

Höhere Analysis für Lehramtsstudierende

Univ.-Prof. Dr. Ansgar Jüngel
Institut für Analysis und Scientific Computing
Technische Universität Wien

Winter 2014

20. Januar 2015

Für dieses Manuskript wurde u.a. folgende Literatur verwendet: H.-W. Alt: Funktionalanalysis (Springer, 2006), J. Frehse: Einführung in die Funktionalanalysis (Universität Bonn, 1999); J. Baumeister: Differentialgleichungen (Universität Frankfurt, 1999).

Es handelt sich um eine vorläufige Version; bitte senden Sie Fehler an juengel@tuwien.ac.at.

Inhaltsverzeichnis

1	Motivation	3
2	Funktionalanalysis	8
2.1	Metrische Räume	8
2.2	Normierte Räume	14
2.3	Prä-Hilberträume	21
2.4	Orthogonalsysteme	28
2.5	Lineare Operatoren	34
2.6	Dualräume und adjungierte Operatoren	41
2.7	Eigenwertprobleme	48
3	Parameterintegrale	54
3.1	Definition und Eigenschaften	54
3.2	Laplace-Transformation	58
4	Gewöhnliche Differentialgleichungen	64
4.1	Differentialgleichungen erster Ordnung	64
4.2	Existenz und Eindeutigkeit von Lösungen	72
4.3	Lineare Differentialgleichungen	81
4.4	Lineare Differentialgleichungen mit konstanten Koeffizienten	87
4.5	Stabilität linearer autonomer Systeme	93
4.6	Stabilität nichtlinearer autonomer Systeme	99
4.7	Räuber-Beute-Modelle	108
4.8	Diskrete Dynamik	111
5	Partielle Differentialgleichungen erster Ordnung	119
5.1	Die Burgers-Gleichung	119
5.2	Methode der Charakteristiken	123

1 Motivation

In der Höheren Analysis für LehramtskandidatInnen werden ausgewählte Kapitel aus der Analysis betrachtet, die auch für den Wahlpflichtunterricht in Schulen oder für Projektgruppen mit SchülerInnen von Interesse sein können. Beispiele, die möglicherweise für den fortgeschrittenen Schulunterricht verwendet werden können, sind mit dem Vermerk "Schule" am Rand markiert (siehe rechts).

Schule

Funktionalanalysis. Eine der Aufgaben in der linearen Algebra ist die Bestimmung der Lösbarkeit von linearen Gleichungssystemen der Form

$$Ax = b,$$

wobei $A \in \mathbb{R}^{m \times n}$ eine Matrix und $b \in \mathbb{R}^m$ ein Vektor seien. Wichtig ist in diesem Zusammenhang, dass es sich um *endlich viele* Gleichungen handelt. Nun gibt es Situationen, in denen die Gleichung $Ax = b$ in *unendlich vielen* Dimensionen gelöst werden muss.

Dazu betrachten wir folgendes Beispiel: Gegeben sei eine Reihe von (unendlich vielen) Feldern mit Inhalt a_n , $n \in \mathbb{N}$. Von einem Zeitschritt zum nächsten werde die Hälfte des Inhalts an das nächste Feld abgegeben, d.h. der Inhalt des Feldes Nr. $n + 1$ ist gleich $a_{n+1} + \frac{1}{2}a_n$. Da auch das Feld Nr. $n + 1$ die Hälfte seines Inhalts an das nächste Feld abgibt, ist der Inhalt von Feld Nr. $n + 1$ dann $\frac{1}{2}(a_{n+1} + a_n)$. Den Übergang von einem Zeitschritt zum nächsten können wir durch eine Matrix mit unendlich vielen Zeilen und Spalten beschreiben:

$$A = \begin{pmatrix} 1/2 & 0 & 0 & 0 & 0 & \dots \\ 1/2 & 1/2 & 0 & 0 & 0 & \dots \\ 0 & 1/2 & 1/2 & 0 & 0 & \dots \\ 0 & 0 & 1/2 & 1/2 & 0 & \dots \\ \vdots & \vdots & & & \vdots & \end{pmatrix}.$$

Falls zu Beginn nur das erste Feld den Inhalt $a_1 = 1$ enthält und alle anderen Felder leer sind, $a_n = 0$ für $n \geq 2$, so können wir dies durch einen Vektor

$$x = (1, 0, 0, 0, 0, \dots)^\top$$

darstellen. (Der Ausdruck x^\top ist der transponierte Vektor; ist x ein Spaltenvektor, so ist x^\top ein Zeilenvektor.) Die Verteilung nach dem ersten Zeitschritt ist dann gleich dem Matrix-Vektor-Produkt Ax , in unserem Fall

$$Ax = (\frac{1}{2}, \frac{1}{2}, 0, 0, 0, \dots)^\top.$$

Eine Frage ist, ob es stationäre Zustände gibt, d.h., Vektoren x , so dass $Ax = x$. Mit anderen Worten: Besitzt A den Eigenwert eins? Gibt es andere reelle Eigenwerte?

Ein anderes Beispiel sind sogenannte Differentialgleichungen, z.B. $u''(x) = b(x)$ für $x \in (0, 1)$. (Eine *Differentialgleichung* ist eine Gleichung für eine Funktion u , in der auch Ableitungen von u auftreten, also Gleichungen der Form $F(x, u, u', u'', \dots) = 0$.) Die entsprechende Gleichung kann ähnlich wie oben als

$$Au = b$$

formuliert werden, wobei $Au = u''$ ein sogenannter *Differentialoperator* und $b(x)$ eine Funktion seien. Gesucht sind etwa alle Funktionen $u \in C^2(\mathbb{R})$, die der obigen Gleichung genügen. Es stellt sich die Frage, ob diese Gleichung überhaupt lösbar ist und wenn ja, wie viele Lösungen existieren.

In unendlichdimensionalen Räumen gibt es einige verblüffende Effekte. Aus der linearen Algebra wissen wir beispielsweise, dass surjektive Abbildungen $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ auch injektiv sind, also einen Kern $N(A) = \{x \in \mathbb{R}^n : Ax = 0\}$ besitzen, der nur aus dem Nullvektor $x = 0$ besteht. In unendlichdimensionalen Räumen muss dies nicht gelten. Betrachte als Beispiel den Raum aller reellen Folgen $X = \{x = (x_n) : x_n \in \mathbb{R}, n \in \mathbb{N}\}$ und die Abbildung $A : X \rightarrow X$, definiert durch

$$Ax = A(x_1, x_2, x_3, \dots) = (x_2, x_3, x_4, \dots), \quad x \in X.$$

Man nennt diese Abbildung den *Linksshift*. Diese Abbildung ist surjektiv, denn zu $y = (y_n) \in X$ liefert $x = (0, y_1, y_2, \dots)$ den Bildvektor $Ax = (y_1, y_2, \dots) = y$. Andererseits ist diese Abbildung *nicht* injektiv, da $A(z, 0, 0, \dots) = (0, 0, \dots)$ für alle $z \in \mathbb{R}$ gilt. Der Kern von A besteht also aus allen Folgen $(x, 0, 0, \dots)$, $x \in \mathbb{R}$:

$$N(A) = \{x = (x_n) : x_1 \in \mathbb{R}, x_n = 0, n \geq 2\}.$$

Ein weiteres ungewohntes Ergebnis lautet, dass Abbildungen, die in unendlichdimensionalen Räumen definiert sind, keine Eigenwerte besitzen müssen. Seien hierfür $C^0([0, 1])$ der Raum aller stetigen Funktionen auf $[0, 1]$ und $(Af)(x) = (x+1)f(x)$ eine Abbildung, die einer Funktion f das Produkt aus der Abbildung $x \mapsto x+1$ und f zuordnet. Man sagt auch, dass A ein *Multiplikationsoperator* ist. Wir behaupten, dass A *keinen* Eigenwert besitzt. Angenommen, A würde doch einen Eigenwert $\lambda \in \mathbb{R}$ und einen Eigenvektor $f \neq 0$ besitzen. Dann existiert ein $x_0 \in [0, 1]$, so dass $f(x_0) \neq 0$. Wegen der Stetigkeit von f existiert eine Umgebung $U(x_0)$ von x_0 , so dass $f(x) \neq 0$ für alle $x \in U(x_0)$. Aus der Eigenwertgleichung

$$\lambda f(x) = (Af)(x) = (x+1)f(x) \quad \text{für alle } x \in [0, 1]$$

folgt dann wegen $f(x) \neq 0$, dass $\lambda = x+1$ für alle $x \in U(x_0)$. Dies kann aber nicht stimmen, und die Annahme, dass A einen Eigenwert besitzt, ist falsch.

Mit Hilfe funktionalanalytischer Hilfsmittel wollen wir folgende Fragestellungen untersuchen:

- Welche Eigenschaften besitzen unendlichdimensionale Räume?
- Welche Eigenschaften haben lineare Abbildungen, die auf diesen Räumen definiert sind?
- Unter welchen Voraussetzungen sind Gleichungen der Form $Au = b$ lösbar?

Parameterintegrale. Parameterintegrale sind Integrale, deren Integrand von einer Variablen (Parameter) abhängt. Es handelt sich also um Funktionen der Form

$$F(x) = \int_a^b f(x, y) dy.$$

Ein wichtiges Beispiel ist die *Laplace-Transformation*

$$(\mathcal{L}[f])(s) = \int_0^{\infty} e^{-st} f(t) dt, \quad s \in \mathbb{C},$$

die in der Elektrotechnik eine große Rolle spielt. Mit ihrer Hilfe können Differentialgleichungen (siehe oben) in algebraische Gleichungen verwandelt werden, deren Lösung häufig einfach möglich ist. Als Beispiel betrachten wir die Spannung $U(t)$, die an einer Spule mit der Induktivität $L > 0$ anfällt. Ist $I(t)$ der Strom, der durch die Spule fließt, so erhalten wir die Spannung aus der Gleichung

$$U(t) = L \frac{dI}{dt}(t), \quad t \geq 0. \quad (1.1)$$

Wir nehmen an, dass zur Zeit $t = 0$ kein Strom fließt: $I(0) = 0$. Wenden wir die Laplace-Transformation auf diese Gleichung an, so folgt

$$\mathcal{L}[U](s) = L \mathcal{L}\left[\frac{dI}{dt}\right](s) = L \int_0^{\infty} e^{-st} \frac{dI}{dt}(t) dt.$$

Mit partieller Integration ergibt sich für $s \neq 0$

$$\mathcal{L}[U](s) = -L \int_0^{\infty} (-s) e^{-st} I(t) dt - [e^{-st} I(t)]_0^{\infty} = sL \int_0^{\infty} e^{-st} I(t) dt = sL \mathcal{L}[I](s). \quad (1.2)$$

Die Laplace-Transformation hat die Ableitung dI/dt in das (einfachere) Produkt $s\mathcal{L}[I]$ verwandelt. Man nennt (1.1) den *Zeitbereich* und (1.2) den *Spektralbereich*.

Die Definition von Parameterintegralen wie die Laplace-Transformation wirft nun einige Fragen auf:

- Unter welchen Voraussetzungen ist das Parameterintegral $F(x) = \int_a^b f(x,y) dy$ definiert, stetig oder sogar differenzierbar?
- Für welche Funktionen f ist die Laplace-Transformation $\mathcal{L}[f]$ definiert? Welche Eigenschaften besitzt die Laplace-Transformation?

Gewöhnliche Differentialgleichungen. Differentialgleichungen sind von enormer Bedeutung in den Naturwissenschaften (aber auch in den Wirtschaftswissenschaften), um zeitliche Vorgänge zu beschreiben und vorherzusagen. Allgemein sind gewöhnliche Differentialgleichungen Gleichungen der Form

$$F(t, u, u', u'', \dots) = 0,$$

und die Funktion $u(t)$ ist gesucht. Als einfaches Beispiel betrachten wir das Wachstum einer Population. Sei $u(t)$ die Anzahl der fortpflanzungsfähigen Lebewesen. Genauer gesagt sei $u(t)$ die Populationsdichte, also die Anzahl der Lebewesen pro Flächeneinheit. Ist diese Anzahl sehr groß, macht es Sinn, $u(t)$ als reellwertig zu betrachten. Wir erwarten, dass die Population umso stärker wächst, je mehr Lebewesen vorhanden sind, die sich vermehren können. Die Rate, mit

der die Population pro Zeiteinheit wächst, du/dt , ist im Allgemeinen proportional zu der Anzahl u der Lebewesen. Bezeichnen wir mit a die Proportionalitätskonstante, so erhalten wir die Beziehung

$$\frac{du}{dt} = au, \quad t > 0, \quad u(0) = u_0,$$

wobei u_0 die Anzahl der Lebewesen zur Zeit $t = 0$ sei. Diese Gleichung ist eine Differentialgleichung. Sie ist einfach zu lösen: Wir suchen eine (nichttriviale) Funktion, deren Ableitung ein Vielfaches ihrer selbst ist; diese Eigenschaft erfüllt (nur) die Exponentialfunktion e^{at} . Die Lösung der obigen Gleichung lautet damit

$$u(t) = u_0 e^{at}, \quad t \geq 0.$$

Sie erfüllt nicht nur die Differentialgleichung, sondern auch den Anfangswert $u(0) = u_0$.

Häufig konkurrieren mehrere Populationen miteinander. Sind beispielsweise $u(t)$ und $v(t)$ die Dichten zweier Populationen (Beute und Räuber), so hängt die Wachstumsrate du/dt der Beutetiere nicht nur von der eigenen Wachstumsrate, sondern auch von der Anzahl der Räubertiere ab: Je mehr Räuber existieren, umso stärker wird die Beutepopulation gejagt. Wir ersetzen daher die Wachstumsrate a durch $a - bv$:

$$\frac{du}{dt} = (a - bv)u, \quad t > 0, \quad u(0) = u_0. \quad (1.3)$$

Wenn sich die Räubertiere ausschließlich von den Beutetieren ernähren, sterben sie aus, wenn keine Beutetiere existieren. Ihre Wachstumsrate ist negativ, $dv/dt < 0$, und wir setzen an: $dv/dt = -cv$. Die Wachstumsrate steigt bei einer größeren Beutepopulation, da dann ein besseres Nahrungsangebot zur Verfügung steht, und wir addieren zu $-c$ den Anteil $+du$ hinzu:

$$\frac{dv}{dt} = (-c + du)v, \quad t > 0, \quad v(0) = v_0. \quad (1.4)$$

Die Gleichungen (1.3)-(1.4) werden die *Lotka-Volterra-Gleichungen* genannt. Sie wurden 1925 von Alfred Lotka und unabhängig 1926 von Vito Volterra formuliert. Im Gegensatz zur Gleichung $du/dt = au$ können wir nicht mehr erwarten, eine explizite Lösung von (1.3)-(1.4) zu finden. Insbesondere stellen sich die folgenden Fragen:

- Unter welchen Voraussetzungen existieren Lösungen von Differentialgleichungen? Sind diese Lösungen eindeutig?
- Wie können Lösungen explizit berechnet werden? Falls dies nicht möglich ist: Wie verhalten sich die Lösungen qualitativ?
- Existiert der Langzeitgrenzwert $\lim_{t \rightarrow \infty} u(t)$ einer Lösung $u(t)$ einer Differentialgleichung?

Partielle Differentialgleichungen erster Ordnung. Partielle Differentialgleichungen treten auf, wenn die Funktionen von mehreren Variablen abhängen. Wir nennen diese Gleichungen von *erster Ordnung*, wenn höchstens erste Ableitungen auftreten:

$$F\left(x_1, \dots, x_n, u, \frac{\partial u}{\partial x_1}, \dots, \frac{\partial u}{\partial x_n}\right) = 0.$$

Da es hiervon eine Vielzahl möglicher Klassen von Differentialgleichungen mit jeweils unterschiedlichen Eigenschaften gibt, werden wir uns hier nur auf eine in den Anwendungen besonders wichtige Klasse von Gleichungen beschränken, nämlich auf sogenannte Erhaltungsgleichungen der Form

$$u_t + f(u)_x = 0, \quad x \in \mathbb{R}, \quad t > 0, \quad u(\cdot, 0) = u_0, \quad (1.5)$$

wobei u_t bzw. $f(u)_x$ die partiellen Ableitungen nach t bzw. x bedeuten und $f(u)$ eine beliebige Funktion sei. Derartige Gleichungen spielen eine wichtige Rolle zur Beschreibung des Verkehrsflusses auf einer Straße mit nur einer Fahrbahn. Wir begründen in Kapitel 5, dass wir in dieser Situation

$$\rho_t + (\rho v)_x = 0$$

schreiben können, wobei $\rho = \rho(x, t)$ die Fahrzeugdichte und $v = v(x, t)$ die Geschwindigkeit der Fahrzeuge sei. Diese Gleichung wird *Erhaltungsgleichung* genannt, weil die Anzahl der Fahrzeuge erhalten bleibt (wir begründen dies später). Wir nehmen an, dass die Geschwindigkeit eine Funktion der Fahrzeugdichte ist. Ist die Fahrbahn leer, so fahren die Autos mit maximaler Geschwindigkeit v_{\max} ; ist die Straße voll, so werden die Autos langsamer fahren und an einem Stauende zum Stehen kommen: $v = 0$, wenn $\rho = \rho_{\max}$. Dies ergibt z.B.

$$v(\rho) = v_{\max} \left(1 - \frac{\rho}{\rho_{\max}} \right), \quad 0 \leq \rho \leq \rho_{\max}.$$

Damit können wir die obige Gleichung in der Form (1.5) mit $u = \rho$ und $f(\rho) = v_{\max} \rho (1 - \rho/\rho_{\max})$ schreiben. Wir erwarten, dass die Autos an einem Stauende abrupt abbremsen müssen; können dann die Lösungen dieser Differentialgleichung differenzierbar sein? Wir werden in Kapitel 5 zeigen, dass es in der Tat *nicht* differenzierbare Lösungen gibt. Dies impliziert die folgenden Fragen:

- Wie können wir den Lösungsbegriff für (1.5) verallgemeinern, so dass er nichtdifferenzierbare Funktionen einschließt?
- Wie können wir die Lösungen von (1.5) berechnen?
- Welche Konsequenzen hat dies für das Verkehrsproblem?

2 Funktionalanalysis

2.1 Metrische Räume

Der Abstand zwischen zwei reellen Zahlen x und y ist durch den Betrag $|x - y|$ gegeben. Dieser Abstandsbegriff kann auf allgemeine Mengen verallgemeinert werden und führt auf den Begriff der Metrik. Eine Metrik gibt der Menge eine lokale und globale Struktur.

Definition 2.1 (Metrischer Raum). *Sei X eine beliebige Menge. Eine Metrik auf X ist eine Abbildung $d : X \times X \rightarrow \mathbb{R}_0^+ = [0, \infty)$ mit den Eigenschaften: Für alle $x, y, z \in X$ gilt*

- (i) $d(x, y) = 0$ genau dann, wenn $x = y$;
- (ii) $d(x, y) = d(y, x)$ (Symmetrie);
- (iii) $d(x, z) \leq d(x, y) + d(y, z)$ (Dreiecksungleichung).

Die Zahl $d(x, y)$ wird der Abstand zwischen x und y genannt. Das Paar (X, d) heißt metrischer Raum.

Beispiel 2.2. (1) Die Abbildung $d(x, y) = |x - y|$ auf $X = \mathbb{R}^n$ ist eine Metrik, wobei $|x - y|^2 = \sum_{k=1}^n (x_k - y_k)^2$ für $x = (x_1, \dots, x_n)$ und $y = (y_1, \dots, y_n)$ der euklidische Abstand ist. Die Dreiecksungleichung ist eine Konsequenz aus der *Minkowski-Ungleichung* für $1 \leq p < \infty$

$$\left(\sum_{k=1}^n |x_k + y_k|^p \right)^{1/p} \leq \left(\sum_{k=1}^n |x_k|^p \right)^{1/p} + \left(\sum_{k=1}^n |y_k|^p \right)^{1/p}, \quad x_k, y_k \in \mathbb{R}, \quad (2.6)$$

denn wir erhalten für $p = 2$

$$\begin{aligned} d(x, z) &= \left(\sum_{k=1}^n |(x_k - y_k) + (y_k - z_k)|^2 \right)^{1/2} \\ &\leq \left(\sum_{k=1}^n |x_k - y_k|^2 \right)^{1/2} + \left(\sum_{k=1}^n |y_k - z_k|^2 \right)^{1/2} = d(x, y) + d(y, z). \end{aligned}$$

(2) Die Abbildung $d(x, y) = |x^3 - y^3|$ ist eine Metrik auf \mathbb{R} , denn mit der Dreiecksungleichung für den Betrag folgt für $x, y, z \in \mathbb{R}$

$$d(x, z) = |(x^3 - y^3) + (y^3 - z^3)| \leq |x^3 - y^3| + |y^3 - z^3| = d(x, y) + d(y, z).$$

(3) Sei X eine Menge. Die Abbildung

$$d(x, y) = \begin{cases} 0 & : x = y \\ 1 & : x \neq y, \end{cases} \quad x, y \in X,$$

ist die *diskrete Metrik*. Dies ist in der Tat eine Metrik, da sie symmetrisch ist und die Dreiecksungleichung gilt.

(4) Sei $X = C_b(\mathbb{R})$ der Raum aller stetigen beschränkten Funktionen $f: \mathbb{R} \rightarrow \mathbb{R}$. Dann ist

$$d(f, g) = \sup_{x \in \mathbb{R}} |f(x) - g(x)|, \quad f, g \in X,$$

wohldefiniert und eine Metrik. (Die Metrikeigenschaften werden durch die Eigenschaften des Betrags vererbt.)

(5) Sei $X = \ell^p(a, b) = \{x = (x_n) : x_n \in (a, b), \|x\|_p < \infty\}$, wobei

$$\|x\|_p = \left(\sum_{n=1}^{\infty} |x_n|^p \right)^{1/p}, \quad 1 \leq p < \infty.$$

Dann ist

$$d(x, y) = \|x - y\|_p, \quad x, y \in X,$$

eine Metrik. Die Dreiecksungleichung entspricht der Minkowski-Ungleichung (2.6) für $n = \infty$. Für $p = \infty$ entspricht der Raum ℓ^p gerade der Menge der beschränkten Folgen. Die Abbildung $d(x, y) = \|x - y\|_\infty = \sup_{n \in \mathbb{N}} |x_n - y_n|$ ist eine Metrik auf ℓ^∞ . Die Dreiecksungleichung ist in diesem Fall eine einfache Konsequenz aus der Dreiecksungleichung für den Betrag. Im Folgenden schreiben wir $\ell^p = \ell^p(\mathbb{R})$.

(6) Im Raum $X = \mathbb{R}^2$ können wir diverse ‘‘Städtemetriken’’ betrachten. Ein Beispiel ist die New-York-Metrik, definiert durch

Schule

$$d(x, y) = |x_1 - x_2| + |y_1 - y_2|, \quad x = (x_1, y_1), y = (x_2, y_2) \in X.$$

Diese Metrik misst die Länge eines Fußweges in einer Stadt mit rechtwinkligen Straßen von einer Kreuzung x zu einer Kreuzung y (siehe Abbildung 2.1). Ein anderes Beispiel ist die Paris-Metrik, definiert durch

$$d(x, y) = |x| + |y| \quad \text{für } x \neq y, \quad d(x, y) = 0 \quad \text{für } x = y,$$

wobei $|x| = \sqrt{x_1^2 + x_2^2}$ und $x, y \in \mathbb{R}^2$. In diesem Beispiel repräsentieren die Vektoren französische Städte, wobei Paris im Mittelpunkt $(0, 0)$ liegt. Es wird angenommen, dass man über Paris fahren muss, um von einer Stadt x zu einer anderen Stadt y zu kommen. Das französische Bahnsystem funktioniert näherungsweise nach dieser Logik. Fallen Ihnen weitere ‘‘Städtemetriken’’ im \mathbb{R}^2 ein? \square

Ein Vorteil metrischer Räume ist, dass in ihnen topologische Eigenschaften wie im \mathbb{R}^n bezüglich der euklidischen Norm definiert werden können. Sei hierzu (X, d) ein metrischer Raum. Wir nennen ähnlich wie in der Analysis

$$B_r(x) = \{y \in X : d(x, y) < r\}$$

die (offene) *Kugel* um $x \in X$ mit Radius $r > 0$. Eine Menge $U \subset X$ heißt *offen*, wenn

$$\forall x \in U : \exists \varepsilon > 0 : B_\varepsilon(x) \subset U.$$

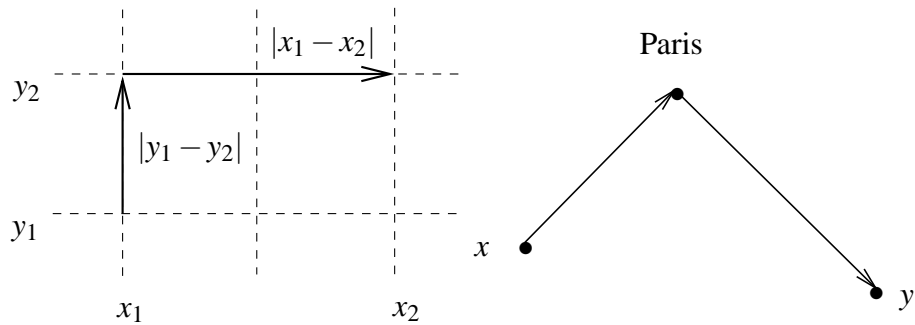


Abbildung 2.1: Illustration der New-York-Metrik (links) und der Paris-Metrik (rechts).

Eine Menge $V \subset X$ heißt *abgeschlossen*, wenn das Komplement $X \setminus V$ offen ist. Der *Abschluss* \bar{U} einer Menge $U \subset X$ ist der Durchschnitt aller abgeschlossenen Teilmengen von X , die U beinhalten:

$$\bar{U} = \bigcap_{M \supset U \text{ abgeschlossen}} M.$$

Schließlich heißt eine Menge $U \subset X$ *dicht* in X , wenn $\bar{U} = X$.

Beispiel 2.3. (1) Sei d die diskrete Metrik auf der Menge X . Dann ist

$$B_r(x) = \begin{cases} \{x\} & : r \leq 1 \\ X & : r > 1, \end{cases}$$

denn es gilt

$$\begin{aligned} B_r(x) &= \{y \in X : d(x, y) < r\} = \{y \in X : d(x, y) = 0\} = \{x\}, \quad r \leq 1, \\ B_r(x) &= \{y \in X : d(x, y) < r\} = \{y \in X : d(x, y) \leq 1\} = X, \quad r > 1. \end{aligned}$$

(2) Sei $X = \mathbb{R}^2$ und betrachte die New-York-Metrik $d(x, y) = |x_1 - x_2| + |y_1 - y_2|$ für $x = (x_1, y_1), y = (x_2, y_2) \in \mathbb{R}^2$. Dann ist $B_r(x)$ ein um $\pi/2$ gedrehtes Quadrat mit (maximalem) Durchmesser $2r$, dessen Mittelpunkt sich in x befindet. Im Falle der Paris-Metrik $d(x, y) = |x| + |y|$ ist die Kugel um $(0, 0)$ gleich einer Kreisscheibe um den Ursprung, denn: Falls $x \neq 0$, erhalten wir

$$B_r(x) = \{y \in \mathbb{R}^2 : |y| < r - |x|\}.$$

Ist also $r \leq |x|$, so folgt $B_r(x) = \emptyset$, anderenfalls ist $B_r(x)$ gleich der Kreisscheibe um $(0, 0)$ mit Radius $r - |x|$. □

Die Definition von Cauchy-Folgen und Folgenkonvergenz in metrischen Räumen ist ähnlich zu der für reelle Zahlenfolgen.

Definition 2.4. Seien (X, d) ein metrischer Raum, $(x_n) \subset X$ eine Folge und $x \in X$. Dann heißt (x_n) konvergent gegen $x \in X$ ($x_n \rightarrow x$ für $n \rightarrow \infty$), wenn

$$\forall \varepsilon > 0 : \exists N \in \mathbb{N} : \forall n \geq N : d(x_n, x) < \varepsilon.$$

In diesem Fall wird x der Grenzwert von (x_n) genannt, und wir schreiben $x = \lim_{n \rightarrow \infty} x_n$. Die Folge (x_n) heißt Cauchy-Folge, wenn

$$\forall \varepsilon > 0 : \exists N \in \mathbb{N} : \forall n, m \geq N : d(x_n, x_m) < \varepsilon.$$

Wie für reelle Zahlenfolgen wird bewiesen, dass jede Folge höchstens einen Grenzwert besitzt (Übungsaufgabe). Konvergente Folgen sind Cauchy-Folgen. Die Umkehrung gilt i.Allg. nicht; Räume, in denen diese Eigenschaft gilt, nennen wir vollständig.

Definition 2.5. Ein metrischer Raum (X, d) heißt vollständig, wenn jede Cauchy-Folge von Elementen aus X einen Grenzwert in X besitzt.

Beispiel 2.6. (1) Der Raum \mathbb{R}^n , versehen mit dem euklidischen Abstand

$$|x - y| = \left(\sum_{i=1}^n (x_i - y_i)^2 \right)^{1/2}, \quad x = (x_1, \dots, x_n), y = (y_1, \dots, y_n) \in \mathbb{R}^n,$$

ist vollständig. Dies wird ähnlich wie für den Raum \mathbb{R} bewiesen.

- (2) Ebenso ist der Raum $\ell^p([a, b])$ und $\ell^p(\mathbb{R})$, definiert in Beispiel 2.2, mit der Norm $\|\cdot\|_p$ für $1 \leq p \leq \infty$ vollständig.
- (3) Der metrische Raum der rationalen Zahlen (mit der Abstandsmetrik) ist nicht vollständig, da der Grenzwert einer Folge rationaler Zahlen nicht rational sein muss.
- (4) Der Raum \mathbb{R}^n , versehen mit der Metrik

$$d(x, y) = \frac{|x - y|}{1 + |x - y|}, \quad x, y \in \mathbb{R}^n,$$

wobei $|\cdot|$ der euklidische Abstand ist, ist vollständig. Sei nämlich $(x_k) = (x_k^1, \dots, x_k^n)$ eine Cauchy-Folge in diesem Raum. Dann ist wegen $d(x_k, x_j) \geq \frac{1}{2}|x_k - x_j|$ für $|x_k - x_j| \leq 1$ die Folge (x_n) auch eine Cauchy-Folge im \mathbb{R}^n , versehen mit dem euklidischen Abstand. Wegen der Vollständigkeit dieses Raumes konvergiert (x_n) , d.h., es existiert $x \in \mathbb{R}^n$, so dass $d_e(x_k, x) = |x_k - x| \rightarrow 0$ für $k \rightarrow \infty$. Dann folgt auch

$$d(x_k, x) = \frac{|x_k - x|}{1 + |x_k - x|} \rightarrow 0 \quad \text{für } k \rightarrow \infty,$$

d.h., (x_k) konvergiert gegen x im Raum (\mathbb{R}^n, d) . □

Ein wichtiges Resultat in vollständigen metrischen Räumen ist der Banachsche Fixpunktsatz, den wir bei der Untersuchung gewöhnlicher Differentialgleichungen anwenden werden.

Satz 2.7 (Fixpunktsatz von Banach). *Seien (X, d) ein vollständiger metrischer Raum und $S : X \rightarrow X$ eine Kontraktion, d.h., es gibt ein $k \in [0, 1)$, so dass für alle $x, y \in X$ folgt*

$$d(S(x), S(y)) \leq kd(x, y).$$

Dann besitzt die Gleichung $S(x) = x$ genau eine Lösung $\hat{x} \in X$, und es gilt für ein beliebiges $x_0 \in X$ und mit $x_n = S(x_{n-1})$ ($n \geq 1$) die Abschätzung

$$d(x_n, \hat{x}) \leq \frac{k^n}{1-k} d(x_1, x_0). \quad (2.7)$$

Der Fixpunktsatz von Banach liefert mit der Folge (x_n) zugleich ein Iterationsverfahren mit der Fehlerabschätzung (2.7).

