden Anfangswertprobleme in Integralgleichungen um (wie wir dies schon häufig getan haben).
- Man verwendet folgendes Hilfsresultat: Wenn  $f$  die lokale Lipschitzbedingung erfüllt und stetig ist, dann erfüllt die Einschränkung von  $f$  auf eine kompakte Teilmenge von  $G$  sogar eine globale Lipschitzbedingung.<sup>4</sup>
- Man verwendet das Gronwall-Lemma um den Abstand von  $x$  und  $\tilde{x}$  abzuschätzen (ähnlich wie wir dies auch im Beweis des Satzes 7.1.4 über die Fehlerfortpflanzung getan hatten).

Anstatt den Beweis von Satz 7.2.4 hier im Detail zu besprechen verweisen wir auf die Literatur – zum Beispiel auf [PW19, Satz 4.1.2] – und diskutieren im Folgenden lieber zwei Beispiele um einige Aspekte der stetigen Abhängigkeit besser zu verstehen.

Zunächst weisen wir aber noch darauf hin, dass Satz 7.2.4 über die stetige Abhängigkeit und Satz 7.1.4 über die Fehlerfortpflanzung eng miteinander zusammenhängen:

**Aufgabe 7.2.5.** Beschreiben Sie den Zusammenhang zwischen den Sätzen 7.1.4 und 7.2.4.

Nun sehen wir uns einige Beispiele an.

**Beispiel 7.2.6 (Stetige Abhängigkeit vom Anfangswert).** Lassen Sie uns das sehr einfache ein-dimensionale Anfangswertproblem

$$(*) \quad \begin{cases} \dot{x}(t) = x(t), \\ x(0) = x_0 \end{cases}$$

auf dem Zeitintervall  $\mathbb{R}$  für ein festes  $x_0 \in \mathbb{R}$  betrachten. Wir befinden uns hier in der Situation der Sätze 7.1.4 und Satz 7.2.4 mit  $G = \mathbb{R}^2$ ,  $f : G \ni (t, y) \mapsto y \in \mathbb{R}$  und  $t_0 = 0$ .<sup>5</sup>

---

<sup>4</sup>Dieses Resultat können Sie zum Beispiel in [PW19, Proposition 2.1.2] nachlesen.

<sup>5</sup>Beachten Sie bitte, dass wir in Satz 7.1.4 über die Fehlerfortpflanzung sogar eine globale Lipschitz-Bedingung benötigen; diese ist in diesem Beispiel aber erfüllt (mit  $L = 1$ ).

Wir sehen uns nun an, was passiert, wenn wir von den Daten  $(t_0, x_0, f)$  nur den Anfangswert  $x_0 = 0$  ein wenig stören,  $t_0$  und  $f$  hingegen gleich lassen – wir ersetzen  $x_0$  also durch einen anderen Wert  $\tilde{x}_0$ .

Für das Anfangwertproblem (\*) kennen wir eine explizite Lösungsformel: Die Lösung  $x : \mathbb{R} \rightarrow \mathbb{R}$  ist gegeben durch

$$x(t) = e^t x_0 \quad \text{für alle } t \in \mathbb{R}$$

Wenn wir  $x_0$  durch einen anderen Startwert  $\tilde{x}_0$  ersetzen, erhalten wir hingegen die Lösung  $\tilde{x} : \mathbb{R} \rightarrow \mathbb{R}$ , die durch

$$\tilde{x}(t) = e^t \tilde{x}_0 \quad \text{für alle } t \in \mathbb{R}$$

gegeben ist. Somit erhalten wir zu jedem Zeitpunkt  $t \in \mathbb{R}$  den Abstand

$$(**) \quad |\tilde{x}(t) - x(t)| = e^t |\tilde{x}_0 - x_0|$$

zwischen  $\tilde{x}(t)$  und  $x(t)$ . Anhand dieses sehr simplen Beispiels kann man einige wichtige Dinge über die Sätze 7.2.4 und 7.1.4 lernen:

- Die Fehlerabschätzung in Satz 7.1.4 lässt sich im Allgemeinen nicht verbessern; dies erkennen Sie in diesem Beispiel an der Formel (\*\*) für  $t \geq 0$ .
- Die Fehlerabschätzung in Satz 7.1.4 ist aber im Allgemeinen auch nicht optimal; die erkennen Sie in diesem Beispiel an der Formel (\*\*) für  $t \leq 0$ .
- Dieses Beispiel zeigt, warum man sich in Satz 7.2.4 auf ein kompaktes Teilintervall  $J$  von  $I_{\max}$  einschränken muss:

In unserem Beispiel ist  $I_{\max} = \mathbb{R}$ , und für  $\tilde{x}_0 \neq x_0$  gilt stets

$$\sup_{t \in \mathbb{R}} |\tilde{x}(t) - x(t)| = \infty,$$

da der Abstand zwischen  $\tilde{x}(t)$  und  $x(t)$  für  $t \rightarrow \infty$  gegen  $\infty$  strebt. Wenn wir uns hingegen auf ein kompaktes Teilintervall  $J = [a, b] \subseteq \mathbb{R}$  einschränken, dann erhalten wir

$$\sup_{t \in \mathbb{R}} |\tilde{x}(t) - x(t)| = e^b |\tilde{x}_0 - x_0|,$$

und dieser Ausdruck geht für  $\tilde{x}_0 \rightarrow x_0$  in der Tat gegen 0.

**Beispiel 7.2.7 (Änderung der Koeffizienten bei linearen Differentialgleichungen).** Sei  $A \in \mathbb{R}^{d \times d}$  und  $x_0 \in \mathbb{R}$ . Lassen Sie uns auf dem Zeitintervall  $\mathbb{R}$  das Anfangswertproblem

$$(*) \quad \begin{cases} \dot{x}(t) = Ax(t), \\ x(0) = x_0 \end{cases}$$

betrachten. Wir befinden uns in der Situation von Satz 7.2.4 mit  $G = \mathbb{R}^{1+d}$ ,  $f : G \ni (t, y) \mapsto Ay \in \mathbb{R}^d$  und  $t_0 = 0$ .<sup>6</sup> Nun stören wir  $f$ , indem wir die Matrix  $A$  durch eine andere Matrix  $\tilde{A} \in \mathbb{R}^{d \times d}$  ersetzen; die Daten  $t_0$  und  $x_0$  lassen wir aber in diesem Beispiel gleich.

Wenn  $J \subseteq \mathbb{R}$  ein nicht-triviales kompaktes Intervall ist und  $\tilde{A}$  nahe genug bei  $A$  liegt, dann muss die Lösung  $\tilde{x}$  des Anfangswertproblems

$$\begin{cases} \dot{\tilde{x}}(t) = \tilde{A}\tilde{x}(t), \\ \tilde{x}(0) = x_0. \end{cases}$$

laut Satz 7.2.4 auf ganz  $J$  nahe an der Lösung  $x$  des Anfangswertproblems (\*) liegen.

In diesem Beispiel können wir dies auch konkret nachrechnen: Es gilt nämlich  $x(t) = e^{tA}x_0$  und  $\tilde{x}(t) = e^{\tilde{A}t}x_0$  für alle  $t \in \mathbb{R}$ , und somit

$$\|\tilde{x}(t) - x(t)\| \leq \|e^{\tilde{A}t} - e^{tA}\| \|x_0\|$$

für alle  $t \in \mathbb{R}$ . Wie in der Lösung von Bonusaufgabe 22 b) auf Übungsblatt 5 kann man sehen, dass

$$\|e^{\tilde{A}t} - e^{tA}\| \leq |t| \|\tilde{A} - A\| e^{|t| \max\{\|\tilde{A}\|, \|A\|\}}$$

für alle  $t \in \mathbb{R}$  gilt. Wenn  $t$  im kompakten Intervall  $J$  liegt und somit im Betrag durch eine Zahl  $s \in [0, \infty)$  beschränkt ist, dann folgt also

$$\|\tilde{x}(t) - x(t)\| \leq s \|\tilde{A} - A\| e^{s \max\{\|\tilde{A}\|, \|A\|\}}$$

für alle  $t \in J$ , und diese obere Schranke geht für  $\tilde{A} \rightarrow A$  gegen 0. Also haben wir die stetige Abhängigkeit von  $f$  aus Satz 7.2.4 in diesem Beispiel explizit nachgerechnet (für den Spezialfall, in dem die gestörte rechte Seite  $\tilde{f}$  ebenfalls autonom, linear und homogen ist.<sup>7</sup>)

**Literaturstellen und verwandte Ergebnisse.** Neben der oben bereits zitierten Quelle [PW19, Abschnitt 4.1] finden Sie Informationen zum Thema *stetige Abhängigkeit* zum Beispiel in [Heu04, Abschnitt 13] (im ein-dimensionalen Fall); das Resultat über stetige Abhängigkeit in dieser Quelle ist technisch etwas anders formuliert als die Ergebnisse, die wir in diesem Abschnitt besprochen haben.

<sup>6</sup>Übrigens kann man auch Satz 7.1.4 anhand dieses Beispiels gut illustrieren.

<sup>7</sup>Die Aussage von Satz 7.2.4 ist aber natürlich viel allgemeiner: Sie lässt auch nicht-lineares und nicht-autonomes  $\tilde{f}$  zu.

### 7.3 Invarianz von Mengen

In diesem Abschnitt beschäftigen wir uns mit folgender Frage: Wenn der Anfangswert  $x_0 \in \mathbb{R}^d$  eines Anfangswertproblems in einer vorgegebenen Menge  $D$  liegt, liegt dann auch die Lösung  $x(t)$  des Anfangswertproblems für alle nachfolgenden Zeiten in  $D$ ? Diese Eigenschaft präzisieren wir in der folgenden Definition.

**Definition 7.3.1 (Positive Invarianz von Mengen).** Sei  $\Omega \subseteq \mathbb{R}^d$  offen und nichtleer, und sei  $I \subseteq \mathbb{R}$  ein offenes und nichtleeres Intervall. Es sei  $f : I \times \Omega \rightarrow \mathbb{R}^d$  stetig und erfülle die lokale Lipschitz-Bedingung. Eine Menge  $D \subseteq \Omega$  heißt *positiv invariant* für die Differentialgleichung  $\dot{x}(t) = f(t, x(t))$ , falls für jedes nicht-triviale Intervall  $J \subseteq I$ , jedes Lösung  $x : J \rightarrow \Omega$  der Differentialgleichung  $\dot{x}(t) = f(t, x(t))$  und jedes  $t_0 \in I$  gilt:

Wenn  $x(t_0) \in D$  ist, dann ist auch  $x(t) \in D$  für alle  $t \in J$  mit  $t \geq t_0$ .

**Bemerkungen 7.3.2.** (a) Das Wort *positiv* taucht im Begriff *positive Invarianz* auf, weil die Menge  $D$  invariant „in positiver Zeit-Richtung“ ist, d.h. wenn  $x(t_0) \in D$  ist, dann ist auch  $x(t) \in D$  für alle  $t$ , die größer (!) als  $t_0$  sind.

(b) Prinzipiell könnte man den Begriff *positive Invarianz* auch für Differentialgleichungen definieren, deren rechte Seite  $f$  nicht auf einem Gebiet der Form  $I \times \Omega$ , sondern auf einer allgemeinen offenen und nichtleeren Menge  $G \subseteq \mathbb{R}^{1+d}$  definiert ist.

Dann ergibt aber die Annahme  $D \subseteq \Omega$  in Definition 7.3.1 keinen Sinn mehr (denn es gibt dann ja gar keine Menge  $\Omega$ ), und es stellt sich somit die Frage, welche Mengen  $D$  man in der Definition überhaupt betrachten will. Wir ersparen uns diese Schwierigkeit, indem wir in Definition 7.3.1 voraussetzen, dass  $f$  auf einer Menge der Form  $I \times \Omega$  definiert ist.

Wir sprechen kurz ein Beispiel für eine besonders wichtige Situation an, die wir im Rest des Abschnitts dann genauer untersuchen werden:

**Beispiel 7.3.3 (Positive Invarianz des positiven Kegels).** Wir betrachten eine rechte Seite  $f : \mathbb{R}^{1+d} \rightarrow \mathbb{R}^d$  (in der Notation von Definition 7.3.1 ausgedrückt ist hier also  $I = \mathbb{R}$  und  $\Omega = \mathbb{R}^d$ ). Als Menge  $D$  wählen wir den sogenannten *positiven Kegel* von  $\mathbb{R}^d$ , d.h. die Menge

$$D := \mathbb{R}_+^d := \{y \in \mathbb{R}^d : y_k \geq 0 \text{ für alle } k \in \{1, \dots, d\}\}.$$

Die positive Invarianz dieser Menge ist in vielen Anwendungen aus folgendem Grund wichtig:

Häufig beschreiben die Komponenten  $x_1, \dots, x_d$  der Lösung  $x$  bestimmte physikalische, chemische oder biologische Größen, die in manchen Modellen

nur einen Sinn ergeben, wenn sie mindestens 0 sind (z.B. Teilchenzahlen, Konzentrationen oder Anzahlen von Individuen mit einem bestimmten Merkmal). Wenn man sich eine Anfangsbedingung  $x(t_0) = x_0$  vorgibt, dann bedeutet  $x_0(t) \in \mathbb{R}_+^d$  also gerade, dass der Anfangswert  $x_0$  im Sinne der Anwendung, die man modellieren möchte, Sinn ergibt. Dann muss aber auch  $x(t) \in \mathbb{R}_+^d$  für alle Zeiten  $t \geq t_0$  gelten, denn ein physikalisch/chemisch/biologisch sinnvoller Anfangswert darf nicht zu einem späteren Zeitpunkt zu einem Wert führen, der aus physikalischer/chemischer/biologischer Sicht keinen Sinn mehr ergibt.

Wenn diese Bedingung der positiven Invarianz von  $\mathbb{R}_+^d$  in solch einer Anwendung nicht erfüllt ist, so ist dies ein deutliches Anzeichen dafür, dass man ein ungeignetes Modell verwendet. Im Umkehrschluss möchte man die positive Invarianz von  $\mathbb{R}_+^d$  in solchen Anwendungen nachweisen können, um sicherzustellen, dass das verwendete Modell nicht aus dem beschriebenen Grund ungeeignet ist.

Vergleichen Sie hierzu bitte auch die Einstiegsfrage (c) des aktuellen Kapitels 7.

**Bemerkung 7.3.4.** Die *Positivitäts-Erhaltung*, die in Beispiel 7.3.3 gefordert wurde, könnte man auch direkt in die Differentialgleichung einbauen – nämlich indem man den Definitionsbereich von  $f$  auf  $\mathbb{R} \times \mathbb{R}_+^d$  einschränkt. Das führt aber zu zwei Problemen:

- Die Menge  $\mathbb{R} \times \mathbb{R}_+^d$  ist nicht offen, aber die komplette Theorie, die wir bisher entwickelt haben, funktioniert nur, wenn die Funktion  $f$  auf einer offenen Menge definiert ist. Wir können  $f$  also gar nicht wirklich auf  $\mathbb{R} \times \mathbb{R}_+^d$  einschränken; wenn, dann müssten wir uns auf die offene Menge  $\mathbb{R} \times C$  einschränken, wobei

$$C := \{y \in \mathbb{R}^d : y_k > 0 \text{ für alle } k \in \{1, \dots, d\}\}$$

das Innere von  $\mathbb{R}_+^d$  ist. Dies ist prinzipiell möglich, führt aber dazu, dass wir keine Lösungen mehr betrachten können, bei denen manche Komponenten 0 sind; diese Einschränkung ist etwas unnatürlich.

- Angenommen, wir schränken  $f$  tatsächlich auf die Menge  $\mathbb{R} \times C$  ein – dann liegen Lösungen der Differentialgleichung  $\dot{x}(t) = f(t, x(t))$  per Definition in  $C$ . Damit haben wir aber das Problem nur verlagert:

Würde nämlich die Lösung der auf ganz  $\mathbb{R}^{1+d}$  definierten Differentialgleichung zu einem Zeitpunkt die Menge  $t_1$  die Menge  $C$  verlassen, so bedeutet dies, dass die Lösung der auf  $\mathbb{R} \times C$  definierten Differentialgleichung an diesem Zeitpunkt  $t_1$  aufhört zu existieren. Durch das Einschränken des Definitionsbereichs von  $f$  erreichen wir also nur, dass

die Frage der positiven Invarianz umgewandelt wird in die Frage, wie weit die Lösung auf dem kleineren Zustandsraum nach rechts existiert.<sup>8</sup>

Wir wollen nun charakterisieren, wann der positive Kegel aus Beispiel 7.3.3 positiv invariant ist. Dazu verwenden wir den folgenden Begriff:

**Definition 7.3.5 (Quasi-positiv Funktion).** Eine Funktion  $f : \mathbb{R}^{1+d} \rightarrow \mathbb{R}^d$  heißt *quasi-positiv*, falls für jedes  $t \in \mathbb{R}$  und jedes  $y \in \mathbb{R}_+^d$  gilt:

Für alle  $k \in \{1, \dots, d\}$  mit  $y_k = 0$  ist  $f_k(t, y) \geq 0$  (wobei  $f_k$  die  $k$ -te Komponenten der Funktion  $f$  bezeichnet).

Quasi-Positivität lässt sich mit Hilfe des Standard-Skalarproduktes auf  $\mathbb{R}^d$  folgendermaßen beschreiben:

**Proposition 7.3.6.** Für jede Funktion  $f : \mathbb{R}^{1+d} \rightarrow \mathbb{R}^d$  sind äquivalent:

(i) Die Funktion  $f$  ist quasi-positiv.

(ii) Für alle  $t \in \mathbb{R}$  und alle  $y, z \in \mathbb{R}_+^d$  mit  $\langle z, y \rangle = 0$  gilt  $\langle z, f(t, y) \rangle \geq 0$ .

*Beweis.* „(i)  $\Rightarrow$  (ii)“ Sei  $f$  quasi-positiv, sei  $t \in \mathbb{R}$  und seien  $y, z \in \mathbb{R}_+^d$  mit  $\langle z, y \rangle = 0$ . Sei  $S \subseteq \{1, \dots, d\}$  die Menge aller Indizes  $k$ , für die gilt  $z_k > 0$  gilt. Dann ist  $y_k = 0$  für alle  $k \in S$ , und somit ist  $f_k(t, y) \geq 0$  für alle  $k \in S$ . Hieraus folgt

$$\langle z, f(t, y) \rangle = \sum_{k \in S} z_k f_k(t, y) \geq 0.$$

Dies zeigt (ii)

„(ii)  $\Rightarrow$  (i)“ Es gelte (ii), und es sei  $t \in \mathbb{R}$  und  $y \in \mathbb{R}_+^d$ . Sei  $k \in \{1, \dots, d\}$  mit  $y_k = 0$ .

Es bezeichne  $e_k \in \mathbb{R}^d$  den  $k$ -ten kanonischen Einheitsvektor. Dann ist  $e_k \in \mathbb{R}_+^d$  und  $\langle e_k, y \rangle = y_k = 0$ , also folgt aus (ii)

$$f_k(t, y) = \langle e_k, f(t, y) \rangle \geq 0.$$

Somit ist  $f$  quasi-positiv. □

Mit Hilfe von quasi-Positivität können wir die positive Invarianz des positiven Kegels charakterisieren:

**Satz 7.3.7.** Sei  $f : \mathbb{R}^{1+d} \rightarrow \mathbb{R}^d$  stetig und erfülle die lokale Lipschitz-Bedingung. Dann sind äquivalent:

<sup>8</sup>Diese Vorgehensweise kann natürlich durchaus zum Ziel führen; wir hatten sie in den Beispielen 4.3.10 und 4.4.4 eingesetzt um die logistische Gleichung, die eine Bakterienpopulation beschreibt, auf dem Zustandsraum  $C = (0, \infty)$  zu untersuchen.

- (i) Der positive Kegel  $\mathbb{R}_+^d$  ist positiv invariant für die Differentialgleichung  $\dot{x}(t) = f(t, x(t))$ .
- (ii) Die Funktion  $f$  ist quasi-positiv.

*Beweis von „(i)  $\Rightarrow$  (ii)“.* Wir nehmen widerspruchshalber an, dass  $f$  nicht quasi-positiv ist. Dann gibt es eine Zeit  $t_0 \in \mathbb{R}$ , einen Vektor  $y \in \mathbb{R}_+^d$  und einen Index  $k \in \{1, \dots, d\}$  mit  $y_k = 0$  und  $f_k(y) < 0$ . Sei nun  $x : I_{\max} \rightarrow \mathbb{R}^d$  die Lösung des Anfangswertproblems

$$\begin{cases} \dot{x}(t) = f(t, x(t)), \\ x(t_0) = y. \end{cases}$$

Dann gilt  $x(t_0) \in \mathbb{R}_+^d$ . Es ist aber  $\dot{x}_k(t_0) = f_k(t, x(t_0)) = f_k(t, y) < 0$ , und somit gilt für alle  $t$  in einer kleinen punktierten rechtsseitigen Umgebung von  $t_0$  die Ungleichung  $x_k(t) < x_k(t_0) = y_k = 0$ , d.h.  $x(t) \notin \mathbb{R}_+^d$ . Dies widerspricht der positiven Invarianz von  $\mathbb{R}_+^d$ .  $\square$

Für den Beweis der Rückrichtung verwenden wir das folgende Hilfsresultat:

**Lemma 7.3.8.** *Es bezeichne*

$$C := \{y \in \mathbb{R}^d : y_k > 0 \text{ für alle } k \in \{1, \dots, k\}\}$$

das Innere des positiven Kegels  $\mathbb{R}_+^d$  in  $\mathbb{R}^d$ . Sei  $f : \mathbb{R}^{1+d} \rightarrow \mathbb{R}^d$  stetig und erfülle die lokale Lipschitz-Bedingung. Außerdem erfülle  $f$  die folgende verschärfte Version von quasi-Positivität: Für jedes  $t \in \mathbb{R}$ , jedes  $y \in \mathbb{R}_+^d$  und jedes  $k \in \{1, \dots, d\}$  mit  $y_k = 0$  gelte  $f_k(t, y) > 0$ .

Dann ist  $C$  positiv invariant für die Differentialgleichung  $\dot{x}(t) = f(t, x(t))$ .

*Beweis.* Angenommen,  $C$  wäre nicht positiv invariant. Dann gibt es ein nicht-triviales Intervall  $J \subseteq \mathbb{R}$ , eine Lösung  $x : J \rightarrow \mathbb{R}^d$  von  $\dot{x}(t) = f(t, x(t))$  sowie Zeiten  $t_0, t \in J$  mit  $t > t_0$  derart, dass  $x(t_0) \in C$ , aber  $x(t) \notin C$  gilt.

Lassen Sie uns für jedes  $k \in \{1, \dots, d\}$  die Menge

$$A_k := \{s \in J : s \geq t_0 \text{ und } x_k(s) \leq 0\}$$

betrachten. Diese Menge ist abgeschlossen in  $J$ , und somit ist auch die endliche Vereinigung  $A := A_1 \cup \dots \cup A_d$  abgeschlossen in  $J$ . Wegen  $x(t) \notin C$  ist mindestens eine der Mengen  $A_k$  nicht-leer, und somit ist auch  $A$  nicht-leer. Weil  $A$  abgeschlossen ist, besitzt  $A$  ein kleinstes Element  $t_1$ .

Es ist  $t_1 > t_0$ , und für alle  $s \in [t_0, t_1]$  sind alle Komponenten von  $x(s)$  größer als 0. Also folgt wegen der Stetigkeit von  $x$ , dass  $x_k(t_1) \geq 0$  für alle  $k \in \{1, \dots, d\}$

gilt – d.h. es ist  $x(t_1) \in \mathbb{R}_+^d$  –, und wegen  $t_1 \in A$  gibt es ein  $k_0 \in \{1, \dots, d\}$  mit  $x_{k_0}(t_1) = 0$ .

Weil  $x_{k_0}$  links von  $t_1$  positiv ist, aber in  $t_1$  verschwindet, muss  $\dot{x}_{k_0}(t_1) \leq 0$  gelten. Andererseits impliziert aber die Voraussetzung an  $f$ , dass

$$\dot{x}_{k_0}(t_1) = f_{k_0}(t_1, x(t_1)) > 0$$

gilt. Dies ist ein Widerspruch. □

Nun können wir mit Hilfe eines Störungsargumentes die noch ausstehende Implikation von Satz 7.3.7 zeigen:

*Beweis von „(ii)  $\Rightarrow$  (i)“ in Satz 7.3.7.* Sei  $J \subseteq \mathbb{R}$  ein nicht-triviales Intervall, sei  $t_0 \in \mathbb{R}$  und sei  $x : J \rightarrow \mathbb{R}^d$  eine Lösung der Differentialgleichung  $\dot{x}(t) = f(t, x(t))$  mit  $x(t_0) \in \mathbb{R}_+^d$ .

Lassen Sie uns mit  $\mathbb{1} \in \mathbb{R}^d$  den Vektor bezeichnen, dessen Einträge alle gleich 1 sind. Für  $\varepsilon > 0$  betrachten wir das Anfangswertproblem

$$\begin{cases} \dot{\tilde{x}}(t) = f(t, \tilde{x}(t)) + \varepsilon \mathbb{1}, \\ \tilde{x}(t_0) = x(t_0) + \varepsilon \mathbb{1}. \end{cases}$$

Wegen Lemma 7.3.8 ist die Lösung dieses Anfangswertproblems für alle Zeiten  $t \geq t_0$ , für die sie existiert, im Inneren von  $\mathbb{R}_+^d$  – und somit insbesondere in  $\mathbb{R}_+^d$  – enthalten. Zugleich zeigt der Satz 7.2.4 über die stetige Abhängigkeit, dass die Lösung dieses Anfangswertproblems auf kompakten Teilintervallen von  $J$  gegen  $x$  strebt, wenn  $\varepsilon \rightarrow 0$  geht; also ist  $x(t)$  für alle  $t \in J$  mit  $t \geq t_0$  in  $\mathbb{R}_+^d$  enthalten. □

Lassen Sie uns nun einige Beispiele für die Anwendung von Satz 7.3.7 besprechen. Wir beginnen mit einem Epidemie-Modell, das bereits in einer Bonusaufgabe in den Übungen erwähnt wurde<sup>9</sup>.

**Beispiel 7.3.9 (Ein SIS-Epidemiemodell).** Für Koeffizienten  $c, d > 0$  betrachten wir die zwei-dimensionale autonome Differentialgleichung

$$\begin{pmatrix} \dot{S}(t) \\ \dot{I}(t) \end{pmatrix} = \begin{pmatrix} -cI(t)S(t) + dI(t) \\ cI(t)S(t) - dI(t) \end{pmatrix}$$

auf dem Zeitintervall  $\mathbb{R}$ . Dies ist ein sehr einfaches Modell, um die Ausbreitung einer Epidemie in einer Population zu beschreiben. Man geht in diesem Modell davon aus, dass alle gesunden Individuen sich mit der Krankheit anstecken können; ihre Anzahl zum Zeitpunkt  $t$  bezeichnet man mit  $S(t)$

<sup>9</sup>Allerdings war das Ziel in der Bonusaufgabe, eine explizite Lösung zu bestimmen; hier haben wir lediglich das Ziel, positive Invarianz des positiven Kegels  $\mathbb{R}_+^2$  zu zeigen.

(für engl. *susceptible*). Die Anzahl der infizierten Individuen zum Zeitpunkt  $t$  bezeichnet man mit  $I(t)$ . Die Konstante  $c > 0$  ist proportional zur Wahrscheinlichkeit, dass ein infiziertes Individuum ein nicht-infiziertes Individuum bei Begegnung ansteckt, und das Produkt  $I(t)S(t)$  ist proportional zur Anzahl der Begegnungen von infizierten und nicht infizierten Individuen. Die Konstante  $d > 0$  beschreibt die Genesungsrate.<sup>10</sup>

Als Zustandsraum wählen wir zunächst  $\Omega = \mathbb{R}^2$ . Allerdings ergeben Vektoren, in denen nicht beide Komponenten größer oder gleich 0 sind, aus biologischer Sicht keinen Sinn<sup>11</sup>. Damit wir wissen, dass die Differentialgleichung aus biologischer Sicht überhaupt Sinn ergeben kann, sollten wir also zeigen, dass der positive Kegel  $\mathbb{R}_+^2$  positiv invariant für diese Gleichung ist. Dazu verwenden wir Satz 7.3.7:

Die Abbildung

$$f : \mathbb{R}^{1+2} \rightarrow \mathbb{R}^2, \\ (t, s, i) \mapsto (-cis + di, cis - di)$$

beschreibt die rechte Seite unserer Differentialgleichung. Laut Satz 7.3.7 müssen wir nur beweisen, dass  $f$  quasi-positiv ist.

Sei also  $t \in \mathbb{R}$ , und sei  $(s, i) \in \mathbb{R}^2$ . Wir müssen zeigen: Wenn  $s = 0$  ist, ist  $f_1(t, s, i) \geq 0$ , und wenn  $i = 0$  ist, ist  $f_2(t, s, i) \geq 0$ . Wir gehen beide Fälle einzeln durch:

- Sei  $s = 0$ . Dann ist  $f_1(t, s, i) = -cis + di = di \geq 0$ , denn wegen  $(s, i) \in \mathbb{R}_+^2$  gilt  $i \geq 0$ .
- Sei  $i = 0$ . Dann ist  $f_2(t, s, i) = cis - di = 0$ .

Also ist  $f$  tatsächlich quasi-positiv, d.h. der positive Kegel  $\mathbb{R}_+^2$  ist tatsächlich positiv invariant für unsere Differentialgleichung. Damit haben wir gezeigt, dass ein Startwert  $(S(0), I(0))$ , der im biologisch sinnvollen Bereich  $\mathbb{R}_+^2$  liegt, zu einer Lösung führt, die für  $t \geq 0$  ebenfalls im biologisch sinnvollen Bereich  $\mathbb{R}_+^2$  liegt.<sup>12</sup>

<sup>10</sup>Bitte beachten Sie, dass dieses Modell so einfach ist, dass es für die Beschreibung der meisten Infektionskrankheiten ungeeignet ist. Das Modell hat unter anderem die folgenden Schwächen: Es ignoriert zum Beispiel eine mögliche Immunisierung von genesenen Individuen, es berücksichtigt keine Inkubationszeit, es ignoriert komplett die räumliche Struktur der Population sowie die Tatsache, dass verschiedene Individuen unterschiedliche Krankheitsverläufe haben können, es modelliert nicht die Tatsache, dass manche Krankheiten in manchen Fällen tödlich verlaufen können, und es enthält die eher fragwürdige Annahme, dass die Genesung eines Individuums von der Infektion zu jedem Zeitpunkt mit zeitunabhängiger Wahrscheinlichkeit erfolgt.

<sup>11</sup>Denn die Anzahl der gesunden Individuen und die Anzahl der infizierten Individuen in einer Population kann nicht negativ sein.

<sup>12</sup>Dies impliziert aber natürlich keineswegs, dass die Gleichung auch bzgl. anderer Aspekte biologisch sinnvolle Ergebnisse liefert.