Beweis. 1. Zuerst zeigen wir, dass (x_n) , definiert durch $x_n = S(x_{n-1})$, eine Cauchy-Folge ist. Da S eine Kontraktion ist, erhalten wir

$$d(x_{n+1}, x_n) = d(S(x_n), S(x_{n-1})) \leq kd(x_n, x_{n-1}) \leq k^2 d(x_{n-1}, x_{n-2}) \leq \dots \leq k^n d(x_1, x_0).$$

Dann folgt aus der Dreiecksungleichung, dass

$$\begin{aligned} d(x_{n+m}, x_n) &\leq d(x_{n+m}, x_{n+m-1}) + d(x_{n+m-1}, x_n) \\ &\leq d(x_{n+m}, x_{n+m-1}) + d(x_{n+m-1}, x_{n+m-2}) + \dots + d(x_{n+1}, x_n) \\ &\leq (k^{n+m-1} + k^{n+m-2} + \dots + k^n) d(x_1, x_0) \\ &= k^n (k^{m-1} + k^{m-2} + \dots + 1) d(x_1, x_0) \\ &= k^n \frac{1-k^m}{1-k} d(x_1, x_0) \leq \frac{k^n}{1-k} d(x_1, x_0). \end{aligned} \quad (2.8)$$

Zu gegebenem $\varepsilon > 0$ existiert nun ein $N \in \mathbb{N}$, so dass $(k^N / (1-k)) d(x_1, x_0) < \varepsilon$. Dann ist für alle $n \geq N, m \geq 0$: $d(x_{n+m}, x_n) < \varepsilon$. Also ist (x_n) eine Cauchy-Folge. Wegen der Vollständigkeit von X konvergiert (x_n) , und wir setzen $\hat{x} = \lim_{n \rightarrow \infty} x_n$.

2. Wir behaupten, dass \hat{x} der gesuchte Fixpunkt ist. Dazu schreiben wir

$$d(\hat{x}, S(\hat{x})) \leq d(\hat{x}, x_n) + d(x_n, S(\hat{x})) \leq d(\hat{x}, x_n) + kd(x_{n-1}, \hat{x}).$$

Im Grenzwert $n \rightarrow \infty$ geht die rechte Seite gegen null, also muss $d(\hat{x}, S(\hat{x})) = 0$ gelten. Dies impliziert $\hat{x} = S(\hat{x})$.

3. Der Fixpunkt \hat{x} ist eindeutig. Ist nämlich \bar{x} ein zweiter Fixpunkt, so folgt aus

$$d(\hat{x}, \bar{x}) = d(S(\hat{x}), S(\bar{x})) \leq kd(\hat{x}, \bar{x}),$$

dass $(1-k)d(\hat{x}, \bar{x}) \leq 0$. Wegen $1-k > 0$ ist dies nur möglich, wenn $d(\hat{x}, \bar{x}) = 0$, also $\hat{x} = \bar{x}$.

4. Schließlich führen wir in (2.8) den Grenzwert $m \rightarrow \infty$ durch, um zu erhalten

$$d(\hat{x}, x_n) \leq \frac{k^n}{1-k} d(x_1, x_0),$$

was die gewünschte Abschätzung liefert. □

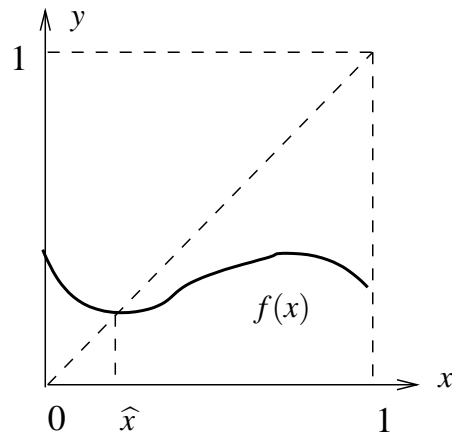


Abbildung 2.2: Eine stetige Funktion in $[0, 1]$ besitzt mindestens einen Fixpunkt \hat{x} .

Beispiel 2.8. (1) Sei $f : [0, 1] \rightarrow [0, 1]$ eine stetige Funktion. Dann ist anschaulich klar, dass der Graph von f die Diagonale im Quadrat $[0, 1]^2$ mindestens einmal schneidet (siehe Abbildung 2.2). Ist f eine Kontraktion, also $|f(x) - f(y)| = d(f(x), f(y)) < d(x, y) = |x - y|$, so gilt für alle Sekantensteigungen

$$-1 < \frac{f(x) - f(y)}{x - y} < 1,$$

also kann der Graph von f die Diagonale höchstens einmal schneiden. In diesem Fall existiert also genau ein Fixpunkt.

(2) Die Abbildung

$$S(x_1, x_2, x_3, \dots) = (x_1, x_2^2, x_3^3, \dots),$$

definiert auf $X = \ell^1([0, 1])$ und versehen mit der Metrik $d(x, y) = \|x - y\|_1$ aus Beispiel 2.2 besitzt unendlich viele Fixpunkte, nämlich

$$\hat{x} = (x, 0, 0, 0, \dots), \quad x \in [0, 1], \quad \hat{x} = (0, 1, 0, 0, \dots), \quad \hat{x} = (0, 0, 1, 0, \dots) \text{ etc.}$$

Dies steht nicht im Widerspruch zum Fixpunktsatz von Banach, weil S keine Kontraktion ist. Es folgt nämlich für $x = (1, 0, 0, \dots)$, $y = 0 \in \ell^1([0, 1])$, dass

$$d(S(x), S(y)) = \|S(x)\|_1 = 1 = \sum_{n=1}^{\infty} x_n = \|x\|_1 = d(x, y).$$

Wir bemerken, dass $d(x, y) \leq kd(S(x), S(y))$ für alle $x, y \in \ell^1([0, 1])$ nur für $k = 1$ gilt.

(3) Sei $S(x) = (2 + 3x - x^2)/3$, $x \in [1, 2]$. Wir behaupten, dass S eine Selbstabbildung ist, also der Wertebereich gleich dem Definitionsbereich ist, und dass S eine Kontraktion ist. Wir müssen zuerst zeigen, dass

$$1 \leq \frac{-x^2 + 3x + 2}{3} \leq 2 \quad \text{für all } x \in [1, 2]$$

gilt. Wir erkennen aus $S(x) = (-x^2 + 3x + 2)/3 = -(x - 3/2)^2/3 + 17/12$, dass das Minimum der Funktion in $[1, 2]$ an den Randpunkten des Intervalls angenommen wird, so dass $S(x) \geq S(1) = S(2) = 4/3 > 1$. Das Maximum von S wird an der Stelle $x = 3/2$ angenommen und $S(x) \leq S(3/2) = 17/12 < 2$. Also ist $S : [1, 2] \rightarrow [1, 2]$.

Als nächstes behaupten wir, dass S eine Kontraktion mit $k = 1/3$ ist. Seien dazu $x, y \in [1, 2]$. Wir müssen zeigen, dass

$$|S(x) - S(y)| = \frac{1}{3} \left| -\left(x - \frac{3}{2}\right)^2 + \left(y - \frac{3}{2}\right)^2 \right| \leq \frac{1}{3} |x - y|.$$

Dies ist äquivalent zu

$$|3 - (x + y)| = \left| \frac{-(x - y)(x + y) + 3(x - y)}{x - y} \right| = \left| \frac{-(x - 3/2)^2 + (y - 3/2)^2}{x - y} \right| \leq 1,$$

und dies ist wegen $2 \leq x + y \leq 4$ erfüllt. Also ist $S : [1, 2] \rightarrow [1, 2]$ eine Kontraktion, und nach dem Banachschen Fixpunktsatz existiert ein eindeutiger Grenzwert \hat{x} von $x = S(x) = (-x^2 + 3x + 2)/3$ oder $-x^2 + 2 = 0$. Also ist $\hat{x} = \sqrt{2}$. Dies können wir nachrechnen, indem wir beispielsweise für $x_0 = 1$ die Folge (x_n) berechnen:

$$x_1 = S(x_0) = \frac{4}{3}, \quad x_2 = \frac{38}{27} = 1,407\dots, \quad x_3 = \frac{3092}{2187} = 1,413\dots$$

Für $n = 10$ erhalten wir $x_{10} = 1,41421356237\dots$ und

$$|\sqrt{2} - x_{10}| \leq \frac{k^{10}}{1 - k} |x_1 - x_0| = \left(\frac{1}{3}\right)^{10} \frac{3}{2} \approx 8,467 \cdot 10^{-6}.$$

Dies garantiert, dass x_{10} der Zahl $\sqrt{2}$ bis auf fünf Nachkommastellen entspricht. Tatsächlich ist diese Abschätzung sehr vorsichtig: x_{10} ist gleich $\sqrt{2}$ sogar bis auf 11 Nachkommastellen. \square

2.2 Normierte Räume

Im vorigen Abschnitt haben wir uns auf die topologische Struktur einer Menge konzentriert. In diesem Abschnitt wollen wir eine algebraische Struktur in den Vordergrund stellen, nämlich für Vektorräume. Dazu sei X ein Vektorraum über \mathbb{R} oder \mathbb{C} (also ein Raum, für den für alle $x, y \in X$ auch $x + y \in X$, für alle $x \in X$ und $\lambda \in \mathbb{R}$ bzw. $\lambda \in \mathbb{C}$ auch $\lambda x \in X$ und die Vektorraumaxiome gelten).

Definition 2.9 (Normierter Raum). *Sei X ein Vektorraum über \mathbb{R} oder \mathbb{C} . Eine Abbildung $\|\cdot\| : X \rightarrow \mathbb{R}_0^+$ heißt Norm und $(X, \|\cdot\|)$ ein normierter Raum, wenn für alle $x, y \in X$ und $\lambda \in \mathbb{R}$ bzw. $\lambda \in \mathbb{C}$ gilt:*

- (i) $\|x\| = 0$ genau dann, wenn $x = 0$;
- (ii) $\|\lambda x\| = |\lambda| \|x\|$;

(iii) $\|x + y\| \leq \|x\| + \|y\|$.

Im Folgenden betrachten wir hauptsächlich *reelle* normierte Räume. Falls der Vektorraum über die komplexen Zahlen definiert ist, vermerken wir dies explizit.

Ist $(X, \|\cdot\|)$ ein normierter Raum, so wird mittels $d(x, y) = \|x - y\|$ eine Metrik auf X induziert. Umgekehrt kann man *nicht* aus jeder Metrik eine Norm konstruieren. Dies ist etwa bei der diskreten Metrik nicht möglich: Anderenfalls könnten wir in dem Vektorraum X $\|x\| = d(x, 0)$ setzen und erhalten für $x \neq 0$ den Widerspruch

$$\|2x\| = d(2x, 0) = 1 \neq 2 = 2d(x, 0) = 2\|x\|.$$

Beispiel 2.10. (1) Die Abbildung

$$\|x\| = \left(\sum_{k=1}^n x_k^2 \right)^{1/2}, \quad x = (x_1, \dots, x_n) \in \mathbb{R}^n,$$

ist eine Norm auf \mathbb{R}^n . Die Dreiecksungleichung ist äquivalent zur *Cauchy-Schwarz-Ungleichung* (die wir nicht beweisen)

$$\sum_{k=1}^n x_k y_k \leq \left(\sum_{k=1}^n x_k^2 \right)^{1/2} \left(\sum_{k=1}^n y_k^2 \right)^{1/2} = \|x\| \|y\|,$$

denn

$$\begin{aligned} \|x + y\|^2 &= \sum_{k=1}^n (x_k + y_k)^2 = \sum_{k=1}^n (x_k^2 + 2x_k y_k + y_k^2) \\ &\leq \|x\|^2 + 2\|x\| \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2. \end{aligned}$$

Wir nennen $\|\cdot\|$ die *euklidische Norm* auf \mathbb{R}^n .

(2) Auf dem \mathbb{R}^n sind die *Summennorm* $\|\cdot\|_1$ und die *Maximumnorm* $\|\cdot\|_\infty$,

$$\|x\|_1 = \sum_{k=1}^n |x_k|, \quad \|x\|_\infty = \max_{k=1, \dots, n} |x_k|$$

auch Normen.

(3) Betrachte den Raum $\ell^p = \{x = (x_n) \subset \mathbb{R} : \|x\|_p < \infty\}$ mit $1 \leq p < \infty$ und

$$\|x\|_p = \left(\sum_{n=1}^{\infty} |x_n|^p \right)^{1/p}$$

aus Beispiel 2.2. Dann ist $\|\cdot\|_p$ eine Norm und $(\ell^p, \|\cdot\|_p)$ ein normierter Raum. Dasselbe gilt für $(\ell^\infty, \|\cdot\|_\infty)$.

(4) Seien $\Omega \subset \mathbb{R}^n$ eine offene Menge und $1 \leq p < \infty$. Der Raum

$$L^p(\Omega) = \left\{ f : \Omega \rightarrow \mathbb{R} \text{ integrierbar} : \int_{\Omega} |f(x)|^p dx < \infty \right\}$$

mit der Abbildung

$$\|f\|_{L^p} = \left(\int_{\Omega} |f(x)|^p dx \right)^{1/p}$$

ist ein normierter Raum. Dasselbe gilt für $L^\infty(\Omega) = \{f : \Omega \rightarrow \mathbb{R} \text{ integrierbar} : \|f\|_{L^\infty} < \infty\}$ mit $\|f\|_{L^\infty} = \sup_{x \in \Omega} |f(x)|$. (Eigentlich mit dem essentiellen Supremum $\|f\|_{L^\infty} = \inf\{K \geq 0 : |f(x)| \leq K \text{ für fast alle } x \in \Omega\}$.)

(5) Seien $\Omega \subset \mathbb{R}^n$ ein beschränktes Gebiet und $X = C^1(\overline{\Omega})$ der Raum aller stetig differenzierbaren Funktionen $f : \overline{\Omega} \rightarrow \mathbb{R}$. Dann sind Funktionen aus X und ihre erste Ableitung integrierbar und

$$\|f\| = (\|f\|_{L^2}^2 + \|\nabla f\|_{L^2}^2)^{1/2} = \left(\int_{\Omega} (f^2 + |\nabla f|^2) dx \right)^{1/2}$$

eine Norm auf X . Die Dreiecksungleichung ist eine Konsequenz der Cauchy-Schwarz-Ungleichung (die wir nach diesem Beispiel beweisen)

$$\int_{\Omega} fg dx \leq \left(\int_{\Omega} f^2 dx \right)^{1/2} \left(\int_{\Omega} g^2 dx \right)^{1/2} = \|f\|_{L^2} \|g\|_{L^2}. \quad (2.9)$$

Wir rechnen

$$\begin{aligned} \|f + g\|^2 &= \int_{\Omega} (f + g)^2 dx + \int_{\Omega} |\nabla(f + g)|^2 dx \\ &= \int_{\Omega} (f^2 + 2fg + g^2) dx + \int_{\Omega} (|\nabla f|^2 + 2\nabla f \cdot \nabla g + |\nabla g|^2) dx \\ &\leq^* \|f\|^2 + 2\|f\|_{L^2} \|g\|_{L^2} + 2\|\nabla f\|_{L^2} \|\nabla g\|_{L^2} + \|g\|^2 \\ &\leq^{**} \|f\|^2 + 2\|f\| \|g\| + \|g\|^2 = (\|f\| + \|g\|)^2. \end{aligned}$$

Die Ungleichung “ \leq^* ” folgt aus der Cauchy-Schwarz-Ungleichung; “ \leq^{**} ” ist äquivalent zu

$$\|f\|_{L^2} \|g\|_{L^2} + \|\nabla f\|_{L^2} \|\nabla g\|_{L^2} \leq (\|f\|_{L^2}^2 + \|\nabla f\|_{L^2}^2)^{1/2} (\|g\|_{L^2}^2 + \|\nabla g\|_{L^2}^2)^{1/2}.$$

Um dies zu beweisen, setzen wir

$$a_1 = \|f\|_{L^2}, \quad a_2 = \|\nabla f\|_{L^2}, \quad b_1 = \|g\|_{L^2}, \quad b_2 = \|\nabla g\|_{L^2}.$$

Dann ist

$$a_1 b_1 + a_2 b_2 \leq \sqrt{a_1^2 + a_2^2} \sqrt{b_1^2 + b_2^2}$$

für alle $a_i, b_i \geq 0$ zu zeigen. Aus der Ungleichung

$$0 \leq (a_1 b_2 - a_2 b_1)^2 = a_1^2 b_2^2 - 2a_1 a_2 b_1 b_2 + a_2^2 b_1^2$$

folgt

$$2a_1 a_2 b_1 b_2 \leq a_1^2 b_2^2 + a_2^2 b_1^2.$$

Wir verwenden diese Ungleichung in der folgenden Rechnung:

$$\begin{aligned} (a_1 b_1 + a_2 b_2)^2 &= a_1^2 b_1^2 + a_2^2 b_2^2 + 2a_1 a_2 b_1 b_2 \\ &\leq a_1^2 b_1^2 + a_2^2 b_2^2 + a_1^2 b_2^2 + a_2^2 b_1^2 = (a_1^2 + a_2^2)(b_1^2 + b_2^2). \end{aligned}$$

Wurzelziehen liefert die Behauptung. Die anderen Normeigenschaften folgen sofort aus der Definition. \square

Es bleibt die Cauchy-Schwarz-Ungleichung (2.9) zu beweisen.

Lemma 2.11 (Cauchy-Schwarz-Ungleichung). *Seien $\Omega \subset \mathbb{R}^n$ eine offene Menge und $f, g \in L^2(\Omega)$. Dann gilt*

$$\int_{\Omega} f g dx \leq \left(\int_{\Omega} f^2 dx \right)^{1/2} \left(\int_{\Omega} g^2 dx \right)^{1/2}.$$

Beweis. Mit den Definitionen

$$a = \int_{\Omega} g^2 dx, \quad b = - \int_{\Omega} f g dx$$

folgt

$$\begin{aligned} 0 &\leq \int_{\Omega} (af + bg)^2 dx = a^2 \int_{\Omega} f^2 dx + 2ab \int_{\Omega} f g dx + b^2 \int_{\Omega} g^2 dx \\ &= a \left(\int_{\Omega} g^2 dx \int_{\Omega} f^2 dx - \left(\int_{\Omega} f g dx \right)^2 \right). \end{aligned}$$

Falls $a = 0$, ist nichts zu zeigen. Anderenfalls folgt die Behauptung. \square

Die Konvergenz von Folgen (x_n) definieren wir ähnlich wie im vorigen Abschnitt: Wir schreiben $x_n \rightarrow x$ für $n \rightarrow \infty$, wenn es ein $x \in X$ gibt, so dass

$$\forall \varepsilon > 0: \exists N \in \mathbb{N}: \forall n \geq N: \|x_n - x\| \leq \varepsilon.$$

Ähnlich wird der Begriff der Cauchy-Folge definiert: (x_n) heißt Cauchy-Folge, wenn

$$\forall \varepsilon > 0: \exists N \in \mathbb{N}: \forall n, m \geq N: \|x_n - x_m\| \leq \varepsilon.$$

Mit konvergenten Folgen kann wie im reellen Fall gerechnet werden. Insbesondere gilt folgendes Resultat.

Lemma 2.12. Seien $(X, \|\cdot\|)$ ein normierter Raum, $(x_n), (y_n)$ konvergente Folgen in X mit $x_n \rightarrow x$ und $y_n \rightarrow y$ für $n \rightarrow \infty$ und $(c_n) \subset \mathbb{R}$ mit $c_n \rightarrow c$. Dann gilt

$$\lim_{n \rightarrow \infty} \|x_n\| = \|x\|, \quad \lim_{n \rightarrow \infty} (x_n + y_n) = x + y, \quad \lim_{n \rightarrow \infty} (c_n x_n) = cx.$$

Beweis. Aus der Dreiecksungleichung folgt

$$|\|x_n\| - \|x\|| \leq \|x_n - x\| + \|x\| - \|x\| \leq \|x_n - x\| \rightarrow 0 \quad \text{für } n \rightarrow \infty,$$

woraus sich die erste Behauptung ergibt. Ferner ist

$$\begin{aligned} \|(x_n + y_n) - (x + y)\| &= \|(x_n - x) + (y_n - y)\| \leq \|x_n - x\| + \|y_n - y\| \rightarrow 0 \\ \|c_n x_n - cx\| &= \|c_n(x_n - x) + (c_n - c)x\| \leq |c_n| \|x_n - x\| + |c_n - c| \|x\| \rightarrow 0 \end{aligned}$$

als das Produkt zweier reeller konvergenter (und folglich beschränkter) Folgen c_n und $\|x_n - x\|$. \square

Wir nennen einen normierten Raum vollständig, wenn er mit der von der Norm erzeugten Metrik vollständig ist. Vollständige normierte Räume zeichnen wir aus.

Definition 2.13 (Banachraum). Einen vollständigen normierten Raum $(X, \|\cdot\|)$ nennen wir einen Banachraum.

Beispiel 2.14. (1) Der Raum \mathbb{R}^n , versehen mit der euklidischen Norm

$$\|x\| = \left(\sum_{i=1}^n x_i^2 \right)^{1/2}, \quad x = (x_1, \dots, x_n) \in \mathbb{R}^n,$$

ist vollständig. Dies folgt aus der Vollständigkeit der reellen Zahlen. Ist nämlich $(x_k) = (x_k^1, \dots, x_k^n)$ eine Cauchy-Folge, so ist jede Komponentenfolge $(x_k^i)_k$ wegen

$$\varepsilon \geq \|x_k - x_j\| \geq |x_k^i - x_j^i|$$

eine reelle Cauchy-Folge, also konvergent gegen x^i . Dann konvergiert (x_k) gegen $x = (x^1, \dots, x^n)$, weil

$$\|x_k - x\|^2 = \sum_{i=1}^n |x_k^i - x^i|^2 \rightarrow 0.$$

(2) Der Raum $(\ell^p, \|\cdot\|_p)$ mit $1 \leq p < \infty$, definiert in Beispiel 2.10 (3), ist ebenfalls ein Banachraum. Der Beweis ist ähnlich wie in (1).

(3) Der Raum $C^1(\overline{\Omega})$, versehen mit der Norm

$$\|f\| = \left(\int_{\Omega} (f^2 + |\nabla f|^2) dx \right)^{1/2}, \quad f \in C^1(\overline{\Omega}),$$

ist *kein* Banachraum. Als Gegenbeispiel definieren wir auf $\Omega = (-1, 1)$ die Funktionenfolge

$$f_n(x) = \sqrt{1 - \frac{1}{n}} - \sqrt{x^2 + \frac{1}{n}}, \quad n \in \mathbb{N}.$$

Die Funktion f_n ist stetig differenzierbar auf $[-1, 1]$, also ein Element von $C^1(\overline{\Omega})$. Außerdem konvergiert $f_n(x)$ für $n \rightarrow \infty$ punktweise gegen $f(x) = 1 - |x|$. Wir behaupten, dass (f_n) auch in der Norm gegen diese Funktion konvergiert. Dies rechnen wir explizit nach:

$$\begin{aligned} \int_{-1}^1 (f_n - f)^2 dx &= \frac{4}{3} \left(1 - \left(1 + \frac{1}{n}\right)^{3/2} + \left(\frac{1}{n}\right)^{3/2} \right) + \frac{2}{n} \\ &\quad + \frac{1}{n} \left(1 - \sqrt{1 + \frac{1}{n}} \right) \left(\ln \left(1 + \sqrt{1 + \frac{1}{n}} \right) - \ln \left(-1 + \sqrt{1 + \frac{1}{n}} \right) \right). \end{aligned}$$

Wegen $x \ln x \rightarrow 0$ für $x \rightarrow 0$ konvergiert dieser Ausdruck für $n \rightarrow \infty$ gegen null. Die Ableitung von f lautet $f'(x) = -1$ für $x > 0$ und $f'(x) = 1$ für $x < 0$. Eine andere längere Rechnung ergibt dann

$$\begin{aligned} \int_{-1}^1 (f'_n - f')^2 dx &= \int_{-1}^0 \left(-\frac{x}{\sqrt{x^2 + 1/n}} - 1 \right)^2 dx + \int_0^1 \left(-\frac{x}{\sqrt{x^2 + 1/n}} + 1 \right)^2 dx \\ &= 2 \int_0^1 \left(-\frac{x}{\sqrt{x^2 + 1/n}} + 1 \right)^2 dx \\ &= 2 \left(-\frac{1}{\sqrt{n}} \arctan \sqrt{n} + 2 - 2\sqrt{1 + \frac{1}{n}} + \frac{2}{\sqrt{n}} \right), \end{aligned}$$

und dies konvergiert wegen $\arctan(x)/x \rightarrow 0$ für $x \rightarrow \infty$ ebenfalls gegen null, falls $n \rightarrow \infty$. Zwar haben wir die Konvergenz von (f_n) in der gewünschten Norm (und (f_n) ist insbesondere eine Cauchy-Folge), aber leider ist der (eindeutige) Grenzwert f *kein* Element des Raumes, da f nicht differenzierbar an $x = 0$ ist. Damit kann $C^1(\overline{\Omega})$ in der obigen Norm nicht vollständig sein.

Bilden wir den Abschluss von $C^1(\overline{\Omega})$ in der obigen Norm,

$$H^1(\Omega) := \overline{C^1(\overline{\Omega})},$$

so ist der resultierende Raum *nach Konstruktion* ein Banachraum. (Genauer gesagt gilt dies, falls $\Omega \subset \mathbb{R}^n$ ein beschränktes Gebiet ist, dessen Rand sich lokal durch eine C^1 -Funktion darstellen lässt.) Der Raum $H^1(\Omega)$ (den man einen *Sobolevraum* nennt) ist die Menge aller Funktionen $f : \Omega \rightarrow \mathbb{R}$, für die eine Folge $(f_n) \subset C^1(\overline{\Omega})$ existiert mit $\|f_n - f\|_{H^1} \rightarrow 0$ für $n \rightarrow \infty$. Man kann zeigen, dass $H^1(\Omega)$ der Raum aller L^2 -Funktionen ist, deren erste Ableitung ebenfalls im Raum $L^2(\Omega)$ liegt. Um dies einzusehen, benötigt man allerdings den Begriff der schwachen Ableitung, da es zunächst nicht klar ist, was unter einer Ableitung einer L^2 -Funktion zu verstehen ist. Wir verfolgen dies hier nicht weiter und verweisen auf die Literatur über Sobolevräume.

- (4) In ähnlicher Weise wie im vorigen Beispiel definieren wir den Raum $L^2(\Omega) = \overline{C^1(\overline{\Omega})}$ (der Abschluss wird in der Norm $\|\cdot\|_{L^2}$ gebildet). Mit dieser Definition wird $L^2(\Omega)$ ein Banachraum. Er ist der Raum aller quadratintegrierbaren Funktionen. \square

In bestimmten Fällen spielt es keine Rolle, welche Norm wir auf einem Raum betrachten, sofern die Normen äquivalent sind im folgenden Sinn.

Definition 2.15 (Äquivalenz von Normen). *Zwei Normen $\|\cdot\|_1$ und $\|\cdot\|_2$ auf einem normierten Raum X heißen äquivalent, wenn Konstanten $C_1, C_2 > 0$ existieren, so dass für alle $x \in X$ gilt:*

$$C_1\|x\|_1 \leq \|x\|_2 \leq C_2\|x\|_1.$$

Ist eine Folge konvergent (divergent) in einer Norm, so ist sie auch konvergent (divergent) in jeder dazu äquivalenten Norm.

Beispiel 2.16. (1) Auf dem Raum \mathbb{R}^n sind die Normen

$$\|x\|_1 = \sum_{i=1}^n |x_i|, \quad \|x\|_2 = \left(\sum_{i=1}^n x_i^2 \right)^{1/2}, \quad \|x\|_\infty = \max_{i=1, \dots, n} |x_i|$$

äquivalent, denn

$$\max_{i=1, \dots, n} |x_i| \leq \left(\sum_{i=1}^n x_i^2 \right)^{1/2} \leq \sum_{i=1}^n |x_i| \leq n \max_{i=1, \dots, n} |x_i|.$$

Für die zweite Ungleichung haben wir die Ungleichung $(a+b)^{1/2} \leq a^{1/2} + b^{1/2}$ für $a, b \geq 0$ verwendet (genauer die auf endliche Summen verallgemeinerte Gleichung), die sofort aus

$$(\sqrt{a+b})^2 = a+b \leq a+b+2\sqrt{ab} = (\sqrt{a} + \sqrt{b})^2$$

folgt. Man kann übrigens zeigen, dass alle Normen auf \mathbb{R}^n (sogar auf allen endlichdimensionalen Räumen) äquivalent sind.

(2) Seien

$$\|f\|_1 = \int_{\Omega} |f| dx, \quad \|f\|_\infty = \max_{x \in \overline{\Omega}} |f(x)|$$

zwei Normen auf $X = C^0(\overline{\Omega})$, wobei $\Omega \subset \mathbb{R}^n$ ein beschränktes Gebiet sei. Die beiden Normen sind *nicht* äquivalent. Es gilt zwar $\|f\|_1 \leq \text{meas}(\Omega)\|f\|_\infty$; wir behaupten, dass es keine Konstante $C > 0$ gibt, so dass $\|f\|_\infty \leq C\|f\|_1$ für alle $f \in X$. Für den Beweis betrachten wir $\Omega = (0, 1)$ und definieren eine Folge

$$f_n(x) = \begin{cases} 1 - nx & : 0 < x < 1/n \\ 0 & : 1/n \leq x < 1. \end{cases}$$

Dann ist $\|f_n\|_\infty = 1$, aber

$$\|f_n\|_1 = \int_0^{1/n} (1 - nx) dx = \left[x - \frac{n}{2} x^2 \right]_0^{1/n} = \frac{1}{2n},$$

so dass $\|f_n\|_\infty = 1 = 2n\|f_n\|_1$ für alle $n \in \mathbb{N}$. Würde nun $\|f_n\|_\infty \leq C\|f_n\|_1$ mit einer von n unabhängigen Konstanten gelten, so ergäbe $2n\|f_n\|_1 = \|f_n\|_\infty \leq C\|f_n\|_1$ den Widerspruch $C \geq 2n \rightarrow \infty$ für $n \rightarrow \infty$. \square

2.3 Prä-Hilberträume

Wir erhalten eine geometrische Struktur, wenn wir Vektorräume mit einem Skalarprodukt versehen.

Definition 2.17 (Prä-Hilbertraum). *Sei H ein Vektorraum über \mathbb{R} (oder \mathbb{C}). Wir nennen H einen Prä-Hilbertraum, wenn es eine Abbildung $(\cdot, \cdot) : H \times H \rightarrow \mathbb{R}$ (oder $\rightarrow \mathbb{C}$) gibt mit den Eigenschaften*

- (i) (\cdot, \cdot) ist (konjugiert) linear in ersten Argument und linear in zweitem Argument, d.h., für alle $a, b \in \mathbb{R}$ (oder $a, b \in \mathbb{C}$) und $u, v, w \in H$ gelten

$$(au + bv, w) = \bar{a}(u, w) + \bar{b}(v, w), \quad (u, av + bw) = a(u, v) + b(u, w);$$

- (ii) (\cdot, \cdot) ist symmetrisch, d.h., für alle $u, v \in H$ gilt $(u, v) = \overline{(v, u)}$;

- (iii) (\cdot, \cdot) ist positiv definit, d.h., für alle $u \in H$ ist $(u, u) \geq 0$ und $(u, u) = 0$ genau dann, wenn $u = 0$.

In diesem Fall nennen wir (\cdot, \cdot) ein Skalarprodukt.

Beispiel 2.18. (1) Der Raum \mathbb{R}^n ist mit dem Skalarprodukt

$$(x, y) = \sum_{j=1}^n x_j y_j, \quad x, y \in \mathbb{R}^n,$$

ein Prä-Hilbertraum. Im \mathbb{C}^n wird durch

$$(x, y) = \sum_{j=1}^n \bar{x}_j y_j, \quad x, y \in \mathbb{C}^n,$$

ein Skalarprodukt definiert.

- (2) Betrachte den Raum ℓ^2 aus Beispiel 2.10. Dann ist

$$(x, y) = \sum_{n=1}^{\infty} x_n y_n, \quad x = (x_n), y = (y_n) \in \ell^2,$$

ein Skalarprodukt und ℓ^2 damit ein Prä-Hilbertraum. Falls ℓ^2 über \mathbb{C} definiert wird, müssen wir das Skalarprodukt wie im vorigen Beispiel zu $(x, y) = \sum_{n=1}^{\infty} \bar{x}_n y_n$ abändern. Für die Räume ℓ^p mit $p \neq 2$ kann allerdings kein (kanonisches) Skalarprodukt wie im Fall $p = 2$ definiert werden.

(3) Der Raum $C^0(\overline{\Omega}; \mathbb{C})$ mit einem beschränkten Gebiet $\Omega \subset \mathbb{R}^n$, versehen mit dem Skalarprodukt

$$(f, g) = \int_{\Omega} \bar{f}g dx, \quad f, g \in C^0(\overline{\Omega}; \mathbb{C}),$$

ist ein Prä-Hilbertraum. Die Linearität und Symmetrie sind klar. Die positive Definitheit gilt, weil aus

$$0 = (f, f) = \int_{\Omega} |f|^2 dx$$

folgt, dass $f = 0$ (f ist ja stetig). Die Abbildung (\cdot, \cdot) ist auch ein Skalarprodukt in dem Raum

$$L^2(\mathbb{R}^n) = \left\{ f: \mathbb{R}^n \rightarrow \mathbb{R} \text{ integrierbar} : \int_{\mathbb{R}^n} |f|^2 dx < \infty \right\}.$$

□

Auf einem Prä-Hilbertraum gelten einige (geometrische) Eigenschaften. Insbesondere sagen wir, dass die Elemente u und v *orthogonal* sind, falls $(u, v) = 0$ erfüllt ist.

Proposition 2.19. Sei H mit (\cdot, \cdot) ein Prä-Hilbertraum über \mathbb{R} mit $\|\cdot\| = \sqrt{(\cdot, \cdot)}$ und seien $u, v \in H$. Dann gilt:

- *Cauchy-Schwarz-Ungleichung:* $|(u, v)| \leq \|u\| \|v\|$;
- *Satz von Pythagoras:* Wenn u und v orthogonal sind, dann $\|u + v\|^2 = \|u\|^2 + \|v\|^2$;
- *Parallelogrammgleichung:* $\|u + v\|^2 + \|u - v\|^2 = 2(\|u\|^2 + \|v\|^2)$ (siehe Abbildung 2.3).

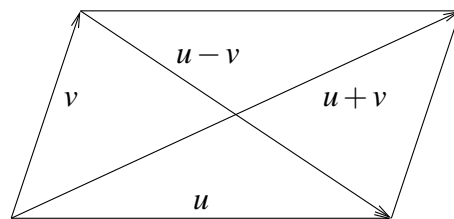


Abbildung 2.3: Illustration der Parallelogramm-Ungleichung.

Beweis. Wir zeigen nur die Cauchy-Schwarz-Ungleichung. Die anderen Aussagen sind eine Übungsaufgabe. Der Beweis der Cauchy-Schwarz-Ungleichung ist eine Verallgemeinerung des Beweises von Lemma 2.11. Falls $v = 0$, ist nichts zu zeigen. Sei also $v \neq 0$. Setze $a = \|v\|^2$ und $b = -\overline{(u, v)}$. Dann folgt

$$\begin{aligned} 0 &\leq (au + bv, au + bv) = |a|^2(u, u) + \underbrace{\bar{a}b(u, v) + a\bar{b}(\overline{u, v})}_{=2\operatorname{Re}(\bar{a}b(u, v))} + |b|^2(v, v) \\ &= \|v\|^2(\|v\|^2\|u\|^2 - 2|(u, v)|^2 + |(u, v)|^2) = \|v\|^2(\|v\|^2\|u\|^2 - |(u, v)|^2). \end{aligned}$$

Nach Division durch $\|v\|^2$ und Wurzelziehen ergibt sich die Behauptung. □

Proposition 2.20. Sei H ein Prä-Hilbertraum. Dann ist $(\cdot, \cdot) : H \times H \rightarrow \mathbb{R}$ (oder $\rightarrow \mathbb{C}$) stetig.

Beweis. Seien $u_0, v_0 \in H$ und $\varepsilon > 0$ gegeben. Sei ferner $\delta > 0$ mit $\delta(\delta + \|v_0\|) + \|u_0\|\delta < \varepsilon$. Dann erhalten wir für alle $\|u - u_0\| + \|v - v_0\| < \delta$ mit Hilfe der Cauchy-Schwarz-Ungleichung

$$\begin{aligned} |(u, v) - (u_0, v_0)| &= |(u - u_0, v) + (u_0, v - v_0)| \\ &\leq \|u - u_0\| \|v\| + \|u_0\| \|v - v_0\| \leq \delta(\delta + \|v_0\|) + \|u_0\|\delta < \varepsilon, \end{aligned}$$

denn $\|v\| \leq \|v - v_0\| + \|v_0\| < \delta + \|v_0\|$. □

Ein Prä-Hilbertraum ist vermittels der Definition $\|u\| = \sqrt{(u, u)}$ auch ein normierter Raum. Wir sagen, dass $\|\cdot\|$ die durch das Skalarprodukt induzierte Norm ist. Es ist einfach einzusehen, dass hierdurch wirklich eine Norm definiert wird: Es gilt $\|u\| \geq 0$ und $0 = \|u\|^2 = (u, u)$ genau dann, wenn $u = 0$. Ferner ist für $c \in \mathbb{R}$ (oder $c \in \mathbb{C}$) wegen der Bilinearität des Skalarprodukts $\|cu\|^2 = (cu, cu) = |c|^2(u, u) = |c|^2\|u\|^2$. Schließlich folgt aus der Cauchy-Schwarz-Ungleichung die Dreiecksungleichung

$$\begin{aligned} \|u + v\|^2 &= (u + v, u + v) = (u, u) + (u, v) + (v, u) + (v, v) \\ &\leq \|u\|^2 + 2\|u\| \|v\| + \|v\|^2 = (\|u\| + \|v\|)^2. \end{aligned}$$

Kann man umgekehrt in einem Prä-Hilbertraum das Skalarprodukt aus der induzierten Norm gewinnen? Die Antwort ist ja: Sei H ein reeller Prä-Hilbertraum. Für alle $u, v \in H$ gilt

$$(u, v) = \frac{1}{4} (\|u + v\|^2 - \|u - v\|^2),$$

wie man durch Ausmultiplizieren der rechten Seite sofort einsieht. In komplexen Prä-Hilberträumen ist diese sogenannte *Polarisationsformel* etwas komplizierter; sie lautet

$$(u, v) = \frac{1}{4} (\|u + v\|^2 - \|u - v\|^2 - i\|u + iv\|^2 + i\|u - iv\|^2).$$

Natürlich können wir *vollständige* Prä-Hilberträumen einführen; solche Räume heißen Hilberträume.

Definition 2.21 (Hilbertraum). Ein Prä-Hilbertraum H heißt Hilbertraum, wenn H mit der induzierten Norm $\|\cdot\| = \sqrt{(\cdot, \cdot)}$ vollständig ist.

Beispiel 2.22. Die Prä-Hilberträume \mathbb{R}^n und ℓ^2 aus Beispiel 2.18 sind Hilberträume. Der Raum $C^1(\bar{\Omega})$ mit dem Skalarprodukt

$$(f, g) = \int_{\Omega} (fg + \nabla f \cdot \nabla g) dx$$

ist jedoch gemäß Beispiel 2.14 (3) *nicht* vollständig und damit *kein* Hilbertraum. □

Ist V ein Unterraum des $H = \mathbb{R}^n$, so können wir den Raum \mathbb{R}^n in V und das orthogonale Komplement V^\perp zerlegen: $\mathbb{R}^n = V \oplus V^\perp$, wobei

$$V^\perp = \{w \in H : \forall v \in V : (v, w) = 0\}.$$

Dies gilt in unendlichdimensionalen Prä-Hilberträumen allerdings nicht notwendigerweise. Ein Gegenbeispiel ist $C^0([0, 1])$ mit dem Skalarprodukt $(f, g) = \int_0^1 f g dx$ und $V = C^\infty([0, 1])$. Angenommen, es gäbe zu $f \in C^0([0, 1])$, $f \notin V$ Funktionen $v \in V$ und $w \in V^\perp$ mit $f = v + w$. Nach Definition gilt $(w, g) = \int_0^1 w g dx = 0$ für alle $g \in V = C^\infty([0, 1])$. Nach einem Satz aus der Analysis ergibt dies $w = 0$, also $f = v \in V$. Widerspruch zu $f \notin V$.

Um eine derartige Zerlegung zu ermöglichen, benötigen wir einen abgeschlossenen Unterraum. Es gilt das folgende Resultat.

Satz 2.23 (Projektionssatz). *Seien H ein Hilbertraum und $V \subset H$ ein abgeschlossener Unterraum. Dann existiert zu jedem $u \in H$ eine eindeutige Zerlegung*

$$u = v + w, \quad v \in V, w \in V^\perp.$$

Wir schreiben in diesem Fall $H = V \oplus V^\perp$ und nennen diese Zerlegung eine direkte Summe.

Beweis. 1. Wir zeigen zuerst, dass es ein eindeutiges Element $v \in V$ gibt mit

$$\|u - v\| = d := \inf_{z \in V} \|u - z\|.$$

Dies bedeutet, dass v dem Element u "am nächsten" ist unter allen anderen Elementen von V (siehe Abbildung 2.4). Falls $u \in V$, ist nichts zu zeigen, da wir dann $v = u$ und $w = 0$ wählen können. Sei also $u \notin V$. Sei ferner $(v_n) \subset V$ eine Minimalfolge, d.h. $\|u - v_n\| \rightarrow d$ für $n \rightarrow \infty$. Wir verwenden die Parallelogrammgleichung

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$$

aus Proposition 2.19, angewendet auf $x = (u - v_n)/2$ und $y = (u - v_m)/2$. Dies ergibt

$$\left\|u - \frac{1}{2}(v_n + v_m)\right\|^2 + \left\|\frac{1}{2}(v_m - v_n)\right\|^2 = \frac{1}{2}\|u - v_n\|^2 + \frac{1}{2}\|u - v_m\|^2.$$

Nach Konstruktion der Minimalfolge gilt $\|u - v_n\| \rightarrow d$ für $n \rightarrow \infty$, also konvergiert die rechte Seite gegen d^2 für $n, m \rightarrow \infty$. Andererseits ist $\|u - (v_n + v_m)/2\|^2 \geq \inf_{z \in V} \|u - z\|^2 = d^2$. Also konvergiert $\|v_n - v_m\|$ für $n, m \rightarrow \infty$ gegen null (genauer: (v_n) ist eine Cauchy-Folge). Da abgeschlossene Unterräume von Hilberträumen wieder Hilberträume und insbesondere vollständig sind (Übungsaufgabe), ist (v_n) konvergent, und wir setzen $v = \lim_{n \rightarrow \infty} v_n$. Wegen $\lim_{n \rightarrow \infty} \|u - v_n\| = d$ folgt $\|u - v\| = d$.

2. Als nächstes zeigen wir die Eindeutigkeit. Sind $v, w \in V$ mit $v \neq w$ und $\|u - v\| = \|u - w\| = d$, so folgt aus der Parallelogrammgleichung, angewendet auf $x = (u - v)/2$ und $y = (u - w)/2$,

$$\begin{aligned} d^2 &\leq \left\|u - \frac{1}{2}(v + w)\right\|^2 = -\left\|\frac{1}{2}(v - w)\right\|^2 + 2\left\|\frac{1}{2}(u - v)\right\|^2 + 2\left\|\frac{1}{2}(u - w)\right\|^2 \\ &= -\frac{1}{4}\|v - w\|^2 + \frac{d^2}{2} + \frac{d^2}{2} < d^2, \end{aligned}$$

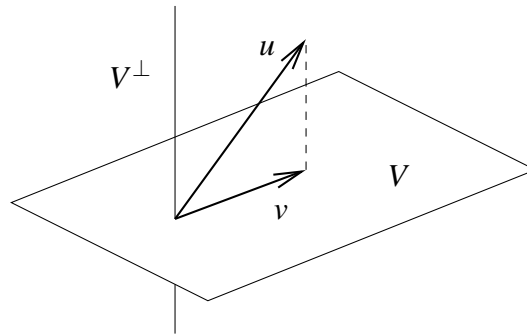


Abbildung 2.4: Illustration für den Projektionssatz.

ein Widerspruch zu $v \neq w$. Also muss $v = w$ gelten.

3. Definiere $w = u - v$. Wir behaupten, dass $w \in V^\perp$. Um dies zu beweisen, bemerken wir zunächst, dass für alle $t \in \mathbb{R}$ und $z \in V$ gilt

$$\|u - v\|^2 \leq \|u - v + tz\|^2,$$

denn $d = \|u - v\|$ beschreibt ja gerade den minimalen Abstand. Multiplizieren wir beide Seiten aus,

$$\|u\|^2 - 2\operatorname{Re}(u, v) + \|v\|^2 \leq \|u\|^2 + \|v\|^2 + t^2\|z\|^2 - 2\operatorname{Re}(u, v) + 2\operatorname{Re}(u - v, tz),$$

so folgt

$$0 \leq |t|^2\|z\|^2 + 2\operatorname{Re}(u - v, tz). \quad (2.10)$$

Division durch $t > 0$ und der Grenzwert $t \rightarrow 0$ ergeben $0 \leq 2\operatorname{Re}(u - v, z)$ für alle $z \in V$. Falls $t < 0$, dreht sich das Ungleichheitszeichen bei der Division durch t um, und der Grenzwert $t \rightarrow 0$ führt auf $0 \geq 2\operatorname{Re}(u - v, z)$ für alle $z \in V$. Insgesamt erhalten wir $\operatorname{Re}(u - v, z) = 0$ für alle $z \in V$.

Es bleibt $\operatorname{Im}(u - v, z) = 0$ für alle $z \in V$ zu zeigen. Wählen wir $t = i\theta$ mit $\theta \in \mathbb{R}$, so folgt aus (2.10), das auch für $t \in \mathbb{C}$ gilt,

$$0 \leq \theta^2\|z\|^2 + 2\operatorname{Re}(i\theta(u - v, z)) = \theta^2\|z\|^2 - 2\theta\operatorname{Im}(u - v, z).$$

Kürzen mit $\theta > 0$ und der Grenzübergang $\theta \rightarrow 0$ liefert $\operatorname{Im}(u - v, z) \leq 0$, während dieselbe Argumentation mit $\theta < 0$ auf $\operatorname{Im}(u - v, z) \geq 0$ führt, also insgesamt $\operatorname{Im}(u - v, z) = 0$. Dies zeigt $(u - v, z) = 0$ für alle $z \in V$, was die Behauptung impliziert. Damit ist der Satz bewiesen. \square

Beispiel 2.24. Betrachte $H = \ell^2$ mit dem Skalarprodukt aus Beispiel 2.18 und sei $V = \{x = (x_n) \in \ell^2 : x_1 = 0, x_n \in \mathbb{R}\}$. Dann ist V abgeschlossen, denn ist $x^{(j)} = (x_n^{(j)})_n$ eine Folge von Zahlenfolgen in ℓ^2 mit $x_1^{(j)} = 0$ und $\|x^{(j)} - x\| \rightarrow 0$ für $j \rightarrow \infty$, so folgt

$$|x_1^{(j)} - x_1| \leq \left(\sum_{n=1}^{\infty} |x_n^{(j)} - x_n|^2 \right)^{1/2} = \|x^{(j)} - x\| \rightarrow 0 \quad \text{für } j \rightarrow \infty$$

und damit $x_1 = 0$, also $x \in V$. Das orthogonale Komplement ist gegeben durch

$$V^\perp = \{(x_n) : x_n = 0 \text{ für alle } n \geq 2\},$$

denn $(x, y) = 0$ für $x \in V$ und $y \in V^\perp$. Außerdem ist $\ell^2 = V \oplus V^\perp$. \square

Mit Hilfe des Projektionssatzes können wir Projektionen definieren.

Definition 2.25 (Projektion). Sei X ein Vektorraum. Dann heißt $P : X \rightarrow X$ eine Projektion, wenn P linear ist und $P^2 = P$.

Hierbei ist P^2 als die Komposition $P \circ P$ definiert, also $P^2(u) = P(Pu)$ für $u \in X$. Projektionen sind eng mit direkten Summen verbunden, wie der folgende Satz zeigt. Dazu definieren wir den Bildraum (engl. *range*) $R(P)$ eines Operators P bzw. den Kern oder Nullraum $N(P)$ von P als

$$R(P) = \{v \in X : \exists u \in X : Pu = v\}, \quad N(P) = \{u \in X : Pu = 0\}.$$

Da P linear ist, sind beide Räume Unterräume von X .

Satz 2.26. Sei X ein Vektorraum. Wenn $P : X \rightarrow X$ eine Projektion ist, dann gilt $X = R(P) \oplus N(P)$. Sind umgekehrt M und N Unterräume von X mit $X = M \oplus N$, dann existiert eine Projektion $P : X \rightarrow X$ mit $R(P) = M$ und $N(P) = N$.

Beweis. 1. Sei P eine Projektion. Wir zeigen zuerst, dass $v \in R(P)$ genau dann gilt, wenn $v = Pv$. Ist $v = Pv$, dann ist nach Definition des Bildraums $v \in R(P)$. Ist umgekehrt $v \in R(P)$, so existiert ein $u \in X$ mit $Pu = v$ und daher nach Definition der Projektion $Pv = P^2u = Pu = v$.

2. Um $X = R(P) \oplus N(P)$ zu zeigen, müssen wir beweisen, dass wir jedes $u \in X$ als eine Summe zweier Element aus $R(P)$ und $N(P)$ schreiben können und dass diese Zerlegung eindeutig ist, d.h., dass $R(P) \cap N(P) = \{0\}$ gilt. Ist $u \in R(P) \cap N(P)$, so gelten $u = Pu$ und $Pu = 0$ und damit $u = 0$, was die letztere Behauptung zeigt. Um die Zerlegung zu zeigen, schreiben wir für $u \in X$

$$u = v + w \quad \text{mit } v = Pu, \quad w = u - Pu.$$

Dann ist $v \in R(P)$ und $w \in N(P)$, denn $Pw = P(u - Pu) = Pu - P^2u = 0$.

3. Definiere für $u \in X$ mit $u = v + w$ und $v \in M$, $w \in N$ die Abbildung $P : X \rightarrow X$, $Pu = v$. Dann ist P linear und eine Projektion, denn $P^2u = Pv = v = Pu$ für alle $u \in X$. Hier haben wir den ersten Teil des Beweises und insbesondere $Pu = v$ genutzt. Es ist klar, dass $R(P) = M$ gilt, denn für $v \in M$ gilt $v = Pv$. Ist schließlich $w \in N$, so folgt aus der eindeutigen Zerlegung $w = 0 + w$ die Gleichung $Pw = 0$, also $w \in N(P)$. \square

Man kann sogar zeigen, dass es zu einem Unterraum $M \subset X$ eine Projektion $P : X \rightarrow X$ auf M gibt (d.h. $R(P) = M$). Dieses Resultat ist viel stärker als die Aussage des obigen Satzes und der Beweis benötigt sogar das Auswahlaxiom.

Beispiel 2.27. (1) Sei H ein Hilbertraum und $V \subset H$ abgeschlossen. Im Beweis des Projektionssatzes 2.23 haben wir gezeigt, dass es zu jedem $u \in H$ ein eindeutiges Element $v \in V$ gibt, so dass $u - v$ den Abstand von u zu V in dem folgenden Sinne minimiert:

$$\|u - v\| = \min_{w \in V} \|u - w\|.$$

Dies definiert die Abbildung $P : H \rightarrow H$, $Pu = v$. Wir behaupten, dass P eine Projektion ist.

Wir zeigen zunächst ein Hilfsresultat: Ist $u \in H$, so ist $w := u - Pu$ orthogonal zu V . Angenommen, dies gilt nicht, d.h., es existiert ein $v \in V$ mit $(w, v) = \alpha \neq 0$. Sei $\lambda = \bar{\alpha}/\|v\|^2$. Dann ist einerseits

$$\begin{aligned} \|w - \lambda v\|^2 &= (w - \lambda v, w - \lambda v) = \|w\|^2 - \bar{\lambda}\alpha - \lambda\alpha + |\lambda|^2\|v\|^2 \\ &= \|w\|^2 - \underbrace{\bar{\lambda}(\alpha - \lambda\|v\|^2)}_{=0} - \lambda\alpha = \|w\|^2 - \frac{|\alpha|^2}{\|v\|^2} < \|w\|^2. \end{aligned}$$

Andererseits ist $w - \lambda v = u - (Pu + \lambda v)$ mit $Pu + \lambda v \in V$, also

$$\|w\| = \|u - Pu\| = \min_{z \in V} \|u - z\| \leq \|u - (Pu + \lambda v)\| = \|w - \lambda v\| < \|w\|,$$

Widerspruch.

Wir zeigen nun die Linearität von P . Seien $u_1, u_2 \in H$ und $w \in V$. Gemäß des obigen Hilfsresultats ist $u_i - Pu_i$ orthogonal zu V ($i = 1, 2$). Aus dem Satz von Pythagoras folgt

$$\begin{aligned} \|(u_1 + u_2) - w\|^2 &= \|(u_1 - Pu_1 + u_2 - Pu_2) + (Pu_1 + Pu_2 - w)\|^2 \\ &= \|u_1 - Pu_1 + u_2 - Pu_2\|^2 + \|Pu_1 + Pu_2 - w\|^2 \\ &\geq \|u_1 + u_2 - (Pu_1 + Pu_2)\|^2. \end{aligned}$$

Das Minimum über alle $w \in V$ ergibt

$$\|u_1 + u_2 - (Pu_1 + Pu_2)\| \leq \min_{w \in V} \|u_1 + u_2 - w\| \leq \|u_1 + u_2 - P(u_1 + u_2)\|,$$

denn $\min_{w \in V} \|u_1 + u_2 - w\| \leq \|u_1 + u_2 - P(u_1 + u_2)\|$. Es folgt $Pu_1 + Pu_2 = P(u_1 + u_2)$, denn $P(u_1 + u_2)$ und $Pu_1 + Pu_2$ realisieren beide das Minimum und müssen daher übereinstimmen. Analog können wir $P(\lambda u) = \lambda Pu$ zeigen.

Es bleibt $P^2 = P$ zu zeigen. Seien $u \in H$ und $v = Pu \in V$. Aus der Minimaleigenschaft von Pu folgt

$$\|Pu - P(Pu)\| = \min_{w \in V} \|Pu - w\| \leq \|Pu - v\| = 0,$$

also $P = P^2$.

Im obigen Sinne ist also Pu die *Bestapproximation* von u im Unterraum V . Mit anderen Worten: Unter allen Elementen des Unterraums V liegt Pu dem Element u am nächsten.

- (2) Sei $H = L^2(\mathbb{R})$ der Raum der quadratintegrierbaren Funktionen auf \mathbb{R} mit dem Skalarprodukt $(f, g) = \int_{\mathbb{R}} fg dx$. Dann ist $P : H \rightarrow H$ mit

$$(Pf)(x) = \frac{1}{2}(f(x) + f(-x)), \quad f \in H, x \in \mathbb{R},$$

eine Projektion auf den Raum M der geraden Funktionen (für die $f(x) = f(-x)$ für alle $x \in \mathbb{R}$ gilt), denn P ist linear,

$$(P^2 f)(x) = \frac{1}{2}((Pf)(x) + (Pf)(-x)) = \frac{1}{4}(f(x) + f(-x)) + \frac{1}{4}(f(-x) + f(x)) = (Pf)(x)$$

und für alle geraden Funktionen f gilt $(Pf)(x) = f(x)$. Damit können wir den Raum $L^2(\mathbb{R})$ zerlegen in den Raum aller geraden Funktionen (der gerade dem Bildraum von P entspricht) und den Raum aller ungeraden Funktionen (der dem Nullraum von P entspricht).

- (3) In der Quantenmechanik ist die folgende Projektion von Bedeutung. Seien H ein Hilbertraum und $u \in H$ ein beliebiger Vektor (in der Quantenmechanik der Zustand des Quantensystems). Wir definieren $P : H \rightarrow H$ durch

$$Pv = (u, v)u, \quad v \in H.$$

Dann ist Pv die ‘‘Orthogonalprojektion’’ von v auf u . In der Physik wird das Skalarprodukt (u, v) durch das Symbol $\langle u|v \rangle$ dargestellt (engl. *bracket*) und $\langle u|$ das ‘‘bra’’ sowie $|v \rangle$ das ‘‘ket’’ genannt. Dann können wir (formal)

$$Pv = \langle u|v \rangle |u \rangle = |u \rangle \langle u|v \rangle$$

oder einfach $P = |u \rangle \langle u|$ schreiben. □

2.4 Orthogonalsysteme

Das Konzept einer orthogonalen Zerlegung eines Vektorraums (als direkte Summe zweier Unterräume) können wir verallgemeinern. Sei dazu H ein Prä-Hilbertraum mit Skalarprodukt (\cdot, \cdot) und induzierter Norm $\|\cdot\|$. Sei im Folgenden I eine (echte) Teilmenge von \mathbb{N} oder $I = \mathbb{N}$.

Definition 2.28 (Orthonormalsystem). *Zwei Elemente u und v des Prä-Hilbertraums H heißen orthogonal, wenn $(u, v) = 0$. Eine Familie $(e_n) \subset H$ mit $e_n \neq 0$ für alle $n \in I$ heißt ein Orthogonalsystem, wenn $(e_n, e_m) = 0$ für alle $n \neq m$. Sind alle e_n normiert, d.h. $\|e_n\| = 1$ für $n \in I$, so nennen wir (e_n) ein Orthonormalsystem.*

Beispiel 2.29. (1) Das kanonische Orthonormalsystem des \mathbb{R}^n (versehen mit dem Standardskalarprodukt) ist gegeben durch die Vektoren

$$e_j = (0, \dots, 0, 1, 0, \dots, 0)^\top, \quad j = 1, \dots, n,$$

wobei die Eins an der j -ten Stelle steht. Eine andere Schreibweise mit dem Kronecker-Symbol δ_{jk} , definiert durch $\delta_{jk} = 0$ für $j \neq k$ und $\delta_{jk} = 1$ für $j = k$, lautet

$$e_j^k = \delta_{jk}, \quad \text{wobei } e_j = (e_j^1, \dots, e_j^n)^\top.$$

- (2) Im Raum ℓ^2 bilden die Zahlenfolgen e_j , definiert durch $e_j^k = \delta_{jk}$, ein Orthonormalsystem. Der einzige Unterschied zu dem obigem Beispiel ist, dass die Indizes j und k alle natürlichen Zahlen durchlaufen.
- (3) Die Funktionen $f_n(x) = \cos(nx)/\sqrt{\pi}$, $n \in \mathbb{N}$, bilden ein Orthonormalsystem in $L^2(-\pi, \pi)$, denn für $n \neq m$ gilt

$$\begin{aligned} (f_n, f_m) &= \frac{1}{\pi} \int_{-\pi}^{\pi} f_n(x) f_m(x) dx = \frac{1}{\pi} \int_{-\pi}^{\pi} \cos(nx) \cos(mx) dx \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} (\cos((m+n)x) + \cos((m-n)x)) dx \\ &= \frac{1}{2\pi} \left[\frac{\sin((m+n)x)}{m+n} + \frac{\sin((m-n)x)}{m-n} \right]_{-\pi}^{\pi} \\ &= \frac{1}{\pi} \left(\frac{\sin((m+n)\pi)}{m+n} + \frac{\sin((m-n)\pi)}{m-n} \right) = 0, \end{aligned}$$

und für $n = m$ ist

$$\begin{aligned} (f_n, f_n) &= \frac{1}{\pi} \int_{-\pi}^{\pi} \cos^2(nx) dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} (1 + \cos(2nx)) dx \\ &= \frac{1}{2\pi} \left[x + \frac{\sin(2nx)}{2n} \right]_{-\pi}^{\pi} = 1. \end{aligned}$$

Damit haben wir $(f_n, f_m) = \delta_{nm}$ gezeigt. □

Elemente eines Orthogonalsystems (e_n) sind linear unabhängig (Übungsaufgabe). Außerdem kann aus ihnen durch Division eines jeden Elements durch seine Norm ein Orthonormalsystem definiert werden. Es gilt sogar: Aus jeder Menge (abzählbar vieler) linear unabhängiger Elemente können wir ein Orthonormalsystem konstruieren. Dies ist der Inhalt des folgenden Satzes.

Satz 2.30 (Orthogonalisierungsverfahren nach Gram-Schmidt). *Seien H ein Prä-Hilbertraum und $(u_n) \subset H$ eine endliche oder abzählbare Menge nichtverschwindender linear unabhängiger Elemente, $n \in I \subset \mathbb{N}$. Dann wird durch*

$$\begin{aligned} v_1 &= u_1, \\ v_{n+1} &= u_{n+1} - \sum_{j=1}^n \frac{(v_j, u_{n+1})}{\|v_j\|^2} v_j, \quad n \geq 2, \end{aligned}$$

ein Orthogonalsystem definiert, das dieselbe lineare Hülle wie (u_n) besitzt.

Beweis. Wir zeigen den Satz mit vollständiger Induktion. Angenommen, (v_1, \dots, v_n) sind paarweise orthogonal. Wir zeigen, dass dann (v_1, \dots, v_n) orthogonal zu v_{n+1} ist. Sei $1 \leq m \leq n$. Dann ist nach Definition

$$\begin{aligned} (v_m, v_{n+1}) &= (v_m, u_{n+1}) - \sum_{j=1, j \neq m}^n \frac{(v_j, u_{n+1})}{\|v_j\|^2} \underbrace{(v_m, v_j)}_{=0} - \frac{(v_m, u_{n+1})}{\|v_m\|^2} \underbrace{(v_m, v_m)}_{=\|v_m\|^2} \\ &= (v_m, u_{n+1}) - (v_m, u_{n+1}) = 0, \end{aligned}$$

wobei wir für die Aussage $(v_m, v_j) = 0$ für $j \neq m$ die Induktionsannahme verwendet haben. Dies beweist, dass (v_n) ein Orthogonalsystem ist. Den Beweis, dass dieses System dieselbe lineare Hülle wie (u_n) besitzt, überlassen wir den LeserInnen. \square

Beispiel 2.31. Die Polynome $(1, x, x^2, \dots)$ auf $(-1, 1)$ sind linear unabhängig, aber bezüglich des Skalarprodukts in $L^2(-1, 1)$ nicht orthogonal. Nach dem Orthogonalisierungsverfahren nach Gram-Schmidt können wir sie jedoch orthogonalisieren:

$$\begin{aligned} v_1(x) &= 1, \\ v_2(x) &= x - \frac{(1, x)}{\|1\|^2} 1 = x, \\ v_3(x) &= x^2 - \frac{(1, x^2)}{\|1\|^2} 1 - \frac{(x, x^2)}{\|x\|^2} x = x^2 - \frac{1}{3} = \frac{1}{3}(3x^2 - 1), \quad \text{usw.} \end{aligned}$$

Üblicherweise werden die Polynome v_n anders skaliert, z.B. $v_3(x) = \frac{1}{2}(3x^2 - 1)$. Dies ändert nichts an der Orthogonalisierungseigenschaft, hat aber den Vorteil, dass die Polynome v_n über die Formel

$$v_{n+1}(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} ((x^2 - 1)^n), \quad n \geq 0, \quad x \in (-1, 1),$$

berechnet werden können. Man nennt diese Funktionen die *Legendre-Polynome*. Sie werden bei der Lösung bestimmter Differentialgleichungen (nämlich sogenannter Legendre-Differentialgleichungen), die bei der Lösung des Wellenanteils der Schrödinger-Gleichung der Quantenmechanik auftreten, benötigt. \square

Wir haben in Beispiel 2.27 gezeigt, dass die Projektion eines Elements eines Hilbertraums auf einem Unterraum gleich der Bestapproximation in diesem Unterraum ist. Diese Idee können wir mit Hilfe von Orthonormalsystemen verallgemeinern. Es ist nämlich $(e_k, u)e_k$ die Projektion von u entlang e_k , und $\sum_{k=1}^n (e_k, u)e_k$ ist die Projektion von u auf die lineare Hülle der (e_1, \dots, e_n) . Können wir den Grenzwert $n \rightarrow \infty$ betrachten und damit alle Elemente $u \in H$ durch die Reihe $\sum_{k=1}^{\infty} (e_k, u)e_k$ darstellen? Die Antwort ist positiv, wie der folgende Satz aussagt.