Als zweites Beispiel betrachten wir lineare, autonome und homogene Differentialgleichungen:

**Beispiel 7.3.10.** Sei  $A \in \mathbb{R}^{d \times d}$ . Wir betrachten die linear, autonome und homogene Differentialgleichung

$$\dot{x}(t) = Ax(t)$$

auf dem Zustandsraum  $\mathbb{R}^d$  und dem Zeitintervall  $\mathbb{R}$ . Für diese Differentialgleichung können wir die positive Invarianz des positiven Kegels  $\mathbb{R}_+^d$  folgendermaßen beschreiben. Die folgenden Aussagen sind äquivalent:

- (i) Der positive Kegel  $\mathbb{R}_+^d$  ist positiv invariant für die Differentialgleichung  $\dot{x}(t) = Ax(t)$ .
- (ii) Für jede Zeit  $t \in [0, \infty)$  gilt: alle Einträge von  $e^{tA}$  sind  $\geq 0$ .
- (iii) Alle nicht-diagonal-Einträge von  $A$  sind  $\geq 0$ .

*Beweis.* Die Äquivalenz von (i) und (ii) kann man direkt aus der Definition von positiver Invarianz folgern; die Äquivalenz von (i) und (iii) kann man aus Satz 7.3.7 folgern. Die Details besprechen wir in den Übungen.  $\square$

Wir veranschaulichen das – sehr allgemein gehaltene – vorangehende Beispiel noch in zwei konkreten Fällen:

**Beispiele 7.3.11.** (a) Betrachten wir die Matrix

$$A = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \in \mathbb{R}^{2 \times 2}.$$

Die beiden nicht-diagonal-Einträge von  $A$  sind 1, also  $\geq 0$ . Somit folgt aus Beispiel 7.3.10: Alle Einträge von  $e^{tA}$  sind  $\geq 0$  für alle  $t \in [0, \infty)$ .

In diesem einfachen Beispiel können wir dies natürlich auch direkt nachrechnen: Laut Aufgaben 20 b) und c) auf Übungsblatt 5 gilt

$$e^{tA} = e^{-t} \begin{pmatrix} \cosh(t) & \sinh(t) \\ \sinh(t) & \cosh(t) \end{pmatrix}$$

für alle  $t \in \mathbb{R}$ . Für  $t \geq 0$  sind tatsächlich alle Einträge dieser Matrix positiv.<sup>13</sup> Für negative Zeiten hingegen besitzt  $e^{tA}$  in diesem Beispiel auch negative Einträge.<sup>14</sup>

<sup>13</sup>Der Punkt an Beispiel 7.3.10 ist aber natürlich, dass man diese Eigenschaft direkt an der Matrix  $A$  ablesen kann – was in komplizierteren Fällen nützlich ist, wenn man  $e^{tA}$  nicht so einfach berechnen kann.

<sup>14</sup>Was sich übrigens auch im Lichte von Beispiel 7.3.10 betrachten lässt: Für  $t < 0$  können wir  $e^{tA}$  einfach schreiben als  $e^{tA} = e^{(-t)(-A)}$  mit  $-t > 0$ ; die Matrix  $-A$  hat aber negative Diagonaleinträge, also muss es Zeiten  $-t > 0$  geben derart, dass manche Einträge von  $e^{(-t)(-A)}$  negativ sind.

(b) Nun betrachten wir noch die Matrix

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathbb{R}^{2 \times 2}.$$

Weil der obere nicht-Diagonal-Eintrag von  $A$  negativ ist, gibt es Zeiten  $t \geq 0$  derart, dass  $e^{tA}$  mindestens einen negativen Eintrag besitzt. Auch dieses Beispiel ist so einfach, dass wir dies konkret nachrechnen können: Für  $t \in \mathbb{R}$  gilt laut Aufgabe 20 d) auf Übungsblatt 5 die Formel

$$A = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}.$$

Wenn wir zum Beispiel  $t = \frac{\pi}{2}$  wählen, ist also der obere rechte Eintrag von  $A$  negativ. Es gibt aber auch Zeiten  $t > 0$ , für die alle Einträge  $\geq 0$  sind; dies ist zum Beispiel für  $t = 2\pi$  der Fall, denn es gilt  $e^{2\pi A} = \text{id}$ .<sup>15</sup>

- Literaturstellen und verwandte Ergebnisse.** (a) Die Behandlung der positiven Invarianz des positiven Kegels in diesem Abschnitt lehnt sich an [PW19, Abschnitt 4.2.2] an, auch wenn wir einige technische Details anders formuliert haben (zum Beispiel im Hinblick auf die genaue Definition des Begriffs *quasi-positiv*).
- (b) Wir haben in diesem Abschnitt nur Kriterien für die positive Invarianz des Kegel  $D = \mathbb{R}_+^d$  besprochen. Der Begriff *positive Invarianz* ist aber natürlich auch für andere Mengen definiert (vgl. Definition 7.3.1).  
Natürlich möchte man gerne auch Kriterien für die positive Invarianz von allgemeinen Mengen  $D$  zur Verfügung haben. Für abgeschlossene Mengen  $D$  lässt sich die positive Invarianz in der Tat charakterisieren; für solche Resultate verweisen wir auf [PW19, Kapitel 7].
- (c) Eine detaillierte Besprechung einiger grundlegender Epidemie-Modellen, von denen einige komplexer und etwas realistischer sind als Beispiel 7.3.9, finden Sie zum Beispiel in [PSZ08, Abschnitt II].

## 7.4 Differenzierbare Abhängigkeit vom Anfangswert

Ähnlich wie wir in Abschnitt 7.2 über die stetige Abhängigkeit der Lösung eines Anfangswertproblems von den Daten gesprochen hatten, können wir auch über differenzierbare Abhängigkeit sprechen. Dies wollen wir in diesem Abschnitt tun – allerdings beschränken wir uns hier darauf, autonome

<sup>15</sup>Dies zeigt, dass in Aussage (ii) in Beispiel 7.3.10 die Quantifizierung über *alle*  $t \in [0, \infty)$  wichtig ist, damit die Äquivalenz dort stimmt.

Gleichungen zu betrachten und für diese kurz die differenzierbare Abhängigkeit vom Anfangswert zu diskutieren. Die wesentliche Voraussetzung um die differenzierbare Abhängigkeit zu erhalten ist, dass die rechte Seite nicht nur Lipschitz-stetig, sondern sogar  $C^1$  ist:

**Satz 7.4.1.** Sei  $\Omega \subseteq \mathbb{R}^d$  nicht-leer und offen, und sei  $h : \Omega \rightarrow \mathbb{R}^d$  stetig differenzierbar. Es sei  $t_0 \in \mathbb{R}$  fest gewählt. Wir betrachten für jedes  $x_0 \in \Omega$  das Anfangswertproblem

$$\begin{cases} \dot{x}(t) = h(x(t)), \\ x(t_0) = x_0. \end{cases}$$

und bezeichnen seine Lösung hier<sup>16</sup> mit

$$I_{\max}(x_0) \ni t \mapsto x(t, x_0) \in \Omega.$$

Dann ist die Menge  $E := \bigcup_{x_0 \in \Omega} (I_{\max}(x_0) \times \{x_0\})$  offen in  $\mathbb{R}^{1+d}$ , und die Abbildung

$$E \ni (t, x_0) \mapsto x(t, x_0) \in \Omega$$

ist stetig differenzierbar. Ihre Ableitung nach  $x_0$ , die wir an jeder Stelle  $(t, x_0) \in E$  mit  $D_{x_0} x(t, x_0) \in \mathbb{R}^{d \times d}$  bezeichnen, ist zudem nach  $t$  differenzierbar, und erfüllt für jedes feste  $x_0 \in \mathbb{R}^d$  die Differentialgleichung

$$\frac{d}{dt} D_{x_0} x(t, x_0) = Dh(x(t, x_0)) D_{x_0} x(t, x_0),$$

sowie die Bedingung  $D_{x_0} x(t_0, x_0) = \text{id}$ .

*Beweis.* Wir verzichten an dieser Stelle auf den Beweis und verweisen stattdessen auf die Literatur, z.B. auf [PW19, Satz 4.3.1].

Lassen Sie uns aber an dieser Stelle noch anmerken, dass die Offenheit der Menge  $E$  leicht aus dem Satz 7.2.4 über die stetige Abhängigkeit von den Daten folgt, und dass die Aussage in der letzten Zeile des Satzes klar ist, weil  $x(t_0, x_0) = x_0$  für alle  $x_0 \in \Omega$  gilt.  $\square$

**Bemerkung 7.4.2.** Dass die Differentialgleichung für  $D_{x_0} x(t, x_0)$  in Satz 7.4.1 stimmt, sollte uns nicht besonders überraschen: Denn wenn wir für den Moment einmal annehmen, dass die Abbildung  $(t, x_0) \mapsto x(t, x_0)$  sogar  $C^2$  ist, dann folgt aus dem Satz von Schwarz  $\frac{d}{dt} D_{x_0} x(t, x_0) = D_{x_0} \dot{x}(t, x_0)$ , und somit folgt die behauptete Gleichung für  $D_{x_0} x(t, x_0)$ , wenn wir die Differentialgleichung

$$\dot{x}(t, x_0) = h(x(t, x_0))$$

<sup>16</sup>Geringfügig abweichend von der Notation, die wir sonst benutzen.

nach  $x_0$  ableiten.

Dies ist aber natürlich kein wirklicher Beweis von Satz 7.4.1, da wir ja a priori nichts über die Differenzierbarkeit der Abbildung  $(t, x_0) \mapsto x(t, x_0)$  wissen.

Anstatt uns lange mit der Theorie hinter Satz 7.4.1 aufzuhalten, wollen wir ihn stattdessen mit folgender Anwendung veranschaulichen. Wir verwenden im Folgenden noch einmal den Begriff des Lösungsflusses  $(\varphi_t)_{t \in [0, \infty)}$ , den wir in Definition 4.5.1 eingeführt hatten.

**Satz 7.4.3.** *Sei  $I \subseteq \mathbb{R}$  ein offenes Intervall, das den Zeitpunkt 0 enthält, und dessen rechte Grenze gleich  $\infty$  ist. Sei  $\Omega \subseteq \mathbb{R}^d$  nicht-leer und offen, und sei  $h : \Omega \rightarrow \mathbb{R}^d$  stetig differenzierbar. Wir nehmen an, dass für jedes  $x_0 \in \Omega$  die Lösung des Anfangswertproblems*

$$\begin{cases} \dot{x}(t) = h(x(t)), \\ x(0) = x_0 \end{cases} \quad (7.1)$$

global nach rechts existiert, d.h. dass  $t_+(0, x_0) = \infty$  gilt. Mit  $(\varphi_t)_{t \in [0, \infty)}$  bezeichnen wir den Lösungsfluss der Differentialgleichung  $\dot{x}(t) = h(x(t))$ , den wir in Definition 4.5.1 eingeführt haben.

Wenn das Vektorfeld  $h$  Divergenz-frei ist – d.h. also wenn  $\operatorname{div} h(y) = 0$  für alle  $y \in \Omega$  gilt<sup>17</sup> – dann ist für jedes  $t \in [0, \infty)$  die Abbildung  $\varphi_t : \Omega \rightarrow \Omega$  volumen-erhaltend, d.h. für jede offene Menge  $U \subseteq \Omega$  gilt

$$\operatorname{vol}(\varphi_t(U)) = \operatorname{vol}(U).$$

*Beweis.* Für jedes  $t \in [0, \infty)$  und jedes  $x_0 \in \Omega$  gilt, wenn wir die Notation aus Satz 7.4.1 verwenden,  $\varphi_t(x_0) = x(t_0, x_0)$ . Also erfüllt  $D\varphi_t(x_0) = D_{x_0} x(t, x_0) \in \mathbb{R}^{d \times d}$  für jedes feste  $x_0 \in \Omega$  laut Satz 7.4.1 die Differentialgleichung

$$\frac{d}{dt} D\varphi_t(x_0) = A(t) D\varphi_t(x_0),$$

wobei wir  $A(t) := Dh(x(t, x_0)) = Dh(\varphi_t(x_0))$  gesetzt haben; zudem ist  $D\varphi_0(x_0) = \operatorname{id}$ . Anders gesprochen bedeutet dies: Die Abbildung  $t \mapsto D\varphi_t(x_0)$  ist die zum Startzeitpunkt 0 gehörende Hauptfundamentalmatrix der Differentialgleichung

$$\dot{\tilde{x}}(t) = A(t)\tilde{x}(t).$$

Insbesondere ist die Matrix  $D\varphi_t(x_0)$  für jedes  $t$  und jedes  $x_0$  invertierbar, und somit ist  $\varphi_t : \Omega \rightarrow \varphi_t(\Omega) \subseteq \Omega$  ein Diffeomorphismus. Wir können somit

<sup>17</sup>Wir erinnern an dieser Stelle daran, dass  $\operatorname{div} h(y)$  definiert ist als  $\operatorname{div} h(y) := \operatorname{spur}(Dh(y)) = \sum_{k=1}^d \partial_k h_k(y)$ .

die Transformationsformel für mehrdimensionale Integrale benutzen und erhalten hieraus für jedes  $t \in [0, \infty)$

$$\text{vol}(\varphi_t(U)) = \int_{\varphi_t(U)} 1 \, dy = \int_U |\det D\varphi_t(x_0)| \, dx_0. \quad (7.2)$$

Wir erinnern an dieser Stelle daran, dass  $\psi(t) := \det D\varphi_t(x_0)$  die sogenannte *Wronski-Determinante* der Lösungsmatrix  $D\varphi_t(x_0)$  von  $\dot{\tilde{x}}(t) = A(t)\tilde{x}(t)$  ist, und somit die Differentialgleichung

$$\dot{\psi}(t) = \text{spur}(A(t))\psi(t)$$

erfüllt – dies hatten wir Bemerkung 5.3.11 besprochen. Weil wir ja  $A(t) = Dh(\varphi_t(x_0))$  gesetzt hatten, ist  $\text{spur}(A(t)) = \text{div}h(\varphi_t(x_0))$ , und da wir  $h$  als Divergenz-frei vorausgesetzt haben, folgt  $\text{spur}(A(t)) = 0$  für alle  $t$ . Also hängt  $\psi$  nicht von der Zeit ab, d.h. es gilt  $\psi(t) = \psi(0) = \det(\text{id}) = 1$  für alle  $t \geq 0$ . Somit erhalten wir aus Gleichung (7.2), dass

$$\text{vol}(\varphi_t(U)) = \int_U 1 \, dx_0 = \text{vol}(U)$$

für alle  $t \in [0, \infty)$  gilt. □

Satz 7.4.3 lässt sich insbesondere auf die Newtonschen Bewegungsgleichungen in der klassischen Mechanik anwenden, und liefert dann einen bekannten Satz von Liouville<sup>18</sup> über die Volumenerhaltung im Phasenraum als Spezialfall. Falls genügend Platz bleibt, werden wir dies auf Übungsblatt 12 besprechen.

**Literaturstellen und verwandte Ergebnisse.** Neben der differenzierbaren Abhängigkeit der Lösung einer Differentialgleichung vom Anfangswert, ist es auch sinnvoll, differenzierbare Abhängigkeit von Parametern zu studieren – denn häufig tauchen in konkreten Modellen ja verschiedene Parameter in Differentialgleichungen auf.

Resultate zur differenzierbaren Abhängigkeit von Parametern finden Sie z.B. in [PW19, Abschnitt 4.3.2] und in [Heu04, Satz 13.2].

## 7.5 Ergänzungen

### Markov-Prozesse mit endlichem Zustandsraum

In Beispiel 7.3.10 hatten wir Matrizen  $A \in \mathbb{R}^{d \times d}$  betrachtet derart, dass der positive Kegel  $\mathbb{R}_+^d$  positiv invariant für die Differentialgleichung

$$\dot{x}(t) = Ax(t)$$

<sup>18</sup>Benannt nach Joseph Liouville (1809 – 1882), französischer Mathematiker.

ist. Wenn man die Entwicklung von zufälligen Größen in kontinuierlicher Zeit – also sogenannte *stochastische Prozesse* – studiert, taucht häufig ein ähnliches Phänomen auf:

Lassen Sie uns ein System betrachten, das  $d$  verschiedene Zustände annehmen kann; wir nummerieren diese Zustände mit den Zahlen  $1, \dots, d$ . In der Wahrscheinlichkeitstheorie betrachtet man Systeme, deren Zustand man nicht präzise kennt; stattdessen kennt man zu jedem Zeitpunkt  $t$  lediglich die Wahrscheinlichkeit, dass sich das System in Zustand 1, Zustand 2, usw. befindet. Wenn wir mit  $p_k(t)$  (für  $k \in \{1, \dots, d\}$ ) die Wahrscheinlichkeit bezeichnen, dass sich das System zum Zeitpunkt  $t$  im Zustand  $k$  befindet, dann ist  $p(t) := (p_1(t), \dots, p_d(t))^T \in \mathbb{R}^d$  also ein Vektor im positiven Kegel  $\mathbb{R}_+^d$  mit der zusätzlichen Eigenschaft

$$1 = \sum_{k=1}^d p_k(t) = \langle \mathbb{1}, p(t) \rangle,$$

wobei  $\mathbb{1} \in \mathbb{R}^d$  den Vektor bezeichnet, dessen Einträge alle gleich 1 sind.

Eine wichtige Klasse von stochastischen Prozessen sind die sogenannten *Markov-Prozesse*. Für diese Prozesse erfüllen die Wahrscheinlichkeits-Vektoren  $p(t)$  eine lineare Differentialgleichung der Form

$$\dot{p}(t) = Qp(t) \tag{7.3}$$

für eine Matrix  $Q \in \mathbb{R}^{d \times d}$ .

Wenn der Startvektor  $p(0)$  eine Wahrscheinlichkeits-Verteilung beschreibt – also in  $\mathbb{R}_+^d$  liegt und  $\langle \mathbb{1}, p(0) \rangle = 1$  erfüllt – dann soll natürlich auch  $p(t)$  für alle  $t \geq 0$  diese Eigenschaft haben, denn ansonsten wäre die Differentialgleichung ja nicht geeignet um die zeitliche Entwicklung von Wahrscheinlichkeiten zu beschreiben.

Also fragt man sich: Welche Eigenschaften muss  $Q$  haben, damit (i)  $\mathbb{R}_+^d$  positiv invariant für die Differentialgleichung (7.3) ist, und damit (ii) die Abbildung  $\mathbb{R}^d \ni y \mapsto \langle \mathbb{1}, y \rangle \in \mathbb{R}$  eine Erhaltungsgröße für (7.3) ist. Diese Fragen kann man leicht mit Hilfe der Theorie beantworten, die wir bereits entwickelt haben:

**Proposition 7.5.1.** *Für jede Matrix  $Q \in \mathbb{R}^{d \times d}$  sind die folgenden Aussagen äquivalent:*

- (i) *Der positive Kegel  $\mathbb{R}_+^d$  ist positiv invariant für die Differentialgleichung (7.3), und die Abbildung  $\mathbb{R}^d \ni y \mapsto \langle \mathbb{1}, y \rangle \in \mathbb{R}$  ist eine Erhaltungsgröße für diese Differentialgleichung.*
- (ii) *Für jedes  $t \in [0, \infty)$  sind alle Einträge der Matrix  $e^{tQ}$  größer oder gleich 0, und die Einträge jeder Spalte von  $e^{tQ}$  summieren sich zu 1.*

(iii) *Alle nicht-diagonal-Einträge von  $Q$  sind  $\geq 0$ , und die Einträge jeder Spalte von  $Q$  summieren sich zu 0.*

*Beweis.* Für den Beweis kann man zum Beispiel die Charakterisierung von Erhaltungsgrößen in Proposition 6.3.3 und die Charakterisierung der positiven Invarianz von  $\mathbb{R}_+^d$  in Beispiel 7.3.10 verwenden.  $\square$

Matrizen  $Q$  mit diesen beiden Eigenschaften beschreiben also Markov-Prozesse mit endlichem Zustandsraum.<sup>19</sup>

---

<sup>19</sup>Bitte beachten Sie, dass in der Stochastik häufig die Wahrscheinlichkeitsverteilungen  $p(t)$  als Zeilenvektoren geschrieben werden, womit man die Differentialgleichung  $\dot{p}(t) = p(t)Q$  anstelle von (7.3) erhält. Dies führt dazu, dass man in der Stochastik fordert, dass die Zeilensummen von  $Q$  gleich 0 sind und dass die Zeilensummen von  $e^{tQ}$  gleich 1 sind (jeweils anstelle der Spaltensummen).

# Langzeitverhalten autonomer Systeme

## Fragen zum Einstieg.

- (a) Stellen Sie sich vor, Sie lassen ein Pendel schwingen. Warum bleibt das Pendel irgendwann stehen?
- (b) Nehmen Sie einen Bleistift und versuchen Sie ihn mit der Spitze nach oben auf Ihrem Schreibtisch aufzustellen. Wenn Ihnen das gelungen ist: Was passiert, wenn sie den Stift leicht stoßen? Wieso passiert das?
- (c) Berechnen Sie die Lösung  $x: \mathbb{R} \rightarrow \mathbb{R}^2$  des Anfangswertproblems

$$\begin{cases} \dot{x}(t) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} x(t) & \text{für } t \in \mathbb{R}, \\ x(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \end{cases}$$

Was passiert mit  $x(t)$  für  $t \rightarrow \infty$ ?

## 8.1 Langzeitverhalten im linearen Fall

In Kapitel 8 beschäftigen wir uns mit der Frage, wie sich die Lösungen einer Differentialgleichung verhalten, wenn die Zeit  $t$  gegen  $\infty$  strebt; wir schränken uns dabei durchgehend auf autonome Differentialgleichungen ein. Im aktuellen Abschnitt 8.1 behandeln wir den wichtigen Spezialfall linearer und homogener Differentialgleichungen, also Gleichungen der Form

$$\dot{x}(t) = Ax(t)$$

für eine Matrix  $A \in \mathbb{R}^{d \times d}$ . Wir wissen aus der linearen Theorie, dass jede Lösung dieser Gleichung global existiert und die Formel  $x(t) = e^{tA}x(0)$  für alle  $t \in \mathbb{R}$  erfüllt. Somit müssen wir, um das Langzeitverhalten der Lösungen von  $\dot{x}(t) = Ax(t)$  zu studieren, nur untersuchen, wie sich die Abbildung

$$\mathbb{R} \ni t \mapsto e^{tA}x(0)$$

für  $t \rightarrow \infty$  verhält. Wir studieren dazu im Folgenden zwei Arten von Verhalten: Konvergenz der Lösungen gegen 0, und Beschränktheit der Lösungen. Zunächst beweisen wir zwei Sätze, die Konvergenz gegen 0 auf unterschiedliche Weise charakterisieren; anschließend behandeln wir Beschränktheit von Lösungen.

**Satz 8.1.1 (Konvergenz gegen Null, Teil 1).** Sei  $A \in \mathbb{R}^{d \times d}$ . Dann sind die folgenden Aussagen äquivalent:

- (i) Für jedes  $x_0 \in \mathbb{R}^d$  gilt  $e^{tA}x_0 \rightarrow 0$  für  $t \rightarrow \infty$ ; oder anders ausgedrückt: jede Lösung der Differentialgleichung

$$\dot{x}(t) = Ax(t)$$

strebt für  $t \rightarrow \infty$  gegen 0.

- (ii) Es gilt  $\|e^{tA}\| \rightarrow 0$  für  $t \rightarrow \infty$ .

- (iii) Es gibt eine Zahl  $\delta > 0$  und eine Zahl  $M \geq 1$  mit der Eigenschaft  $\|e^{tA}\| \leq Me^{-\delta t}$  für alle  $t \in [0, \infty)$ .

- (iv) Die sogenannte Spektralschranke

$$s(A) := \max\{\operatorname{Re} \lambda : \lambda \text{ ist ein Eigenwert von } A\}$$

von  $A$  erfüllt  $s(A) < 0$ ; oder anders ausgedrückt: alle Eigenwerte von  $A$  haben Realteil  $< 0$ .

*Beweis.* Wir zeigen zunächst „(i)  $\Rightarrow$  (iv)  $\Rightarrow$  (ii)  $\Rightarrow$  (i)“, und anschließend „(ii)  $\Leftrightarrow$  (iii)“

„(i)  $\Rightarrow$  (iv)“ Angenommen (iv) gilt nicht. Dann gibt es einen Eigenwert  $\lambda \in \mathbb{C}$  von  $A$  mit  $\operatorname{Re} \lambda \geq 0$ . Sei  $z \in \mathbb{C}^d$  ein zugehöriger Eigenvektor; diesen können wir schreiben als  $z = x + iy$  mit  $x, y \in \mathbb{R}^d$ . Für jede Zahl  $t \in [0, \infty)$  gilt laut Aufgabe 30(a) auf Übungsblatt 7

$$e^{tA}z = e^{t\lambda}z,$$

und somit

$$\|e^{tA}z\| = |e^{t\lambda}| \|z\| = e^{t\operatorname{Re} \lambda} \|z\| \not\rightarrow 0 \quad \text{für } t \rightarrow \infty.$$

Für jedes  $t \in [0, \infty)$  gilt aber natürlich auch

$$\|e^{tA}z\| \leq \|e^{tA}x\| + \|e^{tA}y\|.$$

Also kann nicht sowohl  $\|e^{tA}x\|$  als auch  $\|e^{tA}y\|$  für  $t \rightarrow \infty$  gegen 0 gehen. Dies zeigt, dass (i) nicht erfüllt ist.

„(iv)  $\Rightarrow$  (ii)“ Wir schreiben  $A$  als  $A = TJT^{-1}$ , wobei  $T \in \mathbb{C}^{d \times d}$  invertierbar ist und  $J \in \mathbb{C}^{d \times d}$  die Jordan-Normalform von  $A$  ist. Für alle  $t \in [0, \infty)$  gilt

$$\|e^{tA}\| = \|Te^{tJ}T^{-1}\| \leq \|T\| \|e^{tJ}\| \|T^{-1}\|;$$

also genügt es zu zeigen, dass  $\|e^{tJ}\| \rightarrow 0$  für  $t \rightarrow \infty$  gilt, und dies wiederum ist äquivalent dazu, dass alle Einträge von  $e^{tJ}$  für  $t \rightarrow \infty$  gegen Null streben. Weil aber alle Eigenwerte von  $A$  Realteil  $< 0$  haben, folgt aus Satz 5.5.9, dass tatsächlich alle Einträge von  $e^{tJ}$  für  $t \rightarrow \infty$  gegen 0 gehen.

„(ii)  $\Rightarrow$  (i)“ Wenn (ii) gilt, dann folgt für jedes  $x_0 \in \mathbb{R}^d$

$$\|e^{tA}x_0\| \leq \|e^{tA}\| \|x_0\| \rightarrow 0$$

für  $t \rightarrow \infty$ .

„(ii)  $\Rightarrow$  (iii)“ Wenn (ii) stimmt, dann gilt, wie wir bereits gezeigt haben,  $s(A) < 0$ . Wir wählen nun eine Zahl  $\delta > 0$  derart, dass auch noch  $s(A) + \delta < 0$  gilt. Somit ist

$$s(A + \delta \text{id}) = s(A) + \delta < 0,$$

also folgt aus der Implikation von (iv) nach (ii), die wir bereits gezeigt haben, dass  $\|e^{t(A+\delta \text{id})}\| \rightarrow 0$  für  $t \rightarrow \infty$  gilt. Insbesondere ist  $\|e^{t(A+\delta \text{id})}\|$  für  $t \in [0, \infty)$  beschränkt, d.h. es gibt eine Zahl  $M \geq 1$  derart, dass

$$\|e^{t(A+\delta \text{id})}\| \leq M$$

für alle  $t \in [0, \infty)$  ist<sup>1</sup>. Weil

$$\|e^{t(A+\delta \text{id})}\| = e^{t\delta} \|e^{tA}\|$$

für alle  $t \in [0, \infty)$  gilt, folgt (iii).

„(iii)  $\Rightarrow$  (ii)“ Diese Implikation ist offensichtlich.  $\square$

**Bemerkungen 8.1.2.** (a) Satz 8.1.1 bleibt genauso richtig (mit demselben Beweis), wenn wir auch komplexe Matrizen (d.h.  $A \in \mathbb{C}^{d \times d}$ ) zulassen. Wir haben ihn oben nur für reelle Matrizen formuliert, weil in Aussage (i) eine Differentialgleichung vorkommt, und wir in dieser Vorlesung nur Differentialgleichungen betrachtet haben, deren Lösungen nach  $\mathbb{R}^d$  abbilden.<sup>2</sup>

(b) Der Beweis der Implikation „(ii)  $\Rightarrow$  (iii)“ zeigt sogar, wie man  $\delta$  in Aussage (iii) wählen kann: Man kann ein beliebiges  $\delta \in (0, -s(A))$  verwenden. Beachten Sie aber bitte, dass  $M$  von  $\delta$  abhängt – und wenn  $\delta$  nahe an  $-s(A)$  liegt, kann es passieren, dass  $M$  recht groß wird.

<sup>1</sup>Dass man  $M$  nicht kleiner als 1 wählen kann, sieht man, wenn man  $t = 0$  betrachtet.

<sup>2</sup>Man kann aber ebenso gut auch Differentialgleichungen untersuchen, deren Lösungen nach  $\mathbb{C}^d$  abbilden; dadurch ändert sich nicht viel an der Theorie.

Unsere zweite Charakterisierung von Konvergenz gegen 0 verwendet den Begriff der positiv semi-definiten Matrix. Wir erinnern daran, dass eine Matrix  $B \in \mathbb{C}^{d \times d}$  *selbst-adjungiert* (oder *hermitesch*<sup>3</sup>) heißt, wenn sie gleich ihrer konjugiert komplexen transponierten Matrix  $B^* := \overline{B}^T$  ist; für reelle Matrizen verwendet man anstelle des Begriffs selbst-adjungiert oft auch den Begriff *symmetrisch*. Eine selbst-adjungierte Matrix  $B \in \mathbb{C}^{d \times d}$  heißt *positiv semi-definit*, wenn  $\langle z, Bz \rangle \geq 0$  für alle  $z \in \mathbb{C}^d$  gilt; dies ist äquivalent dazu, dass alle Eigenwerte von  $B$  größer oder gleich 0 sind.<sup>4</sup>

**Satz 8.1.3 (Konvergenz gegen Null, Teil 2).** Sei  $A \in \mathbb{R}^{d \times d}$ . Dann sind die folgenden Aussagen äquivalent:

- (i) Für jedes  $x_0 \in \mathbb{R}^d$  gilt  $e^{tA}x_0 \rightarrow 0$  für  $t \rightarrow \infty$ ; oder anders ausgedrückt: jede Lösung der Differentialgleichung

$$\dot{x}(t) = Ax(t)$$

strebt für  $t \rightarrow \infty$  gegen 0.<sup>5</sup>

- (ii) Für jede selbst-adjungierte und positiv semi-definite Matrix  $Q \in \mathbb{C}^{d \times d}$  gibt es eine selbst-adjungierte und positiv semi-definite Matrix  $P \in \mathbb{C}^{d \times d}$  mit der Eigenschaft

$$A^T P + P A = -Q.$$

*Beweis.* „(i)  $\Rightarrow$  (ii)“ Gelte (i) und sei  $Q \in \mathbb{R}^{d \times d}$  selbst-adjungiert und positiv semi-definit. Laut Satz 8.1.1(iii) gibt es Zahlen  $M \geq 1$  und  $\delta > 0$  derart, dass  $\|e^{tA}\| \leq M e^{-\delta t}$  für alle  $t \in [0, \infty)$  gilt.