Satz 2.32. *Seien H ein Prä-Hilbertraum, $u \in H$ und (e_n) mit $n \in \mathbb{N}$ ein Orthonormalsystem. Dann gilt:*

- (i) *Die Reihe $\sum_{k=1}^{\infty} |(e_k, u)|^2$ konvergiert, d.h. $((e_k, u))$ ist ein Element aus ℓ^2 .*
- (ii) *Die Reihe $\sum_{k=1}^{\infty} (e_k, u)e_k$ konvergiert in H gegen u genau dann, wenn die Parseval-Gleichung*

$$\sum_{k=1}^{\infty} |(e_k, u)|^2 = \|u\|^2$$

erfüllt ist.

Beweis. (i) Definiere die Projektion

$$P_n u = \sum_{k=1}^n (e_k, u) e_k, \quad u \in H.$$

Dann ist $P_n u$ orthogonal zu $(I - P_n)u$ (I die Identität in H), da $u = P_n u + (I - P_n)u$ und $H = \mathbf{R}(P_n) \oplus \mathbf{N}(P_n)$. Folglich ist

$$\begin{aligned} \|u\|^2 &= (P_n u + (I - P_n)u, P_n u + (I - P_n)u) = (P_n u, P_n u) + ((I - P_n)u, (I - P_n)u) \\ &= \|P_n u\|^2 + \|(I - P_n)u\|^2. \end{aligned}$$

Wir schreiben den Summanden $\|P_n u\|^2$ um:

$$\|P_n u\|^2 = (P_n u, P_n u) = \sum_{j,k=1}^n \overline{(e_j, u)} (e_k, u) \underbrace{(e_j, e_k)}_{=\delta_{jk}} = \sum_{k=1}^n |(e_k, u)|^2$$

und schliessen

$$\|u\|^2 = \sum_{k=1}^n |(e_k, u)|^2 + \|(I - P_n)u\|^2 \geq \sum_{k=1}^n |(e_k, u)|^2. \quad (2.11)$$

Die Ungleichung zeigt, dass die Partialsummen $\sum_{k=1}^n |(e_k, u)|^2$ beschränkt sind. Da sie weiterhin monoton wachsend sind, konvergieren sie.

(ii) Gilt die Parseval-Gleichung, so zeigt (2.11) im Grenzwert $n \rightarrow \infty$, dass

$$0 = \lim_{n \rightarrow \infty} (I - P_n)u = u - \sum_{k=1}^{\infty} (e_k, u) e_k.$$

Existiert umgekehrt dieser Grenzwert, also $\lim_{n \rightarrow \infty} (I - P_n)u = 0$, so folgt aus (2.11) im Grenzwert $n \rightarrow \infty$ die Parseval-Gleichung. \square

Wir zeichnen die Reihe $\sum_{k=1}^{\infty} (e_k, u) e_k$ aus und nennen sie eine Fourier-Reihe.

Definition 2.33. Seien H ein Prä-Hilbertraum und (e_k) ein abzählbares Orthonormalsystem. Die Reihe

$$\sum_{k=1}^{\infty} (e_k, u) e_k$$

heißt Fourier-Reihe von u , unabhängig davon, ob sie konvergiert oder nicht. Die Koeffizienten (e_k, u) heißen die Fourier-Koeffizienten von u . Sei H zusätzlich ein Hilbertraum. Konvergiert die Fourier-Reihe von jedem $u \in H$ und ist ihr Grenzwert gleich u , so heißt (e_k) ein vollständiges Orthonormalsystem.

Fourier-Reihen sind spezielle Reihen der Form

$$\sum_{k=1}^{\infty} c_k e_k, \quad c_k \in \mathbb{R} \text{ oder } c_k \in \mathbb{C},$$

nämlich solche, für die $c_k = (e_k, u)$. Eine notwendige Voraussetzung für die Konvergenz solcher Reihen ist $(c_k) \in \ell^2$, d.h. $\sum_{k=1}^{\infty} |c_k|^2$ konvergiert und $(|c_k|^2)$ ist eine Cauchy-Folge. In diesem Fall ist wegen

$$\left\| \sum_{k=j}^m c_k e_k \right\|^2 = \sum_{k=j}^m |c_k|^2$$

auch $\sum_{k=j}^m (e_k, u) e_k$ eine Cauchy-Folge, die jedoch nicht konvergieren muss, wenn H nicht vollständig ist. Ist H vollständig, also ein Hilbertraum, so konvergiert die Reihe $\sum_{k=1}^{\infty} c_k e_k$ genau dann, wenn $(c_k) \in \ell^2$. Für den Fall $c_k = (e_k, u)$ entspricht diese Aussage gerade dem ersten Teil des obigen Satzes.

Aus der Definition folgt, dass für jedes vollständige Orthonormalsystem (e_n) die lineare Hülle $\text{span}(e_n)$ dicht in H liegt, d.h. $\overline{\text{span}(e_n)} = H$. Insbesondere lässt sich jedes $u \in H$ als eine Fourierreihe darstellen, und es gilt die Parseval-Gleichung $\|u\|^2 = \sum_{k=1}^{\infty} |(e_k, u)|^2$.

Beispiel 2.34. (1) Betrachte den Raum $L^2(-\pi, \pi)$ der quadratintegrierbaren Funktionen auf $(-\pi, \pi)$ und sei (e_k) gegeben durch Schule

$$\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}} \cos x, \frac{1}{\sqrt{\pi}} \sin x, \frac{1}{\sqrt{\pi}} \cos(2x), \frac{1}{\sqrt{\pi}} \sin(2x), \dots$$

Wir haben in Beispiel 2.29 (2) gezeigt, dass $(\cos(nx)/\sqrt{\pi})$ ein Orthonormalsystem ist. Analog kann man zeigen, dass $(\sin(nx)/\sqrt{\pi})$ ein Orthonormalsystem bildet, und dasselbe gilt für die obige Funktionenfolge. Die Fourier-Reihe von $f \in L^2(-\pi, \pi)$ bezüglich dieses Orthonormalsystems wird üblicherweise geschrieben als

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx)),$$

wobei die Fourier-Koeffizienten von f definiert sind durch

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(nx) dx, \quad n \geq 0,$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(nx) dx, \quad n \geq 1.$$

Nach Satz 2.32 konvergiert diese Reihe und ihr Wert ist gleich $f(x)$. Man kann zeigen (mit Hilfe des Approximationssatzes von Weierstraß), dass $\text{span}(e_k)$ dicht in $L^2(-\pi, \pi)$ ist, d.h., jede L^2 -Funktion lässt sich durch ihre Fourier-Reihe darstellen.

Die Fourier-Reihe zerlegt eine beliebige L^2 -Funktion in ihr Frequenzspektrum, d.h., sie geben an, mit welcher Amplitude die Frequenzen n auftreten. Dies ist von Bedeutung zur Bestimmung der Frequenzen eines Signals. Ist beispielsweise $f(x) = \sin x + \sin(2x)$, so sind nur die Fourier-Koeffizienten $b_1 = 1$ und $b_2 = 1$ von null verschieden; die Koeffizienten b_n geben gerade die Frequenzen an, mit denen $f(x)$ oszilliert.

- (2) In manchen Anwendungen (Kompression von Audio- oder Bilddateien) werden nur die Kosinusterme verwendet, was auf die *diskrete Kosinustransformation*

$$f_n = \sum_{k=0}^{N-1} a_k \cos\left(\frac{\pi}{N}\left(k + \frac{1}{2}\right)n\right), \quad n = 0, \dots, N-1,$$

für gegebene (a_0, \dots, a_{N-1}) führt. Genau genommen wird in der Bildkompression eine zweidimensionale Transformation verwendet. Die Transformation ist diskret, da die Audiodaten nur zu bestimmten Zeitpunkten gespeichert werden bzw. die Bilddaten in Form von Pixel vorliegen. Die diskrete Kosinustransformation wird bei der Kompression von MPEG- und JPEG-Daten verwendet. Da nur relativ wenige Frequenzanteile mit großen Amplituden verwendet werden, können die Originaldaten erheblich reduziert werden. Außerdem ist es möglich, schnelle Rechenalgorithmen (sogenannte *Fast Fourier Transform*) zu entwickeln.

- (3) Sei w eine auf (a, b) integrierbare Gewichtsfunktion und $L_w^2(a, b) = \{f : (a, b) \rightarrow \mathbb{R} \text{ integrierbar} : \int_a^b f w dx, \int_a^b f^2 w dx \text{ existieren}\}$ ein gewichteter L^2 -Raum mit Skalarprodukt $(f, g) = \int_a^b f(x)g(x)w(x)dx$. Je nach Gewichtsfunktion können wir verschiedene Orthogonalsysteme definieren ($n \geq 0$):

- Legendre-Polynome: $(a, b) = (-1, 1)$, $w(x) = 1$ liefert

$$L_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} ((x^2 - 1)^n).$$

- Hermite-Polynome: $(a, b) = (-\infty, \infty)$, $w(x) = \exp(-x^2)$ liefert

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} (e^{-x^2}).$$

- Laguerre-Polynome: $(a, b) = (0, \infty)$, $w(x) = e^{-x}$ liefert

$$\ell_n(x) = (-1)^n e^x \frac{d^n}{dx^n} (x^n e^{-x})$$

Beachte, dass diese Systeme nicht notwendigerweise normiert sind; es gilt beispielsweise $\|H_n\|_{L^2}^2 = 2^n n! \pi$ und $\|L_n\|_{L^2}^2 = (n!)^2$. Wie die Legendre-Polynome werden auch die Hermite- und Laguerre-Polynome bei der Lösung der Schrödinger-Gleichung mit speziellen Potentialen benötigt. \square

Wir nennen einen metrischen Raum X *separabel*, wenn er eine abzählbare, dichte Teilmenge $M \subset X$ besitzt, d.h. $\overline{M} = X$. Beispiele für separable Räume sind \mathbb{R} (wähle $M = \mathbb{Q}$), $C^0([a, b])$ (wähle die Menge der Polynome mit rationalen Koeffizienten) oder ℓ^p mit $1 \leq p < \infty$. Der Raum ℓ^∞ dagegen ist *nicht* separabel.

Separable Hilberträume haben die bemerkenswerte Eigenschaft, dass sie *isometrisch* zum Hilbertraum ℓ^2 sind, d.h., es existiert eine lineare, invertierbare Abbildung (eine *Isometrie*) $T : H \rightarrow \ell^2$ mit $(u, v) = (Tu, Tv)$ für alle $u, v \in H$. Mit anderen Worten: Separable Hilberträume können in gewisser Weise mit dem Raum ℓ^2 "identifiziert" werden.

Satz 2.35. Sei H ein separabler Hilbertraum. Dann ist H und ℓ^2 isometrisch.

Beweis. Sei (v_n) eine abzählbare, dichte Teilmenge von H . Ohne Einschränkung können wir annehmen, dass (v_1, \dots, v_N) für alle $N \in \mathbb{N}$ linear unabhängig ist (anderenfalls entfernen wir die linear abhängigen Elemente). Führen wir das Orthogonalisierungsverfahren nach Gram-Schmidt durch, so erhalten wir nach Normalisierung ein Orthonormalsystem (e_n) . Es gilt $\text{span}(v_n) = \text{span}(e_n)$, und wegen der Dichtheit ist (e_n) ein vollständiges Orthonormalsystem.

Sei $u \in H$. Dann besitzt u die Darstellung $u = \sum_{k=1}^{\infty} (e_k, u) e_k$. Wir definieren die Abbildung $T : H \rightarrow \ell^2$ durch

$$Tu = ((e_1, u), (e_2, u), \dots), \quad u \in H.$$

Dann ist T wohldefiniert und invertierbar (folgt aus den Eigenschaften der Fourier-Reihe), und es gilt

$$(u, v) = \sum_{k=1}^{\infty} \overline{(e_k, u)} (e_k, v) = (Tu, Tv),$$

d.h., T ist eine Isometrie. □

2.5 Lineare Operatoren

Beschränkte Funktionen $f : (a, b) \rightarrow \mathbb{R}$ haben die Eigenschaft, dass es eine Konstante $C > 0$ gibt, so dass für alle $x \in (a, b)$ die Eigenschaft $|f(x)| \leq C$ gilt. Für *lineare* Funktionen können wir dies nicht erwarten, denn in diesem Fall ist $|f(\lambda x)| = |\lambda| |f(x)| \rightarrow \infty$ für $\lambda \rightarrow \infty$ (sofern $f(x) \neq 0$). Daher definieren wir für solche Funktionen die Beschränktheit anders.

Seien X, Y zwei normierte Räume. Wir bezeichnen die Normen in beiden Räumen mit $\|\cdot\|$, wenn keine Missverständnisse auftreten. Wir nennen eine Abbildung $A : X \rightarrow Y$ einen *Operator*. Ist $Y = \mathbb{R}$, so wird A auch ein *Funktional* genannt.

Definition 2.36 (Beschränkte lineare Operatoren). *Ein linearer Operator $A : X \rightarrow Y$ heißt beschränkt, wenn es ein $C > 0$ gibt, so dass für alle $u \in X$ gilt*

$$\|Au\| \leq C\|u\|.$$

Interessanterweise ist dieser Begriff für lineare Operatoren gleichbedeutend mit der Stetigkeit.

Satz 2.37. *Ein linearer Operator $A : X \rightarrow Y$ ist beschränkt genau dann, wenn er stetig ist.*

Beweis. 1. Sei A beschränkt und sei $u_n \rightarrow u$ für $n \rightarrow \infty$, d.h. $\|u_n - u\| \rightarrow 0$. Dann folgt sofort aus der Beschränktheit von A

$$\|Au_n - Au\| = \|A(u_n - u)\| \leq C\|u_n - u\| \rightarrow 0,$$

was die Stetigkeit von A zeigt.

2. Sei umgekehrt A stetig. Angenommen, A wäre nicht beschränkt, d.h.

$$\forall C > 0 : \exists u \in X : \|Au\| > C\|u\|.$$

Wir wählen $C = n \in \mathbb{N}$ und $u = u_n \in X$, so dass $\|Au_n\| > n\|u_n\|$. Setze $v_n = u_n/(n\|u_n\|)$. Dann folgt

$$\|Av_n\| = \frac{\|Au_n\|}{n\|u_n\|} > 1, \quad \|v_n\| = \frac{1}{n} \rightarrow 0 \text{ für } n \rightarrow \infty.$$

Für $\|v_n\| \rightarrow 0$ folgt jedoch aus der Stetigkeit $\|Av_n\| \rightarrow 0$, Widerspruch. \square

Beispiel 2.38. (1) Matrizen $A = (a_{jk}) \in \mathbb{R}^{n \times n}$ definieren via $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $F(v) = Av$, lineare beschränkte Abbildungen. Verwenden wir die euklidische Norm in \mathbb{R}^n , so folgt

$$\|F(v)\|^2 = \left\| \left(\sum_{k=1}^n a_{jk} v_k \right)_j \right\|^2 = \sum_{j=1}^n \left(\sum_{k=1}^n a_{jk} v_k \right)^2 \leq \sum_{j=1}^n \max_{k=1, \dots, n} a_{jk}^2 \left(\sum_{k=1}^n v_k \right)^2 = C^2 \|v\|^2,$$

wobei $C = n \sum_{j=1}^n \max_{k=1, \dots, n} |a_{jk}|$.

(2) Sei $Y = C^0([0, 1])$ versehen mit der Maximumsnorm $\|f\| = \max_{x \in [0, 1]} |f(x)|$ und sei $X = L^1(0, 1)$. Dann ist der *Fredholmsche Integraloperator* $A : X \rightarrow Y$,

$$(Af)(x) = \int_0^1 k(x, y) f(y) dy, \quad 0 \leq x \leq 1,$$

mit $k \in C^0([0, 1]^2)$ wohldefiniert (d.h. $Af \in C^0([0, 1])$) und stetig, denn

$$\begin{aligned} \|Af\| &= \max_{0 \leq x \leq 1} \left| \int_0^1 k(x, y) f(y) dy \right| \leq \max_{0 \leq x \leq 1} \int_0^1 |k(x, y)| |f(y)| dy \\ &= \max_{(x, z) \in [0, 1]^2} |k(x, z)| \int_0^1 |f(y)| dy \leq C \|f\|_{L^1}, \end{aligned}$$

wobei $C = \max_{(x, z) \in [0, 1]^2} |k(x, z)|$.

(3) Die ‘‘Matrix’’ (genauer: der durch die ‘‘Matrix’’ definierte Operator)

$$A = (a_{jk}) = \begin{pmatrix} 1/2 & 0 & 0 & 0 & 0 & \dots \\ 1/2 & 1/2 & 0 & 0 & 0 & \dots \\ 0 & 1/2 & 1/2 & 0 & 0 & \dots \\ 0 & 0 & 1/2 & 1/2 & 0 & \dots \\ \vdots & \vdots & & & \vdots & \end{pmatrix}$$

aus der Einleitung ist stetig in ℓ^2 , denn

$$Au = \frac{1}{2}(u_1, u_1 + u_2, u_2 + u_3, \dots)^\top$$

und, mit der Cauchy-Schwarz-Ungleichung,

$$\|Au\|^2 = \frac{1}{4}u_1^2 + \frac{1}{4} \sum_{k=1}^{\infty} (u_k + u_{k+1})^2 \leq \frac{1}{4}u_1^2 + \frac{1}{4} \sum_{k=1}^{\infty} 2(u_k^2 + u_{k+1}^2) \leq \sum_{k=1}^{\infty} u_k^2 = \|u\|^2.$$

Also ist $C \leq 1$.

- (4) Sei $V \subset H$ ein abgeschlossener Unterraum eines reellen Hilbertraums H und sei $P : H \rightarrow H$ eine Projektion auf V . Zu gegebenem $u \in H$ existieren nach dem Projektionssatz $v \in V$ und $w \in V^\perp$, so dass $u = v + w$. Dann folgt wegen $(v, w) = 0$

$$\|Pu\|^2 = \|v\|^2 \leq \|v\|^2 + \|w\|^2 = \|v + w\|^2 = \|u\|^2,$$

also ist P stetig mit $C \leq 1$.

- (5) Sei $X = \{f \in C^1(\mathbb{R}) : f, f' \in L^2(\mathbb{R})\}$ versehen mit der Norm

$$\|f\|^2 = \int_{\mathbb{R}} f(x)^2 dx.$$

Wir behaupten, dass der Operator der ersten Ableitung $A : X \rightarrow L^2(\mathbb{R})$, $Au = u'$, nicht stetig ist. Betrachte die Folge

$$u_n(x) = \left(\frac{n}{2\pi}\right)^{1/4} e^{-nx^2/4}, \quad x \in \mathbb{R}, n \in \mathbb{N}.$$

Wir verwenden die Formeln

$$\int_{\mathbb{R}} e^{-y^2/2} dy = \int_{\mathbb{R}} y^2 e^{-y^2/2} dy = \sqrt{2\pi}$$

und die Transformation $y = \sqrt{nx}$, $dy = \sqrt{n} dx$, um die L^2 -Norm von u_n zu berechnen:

$$\|u_n\|^2 = \left(\frac{n}{2\pi}\right)^{1/2} \int_{\mathbb{R}} e^{-nx^2/2} dx = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-y^2/2} dy = 1.$$

Wäre A beschränkt, so würde $\|Au_n\| \leq C\|u_n\| = C$ für alle $n \in \mathbb{N}$ gelten. Allerdings ist

$$\begin{aligned} \|Au_n\|^2 &= \left(\frac{n}{2\pi}\right)^{1/2} \int_{\mathbb{R}} \left(\frac{nx}{2}\right)^2 e^{-nx^2/2} dx \\ &= \frac{n}{4\sqrt{2\pi}} \int_{\mathbb{R}} y^2 e^{-y^2/2} dy = \frac{n}{4} \rightarrow \infty \end{aligned}$$

für $n \rightarrow \infty$, Widerspruch. □

Wir nennen die kleinste Konstante $C > 0$, für die $\|Au\| \leq C\|u\|$ gilt, die Norm von A .

Definition 2.39 (Norm eines Operators). Sei $A : X \rightarrow Y$ ein linearer, beschränkter Operator. Wir nennen die Zahl

$$\|A\| = \sup_{u \neq 0} \frac{\|Au\|}{\|u\|}$$

die Norm von A .

Es ist klar, dass diese Zahl existiert, weil $\|Au\|/\|u\| \leq C$ für alle $u \in X$, $u \neq 0$. Außerdem ist die Norm die kleinste Konstante, die diese Ungleichung erfüllt. Insbesondere gilt

$$\|Au\| \leq \|A\| \|u\| \quad \text{für alle } u \in X.$$

Es handelt sich bei $\|A\|$ wirklich um eine Norm auf dem Raum aller linearen, beschränkten Operatoren $A : X \rightarrow Y$, denn die Dreiecksungleichung folgt aus

$$\|A + B\| = \sup_{u \neq 0} \frac{\|Au + Bu\|}{\|u\|} \leq \sup_{u \neq 0} \frac{\|Au\|}{\|u\|} + \sup_{u \neq 0} \frac{\|Bu\|}{\|u\|} = \|A\| + \|B\|,$$

und die anderen beiden Normeigenschaften gelten ebenfalls. Man kann zeigen, dass es für die Definition der Norm zwei zur ursprünglichen Definition äquivalente Formulierungen gibt:

Proposition 2.40. *Sei $A : X \rightarrow Y$ ein linearer, beschränkter Operator. Dann gilt:*

$$\|A\| = \sup_{\|u\| \leq 1} \|Au\| = \sup_{\|u\|=1} \|Au\|.$$

Beweis. Wir erinnern, dass $\|A\| = \sup_{u \neq 0} \|Au\|/\|u\|$. Sei weiter $\alpha = \sup_{\|u\| \leq 1} \|Au\|$. Wir müssen $\alpha = \|A\|$ zeigen. Es folgt für alle $\|u\| \leq 1$

$$\|Au\| \leq \|A\| \|u\| \leq \|A\|, \quad \text{also } \alpha = \sup_{\|u\| \leq 1} \|Au\| \leq \|A\|.$$

Andererseits existiert nach Definition des Supremums zu jedem $\varepsilon > 0$ ein $u_0 \neq 0$, so dass

$$\frac{\|Au_0\|}{\|u_0\|} > \|A\| - \varepsilon.$$

Dann ergibt sich für $v = u_0/\|u_0\|$ wegen $\|v\| = 1$ nach Definition von α

$$\alpha = \sup_{\|u\| \leq 1} \|Au\| \geq \|Av\| = \left\| A \left(\frac{u_0}{\|u_0\|} \right) \right\| = \frac{\|Au_0\|}{\|u_0\|} > \|A\| - \varepsilon.$$

Da $\varepsilon > 0$ beliebig ist, impliziert dies $\alpha \geq \|A\|$, also insgesamt $\alpha = \|A\|$. Auf dieselbe Art können wir $\|A\| = \sup_{\|u\|=1} \|Au\|$ zeigen. \square

Beispiel 2.41. (1) Sei $P : H \rightarrow H$ die Projektion auf dem abgeschlossenen Unterraum $V \subset H$ mit $V \neq \emptyset$. Wir behaupten, dass $\|P\| = 1$. Wegen Beispiel 2.38 (4) erhalten wir $\|Pu\| = \sup_{u \neq 0} \|Pu\|/\|u\| \leq 1$. Außerdem folgt aus $P^2u = Pu \neq 0$

$$\|P\| = \sup_{v \neq 0} \frac{\|Pv\|}{\|v\|} \geq \frac{\|P(Pu)\|}{\|Pu\|} = 1,$$

so dass $\|P\| = 1$.

(2) Sei (a_n) eine reelle Zahlenfolge und definiere $A : \ell^2 \rightarrow \ell^2$, $Ax = (a_1x_1, a_2x_2, \dots)$. Wir behaupten zuerst, dass A genau dann beschränkt ist, wenn (a_n) beschränkt ist. Ist (a_n) beschränkt, so folgt

$$\|Ax\|^2 = \sum_{n=1}^{\infty} a_n^2 x_n^2 \leq \sup_{n \in \mathbb{N}} a_n^2 \|x\|^2,$$

also $\|A\| \leq \sup_{n \in \mathbb{N}} |a_n|$, und A ist beschränkt. Ist umgekehrt (a_n) unbeschränkt, wählen wir eine Teilfolge (a_{n_k}) , so dass $|a_{n_k}| > k$ für $k \in \mathbb{N}$, und setzen $x = (x_n) = \sum_{k=1}^{\infty} e_{n_k}/k$, wobei e_n der n -te Einheitsvektor im ℓ^2 ist. Dann ist $x_{n_k} = 1/k$, wenn $n = n_k$, und $x_n = 0$ sonst. Es gilt $x \in \ell^2$, denn

$$\|x\|^2 = \sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6} < \infty$$

Andererseits ist

$$\|Ax\|^2 = \sum_{n=1}^{\infty} a_n^2 x_n^2 = \sum_{k=1}^{\infty} \underbrace{a_{n_k}^2}_{>k^2} \underbrace{x_{n_k}^2}_{=1/k^2} > \sum_{k=1}^{\infty} 1 = \infty$$

unbeschränkt. Als nächstes behaupten wir, dass $\|A\| = \sup_{n \in \mathbb{N}} |a_n|$. Wir haben bereits $\|A\| \leq \sup_{n \in \mathbb{N}} |a_n|$ gezeigt. Wegen $\|A\| \geq \|Ae_n\| = |a_n|$ für alle $n \in \mathbb{N}$ folgt nach Supremumsbildung die umgekehrte Ungleichung $\|A\| \geq \sup_{n \in \mathbb{N}} |a_n|$, also Gleichheit. \square

Der Raum aller linearen, beschränkten Operatoren ist unter bestimmten Voraussetzungen ein Banachraum.

Proposition 2.42. *Seien X ein normierter Raum und Y ein Banachraum. Dann ist $L(X, Y) = \{A : X \rightarrow Y : A \text{ linear, beschränkt}\}$ auch ein Banachraum.*

Beweis. Wir müssen die Vollständigkeit von $L(X, Y)$ zeigen. Sei (A_n) eine Cauchy-Folge in $L(X, Y)$, d.h.

$$\forall \varepsilon > 0 : \exists N \in \mathbb{N} : \forall n, m \geq N : \|A_n - A_m\| < \varepsilon.$$

Sei $u \in X$. Dann folgt aus der Definition der Norm von A_n für alle $n, m \geq N$

$$\|A_n u - A_m u\| \leq \|A_n - A_m\| \|u\| \leq \varepsilon \|u\|.$$

Da $\varepsilon > 0$ beliebig ist, bedeutet dies, dass $(A_n u)$ eine Cauchy-Folge in Y ist. Wegen der Vollständigkeit von Y ist diese Folge konvergent, $A_n u \rightarrow v$. Wir definieren den Operator $A : X \rightarrow Y$ durch $Au := v$. Er ist linear (Beweis?). Es bleibt zu zeigen, dass er beschränkt ist und dass $\|A_n - A\| \rightarrow 0$ für $n \rightarrow \infty$.

Nun sind Cauchy-Folgen beschränkt, d.h. $\|A_n\| \leq C$ für alle $n \in \mathbb{N}$. Dann ist aber

$$\|Au\| = \left\| \lim_{n \rightarrow \infty} A_n u \right\| = \lim_{n \rightarrow \infty} \|A_n u\| \leq \limsup_{n \rightarrow \infty} \|A_n\| \|u\| \leq C \|u\|,$$

also ist A beschränkt. Sei schließlich $u \in X$ mit $\|u\| \leq 1$. Aus der Eigenschaft, dass (A_n) eine Cauchy-Folge ist, gilt für $n \geq N$ und $m \geq 0$

$$\|A_{n+m} u - A_n u\| \leq \|A_{n+m} - A_n\| \|u\| \leq \|A_{n+m} - A_n\| < \varepsilon.$$

Im Grenzwert $m \rightarrow \infty$ folgt $\|Au - A_n u\| \leq \varepsilon$. Da u beliebig ist, können wir auch

$$\|A - A_n\| = \sup_{\|u\| \leq 1} \|Au - A_n u\| \leq \varepsilon$$

schreiben. Andererseits ist $\varepsilon > 0$ ebenfalls beliebig, also ist $\|A - A_n\| \rightarrow 0$ für $n \rightarrow \infty$. \square

In vielen Anwendungen sind lineare Gleichungen der Form $Au = b$ zu lösen. Im \mathbb{R}^n wissen wir, dass es eine eindeutige Lösung gibt, wenn $A \in \mathbb{R}^{n \times n}$ invertierbar ist oder wenn A positiv definit im Sinne von $x^\top Ax > 0$ für $x \neq 0$ ist. Dies ist auch bei Hilberträumen eine hinreichende Bedingung. Zunächst benötigen wir die folgende Definition.

Definition 2.43. Seien H_1 und H_2 Hilberträume und $A : H_1 \rightarrow H_2$ linear. Dann heißt A invertierbar, wenn eine lineare Abbildung $B : H_2 \rightarrow H_1$ existiert mit $AB = I$ und $BA = I$, und wir schreiben $B = A^{-1}$.

Eine invertierbare lineare Abbildung ist bijektiv, denn aus $Au = 0$ folgt $0 = B(Au) = Iu = u$ und für $v \in H_2$ setze $u := Bv$, so dass $Au = ABv = v$. Man kann zeigen, dass bijektive lineare Abbildungen invertierbar sind.

Satz 2.44 (Lösbarkeit linearer Gleichungen). Seien H ein Hilbertraum und $A : H \rightarrow H$ linear, beschränkt und koerziv, d.h., es gebe ein $\kappa > 0$, so dass für alle $u \in H$ gilt

$$\operatorname{Re}(u, Au) \geq \kappa \|u\|^2.$$

Dann existiert für alle $b \in H$ eine eindeutige Lösung der Gleichung $Au = b$, und die inverse Abbildung A^{-1} ist beschränkt.

Beweis. 1. Zuerst bemerken wir, dass aus der Koerzivität und der Cauchy-Schwarz-Ungleichung folgt

$$\kappa \|u\|^2 \leq \operatorname{Re}(u, Au) \leq \|u\| \|Au\|, \quad u \in H,$$

also $\kappa \|u\| \leq \|Au\|$. Wir behaupten nun, dass der Bildraum $R(A)$ abgeschlossen ist. Ist nämlich (u_n) eine Folge mit $Au_n \rightarrow z$ für $n \rightarrow \infty$ (insbesondere ist (Au_n) eine Cauchy-Folge), so folgt aus der obigen Folgerung

$$\kappa \|u_n - u_m\| \leq \|A(u_n - u_m)\| = \|Au_n - Au_m\| \leq \varepsilon$$

für hinreichend große Indizes n, m . Daher ist auch (u_n) eine Cauchy-Folge und wegen der Vollständigkeit von H konvergent. Es sei $u_n \rightarrow u$. Aus der Stetigkeit von A folgt, dass $Au_n \rightarrow Au$, und die Eindeutigkeit des Grenzwerts impliziert $z = Au$. Dies bedeutet, dass z ein Element des Bildraums $R(A)$ ist, d.h., $R(A)$ ist abgeschlossen.

2. Weiter behaupten wir, dass A surjektiv ist, d.h. $R(A) = H$. Angenommen, $R(A) \neq H$. Dann können wir nach dem Projektionssatz H in die direkte Summe $R(A) \oplus R(A)^\perp$ zerlegen, da $R(A)$ abgeschlossen ist. Sei $w \in R(A)^\perp$, $w \neq 0$. Dann sind Aw und w nach Konstruktion orthogonal, also

$$0 = \operatorname{Re}(w, Aw) \geq \kappa \|w\|^2,$$

woraus sich $w = 0$ im Widerspruch zu $w \neq 0$ ergibt. Also ist $R(A) = H$.

3. Die Injektivität von A ist eine Konsequenz der Koerzivität, denn ist $Au_1 = Au_2$, so erhalten wir $0 = \|Au_1 - Au_2\| \geq \kappa \|u_1 - u_2\|$, also $u_1 = u_2$. Die Stetigkeit der Inverse ist ebenfalls eine Konsequenz der Koerzivität. Sei nämlich $Au = v$. Dann folgt

$$\|A^{-1}v\| = \|u\| \leq \kappa^{-1}\|Au\| = \kappa^{-1}\|v\|,$$

also $\|A^{-1}\| \leq \kappa^{-1}$. Der Satz ist bewiesen. \square

Bemerkung 2.45. Wir haben in Satz 2.44 bewiesen, dass die Inverse eines linearen, beschränkten und *koerziven* Operators existiert und beschränkt ist. Wir können die Voraussetzungen abschwächen: Sei $A : H \rightarrow H$ ein linearer, beschränkter und *invertierbarer* Operator. Dann ist die Inverse A^{-1} beschränkt. Der Satz basiert auf dem Kategoriesatz von Baire, den wir hier nicht beweisen wollen; wir verweisen auf die Literatur. \square

Beispiel 2.46. Sei $H = L^2(\Omega)$ ($\Omega \subset \mathbb{R}^n$ ein beschränktes Gebiet) und sei k eine quadratintegrierbare Funktion auf Ω^2 mit

$$q^2 := \int_{\Omega} \int_{\Omega} k(x,y)^2 dx dy < 1.$$

Wir suchen Lösungen der Integralgleichung

$$u(x) = \int_{\Omega} k(x,y)u(y)dy, \quad x \in \Omega. \quad (2.12)$$

Dazu definieren wir den linearen Operator $A : H \rightarrow H$, $Au = u - \int_{\Omega} k(\cdot, y)u(y)dy$. Der Operator ist beschränkt, denn mit $(a+b)^2 \leq 2(a^2 + b^2)$ und der Cauchy-Schwarz-Ungleichung erhalten wir

$$\begin{aligned} \|Au\|^2 &= \int_{\Omega} \left(u(x) - \int_{\Omega} k(x,y)u(y)dy \right)^2 dx \\ &\leq 2 \int_{\Omega} u(x)^2 dx + 2 \int_{\Omega} \int_{\Omega} k(x,y)^2 dy \int_{\Omega} u(y)^2 dy dx \\ &= 2(1 + q^2) \int_{\Omega} u(y)^2 dy = 2(1 + q^2)\|u\|^2. \end{aligned}$$

Die Koerzivität von A ist eine Konsequenz der Cauchy-Schwarz-Ungleichung und der Voraussetzung $q < 1$:

$$\begin{aligned} (u, Au) &= (u, u) - \left(u, \int_{\Omega} k(\cdot, y)u(y)dy \right) \geq \|u\|^2 - \|u\| \left\| \int_{\Omega} k(\cdot, y)u(y)dy \right\| \\ &= \|u\|^2 - \|u\| \left(\int_{\Omega} \left(\int_{\Omega} k(x,y)u(y)dy \right)^2 dx \right)^{1/2} \\ &\geq \|u\|^2 - \|u\| \left(\int_{\Omega} \int_{\Omega} k(x,y)^2 dy \underbrace{\int_{\Omega} u(y)^2 dy}_{=\|u\|^2} dx \right)^{1/2} \\ &= \|u\|^2 - \|u\|^2 \left(\int_{\Omega} \int_{\Omega} k(x,y)^2 dy dx \right)^{1/2} = (1 - q)\|u\|^2. \end{aligned}$$

Also ist $\kappa = 1 - q > 0$. Aus Satz 2.44 schließen wir, dass es eine eindeutige Lösung $u \in L^2(\Omega)$ der Integralgleichung (2.12) gibt. \square

Für spätere Verwendung zeigen wir folgendes Resultat über sogenannte Neumann-Reihen.

Lemma 2.47 (Neumann-Reihe). *Seien X ein Banachraum und $A : X \rightarrow X$ ein linearer beschränkter Operator mit $\|A\| < 1$. Dann ist $A - I$ invertierbar, die Inverse $(A - I)^{-1}$ ist beschränkt, und es gilt*

$$(A - I)^{-1} = - \sum_{n=0}^{\infty} A^n,$$

wobei I den Identitätsoperator auf X bezeichne.

Beweis. Definiere die Partialsummen $S_n = \sum_{k=0}^n A^k$. Aus $\|A^2 u\| \leq \|A\| \|Au\| \leq \|A\|^2 \|u\|$, also $\|A^2\| \leq \|A\|^2$, folgern wir allgemein $\|A^k\| \leq \|A\|^k$, $k \in \mathbb{N}$. Für $n > m$ folgt

$$\|S_n - S_m\| = \left\| \sum_{k=m+1}^n A^k \right\| \leq \sum_{k=m+1}^n \|A^k\| \leq \sum_{k=m+1}^n \|A\|^k.$$

Wegen $\|A\| < 1$ ist die Reihe $\sum_{k=0}^{\infty} \|A\|^k$ konvergent. Dann ist wegen der obigen Abschätzung (S_n) eine Cauchy-Folge. Nach Proposition 2.42 ist der Raum $L(X, X)$ aller linearen, beschränkten Abbildungen auf dem Banachraum X vollständig, also ist (S_n) konvergent, und wir können $S = \lim_{n \rightarrow \infty} S_n$ in $L(X, X)$ definieren. Insbesondere ist S beschränkt. Nun ist

$$(A - I)S_n = \sum_{k=0}^n (A^{k+1} - A^k) = A^{n+1} - I.$$

Die linke Seite konvergiert gegen $(A - I)S$, die rechte Seite gegen $-I$, denn $\|A^{n+1}\| \leq \|A\|^{n+1} \rightarrow 0$ für $n \rightarrow \infty$. Also gilt $(A - I)S = -I$. Analog zeigt man $S(A - I) = -I$. Also ist $-S = -\sum_{k=0}^{\infty} A^k$ die (beschränkte) Inverse von $A - I$. \square

2.6 Dualräume und adjungierte Operatoren

Im vorigen Abschnitt haben wir gezeigt, dass die Gleichung $Au = b$ lösbar ist, wenn A koerziv ist. Tatsächlich genügen weit schwächere Voraussetzungen, um diese Gleichung zu lösen. Ist etwa A eine $(n \times n)$ -Matrix, so ist $Au = b$ lösbar (eventuell nicht eindeutig), wenn $b \in \mathbb{R}(A)$ ein Element des Bildraums von A ist. Eine ähnliche Aussage gilt in unendlich-dimensionalen Räumen. Dazu benötigen wir zuerst den Begriff des Dualraums und des adjungierten Operators. Wir erinnern, dass wir Abbildungen $F : X \rightarrow \mathbb{R}$ oder $F : X \rightarrow \mathbb{C}$ (X ein metrischer Raum) *Funktionale* nennen.

Definition 2.48 (Dualraum). *Sei X ein normierter Raum. Der Dualraum X' von X ist die Menge aller linearen, stetigen Funktionale auf X .*

Da Funktionale spezielle Operatoren sind, gelten natürlich alle Resultate aus Abschnitt 2.5 auch für Funktionale. Insbesondere ist die Norm von $F \in X'$ definiert durch

$$\|F\| = \sup_{\|u\| \neq 0} \frac{|F(u)|}{\|u\|} = \sup_{\|u\| \leq 1} |F(u)| = \sup_{\|u\|=1} |F(u)|.$$

Wenn X ein Banachraum ist, so ist X' mit dieser Norm auch ein Banachraum.

Beispiel 2.49. (1) Seien H ein Prä-Hilbertraum und $u \in H$. Dann ist $F(v) = (u, v)$, $v \in H$, ein lineares, stetiges Funktional, denn mit der Cauchy-Schwarz-Ungleichung folgt

$$|F(v)| \leq \|u\| \|v\|, \quad v \in H,$$

also $|F(v)| \leq C\|v\|$ mit $C = \|u\|$.

(2) Sei $X = L^1(0,1)$ versehen mit der Norm $\|f\| = \int_0^1 |f(x)|dx$ und sei $F(f) = \int_0^1 f(x)dx$, $f \in X$. Aus

$$|F(f)| \leq \int_0^1 |f(x)|dx = \|f\|$$

folgt $F \in X'$. □

Interessanterweise können alle linearen, stetigen Funktionale $F \in H'$ in der Form $F(v) = (f, v)$ für ein $f \in H$ geschrieben werden, sofern H ein Hilbertraum ist.

Satz 2.50 (Darstellungssatz von Riesz). *Sei H ein Hilbertraum und $F \in H'$. Dann existiert ein eindeutiges $f \in H$ mit $\|F\| = \|f\|$ und*

$$F(u) = (f, u) \quad \text{für alle } u \in H.$$

Wegen der eindeutigen Relation zwischen F und f können wir einem Hilbertraum mit seinen Dualraum H' identifizieren, und wir schreiben (etwas unpräzise) $H = H'$.

Beweis. Wir bemerken zunächst, dass der Nullraum $N(F)$ von F abgeschlossen ist, denn aus $(u_n) \subset N(F)$ mit $u_n \rightarrow u$, also $F(u_n) = 0$, folgt aus der Stetigkeit von F sofort $0 = F(u_n) \rightarrow F(u)$, also $F(u) = 0$ und $u \in N(F)$. Falls $N(F) = H$, dann ist $F = 0$, und es ist nichts zu beweisen (wähle $f = 0$). Sei also $N(F) \neq H$. Wähle $u_0 \in H$ mit $u_0 \notin N(F)$. Nach dem Projektionssatz existieren $v \in N(F)$ und $w \in N(F)^\perp$, so dass $u_0 = v + w$. Da u_0 kein Element von $N(F)$ ist, muss $w \neq 0$, also $w \notin N(F)$ gelten. Dies bedeutet, dass $F(w) \neq 0$. Sei $u \in H$ und definiere

$$u_1 = u - \frac{F(u)}{F(w)}w.$$

Dann ist $F(u_1) = 0$, d.h. $u_1 \in N(F)$. Folglich sind u_1 und w orthogonal:

$$0 = (w, u_1) = \left(w, u - \frac{F(u)}{F(w)}w \right) = (w, u) - \frac{F(u)}{F(w)}\|w\|^2.$$

Dies ist äquivalent zu

$$F(u) = \left(\frac{\overline{F(w)}}{\|w\|^2} w, u \right),$$

und die Behauptung folgt mit $f = \overline{F(w)}w/\|w\|^2$.

Es bleibt $\|F\| = \|f\|$ zu zeigen. Zunächst bemerken wir, dass

$$|F(u)| = |(f, u)| \leq \|f\| \|u\|, \quad \text{also } \|F\| \leq \|f\|.$$

Andererseits ist

$$\|F\| = \sup_{\|u\| \neq 0} \frac{|F(u)|}{\|u\|} \geq \frac{|F(f)|}{\|f\|} = \frac{|(f, f)|}{\|f\|} = \|f\|.$$

Wir schließen $\|F\| = \|f\|$. □

Beispiel 2.51. Sei $\Omega \subset \mathbb{R}^n$ eine offene Menge. Man kann zeigen (aufwendiger Beweis!), dass der Dualraum von $L^p(\Omega)$ gleich $L^q(\Omega)$ ist:

$$(L^p(\Omega))' = L^q(\Omega), \quad \text{wenn } 1 \leq p < \infty, \quad \frac{1}{p} + \frac{1}{q} = 1.$$

Genauer gesagt sind der Raum $L^p(\Omega)$ und sein Dualraum nicht gleich, aber wir können $(L^p(\Omega))'$ und $L^q(\Omega)$ (isometrisch isomorph) identifizieren, d.h., wir identifizieren jedes $F \in (L^p(\Omega))'$ eindeutig mit einer Funktion $f \in L^q(\Omega)$ vermöge

$$F(u) = \int_{\Omega} f(x)u(x)dx, \quad u \in L^p(\Omega).$$

Der Dualraum von $L^1(\Omega)$ ist gleich $L^\infty(\Omega)$, aber der Dualraum von $L^\infty(\Omega)$ ist *nicht* gleich $L^1(\Omega)$; es gilt nur $L^1(\Omega) \subset (L^\infty(\Omega))'$. □

Mit Hilfe des Darstellungssatzes von Riesz können wir adjungierte Operatoren definieren. Seien dazu H_1 und H_2 Hilberträume und $A : H_1 \rightarrow H_2$ ein linearer, beschränkter Operator. Sei ferner $u \in H_2$. Dann wird durch $\phi_u : H_1 \rightarrow \mathbb{R}$, $\phi_u(v) = (u, Av)$, ein lineares, beschränktes Funktional definiert, d.h. $\phi_u \in H_1'$. Nach dem Darstellungssatz von Riesz existiert ein (eindeutiges) Element aus H_1 , das wir mit A^*u bezeichnen, so dass

$$\phi_u(v) = (A^*u, v). \tag{2.13}$$

Dies definiert einen Operator $A^* : H_2 \rightarrow H_1$.

Definition 2.52 (Adjungierter Operator). Seien H_1, H_2 Hilberträume und $A : H_1 \rightarrow H_2$ linear und beschränkt. Der vermöge (2.13) definierte Operator $A^* : H_2 \rightarrow H_1$ wird der zu A adjungierte Operator genannt. Er erfüllt die Beziehung

$$(u, Av) = (A^*u, v) \quad \text{für alle } u \in H_2, v \in H_1. \tag{2.14}$$

In Gleichung (2.14) steht auf der linken Seite das Skalarprodukt in H_2 , auf der rechten Seite das Skalarprodukt in H_1 . Die obige Argumentation zeigt, dass der Operator A^* wohldefiniert ist. Er ist sogar linear und beschränkt. Um dies einzusehen, genügt es, (2.14) auszunutzen:

$$\begin{aligned}(A^*(u+v), w) &= (u+v, Aw) = (u, Aw) + (v, Aw) \\ &= (A^*u, w) + (A^*v, w) = (A^*u + A^*v, w), \\ (A^*(\lambda u), w) &= (\lambda u, Aw) = \overline{\lambda}(u, Aw) = \overline{\lambda}(A^*u, w) = (\lambda A^*u, w)\end{aligned}$$

für alle $u, v \in H_2$, $w \in H_1$ und $\lambda \in \mathbb{R}$. Dies impliziert $A^*(u+v) = A^*u + A^*v$ und $A^*(\lambda u) = \lambda A^*u$ und damit die Linearität von A^* . Für den Nachweis der Beschränktheit bemerken wir zunächst, dass

$$\|u\| = \sup_{\|v\| \leq 1} |(u, v)|,$$

denn $\sup_{\|v\| \leq 1} |(u, v)| \leq \sup_{\|v\| \leq 1} \|u\| \|v\| = \|u\|$ und $\sup_{\|v\| \leq 1} |(u, v)| \geq |(u, u/\|u\|)| = \|u\|$. Damit erhalten wir

$$\|A^*u\| = \sup_{\|v\| \leq 1} |(A^*u, v)| = \sup_{\|v\| \leq 1} |(u, Av)| \leq \sup_{\|v\| \leq 1} \|u\| \|A\| \|v\| \leq \|A\| \|u\|.$$

Proposition 2.53. Seien H_1, H_2 Hilberträume und $A : H_1 \rightarrow H_2$ linear und beschränkt. Dann ist $A^* : H_2 \rightarrow H_1$ linear und beschränkt, $\|A^*\| = \|A\|$ und $A^{**} = A$.

Beweis. Wir haben bereits bewiesen, dass A^* linear und beschränkt ist. Die obige Rechnung zeigt $\|A^*\| \leq \|A\|$. Andererseits ist

$$\|Au\| = \sup_{\|v\| \leq 1} |(Au, v)| = \sup_{\|v\| \leq 1} |(u, A^*v)| \leq \|A^*\| \|u\|,$$

also $\|A\| \leq \|A^*\|$. Die Beziehung $A^{**} = A$ folgt aus $(A^{**}u, v) = (u, A^*v) = \overline{(A^*v, u)} = \overline{(v, Au)} = (Au, v)$. \square

Beispiel 2.54. (1) Seien $H_1 = H_2 = \mathbb{R}^n$ und $A = (a_{jk}) \in \mathbb{R}^{n \times n}$ eine Matrix, die wir mit einer linearen, beschränkten Abbildung von H_1 nach H_2 identifizieren. Dann ist für alle $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n) \in \mathbb{R}^n$

$$(Ax, y) = \sum_{j,k=1}^n a_{jk} x_k y_j = \sum_{j,k=1}^n x_k a_{kj}^\top y_j = (x, A^\top y),$$

wobei $A^\top = (a_{jk}^\top) = (a_{kj})$ die zu A transponierte Matrix ist. Wir sehen, dass die zu A adjungierte Matrix gleich der transponierten Matrix ist, d.h. $A^* = A^\top$.

(2) Seien $H_1 = H_2 = L^2(\Omega)$ ($\Omega \subset \mathbb{R}^n$ offen) und definiere $A : H_1 \rightarrow H_2$ durch

$$(Au)(x) = \int_{\Omega} k(x, y) u(y) dy, \quad u \in H_1,$$

wobei $k \in L^2(\Omega \times \Omega)$. Wir berechnen mit $u, v \in L^2(\Omega)$

$$\begin{aligned}(Au, v) &= \int_{\Omega} \left(\int_{\Omega} k(x, y) u(y) dy \right) v(x) dx = \int_{\Omega} \left(\int_{\Omega} k(y, x) u(x) dx \right) v(y) dy \\ &= \int_{\Omega} \left(\int_{\Omega} k(y, x) v(y) dy \right) u(x) dx = (u, A^*v),\end{aligned}$$

wobei

$$(A^*v)(x) = \int_{\Omega} k(y, x) v(y) dy.$$

(3) Seien auf $H_1 = H_2 = \ell^2$ der Linksshift L und der Rechtsshift R für $x = (x_n) \in \ell^2$ definiert durch

$$L(x_1, x_2, x_3, \dots) = (x_2, x_3, \dots), \quad R(x_1, x_2, x_3, \dots) = (0, x_1, x_2, \dots).$$

Wir behaupten, dass $R = L^*$ und $L = R^*$. Dies folgt aus

$$(x, Ry) = \sum_{n=1}^{\infty} x_n (Ry)_n = x_2 y_1 + x_3 y_2 + \dots = \sum_{n=1}^{\infty} (Lx)_n y_n = (Lx, y),$$

also $R^* = L$ und $L^* = R^{**} = R$. □

Aus der linearen Algebra wissen wir, dass die direkte Summe des Bildraums einer Matrix $A \in \mathbb{R}^{n \times n}$ und des Nullraums von A^T gleich dem \mathbb{R}^n ist:

$$\mathbb{R}^n = R(A) \oplus N(A^T), \quad A \in \mathbb{R}^{n \times n}.$$

Eine ähnliche Aussage gilt für lineare, beschränkte Operatoren.

Satz 2.55 (Alternativsatz). *Seien H_1, H_2 Hilberträume und $A : H_1 \rightarrow H_2$ linear und beschränkt mit abgeschlossenem Bildraum $R(A)$. Dann gilt*

$$H_2 = R(A) \oplus N(A^*).$$

Beweis. Da $R(A)$ abgeschlossen ist, existiert nach dem Projektionssatz ein Unterraum $V \subset H_2$ mit $H_2 = R(A) \oplus V$. Es gilt $V = R(A)^\perp$. Wir behaupten, dass $V = N(A^*)$. Sei zunächst $v \in N(A^*)$. Dann ist

$$(v, Au) = (A^*v, u) = 0 \quad \text{für alle } u \in H_1,$$

also ist v orthogonal zu $R(A)$, und wir haben $N(A^*) \subset R(A)^\perp$ gezeigt. Sei umgekehrt $v \in R(A)^\perp$. Dann ist v orthogonal zu $R(A)$ und folglich $0 = (v, Au) = (A^*v, u)$ für alle $u \in H_1$. Wir schließen $A^*v = 0$, also $v \in N(A^*)$ und damit $R(A)^\perp \subset N(A^*)$. Daher ist $R(A)^\perp = N(A^*)$. □

Eine Folgerung des Alternativsatzes ist nun eine Aussage, unter welchen Bedingungen die Gleichung $Au = b$ lösbar ist.

Satz 2.56 (Fredholmsche Alternative). *Seien H_1, H_2 Hilberträume und $A : H_1 \rightarrow H_2$ linear und beschränkt mit abgeschlossenem Bildraum $R(A)$. Dann ist die Gleichung $Au = b$ genau dann lösbar, wenn b orthogonal zu $N(A^*)$ ist.*

Beweis. Der Satz ist eine unmittelbare Konsequenz aus dem Alternativsatz, denn $Au = b$ ist für $b \neq 0$ genau dann lösbar, wenn $b \in R(A)$. Wegen $H_2 = R(A) \oplus N(A^*) = R(A) \oplus R(A)^\perp$ und $b \neq 0$ ist dies äquivalent dazu, dass b orthogonal zu $N(A^*)$ ist. \square

Mit Hilfe der Fredholmschen Alternative können wir die Existenz von Lösungen von $Au = b$ aus der eindeutigen Lösbarkeit der Gleichung $A^*v = 0$ schließen, denn unter den Voraussetzungen von Satz 2.56 ist $Au = b$ genau dann lösbar, wenn die Lösung von $A^*v = 0$ eindeutig ist. Ist nämlich $R(A)$ abgeschlossen und ist die Lösung von $A^*v = 0$ eindeutig, so ist $N(A^*) = \{0\}$, und $H_2 = R(A) \oplus N(A^*) = R(A)$, also existiert zu jedem $b \in H_2$ eine Lösung von $Au = b$.

Der Begriff *Alternative* kommt von folgender Dichotomie:

$$\begin{aligned} \text{Entweder } & \forall b \in H_2 : \exists u \in H_1 : Au = b \\ \text{oder } & \exists b \in H_2 : \forall u \in H_1 : Au \neq b. \end{aligned}$$

Die letzte Aussage ist äquivalent zu $R(A) \neq H_2$ oder $N(A^*) = R(A)^\perp \neq \{0\}$, d.h., es gibt ein $v \neq 0$ mit $A^*v = 0$.

Beispiel 2.57. (1) Betrachte den Rechtsshift $R : \ell^2 \rightarrow \ell^2$, $R(x_n) = (0, x_1, x_2, \dots)$. Der Bildraum $R(R)$ ist abgeschlossen, denn eine Folge $x^{(k)} = (x_n^{(k)}) \subset R(R)$ mit $x^{(k)} \rightarrow x$ in ℓ^2 besitzt die Eigenschaft $x_1^{(k)} = 0$, also auch $x_1 = 0$, d.h. $x \in R(R)$. Wir haben weiter oben gezeigt, dass der adjungierte Operator $R^* = L$ gleich dem Linksshift ist. Der Nullraum von L ist der eindimensionale Raum $N(R^*) = N(L) = \{(x_1, 0, 0, \dots) : x_1 \in \mathbb{R}\}$. Unter welchen Bedingungen ist nun die Gleichung

$$Rx = b \quad \Leftrightarrow \quad (0, x_1, x_2, \dots) = (b_1, b_2, b_3, \dots)$$

für $b \in \ell^2$ lösbar? Dies ist offenbar genau dann der Fall, wenn $b_1 = 0$ oder – äquivalent – wenn b orthogonal zu $N(R^*)$ ist. Dies entspricht der Fredholmschen Alternative.

(2) Betrachte den Operator $A : L^2(0, 1) \rightarrow L^2(0, 1)$, $(Au)(x) = xu(x)$, $x \in (0, 1)$. Dann ist A wohldefiniert und $A = A^*$, denn

$$(Au, v) = \int_0^1 (xu(x))v(x)dx = \int_0^1 u(x)(xv(x))dx = (u, Av).$$

Ferner ist $N(A) = \{0\}$, denn aus $0 = (Au)(x) = xu(x)$ folgt sofort $u(x) = 0$ für $0 < x < 1$. Folglich ist jedes $b \in L^2(0, 1)$ orthogonal zu $N(A^*) = N(A)$. Allerdings ist die Gleichung $Au = b$ nicht für jedes $b \in L^2(0, 1)$ lösbar; zum Beispiel ist für $b = 1$ die Gleichung $1 = (Au)(x) = xu(x)$ nicht in $L^2(0, 1)$ lösbar, denn $u(x) = 1/x$ ist nicht quadratintegrierbar in $(0, 1)$. Die Fredholmsche Alternative kann nicht angewendet werden, weil $R(A)$ nicht abgeschlossen in $L^2(0, 1)$ ist (Beweis?). \square

In der linearen Algebra besitzen symmetrische Matrizen spezielle Eigenschaften. Für lineare Operatoren A definieren wir diesen Begriff wie folgt.

Definition 2.58 (Symmetrischer Operator). *Seien H ein Hilbertraum und $A : H \rightarrow H$ ein linearer Operator. Dann heißt A symmetrisch, wenn für alle $u, v \in H$*

$$(Au, v) = (u, Av).$$

Im Falle stetiger Operatoren ist diese Definition zu der Bedingung $A = A^*$ äquivalent.

Beispiel 2.59. (1) Betrachte wie in Beispiel 2.54 den Operator $A : L^2(\Omega) \rightarrow L^2(\Omega)$,

$$(Au)(x) = \int_{\Omega} k(x, y)u(y)dy,$$

wobei $k \in L^2(\Omega \times \Omega)$. Wir haben bereits gezeigt, dass

$$(A^*u)(x) = \int_{\Omega} k(y, x)u(y)dy.$$

Ist k symmetrisch, so ist der Operator A symmetrisch.

(2) Der Operator $A : L^2(0, 1) \rightarrow L^2(0, 1)$, $(Au)(x) = xu(x)$, aus Beispiel 2.57 ist ebenfalls symmetrisch. \square

Interessanterweise ist die Eigenschaft der Symmetrie hinreichend für die Stetigkeit des Operators.

Satz 2.60 (Hellinger-Toeplitz). *Seien H ein Hilbertraum und $A : H \rightarrow H$ ein symmetrischer linearer Operator. Dann ist A stetig.*

Beweis. Wir benutzen das folgende Resultat, das auf dem Kategoriensatz von Baire basiert (für den Beweis verweisen wir auf die Literatur):

Satz vom abgeschlossenen Graphen: Seien H_1, H_2 Hilberträume und $A : H_1 \rightarrow H_2$ ein linearer Operator. Dann ist A stetig genau dann, wenn der Graph $\text{graph}(A) = \{(u, Au) \in H_1 \times H_2 : u \in H_1\}$ von A abgeschlossen ist.

Es genügt also zu zeigen, dass aus $u_n \rightarrow u$ und $Au_n \rightarrow y$ für $n \rightarrow \infty$ die Identität $y = Au$ folgt. Wegen der Linearität genügt es sogar, $u = 0$ zu wählen. Wir nehmen also an, dass $u_n \rightarrow 0$ und $Au_n \rightarrow y$ gilt, und müssen $y = A0 = 0$ beweisen. Dies ist aber eine Konsequenz aus der Stetigkeit des Skalarprodukts:

$$\begin{aligned} \|y\|^2 &= (y, y) = \left(\lim_{n \rightarrow \infty} Au_n, y \right) = \lim_{n \rightarrow \infty} (Au_n, y) = \lim_{n \rightarrow \infty} (u_n, Ay) \\ &= \left(\lim_{n \rightarrow \infty} u_n, Ay \right) = (0, Ay) = 0, \end{aligned}$$

also $y = 0$, wie gewünscht. \square

2.7 Eigenwertprobleme

Aus der linearen Algebra wissen wir, dass symmetrische Matrizen $A \in \mathbb{R}^{n \times n}$ diagonalisierbar sind, d.h., es existiert eine Basis $S = (v_1, \dots, v_n)$ des \mathbb{R}^n aus Eigenvektoren von A mit (reellen) Eigenwerten $(\lambda_1, \dots, \lambda_n)$, so dass

$$S^{-1}AS = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & \ddots & \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}.$$

Wir erinnern, dass $u \neq 0$ ein *Eigenvektor* von A mit zugehörigem *Eigenwert* $\lambda \in \mathbb{C}$ ist, wenn $Au = \lambda u$ gilt. Die Menge aller Eigenwerte wird auch das *Spektrum* der Matrix A genannt.

In diesem Abschnitt wollen wir untersuchen, welche Eigenschaften das Spektrum linearer, beschränkter Operatoren auf (unendlich-dimensionalen) Hilberträumen besitzen. Wir werden weiter unten zeigen, dass lineare, beschränkte Operatoren keine Eigenwerte besitzen müssen. Daher müssen wir den Begriff des Spektrums eines Operators allgemeiner als in der linearen Algebra definieren.

Wir betrachten im gesamten Abschnitt Hilberträume, die über dem Raum der komplexen Zahlen definiert sind.

Definition 2.61 (Spektrum). *Seien H ein Hilbertraum und $A : H \rightarrow H$ ein linearer, beschränkter Operator. Wir definieren die Resolventenmenge $\rho(A)$, das Spektrum $\sigma(A)$ und das Punktspektrum $\sigma_p(A)$:*

$$\begin{aligned} \rho(A) &= \{\lambda \in \mathbb{C} : A - \lambda I \text{ ist invertierbar}\}, \\ \sigma(A) &= \mathbb{C} \setminus \rho(A) = \{\lambda \in \mathbb{C} : A - \lambda I \text{ ist nicht invertierbar}\}, \\ \sigma_p(A) &= \{\lambda \in \mathbb{C} : A - \lambda I \text{ ist nicht injektiv}\}. \end{aligned}$$

Hierbei bezeichnet I die Identität in H . Ist $\lambda \in \sigma_p(A)$, so existiert nach Definition ein $u \in H$ mit $(A - \lambda I)u = 0$ oder $Au = \lambda u$. Elemente des Punktspektrums sind also nichts anderes als Eigenwerte. Das Spektrum umfasst zumindest alle Eigenwerte: $\sigma_p(A) \subset \sigma(A)$, denn eine lineare beschränkte Abbildung ist genau dann bijektiv, wenn sie invertierbar ist. Bei Matrizen bzw. linearen Abbildungen im \mathbb{R}^n ist das Spektrum gleich dem Punktspektrum, da in diesem Fall eine injektive Abbildung automatisch surjektiv, also invertierbar ist. Dies muss bei Operatoren auf unendlichdimensionalen Räumen nicht mehr so sein, siehe die nachfolgenden Beispiele.

Beispiel 2.62. (1) Seien H ein Hilbertraum und $I : H \rightarrow H$ die Identität, definiert durch $Iu = u$, $u \in H$. Die Eigenwertgleichung

$$0 = (I - \lambda I)u = (1 - \lambda)u$$

führt wegen $u \neq 0$ auf $\lambda = 1$. Dies ist der einzige Eigenwert von I ; also ist $\sigma_p(I) = \{1\}$. Um zu überprüfen, ob das Spektrum noch mehr Elemente enthält, müssen wir untersuchen, ob es $\lambda \neq 1$ gibt, so dass $I - \lambda I$ nicht surjektiv ist. Dies ist nicht der Fall; sei nämlich $v \in H$ und setze $u = v/(1 - \lambda) \in H$. Dann ist $(I - \lambda I)u = (1 - \lambda)u = v$, also ist $I - \lambda I$ surjektiv. Dies bedeutet $\sigma(I) = \{1\}$.

- (2) Betrachte den Linksshift $L : \ell^2 \rightarrow \ell^2$, $Lx = (x_2, x_3, \dots)$, $x = (x_n)$. Aus der Eigenwertgleichung

$$0 = (L - \lambda I)x = (x_2 - \lambda x_1, x_3 - \lambda x_2, \dots)$$

folgt $x_2 = \lambda x_1$, $x_3 = \lambda x_2 = \lambda^2 x_1$ etc. und allgemein

$$x_{n+1} = \lambda^n x_1, \quad n \in \mathbb{N}, \quad x_1 \text{ beliebig.}$$

Damit diese Folge in ℓ^2 liegt, muss $|\lambda| < 1$ gelten, denn

$$\|x\|^2 = \sum_{n=0}^{\infty} x_{n+1}^2 = x_1^2 \sum_{n=0}^{\infty} |\lambda|^{2n}$$

konvergiert genau dann, wenn $|\lambda| < 1$. Dies zeigt

$$\sigma_p(L) = \{\lambda \in \mathbb{C} : |\lambda| < 1\}.$$

Das Punktspektrum eines Operators muss also nicht aus isolierten Punkten bestehen. Man kann zeigen, dass das Spektrum $\sigma(L) = \{\lambda : |\lambda| \leq 1\}$ ist.