Für jedes  $n \in \mathbb{N}$  betrachten wir nun die Matrix

$$P_n := \int_0^n e^{tA^T} Q e^{tA} dt \in \mathbb{C}^{d \times d}.$$

Die Folge der  $P_n$  ist eine Cauchy-Folge in  $\mathbb{C}^{d \times d}$ , denn für jede alle  $m, n \in \mathbb{N}$  mit  $m \leq n$  gilt

$$\begin{aligned} \|P_n - P_m\| &= \left\| \int_m^n e^{tA^T} Q e^{tA} dt \right\| \leq \int_m^n \|e^{tA^T}\| \|Q\| \|e^{tA}\| dt \\ &\leq \int_m^n e^{-2\delta t} dt M^2 \|Q\| = \left( e^{-2m\delta} - e^{-2n\delta} \right) \frac{M^2 \|Q\|}{2\delta} \leq e^{-2m\delta} \frac{M^2 \|Q\|}{2\delta}; \end{aligned}$$

<sup>3</sup>Benannt nach Charles Hermite (1822 – 1901), französischer Mathematiker.

<sup>4</sup>Wir erinnern an dieser Stelle daran, dass alle Eigenwerte einer selbst-adjungierten Matrix immer reell sind.

<sup>5</sup>Dies ist einfach nochmals die Aussage (i) aus Satz 8.1.1.

hierbei haben wir für die Ungleichung zwischen den beiden Zeilen verwendet, dass  $e^{tA^T} = (e^{tA})^T$  gilt, und dass jede Matrix dieselbe induzierte 2-Norm hat wie ihre Transponierte.

Also ist  $(P_n)$  eine Cauchy-Folge und konvergiert somit gegen eine Matrix  $P \in \mathbb{C}^{d \times d}$ . Man sieht leicht, dass jede Matrix  $P_n$  selbst-adjungiert ist, also ist auch  $P$  selbst-adjungiert. Außerdem gilt für jedes  $z \in \mathbb{C}^d$

$$\langle z, Pz \rangle = \lim_{n \rightarrow \infty} \int_0^n \langle z, (e^{tA})^T Q e^{tA} z \rangle dt = \lim_{n \rightarrow \infty} \int_0^n \underbrace{\langle e^{tA} z, Q e^{tA} z \rangle}_{\geq 0} dt \geq 0;$$

also ist  $P$  positiv semi-definit.

Nun zeigen wir noch, dass  $P$  die gewünschte Gleichung erfüllt: Für jedes  $n \in \mathbb{N}$  gilt

$$A^T P_n + P_n A = \int_0^n \underbrace{A^T e^{tA^T} Q e^{tA} + e^{tA^T} Q e^{tA} A}_{= \frac{d}{dt} (e^{tA^T} Q e^{tA})} dt = e^{nA^T} Q e^{nA} - Q \rightarrow -Q$$

für  $n \rightarrow \infty$ . Zugleich gilt aber natürlich  $A^T P_n + P_n A \rightarrow A^T P + P A$  für  $n \rightarrow \infty$ , womit tatsächlich

$$A^T P + P A = -Q$$

gezeigt ist.

„(ii)  $\Rightarrow$  (i)“ Wir nehmen an, dass (i) falsch ist. Aufgrund von Satz 8.1.1 gibt es dann einen Eigenwert  $\lambda \in \mathbb{C}$  von  $A$  mit Realteil  $\operatorname{Re} \lambda \geq 0$ . Sei  $z \in \mathbb{C}^d \setminus \{0\}$  ein zugehöriger Eigenvektor. Wir betrachten die Matrix  $Q := z \cdot \bar{z}^T \in \mathbb{C}^{d \times d}$ ; sie ist selbst-adjungiert und positiv semi-definit. Wenn es nun eine selbst-adjungierte und positiv semi-definite Matrix  $P \in \mathbb{C}^{d \times d}$  mit  $A^T P + P A = -Q$  gäbe, dann würde

$$\langle z, (A^T P + P A) z \rangle = -\langle z, Q z \rangle$$

folgen. Die rechte Seite erfüllt<sup>6</sup>  $\langle z, Q z \rangle = -\|z\|^4 < 0$ , während die linke Seite gleich

$$\langle A z, P z \rangle + \langle z, P A z \rangle = \bar{\lambda} \langle z, P z \rangle + \lambda \langle z, P z \rangle = 2 \operatorname{Re} \lambda \langle z, P z \rangle$$

und somit  $\geq 0$  ist. Widerspruch. Also gibt es kein  $P$ , welches die gewünschte Gleichung erfüllt, d.h. die Aussage (ii) ist falsch.  $\square$

<sup>6</sup>Wir verwenden hier die in der Physik übliche Konvention, dass das komplexe Skalarprodukt linear in der zweiten Komponente und anti-linear in der ersten ist, d.h. dass  $\langle a, b \rangle = \bar{a}^T b$  für alle  $a, b \in \mathbb{C}^d$  gilt; in der Literatur über lineare Algebra und Funktionalanalysis wird aber auch häufig die umgekehrte Konvention (d.h.  $\langle a, b \rangle = a^T \bar{b}$ ) verwendet – in diesem Fall müsste man  $Q = \bar{z} \cdot z^T$  definieren, damit das Argument funktioniert.

**Bemerkungen 8.1.4.** (a) Auch Satz 8.1.3 bleibt für komplexe Matrizen  $A$  richtig. In diesem Fall muss man in Aussage (ii) anstelle der Matrix  $A^T$  die Matrix  $A^* := \overline{A}^T$  benutzen.

(b) Anstelle von positiv semi-definiten Matrizen  $P$  und  $Q$  in Satz 8.1.3 kann man auch positiv definite Matrizen betrachten.<sup>7</sup> Wenn man positiv-definite Matrizen verwendet, erhält man sogar noch eine Äquivalenz zu einer a priori schwächeren Aussage, in der man nur die Existenz einer Matrix  $Q$  und einer Matrix  $P$  fordert.

Genauer sind für jedes  $A \in \mathbb{R}^{d \times d}$  äquivalent:

(i) Für jedes  $x_0 \in \mathbb{R}^d$  gilt  $e^{tA}x_0 \rightarrow 0$  für  $t \rightarrow \infty$ ; oder anders ausgedrückt: jede Lösung der Differentialgleichung

$$\dot{x}(t) = Ax(t)$$

strebt für  $t \rightarrow \infty$  gegen 0.

(ii) Für jede selbst-adjungierte und positiv definite Matrix  $Q \in \mathbb{R}^{d \times d}$  gibt es eine selbst-adjungierte und positiv definite Matrix  $P \in \mathbb{R}^{d \times d}$  mit der Eigenschaft

$$A^T P + PA = -Q$$

(iii) Es gibt eine selbst-adjungierte und positiv definite Matrix  $Q \in \mathbb{R}^{d \times d}$  und selbst-adjungierte und positiv definite Matrix  $P \in \mathbb{R}^{d \times d}$  mit der Eigenschaft

$$A^T P + PA = -Q.$$

Die Implikation von (i) nach (ii) zeigt man fast genau wie im Beweis von Satz 8.1.3, und die Implikation von (ii) nach (iii) ist klar. Wir verzichten an dieser Stelle auf einen Beweis der Implikation (iii) nach (i).

Nun betrachten wir noch die Frage, unter welchen Bedingungen alle Lösungen von  $\dot{x}(t) = Ax(t)$  für  $t \rightarrow \infty$  beschränkt sind. Diese Frage beantwortet der folgende Satz.

**Satz 8.1.5 (Beschränktheit der Lösungen).** Sei  $A \in \mathbb{R}^{d \times d}$ . Dann sind die folgenden Aussagen äquivalent:

<sup>7</sup>Wir erinnern daran, dass eine selbst-adjungierte Matrix  $B \in \mathbb{C}^{d \times d}$  positiv definit heißt, falls  $\langle z, Bz \rangle > 0$  für alle  $z \in \mathbb{C}^{d \times d}$  gilt; dies ist äquivalent dazu, dass jeder Eigenwert von  $B$  echt größer als 0 ist.

- (i) Für jedes  $x_0 \in \mathbb{R}^d$  gilt  $\sup_{t \in [0, \infty)} \|e^{tA} x_0\| < \infty$ ; oder anders ausgedrückt: jede Lösung der Differentialgleichung

$$\dot{x}(t) = Ax(t)$$

ist auf  $[0, \infty)$  beschränkt.

- (ii) Es gilt  $\sup_{t \in [0, \infty)} \|e^{tA}\| < \infty$ .
- (iii) Die Spektralschranke  $s(A)$  von  $A$  erfüllt  $s(A) \leq 0$  (anders ausgedrückt: alle Eigenwerte von  $A$  haben Realteil höchstens 0), und alle Eigenwerte von  $A$ , die in  $i\mathbb{R}$  liegen, sind semi-simpel.

*Beweis.* Den Beweis behandeln wir in den Übungen auf Blatt 12. □

Wir schließen Abschnitt 8.1 mit einigen einfachen Beispielen:

**Beispiele 8.1.6.** (a) Sei

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathbb{R}^{2 \times 2}.$$

Die Matrix  $A$  besitzt die Eigenwerte  $i$  und  $-i$  und beide haben jeweils die geometrische und die algebraische Vielfachheit 1, sind also semi-simpel. Somit folgt aus Satz 8.1.5, dass  $\|e^{tA}\|$  für  $t \in [0, \infty)$  beschränkt ist, und aus Satz 8.1.1, dass  $e^{tA}$  für  $t \rightarrow \infty$  nicht gegen 0 strebt (wegen  $s(A) = 0$ ). In diesem einfachen Beispiel können wir dies natürlich auch direkt überprüfen, denn wir wissen ja bereits, dass

$$e^{tA} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$$

für alle Zeiten  $t$  gilt.<sup>8</sup>

(b) Sei

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{2 \times 2}.$$

Dann hat  $A$  nur den Eigenwert 0; seine geometrische Vielfachheit ist 1 und seine algebraische Vielfachheit ist 2, also ist dieser Eigenwert nicht semi-simpel.

Somit hat  $A$  zwar Spektralschranke 0, aber auf der imaginären Achse liegt ein nicht-semi-simpler Eigenwert. Laut Satz 8.1.5 gilt also in diesem Beispiel  $\sup_{t \in [0, \infty)} \|e^{tA}\| = \infty$ .

<sup>8</sup>Vgl. zu diesem Beispiel übrigens auch die Einstiegsfrage (c) des Kapitels 8.

Dieses Beispiel ist ebenfalls so einfach, dass wir diese Aussage auch direkt nachrechnen können: Für alle Zeiten  $t$  gilt

$$e^{tA} = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix},$$

also sehen wir hier ebenfalls, dass  $e^{tA}$  für  $t \rightarrow \infty$  nicht beschränkt ist.

(c) Sei nun

$$A = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix} \in \mathbb{R}^{2 \times 2};$$

dies ist fast dasselbe Beispiel wie in (b), lediglich um  $-id$  verschoben. Also ist nun  $-1$  der einzige Eigenwert von  $A$ , d.h.  $A$  hat die Spektralschranke  $-1$ . Somit folgt aus Satz 8.1.1, dass es ein  $\delta > 0$  und ein  $M \geq 1$  gibt mit der Eigenschaft

$$\|e^{tA}\| \leq M e^{-\delta t}$$

für alle  $t \in [0, \infty)$ . Auch in diesem einfachen Beispiel können wir dies direkt nachrechnen: Für alle Zeiten  $t$  gilt

$$e^{tA} = e^{-t} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix},$$

also gilt für jedes  $\delta \in (0, 1)$ , dass  $\|e^{\delta t} e^{tA}\|$  für  $t \in [0, \infty)$  beschränkt ist durch eine Zahl  $M \geq 1$  (welche von der Wahl von  $\delta$  abhängt). Folglich ist  $\|e^{tA}\| \leq e^{-\delta t} M$  für alle  $t \in [0, \infty)$ . Dass wir  $\delta \in (0, 1)$  wählen können, ist auch konsistent mit Bemerkung 8.1.2(b).

## 8.2 Stabilität und Attraktivität

Wir wollen im Rest von Kapitel 8 auch das Langzeitverhalten der Lösungen von nichtlinearen Differentialgleichungen studieren. Dazu benötigen wir ein paar Begriffe, die wir im aktuellen Abschnitt 8.2 einführen. Wie zu Beginn von Kapitel 8 angekündigt betrachten wir im ganzen Kapitel nur autonome Systeme.

**Definition 8.2.1 (Equilibrium).** Sei  $\Omega \subseteq \mathbb{R}^d$  offen und nichtleer und sei  $h : \Omega \rightarrow \mathbb{R}^d$  lokal Lipschitz stetig. Ein Punkt  $x^* \in \Omega$  heißt *Equilibrium* (oder auch *Gleichgewichtspunkt* oder *Ruhelage*) der Differentialgleichung

$$\dot{x}(t) = h(x(t)),$$

wenn die konstante Funktion  $\mathbb{R} \ni t \mapsto x^* \in \mathbb{R}^d$  eine Lösung der Differentialgleichung ist.

Anders formuliert: Der Punkt  $x^*$  ist ein Equilibrium genau dann, wenn  $h(x^*) = 0$  gilt.

**Aufgabe 8.2.2.** Zeigen Sie, dass die beiden Aussagen, deren Äquivalenz in Definition 8.2.1 behauptet wird, tatsächlich äquivalent sind.

Wir erwähnen ein paar einfache Beispiele:

**Beispiele 8.2.3.** (a) Sei  $A \in \mathbb{R}^{d \times d}$ . Dann ist der Nullpunkt ein Equilibrium der linearen Differentialgleichung

$$\dot{x}(t) = Ax(t).$$

Falls die Matrix  $A$  vollen Rang hat, ist der Nullpunkt das einzige Equilibrium.

(b) Seien  $c, d > 0$ . Wir betrachten wieder einmal die logistische Gleichung

$$\dot{b}(t) = (c - db(t))b(t),$$

die wir in Beispiel 2.1.3 eingeführt hatten. Offensichtlich besitzt die Gleichung genau zwei Equilibria, nämlich die beiden Zahlen 0 und  $\frac{c}{d}$ .

(c) Seien  $c, d > 0$ . Wir betrachten noch einmal das SIS-Epidemie-Modell aus Beispiel 7.3.9, d.h. die Differentialgleichung

$$\begin{pmatrix} \dot{S}(t) \\ \dot{I}(t) \end{pmatrix} = \begin{pmatrix} -cI(t)S(t) + dI(t) \\ cI(t)S(t) - dI(t) \end{pmatrix}.$$

Ein Punkt  $(s, i) \in \mathbb{R}^2$  ist genau dann ein Equilibrium dieser Gleichung, wenn  $-cis + di = 0$  gilt. Somit sind alle Punkte der Form  $(s, 0)$  mit  $s \in \mathbb{R}$  – also die komplette  $s$ -Achse – Equilibria<sup>9</sup>, und ebenso sind alle Punkte der Form  $(\frac{d}{c}, i)$  mit  $i \in \mathbb{R}$  Equilibria<sup>10</sup>.

Wir interessieren uns insbesondere für die *Stabilität* und die *Attraktivität* eines Equilibriums. In der folgenden Definition präzisieren wir den Begriff *Stabilität*; über *Attraktivität* sprechen wir weiter unten in Definition 8.2.7.

**Definition 8.2.4 (Stabilität eines Equilibriums).** Sei  $\Omega \subseteq \mathbb{R}^d$  offen und nicht-leer und sei  $h: \Omega \rightarrow \mathbb{R}^d$  lokal Lipschitz stetig. Sei  $x^* \in \Omega$  ein Equilibrium der Differentialgleichung

$$\dot{x}(t) = h(x(t)).$$

Das Equilibrium  $x^*$  heißt...

<sup>9</sup>Wovon aber natürlich nur die Punkte mit  $s \geq 0$  biologisch relevant sind.

<sup>10</sup>Wovon natürlich nur die Punkte mit  $i \geq 0$  biologisch relevant sind.

- (i) ...*stabil*, falls es für jede Umgebung  $W \subseteq \Omega$  von  $x^*$  eine Umgebung  $U \subseteq \Omega$  von  $x^*$  mit folgender Eigenschaft gibt: Für jeden Startwert  $x_0 \in U$  existiert die Lösung  $x$  des Anfangswertproblems

$$\begin{cases} \dot{x}(t) = h(x(t)), \\ x(0) = x_0 \end{cases}$$

global nach rechts und liegt für alle  $t \in [0, \infty)$  in  $W$ .

- (ii) ...*instabil*, wenn es nicht stabil ist.

Für lineare Differentialgleichungen kann man Stabilität des Equilibriums 0 ganz leicht beschreiben; dies tun wir in folgender Proposition:

**Proposition 8.2.5.** Sei  $A \in \mathbb{R}^d$ . Das Equilibrium 0 der Differentialgleichung

$$\dot{x}(t) = Ax(t)$$

ist genau dann stabil, wenn die äquivalenten Aussagen aus Satz 8.1.5 erfüllt sind.

*Beweis.* „ $\Rightarrow$ “ Sei die Ruhelage 0 stabil, und sei  $W \subseteq \mathbb{R}^d$  die abgeschlossene Einheitskugel; diese ist eine Umgebung von 0. Dann gibt es laut Definition der Stabilität eine Umgebung  $U$  von 0 mit der Eigenschaft  $e^{tA}x_0 \in W$  für alle  $t \in [0, \infty)$  und alle  $x_0 \in U$ . Wir wählen nun ein  $\varepsilon > 0$  derart, dass die abgeschlossene Kugel mit Radius  $\varepsilon$  um 0 in  $U$  enthalten ist.

Dann gilt für alle  $x_0 \in \mathbb{R}^d$  mit Norm  $\|x_0\| \leq \varepsilon$ , dass  $x_0 \in U$  ist und somit  $\|e^{tA}x_0\| \leq 1$  für alle  $t \in [0, \infty)$  ist. Für alle  $x_0 \in \mathbb{R}^d$  mit Norm höchstens 1 folgt daraus aber, dass  $\|e^{tA}x_0\| \leq \frac{1}{\varepsilon}$  für alle  $t \in [0, \infty)$  ist. Somit ist

$$\|e^{tA}\| \leq \frac{1}{\varepsilon}$$

für alle  $t \in [0, \infty)$  gezeigt.

„ $\Leftarrow$ “ Sei  $C := \sup_{t \in [0, \infty)} \|e^{tA}\| \in [1, \infty)$ , und sei  $W \subseteq \mathbb{R}^d$  eine Umgebung von 0. Dann gibt es ein  $\varepsilon > 0$  derart, dass die abgeschlossene Kugel mit Radius  $\varepsilon$  um 0 in  $W$  liegt. Wir wählen nun  $U$  als die abgeschlossene Kugel um 0 mit Radius  $\frac{\varepsilon}{C}$ . Dann gilt für jedes  $x_0 \in U$  und jedes  $t \in [0, \infty)$  die Abschätzung

$$\|e^{tA}x_0\| \leq \|e^{tA}\| \|x_0\| \leq C \frac{\varepsilon}{C} = \varepsilon,$$

und somit  $e^{tA}x_0 \in W$ . □

**Beispiel 8.2.6.** Für die Matrix

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

hatten wir in Beispiel 8.1.6(a) bereits besprochen, dass die äquivalenten Aussagen von Satz 8.1.5 erfüllt sind. Also ist das Equilibrium 0 der Differentialgleichung  $\dot{x}(t) = Ax(t)$  laut Proposition 8.2.5 stabil.

Zugleich verhalten sich alle Lösungen der Gleichung

$$\dot{x}(t) = Ax(t)$$

periodisch (mit Periode  $2\pi$ ), d.h. die Lösungen konvergieren nicht gegen das Equilibrium 0 für  $t \rightarrow \infty$  (es sei denn der Startwert ist bereits 0).

Der letzte Satz in vorangehendem Beispiel zeigt, dass Stabilität etwas anderes ist als Konvergenz der Lösungen gegen das Equilibrium. Mit Konvergenz der Lösungen gegen Equilibria beschäftigen wir uns in der folgenden Definition.

**Definition 8.2.7 (Attraktivität und asymptotische Stabilität eines Equilibriums).** Sei  $\Omega \subseteq \mathbb{R}^d$  offen und nichtleer und sei  $h : \Omega \rightarrow \mathbb{R}^d$  lokal Lipschitz stetig. Sei  $x^* \in \Omega$  ein Equilibrium der Differentialgleichung

$$\dot{x}(t) = h(x(t)).$$

Das Equilibrium  $x^*$  heißt...

- (i) ...*attraktiv*, falls es eine Umgebung  $U \subseteq \Omega$  von  $x^*$  mit folgender Eigenschaft gibt: Für jeden Startwert  $x_0 \in U$  existiert die Lösung  $x$  des Anfangswertproblems

$$\begin{cases} \dot{x}(t) = h(x(t)), \\ x(0) = x_0 \end{cases}$$

global nach rechts und konvergiert für  $t \rightarrow \infty$  gegen  $x^*$ .

- (ii) ...*asymptotisch stabil*, wenn es stabil und attraktiv ist.

Für lineare Differentialgleichungen haben wir Attraktivität der Ruhelage 0 im Grunde bereits in Abschnitt 8.1 charakterisiert. Wir fassen dies in der folgenden Proposition noch einmal unter Verwendung der soeben eingeführten Terminologie zusammen:

**Proposition 8.2.8.** Sei  $A \in \mathbb{R}^d$ . Das Equilibrium 0 der Differentialgleichung

$$\dot{x}(t) = Ax(t)$$

ist attraktiv genau dann, wenn es asymptotisch stabil ist, genau dann, wenn es die äquivalenten Aussagen aus den Sätzen 8.1.1 und 8.1.3 erfüllt.

*Beweis.* Der Beweis ist leicht, wenn man die Ergebnisse aus Abschnitt 8.1 verwendet. Die Details besprechen wir in den Übungen.  $\square$

Für die Ruhelage 0 von linearen Differentialgleichungen ist Attraktivität also dasselbe wie asymptotische Stabilität<sup>11</sup>. Es gibt aber tatsächlich Beispiele nicht-linearer Differentialgleichungen mit einem Equilibrium, das zwar attraktiv, aber nicht stabil ist<sup>12</sup>; aus diesem Grund ist der Begriff *asymptotisch stabil* im Allgemeinen tatsächlich etwas stärker als der Begriff *attraktiv*.<sup>13</sup>

In den nachfolgenden beiden Abschnitten werden wir Methoden besprechen, mit denen man nachweisen kann, dass ein Equilibrium asymptotisch stabil ist.

### 8.3 Linearisierte Stabilität

In diesem Abschnitt betrachten wir ein Equilibrium  $x^*$  einer autonomen Differentialgleichung mit rechter Seite  $h$ . Wir nehmen zusätzlich an, dass  $h$  stetig differenzierbar ist und zeigen, dass man mit Hilfe der Ableitung von  $h$  im Punkt  $x^*$  – also anhand der Matrix  $Dh(x^*) \in \mathbb{R}^{d \times d}$  – in manchen Fällen nachweisen kann, dass das Equilibrium asymptotisch stabil ist  $x^*$ . Dies ist der Inhalt des folgenden Satzes:

**Satz 8.3.1 (Linearisierte Stabilität).** Sei  $\Omega \subseteq \mathbb{R}^d$  nicht-leer und offen, und sei  $h : \Omega \rightarrow \mathbb{R}^d$  stetig differenzierbar. Außerdem sei  $x^* \in \Omega$  ein Equilibrium der Differentialgleichung

$$\dot{x}(t) = h(x(t)).$$

Wenn  $s(Dh(x^*)) < 0$  gilt<sup>14</sup> – d.h. wenn alle Eigenwerte der Matrix  $Dh(x^*)$  Realteil  $< 0$  haben –, dann ist das Equilibrium  $x^*$  asymptotisch stabil.

Bevor wir den Satz beweisen, ist es sinnvoll ein paar Worte zur Bezeichnung *Linearisierte Stabilität* zu verlieren: Erinnern Sie sich zunächst daran, dass die Definition der Ableitung von  $h$  im Punkt  $x^*$  – also der Matrix  $Dh(x^*)$  – gerade so ist, dass man die Funktion  $h$  für Punkte  $y \in \Omega$ , die nahe an  $x^*$  liegen, gut approximieren kann durch die Formel

$$h(y) \approx h(x^*) + Dh(x^*)(y - x^*); \quad (8.1)$$

<sup>11</sup>In anderen Worten: Attraktivität der Ruhelage 0 impliziert für lineare Differentialgleichungen bereits Stabilität.

<sup>12</sup>Auch wenn dies auf den ersten Blick vermutlich etwas überraschend wirkt.

<sup>13</sup>Bitte beachten Sie auch nochmals, dass der Begriff *asymptotisch stabil* auch stärker ist als der Begriff *stabil*; dies gilt sogar im linearen Fall, wie wir in Beispiel 8.2.6 gesehen hatten.

<sup>14</sup>An dieser Stelle sei daran erinnert, dass die *Spektralschranke*  $s(A)$  einer Matrix  $A \in \mathbb{R}^{d \times d}$  definiert ist als das Maximum der Realteile aller Eigenwerte von  $A$ .

hierbei bedeutet „gut approximieren“, dass der Fehler zwischen der linken und der rechten Seite dieser „approximativen Gleichheit“ schneller als  $\|y - x^*\|$  gegen 0 strebt, wenn  $y$  gegen  $x^*$  geht.

Nun ist  $x^*$  in Satz 8.3.1 aber ein Equilibrium, d.h. es gilt  $h(x^*) = 0$ . Damit fällt in der Approximation (8.1) der Term nullter Ordnung weg – es ist also

$$h(y) \approx Dh(x^*)(y - x^*)$$

für  $y$  nahe bei  $x^*$ . Stellen Sie sich jetzt zum Schluss noch der Einfachheit halber vor, dass das Equilibrium  $x^*$  gleich dem Nullpunkt ist<sup>15</sup>; dann ist  $h(y) \approx D(h^*)y$  für alle  $y$ , die nahe am Equilibrium  $x^* = 0$  liegen – und somit stimmt die Differentialgleichung

$$\dot{x}(t) = h(x(t))$$

zumindest in einer Umgebung des Nullpunkts näherungsweise mit der linearen Differentialgleichung

$$\dot{x}(t) = Dh(x^*)x(t)$$

überein. Die Bedingung  $s(Dh(x^*)) < 0$  in Satz 8.3.1 besagt gerade, dass das Equilibrium 0 dieser linearen Differentialgleichung asymptotisch stabil ist (siehe Proposition 8.2.8 und Satz 8.1.1(iv)). Damit besagt der Satz 8.3.1 von der linearisierten Stabilität also:

Das Equilibrium  $x^*$  von  $\dot{x}(t) = h(x(t))$  (das wir nach Verschiebung als gleich 0 annehmen können) ist asymptotisch stabil, falls das Equilibrium 0 der *linearisierten* Differentialgleichung  $\dot{x}(t) = Dh(x^*)x(t)$  asymptotisch stabil ist.<sup>16</sup>

Mit den oben stehenden Erläuterungen ist übrigens auch schon die Beweisstrategie für Satz 8.3.1 beschrieben. Lassen Sie uns diese Strategie nun im Detail ausführen.

*Beweis von Satz 8.3.1.* Wir beginnen mit zwei einfachen Beobachtungen:

- Ohne Einschränkung können (und werden) wir  $x^* = 0$  annehmen (ansonsten verschieben wir  $\Omega$  und  $h$  entsprechend).
- Wenn wir die rechte Seite  $h$  auf eine kleinere Menge  $\tilde{\Omega} \subseteq \Omega$  einschränken<sup>17</sup> und für die resultierende Differentialgleichung die asymptotische Stabilität des Equilibriums 0 zeigen, so gilt für die ursprüngliche Differentialgleichung ebenfalls die asymptotische Stabilität des Equilibriums 0; dies folgt sofort aus der Definition der Begriffe *stabil* und *attraktiv*.

<sup>15</sup>Diese Vorstellung ist gar nicht so weit weg von der Realität: Wir können immer  $x^* = 0$  erreichen, indem wir die ganze Situation um den Vektor  $-x^*$  verschieben.

<sup>16</sup>Die umgekehrte Implikation gilt übrigens nicht.

<sup>17</sup>Wobei  $\tilde{\Omega}$  ebenfalls nicht-leer und offen ist.

Also können (und werden) wir ab sofort annehmen, dass  $\Omega$  eine offene Kugel mit Mittelpunkt 0 und Radius  $R > 0$  ist. Wir können (und werden) außerdem annehmen, dass  $h$  auf  $\Omega$  beschränkt ist.<sup>18</sup>

Wir verwenden von nun an die Abkürzung  $A := Dh(x^*) = Dh(0)$ . Also können wir  $h$  schreiben als

$$h(y) = Ay + r(y) \quad \text{für } y \in \Omega,$$

wobei der Restterm  $r : \Omega \rightarrow \mathbb{R}^d$  eine Funktion ist derart, dass  $\|r(y)\|$  für  $y \rightarrow x^* = 0$  schneller als  $\|y\|$  gegen 0 geht. Anders ausgedrückt bedeutet dies: Es gibt eine stetige und monoton wachsende Funktion  $\eta : [0, R] \rightarrow [0, \infty)$  mit  $\eta(0) = 0$  und  $\|r(y)\| \leq \eta(\|y\|)\|y\|$  für alle  $y \in \Omega$ .<sup>19</sup>

Laut Voraussetzung ist  $s(A) < 0$ , also gibt es laut Satz 8.1.1(iii) und (iv) Zahlen  $M \geq 1$  und  $\delta > 0$  derart, dass  $\|e^{tA}\| \leq Me^{-\delta t}$  für alle  $t \in [0, \infty)$  ist. Die Idee ist nun mit Hilfe des Gronwall-Lemmas die Norm der Lösungen unserer Differentialgleichung abzuschätzen.

Sei also  $I \subseteq \mathbb{R}$  ein nichtleeres Intervall mit  $0 \in I$  und sei  $x : I \rightarrow \Omega$  eine Lösung von  $\dot{x}(t) = h(x(t))$ . Dies bedeutet anders geschrieben, dass  $x$  die Differentialgleichung

$$\dot{x}(t) = Ax(t) + r(x(t))$$

auf  $I$  erfüllt. Aus der Faltungsformel in Korollar 5.6.3 folgt nun (mit  $b(t) = r(x(t))$ )

$$x(t) = e^{tA}x(0) + \int_0^t e^{(t-s)A}r(x(s)) \, ds$$

für alle  $t \in I$ . Für diejenigen Zeiten  $t \in I$ , die  $\geq 0$  sind, folgt hieraus

$$\begin{aligned} \|x(t)\| &\leq Me^{-\delta t} \|x(0)\| + \int_0^t Me^{-\delta(t-s)} \|r(x(s))\| \, ds \\ &\leq Me^{-\delta t} \|x(0)\| + \int_0^t Me^{-\delta(t-s)} \eta(\|x(s)\|) \|x(s)\| \, ds \\ &\leq Me^{-\delta t} \|x(0)\| + \int_0^t Me^{-\delta(t-s)} \eta(\sigma) \|x(s)\| \, ds, \end{aligned}$$

<sup>18</sup>Denn wenn  $h$  nicht beschränkt ist, verkleinern wir den Radius  $R$  noch ein wenig, wodurch wir die Beschränktheit von  $h$  erreichen.