- (3) Es gibt auch Operatoren, die ein leeres Punktspektrum besitzen. Ein Beispiel ist der Rechtsshift $R : \ell^2 \rightarrow \ell^2$, $Rx = (0, x_1, x_2, \dots)$. Aus der Eigenwertgleichung folgt nämlich

$$0 = (R - \lambda I)x = (-\lambda x_1, x_1 - \lambda x_2, x_2 - \lambda x_3, \dots).$$

Falls $\lambda = 0$, so folgt sofort $x_n = 0$ für alle $n \in \mathbb{N}$, Widerspruch zu $x \neq 0$. Ist $\lambda \neq 0$, so erhalten wir wieder $x_1 = 0$, $x_2 = 0$ etc., Widerspruch. Daher ist $\sigma_p(R) = \emptyset$. Da R nicht surjektiv und damit nicht invertierbar ist, gilt $0 \in \sigma(R)$, aber 0 ist kein Eigenwert von R . Man kann zeigen, dass $\sigma(R) = \{\lambda : |\lambda| \leq 1\}$. \square

Wir können für die Resolventenmenge und das Spektrum topologische Eigenschaften herleiten.

Satz 2.63. *Seien H ein Hilbertraum und $A : H \rightarrow H$ ein linearer, beschränkter Operator. Dann ist $\rho(A)$ offen und $\sigma(A)$ kompakt. Außerdem gilt $\sigma(A) \subset \{\lambda \in \mathbb{C} : |\lambda| \leq \|A\|\}$.*

Beweis. 1. Sei $\lambda \in \rho(A)$, d.h., $(A - \lambda I)^{-1}$ existiert und ist beschränkt (siehe Bemerkung 2.45). Wir müssen zeigen, dass für alle μ aus einer Umgebung von λ die Differenz $\lambda - \mu$ aus $\rho(A)$ ist. Dazu wählen wir $|\mu| < 1/\|(A - \lambda I)^{-1}\|$ und schreiben

$$A - (\lambda - \mu)I = -(A - \lambda I)(-\mu(A - \lambda I)^{-1} - I).$$

Nach dem Lemma 2.47 über die Neumann-Reihe ist $-\mu(A - \lambda I)^{-1} - I$ invertierbar, denn wir haben $\|\mu(A - \lambda I)^{-1}\| < 1$ vorausgesetzt. Dann ist auch $A - (\lambda - \mu)I$ invertierbar, d.h. $\lambda - \mu \in \rho(A)$. Also ist $\rho(A)$ offen.

2. Aus Obigem folgt, dass $\sigma(A) = \mathbb{C} \setminus \rho(A)$ abgeschlossen ist. Wir behaupten, dass für alle $\lambda \in \sigma(A)$ gilt $|\lambda| \leq \|A\|$. Dann ist $\sigma(A)$ abgeschlossen und beschränkt und folglich kompakt. Um die Behauptung zu zeigen, beweisen wir die äquivalente Formulierung $\{|\lambda| > \|A\|\} \subset \rho(A)$. Sei also $|\lambda| > \|A\|$. Dann ist $\|A/\lambda\| < 1$, und nach Lemma 2.47 über die Neumann-Reihe ist $A/\lambda - I = (A - \lambda I)/\lambda$ invertierbar und $\lambda \in \rho(A)$. \square

Man kann sogar zeigen, dass das Spektrum $\sigma(A)$ linearer, beschränkter Operatoren A nicht-leer ist. Der Beweis benötigt Eigenschaften analytischer Funktionen und den Satz von Liouville aus der Funktionentheorie; wir verweisen daher auf die Literatur. Dieses Resultat gilt nur für Hilberträume über \mathbb{C} ; anderenfalls ist es sogar in endlichdimensionalen Räumen falsch. Betrachte nämlich in $H = \mathbb{R}^2$ die lineare Abbildung, die durch die Matrix

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

definiert wird. In diesem Fall gibt es keine (reellen) Eigenwerte.

Beispiel 2.64. Betrachte die unendliche Matrix

Schule

$$A = \begin{pmatrix} 1/2 & 0 & 0 & 0 & 0 & \cdots \\ 1/2 & 1/2 & 0 & 0 & 0 & \cdots \\ 0 & 1/2 & 1/2 & 0 & 0 & \cdots \\ 0 & 0 & 1/2 & 1/2 & 0 & \cdots \\ \vdots & \vdots & & & \vdots & \ddots \end{pmatrix},$$

aufgefasst als ein linearer Operator $A : \ell^\infty \rightarrow \ell^\infty$, die wir in der Einleitung definiert haben, um die Dynamik einer Reihe von Feldinhalten, die sich von links nach rechts bewegen, zu beschreiben. Wir möchten untersuchen, ob es stationäre Zustände gibt, d.h. Vektoren x mit $Ax = x$. Wir müssen also nachprüfen, ob $\lambda = 1$ ein Eigenwert von A ist. Die Eigenwertgleichung $Ax = \lambda x$ lautet komponentenweise

$$\begin{aligned} \frac{1}{2}x_1 &= \lambda x_1, \\ \frac{1}{2}(x_1 + x_2) &= \lambda x_2, \\ \frac{1}{2}(x_2 + x_3) &= \lambda x_3, \quad \text{etc.} \end{aligned}$$

Aus der ersten Gleichung folgt entweder $x_1 = 0$ oder $\lambda = 1/2$. Falls $\lambda = 1/2$, ist die erste Gleichung trivial, aus der zweiten folgt $x_1 = 0$, die dritte impliziert $x_2 = 0$ etc., also $x = 0$. Falls $\lambda \neq 1/2$, muss $x_1 = 0$ gelten. Aus der zweiten Gleichung folgern wir $x_2 = 0$, aus der dritten $x_3 = 0$ etc. Daher ist $x = 0$. Es kann also keine Eigenwerte geben.

Dieses Ergebnis ist verständlich: Die Anwendung von A auf einen Vektor x beschreibt die Dynamik der Feldinhalte, die langsam von links ($n = 1$) nach rechts ($n \rightarrow \infty$) wandern. Nach unendlich langer Zeit sind alle Felder leer. Eine interessantere Dynamik erhalten wir, wenn wir in jedem Zeitschritt in das erste Feld einen neuen Inhalt ablegen. Wir ändern die Matrix daher ab zu

$$B = \begin{pmatrix} p+q & 0 & 0 & 0 & 0 & \cdots \\ 1-q & q & 0 & 0 & 0 & \cdots \\ 0 & 1-q & q & 0 & 0 & \cdots \\ 0 & 0 & 1-q & q & 0 & \cdots \\ \vdots & \vdots & & & \vdots & \ddots \end{pmatrix},$$

wobei $0 < p, q < 1$ zwei Zahlen sind. In das erste Feld wird in jedem Zeitschritt der Inhalt p nachgelegt; ein Teil des Inhalts, nämlich $1 - q$, wird von einem Feld in das rechts benachbarte gelegt, der Inhalt q verbleibt im Feld. Für $p = 0$ und $q = 1/2$ erhalten wir die oben untersuchte Matrix. Die Eigenwertgleichung $Bx = \lambda x$ lautet komponentenweise

$$\begin{aligned}(p + q)x_1 &= \lambda x_1, \\ (1 - q)x_1 + qx_2 &= \lambda x_2, \\ (1 - q)x_2 + qx_3 &= \lambda x_3, \quad \text{etc.}\end{aligned}$$

Falls $x_1 = 0$, folgt aus der zweiten Gleichung $\lambda = q$, aus der dritten Gleichung $(1 - q)x_2 = 0$, also $x_2 = 0$. Die weiteren Gleichungen implizieren $x_n = 0$ für $n \geq 3$ und insbesondere $x = 0$. Sei also $x_1 \neq 0$. Dann muss $\lambda = p + q$ gelten. Aus der zweiten Gleichung folgern wir $(1 - q)x_1 = px_2$ oder $x_2 = ((1 - q)/p)x_1$, aus der dritten Gleichung $(1 - q)x_2 = px_3$ oder $x_3 = ((1 - q)/p)x_2 = ((1 - q)/p)^2 x_1$ etc. Wir erhalten den Eigenwert λ zum Eigenvektor $x = (x_n)$ mit

$$\lambda = p + q, \quad x_{n+1} = \left(\frac{1 - q}{p}\right)^n x_1, \quad x_1 \neq 0.$$

Damit $(x_n) \in \ell^\infty$ ist, muss $(1 - q)/p \leq 1$ gelten.

Für unsere Anwendung ist der Fall $\lambda = 1$ interessant. Dann muss $p = 1 - q$ gewählt werden, und der stationäre Zustand lautet $x = x_1(1, 1, 1, \dots)$ mit beliebigem $x_1 \neq 0$. Nach unendlich langer Zeit ist der Inhalt aller Felder also gleich. \square

Für symmetrische Operatoren können wir weitere Aussagen über das Spektrum machen.

Proposition 2.65. *Seien H ein Hilbertraum und $A : H \rightarrow H$ ein linearer, beschränkter und symmetrischer Operator. Dann gilt: Alle Eigenwerte sind reell, d.h. $\sigma_p(A) \subset \mathbb{R}$, und Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal.*

Beweis. Sei $\lambda \in \sigma_p(A)$, d.h., es existiert ein $u \neq 0$ mit $Au = \lambda u$. Wegen

$$\lambda(u, u) = (u, \lambda u) = (u, Au) = (Au, u) = (\lambda u, u) = \bar{\lambda}(u, u)$$

folgt $\lambda = \bar{\lambda}$, also $\lambda \in \mathbb{R}$. Sind $\lambda \neq \mu \in \sigma_p(A)$ mit $Au = \lambda u$ und $Av = \mu v$, so folgt

$$(\lambda - \mu)(u, v) = (\lambda u, v) - (u, \mu v) = (Au, v) - (u, Av) = 0$$

und wegen $\lambda - \mu \neq 0$ dann $(u, v) = 0$. \square

Man kann sogar zeigen, dass das Spektrum eines symmetrischen Operators A reell ist. Dann folgt mit Satz 2.63, dass $\sigma(A) \subset [-\|A\|, \|A\|]$.

Sei $A \in \mathbb{R}^{n \times n}$ eine symmetrische Matrix. Dann gibt es eine Orthonormalbasis $S = (u_1, \dots, u_n)$ aus Eigenvektoren von A mit zugehörigen Eigenwerten $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. Ist weiterhin $u \in \mathbb{R}^n$ mit $u = \sum_{k=1}^n (u_k, u)u_k \in \mathbb{R}^n$, so folgt

$$Au = \sum_{k=1}^n (u_k, u)Au_k = \sum_{k=1}^n \lambda_k (u_k, u)u_k.$$

Wir sagen, dass wir den Operator A in seine Eigenvektoren “entwickelt” haben. Gilt dies auch für lineare, beschränkte Operatoren in (unendlich-dimensionalen) Hilberträumen? Im Allgemeinen nein. Besitzt der Operator jedoch eine zusätzliche Eigenschaft (Kompaktheit), so gilt die obige Entwicklung, wobei die Summe durch eine Reihe ersetzt werden muss. Zuerst definieren wir den Begriff der Kompaktheit.

Definition 2.66 (Kompakter Operator). *Seien H_1, H_2 Hilberträume und $A : H_1 \rightarrow H_2$ ein linearer Operator. Dann heißt A kompakt, wenn es zu jeder beschränkten Folge $(u_n) \subset H_1$ eine Teilfolge (u_{n_k}) gibt, so dass (Au_{n_k}) in H_2 konvergiert.*

Beispiel 2.67. Sei $A : L^2(0, 1) \rightarrow \mathbb{R}$, $Au = \int_0^1 u(x) dx$, ein linearer Operator (sogar ein lineares Funktional). Wir behaupten, dass A kompakt ist. Sei nämlich (u_n) eine beschränkte Folge in $L^2(0, 1)$. Dann ist nach der Cauchy-Schwarz-Ungleichung

$$|Au_n| \leq \int_0^1 |u_n(x)| dx \leq \left(\int_0^1 dx \right)^{1/2} \left(\int_0^1 |u_n(x)|^2 dx \right)^{1/2} = \|u_n\|_{L^2}$$

eine beschränkte Zahlenfolge. Nach dem Satz von Bolzano-Weierstraß enthält (Au_n) eine konvergente Teilfolge. Also ist A kompakt. \square

Es stellt sich die Frage, ob lineare Operatoren, deren Bildraum “klein” ist, immer kompakt sind. Dies gilt in der Tat für endlichdimensionale Bildräume.

Proposition 2.68. *Seien H_1, H_2 Hilberträume und $A : H_1 \rightarrow H_2$ ein linearer und stetiger Operator. Sei ferner $\dim H_2 < \infty$. Dann ist A kompakt.*

Beweis. Sei (u_n) eine beschränkte Folge. Dann ist wegen

$$\|Au_n\| \leq \|A\| \|u_n\|$$

auch (Au_n) eine beschränkte Folge in einem endlichdimensionalen Raum. Wir wissen aus der Analysis, dass in solchen Räumen die abgeschlossenen Kugeln $\overline{B_r(0)}$ für alle $r > 0$ kompakt sind, d.h., jede Folge aus $\overline{B_r(0)}$ enthält eine konvergente Teilfolge. Die Folge (Au_n) liegt in einer solchen Kugel, also existiert eine Teilfolge (Au_{n_k}) mit $Au_{n_k} \rightarrow y$ für ein $y \in H_2$. Dies beweist die Kompaktheit von A . \square

Bemerkung 2.69. Kompakte Operatoren $A : H_1 \rightarrow H_2$ sind immer stetig, denn das Bild der Einheitskugel ist relativ kompakt und somit beschränkt. Daraus folgt sofort die Beschränktheit des Operators und damit seine Stetigkeit. \square

Das Hauptresultat ist der folgende Satz (für den aufwendigen Beweis verweisen wir auf die Literatur).

Satz 2.70 (Spektralsatz für kompakte, symmetrische Operatoren). *Sei $A : H \rightarrow H$ ein linearer, kompakter und symmetrischer Operator auf einem Hilbertraum H . Dann gilt:*

(i) Das Spektrum $\sigma(A)$ enthält höchstens abzählbar viele Zahlen:

$$\sigma(A) \subset \{0\} \cup \{\lambda_n \in \mathbb{R} : \lambda_n \neq 0, n \in N\},$$

wobei $N = \mathbb{N}$ oder N ist eine endliche Menge, und die Folge (λ_n) besitzt keinen Häufungspunkt oder häuft sich höchstens bei null.

(ii) Es gibt ein vollständiges Orthonormalsystem (u_n) aus Eigenvektoren von A mit dazugehörigen Eigenwerten (λ_n) , so dass

$$Au = \sum_{k \in N} \lambda_k (u_k, u) u_k, \quad u \in H.$$

Beispiel 2.71. Definiere den Operator $A : L^2(0, 1) \rightarrow L^2(0, 1)$, $Af = u$, wobei u die eindeutige Lösung der Differentialgleichung

$$u''(x) = f(x), \quad x \in (0, 1), \quad u(0) = u(1) = 0,$$

sei. Diese Lösung können wir durch zweimaliges Integrieren von f erhalten; die beiden Integrationskonstanten können wir mit Hilfe der Randbedingungen $u(0) = 0$ und $u(1) = 0$ berechnen. Um die Eigenwerte von A zu berechnen, sei $f \in L^2(0, 1)$ ein Eigenvektor zum Eigenwert λ . Dann folgt

$$u = Af = \lambda f = \lambda u'',$$

also ist die Gleichung

$$u''(x) = \frac{1}{\lambda} u(x), \quad x \in (0, 1), \quad u(0) = u(1) = 0,$$

zu lösen. Wegen der Randbedingungen $u(0) = u(1) = 0$ kommen als Lösungen nur trigonometrische Funktionen in Frage, nämlich $u(x) = \sin(n\pi x)$ mit $n \in \mathbb{N}$. Setzen wir diesen Ansatz in die Differentialgleichung $u'' = \lambda^{-1}u$ ein, so folgt $\lambda_n = -1/(n\pi)^2$.

Wir fassen zusammen: Der Operator A besitzt die Eigenwerte und Eigenvektoren

$$\lambda_n = -\frac{1}{(n\pi)^2}, \quad u_n(x) = \sin(n\pi x), \quad n \in \mathbb{N}.$$

Die Folge (λ_n) konvergiert gegen null für $n \rightarrow \infty$, also ist null ihr einziger Häufungspunkt. Man kann übrigens zeigen, dass A ein kompakter Operator ist (dies benötigt allerdings vertiefte Kenntnisse der Funktionalanalysis), so dass der Spektralsatz anwendbar ist. \square

3 Parameterintegrale

3.1 Definition und Eigenschaften

In diesem Kapitel betrachten wir Integrale, die von einem Parameter abhängen:

$$F(y) = \int_a^b f(x,y)dx, \quad y \in I.$$

Der Parameter lautet hier y , und I ist das Parameterintervall. Damit das obige Integral definiert ist, nehmen wir an, dass für jedes $y \in I$ die Funktion $f(\cdot, y) : (a, b) \rightarrow \mathbb{R}$ integrierbar ist. Wir untersuchen nun, unter welchen Bedingungen F stetig oder differenzierbar ist. Eine notwendige Bedingung für die Stetigkeit von F ist sicherlich die Stetigkeit von $y \mapsto f(x, y)$. Genügt diese Bedingung? Um die Antwort zu motivieren, formulieren wir die Stetigkeit von F genauer: Die Funktion F ist stetig in I , wenn für alle $y \in I$ und alle Folgen $y_n \rightarrow y$ ($n \rightarrow \infty$)

$$\lim_{n \rightarrow \infty} F(y_n) = \lim_{n \rightarrow \infty} \int_a^b f(x, y_n)dx = \int_a^b \lim_{n \rightarrow \infty} f(x, y_n)dx = \int_a^b f(x, y)dx = F(y), \quad (3.1)$$

denn die Stetigkeit von f impliziert $\lim_{n \rightarrow \infty} f(x, y_n) = f(x, y)$. Im Allgemeinen können wir nicht einfach den Grenzwert $n \rightarrow \infty$ in das Integral hineinziehen. Die Antwort auf die obige Frage ist also i.Allg. nein. Aus der Analysis ist bekannt, dass dies doch möglich ist, wenn die Funktion im Integranden *gleichmäßig stetig* ist. Dazu definieren wir:

Definition 3.1 (Gleichmäßige Stetigkeit). *Seien X, Y zwei normierte Räume. Die Funktion $f : X \rightarrow Y$ heißt gleichmäßig stetig, wenn*

$$\forall \varepsilon > 0 : \exists \delta > 0 : \forall x, x_0 \in X : \|x - x_0\| < \delta \Rightarrow \|f(x) - f(x_0)\| < \varepsilon.$$

Im Gegensatz zur punktweisen Stetigkeit hängt δ nur von ε , aber *nicht* vom Punkt x_0 ab. Gleichmäßige Stetigkeit ist stärker als punktweise Stetigkeit: Jede gleichmäßig stetige Funktion ist stetig in jedem Punkt. Aus der Analysis ist bekannt, dass stetige Funktionen auf einer *kompakten* Menge gleichmäßig stetig sind (dies ist der Satz von Heine). Insbesondere sind stetige Funktionen $f : I \times J \rightarrow \mathbb{R}$ gleichmäßig stetig, wenn I und J beschränkte, abgeschlossene Intervalle sind. In diesem Fall folgt aus der Stetigkeit bzw. Differenzierbarkeit von f bereits die Stetigkeit bzw. Differenzierbarkeit von F , wie der folgende Satz zeigt.

Satz 3.2 (Parameterintegrale). *Seien $a < b \in \mathbb{R}$, $I \subset \mathbb{R}$ ein Intervall, $f : [a, b] \times I \rightarrow \mathbb{R}$ eine gegebene Funktion, so dass $F : I \rightarrow \mathbb{R}$, definiert durch*

$$F(y) = \int_a^b f(x,y)dx, \quad y \in I,$$

wohldefiniert ist.

(i) *Sei f stetig. Dann ist F auch stetig. (Insbesondere gilt (3.1).)*

(ii) Seien f und $\partial f/\partial y$ stetig. Dann ist F differenzierbar, und es gilt

$$F'(y) = \int_a^b \frac{\partial f}{\partial y}(x, y) dx, \quad y \in I.$$

Beweis. Wir nehmen an, dass I ein beschränktes, abgeschlossenes Intervall ist. Anderenfalls können wir den Satz für beschränkte, abgeschlossene Intervalle beweisen und I durch solche Intervalle approximieren.

(i) Da $J = [a, b] \times I$ beschränkt und abgeschlossen im \mathbb{R}^2 ist, ist diese Menge kompakt, und f ist gleichmäßig stetig auf J . Daher gibt es zu $\varepsilon > 0$ ein $\delta > 0$, so dass für alle $(x, y), (x_0, y_0) \in J$ mit $|x - x_0| + |y - y_0| < \delta$ folgt $|f(x, y) - f(x_0, y_0)| < \varepsilon$. Dies impliziert für alle $y, y_0 \in I$ mit $|y - y_0| < \delta$

$$|F(y) - F(y_0)| \leq \int_a^b |f(x, y) - f(x, y_0)| dx < \varepsilon(b - a).$$

Da $\varepsilon > 0$ beliebig ist, folgt die Stetigkeit von F .

(ii) Seien $y, y + h \in I$ mit $h \neq 0$. Dann ist nach dem Mittelwertsatz

$$\begin{aligned} \frac{1}{h}(F(y+h) - F(y)) - \int_a^b \frac{\partial f}{\partial y}(x, y) dx &= \int_a^b \left(\frac{1}{h}(f(x, y+h) - f(x, y)) - \frac{\partial f}{\partial y}(x, y) \right) dx \\ &= \int_a^b \left(\frac{\partial f}{\partial y}(x, z_h) - \frac{\partial f}{\partial y}(x, y) \right) dx, \end{aligned}$$

wobei die Zwischenstelle z_h zwischen y und $y + h$ liegt. Klarerweise gilt $z_h \rightarrow y$ für $h \rightarrow 0$. Nun ist auch $\partial f/\partial y$ gleichmäßig stetig auf J ; daher können wir gemäß (i) den Grenzwert $h \rightarrow 0$ mit dem Integral vertauschen und erhalten

$$\lim_{h \rightarrow 0} \left(\frac{1}{h}(F(y+h) - F(y)) - \int_a^b \frac{\partial f}{\partial y}(x, y) dx \right) = \int_a^b \lim_{h \rightarrow 0} \left(\frac{\partial f}{\partial y}(x, z_h) - \frac{\partial f}{\partial y}(x, y) \right) dx = 0.$$

Dies beweist (ii). □

Beispiel 3.3. Die Funktion $f(x, y) = x^y$, $(x, y) \in [a, b]^2$, ist differenzierbar für jedes $0 < a < b$ und

$$\frac{\partial f}{\partial y}(x, y) = \frac{\partial}{\partial y} e^{y \ln x} = e^{y \ln x} \ln x = x^y \ln x.$$

Daher ist auch $F(y) = \int_a^b f(x, y) dx$ differenzierbar und

$$F'(y) = \int_a^b x^y \ln x dx.$$

Nun kann das Integral F explizit berechnet werden:

$$F(y) = \left[\frac{x^{y+1}}{y+1} \right]_{x=a}^{x=b},$$

woraus wir unmittelbar die Ableitung

$$F'(y) = \frac{d}{dy} \left[\frac{x^{y+1}}{y+1} \right]_{x=a}^{x=b} = \left[\frac{x^{y+1} \ln x}{y+1} - \frac{x^{y+1}}{(y+1)^2} \right]_{x=a}^{x=b}$$

erhalten. Dies impliziert die Formel

$$\int x^y \ln x dx = \frac{x^{y+1} \ln x}{y+1} - \frac{x^{y+1}}{(y+1)^2}.$$

Parameterintegrale können also helfen, komplizierte Integrale zu berechnen. \square

Viele Parameterintegrale in den Anwendungen sind über unbeschränkte Intervalle definiert, und der obige Satz kann nicht angewendet werden. Wir betrachten daher im Folgenden uneigentliche Parameterintegrale. Hierfür benötigen wir, dass das Parameterintegral gleichmäßig bezüglich y konvergiert. Genauer: Sei $f : [a, b) \times I \rightarrow \mathbb{R}$ eine Funktion, so dass für alle $y \in I$ die Abbildung $x \mapsto f(x, y)$ auf jedem kompakten Teilintervall von $[a, b)$ integrierbar ist und $\int_a^b f(x, y) dx$ existiert. Wir sagen, dass das Integral $\int_a^b f(x, y) dx$ *gleichmäßig in $y \in I$ konvergiert*, wenn

$$\forall \varepsilon > 0 : \exists c \in (a, b) : \forall y \in I : \left| \int_c^b f(x, y) dx \right| < \varepsilon.$$

Satz 3.4 (Uneigentliche Parameterintegrale). *Seien $-\infty < a < b \leq \infty$, $f : [a, b) \times I \rightarrow \mathbb{R}$ eine gegebene Funktion und $F : I \rightarrow \mathbb{R}$, definiert durch*

$$F(y) = \int_a^b f(x, y) dx, \quad y \in I,$$

ein Parameterintegral.

(i) *Sei f stetig und $F(y)$ konvergiere gleichmäßig in $y \in I$. Dann ist F stetig.*

(ii) *Seien f und $\partial f / \partial y$ stetig. Es konvergiere F für ein $y_0 \in I$ und es konvergiere das Integral $\int_a^b (\partial f / \partial y)(x, y) dx$ gleichmäßig in $y \in I$. Dann konvergiert F gleichmäßig in $y \in I$, F ist differenzierbar, und es gilt*

$$F'(y) = \int_a^b \frac{\partial f}{\partial y}(x, y) dx, \quad y \in I.$$

Beweis. Der Beweis ist ähnlich wie der des vorherigen Satzes. Wir zeigen daher nur den ersten Teil. Sei I wieder ohne Einschränkung ein beschränktes, abgeschlossenes Intervall. Sei weiter $\varepsilon > 0$. Die Funktion f ist gleichmäßig stetig auf $[a, c] \times I$ für jedes $a < c < b$. Sei c so, dass nach Definition der gleichmäßigen Konvergenz des Integrals

$$\left| \int_c^b f(x, y) dx \right| < \varepsilon \quad \text{für alle } y \in I.$$

Aus der gleichmäßigen Stetigkeit von f folgt die Existenz von $\delta > 0$, so dass für alle (x, y) , $(x_0, y_0) \in [a, c] \times I$ mit $|x - x_0| + |y - y_0| < \delta$ folgt $|f(x, y) - f(x_0, y_0)| < \varepsilon/(c - a)$. Dies ergibt insbesondere für $|y - y_0| < \delta$

$$|F(y) - F(y_0)| \leq \int_a^c \underbrace{|f(x, y) - f(x, y_0)|}_{< \varepsilon/(c-a)} dx + \underbrace{\int_c^b (|f(x, y)| + |f(x, y_0)|) dx}_{\leq 2\varepsilon} < 3\varepsilon.$$

Da $\varepsilon > 0$ beliebig ist, folgt die Stetigkeit von F . □

Beispiel 3.5. Die *Gamma-Funktion* ist ein wichtiges Beispiel für ein uneigentliches Parameterintegral:

$$\Gamma(y) = \int_0^\infty x^{y-1} e^{-x} dx, \quad y \in (0, \infty).$$

Wir behaupten, dass Γ gleichmäßig in $y \in [a, b]$ für alle $0 < a < b < \infty$ konvergiert. Das Integral ist uneigentlich an der oberen und, falls $y < 1$, an der unteren Integralgrenze. Wir zeigen die gleichmäßige Konvergenz nur für die obere Grenze; der Beweis für die untere Grenze geht analog. Seien also $y \in [a, b]$ und $R \geq 1$. Wir integrieren mehrfach partiell:

$$\begin{aligned} \int_R^\infty x^{y-1} e^{-x} dx &= \int_R^\infty x^{y-1} (-e^{-x})' dx = \underbrace{[-x^{y-1} e^{-x}]_R^\infty}_{=R^{y-1} e^{-R}} + (y-1) \int_R^\infty x^{y-2} e^{-x} dx \\ &= R^{y-1} e^{-R} + (y-1) \underbrace{[-x^{y-2} e^{-x}]_R^\infty}_{=R^{y-2} e^{-R}} + (y-1)(y-2) \int_R^\infty x^{y-3} e^{-x} dx \\ &= (R^{y-1} + (y-1)R^{y-2}) e^{-R} + (y-1)(y-2) \int_R^\infty x^{y-3} e^{-x} dx. \end{aligned}$$

Hierbei haben wir für den Grenzwert $\lim_{x \rightarrow \infty} x^{y-1} e^{-x} = 0$ die Regel von de l'Hospital verwendet. Führen wir die partielle Integration m -mal durch, so erhalten wir für eine gewisse höchstens polynomiell wachsende Funktion $g(R)$ und eine Konstante $C > 0$, die von a und b abhängt,

$$\int_R^\infty x^{y-1} e^{-x} dx = g(R) e^{-R} + C \int_R^\infty x^{y-m} e^{-x} dx.$$

Wählen wir $m \geq b + 2$, also $y - m \leq b - m \leq -2$, so können wir folgendermaßen abschätzen:

$$\begin{aligned} 0 &\leq \int_R^\infty x^{y-1} e^{-x} dx \leq |g(R)| e^{-R} + C \int_R^\infty x^{-2} e^{-x} dx \leq C_1 R^{b-1} e^{-R} + C R^{-2} \int_R^\infty e^{-x} dx \\ &\leq C_1 R^{b-1} e^{-R} + C_2 R^{-2} e^{-R}, \end{aligned}$$

wobei C_1 und $C_2 > 0$ unabhängig von $y \in [a, b]$ sind (denn wir können y durch $|y| \leq \max\{|a|, |b|\}$ abschätzen). Die rechte Seite konvergiert gleichmäßig in $y \in [a, b]$ gegen null für $R \rightarrow \infty$.

Also sind die Voraussetzungen von Satz 3.4 erfüllt, und die Gamma-Funktion ist stetig in jedem $[a, b]$, also in $[0, \infty)$. Sie ist sogar auf $(0, \infty)$ unendlich oft differenzierbar mit

$$\Gamma^{(n)}(y) = \int_0^\infty x^{y-1} (\ln x)^n e^{-x} dx,$$

rwias mit vollständiger Induktion bewiesen werden kann. □

Die Gamma-Funktion ist eine Verallgemeinerung der Fakultät, wie das folgende Resultat zeigt (Übungsaufgabe).

Proposition 3.6. Für alle $y > 0$ gilt

$$\Gamma(y+1) = y\Gamma(y).$$

Insbesondere ist $\Gamma(n+1) = n!$ für alle $n \in \mathbb{N}$.

Die Gamma-Funktion ist nützlich, um eine einfache Formel für das Volumen der n -dimensionalen Sphäre mit Radius $R > 0$ zu formulieren. Es gilt nämlich

$$V_n = \frac{\pi^{n/2} R^n}{\Gamma(n/2 + 1)}, \quad n \in \mathbb{N}.$$

Für $n = 3$ folgt wegen $\Gamma(5/2) = 3\sqrt{\pi}/4$ die bekannte Formel $V_3 = 4\pi^3/2 R^3 / (3\sqrt{\pi}) = (4/3)\pi R^3$.

3.2 Laplace-Transformation

In den Anwendungen der Elektrotechnik spielt ein bestimmtes uneigentliches Parameterintegral eine wichtige Rolle, nämlich die Laplace-Transformation.

Definition 3.7 (Laplace-Transformation). Die Laplace-Transformation einer Funktion $f : [0, \infty) \rightarrow \mathbb{C}$ ist definiert durch

$$(\mathcal{L}[f])(s) = \int_0^\infty e^{-st} f(t) dt, \quad s \in S,$$

und $S \subset \mathbb{C}$ ist eine Menge, für die das Integral existiert.