<sup>19</sup>Hierbei verwenden wir, dass  $\Omega$  eine Kugel mit Radius  $R$  und Mittelpunkt 0 ist. Man kann zum Beispiel  $\eta(a) = \sup\{\frac{\|r(y)\|}{\|y\|} : y \in \Omega, 0 < \|y\| \leq a\}$  für  $a \in (0, R)$  wählen. Im Punkt 0 wählt man  $\eta$  gleich 0, und im Punkt  $R$  setzt man  $\eta$  durch  $\eta(R) = \lim_{a \uparrow R} \eta(a)$  fort. Die Stetigkeit von  $\eta$  in 0 rührt gerade daher, dass  $r(y)$  für  $\|y\| \rightarrow 0$  schneller als  $\|y\|$  gegen 0 geht.

Dass wir  $\eta$  auch an der Stelle  $R$  – und nicht nur auf dem halboffenen Intervall  $[0, R)$  – definieren, ist nur ein notationeller Trick; er erleichtert nachher die Rechnung ein wenig.

wobei  $\sigma := \sup\{\|x(t)\| : t \in I\}$  ist<sup>20</sup>. Wenn wir  $\psi(t) := e^{\delta t} \|x(t)\|$  für alle  $t \in I$  mit  $t \geq 0$  definieren, dann folgt also für all solche  $t$  die Abschätzung

$$\psi(t) \leq M \|x(0)\| + \int_0^t M \eta(\sigma) \psi(s) \, ds.$$

Das Grownwall-Lemma 7.1.1 zeigt, dass aus dieser Integralungleichung die explizite Ungleichung

$$\psi(t) \leq e^{M\eta(\sigma)t} M \|x(0)\|,$$

und somit

$$\|x(t)\| \leq e^{(M\eta(\sigma)-\delta)t} M \|x(0)\| \quad (8.2)$$

für alle  $t \in I$  mit  $t \geq 0$  folgt.

Nun wählen wir ein  $a \in (0, R)$ , für welches  $\eta(a)$  so klein ist, dass  $M\eta(a) - \delta < 0$  gilt; dies ist möglich, da  $\eta$  im Punkt 0 stetig und gleich 0 ist. Für jedes  $b \in (0, a]$  folgt dann die folgende Eigenschaft unserer Differentialgleichung:

(\*) Wenn  $x_0 \in \Omega$  mit  $M \|x_0\| < b$  ist und  $x : I_{\max} \rightarrow \Omega$  die Lösung des Anfangswertproblems

$$\begin{cases} \dot{x}(t) = h(x(t)), \\ x(0) = x_0 \end{cases}$$

bezeichnet, die auf dem maximalen Existenzintervall definiert ist, dann gilt  $\|x(t)\| < b$  für alle  $t \in I_{\max}$  mit  $t \geq 0$ :

Andernfalls könnten wir nämlich das kleinste  $\hat{t} \in I_{\max} \cap [0, \infty)$  wählen, für welches  $\|x(\hat{t})\| \geq b$  gilt. Dann wäre  $\hat{t} > 0$  (wegen  $\|x(0)\| = \|x_0\| \leq M \|x_0\| < b$ ) und für die Einschränkung der Lösung  $x$  auf das Intervall  $[0, \hat{t})$  würde  $\sigma := \sup_{t \in [0, \hat{t})} \|x(t)\| \leq b$  und folglich  $\eta(\sigma) \leq \eta(b) \leq \eta(a)$  gelten. Dann erhielten wir aber aus der oben gezeigten Ungleichung (8.2), dass

$$\|x(t)\| \leq M \|x_0\|$$

für alle  $t \in [0, \hat{t})$  ist. Wegen der Stetigkeit von  $x$  folgt dann aber auch  $\|x(\hat{t})\| \leq M \|x_0\| < b$ , was im Widerspruch zur Definition von  $\hat{t}$  steht. Also ist (\*) bewiesen.

Aus der Aussage (\*) können wir jetzt leicht die Stabilität und die Attraktivität des Equilibriums 0 folgern:

*Stabilität:* Sei  $W \subseteq \Omega$  eine Umgebung von  $x^* = 0$ . Dann gibt es ein  $\varepsilon > 0$  derart, dass die abgeschlossene Kugel mit Radius  $\varepsilon$  und Mittelpunkt 0 in  $W$

<sup>20</sup>An dieser Stelle ist es praktisch, dass wir  $\eta$  auch an der Stelle  $R$  definiert hatten: Obwohl nämlich  $\|x(t)\| < R$  für alle  $t \in I$  gilt, kann das Supremum  $\sigma$  unter Umständen dennoch gleich  $R$  sein.

enthalten ist. Wir setzen  $b := \min\{a, \varepsilon\}$  und wählen  $U$  als die offene Kugel um 0 mit Radius  $\frac{b}{M}$ . Für einen Startwert  $x_0 \in U$  erfüllt dann die Lösung  $x : I_{\max} \rightarrow \Omega$  des zugehörigen Anfangswertproblems wegen (\*) die Eigenschaft  $\|x(t)\| < b \leq a < R$  für alle  $0 \leq t \in I_{\max}$ . Die Charakterisierung des Randverhaltens von Lösungen, die wir in Korollar 4.3.8 bewiesen haben, zeigt somit, dass der rechte Randpunkt von  $I_{\max}$  gleich  $\infty$  ist – d.h. die Lösung  $x$  existiert global nach rechts. Außerdem ist wegen  $\|x(t)\| < b \leq \varepsilon$  der Vektor  $x(t)$  in  $W$  enthalten für jedes  $t \geq 0$ . Dies zeigt die Stabilität des Equilibriums 0.

*Attraktivität:* Es sei  $U$  die offene Kugel um 0 mit Radius  $\frac{a}{M}$ . Sei  $x_0 \in U$  und sei  $x : I_{\max} \rightarrow \Omega$  die Lösung des Anfangswertproblems mit Startwert  $x_0$ . Wegen (\*) gilt dann  $\|x(t)\| \leq a < R$  für alle  $0 \leq t \in I_{\max}$ , also folgt auch hier aus der Charakterisierung des Randverhaltens in Korollar 4.3.8, dass die Lösung  $x$  global nach rechts existiert. Außerdem können wir nun die Ungleichung (8.2) auf die Lösung  $x$  anwenden; wegen  $\sigma \leq a$  ist der Faktor  $M\eta(\sigma) - \delta$  im Exponenten negativ, also folgt  $x(t) \rightarrow 0$  für  $t \rightarrow \infty$ .  $\square$

Wir diskutieren zwei Beispiele für die Anwendungen dieses Satzes.

**Beispiele 8.3.2.** (a) Lassen Sie uns wieder einmal die logistische Gleichung

$$\dot{b}(t) = ((c - db(t))b(t))$$

mit Konstanten  $c, d > 0$  betrachten. Diese Differentialgleichung besitzt offenbar die Equilibria 0 und  $\frac{c}{d}$  (dies hatten wir auch schon in Beispiel 8.2.3(b) kurz besprochen).

Die rechte Seite können wir schreiben als  $h(b(t))$  schreiben mit

$$h : \mathbb{R} \ni y \mapsto (c - dy)y \in \mathbb{R}.$$

Für die Ableitung von  $h$  gilt  $Dh(y) = c - 2dy$  für alle  $y \in \mathbb{R}$ . Somit ist

$$Dh\left(\frac{c}{d}\right) = c - 2d\frac{c}{d} = -c < 0.$$

Der einzige Eigenwert der Matrix  $Dh\left(\frac{c}{d}\right) \in \mathbb{R}^{1 \times 1} = \mathbb{R}$  ist somit gleich  $-c$ , also negativ. Deshalb ist das Equilibrium  $\frac{c}{d}$  laut Satz 8.3.1 asymptotisch stabil.

Das Equilibrium 0 werden wir in Beispiel 8.3.6 besprechen.

(b) Lassen Sie uns die zwei-dimensionale Differentialgleichung

$$\dot{x}(t) = h(x(t))$$

betrachten, wobei

$$h : \mathbb{R}^2 \rightarrow \mathbb{R}^2,$$

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \mapsto \begin{pmatrix} y_1^2 - y_1 - y_2 + y_1 y_2 \\ y_1 + y_2^2 \end{pmatrix}$$

ist.<sup>21</sup>

Es ist  $h(0) = 0$ , d.h. der Nullpunkt ist ein Equilibrium der Differentialgleichung. Die Ableitung von  $h$  ist gegeben durch

$$Dh(y) = \begin{pmatrix} 2y_1 - 1 + y_2 & -1 + y_1 \\ 1 & 2y_2 \end{pmatrix}$$

für alle  $y \in \mathbb{R}^2$ , und somit ist

$$Dh(0) = \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}.$$

Man kann leicht ausrechnen, dass diese Matrix die Eigenwerte  $\frac{-1+i\sqrt{3}}{2}$  und  $\frac{-1-i\sqrt{3}}{2}$  besitzt. Also ist die Spektralschranke von  $Dh(0)$  gleich  $-\frac{1}{2}$ ; deshalb können wir Satz 8.3.1 anwenden und erhalten, dass das Equilibrium 0 der Differentialgleichung asymptotisch stabil ist.

In Beispiel 8.3.2(b) konnten wir die Eigenwerte von  $Dh(0)$  leicht ausrechnen. Es gibt aber auch Beispiele, in denen dies nicht so leicht möglich ist – insbesondere dann, wenn in der Differentialgleichung auch Parameter vorkommen. In solchen Fällen ist es natürlich hilfreich, wenn wir andere Kriterien dafür haben, dass die Spektralschranke einer Matrix kleiner als 0 ist. Ein einfaches Beispiel für solch ein Kriterium im Fall von  $2 \times 2$ -Matrizen ist die folgende Proposition:

**Proposition 8.3.3 (Routh–Hurwitz-Kriterium<sup>22</sup> im 2D-Fall<sup>23</sup>).** Sei  $A \in \mathbb{R}^{2 \times 2}$ . Dann sind äquivalent:

- (i) Es gilt  $s(A) < 0$ .
- (ii) Es gilt  $\text{spur}(A) < 0$  und  $\det(A) > 0$ .

*Beweis.* Der Beweis ist eine schöne Übungsaufgabe, die wir auf Blatt 13 diskutieren. □

<sup>21</sup>Dies ist ein konstruiertes Beispiel um die Anwendung von Satz 8.3.1 mit möglichst einfachen Zahlen zu demonstrieren; das Beispiel ist also nicht durch irgendein konkretes Modell motiviert.

<sup>22</sup>Benannt nach Edward John Routh (1831 – 1907), englischer Mathematiker, und Adolf Hurwitz (1859 – 1919), deutscher Mathematiker.

<sup>23</sup>Es gibt auch ein Routh–Hurwitz-Kriterium für höher-dimensionale Matrizen, welches aber komplizierter ist; siehe die Literaturverweise am Ende von Abschnitt 8.3.

Nun, da wir dieses Kriterium angesprochen haben, lohnt es sich nochmals einen Blick auf Beispiel 8.3.2(b) zu werfen,

**Beispiel 8.3.4.** In Beispiel 8.3.2(b) war die Ableitung von  $h$  im Equilibrium 0 gegeben durch

$$Dh(0) = \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}.$$

Es gilt  $\text{spur}(Dh(0)) = -1 < 0$  und  $\det(Dh(0)) = 1 > 0$ , also folgt aus dem zwei-dimensionalen Routh–Hurwitz-Kriterium in Proposition 8.3.3, dass die Spektralschranke von  $Dh(0)$  kleiner als 0 ist.

Selbst für solch ein simples Zahlenbeispiel kann es also einfacher sein das Routh–Hurwitz-Kriterium anzuwenden, als die Eigenwerte konkret auszurechnen.

Es gibt auch eine umgekehrte Variante des Satzes von der linearisierten Stabilität: Wenn die Spektralschranke  $s(Dh(x^*))$  strikt größer als 0 ist, dann ist das Equilibrium  $x^*$  instabil. Die Details hierzu finden Sie in folgendem Satz.

**Satz 8.3.5 (Linearisierte Instabilität<sup>24</sup>).** Sei  $\Omega \subseteq \mathbb{R}^d$  nicht-leer und offen, und sei  $h : \Omega \rightarrow \mathbb{R}^d$  stetig differenzierbar. Außerdem sei  $x^* \in \Omega$  ein Equilibrium der Differentialgleichung

$$\dot{x}(t) = h(x(t)).$$

Wenn  $s(Dh(x^*)) > 0$  gilt – d.h. wenn die Matrix  $Dh(x^*)$  einen Eigenwert mit Realteil  $> 0$  besitzt –, dann ist das Equilibrium  $x^*$  instabil.

*Beweis.* Der Beweis ist etwas aufwendiger als der Beweis von Satz 8.3.1. Wir veweisen für den Beweis auf die Literatur, z.B. auf [PW19, Satz 5.4.1(2.)].  $\square$

**Beispiel 8.3.6.** Wie in Beispiel 8.3.2(a) angekündigt kommen wir noch einmal zur logistischen Gleichung

$$\dot{b}(t) = (c - db(t))b(t)$$

(mit  $c, d > 0$ ) zurück und untersuchen nun die Stabilitätseigenschaften des Equilibriums 0: Es gilt  $Dh(0) = c > 0$ , und somit folgt aus Satz 8.3.5, dass das Equilibrium 0 instabil ist.

Anschaulich sollte dies keine Überraschung sein: Wenn es 0 Bakterien ins unserer Petrischale gibt, können diese sich auch nicht vermehren; somit

<sup>24</sup>Die Bezeichnung *Linearisierte Instabilität* scheint in der Literatur allerdings recht unüblich zu sein; wenn Sie eine Bezeichnung verwenden wollen, die möglichst weitläufig verstanden wird, ist es wahrscheinlich besser, wenn Sie beide Sätze 8.3.1 und 8.3.5 als *Satz von der linearisierten Stabilität* bezeichnen.

bleibt es auch bei der Anzahl 0 – dies besagt, dass 0 ein Equilibrium ist. Wenn wir aber diese Zahl 0 ein klein wenig stören – also ein paar Bakterien in der Petrischale platzieren – beginnen diese, sich zu vermehren, und die Bakterienanzahl entwickelt sich somit weg von 0; dies bedeutet, dass das Equilibrium 0 instabil ist.

Die beiden Sätze 8.3.1 und 8.3.5 lassen die Frage offen, was im Falle  $s(Dh(x^*)) = 0$  passiert. Tatsächlich ist es so, dass man nur anhand der Information  $s(Dh(x^*)) = 0$  nicht entscheiden kann, welche Stabilitätseigenschaften das Equilibrium  $x^*$  besitzt. Hierzu kann man entsprechende Beispiele konstruieren, die wir an dieser Stelle aber nicht besprechen.

**Literaturstellen und verwandte Ergebnisse.** Auch für Matrizen  $A \in \mathbb{R}^{d \times d}$  in höherer Dimension lässt sich die Frage, ob  $s(A) < 0$  gilt, mit Hilfe eines Kriteriums von *Routh–Hurwitz* untersuchen. Mehr dazu können Sie zum Beispiel in [PW19, Seiten 119–120] nachlesen.

Das Resultat in Proposition 8.3.3 ist ein Spezialfall dieses allgemeinen Routh–Hurwitz-Kriteriums.

## 8.4 Ljapunov-Funktionen

Der Satz 8.3.1 über die linearisierte Stabilität ist sehr nützlich, weil er ein recht allgemeines Verfahren liefert, um die asymptotische Stabilität eines Equilibriums zu zeigen: Man linearisiere die rechte Seite  $h$  der Differentialgleichung und bestimme die Spektralschranke. Allerdings hat der Satz die folgenden Nachteile:

- Manchmal ist es nicht leicht, die Spektralschranke einer Matrix zu bestimmen. Dies gilt insbesondere dann, wenn die Matrix hohe Dimension hat oder wenn Parameter in ihr vorkommen. Auch das Routh–Hurwitz-Kriterium kann hier nicht immer Abhilfe schaffen – denn es gibt zwar eine Variante dieses Kriteriums, die in beliebiger Dimension gilt, aber diese wird mit steigender Dimension sehr komplex.
- Der Satz macht keine Aussage, wenn die Spektralschranke gleich 0 ist. Manchmal hat man aber auch unter diesen Umständen asymptotische Stabilität des Equilibriums; also möchte man gerne eine Methode haben, mit der man auch solche Fälle behandeln kann.
- Der Satz liefert nur ein *lokales* Ergebnis: Über Startwerte, die weit weg vom Equilibrium liegen, sagt der Satz nichts auf.<sup>25</sup> Manchmal ist es aber

<sup>25</sup>Bitte sehen Sie sich hierzu noch einmal die Definition von *Attraktivität* an (Definition 8.2.7) und machen Sie sich dabei klar, weshalb dies eine lokale Eigenschaft ist.

naheliegend zu erwarten, dass ein Equilibrium *global attraktiv* ist – d.h. dass alle Lösungen der Differentialgleichung gegen das Equilibrium konvergieren, und zwar unabhängig davon, wo der Startwert liegt.

In diesem Abschnitt werden wir zwei Sätze besprechen, die für die oben stehenden Punkte manchmal Abhilfe schaffen – Voraussetzung dieser Sätze ist aber jeweils, dass man eine Ljapunov-Funktion<sup>26</sup> findet, die geeignete Zusatzeigenschaften besitzt<sup>27</sup>.

Der erste der zwei Sätze in diesem Abschnitt gibt ein hinreichendes Kriterium für Stabilität mit Hilfe einer Ljapunov-Funktion an:

**Satz 8.4.1 (Stabilität mittels Ljapunov-Funktion).** *Sei  $\Omega \subseteq \mathbb{R}^d$  offen und nicht-leer und sei  $h : \Omega \rightarrow \mathbb{R}^d$  lokal Lipschitz-stetig. Zudem sei  $V : \Omega \rightarrow \mathbb{R}$  eine Ljapunov-Funktion für die Differentialgleichung  $\dot{x}(t) = h(x(t))$ .*

*Wenn  $x^* \in \Omega$  ein Equilibrium der Differentialgleichung  $\dot{x}(t) = h(x(t))$  und zugleich ein striktes (lokales) Minimum von  $V$  ist, dann ist  $x^*$  stabil.*

*Beweis.* Sei  $W \subseteq \Omega$  eine Umgebung von  $x^*$ . Weil  $x^*$  ein striktes lokales Minimum von  $V$  ist, gibt es ein  $\varepsilon > 0$  derart, dass die abgeschlossene Kugel  $B$  mit Radius  $\varepsilon$  und Mittelpunkt  $x^*$  einerseits komplett in  $W$  liegt, und dass andererseits  $V(y) > V(x^*)$  für alle  $y \in B$  gilt. Sei nun  $\beta \in \mathbb{R}$  der minimale Wert von  $V$  auf dem Rand der Kugel  $B$ . Aus Kompaktheitsgründen existiert dieser minimale Wert und ist zudem größer als  $V(x^*)$ .

Mit  $U$  bezeichnen wir den Durchschnitt des Urbilds  $V^{-1}((-\infty, \beta))$  mit der offenen Kugel mit Radius  $\varepsilon$  und Mittelpunkt  $x^*$ . Dann ist  $U$  eine offene Umgebung von  $x^*$ , die komplett in  $W$  enthalten ist. Weil  $V$  eine Ljapunov-Funktion ist, ist das Urbild  $V^{-1}((-\infty, \beta))$  positiv invariant für unsere Differentialgleichung, und weil  $V$  überall auf dem Rand von  $B$  mindestens gleich  $\beta$  ist, kann eine Lösung, die in  $V^{-1}((-\infty, \beta))$  startet, für positive Zeiten niemals den Rand von  $B$  erreichen. Also ist auch  $U$  positiv invariant für unsere Differentialgleichung.

Weil  $U$  in der kompakten Teilmenge  $B$  von  $\Omega$  enthalten ist, existiert also jede Lösung, die in  $U$  startet, global nach rechts – dies folgt aus der Charakterisierung des Randverhaltens in Satz 4.3.6. Zudem bleibt eine Lösung, die in  $U$  startet, wegen der positiven Invarianz für alle Zeiten  $t \geq 0$  in  $U$  und somit in  $W$ . Damit ist die Stabilität des Equilibriums  $x^*$  gezeigt.  $\square$

Für den zweiten Satz, den wir in diesem Abschnitt behandeln wollen, benötigen wir den Begriff der *strikten Ljapunov-Funktion*, welcher eine Ver-

<sup>26</sup>Zur Erinnerung: Ljapunov-Funktionen hatten wir in Definition 4.4.1 eingeführt. Es ist empfehlenswert, dass Sie sich den Unterabschnitt *Ljapunov-Funktionen* von Abschnitt 4.4 noch einmal durchlesen, damit Sie sich an diesen Begriff und seine Eigenschaften erinnern.

<sup>27</sup>Und dies ist nicht immer einfach.

schärfung des Begriffs *Ljapunov-Funktion* ist. Um der Konsistenz mit Definition 4.4.1 Willen führen wir den Begriff unter denselben allgemeinen Bedingungen wie in dieser Definition ein – auch wenn wir ihn im Folgenden nur für autonome Differentialgleichungen benötigen.

**Definition 8.4.2 (Strikte Ljapunov-Funktion).** Seien  $I \subseteq \mathbb{R}$  und  $\Omega \subseteq \mathbb{R}^d$  nichtleer und offen, wobei  $I$  ein Intervall sei. Sei  $f : I \times \Omega \rightarrow \mathbb{R}^d$  stetig und erfülle die lokale Lipschitz-Bedingung.

Eine stetige Funktion  $V : \Omega \rightarrow \mathbb{R}$  heißt *strikte Ljapunov-Funktion* für die Differentialgleichung

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

falls für jedes nicht-triviale Intervall  $J \subseteq I$  und jede nicht-konstante Lösung  $x : J \rightarrow \Omega$  dieser Differentialgleichung gilt: Die Funktion

$$V \circ x : J \ni t \mapsto V(x(t)) \in \mathbb{R}$$

ist streng monoton fallend.

Jede strikte Ljapunov-Funktion ist also eine Ljapunov-Funktion; umgekehrt ist eine Ljapunov-Funktion sogar strikt, wenn Sie entlang nicht konstanter Lösungen nicht nur monoton fallend, sondern sogar *streng* monoton fallend ist.

In Proposition 4.4.2 hatten wir Ljapunov-Funktionen mit Hilfe der sogenannten *orbitalen Ableitung*  $\langle \nabla V(y), f(t, y) \rangle$  charakterisiert. Für strikte Ljapunov-Funktionen ist eine Charakterisierung nicht so einfach möglich – aber zumindest liefert die orbitale Ableitung ein hinreichendes Kriterium dafür, dass eine Funktion  $V$  eine Ljapunov-Funktion ist.

**Proposition 8.4.3.** Seien  $I, \Omega$  und  $f$  wie in Definition 8.4.2. Es sei  $V : \Omega \rightarrow \mathbb{R}$  stetig differenzierbar.

Für alle  $y \in \Omega$ , die kein Equilibrium der Differentialgleichung  $\dot{x}(t) = f(t, x(t))$  sind, und für alle  $t \in I$  gelte  $\langle \nabla V(y), f(t, y) \rangle < 0$ . Dann ist  $V$  eine strikte Ljapunov-Funktion für die Differentialgleichung  $\dot{x}(t) = f(t, x(t))$ .

*Beweis.* Der Beweis ist genauso einfach wie der Beweis der Implikation „(ii)  $\Rightarrow$  (i)“ in Proposition 4.1.11. Man muss lediglich zusätzlich beachten, dass eine nicht-konstante Lösung niemals durch ein Equilibrium verlaufen kann (dies folgt aus dem globalen Eindeutigkeitssatz von Picard–Lindelöf, den wir in Korollar 4.1.12 besprochen haben).  $\square$

Die umgekehrte Implikation gilt in obiger Proposition nicht. Ein Beispiel, welches dies zeigt, werden wir in den Übungen diskutieren.

Nun folgt der zweite Satz in diesem Abschnitt; er zeigt, dass man strikte Ljapunov-Funktionen verwenden kann, um für manche Differentialgleichung Attraktivität<sup>28</sup> eines Equilibriums zu zeigen.

**Satz 8.4.4 (Globale Attraktivität mittels Ljapunov-Funktion).** Sei  $\Omega \subseteq \mathbb{R}^d$  offen und nicht-leer und sei  $h : \Omega \rightarrow \mathbb{R}^d$  lokal Lipschitz-stetig. Es sei  $V : \Omega \rightarrow \mathbb{R}$  eine strikte Ljapunov-Funktion für die Differentialgleichung  $\dot{x}(t) = h(x(t))$ , welche die folgenden beiden Eigenschaften erfüllt:<sup>29</sup>

(a) Es gilt  $\lim_{\|y\| \rightarrow \infty} V(y) = \infty$ .

(b) Es gilt  $\lim_{y \rightarrow \partial\Omega} V(y) = \infty$ .

Wenn die Differentialgleichung  $\dot{x}(t) = h(x(t))$  genau ein Equilibrium  $x^* \in \Omega$  besitzt, dann konvergieren alle Lösungen<sup>30</sup> der Differentialgleichung  $\dot{x}(t) = h(x(t))$  für  $t \rightarrow \infty$  gegen  $x^*$ .

Einen Beweis des Satzes können Sie, wenn Sie möchten, in den Ergänzungen 8.5 nachlesen. Innerhalb der Vorlesung verzichten wir auf den Beweis; stattdessen machen wir die folgende Bemerkung und diskutieren anschließend ein Beispiel.

**Bemerkung 8.4.5.** Im Gegensatz zu Satz 8.4.1 haben wir in Satz 8.4.4 nicht explizit angenommen, dass  $x^*$  ein Minimum von  $V$  ist. Man kann sich aber leicht überlegen, dass die übrigen Voraussetzungen in Satz 8.4.4 implizieren, dass  $x^*$  automatisch ein globales Minimum von  $V$  ist.

Als einfaches Anwendungsbeispiel für Satz 8.4.4 diskutieren wir noch einmal die logistische Gleichung auf dem Zustandsraum  $(0, \infty)$ :

**Beispiel 8.4.6.** Es seien wieder  $c, d > 0$  und wir betrachten erneut die logistische Gleichung

$$\dot{b}(t) = (c - db(t))b(t)$$

auf dem Zustandsraum  $(0, \infty)$ . Die rechte Seite bezeichnen wir mit  $h : (0, \infty) \rightarrow \mathbb{R}$ , d.h. es sei  $h(y) = (c - dy)y$  für alle  $y \in (0, \infty)$ . Offensichtlich ist das einzige Equilibrium der Differentialgleichung im Zustandsraum  $(0, \infty)$  der Punkt  $x^* := \frac{c}{d}$ .

In Beispiel 4.4.4 hatten wir bereits gezeigt, dass

$$\tilde{V} : (0, \infty) \rightarrow \mathbb{R}, \quad \tilde{V}(y) = \frac{(c - dy)^2}{y} \text{ für } y \in (0, \infty)$$

<sup>28</sup>Sogar globale Attraktivität.

<sup>29</sup>Diese beiden Eigenschaften hatten wir bereits in Satz 4.4.3 betrachtet um globale Existenz der Lösungen nach rechts zu erhalten.

<sup>30</sup>Die laut Satz 4.4.3 alle global nach rechts existieren.

eine Ljapunov-Funktion ist. Diese geht außerdem bei  $\infty$  und am Rand des Zustandsraums  $(0, \infty)$  gegen  $\infty$ . Für die orbitale Ableitung von  $\tilde{V}$  hatten wir in Beispiel 4.4.4 die Formel

$$\langle \nabla \tilde{V}(y), h(y) \rangle = -2c(c - dy)^2 \frac{1}{y}$$

für alle  $y \in (0, \infty)$  berechnet, und dieser Werte ist negativ für alle  $y$  im Zustandsraum, mit Ausnahme des Punktes  $x^*$ . Also ist  $\tilde{V}$  eine strikte Ljapunov-Funktion für unsere Differentialgleichung (siehe Proposition 8.4.3).

Somit folgt aus Satz 8.4.4, dass alle Lösungen der logistischen Gleichung (mit Startwerten in  $(0, \infty)$ ) für  $t \rightarrow \infty$  gegen  $x^* = \frac{c}{d}$  konvergieren.

Während wir in Beispiel 8.3.2 mit Hilfe des Prinzips der linearisierten Stabilität also nur zeigen konnte, dass das Equilibrium  $\frac{c}{d}$  asymptotisch stabil ist, haben wir nun mit Hilfe einer Ljapunov-Funktion sogar gezeigt, dass dieses Equilibrium *global* attraktiv ist.<sup>31</sup>

**Literaturstellen und verwandte Ergebnisse.** In diesem Abschnitt haben wir nur beispielhaft zwei Sätze bewiesen, die das Langzeitverhalten der Lösungen von Differentialgleichungen mit Ljapunov-Funktionen in Beziehung setzen. Es gibt noch weitere Sätze von diesem Typ – siehe z.B. [PW19, Abschnitt 5.5 und Kapitel 8].

## 8.5 Ergänzungen

### Beweis von Satz 8.4.4

In diesem Unterabschnitt geben wir einen Beweis von Satz 8.4.4; in der Vorlesung hatten wir diesen Beweis ausgespart.

*Beweis von Satz 8.4.4.* Der Beweis wird besonders übersichtlich, wenn wir den Begriff des *Lösungsflusses* verwenden, den wir in Definition 4.5.1 eingeführt haben.<sup>32</sup> Wir bezeichnen den Lösungsfluss mit  $(\varphi_t)_{t \in [0, \infty)}$ , und wir müssen somit  $\varphi_t(x_0) \rightarrow x^*$  für alle  $x_0 \in \Omega$  zeigen. Sei also  $x_0 \in \Omega$  fest.

Wir zeigen nun zunächst die folgende Aussage:

(\*) Ist  $(t_k)_{k \in \mathbb{N}} \subseteq [0, \infty)$  eine monoton steigende, gegen  $\infty$  konvergente Folge, und konvergiert  $(\varphi_{t_k}(x_0))_{k \in \mathbb{N}}$  gegen einen Punkt  $x_{\text{lim}}$  in  $\Omega$ , so gilt  $x_{\text{lim}} = x^*$ .

Dazu bemerken wir zunächst, dass die Funktion  $[0, \infty) \ni t \mapsto V(\varphi_t(x_0)) \in \mathbb{R}$  monoton fallend und nach unten beschränkt ist, und somit für  $t \rightarrow \infty$  gegen eine Zahl  $\alpha \in \mathbb{R}$  konvergiert. Für jedes  $t \in [0, \infty)$  folgt nun

$$\alpha = \lim_{k \rightarrow \infty} V(\varphi_{t+t_k}(x_0)) = \lim_{k \rightarrow \infty} V(\varphi_t(\varphi_{t_k}(x_0)))$$

<sup>31</sup>Stabilität dieses Equilibriums kann man auch mit Hilfe der gegebenen Ljapunov-Funktion erhalten, nämlich indem man Satz 8.4.1 anwendet.