Wir müssen klären, unter welchen Voraussetzungen an f die Laplace-Transformation $\mathcal{L}[f]$ existiert.

Proposition 3.8. Die Funktion $f : [0, \infty) \rightarrow \mathbb{C}$ sei stückweise stetig und es existieren $M, s_0 > 0$, so dass

$$|f(t)| \leq M e^{s_0 t} \quad \text{für alle } t \geq 0.$$

Dann existiert die Laplace-Transformation $\mathcal{L}[f]$ in der Halbebene $\operatorname{Re}(s) > s_0$. Sie ist dort unendlich oft stetig differenzierbar.

Beweis. Sei $s = \alpha + i\beta$, $\alpha, \beta \in \mathbb{R}$, mit $\alpha > s_0$. Die Existenzaussage folgt sofort aus der Abschätzung

$$\int_0^T |e^{-st} f(t)| dt = \int_0^T e^{-\alpha t} |f(t)| dt \leq M \int_0^T e^{(s_0 - \alpha)t} dt = \frac{M}{\alpha - s_0} (1 - e^{-(\alpha - s_0)T})$$

für alle $T > 0$. Das Integral konvergiert gleichmäßig in s , denn

$$\left| \int_T^\infty e^{-st} f(t) dt \right| \leq M \int_T^\infty e^{(s_0 - \alpha)t} dt = \frac{M}{\alpha - s_0} e^{-(\alpha - s_0)T}$$

ist unabhängig von s kleiner als ein vorgegebenes $\varepsilon > 0$, sofern $T > 0$ hinreichend groß gewählt wird. Also ist $\mathcal{L}[f]$ stetig. Die Aussage über die Differenzierbarkeit wird analog bewiesen. \square

Beispiel 3.9. (1) Sei $f(t) = e^{at}$ mit $a > 0$. Sei $s = \alpha + i\beta$ mit $\alpha, \beta \in \mathbb{R}$. Dann ist

$$(\mathcal{L}[f])(s) = \int_0^{\infty} e^{-(s-a)t} dt = \left[\frac{e^{-(s-a)t}}{-(s-a)} \right]_0^{\infty} = \left[\frac{e^{-(\alpha-a)t} e^{-i\beta t}}{-(s-a)} \right]_0^{\infty}.$$

Der Grenzwert $t \rightarrow \infty$ existiert genau dann, wenn $\alpha > a$. In diesem Fall lautet die Laplace-Transformation von f gerade $(\mathcal{L}[f])(s) = 1/(s-a)$, wobei $\operatorname{Re}(s) > a$.

(2) Sei $f(t) = 1$ für $0 \leq t \leq 1$ und $f(t) = 0$ für $t > 1$. Es folgt

$$(\mathcal{L}[f])(s) = \int_0^1 e^{-st} dt = \left[\frac{e^{-st}}{-s} \right]_0^1 = \frac{1}{s}(1 - e^{-s}),$$

wobei $\operatorname{Re}(s) > 0$.

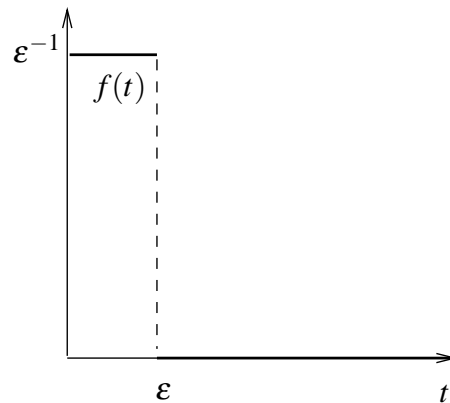


Abbildung 3.5: Illustration der Funktion f aus Beispiel 3.9 (3).

(3) Die Laplace-Transformation bildet den “Zeitbereich” in den “Spektralbereich” (oder “Frequenzbereich”) ab. Um diese Aussage zu verstehen, berechnen wir die Laplace-Transformation des folgenden Signals (siehe Abbildung 3.5):

$$f(t) = \begin{cases} \epsilon^{-1} & : 0 < t < \epsilon \\ 0 & : t \geq \epsilon. \end{cases}$$

Für sehr kleines $\epsilon > 0$ beschreibt f ein kurzes Signal mit Amplitude $1/\epsilon$. Die Laplace-Transformation lautet

$$(\mathcal{L}[f])(s) = \int_0^{\epsilon} e^{-st} f(t) dt = \left[\frac{e^{-st}}{-\epsilon s} \right]_0^{\epsilon} = \frac{1}{\epsilon s}(1 - e^{-\epsilon s}),$$

wobei $s \neq 0$. Diese Funktion können wir besser verstehen, wenn wir den Grenzwert $\epsilon \rightarrow 0$ mit der Regel von de l’Hospital durchführen:

$$\lim_{\epsilon \rightarrow 0} (\mathcal{L}[f])(s) = \lim_{\epsilon \rightarrow 0} \frac{(1 - e^{-\epsilon s})'}{(\epsilon s)'} = \lim_{\epsilon \rightarrow 0} \frac{s e^{-\epsilon s}}{s} = 1.$$

Die Laplace-Transformation eines ultrakurzen Signals mit “unendlich” großer Amplitude entspricht also der Frequenz eins. \square

Die Laplace-Transformation besitzt einige für die Anwendungen wichtige Eigenschaften.

Proposition 3.10. Die Laplace-Transformation $f \mapsto \mathcal{L}[f]$ ist linear und erfüllt in ihrem Definitionsbereich die folgenden Eigenschaften:

- (i) Ähnlichkeitssatz: $\mathcal{L}[f(at)] = (1/a)(\mathcal{L}[f])(s/a)$, $a > 0$;
- (ii) Verschiebungssatz: $\mathcal{L}[e^{-at}f(t)] = (\mathcal{L}[f])(s+a)$, $a > 0$;
- (iii) Erster Ableitungssatz: $(\mathcal{L}[f'])(s) = s(\mathcal{L}[f])(s) - f(0)$;
- (iv) Zweiter Ableitungssatz: $(\mathcal{L}[f''])(s) = s^2(\mathcal{L}[f])(s) - sf(0) - f'(0)$.

In der Elektrotechnik werden diese Beziehungen mit $F = \mathcal{L}[f]$ wie folgt dargestellt:

Ähnlichkeitssatz:	$f(at)$	$\circ \bullet$	$(1/a)F(s/a)$
Verschiebungssatz:	$e^{-at}f(t)$	$\circ \bullet$	$F(s+a)$
Erste Ableitung:	$f'(t)$	$\circ \bullet$	$sF(s) - f(0)$
Zweite Ableitung:	$f''(t)$	$\circ \bullet$	$s^2F(s) - sf(0) - f'(0)$

Beweis. Die Punkte (i), (ii) und (iv) sind Übungsaufgaben. Wir beweisen nur (iii) zur Illustration. Dazu integrieren wir partiell:

$$(\mathcal{L}[f'])(s) = \int_0^\infty e^{-st} f'(t) dt = [e^{-st} f(t)]_0^\infty + s \int_0^\infty e^{-st} f(t) dt.$$

Der erste Summand verschwindet im Grenzwert $t \rightarrow \infty$, da für $\text{Re}(s) > s_0$ und $|f(t)| \leq Me^{s_0 t}$ folgt $|e^{-st} f(t)| \leq M \exp(-(\text{Re}(s) - s_0)t) \rightarrow 0$ für $t \rightarrow \infty$. Damit ergibt sich wegen

$$(\mathcal{L}[f'])(s) = -f(0) + s \int_0^\infty e^{-st} f(t) dt = -f(0) + s(\mathcal{L}[f])(s)$$

die Behauptung. \square

Der große Vorteil der Laplace-Transformation ist, dass mit ihrer Hilfe die Lösung gewöhnlicher Differentialgleichungen auf die Manipulation algebraischer Gleichungen zurückgeführt werden kann. Wir betrachten als Anwendungsbeispiel einen RLC-Schaltkreis, bestehend aus einem Widerstand R , einer Induktivität L und einer Kapazität C in Reihe mit angelegter Spannung $U = U(t)$ (siehe Abbildung 3.6). Bevor wir eine Gleichung herleiten, die die zeitliche Entwicklung der Spannung bzw. des Stromes beschreibt, erklären wir die elektrotechnischen Begriffe.

Schule

Ein Widerstand ist ein Bauteil, an dem eine Spannung abfällt, z.B. ein Draht. Nach dem Ohmschen Gesetz ist der durch einen Widerstand fließende Strom I_R proportional zur Spannung U_R :

$$U_R = RI_R.$$

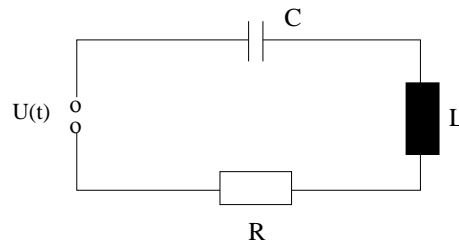


Abbildung 3.6: Ein RLC -Schaltkreis.

Die Proportionalitätskonstante nennt man auch *Widerstand*. Kondensatoren speichern elektrische Ladungen. Der elektrische Strom I_C durch einen Kondensator ist proportional zu Spannungsänderungen:

$$I_C = C \frac{dU_C}{dt},$$

wobei die Proportionalitätskonstante *Kapazität* genannt wird. Spulen speichern Energie in ihren magnetischen Feldern. Hier ist die abfallende Spannung U_L proportional zu den Stromänderungen

$$U_L = L \frac{dI_L}{dt}$$

mit der *Induktivität* als Proportionalitätskonstante.

Nach der Maschenregel ist die Summe der zu einer Masche gehörenden Spannungen null, d.h., die Summe der an den Bauteilen abfallenden Spannungen ist gleich der angelegten Spannung:

$$U = U_R + U_L + U_C.$$

Nach der Knotenregel fließt durch jedes der drei Bauteile der gleiche Strom:

$$I := I_R = I_L = I_C.$$

Leiten wir die Gleichung für die Summe der Spannungen einmal ab und setzen die obigen Gleichungen für U_R , U_L und U_C ein, so erhalten wir

$$U' = RI' + LI'' + \frac{I}{C}.$$

Setzen wir $2d = R/L$ und $\omega^2 = 1/(CL)$, so ergibt sich die Differentialgleichung

$$I'' + 2dI' + \omega^2 I = \frac{U'(t)}{L}. \quad (3.2)$$

Zur Zeit $t = 0$ seien der Strom und seine zeitliche Änderung an der Spule bekannt:

$$I(0) = I_0, \quad I'(0) = \frac{U_L(0)}{L}.$$

Prinzipiell könnten wir die Differentialgleichung (3.2) mit den Methoden aus Kapitel 4 lösen. In der Elektrotechnik kann die angelegte Spannung jedoch auch unstetig sein, und in diesem Fall sind diese Methoden nicht ohne Weiteres anwendbar. Wir verwenden zur Lösung von (3.2) die Laplace-Transformation.

Beispiel 3.11. (1) Wir betrachten einen Schaltkreis, der nur aus einer Spule mit der Induktivität $L = 1$ besteht. Die an der Spule abfallende Spannung betrage

$$U_L = I_L'.$$

Die Spulenspannung sei durch $U_L(t) = \cos t$ gegeben, und der Anfangsstrom sei $I_L(0) = 0$. Dann müssen wir die Differentialgleichung

$$I_L' = \cos t, \quad t > 0, \quad I_L(0) = 0,$$

lösen. Wir können hier die Lösung erraten; sie lautet $I_L(t) = \sin t$. Um die Methode der Laplace-Transformation zu üben, formulieren wir die Differentialgleichung im Spektralbereich. Sei $u = \mathcal{L}[I_L]$. Mit dem Ableitungssatz $(\mathcal{L}[I_L'])(s) = su(s) - I_L(0) = su(s)$ und $(\mathcal{L}[\cos])(s) = s/(s^2 + 1)$ (siehe Tabelle 1) folgt

$$su(s) = (\mathcal{L}[I_L'])(s) = (\mathcal{L}[\cos t])(s) = \frac{s}{s^2 + 1},$$

und nach u aufgelöst:

$$u(s) = \frac{1}{s^2 + 1}.$$

Nach Tabelle 1 lautet die Rücktransformation $I_L(t) = \sin t$ – wie erwartet.

$f(t)$	$(\mathcal{L}[f])(s)$
t^n	$\frac{n!}{s^{n+1}}, n \in \mathbb{N}$
e^{at}	$\frac{1}{s-a}$
$\sin(at)$	$\frac{a}{s^2 + a^2}$
$\cos(at)$	$\frac{s}{s^2 + a^2}$

Tabelle 1: Einige Funktionen und ihre Laplace-Transformation.

(2) Als zweites Beispiel untersuchen wir die Differentialgleichung (3.2) mit $d = 0$, $\omega = 1$ und der Ableitung der angelegten Spannung $U'(t) = L$ für $0 \leq t \leq 1$ und $U'(t) = 0$ für $t > 1$:

$$I'' + I = f(t) := \begin{cases} 1 & : 0 \leq t \leq 1 \\ 0 & : t > 1, \end{cases} \quad t > 0, \quad I(0) = I'(0) = 0. \quad (3.3)$$

Wir ersetzen I durch $u = \mathcal{L}[I]$, I' durch $su(s) - I(0) = su(s)$ und I'' durch $s^2u(s) - sI(0) - I'(0) = s^2u(s)$ (gemäß dem Ableitungssatz). Nach Beispiel 3.9 (2) folgt $F(s) = (\mathcal{L}[f])(s) = (1 - e^{-s})/s$. Dies ergibt die transformierte Differentialgleichung

$$(\mathcal{L}[I'' + I])(s) = (s^2 + 1)u(s) = F(s) = \frac{1}{s}(1 - e^{-s}),$$

und, nach $u(s)$ aufgelöst,

$$u(s) = \frac{1 - e^{-s}}{s(s^2 + 1)}.$$

Eine Partialbruchzerlegung liefert die Darstellung

$$u(s) = \frac{1}{s} - \frac{s}{s^2 + 1} - \frac{e^{-s}}{s} + \frac{se^{-s}}{s^2 + 1}.$$

Man kann zeigen, dass dies die Laplace-Transformation von

$$I(t) = \begin{cases} 1 - \cos t & : 0 \leq t < 1 \\ \cos(t-1) - \cos t & : t \geq 1 \end{cases}$$

ist. Dies ist eine einmal differenzierbare Lösung der Differentialgleichung (3.3). Sie ist *nicht* zweimal stetig differenzierbar, denn

$$I''(t) = \begin{cases} \cos t & : 0 \leq t < 1 \\ -\cos(t-1) + \cos t & : t \geq 1 \end{cases}$$

und $I''(1-0) = (\cos t)_{t=1} = \cos 1$ sowie $I''(1+0) = (-\cos(t-1) + \cos t)_{t=1} = \cos 1 - 1$. Es ist übrigens zu erwarten, dass I nicht zweimal stetig differenzierbar ist, denn $I'' = -I + f(t)$ ist unstetig. \square

4 Gewöhnliche Differentialgleichungen

Eine gewöhnliche Differentialgleichung ist eine Gleichung der Form

$$F(x^{(n)}, x^{(n-1)}, \dots, x', x, t) = 0, \quad t \in \mathbb{R},$$

wobei $F : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$ eine beliebige Funktion sei. Im diesem Kapitel entwickeln wir eine Lösungstheorie für spezielle Klassen von Differentialgleichungen, untersuchen die Stabilität der Lösungen und stellen einige Anwendungsbeispiele vor.

4.1 Differentialgleichungen erster Ordnung

In diesem Abschnitt geben wir einige Rezepte für die Lösung spezieller Klassen von Differentialgleichungen. Eine Existenztheorie und weiterführende Untersuchungen stellen wir in den folgenden Abschnitten vor.

Definition 4.1 (Differentialgleichung erster Ordnung). *Sei $F : \mathbb{R}^2 \times \mathbb{R} \rightarrow \mathbb{R}$ eine Funktion.*

(i) *Wir nennen die Gleichung*

$$F(x', x, t) = 0, \quad t \in \mathbb{R},$$

eine implizite Differentialgleichung erster Ordnung.

(ii) *Ist $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ eine Funktion, so nennen wir*

$$x' = f(x, t), \quad t \in \mathbb{R}, \tag{4.1}$$

eine explizite Differentialgleichung erster Ordnung.

(iii) *Die Angabe*

$$x(t_0) = x_0 \tag{4.2}$$

für ein Paar $(x_0, t_0) \in \mathbb{R}^2$ heißt ein Anfangswert. Typischerweise wird die entsprechende Differentialgleichung dann für $t > t_0$ betrachtet. In diesem Fall nennen wir (4.1)-(4.2) ein Anfangswertproblem.

(iv) *Hängt die Differentialgleichung (4.1) nicht explizit von t ab, so nennen wir die Differentialgleichung $x' = f(x)$ autonom.*

Gleichung (4.1) heißt explizit, weil die höchste Ableitung (in diesem Fall die erste Ableitung) explizit als Funktion von $x = x(t)$ und t dargestellt wird. Die Funktion f heißt die *rechte Seite* von (4.1). Wir werden uns hauptsächlich mit expliziten Differentialgleichungen befassen. Bevor wir uns der Lösung von (4.1) zuwenden, definieren wir, was wir unter einer Lösung verstehen.

Definition 4.2 (Lösung einer Differentialgleichung). *Sei $I \subset \mathbb{R}$ ein (nichtleeres) Intervall. Wir nennen $x : I \rightarrow \mathbb{R}$ eine Lösung der Differentialgleichung (4.1), wenn*

(i) *$x : I \rightarrow \mathbb{R}$ ist differenzierbar,*

(ii) $x'(t) = f(x(t), t)$ ist für alle $t \in I$ erfüllt.

Außerdem nennen wir x eine Lösung des Anfangswertproblems (4.1)-(4.2), wenn (i), (ii) und $x(t_0) = x_0$ gilt.

Beispiel 4.3. (1) Die Differentialgleichung

$$xx' + t = 0, \quad t \in \mathbb{R},$$

ist implizit. Wir können sie einfach lösen: Bemerken wir nämlich, dass wir sie in der Form

$$\frac{1}{2} \frac{d}{dt}(x^2 + t^2) = 0$$

schreiben können, so folgt $x^2 + t^2 = c$ mit einer beliebigen Konstante $c > 0$ und $x(t) = \pm\sqrt{c - t^2}$. Diese Lösungen existieren, solange $c - t^2$ nicht negativ wird. Das (maximale) Lösungsintervall lautet also $I = [-\sqrt{c}, \sqrt{c}]$.

(2) In der Einleitung haben wir motiviert, dass das unbegrenzte Wachstum einer Population durch das Anfangswertproblem Schule

$$x' = ax, \quad t > 0, \quad x(0) = x_0,$$

modelliert wird, wobei $x(t)$ die Populationsdichte, $t > 0$ die Zeit und $a \in \mathbb{R}$ die Wachstumsrate bedeute. Eine einfache Rechnung zeigt, dass die Funktion

$$x(t) = x_0 e^{at}, \quad t \geq 0,$$

eine Lösung des Anfangswertproblems ist. Eine Beschränkung des Wachstums ist im Populationsmodell von Verhulst (auch *logistische Gleichung* genannt) gegeben:

$$x' = (a - bx)x, \quad t > 0, \quad x(0) = x_0,$$

wobei $b > 0$, und die Wachstumsrate $a - bx$ nimmt mit wachsender Population ab. Um biologisch sinnvolle Lösungen zu erhalten, fordern wir $x \geq 0$. Die Lösung lautet

$$x(t) = \frac{a}{b(1 - c \exp(-at))}, \quad t \geq 0, \quad c := 1 - \frac{a}{bx_0}.$$

Weiter unten zeigen wir, wie man diese Lösung berechnen kann (siehe Beispiel 4.5).

(3) Der Preis $p = p(t)$ hängt in einer freien Marktwirtschaft von Angebot A und Nachfrage N ab. Andererseits sind auch Angebot und Nachfrage Funktionen des Preises: $A = A(p)$, $N = N(p)$. Übersteigt die Nachfrage das Angebot, $N > A$, so erwarten wir, dass der Preis steigt: $p' > 0$; ist die Nachfrage geringer als das Angebot, so wird der Preis sinken: $p' < 0$. Wir modellieren dieses Verhalten mit einer Funktion $g(x)$, die die Eigenschaften $g(x) < 0$ für $x < 0$, $g(x) > 0$ für $x > 0$ und $g(0) = 0$ besitzt: Schule

$$p' = g(N(p) - A(p)), \quad t > 0, \quad p(0) = p_0.$$

Ein einfaches Beispiel ist die lineare Funktion $g(x) = x$. Hängen Angebot und Nachfrage linear vom Preis ab, so können wir

$$N(p) = a + bp, \quad A(p) = c + dp, \quad a, b, c, d \in \mathbb{R},$$

setzen. Dies ergibt die Differentialgleichung

$$p' = (a - c + (b - d)p) = (b - d)(p - p_G), \quad p_G = -\frac{a - c}{b - d},$$

falls $b \neq d$. Da die rechte Seite linear in p ist, ist dies eine lineare Differentialgleichung. Falls $p(t) = p_G$, so ist $p'(t) = 0$, d.h., der Preis verändert sich nicht. Wir können daher p_G als den Gleichgewichtspreis identifizieren. \square

Eine Differentialgleichung besitzt in der Regel unendlich viele Lösungen. Beispielsweise ist $x(t) = ce^{at}$ für jedes $c \in \mathbb{R}$ eine Lösung von $x' = ax$. Wir nennen $x(t) = ce^{at}$ die *allgemeine Lösung* der Differentialgleichung. Es ist verständlich, dass wir unendlich viele Lösungen erhalten, da wir sie aus der Differentialgleichung durch Integration erhalten und hierbei eine beliebige Integrationskonstante anfällt. Diese Konstante können wir bestimmen, indem wir einen Anfangswert $x(t_0) = x_0$ fordern. Dann ist $x(t) = x_0 e^{a(t-t_0)}$, $t \geq 0$, die eindeutige Lösung des entsprechenden Anfangswertproblems. Dass diese Lösung wirklich eindeutig ist, werden wir später beweisen.

Allerdings ist es möglich, dass auch ein Anfangswertproblem keine oder unendlich viele Lösungen besitzt. Ein Beispiel ist das Problem

$$x' = \sqrt{|x|}, \quad t > 0, \quad x(0) = 0.$$

Klarerweise ist $x_1(t) = 0$ für alle $t \geq 0$ eine Lösung. Neben dieser trivialen Lösung ist aber auch die Funktion

$$x_2(t) = \frac{1}{4}t^2$$

eine Lösung des Anfangswertproblems.

Um sich einen ersten Überblick zu verschaffen, kann man die Lösungen der Differentialgleichung (4.1) graphisch angeben. Sei hierfür $x: I \rightarrow \mathbb{R}$ eine Lösung von (4.1) mit $x(t_0) = x_0$, die wir mit $\phi_t(x_0)$ bezeichnen. Wir nennen die Menge $\{\phi_t(x_0) : t \in I\}$ eine *Trajektorie* von (4.1) durch x_0 . Das *Richtungsfeld* von (4.1) ist die Gesamtheit aller $p = f(x, t)$. Einpunktige Trajektorien sind konstante Lösungen; diese nennen wir *Gleichgewichtspunkte* oder einfacher *Gleichgewichte*.

Beispiel 4.4. Betrachte die Differentialgleichung $x' = ax$, wobei $a > 0$, $t \in \mathbb{R}$, mit Anfangswert $x(t_0) = x_0$. Die Trajektorie durch x_0 ist gegeben durch

$$\phi_t(x_0) = x_0 e^{a(t-t_0)}, \quad t \in \mathbb{R}.$$

Wir können jedem Punkt in der Ebene einen Vektor mit Fußpunkt (x_0, t_0) und Steigung $x' = ax$ zuordnen (die Länge ist willkürlich; siehe Abbildung 4.7). Fixieren wir den Punkt (x_0, t_0) und gehen entlang dieser Vektoren, so können wir graphisch eine Lösung der Differentialgleichung zeichnen. \square

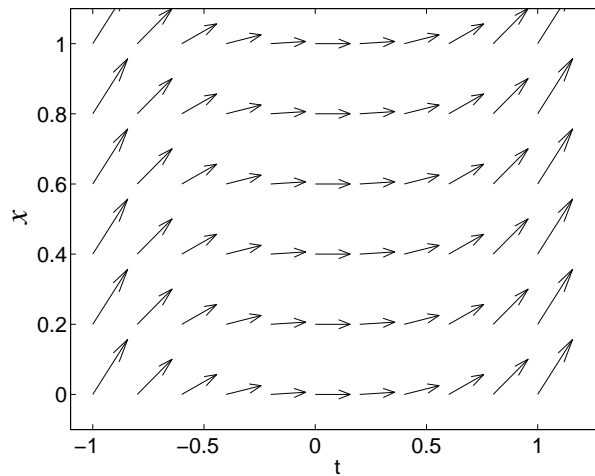


Abbildung 4.7: Richtungsfeld der Differentialgleichung $x' = x^2$.

Im Folgenden stellen wir einige direkte Lösungsmethoden für spezifische Klassen von Differentialgleichungen vor.

Trennung der Variablen. Eine Lösung der Differentialgleichung

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

können wir folgendermaßen bestimmen. Sei x eine Lösung mit $g(x(t)) \neq 0$ für alle t . Dann können wir die Differentialgleichung durch $g(x)$ dividieren und erhalten nach Integration

$$\int_{t_0}^t \frac{x'(s)}{g(x(s))} ds = \int_{t_0}^t f(s) ds.$$

Mit der Substitution $r = x(s)$, also *in formaler Schreibweise* $dr = x'(s)ds$, folgt

$$G(x) := \int_{x_0}^x \frac{dr}{g(r)} = \int_{t_0}^t f(s) ds =: F(t), \quad (4.3)$$

wobei wir $x(t_0) = x_0$ und $x(t) = x$ verwendet haben. Damit erhält man eine Lösung in impliziter Form, die man versuchen kann, nach x aufzulösen. Das Rezept lautet wie folgt:

- Bestimme die Gleichung (4.3).
- Löse (4.3) ggf. nach x auf.
- Überprüfe, ob $x = x(t)$ eine Lösung der Differentialgleichung ist.

Beispiel 4.5. Das Populationsmodell von Verhulst ist gemäß Beispiel 4.3 gegeben durch

$$x' = (a - bx)x, \quad t > 0, \quad x(0) = x_0,$$

wobei $a, b, x_0 > 0$. Falls $(a - bx)x \neq 0$, können wir

$$\int_{x_0}^x \frac{dr}{(a - br)r} = \int_0^t ds = t \quad (4.4)$$

schreiben. Das Integral auf der linken Seite berechnen wir mit Partialbruchzerlegung. Aus

$$\frac{1}{(a - br)r} = \frac{b}{a} \frac{1}{a - br} + \frac{1}{a} \frac{1}{r}$$

folgt

$$\begin{aligned} \int_{x_0}^x \frac{dr}{(a - br)r} &= \frac{b}{a} \int_{x_0}^x \frac{dr}{a - br} + \frac{1}{a} \int_{x_0}^x \frac{dr}{r} = \frac{b}{a} \left[-\frac{1}{b} \ln(a - br) \right]_{x_0}^x + \frac{1}{a} [\ln r]_{x_0}^x \\ &= \frac{1}{a} \ln \frac{a - bx_0}{a - bx} + \frac{1}{a} \ln \frac{x}{x_0} = \frac{1}{a} \ln \frac{(a - bx_0)x}{(a - bx)x_0}. \end{aligned}$$

Setzen wir dies in (4.4) ein, erhalten wir

$$\frac{(a - bx_0)x}{(a - bx)x_0} = e^{at}$$

und nach Auflösen nach x schließlich

$$x(t) = x = \frac{a}{b(1 - c \exp(-at))}, \quad c = 1 - \frac{a}{bx_0}.$$

Setzen wir diese Formel in die Differentialgleichung ein, sehen wir, dass es sich in der Tat um eine Lösung handelt.

Wir haben gefordert, dass $(a - bx)x \neq 0$ erfüllt ist, um die Division durchführen zu können. Dies könnte Lösungen der Form $a - bx = 0$ (bzw. $x = a/b$) oder $x = 0$ ausschließen. Die Lösung $x(t) = a/b$ ist in der obigen Formel mit $c = 0$ enthalten. Die Funktion $x(t) = 0$ ist keine Lösung, da wir $x_0 > 0$ gefordert haben. \square

Homogene Differentialgleichungen. Eine Differentialgleichung der Form

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

wird *homogen* genannt. Mit der Substitution $y = x/t$ folgt

$$y' = \frac{x't - x}{t^2} = \frac{f(y) - y}{t}.$$

Dies ist eine Gleichung mit getrennten Variablen, die wir mit der obigen Methode lösen können:

$$\int \frac{dy}{f(y) - y} = \int \frac{dt}{t} = \ln|t| + k, \quad k \in \mathbb{R}.$$

Die Stellen mit $f(y) = y$ müssen separat untersucht werden.

Beispiel 4.6. Betrachte die Gleichung

$$x' = \frac{x^2}{t^2} - \frac{x}{t}, \quad t \neq 0.$$

Mit $y = x/t$ und $f(y) = y^2 - y$ erhalten wir

$$y' = \frac{y^2 - 2y}{t}$$

und daraus die Lösungsformel

$$\ln|t| + k = \int \frac{dy}{y(y-2)} = \frac{1}{2} \int \left(\frac{1}{y-2} - \frac{1}{y} \right) dy = \frac{1}{2} \ln \frac{|y-2|}{|y|}$$

mit der Integrationskonstanten $k \in \mathbb{R}$, wobei wir $y \neq 0$ und $y \neq 2$ fordern. Daraus ergibt sich

$$\left| 1 - \frac{2}{y} \right| = \left| \frac{y-2}{y} \right| = e^{2k} e^{2 \ln|t|} = e^{2k} t^2,$$

also nach Auflösen der Betragstriche

$$1 - \frac{2}{y} = \pm e^{2k} t^2.$$

Wählen wir das Pluszeichen, erhalten wir für $k \in \mathbb{R}$ aus $+e^{2k}$ alle positiven Konstanten; wählen wir dagegen das Minuszeichen, so umfasst $-e^{2k}$ alle negativen Konstanten. Insgesamt ergibt dies alle Konstanten $c \neq 0$. Wir können also setzen

$$1 - \frac{2}{y} = ct^2 \quad \text{oder} \quad y = \frac{2}{1 - ct^2}.$$

Damit dies Sinn macht, muss $ct^2 \neq 1$ gelten. Nach $x = x(t)$ aufgelöst bedeutet dies

$$x(t) = \frac{2t}{1 - ct^2}, \quad c \neq 0.$$

Diese Lösungen existieren auf jedem Intervall mit $t \neq 0$ und, falls $c > 0$, $t \neq \pm 1/\sqrt{c}$. Für $c = 0$ erhalten wir $x(t) = 2t$; es zeigt sich, dass dies auch eine Lösung der Differentialgleichung ist.

Wir haben die Werte $y = 0$, $y = 2$ und $ct^2 = 1$ ausgeschlossen. Dies entspricht $x = 0$, $x = 2t$ und $ct^2 = 1$. Die Funktion $x(t) = 0$ ist eine Lösung, die bei der Division "verloren" gegangen ist. Der Parameter t erfüllt $t \neq 0$; also sind die Definitionsbereiche der Lösungen $x(t) = 0$ und $x(t) = 2t$ entweder $(-\infty, 0)$ oder $(0, \infty)$, denn wir haben in unserem Lösungsbegriff gefordert, dass der Definitionsbereich ein Intervall ist.