<sup>32</sup>Dies ist hier möglich, da alle Lösungen global nach rechts existieren.

$$= V\left(\varphi_t\left(\lim_{k \rightarrow \infty} \varphi_{t_k}(x_0)\right)\right) = V(\varphi_t(x_{\text{lim}})),$$

wobei wir für die vorletzte Gleichheit benutzt haben, dass die Abbildung  $\varphi_t : \Omega \rightarrow \Omega$  stetig ist (dies folgt aus Satz 7.2.4 über die stetige Abhängigkeit). Dies bedeutet aber, dass  $V$  entlang der Lösung  $[0, \infty) \ni t \mapsto \varphi_t(x_{\text{lim}}) \in \Omega$  konstant gleich  $\alpha$  ist. Da  $V$  eine strikte Ljapunov-Funktion ist, muss also die Lösung  $[0, \infty) \ni t \mapsto \varphi_t(x_{\text{lim}}) \in \Omega$  selbst konstant sein, d.h. es gilt  $\varphi_t(x_{\text{lim}}) = \varphi_0(x_{\text{lim}}) = x_{\text{lim}}$  für alle  $t \in [0, \infty)$ . Dies bedeutet gerade, dass  $x_{\text{lim}}$  ein Equilibrium ist.

Es gibt aber laut Voraussetzung nur ein Equilibrium, nämlich  $x^*$ . Also ist  $x_{\text{lim}} = x^*$ . Damit ist die Aussage (\*) gezeigt.

Nun können wir leicht die Behauptung zeigen: Aufgrund der Annahmen (a) und (b) liegt der komplette Orbit  $\{\varphi_t(x_0) : t \in [0, \infty)\}$  in einer kompakten Teilmenge von  $\Omega$ .

Wäre nun  $\varphi_t(x_0)$  für  $t \rightarrow \infty$  nicht gegen  $x^*$  konvergent, dann gäbe es eine monoton steigende gegen  $\infty$  konvergente Folge  $(t_k) \subseteq [0, \infty)$  derart, dass die Folge  $(\varphi_{t_k}(x_0))$  den Punkt  $x^*$  nicht als Häufungspunkt besitzt. Wir können aber eine Teilfolge  $(\varphi_{t_{k_j}}(x_0))$  wählen, die gegen einen Punkt in  $\Omega$  konvergiert, und laut (\*) ist der Grenzwert dieser Teilfolge gleich  $x^*$ ; Widerspruch.  $\square$

### $\omega$ -Limes-Mengen

Wenn man weiterführenden Sätze über das Langzeitverhalten der Lösungen von Differentialgleichungen beweisen will, ist die sogenannte  $\omega$ -Limesmenge – oder kurz: *Limesmenge* – ein wichtiges Konzept: Existiert die Lösung  $x$  einer Differentialgleichung global nach rechts, so nennt man die Menge

$$\{y \in \mathbb{R}^d : \text{es gibt eine gegen } \infty \text{ konvergente Folge } (t_k) \text{ mit } x(t_k) \rightarrow y\}$$

die  $\omega$ -Limesmenge der Lösung  $x$ . Diese Menge ist also – anschaulich gesprochen – die Menge der Häufungspunkte der Lösung für unendlich große Zeiten.

Mehr über dieses Konzept und seine Anwendungen können Sie zum Beispiel [PW19, ab Abschnitt 8.4] nachlesen.

### Ebene autonome Systeme und der Satz von Poincaré–Bendixon

Das Langzeitverhalten autonomer Differentialgleichungen in zwei Dimensionen – also *ebener autonomer Systeme* – besitzt eine besonders elegante Theorie.<sup>33</sup> Das wesentliche Resultat über solche Systeme ist der Satz von Poincaré–Bendixon<sup>34</sup>. Sie können diese Theorie bei Interesse zum Beispiel in [PW19, Kapitel 9] nachlesen.

<sup>33</sup>Was im Wesentlichen daran liegt, dass Kurven in der Ebene geometrisch betrachtet weniger Freiheiten besitzen als in höherer Dimension.

<sup>34</sup>Benannt nach Jules Henri Poincaré (1854 – 1912), französischer Mathematiker und theoretischer Physiker, und Ivar Otto Bendixson (1861 – 1935), schwedischer Mathematiker.

## Zwei Beispiele im Detail

### Fragen zum Einstieg.

- (a) Was haben Differentialgleichungen mit der Ausbreitung einer Epidemie zu tun?
- (b) Kann man die Wärmeleitung, die wir in Beispiel 1.3.1 diskutiert haben, auch mit einer gewöhnlichen anstelle einer partiellen Differentialgleichung beschreiben?

### 9.1 Ein Epidemie-Modell

In diesem Abschnitt besprechen wir ein Beispiel für ein Epidemie-Modell, das etwas komplexer ist als diejenigen Modelle, die wir bisher diskutiert haben. Der Übersicht halber unterteilen wir den Abschnitt in einige (nicht-nummerierte) Unterabschnitte.

#### Modellbeschreibung

Wir betrachten eine Infektionskrankheit, die sich über lange Zeit in einer Population ausbreitet. Dabei unterteilen wir die Population zum Zeitpunkt  $t$  in drei Klassen von Individuen:

- Mit  $S(t)$  bezeichnen wir die Anzahl der Individuen, die gesund und infizierbar sind.
- Mit  $I(t)$  bezeichnen wir die Anzahl der infizierten Individuen.
- Mit  $R(t)$  bezeichnen wir die Anzahl der immunen Individuen.

Wir treffen folgende Annahmen bezüglich der Verbreitung der Infektionskrankheit und der Entwicklung der Bevölkerung:

1. *Neuinfektionen*: Die Anzahl der Neuinfektionen ist proportional zum Produkt aus infizierbaren und infizierten Individuen; die Proportionalitätskonstante bezeichnen wir mit  $c > 0$ .
2. *Gesundung*: Infizierte Individuen werden mit einer Rate  $a > 0$  wieder gesund, und die Infektion verläuft nie tödlich.
3. *Dauerhafte Immunität*: Individuen, die die Krankheit überstanden haben, sind immun und bleiben dies für immer – d.h. es gibt keinen Immunitätsverlust.

4. *Geburts- und Sterbefälle:* Wir betrachten lange Zeiträume und berücksichtigen deshalb auch Geburten und Sterbefälle (welche laut Annahme 2. nicht durch die hier modellierte Infektion verursacht werden).
5. *Sterberate:* Wir nehmen dabei an, dass die Sterberate der Population durch eine Zahl  $b > 0$  gegeben ist, und dass die Sterberate für Infizierbare, Infizierte und Immune gleich ist.
6. *Geburtenrate:* Wir nehmen an, dass die Geburtenrate in der Bevölkerung gleich der Sterberate, also gleich  $b$  ist. Außerdem nehmen wir an, dass die Wahrscheinlichkeit sich fortzupflanzen nicht davon abhängt, ob ein Individuum infizierbar, infiziert oder immun ist.
7. *Gesunde Geburten und Impfungen:* Zuletzt nehmen wir an, dass jedes neu geborene Individuum gesund auf die Welt kommt (selbst wenn ein Elternteil oder beide Elternteile infiziert waren) und dass ein fester Anteil  $p \in [0, 1]$  der Neugeborenen direkt nach der Geburt geimpft wird und anschließend für immer immun ist.

Auch dieses Modell ist natürlich – obwohl es bereits viele Effekte berücksichtigt – weit davon entfernt, eine quantitativ realistische Beschreibung des genauen Verlaufs einer Krankheitsausbreitung zu liefern. Dennoch lohnt sich eine genauere Betrachtung, da das Modell einerseits mathematisch recht lehrreich ist, und da andererseits zumindest qualitativ einige Effekte im Modell auftreten, die man auch in der Realität beobachten kann.

Berücksichtigt man alle obigen Annahmen, so gelangt man zu folgendem Anfangswertproblem, das die Ausbreitung der Krankheit beschreibt:

$$\begin{cases} \begin{pmatrix} \dot{S}(t) \\ \dot{I}(t) \\ \dot{R}(t) \end{pmatrix} = \begin{pmatrix} b(1-p)(S(t)+I(t)+R(t)) - cS(t)I(t) - bS(t) \\ cS(t)I(t) - aI(t) - bI(t) \\ bp(S(t)+I(t)+R(t)) + aI(t) - bR(t) \end{pmatrix} \\ \begin{pmatrix} S(0) \\ I(0) \\ R(0) \end{pmatrix} = \begin{pmatrix} S_0 \\ I_0 \\ R_0 \end{pmatrix}, \end{cases} \quad (9.1)$$

wobei  $S_0, I_0, R_0 \geq 0$  reelle Zahlen sind. Wir betrachten die Differentialgleichung hier mathematisch zunächst auf dem Zeitintervall  $\mathbb{R}$  und dem Zustandsraum  $\mathbb{R}^3$ .<sup>1</sup>

---

<sup>1</sup>Wobei aus biologischer Sicht natürlich in erster Linie Zeiten  $t \geq 0$  und nur Vektoren im positiven Kegel  $\mathbb{R}_+^3$  eine Rolle spielen.

Um die rechte Seite in der Differentialgleichung (9.1) zu bezeichnen, definieren wir die Funktion

$$h: \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad (9.2)$$

$$\begin{pmatrix} s \\ i \\ r \end{pmatrix} \mapsto \begin{pmatrix} b(1-p)(s+i+r) - csi - bs \\ csi - ai - bi \\ bp(s+i+r) + ai - br \end{pmatrix}.$$

Viele unserer Resultate hatten wir für Situationen definiert, in denen die rechte Seite der Differentialgleichung eine Funktion  $f$  ist, die auch von der Zeit abhängen kann. Der notationellen Konsistenz wegen definieren wir deshalb auch noch die Funktion<sup>2</sup>

$$f: \mathbb{R}^{1+3} \rightarrow \mathbb{R}^3, \quad (9.3)$$

$$\begin{pmatrix} t \\ s \\ i \\ r \end{pmatrix} \mapsto h(s, i, r).$$

### Existenz und Eindeutigkeit von Lösungen

Wir interessieren uns zunächst dafür, ob das Anfangswertproblem (9.1) eine eindeutige Lösung besitzt. Dies lässt sich natürlich mit Hilfe der Resultate von Picard–Lindelöf untersuchen:

**Proposition 9.1.1.** *Die Abbildung  $f$  erfüllt die lokale Lipschitz-Bedingung.<sup>3</sup> Folglich besitzt das Anfangswertproblem (9.1) eine eindeutig bestimmte Lösung, die auf einem maximalen Existenzintervall  $I_{\max}$  definiert ist.*

*Beweis.* Die Funktionen  $f$  ist stetig differenzierbar, da jede Komponente von  $f$  nur eine Kombination von Produkten und Summen von konstanten Parametern und den Argumenten  $s, i, r$  ist; deshalb erfüllt  $f$  die lokale Lipschitz-Bedingung (Proposition 4.1.8). Also folgt die Existenz einer Lösung aus dem lokalen Existenzsatz von Picard–Lindelöf (Korollar 4.1.12), die Eindeutigkeit aus dem globalen Eindeutigkeitssatz von Picard–Lindelöf (Korollar 4.1.11) und Existenz des maximalen Existenzintervalls aus Proposition 4.3.2.  $\square$

### Positive Invarianz des positiven Kegels

Weil Werte für  $S, I$  und  $R$  nur dann biologisch Sinn ergeben, wenn sie  $\geq 0$  sind, interessieren wir uns dafür, ob der positive Kegel  $\mathbb{R}_+^3$  positiv invariant für die

<sup>2</sup>Die natürlich in Wirklichkeit gar nicht von  $t$  abhängt.

<sup>3</sup>Da unsere Differentialgleichung autonom ist, könnte man äquivalent dazu auch sagen: Die Abbildung  $h$  ist lokal Lipschitz-stetig.

Differentialgleichung in (9.1) ist. Dass dies tatsächlich der Fall ist, beweisen wir in der nächsten Proposition.

**Proposition 9.1.2.** *Der positive Kegel  $\mathbb{R}_+^3$  ist positiv invariant für die Differentialgleichung in (9.1).*

*Beweis.* Laut der Charakterisierung der positiven Invarianz von  $\mathbb{R}_+^3$ , die wir in Satz 7.3.7 bewiesen haben, genügt es zu zeigen, dass die Funktion  $f$  quasi-positiv ist. Weil  $f$  nicht von der Zeit abhängt, können wir stattdessen auch einfach die Funktion  $h$  betrachten.

Sei also  $(s, i, r) \in \mathbb{R}_+^3$ , sei  $k \in \{1, 2, 3\}$  und sei die  $k$ -te Komponente von  $(s, i, r)$  gleich 0. Wir müssen  $h_k(s, i, r) = 0$  zeigen, und dazu unterscheiden wir, welchen Wert der Index  $k$  hat:

- Sei  $k = 1$ . Dann ist  $s = 0$ , und es gilt somit

$$h_1(s, i, r) = b(1 - p)(i + r) \geq 0$$

wegen  $b > 0$ ,  $p \in [0, 1]$  und  $i, r \geq 0$ .

- Sei  $k = 2$ . Dann ist  $i = 0$ , und es gilt somit

$$h_2(s, i, r) = 0 \geq 0.$$

- Sei  $k = 3$ . Dann ist  $r = 0$ , und es gilt somit

$$h_3(s, i, r) = bp(s + i) + ai \geq 0$$

wegen  $b, a > 0$ ,  $p \geq 0$  und  $s, i \geq 0$ .

Wegen dieser Eigenschaften von  $h$  ist  $f$  tatsächlich quasi-positiv. □

### Erhaltung der Gesamtpopulation und Reduktion auf ein zwei-dimensionales System

Weil wir angenommen haben, dass die Geburts- und Sterberate gleich groß sind und dass niemand an der modellierten Krankheit verstirbt, sollten wir erwarten, dass die Gesamtpopulation – also die Zahl  $S(t) + I(t) + R(t)$  – konstant ist. Die folgende Proposition zeigt, dass dies tatsächlich der Fall ist.

**Proposition 9.1.3.** *Die Funktion  $N : \mathbb{R}^3 \rightarrow \mathbb{R}$ , die durch  $N(s, i, r) = s + i + r$  für alle  $(s, i, r) \in \mathbb{R}^3$  gegeben ist, ist eine Erhaltungsgröße für die Differentialgleichung in (9.1).*

*Beweis.* Weil  $N$  stetig differenzierbar ist, müssen wir laut der Charakterisierung von Erhaltungsgrößen, die wir in Proposition 6.3.3 bewiesen haben, nur nachrechnen, dass die orbitale Ableitung  $\langle \nabla N(s, i, r), h(s, i, r) \rangle$  für alle  $(s, i, r) \in \mathbb{R}^3$  verschwindet. Und tatsächlich gilt für jeden Vektor  $(s, i, r) \in \mathbb{R}^3$

$$\begin{aligned} \langle \nabla N(s, i, r), h(s, i, r) \rangle &= \left\langle \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} b(1-p)(s+i+r) - csi - bs \\ csi - ai - bi \\ bp(s+i+r) + ai - br \end{pmatrix} \right\rangle \\ &= b \underbrace{\left( (1-p) + p \right)}_{=b(s+i+r)} (s+i+r) - csi - bs + csi - ai - bi + ai - br = 0. \end{aligned}$$

Also folgt aus Proposition 6.3.3, dass  $N$  tatsächlich eine Erhaltungsgröße ist.  $\square$

Die Gesamtgröße der Population hängt in unserem Modell also nicht von der Zeit ab. Wir verwenden im Folgenden die Bezeichnung

$$N_0 := N(S_0, I_0, R_0),$$

d.h.  $N_0$  bezeichnet die Größe der Population zum Zeitpunkt 0 – und somit auch die Größe der Population zu jedem anderen Zeitpunkt. Die Differentialgleichung in unserem Anfangswertproblem (9.1) können wir somit umschreiben in die Differentialgleichung

$$\begin{pmatrix} \dot{S}(t) \\ \dot{I}(t) \end{pmatrix} = \begin{pmatrix} b(1-p)N_0 - cS(t)I(t) - bS(t) \\ cS(t)I(t) - aI(t) - bI(t) \end{pmatrix}. \quad (9.4)$$

Die Differentialgleichung für  $R$  können wir hierbei weglassen, da  $R(t)$  in den anderen beiden Gleichungen nicht mehr explizit vorkommt, und da  $R(t)$  sich zu jedem Zeitpunkt  $t$  mit Hilfe der Formel

$$R(t) = N_0 - S(t) - I(t)$$

aus den Größen  $S(t)$  und  $I(t)$  berechnen lässt.

Zur Bezeichnung der rechten Seite der Differentialgleichung (9.4) führen wir die Funktion

$$\begin{aligned} \tilde{h} : \mathbb{R}^2 &\rightarrow \mathbb{R}^2, \\ \begin{pmatrix} s \\ i \end{pmatrix} &\mapsto \begin{pmatrix} b(1-p)N_0 - csi - bs \\ csi - ai - bi \end{pmatrix} \end{aligned}$$

ein. Diese werden wir im Folgenden verwenden um das Langzeitverhalten der Lösungen von (9.4) zu untersuchen.

## Equilibria und Langzeitverhalten

Wir bestimmen zunächst die Equilibria der Differentialgleichung (9.4):

**Proposition 9.1.4.** Die Differentialgleichung (9.4) besitzt in  $\mathbb{R}^2$  genau zwei Equilibria:

(a) Das Krankheits-freie Equilibrium  $x^* := ((1-p)N_0, 0)$ .

(b) Das endemische<sup>4</sup> Equilibrium  $x^{**} := \left( \frac{a+b}{c}, \frac{b(1-p)N_0}{a+b} - \frac{b}{c} \right)$ .

*Beweis.* Das kann man sofort nachrechnen, wenn man  $\tilde{h}(s, i)$  gleich 0 setzt; man muss lediglich die beiden Fälle  $i = 0$  und  $i \neq 0$  unterscheiden.  $\square$

Das endemische Equilibrium ist natürlich nur dann biologisch relevant, wenn  $\frac{b(1-p)N_0}{a+b} - \frac{b}{c} \geq 0$  gilt, was wiederum äquivalent zur Ungleichung

$$c(1-p)N_0 \geq a+b$$

ist. Im Folgenden betrachten wir nun zwei Fälle:

- 1. Fall: Hohe Impfquote.<sup>5</sup> Hiermit meinen wir den Fall  $c(1-p)N_0 < a+b$ , in welchem das endemische Equilibrium nicht biologisch relevant ist. Diese Ungleichung ist äquivalent dazu, dass

$$1 - \frac{a+b}{cN_0} < p$$

gilt, d.h. sie ist ab einer gewissen Impfquote für die Neugeborenen erfüllt.

- 2. Fall: Niedrige Impfquote. Damit meinen wir den umgekehrten Fall  $c(1-p)N_0 > a+b$ , in welchem das endemische Equilibrium sogar im Inneren des positiven Kegels  $\mathbb{R}_+^2$  liegt, und welcher äquivalent zur oberen Abschätzung

$$1 - \frac{a+b}{cN_0} > p$$

für die Impfquote ist.

<sup>4</sup>Mit dem Begriff *Endemie* bezeichnet man die Situation, dass eine Krankheit dauerhaft in einer Population präsent ist, d.h. dass immer ein Teil der Bevölkerung mit der Krankheit infiziert ist.

<sup>5</sup>Die Begriffe *hohe* und *niedrige* Impfquote wollen wir an dieser Stelle nicht vage verwenden, sondern je im Sinne der konkreten Ungleichung, die wir in beiden Fällen angeben.

Den Fall, in welchem die Gleichheit  $c(1-p)N_0 = a + b$  gilt<sup>6</sup>, betrachten wir hier nicht.

Wir beginnen mit dem ersten Fall:

**Proposition 9.1.5 (Hohe Impfquote).** *Wenn  $c(1-p)N_0 < a + b$  gilt, dann liegt im positiven Kegel  $\mathbb{R}_+^2$  nur das Krankheits-freie Equilibrium  $x^*$ , und dieses ist asymptotisch stabil.*

*Beweis.* Offensichtlich liegt dies endemische Equilibrium in diesem Fall nicht in  $\mathbb{R}_+^2$ . Um zu zeigen, dass das Krankheits-freie Equilibrium asymptotisch stabil ist, verwenden wir linearisierte Stabilität: Wir alle  $(s, i) \in \mathbb{R}^2$  gilt

$$D\tilde{h}(s, i) = \begin{pmatrix} -ci - b & -cs \\ ci & cs - a - b \end{pmatrix};$$

somit ist insbesondere

$$D\tilde{h}(x^*) = \begin{pmatrix} -b & -c(1-p)N_0 \\ 0 & c(1-p)N_0 - (a + b) \end{pmatrix}.$$

Aufgrund der Voraussetzung  $c(1-p)N_0 < a + b$  ist der Eintrag dieser Matrix an der Stelle (2, 2) negativ, d.h. die Matrix  $D\tilde{h}(x^*)$  besitzt negative Spur und positive Determinante. Laut des zwei-dimensionalen Routh–Hurwitz-Kriteriums (Proposition 8.3.3) hat  $D\tilde{h}(x^*)$  somit negative Spektralschranke. Also ist das Equilibrium  $x^*$  aufgrund des Satzes 8.3.1 von der linearisierten Stabilität asymptotisch stabil.  $\square$

Nun betrachten wir den zweiten Fall.

**Proposition 9.1.6 (Niedrige Impfquote).** *Wenn  $c(1-p)N_0 > a + b$  gilt, dann liegen beide Equilibria  $x^*$  und  $x^{**}$  im positiven Kegel  $\mathbb{R}_+^2$ ; das Krankheits-freie Equilibrium  $x^*$  ist in diesem Fall instabil, und das endemische Equilibrium  $x^{**}$  ist asymptotisch stabil.*

*Beweis.* An der Formel für  $x^{**}$  erkennt man sofort, dass  $x^{**}$  unter der gegebenen Voraussetzung in  $\mathbb{R}_+^2$  (sogar im Inneren dieser Menge) liegt.

Wir zeigen nun zunächst die Instabilität des Krankheits-freien Equilibriums: Wir erhalten wieder

$$D\tilde{h}(x^*) = \begin{pmatrix} -b & -c(1-p)N_0 \\ 0 & c(1-p)N_0 - (a + b) \end{pmatrix},$$

<sup>6</sup>In diesem Fall liegt das endemische Equilibrium gerade auf dem Rand des positiven Kegels, was Null Infizierte bedeutet – es handelt sich hierbei also nicht wirklich um einen endemischen Zustand.

aber dieses Mal ist der Eintrag an der Stelle  $(2, 2)$  positiv. Somit besitzt die Matrix  $D\tilde{h}(x^*)$  negative Determinante, d.h. beide Eigenwerte der Matrix sind reell und einer der beiden Eigenwerte muss positiv sein. Aus Satz 8.3.5 über die linearisierte Instabilität folgt somit, dass das Equilibrium  $x^*$  instabil ist.

Nun zeigen wir die asymptotische Stabilität des endemischen Equilibriums  $x^{**}$ . Es gilt

$$D\tilde{h}(x^{**}) = \begin{pmatrix} -b \frac{c(1-p)N_0}{a+b} & -(a+b) \\ b \left( \frac{c(1-p)N_0}{a+b} - 1 \right) & 0 \end{pmatrix}.$$

Aufgrund der Voraussetzung ist der Eintrag an der Stelle  $(2, 1)$  positiv, d.h. die Matrix  $D\tilde{h}(x^{**})$  besitzt negative Spur und positive Determinante. Aufgrund des zwei-dimensionalen Routh–Hurwitz-Kriterium (Proposition 8.3.3) hat  $D\tilde{h}(x^{**})$  also negative Spektralschranke. Somit ist das Equilibrium  $x^{**}$  laut des Satzes 8.3.1 von der linearisierten Stabilität asymptotisch stabil.  $\square$

- Literaturstellen und verwandte Ergebnisse.** (a) Die obige Darstellung orientiert sich grob an [PSZ08, Kapitel 11]. Dort finden Sie auch noch ein weitergehendes Resultat zum Lanzeitverhalten im Falle einer hohen Impfquote: Das Krankheits-freie Gleichgewicht ist nämlich nicht nur asymptotisch stabil, besitzt sogar eine gewisse globale Attraktivitätseigenschaft.
- (b) Weitere Beispiele für Differentialgleichungen, die epidemische Modelle beschreiben, finden Sie unter anderen in den Kapiteln 7–10 und 11 von [PSZ08].

## 9.2 Endliche Approximation der Wärmeleitungsgleichung auf einem Intervall

In diesem letzten Abschnitt der Vorlesung kommen wir noch einmal auf ein Beispiel einer *partiellen* Differentialgleichung zurück, das wir zu Beginn der Vorlesung kurz angesprochen hatten: In Beispiel 1.3.1 haben wir uns die *Wärmeleitungsgleichung*

$$\frac{\partial}{\partial t} u(t, z) = c \frac{\partial^2}{\partial z^2} u(t, z) \quad (9.5)$$

angesehen. Hier ist die Zeit  $t$  eine reelle Zahl im Intervall  $[0, \infty)$ , der Ort  $z$  eine reelle Zahl aus dem Intervall  $[0, L]$  (für ein  $L > 0$ ), und  $u : [0, \infty) \times [0, L] \rightarrow \mathbb{R}$  die gesuchte Funktion. Die Zahl  $c > 0$  ist eine Konstante.

In diesem kurzen Abschnitt wollen wir darüber sprechen, wie man diese partielle Differentialgleichung mit einer einfachen Methode durch eine gewöhnliche Differentialgleichung approximieren kann. Wir unterteilen den Abschnitt wieder in nicht-nummerierte Unterabschnitte.

### Anfangs- und Randwerte

Wie auch bei einer gewöhnlichen Differentialgleichung benötigt man bei der Wärmeleitungsgleichung (9.5) einen Anfangswert um Eindeutigkeit der Lösung zu erhalten – man muss also den Wert  $u(0, z)$  für alle  $z \in [0, L]$  vorgeben.<sup>7</sup>

Weil bei partiellen Differentialgleichungen aber Ableitungen nach mehreren Variablen vorkommen, genügt auch die Festlegung von  $u(0, z)$  für alle  $z \in [0, L]$  noch nicht um Eindeutigkeit der Lösung zu erhalten. Stattdessen benötigt man auch noch sogenannte *Randwerte* – d.h. man muss zum Beispiel  $u(t, 0)$  und  $u(t, L)$  für alle Zeiten  $t \geq 0$  vorgeben. Der Einfachheit halber wollen wir hier durchgehend die Randbedingungen

$$u(t, 0) = 0 \quad \text{und} \quad u(t, L) = 0 \quad \text{für alle } t \in [0, \infty)$$

verwenden; man bezeichnet diese Randbedingungen als *homogene Dirichlet-Randbedingungen*.

In Beispiel 1.3.1 hatten wir angesprochen, dass die Wärmeleitungsgleichung physikalisch die zeitliche Entwicklung der Temperatur in einem langen dünnen Stab beschreibt. Die homogenen Dirichlet-Randbedingungen bedeuten physikalisch, dass die beiden Enden des Stabs (an den Positionen 0 und  $L$ ) durchgehend auf der Temperatur 0 gehalten werden.<sup>8</sup>

### Finite-Differenzen-Approximation

Wir wollen nun aus der partiellen Differentialgleichung (9.5) eine gewöhnliche machen, indem wir den Stab  $L$  durch endlich viele Punkte *diskretisieren* und die zweite räumliche Ableitung  $\frac{\partial^2}{\partial z^2} u(t, z)$  durch Differenzenquotienten annähern. Dieses Annähern von Ableitungen durch Differenzenquotienten zur approximativen Behandlung von Differentialgleichungen bezeichnet man, insbesondere in der Numerik, als *Finite-Differenzen-Methode*. Zur numerischen Lösung könnte man zum Beispiel auch noch die Zeit diskretisieren und

---

<sup>7</sup>Technisch betrachtet handelt es sich bei dem „Anfangswert“ natürlich nicht nur um einen oder um endliche viele Werte, sondern um eine *Anfangsfunktion* – funktionalanalytisch gesprochen bedeutet dies, dass der Anfangswert ein Vektor aus einem geeigneten Funktionenraum ist.

<sup>8</sup>Die Zahl 0 muss hier aber nicht unbedingt den absoluten Temperatur-Nullpunkt bezeichnen, sondern kann zum Beispiel den Nullpunkt in einer gewählten Temperaturskala bezeichnen.

die zeitliche Ableitung ebenfalls durch Differenzenquotienten annähern; dies werden wir hier jedoch nicht tun.<sup>9</sup>

Wir fixieren nun also eine natürliche Zahl  $d$  und fügen  $d$  äquidistante Punkte

$$z_1 := \frac{1}{d+1}L, \quad \dots, \quad z_d := \frac{d}{d+1}L$$

innerhalb des Intervalls  $[0, L]$  ein. Zudem definieren wir die beiden Randpunkte  $z_0 := 0$  und  $z_{d+1} := L$ .

Zu jedem Zeitpunkt  $t \in [0, \infty)$  werten wir nun die gesuchte Temperaturverteilung  $u(t, \cdot)$  an den Punkten  $z_0, \dots, z_{d+1}$  aus und erhalten somit Zahlen

$$\begin{aligned} u_0(t) &:= u(t, z_0), \\ u_1(t) &:= u(t, z_1), \\ &\vdots \\ u_d(t) &:= u(t, z_d), \\ u_{d+1}(t) &:= u(t, z_{d+1}). \end{aligned}$$

Die homogenen Dirichlet-Randbedingungen, die wir im vorangehenden Unterabschnitt besprochen haben, implizieren, dass

$$u_0(t) = 0 \quad \text{und} \quad u_{d+1}(t) = 0$$

für alle Zeiten  $t$  gilt.

Als nächstes approximieren wir die zweite räumliche Ableitung von  $u(t, \cdot)$  an den Punkten  $z_1, \dots, z_d$  an. Sei also  $k \in \{1, \dots, d\}$ . Die zweite räumliche Ableitung von  $u(t, \cdot)$  an der Stelle  $z_k$  kann man durch den Differenzenquotienten

$$\begin{aligned} \frac{\partial^2}{\partial z^2} u(t, z_k) &\approx \frac{d^2}{L^2} \left( u(t, z_{k-1}) - 2u(t, z_k) + u(t, z_{k+1}) \right) \\ &= \frac{d^2}{L^2} \left( u_{k-1}(t) - 2u_k(t) + u_{k+1}(t) \right) \end{aligned}$$

annähern. Die Wärmeleitungsgleichung (9.5) besagt, wenn wir diese Näherung verwenden, gerade, dass

$$\dot{u}_k(t) = c \frac{d^2}{L^2} \left( u_{k-1}(t) - 2u_k(t) + u_{k+1}(t) \right)$$

<sup>9</sup>Bitte beachten Sie aber unbedingt, dass das soeben angedeutete Verfahren zur numerischen Lösung einer partiellen Differentialgleichung stark simplifiziert dargestellt ist. Wenn man partielle Differentialgleichungen tatsächlich numerisch lösen möchte, muss man eine Vielzahl von Subtilitäten beachten. Zudem gibt es neben der Finite-Differenzen-Methode auch noch anderen Verfahren wie zum Beispiel die *Finite-Elemente-Methode*, die nicht direkt auf einer Approximation der Ableitungen beruht, sondern darauf, die Differentialgleichung zu *Testen*, indem man sie gegen geeignete Funktionen integriert.

für alle  $t \in [0, \infty)$  und alle  $k \in \{1, \dots, d\}$  gilt.<sup>10</sup>

Für  $k = 1$  und  $k = d$  können wir obige Differentialgleichung noch etwas vereinfachen, denn die dort auftretenden Terme  $u_0(t)$  und  $u_{d+1}(t)$  sind ja gleich 0. Deshalb erfüllen  $\dot{u}_1(t)$  und  $\dot{u}_d(t)$  die etwas vereinfachten Differentialgleichungen

$$\begin{aligned}\dot{u}_1(t) &= c \frac{d^2}{L^2} (-2u_1(t) + u_2(t)), \\ \dot{u}_d(t) &= c \frac{d^2}{L^2} (u_{d-1}(t) - 2u_d(t)).\end{aligned}$$

Insgesamt erhalten wir also die Differentialgleichung

$$\begin{pmatrix} \dot{u}_1(t) \\ \dot{u}_2(t) \\ \vdots \\ \dot{u}_{d-1}(t) \\ \dot{u}_d(t) \end{pmatrix} = c \frac{d^2}{L^2} \underbrace{\begin{pmatrix} -2 & 1 & & & \\ 1 & -2 & 1 & & \\ & & \ddots & & \\ & & & 1 & -2 & 1 \\ & & & & 1 & -2 \end{pmatrix}}_{=:A} \underbrace{\begin{pmatrix} u_1(t) \\ u_2(t) \\ \vdots \\ u_{d-1}(t) \\ u_d(t) \end{pmatrix}}_{=: \tilde{u}(t)}. \quad (9.6)$$

Kurz geschrieben lautet diese Differentialgleichung also

$$\dot{\tilde{u}}(t) = A \tilde{u}(t)$$

mit der soeben definierten Matrix  $A \in \mathbb{R}^{d \times d}$  und der Funktion  $\tilde{u} : [0, \infty) \rightarrow \mathbb{R}^d$ . Als Anfangsbedingung haben wir natürlich

$$\tilde{u}(0) = \begin{pmatrix} u_1(0) \\ \vdots \\ u_d(0) \end{pmatrix} = \underbrace{\begin{pmatrix} u(0, z_1) \\ \vdots \\ u(0, z_d) \end{pmatrix}}_{=: \tilde{u}_0};$$

der rechts stehende Vektor ist hierbei bekannt, da wir ja als Anfangsbedingung für die partielle Differentialgleichung (9.5) die Werte  $u(0, z)$  für alle  $z \in [0, L]$  kennen müssen.

### Eigenschaften der gewöhnlichen Differentialgleichung

Nachdem wir im vorangehenden Unterabschnitt geklärt haben, wie man die Wärmeleitungsgleichung durch die gewöhnliche Differentialgleichung (9.6) approximieren kann, wollen wir nun noch ein paar Worte dazu sagen, welche

<sup>10</sup>Für  $u_0(t)$  und  $u_{d+1}(t)$  brauchen wir keine Differentialgleichung anzugeben, denn diese Werte müssen ja aufgrund der Randbedingungen ohnehin zu jedem Zeitpunkt gleich 0 sein.

Eigenschaften die Gleichung (9.6) sowie ihre Lösungen haben. Wir tun dies in der folgenden Proposition; damit wiederholen wir zugleich einige Resultate, die wir im Laufe der Vorlesung über lineare Differentialgleichungen besprochen haben.

**Proposition 9.2.1.** *Wir betrachten die Differentialgleichung (9.6) mit Zustandsraum  $\mathbb{R}^d$  und auf dem Zeitintervall  $\mathbb{R}$ .<sup>11</sup> Dann gilt:*

- (a) *Für jeden Startwert  $\tilde{u}_0 \in \mathbb{R}^d$  existiert die Lösung  $\tilde{u}$  von (9.6) mit Anfangsbedingung  $\tilde{u}(0) = \tilde{u}_0$  global – d.h. das maximale Existenzintervall ist gleich  $\mathbb{R}$ .*
- (b) *Die in (a) erwähnte Lösung ist gegeben durch  $\tilde{u}(t) = e^{tA}\tilde{u}_0$  für alle  $t \in \mathbb{R}$ .*
- (c) *Liegt  $\tilde{u}_0$  im positiven Kegel  $\mathbb{R}_+^d$ , so liegt auch  $\tilde{u}(t)$  für jeden Zeitpunkt  $t \geq 0$  im positiven Kegel  $\mathbb{R}_+^d$ .*
- (d) *Es gilt  $\|e^{tA}\| \rightarrow 0$  für  $t \rightarrow \infty$ .*

*Beweis.* (a) Die globale Existenz der Lösungen für lineare Differentialgleichungen haben wir in Proposition 5.1.5 bewiesen.

(b) Laut Korollar 5.4.7 ist die Abbildung

$$\mathbb{R} \ni t \mapsto e^{tA} \in \mathbb{R}^{d \times d}$$

die Hauptfundamentalmatrix der Differentialgleichung (9.6) zum Startzeitpunkt 0. Die behauptete Formel für die Lösung  $\tilde{u}$  folgt somit auf Proposition 5.3.5(d).

(c) Weil alle nicht-diagonal-Einträge der Matrix  $A$  größer oder gleich 0 sind, folgt diese Behauptung aus Beispiel 7.3.10.

(d) Wir zeigen, dass die Spektralschranke von  $A$  negativ ist: Weil  $A$  reell und symmetrisch und somit selbst-adjungiert ist, sind alle Eigenwerte von  $A$  reell. Außerdem kann man zum Beispiel mit Hilfe der Gerschgorin-Kreise erkennen, dass jeder Eigenwert von  $A$  höchstens die Entfernung  $2c\frac{d^2}{L^2}$  vom Punkt  $-2c\frac{d^2}{L^2}$  hat – also liegen alle Eigenwerte im Intervall  $[-4c\frac{d^2}{L^2}, 0]$ . Zuletzt kann man leicht nachrechnen, dass die Gleichung  $Ay = 0$  nur die Lösung  $y = 0$  besitzt – und somit ist 0 kein Eigenwert von  $A$ . Also ist tatsächlich  $s(A) < 0$ . Daher folgt die Behauptung aus der Äquivalenz der Aussagen (ii) und (iv) in Satz 8.1.1.  $\square$

<sup>11</sup>Für diese lineare gewöhnliche Differentialgleichung spricht nicht wirklich etwas dagegen, ganz  $\mathbb{R}$  als Zeitintervall zu betrachten. Physikalisch sind aber vor allem Zeiten  $t \geq 0$  interessant – und wenn man anstelle der approximierenden gewöhnlichen Differentialgleichung (9.6) die partielle Differentialgleichung (9.5) untersucht, stellt man fest, dass negative Zeiten hierfür erhebliche Probleme bereiten. Dies ist aber ein Phänomen aus der Theorie der (sogenannten *parabolischen*) partiellen Differentialgleichungen, auf das wir hier nicht weiter eingehen.

# Appendices





## Metrische und normierte Räume

### Fragen zum Einstieg.

- (a) Können Sie sich verschiedene sinnvolle Möglichkeiten vorstellen, den *Abstand* zwischen zwei Städten (sagen wir der Einfachheit halber: innerhalb Deutschlands) zu definieren?
- (b) Nehmen Sie sich einen Globus und suchen Sie die beiden Städte Berlin und Vancouver. Können Sie die geographisch kürzest mögliche Flugroute erkennen, die ein Flugzeug von Berlin nach Vancouver nehmen kann (unabhängig von Erwägungen wie Windrichtung oder Sicherheit)? Falls ja, woran erkennen Sie diese Flugroute?
- (c) Skizzieren Sie die drei Funktionen  $f, g, h: [0, 1] \rightarrow \mathbb{R}$ , die durch

$$f(t) = 1, \quad g(t) = 1 - t^5, \quad \text{und} \quad h(t) = 1 + \frac{1}{10} \sin(100\pi t)$$

für alle  $t \in [0, 1]$  gegeben sind.

Finden Sie, dass die Funktionen  $f$  und  $g$  nahe beisammen liegen? Wie sieht es mit den Funktionen  $f$  und  $h$  aus?

Haben Sie allgemein eine Idee (oder sogar mehrere Ideen), wie man den Abstand zweier Funktionen definieren kann?

- (d) Wann nochmal heißt eine Menge  $M \subseteq \mathbb{R}^d$  *kompakt*? Versuchen Sie sich an die Analysis 2 zu erinnern: Was bringt es Ihnen, wenn Sie wissen, dass eine Menge kompakt ist?

### A.1 Metrische Räume: Grundbegriffe

Wir wiederholen in diesem Anhang einige Grundbegriff zu *metrischen Räumen*, die Sie aus der Analysis 2 kennen.

**Definition A.1.1 (Metrik und metrischer Raum).** Sei  $M$  eine Menge. Eine Abbildung  $d: M \times M \rightarrow [0, \infty)$  heißt eine *Metrik* auf  $M$ , falls Sie die folgenden Axiome erfüllt:

- (i) *Positivite Definitheit:* Für alle  $x, y \in M$  gilt  $d(x, y) = 0$  genau dann, wenn  $x = y$  ist.
- (ii) *Symmetrie:* Für alle  $x, y \in M$  gilt  $d(x, y) = d(y, x)$ .
- (iii) *Dreiecks-Ungleichung:* Für alle  $x, y, z \in M$  gilt  $d(x, z) \leq d(x, y) + d(y, z)$ .

Ein *metrischer Raum* ist ein Paar  $(M, d)$ , wobei  $M$  eine Menge und  $d$  eine Metrik auf  $M$  ist.

Anschaulich ist eine Metrik ein Abstandsmaß. Die Dreiecksungleichung besagt, dass der Abstand von einem Punkt  $x$  zu einem Punkt  $z$  höchstens so groß ist, wie der Abstand von  $x$  zu einem beliebigen dritten Punkt  $y$  plus den Abstand von diesem  $y$  zu  $z$ .

**Beispiele A.1.2 (Einige metrische Räume).** (a) Sie  $M \subseteq \mathbb{R}^d$  für eine feste Dimension  $d \in \mathbb{N}$  und bezeichne  $\|\cdot\|$  die Euklidische Norm auf  $\mathbb{R}^d$ . Dann ist durch

$$d(x, y) := \|x - y\| \quad \text{für } x, y \in M$$

eine Metrik auf  $M$  definiert, die sogenannte *Euklidische Metrik*. Im Falle  $d = 2$  beschreibt diese Metrik den üblichen geometrischen Abstandsbegriff in der Ebene, und im Falle  $d = 3$  den üblichen geometrischen Abstandsbegriff im drei-dimensionalen Raum.

(b) Für alle  $x, y \in \mathbb{Q}$  sei  $d(x, y) = |x - y|$ . Dann ist  $d$  eine Metrik auf  $\mathbb{Q}$ . Dies ist ein Spezialfall von Beispiel (a).

(c) Sei  $S$  eine Menge von endlichen vielen Städten in einem Land, die durch ein Straßennetz verbunden sind. Wir nehmen der Einfachheit halber an, dass alle Straßen in beide Richtungen gleich schnell durchfahren werden können, und dass die Zeit zum Abfahren einer Straße nicht von Uhrzeit, Verkehrslage oder gewähltem Fahrzeug abhängt (äußerst realistisch!).

Für je zwei Städte  $s, t \in S$  bezeichne  $d(s, t)$  die kürzeste Zeit, in der es möglich ist von  $s$  nach  $t$  zu gelangen. Dann ist  $d$  eine Metrik auf  $S$ .

(d) Betrachten Sie eine Menge  $F = \{p, s_1, \dots, s_n\}$  mit paarweise verschiedenen Objekten  $p, s_1, \dots, s_n$  (und  $n \in \mathbb{N}$ ). Für jedes  $k \in \{1, \dots, n\}$  sei eine reelle Zahl  $r_k > 0$  gegeben. Wir definieren eine Abbildung  $d : F \times F \rightarrow [0, \infty)$  folgendermaßen: Es sei

$$d(p, s_k) = d(s_k, p) = r_k$$

für alle  $k \in \{1, \dots, n\}$ ,

$$d(s_j, s_k) = r_j + r_k$$

für alle  $j, k \in \{1, \dots, n\}$  mit  $j \neq k$ , und  $d(t, t) = 0$  für alle  $t \in F$ .

Dann ist  $d$  eine Metrik auf  $F$ , die sogenannten *französische Eisenbahn-Metrik* oder auch *SNCF-Metrik*.

(Leider wird dieser Witz mit jedem weiteren Ausbau des französischen TGV-Netzes weniger lustig.)

(e) Es bezeichne  $E$  die Oberfläche einer Kugel im  $\mathbb{R}^3$  (zum Beispiel der Erde). Für zwei Punkte  $p, q$  definieren wir  $d(p, q)$  folgendermaßen:

- Wenn  $p = q$  ist, sei  $d(p, q) = 0$ .
- Wenn  $p \neq q$  ist, bezeichne  $G$  den eindeutig bestimmten Großkreis, der durch  $p$  und  $q$  verläuft. Auf diesem Großkreis gibt es zwei Verbindungen zwischen  $p$  und  $q$ . Wir definieren  $d(p, q)$  als die Länge der kürzeren dieser beiden Verbindungen (bzw. als Länge einer der beiden Verbindungen, wenn beide gleich lang sind).

Dann ist  $d$  eine Metrik auf  $E$ .

**Aufgabe A.1.3.** (a) Veranschaulichen Sie die französische Eisenbahn-Metrik anhand eines ebenen Graphen.

(b) Was hat Beispiel A.1.2(e) mit der Einstiegsfrage (b) dieses Kapitels zu tun?

Ein wichtiges Prinzip in der Mathematik ist es, aus gegebenen Objekten neue zu konstruieren. Ein einfaches Beispiel für dieses Prinzip ist in der folgenden Bemerkung erläutert:

**Bemerkung A.1.4.** Sei  $(M, d)$  ein metrischer Raum und sei  $N \subseteq M$ . Es bezeichne  $d_N$  die *Einschränkung* der Abbildung  $d : M \times M \rightarrow [0, \infty)$  auf die Teilmenge  $N \times N$  von  $M \times M$ , d.h. es sei  $d_N : N \times N \rightarrow [0, \infty)$  durch

$$d_N(x, y) = d(x, y)$$

für alle  $x, y \in N$  gegeben. Dann ist  $d_N$  eine Metrik auf  $N$ ; man bezeichnet Sie oft als die *von  $d$  induzierte Metrik* auf  $N$ .

Kurzum: Aus jeder Teilmenge eines metrischen Raumes  $(M, d)$  kann man durch Einschränkung der Metrik  $d$  selbst einen metrischen Raum machen.

## A.2 Konvergenz und Vollständigkeit

Auf metrischen Räumen kann man über Konvergenz von Folgen sprechen. Wir wiederholen diesen Begriff als nächstes.

**Definition A.2.1 (Konvergente Folgen und Cauchy-Folgen).** Sei  $(M, d)$  ein metrischer Raum und sei  $(x_n)_{n \in \mathbb{N}}$  eine Folge von Elementen von  $M$ .

(i) Sei  $x \in M$ . Die Folge  $(x_n)_{n \in \mathbb{N}}$  heißt *konvergent gegen  $x$*  (oder in anderen Worten: die Folge *konvergiert gegen  $x$* ), falls gilt:

Für jede reelle Zahl  $\varepsilon > 0$  gibt es einen Index  $n_0 \in \mathbb{N}$  derart, dass  $d(x_n, x) < \varepsilon$  für alle nachfolgenden Indizes  $n \geq n_0$  gilt.

Wir verwenden die Notation „ $x_n \rightarrow x$  für  $n \rightarrow \infty$ “ oder „ $\lim_{n \rightarrow \infty} x_n = x$ “ um zum Ausdruck zu bringen, dass die Folge  $(x_n)_{n \in \mathbb{N}}$  gegen  $x$  konvergiert.

(ii) Die Folge  $(x_n)_{n \in \mathbb{N}}$  heißt *konvergent* (oder in anderen Worten: die Folge *konvergiert*), wenn es ein  $x \in M$  gibt derart, dass  $(x_n)_{n \in \mathbb{N}}$  gegen  $x$  konvergiert.

(iii) Man nennt die Folge  $(x_n)_{n \in \mathbb{N}}$  eine *Cauchy-Folge*, falls gilt:

Für jedes reelle Zahl  $\varepsilon > 0$  gibt es einen Index  $n_0 \in \mathbb{N}$  derart, dass  $d(x_m, x_n) < \varepsilon$  für alle nachfolgenden Indizes  $m, n \geq n_0$  gilt.

Die folgenden einfachen Aufgaben können Ihnen bei der Wiederholung der oben besprochenen Begriffe helfen.

**Aufgaben A.2.2.** Sei  $(M, d)$  ein metrischer Raum. Zeigen Sie:

(a) Jede konvergente Folge in  $(M, d)$  ist eine Cauchy-Folge.

(b) Eine Folge  $(x_n)_{n \in \mathbb{N}}$  in  $M$  ist genau dann eine Cauchy-Folge, wenn gilt:

Für jede reelle Zahl  $\varepsilon > 0$  gibt es einen Index  $n_0 \in \mathbb{N}$  derart, dass  $d(x_n, x_{n_0}) < \varepsilon$  für jeden nachfolgenden Index  $n \geq n_0$  gilt.

(c) Sei  $M$  endlich. Dann ist eine Folge  $(x_n)_{n \in \mathbb{N}}$  in  $M$  genau dann konvergent, wenn sie schließlich konstant ist.

Während eine konvergente Folge stets eine Cauchy-Folge ist, gibt es metrische Räume, in denen die umgekehrte Implikation nicht gilt – zum Beispiel den Raum aus Beispiel A.1.2(b):

**Aufgabe A.2.3.** Wir versehen  $\mathbb{Q}$  mit der Metrik  $d$  aus Beispiel A.1.2. Finden Sie eine Cauchy-Folge in  $\mathbb{Q}$ , die nicht konvergiert.

Metrische Räume, in denen jede Cauchy-Folge konvergiert, spielen eine enorm wichtige Rolle in der Analysis. Deshalb haben Sie einen eigenen Namen:

**Definition A.2.4 (Vollständiger metrischer Raum).** Ein metrischer Raum  $(M, d)$  heißt *vollständig*, wenn jede Cauchy-Folge in  $M$  konvergiert.

Um zu erkennen, dass Vollständigkeit ein wichtiger Begriff ist, können Sie zum Beispiel einen Blick auf den *Banachschen Fixpunktsatz* in Abschnitt B.2 werden: Dieser Satz gilt nur auf metrischen Räumen, die vollständig sind! Zugleich spielt der Satz eine zentrale Rolle für das Studium gewöhnlicher Differentialgleichungen, weil man ihn (bzw. eine Variante davon, vgl. Satz 4.1.4) zum Beweis des Satzes von Picard–Lindelöf benötigt.

Zum Abschluss dieses Unterabschnitts noch eine äußerst spaßige Aufgabe – Sie sollten diese Aufgabe aber nur bearbeiten, wenn Sie das Gefühl haben, bereits recht sicher mit metrischen Räumen und den Begriffen *Konvergenz* und *Cauchy-Folge* umgehen zu können.

**Aufgabe A.2.5.** Sei  $(x_n)_{n \in \mathbb{N}}$  eine Folge in einem metrischen Raum  $(M, d)$ . Die Aussage, dass die Folge  $(x_n)_{n \in \mathbb{N}}$  konvergiert, kann man in Quantor-Schreibweise folgendermaßen formulieren:

$$\exists x \in M \forall \varepsilon > 0 \exists n_0 \in \mathbb{N} \forall n \geq n_0 : d(x_n, x) < \varepsilon.$$

Sei nun  $(M, d)$  vollständig. Zeigen Sie, dass  $(x_n)_{n \in \mathbb{N}}$  genau dann konvergiert, wenn die – formal schwächere – Aussage

$$\forall \varepsilon > 0 \exists x \in M \exists n_0 \in \mathbb{N} \forall n \geq n_0 : d(x_n, x) < \varepsilon$$

gilt.

### A.3 Topologische Begriffe und Stetigkeit

Auf metrischen Räumen lassen sich topologische Begriffe wie z.B. *offene Menge* und *abgeschlossene Menge* definieren, und es gibt verschiedene (aber äquivalente) Möglichkeiten dies zu tun. Wir gehen hier nicht sehr weit in die Tiefe, sondern begnügen uns damit kurz zu wiederholen, dass abgeschlossene Mengen sich zum Beispiel mit Hilfe von Folgenkonvergenz beschreiben lassen:

**Definition A.3.1 (Abgeschlossene und offene Mengen).** Sei  $(M, d)$  ein metrischer Raum.

- (i) Eine Teilmenge  $A \subseteq M$  heißt *abgeschlossen*, wenn für jede konvergente Folge  $(x_n)_{n \in \mathbb{N}}$  in  $M$ , deren Glieder  $x_n$  alle in  $A$  liegen, auch der Grenzwert in  $A$  liegt.
- (ii) Eine Teilmenge  $U \subseteq M$  heißt *offen*, wenn ihr Komplement  $M \setminus U$  abgeschlossen ist.

Aus topologischer Sicht ist die obige Definition etwas schräg: Üblicherweise geht man in der Topologie zunächst vom Begriff der offenen Menge aus und definiert dann eine Menge als abgeschlossen, wenn ihr Komplement offen ist. Obige Definition ist hierzu allerdings äquivalent.

**Werbung (Topologie-Vorlesung).** Im Wintersemester 2020/21 werde ich eine Vorlesung über *Topologie* anbieten. Hörerinnen und Hörer sind herzlich willkommen!

Die folgende Standardausage können Sie zu Ihrer Übung beweisen.

**Aufgabe A.3.2.** Sei  $(M, d)$  ein vollständiger metrischer Raum und sei  $N \subseteq M$  mit der Metrik  $d_N$  versehen, die von  $d$  induziert wird (vgl. Bemerkung A.1.4).

Zeigen Sie: Die Menge  $N$  ist genau dann abgeschlossen (im metrischen Raum  $M$ ), wenn der metrische Raum  $(N, d_N)$  vollständig ist.

Zuletzt gehen wir noch auf den – äußerst wichtigen – Begriff der *Stetigkeit* einer Abbildung zwischen zwei metrischen Räumen ein.

**Definition A.3.3 (Stetigkeit).** Seien  $(L, d_L)$  und  $(M, d_M)$  metrische Räume und sei  $f : L \rightarrow M$  eine Abbildung.

- (i) Sei  $x \in L$ . Die Abbildung  $f$  heißt *stetig in  $x$* , wenn für jede gegen  $x$  konvergente Folge  $(x_n)_{n \in \mathbb{N}}$  in  $L$  gilt, dass die Folge  $(f(x_n))_{n \in \mathbb{N}}$  in  $M$  gegen  $f(x)$  konvergiert.
- (ii) Die Abbildung  $f$  heißt *stetig*, wenn sie in jedem Punkt  $x \in L$  stetig ist.

Auch die Wahl der obigen Definition von Stetigkeit für diesen Anhang ist eher ihrer Kürze denn ihrer topologischen Systematik geschuldet. In der Topologie würde man eher eine (äquivalente) Definition der Stetigkeit mit Hilfe von sogenannten *Umgebungen* wählen.

Hier ist eine Aufgabe um den Umgang mit Stetigkeit zu üben:

**Aufgabe A.3.4.** Seien  $(L, d_L)$ ,  $(M, d_M)$  und  $(N, d_N)$  metrische Räume und seien  $f : L \rightarrow M$  und  $g : M \rightarrow N$  stetige Abbildungen. Zeigen Sie, dass  $g \circ f : L \rightarrow N$  ebenfalls stetig ist.

## A.4 Normierte Räume

Eine besondere Klasse metrischer Räume sind die *normierten Räume*. Dies sind Vektorräumen über  $\mathbb{R}$  oder  $\mathbb{C}$ , welche mit einer *Norm* ausgestattet sind, die wiederum eine Metrik induziert. Es folgt die präzise Definition:

**Definition A.4.1 (Normierter Raum).** Sei  $V$  ein Vektorraum über  $\mathbb{R}$  oder  $\mathbb{C}$ . Eine Abbildung  $\|\cdot\| : V \rightarrow [0, \infty)$  heißt *Norm* auf  $V$ , wenn sie die folgenden Axiome erfüllt:

- (i) *Positive Definitheit:* Für jedes  $v \in V$  gilt  $\|v\| = 0$  genau dann, wenn  $v = 0$  ist.
- (ii) *Positive Homogenität:* Für jedes  $v \in V$  und jeden Skalar  $\lambda$  gilt  $\|\lambda v\| = |\lambda| \|v\|$ .
- (iii) *Dreiecksungleichung:* Für alle  $v, w \in V$  gilt  $\|v + w\| \leq \|v\| + \|w\|$ .

Ein *normierter Raum* ist ein Paar  $(V, \|\cdot\|)$ , wobei  $V$  ein Vektorraum über  $\mathbb{R}$  oder  $\mathbb{C}$  ist und  $\|\cdot\|$  eine Norm auf  $V$  ist.

Wenn wir den Skalkörper spezifizieren wollen, sagen wir oft „ $(V, \|\cdot\|)$  ist eine normierter Raum über  $\mathbb{R}$ “ bzw. „ $(V, \|\cdot\|)$  ist eine normierter Raum über  $\mathbb{C}$ “.

Der Zusammenhang zwischen normierten Räumen und metrischen Räumen wird durch die folgende Proposition beschrieben.

**Proposition A.4.2 (Die Metrik auf einem normierten Raum).** Sei  $(V, \|\cdot\|)$  ein normierter Raum. Wir definieren eine Abbildung  $d : V \times V \rightarrow [0, \infty)$  durch

$$d(v, w) = \|w - v\|$$

für alle  $v, w \in V$ . Dann ist  $d$  eine Metrik auf  $V$  (die man oft als die von  $\|\cdot\|$  induzierte Metrik bezeichnet).

**Aufgabe A.4.3.** Beweisen Sie Proposition A.4.2.

Proposition A.4.2 zeigt, dass wir jeden normierten Raum auch immer automatisch als metrischen Raum auffassen können – und dies werden wir auch stets tun. Insbesondere sind alle Begriffe, die für metrische Räume definiert sind – wie z.B. konvergente Folge, Cauchy-Folge, Vollständigkeit, Abgeschlossenheit, Offenheit – auch auf jedem normierten Raum automatisch definiert. Vollständige normierte Räume sind – in der Funktionalanalysis und weiter darüber hinaus – von so zentraler Bedeutung, dass sie einen eigenen Namen erhalten:

**Definition A.4.4 (Banachraum).** Einen normierten Raum bezeichnet man als *Banachraum*<sup>1</sup>, wenn er vollständig ist.

Wir wiederholen einige Beispiele für normierte Räume und geben jeweils an, ob sie vollständig sind:

**Beispiele A.4.5 (Einige endlich-dimensionale normierte Räume).** Sei  $\mathbb{K}$  gleich  $\mathbb{R}$  oder gleich  $\mathbb{C}$  und sei  $d \in \mathbb{N}$ . Wir betrachten einige verschiedene Normen auf  $\mathbb{K}^d$ :

- (a) *Euklidische Norm:* Für jedes  $x \in \mathbb{K}^d$  sei  $\|x\| = \left(\sum_{n=1}^d |x_n|^2\right)^{1/2}$ . Dann ist  $(\mathbb{K}^d, \|\cdot\|)$  ein normierter Raum. (Häufig schreibt man auch  $\|\cdot\|_2$  für die Euklidische Norm.)
- (b) *1-Norm:* Für jedes  $x \in \mathbb{K}^d$  sei  $\|x\|_1 = \sum_{n=1}^d |x_n|$ . Dann ist  $(\mathbb{K}^d, \|\cdot\|_1)$  ein normierter Raum.

<sup>1</sup>Benannt nach Stefan Banach, ebenso wie der Banachsche Fixpunktsatz (vgl. die Fußnote in Satz 4.1.4) und noch eine lange Liste weiterer Objekte und Resultate in der Funktionalanalysis.

- (c)  $\infty$ -Norm: Für jedes  $x \in \mathbb{K}^d$  sei  $\|x\|_\infty = \max\{|x_n| : n = 1, \dots, d\}$ . Dann ist  $(\mathbb{K}^d, \|\cdot\|_\infty)$  ein normierter Raum.

Alle endlich-dimensionalen normierten Räumen sind vollständig; also sind alle oben stehenden Räume Banachräume.

**Beispiele A.4.6 (Einige Folgenräume).** Sei  $\mathbb{K}$  gleich  $\mathbb{R}$  oder gleich  $\mathbb{C}$ .

- (a) *Raum der abbrechenden Folgen:* Es sei

$$c_{00} := \{(x_n)_{n \in \mathbb{N}} \subseteq \mathbb{K} : \text{nur endliche viele } x_n \text{ sind ungleich } 0\},$$

und für jede Folge  $(x_n)_{n \in \mathbb{N}} \in c_{00}$  sei  $\|(x_n)_{n \in \mathbb{N}}\|_\infty := \sup\{|x_n| : n \in \mathbb{N}\}$ . Dann ist  $(c_{00}, \|\cdot\|_\infty)$  ein normierter Raum, der aber nicht vollständig ist.

- (b) *Raum der Nullfolgen:* Es sei

$$c_0 := \{(x_n)_{n \in \mathbb{N}} \subseteq \mathbb{K} : \lim_{n \rightarrow \infty} x_n = 0\}$$

und  $\|\cdot\|_\infty$  sei definiert wie in (a) (aber dieses mal für jede Folge  $(x_n)_{n \in \mathbb{N}} \in c_0$  statt nur für Folgen aus  $c_{00}$ ). Dann ist  $(c_0, \|\cdot\|_\infty)$  ein Banachraum.

- (c) *Raum der  $p$ -summierbaren Folgen:* Wir fixieren eine reelle Zahl  $p \in [1, \infty)$ . Es sei

$$\ell^p := \{(x_n)_{n \in \mathbb{N}} \subseteq \mathbb{K} : \sum_{n=1}^{\infty} |x_n|^p < \infty\}$$

und für jede Folge  $(x_n)_{n \in \mathbb{N}} \in \ell^p$  setzen wir  $\|(x_n)_{n \in \mathbb{N}}\|_p := (\sum_{n=1}^{\infty} |x_n|^p)^{1/p}$ . Dann ist  $(\ell^p, \|\cdot\|_p)$  ein Banachraum.

Das nachfolgende Beispiel (b) spielt eine wichtige Rolle in der Theorie der gewöhnlichen Differentialgleichungen.

**Beispiele A.4.7 (Räume stetiger Funktionen).** Sei  $\mathbb{K}$  gleich  $\mathbb{R}$  oder gleich  $\mathbb{C}$ . Es sei  $\mathbb{K}$  mit der Euklidischen Metrik ausgestattet und es sei  $I \subseteq \mathbb{R}$  ein kompaktes Intervall.

- (a) *Raum stetiger Funktionen mit  $\infty$ -Norm:* Es sei

$$C(I) := \{f : I \rightarrow \mathbb{K} : f \text{ ist stetig}\}$$

der Raum der stetigen  $\mathbb{K}$ -wertigen Funktionen auf  $I$ . Für jedes  $f \in C(I)$  setzen wir

$$\|f\|_\infty := \sup_{t \in I} |f(t)|.$$

Dann ist  $(C(I), \|\cdot\|_\infty)$  ein Banachraum.

- (b) *Raum stetiger Vektor-wertiger Funktionen mit  $\infty$ -Norm:* Wir verallgemeinern das Beispiel aus (a) nun ein wenig und betrachten anstelle von skalar-wertigen Funktionen solche, die Werte in  $\mathbb{K}^d$  annehmen.

Sei  $d \in \mathbb{N}$  und bezeichne  $\|\cdot\|$  die Euklidische Norm auf  $\mathbb{K}^d$ . Wir setzen

$$C(I; \mathbb{K}^d) := \{f : M \rightarrow \mathbb{K}^d : f \text{ ist stetig}\}.$$

Für jedes  $f \in C(I; \mathbb{K}^d)$  sei  $\|f\|_\infty := \sup_{t \in I} \|f(t)\|$ . Dann ist  $(C(I; \mathbb{K}^d), \|\cdot\|_\infty)$  ein Banachraum.

- (c) *Raum stetiger Funktionen mit 1-Norm:* Wir betrachten wieder den Raum  $C(I)$  aus Beispiel (a), aber dieses Mal mit der Norm  $\|\cdot\|_1$ , die durch

$$\|f\|_1 := \int_I |f(t)| \, dt$$

für alle  $f \in C(I)$  gegeben ist. Dann ist  $(C(I), \|\cdot\|_1)$  ein normierter Raum, der aber nicht vollständig ist.

**Aufgabe A.4.8.** Was haben die Beispiele A.4.7(a) and (c) mit der Einstiegsfrage (c) dieses Kapitels zu tun?

Zum Schluss geben wir noch eine Charakterisierung der Vollständigkeit von normierten Räumen mit Hilfe von absolut konvergenten Reihen an. Dazu benötigen wir den folgenden Begriff:

**Definition A.4.9 (Konvergente und absolut konvergente Reihen).** Sei  $(V, \|\cdot\|)$  ein normierter Raum, und für jedes  $k \in \mathbb{N}$  sei  $v_k \in V$ .

- (i) Wir sagen, dass die Reihe  $\sum_{n=0}^{\infty} v_n$  absolut konvergiert, wenn

$$\sum_{k=0}^{\infty} \|v_k\| < \infty$$

ist, d.h. wenn die Reihe über die Normen von  $v_k$  in  $\mathbb{R}$  konvergiert.

- (ii) Wir sagen, dass die Reihe  $\sum_{n=0}^{\infty} v_n$  konvergiert, wenn die Folge der Partialsummen

$$\left( \sum_{k=0}^n v_k \right)_{n \in \mathbb{N}_0}$$

gegen einen Vektor aus  $V$  konvergiert.

Die folgende Charakterisierung von Banachräumen ist ein wesentlicher Grund, warum man Vollständigkeit von normierten Räumen in der Funktionalanalysis andauernd benötigt.

**Proposition A.4.10.** Für jeden normierten Vektorraum  $(V, \|\cdot\|)$  sind die folgenden beiden Aussagen äquivalent:

- (i) Der Raum  $(V, \|\cdot\|)$  ist vollständig, also ein Banachraum.
- (ii) Jede absolute konvergente Reihe in  $V$  ist konvergent.

*Beweis.* Der Beweis ist eine Standard-Übungsaufgabe am Anfang der meisten Funktionalanalysis-Vorlesungen; man benötigt dafür nur die Begriffe, die in diesem Abschnitt wiederholt wurde, und Sie können den Beweis gerne selbst führen um sich mit den Konvergenzbegriffen für Reihen vertraut zu machen.  $\square$

Bitte beachte Sie unbedingt, dass die umgekehrte Aussage in Proposition A.4.10 nicht stimmt, d.h. eine konvergente Reihe muss im Allgemeinen nicht absolut konvergieren (und zwar auch dann nicht, wenn der Raum vollständig ist). Sie kennen sogar im ein-dimensionalen Raum  $\mathbb{R}$  aus der Analysis 1 Beispiele für konvergente Reihen, die nicht absolut konvergieren.<sup>2</sup>

**Literaturstellen und verwandte Ergebnisse.** Das tiefere Studium von normierten Räumen, und insbesondere von Banachräumen, ist Teil der *Funktionalanalysis*. Es gibt zahlreiche einführende Lehrbücher zur Funktionalanalysis; wir verweisen hier beispielhaft auf [Kab11].

## A.5 Kompaktheit

Kompaktheit ist eines der zentralsten Konzepte in der gesamten Analysis. Wir geben hier eine kurze Wiederholung des Begriffs im Rahmen metrischer Räume.

Wir starten zunächst mit zwei Beschränktheitsbegriffen:

**Definition A.5.1 (Beschränkte und total beschränkte Mengen).** Sei  $(M, d)$  ein metrischer Raum und sei  $C \subseteq M$ .

- (i) Die Menge  $C$  heißt *beschränkt*, wenn für alle  $x \in M$  eine Zahl  $R \in [0, \infty)$  existiert derart, dass  $d(c, x) \leq R$  für alle  $c \in C$  gilt.
- (ii) Die Menge  $C$  heißt *total beschränkt*, wenn es für jedes  $\varepsilon > 0$  eine endliche Menge  $F \subseteq M$  mit der folgenden Eigenschaft gibt:  
Für jedes  $c \in C$  gibt es ein  $f \in F$  mit  $d(c, f) \leq \varepsilon$ .

Wenn Ihnen diese Begriffe aus der Analysis 2 nicht mehr vertraut sind, können Sie zum Beispiel die folgenden Aufgaben zur Wiederholung lösen:

---

<sup>2</sup>Welche zum Beispiel?

**Aufgabe A.5.2.** Sei  $(M, d)$  ein metrischer Raum und  $C \subseteq M$ .

- (a) Zeigen Sie:  $C$  ist genau dann beschränkt, wenn  $C$  leer ist oder es ein  $x_0 \in C$  und ein  $R \in [0, \infty)$  gibt derart, dass  $d(c, x_0) \leq R$  für alle  $c \in C$ .
- (b) Zeigen Sie: Die Menge  $C$  ist genau dann total beschränkt, wenn es für jedes  $\varepsilon > 0$  eine endliche Menge  $F \subseteq C$  mit der folgenden Eigenschaft gibt:  
Für jedes  $c \in C$  gibt es ein  $f \in F$  mit  $d(c, f) \leq \varepsilon$ .

*Hinweis:* Der Unterschied dieser Aussage zur Aussage in der Definition von totaler Beschränktheit besteht (a priori) darin, dass nun nur Mengen  $F \subseteq C$  (anstelle von Mengen  $F \subseteq M$ ) betrachtet werden.

**Aufgabe A.5.3.** (a) Sei  $(M, d)$  ein metrischer Raum und  $C \subseteq M$ . Zeigen Sie: Wenn  $C$  total beschränkt ist, dann ist  $C$  beschränkt.

- (b) Betrachten Sie nun den normierten Raum  $\ell^1$  (siehe Beispiel A.4.6(c)), sagen wir über dem Skalkörper  $\mathbb{R}$ . Finden Sie eine Menge  $C \subseteq \ell^1$ , die beschränkt, aber nicht total beschränkt ist.

Der Begriff *Kompaktheit* ist in metrischen Räumen folgendermaßen definiert:

**Definition A.5.4 (Kompakte Menge).** Sei  $(M, d)$  ein metrischer Raum. Eine Teilmenge  $C \subseteq M$  heißt *kompakt*, wenn die folgende Aussage gilt:

Wann immer ein System von offenen Mengen  $(U_i)_{i \in I}$  in  $M$  mit beliebiger Indexmenge  $I$  die Eigenschaft  $\bigcup_{i \in I} U_i \supseteq C$  erfüllt, gibt es eine endliche Teilmenge  $J \subseteq I$  mit  $\bigcup_{i \in J} U_i \supseteq C$ .

Kurz gesprochen bedeutet Kompaktheit von  $C$  also: *Jede offene Überdeckung von  $C$  in  $M$  besitzt eine endliche Teilüberdeckung.*

Auf den ersten Blick mag der Begriff der Kompaktheit etwas merkwürdig anmuten, und es scheint als hätte er nur wenig mit den Begriffen zu tun, die wir zuvor in diesem Unterabschnitt wiederholt haben. Diese Unklarheiten sollten sich durch den folgenden Satz klären:

**Satz A.5.5 (Kompaktheit in metrischen Räumen).** Sei  $(M, d)$  ein metrischer Raum und sei  $C \subseteq M$ . Dann sind äquivalent:

- (i) Die Menge  $C$  ist kompakt.
- (ii) Die Menge  $C$  ist total beschränkt und  $C$  ist als metrischer Raum (mit der von  $M$  induzierten Metrik) vollständig.
- (iii) Jede Folge in  $C$  besitzt eine Teilfolge, die gegen einen Punkt in  $C$  konvergiert.

Die Aussage (c) wirkt für Sie auf den ersten Blick vielleicht etwas intuitiver als die abstrakt anmutende Definition von Kompaktheit. Es gibt aber einige Gründe, dennoch die „endliche Überdeckungseigenschaft“ als Definition für Kompaktheit zu verwenden. Zum einen ist letztere Eigenschaft in einigen Beweisen tatsächlich einfacher handhabbar als die Folgenbedingung aus Satz A.5.5(c); zum anderen funktioniert die Definition von Kompaktheit mittels „endlicher Überdeckungseigenschaft“ auch noch im allgemeineren Rahmen topologischer Räume (die Folgenbedingung aus Satz A.5.5(c) jedoch nicht mehr).

In vollständigen metrischen Räumen lässt sich die Aussage (b) aus obigem Satz etwas vereinfachen:

**Korollar A.5.6 (Kompaktheit in vollständigen metrischen Räumen).** Sei  $(M, d)$  ein vollständiger metrischer Raum und sei  $C \subseteq M$ . Dann sind äquivalent:

- (i) Die Menge  $C$  ist kompakt.
- (ii) Die Menge  $C$  ist total beschränkt und abgeschlossen.

Man beachte, dass die Beschränktheitsbedingung in Teilaussage (b) „ $C$  ist total beschränkt“ lautet, und nicht etwas „ $C$  ist beschränkt“. Dass eine Menge  $C \subseteq M$  genau dann kompakt ist, wenn Sie beschränkt und abgeschlossen ist, gilt nur in wenigen Spezialfällen, zum Beispiel wenn  $M$  ein endlich-dimensionaler normierter Vektorraum ist.

**Aufgabe A.5.7.** Geben Sie eine Teilmenge von  $\ell^1$  an, die beschränkt und abgeschlossen, aber nicht kompakt ist.

Die folgende Begriffsbildung ist oft nützlich, und wir verwenden sie zum Beispiel bei der Formulierung des Schauderschen Fixpunktsatzes (Satz B.3.6).

**Definition A.5.8 (Relativ kompakte Menge).** Sei  $(M, d)$  ein metrischer Raum und sei  $C \subseteq M$ . Die Menge  $C$  heißt *relativ kompakt*, wenn ihr Abschluss  $\overline{C}$  kompakt ist.

Um die obige Definition zu verstehen ist es wichtig, sich daran zu erinnern, wie der Abschluss  $\overline{C}$  einer Menge  $C$  in einem metrischen Raum  $(M, d)$  definiert ist: Es ist  $\overline{C}$  per Definition der Durchschnitt aller abgeschlossenen Teilmengen von  $M$ , die  $C$  enthalten – also die *kleinste* abgeschlossene Teilmenge von  $M$ , die  $C$  enthält.

Wir schließen diesen Abschnitt mit einer Charakterisierung der relativ kompakten Teilmengen des Banachraumes  $C(I; \mathbb{K}^d)$  (wobei dieser Raum mit der Norm  $\|\cdot\|_\infty$  ausgestattet ist; siehe Beispiel A.4.7(b)); es handelt sich um eine Variante des sehr bekannten *Satzes von Arzelà–Ascoli*.

**Satz A.5.9 (Satz von Arzelà–Ascoli<sup>3</sup>).** Sei  $\mathbb{K}$  gleich  $\mathbb{R}$  oder gleich  $\mathbb{C}$ , sei  $d \in \mathbb{N}$  und sei  $\emptyset \neq I \subseteq \mathbb{R}$  ein kompaktes Intervall. Für jede Teilmenge  $C \subseteq C(I; \mathbb{K}^d)$  sind äquivalent:

- (i) Die Menge  $C$  ist relativ kompakt.
- (ii) Die Menge  $C$  ist beschränkt und gleichgradig stetig, d.h. für jedes  $\varepsilon > 0$  und jedes  $t \in I$  gibt es ein  $\delta > 0$  derart, dass  $\|f(s) - f(t)\| < \varepsilon$  für alle  $f \in C$  und alle  $s \in I$  mit  $|s - t| < \delta$  gilt.

Das folgende Korollar des Satzes von Arzelà-Ascoli verwenden wir im Beweis des Fixpunktsatzes von Peano.

**Korollar A.5.10.** Sei  $\mathbb{K}$  gleich  $\mathbb{R}$  oder gleich  $\mathbb{C}$ , sei  $d \in \mathbb{N}$  und sei  $\emptyset \neq I \subseteq \mathbb{R}$  ein kompaktes Intervall. Sei  $C \subseteq C(I; \mathbb{K}^d)$  beschränkt und es gebe eine Zahl  $L \in [0, \infty)$  derart, dass jedes  $f \in C$  Lipschitz-stetig mit Lipschitz-Konstante höchstens  $L$  ist. Dann ist  $C$  relativ kompakt.

**Aufgabe A.5.11.** Zeigen Sie, dass Korollar A.5.10 aus Satz A.5.9 folgt.

## A.6 Zusammenhängende Mengen und Zusammenhangsargumente

In diesem Abschnitt wiederholen wir kurz den Begriff der *zusammenhängenden Menge* in einem metrischen Raum und erläutern, warum dieses Konzept in Beweisen oft sehr nützlich ist. In der Vorlesung kommt dieses Konzept im Beweis des globalen Eindeutigkeitssatzes von Picard–Lindelöf (Korollar 4.1.11) vor; dies ist auch der Grund, warum wir nur Lösungen von Differentialgleichungen betrachten, die auf Intervallen definiert sind (vgl. Bemerkung A.6.11 am Ende dieses Abschnitts).

**Definition A.6.1 (Zusammenhängende Menge).** Sei  $(M, d)$  ein metrischer Raum. Eine Teilmenge  $N \subseteq M$  heißt *zusammenhängend*, falls sie die folgende Eigenschaft erfüllt:

Für alle offenen und disjunkten Mengen  $U_1, U_2 \subseteq M$ , die  $N$  überdecken (im Sinne von  $U_1 \cup U_2 \supseteq N$ ) ist  $N$  eine Teilmenge von  $U_1$  oder eine Teilmenge von  $U_2$ .

**Bemerkung A.6.2.** Es ist nicht schwer zu zeigen, dass *zusammenhängend* eine *intrinsische Eigenschaft* ist, d.h. eine Teilmenge  $N$  eines metrischen Raumes  $(M, d)$  ist genau dann zusammenhängend, wenn sie als Teilmenge des metrischen Raumes  $(N, d_N)$  zusammenhängend ist.

<sup>3</sup>Benannt nach Cesare Arzelà (1847 – 1912) und Giulio Ascoli (1843 – 1896), beide italienische Mathematiker.

Aufgrund der vorangehenden Bemerkung kann man sich, wenn man wissen möchte, ob eine Menge  $N \subseteq M$  zusammenhängend ist, darauf zurückziehen nur den Raum  $(N, d_N)$  zu betrachten und sich zu fragen, ob  $N$  in diesem Raum zusammenhängend ist. Für diese Frage – ob ein gegebener metrischer Raum selbst zusammenhängend ist – ist auch die folgende Proposition häufig nützlich:

**Proposition A.6.3.** *Sei  $(M, d)$  ein metrischer Raum. Dann sind äquivalent:*

- (i) *Die Menge  $M$  ist zusammenhängend.*
- (ii) *Mit Ausnahme von  $\emptyset$  und  $M$  selbst ist keine Teilmenge von  $M$  sowohl offen als auch abgeschlossen.*

Proposition A.6.3 spielt auch für sogenannte *Zusammenhangs-Argumente* eine wichtige Rolle – siehe Proposition A.6.10 weiter unten. Nun wollen wir aber zunächst einige Beispiele ansprechen.

**Beispiele A.6.4.** (a) Eine Teilmenge  $N \subseteq \mathbb{R}$  ist genau dann zusammenhängend, wenn  $N$  ein Intervall<sup>4</sup> ist.

- (b) Jede Kreislinie, jede offene Kreisscheibe und jede abgeschlossene Kreisscheibe in  $\mathbb{R}^2$  ist zusammenhängend.
- (c) Jede Strecke zwischen zwei Punkten in  $\mathbb{R}^d$  ist zusammenhängend.
- (d) Die Vereinigung von zwei nichtleeren disjunkten offenen Teilmengen in  $\mathbb{R}^2$  ist nicht zusammenhängend.
- (e) Die Vereinigung von zwei nichtleeren disjunkten kompakten Teilmengen in  $\mathbb{R}^2$  ist nicht zusammenhängend.

Die Eigenschaft *zusammenhängend* bleibt unter stetigen Abbildungen erhalten; dies ist der Inhalt der folgenden Proposition:

**Proposition A.6.5.** *Seien  $(L, d_L)$  und  $(M, d_M)$  metrische Räume und sei  $f : L \rightarrow M$  stetig. Wenn  $N \subseteq L$  zusammenhängend ist, dann ist  $f(N)$  ebenfalls zusammenhängend.*

Als Konsequenz dieser Proposition kann man zum Beispiel ganz leicht die Aussage aus Beispiel A.6.4(c) beweisen: Wenn  $x, y$  zwei Punkte in  $\mathbb{R}^d$  sind, dann ist

$$f : [0, 1] \ni t \mapsto x + t(y - x) \in \mathbb{R}^d$$

---

<sup>4</sup>Wobei wir hierbei auch die leere Menge als Intervall auffassen (was in manchen Vorlesungen und Lehrbüchern anders gehandhabt wird).

eine stetige Abbildung, und das Bild der zusammenhängenden Menge  $[0, 1]$  unter dieser Abbildung ist genau die Strecke zwischen  $x$  und  $y$ .

Die folgende Proposition ist nützlich um für viele weitere Mengen zu zeigen, dass Sie zusammenhängend sind:

**Proposition A.6.6.** *Sei  $(M, d)$  ein metrischer Raum und sei  $N \subseteq M$ . Die folgenden beiden Aussagen sind äquivalent:*

- (i) *Die Menge  $N$  ist zusammenhängend.*
- (ii) *Für alle Punkte  $x, y \in N$  gibt es eine Teilmenge  $L \subseteq N$ , die zusammenhängend ist und die  $x$  und  $y$  enthält.*

Als einfache, aber wichtige Konsequenz von Proposition A.6.6 erhält man das folgende Korollar:

**Korollar A.6.7.** *Sei  $(M, d)$  ein metrischer Raum und sei  $N \subseteq M$  Weg-zusammenhängend, d.h. für alle Punkte  $x, y \in N$  gebe es eine stetige Abbildung  $\gamma : [0, 1] \rightarrow N$  mit  $\gamma(0) = x$  und  $\gamma(1) = y$ . Dann ist  $N$  zusammenhängend.*

Das Korollar folgt sofort aus den Propositionen A.6.6 und A.6.3.

**Beispiel A.6.8.** Sei  $(V, \|\cdot\|)$  ein normierter Raum über  $\mathbb{R}$  mit  $\dim V \geq 2$ . Dann ist die *Einheitssphäre*

$$S := \{v \in V : \|v\| = 1\}$$

zusammenhängend.

*Beweis.* Seien  $x, y \in S$ . Wir wollen Korollar A.6.7 anwenden und betrachten hierzu zunächst den Fall, dass  $(x, y)$  linear unabhängig ist. In diesem Fall enthält das Verbindungsstück

$$\{x + (1 - t)(y - x) : t \in [0, 1]\} \subseteq V$$

nicht den Nullvektor. Somit ist

$$\gamma : [0, 1] \ni t \mapsto \frac{x + (1 - t)(y - x)}{\|x + (1 - t)(y - x)\|} \in S$$

eine stetige Abbildung mit  $\gamma(0) = x$  und  $\gamma(1) = y$ .

Nun betrachten wir noch den Fall, dass  $(x, y)$  linear abhängig ist. Dann ist  $y = x$  oder  $y = -x$ . Im Falle  $y = x$  gibt es offensichtlich einen stetigen Weg von  $x$  nach  $y$ , also sei nun  $y = -x$ . Wegen  $\dim V \geq 2$  gibt es ein  $z \in S$  derart, dass  $(x, z)$  linear unabhängig ist, und somit ist auch  $(z, y)$  linear unabhängig. Wie im vorangehenden Schritt finden wir somit einen stetigen Weg in  $S$ , der  $x$  und  $z$  verbindet und einen stetigen Weg in  $S$ , der  $z$  und  $y$  verbindet. Indem wir

diese beiden Weg zusammensetzen, erhalten wir einen stetigen Weg in  $S$ , der  $x$  und  $y$  verbindet.

Also folgt aus Korollar A.6.7, dass  $S$  zusammenhängend ist.  $\square$

Man beachte, dass die Einheitsphäre im ein-dimensionalen normierten Vektorraum  $\mathbb{R}$  (mit Skalkörper  $\mathbb{R}$  und dem Betrag als Norm) nicht zusammenhängend ist – d.h. die Voraussetzung  $\dim V \geq 2$  in Beispiel A.6.8 kann man nicht einfach weglassen.

**Aufgabe A.6.9.** Sei  $(V, \|\cdot\|)$  ein normierter Vektorraum über  $\mathbb{C}$ . Zeigen Sie, dass die Einheitsphäre in  $V$  immer zusammenhängend ist – egal, welche Dimension  $V$  besitzt.

Zusammenhängende Menge spielen eine wichtige Rolle bei sogenannten *Zusammenhangs-Argumenten*; in typisches solches Argument beschrieben wir in der nächsten Proposition:

**Proposition A.6.10 (Zusammenhangs-Argument).** *Seien  $(L, d_L)$  und  $(M, d_M)$  metrische Räume und sei  $L$  zusammenhängend. Seien  $f, g : L \rightarrow M$  stetige Funktionen. Wenn die Menge*

$$C := \{x \in L : f(x) = g(x)\}$$

*nicht-leer und offen ist, dann sind die beiden Funktionen  $f$  und  $g$  gleich.*

*Beweis.* Wir müssen nur zeigen, dass  $C = L$  gilt. Weil  $f$  und  $g$  stetig sind, ist  $C$  abgeschlossen, und nach Voraussetzung ist  $C$  offen. Also folgt, weil  $L$  zusammenhängend ist, aus Proposition A.6.3, dass  $C = \emptyset$  oder  $C = L$  gilt. Weil  $C$  nach Voraussetzung nicht-leer ist, folgt also  $C = L$ .  $\square$

Genau dieses Argument kommt im Beweis des globalen Eindeutigkeitsatzes von Picard–Lindelöf (Korollar 4.1.11) vor.

Es gibt auch etwas subtilere Zusammenhangsargumente. In der Funktionentheorie betrachtet man zum Beispiel, wenn man den *Identitätssatz für holomorphe Funktionen* beweist, nicht einfach die Menge  $C$  aller Punkte, auf denen zwei gegebenen Funktionen  $f$  und  $g$  übereinstimmen, sondern man betrachtet die Menge aller Häufungspunkte von  $C$  – nennen wir sie  $H$  – und zeigt, dass die abgeschlossene Menge  $H$  nicht-leer und offen ist.

Wir schließen diesen Abschnitt mit einer Bemerkung dazu, weshalb wir in der Theorie gewöhnlicher Differentialgleichungen stets Lösungen betrachten, die auf Intervallen definiert sind.

**Bemerkung A.6.11.** Die Tatsache, dass Intervalle die einzigen zusammenhängenden Teilmengen von  $\mathbb{R}$  sind, ist auch der Grund, warum wir in der

Vorlesung nur Lösungen von Differentialgleichungen betrachten, die auf Intervallen definiert sind: Andere Teilmengen von  $\mathbb{R}$  sind nicht zusammenhängend, und somit können wir auf solchen Mengen keine Eindeutigkeit der Lösung eines Anfangswertproblems erwarten.

Dies sind man schon im folgenden sehr einfachen Beispiel: Betrachten wir das ein-dimensionale Anfangswertproblem

$$(*) \quad \begin{cases} \dot{x}(t) = 0, \\ x(0) = 1. \end{cases}$$

Wenn  $I \subseteq \mathbb{R}$  ein nicht-triviales Intervall ist, das den Startzeitpunkt 0 enthält, dann sind zwei Lösungen  $x : I \rightarrow \mathbb{R}$  und  $\tilde{x} : I \rightarrow \mathbb{R}$  laut Korollar 4.1.11 gleich (und somit an jedem Zeitpunkt  $t \in I$  gleich 1).

Wenn wir hingegen auch Lösungen zulassen, die nicht auf einem Intervall definiert sind, so können wir zum Beispiel die beiden Funktionen

$$x, \tilde{x} : [-1, 1] \cup [3, 4] \rightarrow \mathbb{R}$$

betrachten, wobei  $x(t) = 1$  für alle  $t \in [-1, 1] \cup [3, 4]$  sei und wobei

$$\tilde{x}(t) = \begin{cases} 1 & \text{für } t \in [-1, 1], \\ 2 & \text{für } t \in [3, 4] \end{cases}$$

sei. Beiden Funktionen erfüllen die Anfangsbedingung und die Differentialgleichung in (\*); aber  $x$  und  $\tilde{x}$  sind nicht gleich.



## Fixpunktsätze

### Fragen zum Einstieg.

- (a) Was besagt noch einmal der *Zwischenwertsatz* aus der Analysis 1? Haben Sie eine anschauliche Vorstellung davon, weshalb der Satz gilt?
- (b) Gibt es eine stetige Funktion  $f : [0, 1] \rightarrow \mathbb{R}$  mit der Eigenschaft

$$f(t) = 1 + \int_0^t f(s) \, ds \quad \text{für alle } t \in [0, 1]?$$

- (c) Suchen Sie im Internet nach dem *Satz vom Fußball* (sofern Sie ihn nicht schon kennen).

### B.1 Aperitif: Wozu Fixpunktsätze?

Eine grundlegende Problemstellung in der Mathematik ist es, Gleichungen zu lösen. Betrachten Sie z.B. die folgende Gleichung: Wir suchen eine reelle Zahl  $x$  mit der Eigenschaft

$$e^{-x} = x^5. \tag{B.1}$$

Diese Lösung lässt sich nicht ohne weiteres „explizit“ (oder „analytisch“) lösen – was auch immer man mit diesen Ausdrücken überhaupt meinen mag. Man kann aber sehr leicht sehen, dass die Gleichung (B.1) zumindest eine Lösung besitzt, und zwar sogar eine in dem Intervall  $[0, 1]$ .

*Beweis.* Wir betrachten die stetige Abbildung  $f : [0, 1] \rightarrow \mathbb{R}$ ,  $x \mapsto e^{-x} - x^5$ . Dann gilt  $f(0) = 1 \geq 0$  und  $f(1) = 1/e - 1 \leq 0$ . Also besitzt  $f$  aufgrund des Zwischenwertsatzes eine Nullstelle in  $[0, 1]$ .  $\square$

Der Zwischenwertsatz liefert also eine Existenzaussage über die Lösung von Gleichungen.<sup>1</sup> Er ist allerdings eine sehr ein-dimensionale Angelegenheit, und wir würden gerne auch Existenzaussagen über Lösungen von Gleichungen in höherer Dimension (oder sogar in unendlicher Dimension) erhalten.

Wenn wir dieses Ziel haben, ist es hilfreich, den Zwischenwertsatz ein wenig umzuformulieren. Es gilt:

**Satz B.1.1 (Fixpunkt-Variante des Zwischenwertsatzes).** *Seien  $a < b$  reelle Zahlen und sei  $f : [a, b] \rightarrow \mathbb{R}$  stetig. Wenn  $f(a) \geq a$  und  $f(b) \leq b$  gilt, dann gibt es ein  $x \in \mathbb{R}$  mit  $f(x) = x$ .*

<sup>1</sup>Übrigens kann man den Zwischenwertsatz sehr leicht aus Proposition A.6.5 und aus der Charakterisierung zusammenhängender Teilmengen von  $\mathbb{R}$  in Beispiel A.6.4(a) folgern.

**Aufgabe B.1.2.** Zeigen Sie, dass Satz B.1.1 äquivalent zum Zwischenwertsatz ist.

Man sagt, dass  $x$  ein *Fixpunkt* einer Funktion  $f$  ist, wenn  $f(x) = x$  gilt. Satz B.1.1 macht also nicht eine Aussage über die Existenz einer Nullstelle von  $f$ , sondern über die Existenz eines Fixpunkts von  $f$ . Aufgabe B.1.2 zeigen aber, dass beide Aussagen (für stetige Funktionen, die auf einem Intervall definiert sind) im Wesentlichen äquivalent sind.

In konkreten Situationen ist es oft leicht eine gegebene Gleichung in eine Fixpunktgleichung umzuschreiben. Betrachten Sie zum Beispiel noch einmal die Gleichung B.1. Sie ist offenbar äquivalent zur Fixpunktgleichung

$$e^{-x} - x^5 + x = x. \quad (\text{B.2})$$

Dies Gleichung sieht auf den ersten Blick etwas komplizierter aus als (B.1), aber mit Satz B.1.1 lässt sich die Existenz einer Lösung natürlich genauso leicht zeigen:

Sei  $f : [0, 1] \rightarrow \mathbb{R}$ ,  $x \mapsto e^{-x} - x^5 + x$ . Dann ist  $f(0) = 1 \geq 0$  und  $f(1) = 1/e \leq 1$ , also besitzt  $f$  laut Satz B.1.1 einen Fixpunkt.

Der entscheidende Punkt ist nun: Fixpunktsätze lassen sich sehr gut auf höher-dimensionale Räume verallgemeinern! Ein Beispiel sind die Fixpunktsätze von Brouwer und Schauder, die wir in Abschnitt B.3 besprechen. Sie gelten in mehrdimensionalen bzw. sogar unendlich-dimensionalen Räumen. Der Banachsche Fixpunktsatz, den wir nun zuerst besprechen (Abschnitt B.2) gilt sogar in vollständigen metrischen Räumen.

Wir sollten im Hinterkopf behalten, dass Fixpunktsätze im Wesentlichen eine Möglichkeit sind, um die Existenz einer Lösung für Gleichungen zu zeigen (und im Falle des Banachschen Fixpunktsatzes sogar noch dazu die Eindeutigkeit!).

## B.2 Der Banachsche Fixpunktsatz

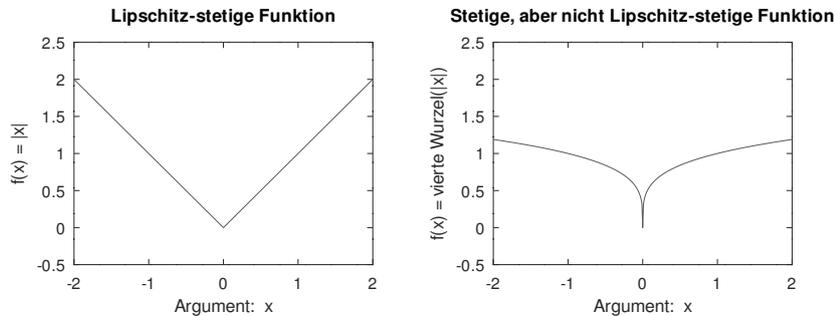
Um den Banachschen Fixpunktsatz zu formulieren, benötigen wir den Begriff der *Lipschitz-Stetigkeit*.

**Definition B.2.1 (Lipschitz-Stetigkeit).** Seien  $(M, d_M)$  und  $(N, d_N)$  metrische Räume. Eine Abbildung  $f : M \rightarrow N$  heißt *Lipschitz-stetig*, wenn es eine reelle Zahl  $L \geq 0$  gibt derart, dass

$$d_N(f(x), f(y)) \leq L d_M(x, y) \quad \text{für alle } x, y \in M \quad (\text{B.3})$$

gilt. In diesem Fall heißt die kleinste reelle Zahl  $L \geq 0$ , für die (B.3) gilt, die *Lipschitz-Konstante* von  $f$ .

Abbildung B.1: Graphische Darstellung der Lipschitz-stetigen Funktion  $[-2, 2] \ni x \mapsto |x| \in \mathbb{R}$  und der Funktion  $[-2, 2] \ni x \mapsto |x|^{1/4}$ , die zwar stetig, aber nicht Lipschitz-stetig ist.



**Aufgaben B.2.2.** (a) Woher wissen wir in Definition B.2.1 überhaupt, dass es eine *kleinste* Zahl  $L \geq 0$  gibt, für die (B.3) gilt?

(b) Zeigen Sie, dass jede Lipschitz-stetige Funktion stetig ist.

Anschaulich bedeutet Lipschitz-Stetigkeit, dass eine Funktion „sich nirgends schneller als linear ändert“. Ein einfaches Beispiel einer Lipschitz-stetigen Funktion ist

$$[-2, 2] \ni x \mapsto |x| \in \mathbb{R};$$

dass diese Funktion Lipschitz-stetig ist, kann man mit Hilfe der unteren Dreiecksungleichung beweisen. Ein Beispiel einer Funktion, die stetig, aber nicht Lipschitz-stetig ist, ist

$$[-2, 2] \ni x \mapsto |x|^{1/4} \in \mathbb{R};$$

der Grund, weshalb diese Funktion nicht Lipschitz-stetig ist, ist ihr Verhalten im Punkt  $x = 0$ : je näher man diesem Punkt kommt, desto schneller ändert sich die Funktion – man kann, anschaulich gesprochen, ihre Änderung also nicht durch eine lineare Funktion abschätzen. Die beiden soeben diskutierten Funktionen sind in der Abbildung B.1 graphisch dargestellt.

Nun kommen wir zum Banachschen Fixpunktsatz, der diesem Abschnitt seinen Namen gibt. Wir geben im Folgenden eine Variante des Satzes an, die man häufig in der Literatur findet.

**Satz B.2.3 (Banachscher Fixpunktsatz).** Sei  $(M, d)$  ein nicht-leerer vollständiger metrischer Raum und sei  $f : M \rightarrow M$  eine Lipschitz-stetige Funktion mit Lipschitzkonstante  $L < 1$ .

Dann besitzt  $f$  genau einen Fixpunkt  $\hat{x} \in M$ . Außerdem konvergiert für jedes  $x \in M$  die Folge der Iterierten  $(f^n(x))_{n \in \mathbb{N}}$  gegen  $\hat{x}$ .

*Beweis.* Wir bemerken zunächst, dass  $f$  höchstens einen Fixpunkt haben kann: Wenn nämlich  $\hat{x}_1$  und  $\hat{x}_2$  Fixpunkte von  $f$  sind, dann gilt

$$d(\hat{x}_2, \hat{x}_1) = d(f(\hat{x}_2), f(\hat{x}_1)) \leq L d(\hat{x}_2, \hat{x}_1);$$

wegen  $L < 1$  folgt daraus  $d(\hat{x}_2, \hat{x}_1) = 0$ , also  $\hat{x}_2 = \hat{x}_1$ .

Sei nun  $x \in M$  beliebig. Wir zeigen, dass  $(f^n(x))_{n \in \mathbb{N}}$  eine Cauchy-Folge ist. Für jedes  $n \in \mathbb{N}$  gilt die Abschätzung

$$d(f^{n+1}(x), f^n(x)) \leq L d(f^n(x), f^{n-1}(x));$$

somit folgt

$$d(f^{n+1}(x), f^n(x)) \leq L^n d(f(x), x).$$

für alle  $n \in \mathbb{N}_0$ . Aus der Dreieckungleichung folgt somit für jedes  $n \in \mathbb{N}_0$  die Ungleichung

$$d(f^n(x), x) \leq \sum_{k=0}^{n-1} d(f^{k+1}(x), f^k(x)) \leq \sum_{k=0}^{n-1} L^k d(f(x), x) \leq \frac{1}{1-L} d(f(x), x).$$

Für festes  $n_0 \in \mathbb{N}$  und jede natürliche Zahl  $n \geq n_0$  gilt deshalb

$$d(f^n(x), f^{n_0}(x)) \leq L^{n-n_0} d(f^{n-n_0}(x), x) \leq L^{n_0} \frac{1}{1-L} d(f(x), x).$$

Also ist  $(f^n(x))_{n \in \mathbb{N}}$  in der Tat eine Cauchy-Folge. Aus der Vollständigkeit von  $(M, d)$  folgt somit, dass diese Folge gegen einen Punkt  $\hat{x} \in M$  konvergiert.

Nun verwenden wir die Stetigkeit von  $f$  um zu sehen, dass  $\hat{x}$  ein Fixpunkt ist: Es gilt

$$f(\hat{x}) = f\left(\lim_{n \rightarrow \infty} f^n(x)\right) = \lim_{n \rightarrow \infty} f(f^n(x)) = \lim_{n \rightarrow \infty} f^{n+1}(x) = \hat{x}.$$

Insbesondere folgt, dass  $f$  einen Fixpunkt  $\hat{x}$  besitzt. □

Die oben stehende Variante des Banachschen Fixpunktsatzes kennen Sie wahrscheinlich aus Ihrer Analysis 2-Vorlesung oder aus der Literatur. In Abschnitt 4.1 beweisen wir eine etwas allgemeinere Variante des Banachschen Fixpunktsatzes, die sehr nützlich für den Beweis des Satzes von Picard-Lindelöf ist. Den Beweis von Satz B.2.3 haben wir trotzdem in dieses Manuskript aufgenommen, das es lehrreich ist, ihn mit dem Beweis der allgemeineren Version des Resultats in Satz 4.1.4 zu vergleichen.

### B.3 Die Fixpunktsätze von Brouwer und Schauder

Der Banachsche Fixpunktsatz liefert sowohl die Existenz als auch die Eindeutigkeit eines Fixpunktes einer Funktion  $f$ . In diesem Abschnitt wollen wir einen anderen Typ von Fixpunktsatz besprechen. Wir benötigen hier keine Lipschitz-Stetigkeit der Funktion  $f$ , sondern nur ihre Stetigkeit – dafür muss allerdings der Definitions- und Wertebereich von  $f$  eine konvexe Teilmenge eines Vektorraums sein, und  $f$  muss noch dazu eine Kompaktheitsvoraussetzung erfüllen. Außerdem erhalten wir nur die Existenz, aber keine Eindeutigkeit des Fixpunktes.

Wir beginnen mit einer kurzen Wiederholung des Begriffs *konvexe Menge*.

**Definition B.3.1 (Konvexe Menge).** Sei  $V$  ein Vektorraum über  $\mathbb{R}$  oder  $\mathbb{C}$ . Eine Teilmenge  $K \subseteq V$  heißt *konvex*, wenn für alle Punkte  $x, y \in K$  und alle reellen Zahlen  $\lambda \in [0, 1]$  die sogenannte *Konvexkombination*  $\lambda x + (1 - \lambda)y$  wieder in  $K$  liegt.

**Aufgabe B.3.2.** (a) Wählen Sie zwei beliebige Punkte  $x, y \in \mathbb{R}^2$  und zeichnen Sie die Menge  $\{\lambda x + (1 - \lambda)y : \lambda \in [0, 1]\}$  aller Konvexkombinationen von  $x$  und  $y$ .

(b) Ist eine Kreisscheibe im  $\mathbb{R}^2$  konvex? Ist eine Kreislinie im  $\mathbb{R}^2$  konvex?

(c) Zeichnen Sie drei verschiedene Mengen im  $\mathbb{R}^2$ , die konvex sind.

(d) Zeichnen Sie drei verschiedene Mengen im  $\mathbb{R}^2$ , die nicht konvex sind.

Für eine konvexe Menge liegen auch die Konvexkombinationen von mehr als zwei Punkten wieder in der Menge:

**Proposition B.3.3.** Sei  $V$  ein Vektorraum über  $\mathbb{R}$  oder  $\mathbb{C}$  und sei  $K \subseteq V$  konvex. Dann sind äquivalent:

(i) Die Menge  $K$  ist konvex.

(ii) Für jedes  $n \in \mathbb{N}$ , alle Punkte  $x_1, \dots, x_n \in K$  und alle Zahlen  $\lambda_1, \dots, \lambda_n \in [0, 1]$  mit der Eigenschaft  $\sum_{k=1}^n \lambda_k = 1$  liegt auch die Konvexkombination  $\sum_{k=1}^n \lambda_k x_k$  in  $K$ .

*Beweis.* Dies ist eine gute Übungsaufgabe, wenn Sie ihr Verständnis des Konvexitätsbegriffes auffrischen wollen!  $\square$

Die folgende wichtige Aussage ist der sogenannte *Fixpunktsatz von Brouwer*:

**Satz B.3.4 (Brouwerscher<sup>2</sup> Fixpunktsatz).** Sei  $d \in \mathbb{N}$ , sei  $\emptyset \neq K \subseteq \mathbb{R}^d$  konvex und abgeschlossen und sei  $f : K \rightarrow K$  stetig. Dann gibt es ein  $x \in K$  mit  $f(x) = x$ .

**Aufgabe B.3.5.** Beweisen Sie den Brouwerschen Fixpunktsatz im Spezialfall  $d = 1$ .

Der Brouwersche Fixpunktsatz besitzt eine unendlich-dimensionale Verallgemeinerung. Es handelt sich um den *Fixpunktsatz von Schauder*, von welchem wir die folgende Variante im Beweis des Satzes von Peano verwenden:

**Satz B.3.6 (Schauderscher<sup>3</sup> Fixpunktsatz).** Sei  $(V, \|\cdot\|)$  ein Banachraum, sei  $\emptyset \neq K \subseteq V$  konvex und abgeschlossen und sei  $f : K \rightarrow K$  stetig und  $f(K)$  relativ kompakt. Dann gibt es ein  $x \in K$  mit  $f(x) = x$ .

**Aufgabe B.3.7.** Zeigen Sie, dass der Brouwersche Fixpunktsatz ein Spezialfall des Schauderschen Fixpunktsatzes ist.

**Literaturstellen und verwandte Ergebnisse.** (a) Für einen Beweis des Fixpunktsatzes von Brouwer verweisen wir beispielsweise auf die Referenz [SW93, Satz 16.A.3 auf Seite 539].

(b) Den Schauderschen Fixpunktsatz finden Sie zum Beispiel in [Zei86, Theorem 2.A auf Seite 56].

---

<sup>2</sup>Benannt nach Luitzen E. J. Brouwer (1881 – 1966), niederländischer Mathematiker.

<sup>3</sup>Benannt nach Juliusz Schauder (1899 – 1943), polnischer Mathematiker



# Matrix-Analysis

## Fragen zum Einstieg.

- (a) Wann nennt man eine Matrix  $A \in \mathbb{R}^{d \times d}$  diagonalisierbar? Und wozu braucht man diesen Begriff überhaupt?
- (b) Können Sie ohne Computerunterstützung die 1000-te Potenz der Matrix

$$A = \begin{pmatrix} 1.5 & 1 \\ -2 & -1.5 \end{pmatrix} \in \mathbb{R}^{2 \times 2}$$

berechnen? Können Sie mit Computerunterstützung die 1000-te Potenz von  $A$  berechnen?

- (c) Betrachten Sie eine quadratische Matrix  $A$  mit reellen (oder gerne auch komplexen) Einträgen. Wie ist die Aussage

$$A^n \rightarrow 0 \quad \text{für} \quad n \rightarrow \infty$$

zu verstehen?

Ist diese Aussage für die Matrix  $A$  aus Frage (b) wahr?

## C.1 Matrix-Normen

Betrachten Sie eine Folge von Matrizen  $A_n \in \mathbb{R}^{d \times d}$  und eine feste Matrix  $A \in \mathbb{R}^{d \times d}$ . Man sagt, dass die Folge  $(A_n)$  gegen die Matrix  $A$  konvergiert, falls jeder Eintrag von  $(A_n)$  gegen den entsprechenden Eintrag (an derselben Stelle) von  $A$  konvergiert. Sei zum Beispiel  $d = 2$  und

$$A_n = \begin{pmatrix} \frac{n+1}{n} & 0 \\ -1 + \frac{n}{n^2+n} & 2^{-n} \end{pmatrix}$$

für jedes  $n \in \mathbb{N}$ . Dann konvergiert die Folge  $(A_n)$  gegen die Matrix

$$A = \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}.$$

Wenn wir hingegen die Folge  $(\tilde{A}_n)$  mit

$$\tilde{A}_n = \begin{pmatrix} \frac{n+1}{n} & 0 \\ -1 + \frac{n}{n^2+n} & (-1)^{-n} \end{pmatrix}$$

für alle  $n \in \mathbb{N}$  betrachten, dann stellen wir fest, dass diese Folge nicht gegen eine Matrix konvergiert, denn die Folge der Einträge an der Position  $(2, 2)$  – also die Folge von Zahlen  $((-1)^{-n})$  – ist nicht konvergent.

Sie wissen bereits, dass man Konvergenz auch allgemein in metrischen – und somit insbesondere in normierten – Räumen definieren kann. Deshalb liegt es nahe, sich zu fragen, ob man auf dem  $\mathbb{R}$ -Vektorraum  $\mathbb{R}^{d \times d}$  vielleicht auch eine Norm einführen kann – und zwar derart, dass Konvergenz bezüglich dieser Norm genau die oben beschriebene eintragsweise Konvergenz ist.

Dies ist in der Tat möglich, und es gibt sogar viele verschiedene solcher Normen; wir diskutieren im Folgenden eine spezielle solche Norm, die besonders wichtig ist. Außerdem diskutieren wir sie nicht nur für den Skalkörper  $\mathbb{R}$ , sondern auch für den Skalkörper  $\mathbb{C}$ .

**Definition C.1.1 (Induzierte 2-Norm von Matrizen).** Sei  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$ . Für jeden Vektor  $x \in \mathbb{K}^d$  bezeichne

$$\|x\| := \left( \sum_{k=1}^d |x_k|^2 \right)^{1/2}$$

die Euklidische Norm von  $x$  (manchmal bezeichnet man die Euklidische Norm von  $x$  auch als *2-Norm* von  $x$ ).

Für jede Matrix  $A \in \mathbb{K}^{d \times d}$  bezeichnen wir dann die Zahl

$$\|A\| := \sup\{\|Ax\| : x \in \mathbb{K}^d, \|x\| \leq 1\} \in [0, \infty)$$

als die *induzierte 2-Norm* von  $A$ .

**Bemerkungen C.1.2.** (a) Man kann zeigen, dass tatsächlich stets  $\|A\| < \infty$  ist; dies hängt mit der Stetigkeit der linearen Abbildung  $\mathbb{K}^d \ni x \mapsto Ax \in \mathbb{K}^d$  zusammen.

- (b) Anstelle des Supremums in der Definition von  $\|A\|$  darf man aus Kompaktheitsgründen auch ein Maximum setzen.
- (c) Man kann zur Definition von Matrix-Normen auch andere Normen auf  $\mathbb{K}^d$  als die Euklidische Norm verwenden (und erhalten dann natürlich jeweils eine andere Matrix-Norm); dies diskutieren wir an dieser Stelle aber nicht weiter.
- (d) Man kann auf dieselbe Weise auch die induzierte 2-Norm von nicht-quadratischen Matrizen (d.h. von Matrizen  $A \in \mathbb{K}^{d_1 \times d_2}$  zeigen); das ist nicht schwieriger, aber wir benötigen es in dieser Vorlesung nicht.
- (e) Auf dieselbe Weise kann man auch die sogenannte *Operatornorm* von stetigen linearen Abbildungen zwischen normierten Räumen definieren; dies ist Teil der *Funktionalanalysis*.

Im Folgenden führen wir einige Eigenschaften der induzierten 2-Norm von Matrizen auf.

**Proposition C.1.3.** Sei  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$ . Für die induzierte 2-Norm auf  $\mathbb{K}^{d \times d}$  gilt:

- (a) Sie ist tatsächlich eine Norm auf  $\mathbb{K}^{d \times d}$  (und da  $\mathbb{K}^{d \times d}$  endlich-dimensional ist, ist dieser Raum zusammen mit der induzierten 2-Norm somit ein Banachraum).
- (b) Ein Folge  $(A_n) \subseteq \mathbb{K}^{d \times d}$  konvergiert bezüglich der induzierten 2-Norm gegen eine Matrix  $A \in \mathbb{K}^{d \times d}$  genau dann, wenn die Folge eintragsweise gegen  $A$  konvergiert.
- (c) Die induzierte 2-Norm ist verträglich mit der Euklidischen Norm auf  $\mathbb{K}^d$ , d.h. für alle  $A \in \mathbb{K}^{d \times d}$  und alle  $x \in \mathbb{K}^d$  gilt  $\|Ax\| \leq \|A\| \|x\|$ .
- (d) Die induzierte 2-Norm der Einheitsmatrix in  $\mathbb{K}^{d \times d}$  ist gleich 1.
- (e) Die induzierte 2-Norm ist submultiplikativ, d.h. für alle  $A, B \in \mathbb{K}^{d \times d}$  gilt  $\|AB\| \leq \|A\| \|B\|$ .  
Insbesondere gilt für alle  $A \in \mathbb{R}^{d \times d}$  und für alle  $n \in \mathbb{N}_0$  die Ungleichung  $\|A^n\| \leq \|A\|^n$ .
- (f) Man kann die induzierte 2-Norm folgendermaßen ausrechnen: Für jedes  $A \in \mathbb{K}^{d \times d}$  ist  $\|A\|$  gleich der Wurzel aus dem größten Eigenwert der Matrix  $\overline{A}^T A$  (beachten Sie, dass die Matrix  $\overline{A}^T A$  stets selbstadjungiert und positiv semi-definit ist, sodass es tatsächlich sinnvoll ist, von ihrem größten Eigenwert zu sprechen).<sup>1,2</sup>
- (g) Man kann die induzierte 2-Norm durch eine leicht zu berechnende Größe nach oben abschätzen: Ist  $A \in \mathbb{K}^{d \times d}$  und bezeichnet  $A_{j,k}$  den Eintrag von  $A$  an der Stelle  $(j, k)$ , so gilt

$$\|A\| \leq \left( \sum_{j,k=1}^d |A_{j,k}|^2 \right)^{1/2} \leq \sum_{j,k=1}^d |A_{j,k}|.$$

Beachten Sie bitte, dass man die induzierte 2-Norm einer Matrix  $A \in \mathbb{K}^{d \times d}$  zwar aus den Eigenwerten von  $\overline{A}^T A$  bestimmen kann (indem man die Wurzel des größten von ihnen nimmt), aber im Allgemeinen nicht direkt aus den Eigenwerten von  $A$ . Zum Beispiel besitzt die Matrix

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in \mathbb{C}^{2 \times 2}$$

<sup>1</sup>Hierbei benutzen wir die folgende Notation: Für jedes  $A \in \mathbb{K}^d$  bezeichnen wir mit  $\overline{A} \in \mathbb{K}^d$  diejenige Matrix, deren Einträge die konjugiert-komplexen Zahlen der Einträge von  $A$  sind.

<sup>2</sup>Dass man die induzierte 2-Norm im Prinzip ausrechnen *kann* bedeutet nicht unbedingt, dass es immer sinnvoll oder zielführend ist, sie tatsächlich konkret auszurechnen.

nur den Eigenwert 0, aber ihre induzierte 2-Norm  $\|A\|$  ist laut Proposition C.1.3(f) gleich der Wurzel des größten Eigenwerts der Matrix

$$\overline{A}^T A = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

also gleich 1.

Wir schließen diesen Abschnitt mit der folgenden Bemerkung zu einer Subtilität bezüglich des Zusammenhangs zwischen induzierter Matrixnorm und Skalarkörper:

**Bemerkung C.1.4.** Betrachten Sie eine quadratische Matrix  $A$  mit reellen Einträgen, d.h.  $A \in \mathbb{R}^{d \times d}$ . Die induzierte 2-Norm von  $A$  ist laut Definition C.1.1 (mit  $\mathbb{K} = \mathbb{R}$ ) gleich

$$\|A\| := \sup\{\|Ax\| : x \in \mathbb{R}^d, \|x\| \leq 1\} \in [0, \infty).$$

Nun ist jede Matrix in  $\mathbb{R}^{d \times d}$  aber natürlich auch ein Element von  $\mathbb{C}^{d \times d}$ ; wenn wir aber  $A$  als Element von  $\mathbb{C}^{d \times d}$  auffassen – sowie man dies mit reellen Matrizen häufig tut – dann erhalten wir aus Definition C.1.1 (mit  $\mathbb{K} = \mathbb{C}$ ) die Formel

$$\|A\| := \sup\{\|Ax\| : x \in \mathbb{C}^d, \|x\| \leq 1\} \in [0, \infty)$$

für die induzierte 2-Norm von  $A$ . Wir haben also zwei verschiedene Definitionen für die induzierte 2-Norm einer reellen Matrix  $A$  – je nachdem, ob wir als Skalarkörper  $\mathbb{R}$  oder  $\mathbb{C}$  wählen.<sup>3</sup>

Dies ist aber nur ein scheinbares Problem; man kann nämlich zeigen, dass die beiden obenstehenden Werte für  $\|A\|$  immer übereinstimmen.<sup>4</sup>

## C.2 Diagonalisierbarkeit und die Jordansche Normalform

Das Diagonalisieren von Matrizen sowie ihre Jordan-Normalform sind Konzepte aus der Linearen Algebra, die wir an zwei Stellen in der Vorlesung benötigen: Zum einen zur Berechnung von Matrix-Exponentialfunktionen in Abschnitt 5.5, und zum anderen zum Studium des Langzeitverhaltens der Lösungen von linearen Differentialgleichungen in Abschnitt 8.1. Deshalb

---

<sup>3</sup>Für Matrizen mit mindestens einem nicht-reellen Eintrag gibt es diese Ambiguität hingegen nicht: Denn wenn nicht alle Einträge der Matrix  $A$  reell sind, kann man  $A$  nicht als Element von  $\mathbb{R}^{d \times d}$  auffassen, also ist für solche Matrizen in Definition C.1.1 immer  $\mathbb{K} = \mathbb{C}$  zu wählen.

<sup>4</sup>Allerdings ist das keineswegs offensichtlich; einen Beweis findet man zum Beispiel in [Fen98, Proposition 2.1.1] (die dort bewiesene Aussage ist sogar noch deutlich allgemeiner).

wiederholen wir hier einige wichtige Resultate zur Diagonalisierung und zur Jordan-Normalform noch einmal. Der besseren Übersicht halber unterteilen wir den Abschnitt in zwei (nicht-nummerierte) Unterabschnitte.

### Diagonalisierbarkeit

Wir erinnern zunächst an die folgenden Begriffe: Sei  $A \in \mathbb{C}^{d \times d}$  und sei  $\lambda \in \mathbb{C}$  ein Eigenwert von  $A$ . Die *geometrische Vielfachheit* des Eigenwertes  $\lambda$  ist per Definition die Dimension des *Eigenraums*

$$\begin{aligned} \ker(\lambda \text{id} - A) &= \{y \in \mathbb{R}^d : Ay = \lambda y\} \\ &= \{y \in \mathbb{R}^d : y \text{ ist Eigenvektor von } A \text{ zum Eigenwert } \lambda\} \cup \{0\}. \end{aligned}$$

Die *algebraische Vielfachheit* des Eigenwertes  $\lambda$  ist per Definition die Vielfachheit von  $\lambda$  als Nullstelle des charakteristischen Polynoms von  $A$ .<sup>5</sup> Die algebraische Vielfachheit ist immer mindestens so groß wie die geometrische Vielfachheit; wenn die beiden Vielfachheiten übereinstimmen, dann heißt der Eigenwert  $\lambda$  *semi-simpel*<sup>6</sup>.

Die Matrix  $A$  heißt *diagonalisierbar*, wenn sie *ähnlich* zu einer Diagonalmatrix ist, d.h. präziser gesprochen, wenn es eine Diagonalmatrix  $D \in \mathbb{C}^{d \times d}$  und eine invertierbare Matrix  $T \in \mathbb{C}^{d \times d}$  mit der Eigenschaft  $A = TDT^{-1}$  gibt. Der folgende Satz charakterisiert Diagonalisierbarkeit auf verschiedene Weisen:

**Satz C.2.1.** *Für jede Matrix  $A \in \mathbb{C}^{d \times d}$  sind die folgenden Aussagen äquivalent:*

- (i) *Die Matrix  $A$  ist diagonalisierbar.*
- (ii) *Jeder Eigenwert von  $A$  ist semi-simpel.*
- (iii) *Es gibt eine Basis von  $\mathbb{C}^d$ , die aus Eigenvektoren von  $A$  besteht.*

Wenn  $A$  diagonalisierbar ist und somit  $A = TDT^{-1}$  für eine Diagonalmatrix  $D \in \mathbb{C}^{d \times d}$  und eine invertierbare Matrix  $T \in \mathbb{C}^{d \times d}$  gilt, dann stehen auf der Diagonalen von  $D$  immer die Eigenwerte von  $A$  – jeder Eigenwert so oft, wie es seine geometrische Vielfachheit<sup>7</sup> vorgibt – und die Spalten von  $T$  sind zugehörige Eigenvektoren. Die Reihenfolge der Eigenwerte auf der Diagonalen von  $D$  ist dabei im Prinzip willkürlich; man muss lediglich darauf achten, dass sie zur Reihenfolge der zugehörigen Eigenvektoren in  $T$  passt.

<sup>5</sup>Zur Erinnerung: Das *charakteristische Polynom*  $p_A$  von  $A$  ist gegeben durch  $p_A(\lambda) = \det(\lambda I - A)$  für alle  $\lambda \in \mathbb{C}$ .

<sup>6</sup>Oder auch *halb-einfach*.

<sup>7</sup>Die wegen der Diagonalisierbarkeit von  $A$  gleich der algebraischen Vielfachheit ist.

### Jordan-Normalform

Matrizen, die nicht diagonalisierbar sind, kann man trotzdem immer mittels Ähnlichkeitstransformation in eine besonders einfache Form bringen, nämlich in die sogenannte *Jordan-Normalform*<sup>8</sup>. Präziser gesprochen bedeutet das:

Sei  $A \in \mathbb{C}^{d \times d}$ . Dann kann man  $A$  in der Form

$$A = TJT^{-1},$$

schreiben, wobei  $T \in \mathbb{C}^{d \times d}$  invertierbar ist, und wobei  $J \in \mathbb{C}^{d \times d}$  nun nicht unbedingt eine Diagonalmatrix ist, sondern allgemeiner eine *Block-Diagonalmatrix*, deren Diagonalblöcke sogenannte *Jordan-Blöcke* sind. Dies bedeutet, dass  $J$  von der Form

$$J = \begin{pmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_m \end{pmatrix}$$

ist, mit Matrizen  $J_1 \in \mathbb{C}^{d_1 \times d_1}, \dots, J_m \in \mathbb{C}^{d_m \times d_m}$ , und die  $k$ -te Block-Matrix  $J_k$  hat die spezielle Gestalt

$$J_k = \begin{pmatrix} \lambda_k & 1 & 0 & \cdots & 0 \\ 0 & \lambda_k & 1 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & & & \lambda_k & 1 \\ 0 & \cdots & \cdots & 0 & \lambda_k \end{pmatrix}.$$

Die Zahlen  $\lambda_1, \dots, \lambda_m$ , von denen die  $k$ -te  $d_k$ -mal auf der Diagonalen des Jordanblocks  $J_k$  steht, sind dabei genau die Eigenwerte von  $A$ .

An der Jordan-Normalform kann man auch die algebraische und geometrische Vielfachheit eines Eigenwertes ablesen:

**Proposition C.2.2.** *Sei  $A \in \mathbb{C}^{d \times d}$  und sei  $\lambda \in \mathbb{C}$  ein Eigenwert von  $A$ . Dann gilt:*

- Die geometrische Vielfachheit von  $\lambda$  ist gleich der Anzahl der Jordanblöcke, bei denen  $\lambda$  auf der Diagonalen steht.*
- Die algebraische Vielfachheit von  $\lambda$  ist die Summe der Dimensionen aller Jordanblöcke, bei denen  $\lambda$  auf der Diagonalen steht.*
- Der Eigenwert  $\lambda$  ist genau dann semi-simpel, wenn alle Jordanblöcke, bei denen  $\lambda$  auf der Diagonalen steht, die Größe  $1 \times 1$  besitzen.*

Insbesondere gilt somit: Die Matrix  $A \in \mathbb{C}^{d \times d}$  ist genau dann diagonalisierbar, wenn in ihrer Jordan-Normalform alle Jordanblöcke die Größe  $1 \times 1$  besitzen (d.h. also wenn die Jordan-Normalform eine Diagonalmatrix ist).

<sup>8</sup>Benannt nach Camille Jordan (1838 – 1922), französischer Mathematiker.

## Literaturverzeichnis

- [Arn92] Vladimir I. Arnol'd. *Ordinary differential equations. Transl. from the 3rd Russian ed. by Roger Cooke.* Berlin etc.: Springer-Verlag, 1992.
- [Fen98] G. Fendler. On dilations and transference for continuous one-parameter semigroups of positive contractions on  $\mathcal{L}^p$ -spaces. *Ann. Univ. Sarav., Ser. Math.*, 9(1):1–97, 1998.
- [Har02] Philip Hartman. *Ordinary differential equations. 2nd ed., unabridged, corrected republication of the 1982 original.*, volume 38. Philadelphia, PA: SIAM, 2nd ed., unabridged, corrected republication of the 1982 original edition, 2002.
- [Heu04] Harro Heuser. *Gewöhnliche Differentialgleichungen. Einführung in Lehre und Gebrauch.* Stuttgart: Teubner, 4th revised ed. edition, 2004.
- [HS74] Morris W. Hirsch and Stephen Smale. *Differential equations, dynamical systems, and linear algebra.* New York-San Francisco-London: Academic Press, 1974.
- [Kab11] Winfried Kabbalo. *Grundkurs Funktionalanalysis.* Heidelberg: Spektrum Akademischer Verlag, 2011.
- [PSZ08] Jan W. Prüß, Roland Schnaubelt, and Rico Zacher. *Mathematische Modelle in der Biologie.* Basel: Birkhäuser, 2008.
- [PW19] Jan W. Prüß and Mathias Wilke. *Gewöhnliche Differentialgleichungen und dynamische Systeme.* Basel: Birkhäuser, 2019. Zweite Auflage.
- [SW93] Uwe Storch and Hartmut Wiebe. *Lehrbuch der Mathematik. Für Mathematiker, Informatiker und Physiker. Band III: Analysis mehrerer Veränderlicher - Integrationstheorie.* Mannheim: B.I. Wissenschaftsverlag, 1993.
- [Tes12] Gerald Teschl. *Ordinary differential equations and dynamical systems.*, volume 140. Providence, RI: American Mathematical Society (AMS), 2012.
- [Zei86] Eberhard Zeidler. *Nonlinear functional analysis and its applications. I: Fixed-point theorems. Transl. from the German by Peter R. Wadsack.* New York etc.: Springer-Verlag, 1986.