Zusammengefasst erhalten wir mit $c \geq 0$ die folgenden Lösungen:

$$\begin{aligned} x(t) &= 0 && \text{auf } (-\infty, 0) \text{ und } (0, \infty), \\ x(t) &= \frac{2t}{1 + ct^2} && \text{auf } \mathbb{R}, \\ x(t) &= \frac{2t}{1 - ct^2} && \text{auf } (-\infty, -1/\sqrt{c}), (-1/\sqrt{c}, 0), (0, 1/\sqrt{c}) \text{ und } (1/\sqrt{c}, \infty). \end{aligned}$$

Dies sind alle Lösungen. □

Lineare Differentialgleichungen. Es handelt sich um Differentialgleichungen der Form

$$x' = a(t)x + b(t), \quad t \in I, \quad x(t_0) = x_0, \quad (4.5)$$

wobei $a, b : I \rightarrow \mathbb{R}$ stetige Funktionen auf einem Intervall $I \subset \mathbb{R}$ seien. Wir nennen diese Gleichung *homogen*, falls $b = 0$, anderenfalls *inhomogen*. Die homogene Gleichung

$$x' = a(t)x$$

können wir mit Hilfe der Trennung der Variablen lösen. Aus

$$\ln|x| + k = \int \frac{dx}{x} = \int a(t)dt$$

folgt die allgemeine Lösung der homogenen Gleichung

$$x(t) = c \exp\left(\int a(t)dt\right), \quad t \in I, \quad c \in \mathbb{R}.$$

Ist ein Anfangswert $x(t_0) = x_0$ vorgegeben, erhalten wir folgendes Resultat.

Proposition 4.7 (Lösung der homogenen linearen Differentialgleichung). *Sei $a : I \rightarrow \mathbb{R}$ stetig und $t_0 \in I$. Setze $A(t) = \int_{t_0}^t a(s)ds$. Dann ist*

$$x(t) = x_0 e^{A(t)}, \quad t \in I,$$

die eindeutige Lösung von

$$x' = a(t)x, \quad t \in I, \quad x(t_0) = x_0. \quad (4.6)$$

Beweis. Der Nachweis, dass die in der Proposition definierte Funktion wirklich eine Lösung ist, folgt aus einer leichten Rechnung. Es bleibt zu zeigen, dass sie die einzige Lösung ist. Sei $y(t)$ eine beliebige Lösung von (4.6). Wir rechnen für $t \in I$

$$\frac{d}{dt}(y(t)e^{-A(t)}) = y'(t)e^{-A(t)} - y(t)A'(t)e^{-A(t)} = e^{-A(t)}(y'(t) - a(t)y(t)) = 0.$$

Also existiert eine Konstante $c \in \mathbb{R}$ mit $y(t)e^{-A(t)} = c$. Wegen $x_0 = y(t_0) = ce^{A(0)} = c$ folgt die Behauptung. \square

Wir kehren zur inhomogenen Gleichung (4.5) zurück. Wir behaupten, dass die allgemeine Lösung der inhomogenen Gleichung gleich der Summe aus der allgemeinen Lösung der homogenen Gleichung $x_h(t) = c \exp(A(t))$ und einer partikulären Lösung der inhomogenen Gleichung $x_p(t)$ ist.

Lemma 4.8. *Seien $a, b : I \rightarrow \mathbb{R}$ stetig, x_h die allgemeine Lösung der homogenen Gleichung (4.6) und x_p eine partikuläre Lösung der inhomogenen Gleichung (4.5). Dann ist $x = x_h + x_p$ die allgemeine Lösung der inhomogenen Gleichung (4.5).*

Beweis. Die Summe $x = x_h + x_p$ ist tatsächlich eine Lösung von (4.5), denn

$$x' - a(t)x = x'_h + x'_p - a(t)(x_h + x_p) = \underbrace{(x'_h - a(t)x_h)}_{=0} + \underbrace{(x'_p - a(t)x_p)}_{=b(t)} = b(t).$$

Es bleibt nachzuweisen, dass durch $x = x_h + x_p$ alle Lösungen erfasst sind. Sei dazu $y(t)$ eine beliebige Lösung der inhomogenen Gleichung (4.5) und betrachte die Differenz $y - x_p$. Wegen

$$(y - x_p)' - a(t)(y - x_p) = (y' - a(t)y) - (x'_p - a(t)x_p) = b(t) - b(t) = 0$$

ist $y - x_p$ eine Lösung der homogenen Gleichung und kann als $x_h = y - x_p$ geschrieben werden (denn x_h ist die allgemeine Lösung der homogenen Gleichung). Dies ergibt aber mit $y = x_h + x_p$ die gewünschte Darstellung. \square

Wir können eine partikuläre Lösung der inhomogenen Gleichung mit der *Methode der Variation der Konstanten* berechnen. Dazu machen wir den Lösungsansatz

$$x_p(t) = c(t) \exp(A(t))$$

mit einer zu bestimmenden Funktion $c(t)$. Setzen wir ihn in die inhomogene Gleichung ein, so folgt

$$b(t) = x'_p - a(t)x_p = e^{A(t)}(c' + cA' - ac) = c'e^{A(t)},$$

also

$$c = \int b(s)e^{-A(s)} ds.$$

Damit ist $x_p(t) = \int b(s)e^{A(t)-A(s)} ds$. Wir erhalten das folgende Resultat.

Proposition 4.9 (Lösung der inhomogenen linearen Gleichung). *Seien $a, b : I \rightarrow \mathbb{R}$ stetig und $t_0 \in I$. Setze $A(t) = \int_{t_0}^t a(s) ds$. Dann ist*

$$x(t) = x_0 e^{A(t)} + \int_{t_0}^t e^{A(t)-A(s)} b(s) ds, \quad t \in I,$$

die eindeutige Lösung der inhomogenen Gleichung (4.5) mit Anfangswert $x(t_0) = x_0$.

Der Beweis der Eindeutigkeit ist eine Übungsaufgabe.

Beispiel 4.10. Wir betrachten das Preismodell aus Beispiel 4.3:

$$p' = (b - d)(p - p_G), \quad t > 0, \quad p(0) = p_0.$$

Diese Gleichung ist linear mit $a(t) = b - d$ und $b(t) = -(b - d)p_G$. Wir berechnen die Lösung mit Hilfe von Proposition 4.9. Aus

$$A(t) = \int_0^t a(s) ds = (b - d)t$$

folgt

$$\begin{aligned} p(t) &= p_0 e^{(b-d)t} - \int_0^t e^{(b-d)(t-s)} (b-d) p_G ds \\ &= p_0 e^{(b-d)t} + p_G [e^{(b-d)(t-s)}]_0^t = (p_0 - p_G) e^{(b-d)t} + p_G. \end{aligned}$$

Wir erinnern, dass $N(p) = a + bp$ (Nachfrage) und $A(p) = c + dp$ (Angebot). Wir erwarten, dass die Nachfrage bei höheren Preisen fällt, $N'(p) < 0$. Andererseits wird das Angebot bei höheren Preisen ausgebaut werden, so dass $A'(p) > 0$. Diese Bedingungen implizieren $b = N'(p) < 0 < A'(p) = d$. Folglich konvergiert der Preis $p(t)$ für große Zeiten $t \rightarrow \infty$ gegen den Gleichgewichtspreis p_G , d.h., in diesem einfachen Modell stabilisiert das Wechselspiel zwischen Angebot und Nachfrage den Preis. \square

4.2 Existenz und Eindeutigkeit von Lösungen

In diesem Abschnitt entwickeln wir eine Existenztheorie für Systeme von Differentialgleichungen erster Ordnung. Seien $D \subset \mathbb{R}^{n+1}$ ein Gebiet und $f : D \rightarrow \mathbb{R}^n$ eine Funktion. Wir betrachten Differentialgleichungen der Form

$$x' = f(x, t), \quad t \in \mathbb{R}, \quad x(t_0) = x_0. \quad (4.7)$$

Im Gegensatz zum vorigen Abschnitt sind $f = (f_1, \dots, f_n)$ und $x = (x_1, \dots, x_n)$ vektorwertige Funktionen und $x_0 = (x_{0,1}, \dots, x_{0,n})$ ein Vektor. In Komponenten geschrieben lautet das System:

$$\begin{aligned} x'_1 &= f_1(x_1, \dots, x_n, t), \\ &\vdots \\ x'_n &= f_n(x_1, \dots, x_n, t). \end{aligned}$$

Wir nennen wie im vorigen Abschnitt $x : I \rightarrow \mathbb{R}$ eine Lösung dieses Systems, wenn $I \subset \mathbb{R}$ ein Intervall ist, x differenzierbar ist und die Differentialgleichung samt Anfangswert löst. Wir setzen stets voraus, dass $R(x) \times I \subset D$ gilt, wobei $R(x)$ den Bildraum von x bezeichne.

Beispiel 4.11. Das Lotka-Volterra-Modell aus der Einleitung zur Beschreibung der zeitlichen Entwicklung einer Beutepopulation $x_1(t)$ und einer Räuberpopulation $x_2(t)$ lautet

$$x'_1 = (a - bx_2)x_1, \quad x'_2 = (-c + dx_1)x_2,$$

oder in kompakter Formulierung

$$x' = \frac{d}{dt} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} (a - bx_2)x_1 \\ (-c + dx_1)x_2 \end{pmatrix} =: f(x),$$

wobei $x = (x_1, x_2)^\top$. \square

Die Existenztheorie für das Anfangswertproblem (4.7) basiert auf einer Formulierung der Differentialgleichung als Integralgleichung, auf die der Fixpunktsatz von Banach angewendet werden kann.

Lemma 4.12. Sei $f : D \rightarrow \mathbb{R}^n$ stetig. Eine differenzierbare Funktion $x : I \rightarrow \mathbb{R}^n$ ist Lösung des Anfangswertproblems

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

genau dann, wenn x eine stetige Lösung der folgenden Integralgleichung ist:

$$x(t) = x_0 + \int_{t_0}^t f(x(s), s) ds, \quad t \in I. \quad (4.8)$$

Beweis. 1. Sei x eine Lösung des Anfangswertproblems. Die k -te Komponente lautet

$$x'_k = f_k(x, t).$$

Die rechte Seite ist nach Voraussetzung stetig; insbesondere ist x_k stetig differenzierbar, und nach dem Hauptsatz der Differential- und Integralrechnung erhalten wir nach Integration

$$x_k(t) - x_k(t_0) = \int_{t_0}^t f_k(x(s), s) ds, \quad t \in I.$$

Also löst x_k für alle $k = 1, \dots, n$ die Integralgleichung.

2. Sei umgekehrt x_k eine Lösung der Integralgleichung. Klarerweise gilt $x_k(t_0) = x_{0,k}$. Die Funktion x_k ist als Stammfunktion einer stetigen Funktion differenzierbar, so dass

$$x'_k(t) = \frac{d}{dt} \left(x_{0,k} + \int_{t_0}^t f_k(x(s), s) ds \right) = f_k(x(t), t).$$

Also löst x_k das Anfangswertproblem. □

Mit diesen Vorbereitungen können wir den Existenzsatz formulieren und beweisen.

Satz 4.13 (Picard-Lindelöf; globale Version). Sei $D = \mathbb{R}^n \times [t_0, t_0 + a]$ mit $a > 0$. Die Funktion $f : D \rightarrow \mathbb{R}^n$ sei stetig in D und (global) lipschitzstetig in x gleichmäßig in t , d.h.,

$$\exists L > 0 : \forall (x, t), (y, t) \in D : |f(x, t) - f(y, t)| \leq L|x - y|.$$

Dann besitzt das Anfangswertproblem (4.7) genau eine Lösung $x : [t_0, t_0 + a] \rightarrow \mathbb{R}^n$.

Beweis. Die Idee des Beweises ist die Verwendung des Fixpunktsatzes von Banach. Wir zeigen, dass die Integralgleichung (4.8) eine eindeutige Lösung besitzt. Nach Lemma 4.12 folgt dann die Behauptung des Satzes. Sei $I = [t_0, t_0 + a]$. Für den Fixpunktsatz wählen wir den Banachraum $X = C^0(I; \mathbb{R}^n)$ mit der Norm

$$\|x\| = \sup_{t \in I} e^{-2Lt} |x(t)|, \quad x \in X.$$

Den Nachweis, dass dies wirklich eine Norm ist, überlassen wir den LeserInnen. Die Fixpunkt-abbildung sei definiert durch

$$S: X \rightarrow X, \quad (Sx)(t) = x_0 + \int_{t_0}^t f(x(s), s) ds, \quad t \in I.$$

Wir müssen zeigen, dass S eine Kontraktion ist. Seien $x, y \in X$ und $t \in I$. Dann ist $(x(t), t), (y(t), t) \in D$, und mit der Lipschitzstetigkeit von f folgt

$$\begin{aligned} |(Sx)(t) - (Sy)(t)| &\leq \int_{t_0}^t |f(x(s), s) - f(y(s), s)| ds \leq L \int_{t_0}^t e^{-2Ls} |x(s) - y(s)| e^{2Ls} ds \\ &= L \|x - y\| \int_{t_0}^t e^{2Ls} ds \leq L \|x - y\| \frac{e^{2Lt}}{2L} = \frac{1}{2} e^{2Lt} \|x - y\|. \end{aligned}$$

Gemäß der Definition der Norm erhalten wir damit

$$\|Sx - Sy\| = \sup_{t \in I} e^{-2Lt} |(Sx)(t) - (Sy)(t)| \leq \frac{1}{2} \|x - y\|.$$

Nach dem Fixpunktsatz von Banach existiert genau ein Fixpunkt $x \in X$ mit $Sx = x$, also genau eine Lösung der Integralgleichung (4.8). \square

Beispiel 4.14. (1) Betrachte das Anfangswertproblem

$$x' = \frac{t}{t+1} \sin x, \quad t > 0, \quad x(0) = x_0.$$

Wir überprüfen, ob die Funktion $f(x, t) = (t/(t+1)) \sin x$ lipschitzstetig in x gleichmäßig in t ist. Seien $x, y \in \mathbb{R}$ und $t \geq 0$. Dann existiert nach dem Mittelwertsatz ein z zwischen x und y mit

$$|f(x, t) - f(y, t)| = \underbrace{\frac{t}{t+1}}_{\leq 1} \underbrace{|\sin x - \sin y|}_{=\cos z(x-y)} \leq |x - y|,$$

denn $|\cos z| \leq 1$. Nach dem Satz von Picard-Lindelöf existiert eine eindeutige Lösung $x: [0, \infty) \rightarrow \mathbb{R}$ des obigen Anfangswertproblems. Die Lösung existiert also auf dem gesamten Intervall $[0, \infty)$.

(2) Die Schwingung eines Pendels kann durch das Anfangswertproblem

$$\frac{d}{dt} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ -\sin x_1 \end{pmatrix}, \quad t > 0, \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} (0) = \begin{pmatrix} x_{01} \\ x_{02} \end{pmatrix}$$

modelliert werden. Die Variable x_1 stellt die Auslenkung des Pendels dar, die Variable x_2 dessen Geschwindigkeit. Wir überprüfen, ob die Funktion

$$f: \mathbb{R}^2 \times [0, \infty) \rightarrow \mathbb{R}^2, \quad f(x_1, x_2, t) = \begin{pmatrix} x_2 \\ -\sin x_1 \end{pmatrix},$$

lipschitzstetig in $x = (x_1, x_2)$ ist. Ähnlich wie im vorigen Beispiel folgt

$$\begin{aligned} |f(x_1, x_2, t) - f(y_1, y_2, t)| &= \sqrt{(x_2 - y_2)^2 + (\sin x_1 - \sin y_1)^2} \\ &\leq \sqrt{(x_2 - y_2)^2 + (x_1 - y_1)^2} = |x - y|, \end{aligned}$$

also ist f lipschitzstetig in x gleichmäßig in t (denn f hängt gar nicht von t ab). Also besitzt das obige Anfangswertproblem eine eindeutige Lösung auf $[0, \infty)$.

(3) Das Anfangswertproblem

$$x' = x^2, \quad t > 0, \quad x(0) = 1,$$

besitzt dagegen nur die lokale Lösung $x(t) = 1/(1-t)$ für $t \in [0, 1)$. Was ist schiefgegangen? Die Funktion $f(x, t) = x^2$ ist nicht global lipschitzstetig in x ; wegen $|f(x, t) - f(y, t)| = |x+y||x-y|$ kann die Lipschitzkonstante nicht unabhängig von x und y gewählt werden. \square

Das letzte Beispiel zeigt, dass ein Anfangswertproblem eine *lokale* Lösung besitzen kann, selbst wenn der Satz von Picard-Lindelöf nicht anwendbar ist. Das Problem im Beispiel war, dass die rechte Seite nicht *global* lipschitzstetig ist. Wir zeigen nun, dass *lokale* Lipschitzstetigkeit ausreicht, um zumindest eine lokale Lösung zu erhalten.

Satz 4.15 (Picard-Lindelöf; lokale Version). *Seien $B = B_r(x_0) \subset \mathbb{R}^n$ die Kugel um $x_0 \in \mathbb{R}^n$ mit Radius $r > 0$, $a > 0$ und $D = B \times [t_0, t_0 + a]$. Die Funktion $f : D \rightarrow \mathbb{R}^n$ sei stetig in D und lipschitzstetig in x gleichmäßig in t , d.h.,*

$$\exists L > 0 : \forall (x, t), (y, t) \in D : |f(x, t) - f(y, t)| \leq L|x - y|.$$

Dann besitzt das Anfangswertproblem (4.7) genau eine Lösung $x : [t_0, t_0 + b] \rightarrow \mathbb{R}^n$, wobei $b = \min\{a, r/M\}$ und $M = \max_{(x,t) \in D} |f(x, t)|$.

Im Unterschied zur globalen Version des Satzes von Picard-Lindelöf fordern wir die Lipschitzstetigkeit im obigen Satz nur auf der Kugel B , nicht mehr auf \mathbb{R}^n .

Beweis. Wir wenden wieder den Fixpunktsatz von Banach an und wählen den Banachraum $X = C^0(I; \mathbb{R}^n)$ (wobei $I = [t_0, t_0 + b]$) mit der Norm $\|x\| = \sup_{t \in I} e^{-2Lt}|x(t)|$. Die Fixpunktabbildung sei wieder

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

Im Unterschied zum vorherigen Beweis sei S allerdings nicht auf dem gesamten Banachraum X definiert, sondern nur auf der Teilmenge

$$Y = \{x \in X : \forall t \in I : |x(t) - x_0| \leq M|t - t_0|\}.$$

Da diese Menge abgeschlossen ist (Beweis?), ist $(Y, \|\cdot\|)$ ein vollständiger metrischer Raum (kein normierter Raum, weil wir keine Vektorraumstruktur haben). Wir müssen zeigen, dass S

wohldefiniert und eine Selbstabbildung ist, also $(x(t), t)$ für $x \in Y$ im Definitionsbereich von f liegt und $S : Y \rightarrow Y$ gilt. Die erste Behauptung folgt gemäß der Definition von b aus

$$|x(t) - x_0| \leq M|t - t_0| \leq Mb \leq M \frac{r}{M} = r$$

und damit $(x(t), t) \in B \times I \subset D$. Die zweite Behauptung ist eine Konsequenz der Definition von M und

$$|(Sx)(t) - x_0| \leq \int_{t_0}^t |f(x(s), s)| ds \leq M \int_{t_0}^t ds = M(t - t_0),$$

also $Sx \in Y$. Die Lipschitzstetigkeit von f in x folgt wie im Beweis von Satz 4.13 mit der Lipschitzkonstanten $L = 1/2$. Der Fixpunktsatz von Banach liefert nun genau einen Fixpunkt, also genau eine Lösung der Integralgleichung (4.8). \square

Beispiel 4.16. (1) Wir wissen bereits aus Beispiel 4.14, dass das Anfangswertproblem

$$x' = x^2, \quad t > 0, \quad x(0) = 1,$$

für $t \in [0, 1)$ lösbar ist. Wir wollen die Voraussetzungen des Satzes 4.15 von Picard-Lindelöf überprüfen. Die Funktion $f(x, t) = x^2$ ist lipschitzstetig in x : Mit $x, y \in B_r(1)$ folgt

$$|f(x, t) - f(y, t)| = |x + y||x - y| \leq 2(r + 1)|x - y|, \quad \text{also } L = 2(r + 1).$$

Wir berechnen ferner $M = \max_{x \in B_r(1)} |f(x)| = (r + 1)^2$. Nach Satz 4.15 existiert eine eindeutige Lösung auf dem Intervall $[0, b]$ mit $b = r/(r + 1)^2 < 1$.

(2) Das Lotka-Volterra-Modell $x' = f(x)$, $x = (x_1, x_2) \in [0, \infty)^2$, mit der rechten Seite

$$f(x) = \begin{pmatrix} (a - bx_2)x_1 \\ (-c + dx_1)x_2 \end{pmatrix},$$

wobei $a, b, c, d > 0$, besitzt ebenfalls eine lokale Lösung, denn

$$|f(x) - f(y)|^2 = ((a - bx_2)x_1 - (a - by_2)y_1)^2 + ((-c + dx_1)x_2 - (-c + dy_1)y_2)^2,$$

wobei $x = (x_1, x_2), y = (y_1, y_2) \in B_r(x_0)$, kann nach oben durch

$$L((x_1 - y_1)^2 + (x_2 - y_2)^2) = L|x - y|^2$$

abgeschätzt werden, wobei $L > 0$ von r und x_0 abhängt. \square

Für späteren Gebrauch beweisen wir das folgende Hilfsresultat.

Lemma 4.17 (Gronwall). *Seien $g, h \in C^0([a, b])$ nichtnegative Funktionen und $g_0 \geq 0$. Es gelte für alle $t \in [a, b]$*

$$g(t) \leq g_0 + \int_a^t g(s)h(s)ds.$$

Dann ist

$$g(t) \leq g_0 \exp\left(\int_a^t h(s)ds\right), \quad t \in [a, b].$$

Beweis. 1. Sei zunächst $g_0 > 0$. Dann besitzt

$$F(t) = g_0 + \int_a^t g(s)h(s)ds > 0$$

die Ableitung $F'(t) = g(t)h(t) \leq F(t)h(t)$ (folgt aus der Voraussetzung des Lemmas). Dies impliziert

$$(\ln F(t))' = \frac{F'(t)}{F(t)} \leq h(t), \quad t \in [a, b],$$

und Integration dieser Ungleichung über (a, t) liefert

$$\ln F(t) - \ln F(a) \leq \int_a^t h(s)ds.$$

(Hier benötigen wir $F(a) = g_0 > 0$.) Wir lösen dies nach $F(t)$ auf und verwenden $F(a) = g_0$:

$$g(t) \leq F(t) \leq g_0 \exp\left(\int_a^t h(s)ds\right).$$

2. Falls $g_0 = 0$, so gilt für alle $\varepsilon > 0$

$$g(t) \leq \int_a^t g(s)h(s)ds \leq \varepsilon + \int_a^t g(s)h(s)ds.$$

Nach Schritt 1 des Beweises ergibt sich

$$g(t) \leq \varepsilon \exp\left(\int_a^t h(s)ds\right)$$

für alle $\varepsilon > 0$, und im Grenzwert $\varepsilon \rightarrow 0$ erhalten wir $g(t) \leq 0$, also $g(t) = 0$ für $t \in [a, b]$. \square

Variieren die Anfangswerte nur wenig, so erwarten wir, dass sich auch die entsprechenden Lösungen zumindest für kleine Zeiten nur wenig voneinander unterscheiden.

Satz 4.18 (Stetige Abhängigkeit von den Anfangswerten). *Seien $D \subset \mathbb{R}^{n+1}$ und $f : D \rightarrow \mathbb{R}^n$ stetig in D und lipschitzstetig in x gleichmäßig in t mit der Lipschitzkonstante $L > 0$. Seien weiter $x, y : [t_0, t_0 + a] \rightarrow \mathbb{R}^n$ ($a > 0$) Lösungen von*

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

mit Anfangswerten $x(t_0) = x_0$ bzw. $y(t_0) = y_0$, wobei $(t_0, x_0), (t_0, y_0) \in D$. Dann gilt

$$|x(t) - y(t)| \leq |x_0 - y_0| e^{L(t-t_0)}, \quad t \in [t_0, t_0 + a].$$

Der Beweis verwendet das Lemma von Gronwall und ist eine Übungsaufgabe.

Wir haben bereits in Abschnitt 4.1 gesehen, dass das Anfangswertproblem $x' = \sqrt{|x|}$, $t > 0$, $x(0) = 0$, mehrere Lösungen besitzt. Der Verlust der Eindeutigkeit ist im Rahmen der Sätze von Picard-Lindelöf verständlich, da die Funktion $f(x) = \sqrt{|x|}$ nicht lipschitzstetig ist (sondern nur hölderstetig mit Exponenten $1/2$). Um die Existenz lokaler Lösungen zu erhalten, ist offenbar die Lipschitzstetigkeit eine zu starke Voraussetzung. Unter welchen Bedingungen können wir die Existenz von Lösungen sicherstellen? Es zeigt sich, dass hierfür die Stetigkeit der rechten Seite genügt.

Satz 4.19 (Peano). Seien $B \subset \mathbb{R}^n$ offen, $f : B \times [t_0, t_0 + a] \rightarrow \mathbb{R}^n$ eine stetige Funktion und $x_0 \in B$. Dann existiert ein $0 < b \leq a$ und eine Lösung $x : [t_0, t_0 + b] \rightarrow \mathbb{R}^n$ des Anfangswertproblems (4.7).

Der Satz von Peano garantiert die Existenz von *lokalen* Lösungen, aber nicht deren Eindeutigkeit. Das Beispiel $x' = \sqrt{|x|}$ zeigt, dass Eindeutigkeit von Lösungen auch nicht zu erwarten ist. Der Beweis des Satzes ist etwas aufwendiger und basiert auf dem Satz von Arzelà-Ascoli; wir verwiesen auf die Literatur.

Wir können die Frage stellen, wie groß das *maximale* Existenzintervall I ist, auf dem eine Lösung definiert ist. Dazu zunächst die folgenden Beispiele.

Beispiel 4.20. (1) Die Lösung $x(t) = e^{at}$ des Anfangswertproblems $x' = ax$, $x(0) = 1$ auf $(x, t) \in \mathbb{R}^2$ ist für alle $t \in \mathbb{R}$ erklärt. Das maximale Existenzintervall lautet $I = \mathbb{R}$.

(2) Die Differentialgleichung $x' = x^2$, $x(0) = 1$ auf $(x, t) \in D = \mathbb{R}^2$ besitzt die Lösung $x(t) = 1/(1-t)$, die für alle $t \in I = (-\infty, 1)$ erklärt ist. Für $t \rightarrow 1$ “explodiert” die Lösung: $\lim_{t \rightarrow 1} |x(t)| = \infty$.

(3) Die Differentialgleichung $x' = 1/x$, $x(0) = 1$, auf $D = \{(x, t) \in \mathbb{R}^2 : x > 0\}$ besitzt die Lösung $x(t) = \sqrt{2t+1}$ mit dem maximalen Existenzintervall $I = (-1/2, \infty)$. Für $t \rightarrow -1/2$ folgt

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

Also konvergiert die Lösung bis zum Rand ∂D des Definitionsbereichs D . □

Die Beispiele zeigen, dass es drei Situationen für das Verhalten der Lösung im maximalen Existenzintervall zu geben scheint:

- Die Lösung $x(t)$ existiert für alle $t \geq t_0$.
- Die Lösung “explodiert”, d.h., es existiert $t^* > t_0$ mit $\lim_{t \rightarrow t^*} |x(t)| = \infty$.
- Die Lösung kann bis zum Rand ∂D des Definitionsbereichs fortgesetzt werden, d.h., es existiert $t^* > t_0$ mit $\lim_{t \rightarrow t^*} \text{dist}((x(t), t), \partial D) = 0$.

Hierbei ist $\text{dist}(x, M) = \inf\{|x - y| : y \in M\}$ der (minimale) Abstand eines Punktes x zur Menge M . Der folgende Satz zeigt, dass dadurch alle möglichen Fälle beschrieben sind.

Satz 4.21. Seien $D \subset \mathbb{R}^{n+1}$ offen, $(x_0, t_0) \in D$, $f : D \rightarrow \mathbb{R}^n$ stetig und lokal lipschitzstetig in x gleichmäßig in t , und $x : I \rightarrow \mathbb{R}^n$ eine Lösung des Anfangswertproblems

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

mit maximalem Existenzintervall $I \subset \mathbb{R}$. Dann enthält keine kompakte Teilmenge von D den Graphen $\{(x(t), t) : t \in I\}$ von x (siehe Abbildung 4.8).

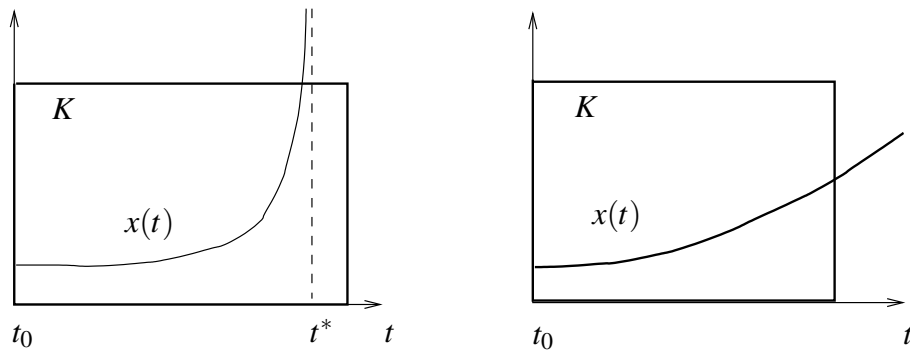


Abbildung 4.8: Die Graphen von $x(t)$ verlassen die kompakte Menge K .

Beweis. Das maximale Existenzintervall sei $I = (a, b)$. Wir machen eine Widerspruchsannahme: Sei $K \subset D$ eine kompakte Menge mit $\{(x(t), t) : t \in I\} \subset K$. Da K als kompakte Menge beschränkt ist, folgt $-\infty < a < b < \infty$. Die Funktion f ist stetig und damit auf jedem Kompaktum beschränkt; wir setzen $M = \max_{(x,t) \in K} |f(x,t)|$. Dann gilt für $t, \tau \in [t_0, b)$:

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

Sei nun (t_n) eine Folge mit $t_n \rightarrow b$ ($n \rightarrow \infty$). Die obige Abschätzung zeigt, dass $(x(t_n))$ eine Cauchy-Folge und damit konvergent ist. Wir setzen $x_b = \lim_{n \rightarrow \infty} x(t_n)$. Dieser Grenzwert ist eindeutig: Ist nämlich (s_n) eine weitere Folge mit $s_n \rightarrow b$, so folgt aus der obigen Abschätzung

$$|x(t_n) - x(s_n)| \leq M|t_n - s_n| \rightarrow 0,$$

also $\lim_{n \rightarrow \infty} x(s_n) = x_b$. Dann können wir x durch die Funktion

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

eindeutig fortsetzen, und es gilt

$$y(t) = x_b + \int_b^t f(y(s), x) ds, \quad t \in [t_0, b].$$

Die kompakte Menge K ist abgeschlossen, so dass $(x_b, b) \in K \subset D$ folgt. Nach dem Satz von Picard-Lindelöf existiert dann eine (lokale) Lösung von

$$x' = f(x, t), \quad x(b) = x_b$$

auf einem Intervall $[b, b + c]$ für ein $c > 0$. Dies widerspricht der Definition von b (maximales Existenzintervall). \square

In vielen Anwendungen treten Differentialgleichungen mit Ableitungen höherer Ordnung auf.

Beispiel 4.22. Nach dem Newtonschen Gesetz ist die Beschleunigung $x'' = d^2x/dt^2$ eines Gegenstands proportional zur auf den Gegenstand wirkenden Kraft F :

$$mx'' = F,$$

wobei $m > 0$ die Masse des Gegenstands bezeichne. Wir nehmen an, dass eine Gravitationskraft auf den Gegenstand wirkt:

$$F(t) = -\frac{GmM}{(x(t) + R)^2},$$

wobei M die Erdmasse, G die Gravitationskonstante und R der Erdradius seien. Hierbei haben wir die Annahme gemacht, dass die Erde eine perfekte Kugel mit homogener Massenverteilung ist. Setzen wir diese Beziehung in das Newtonsche Gesetz ein, so erhalten wir mit der Gravitationsfeldstärke der Erde $g := GM/R^2$ die Differentialgleichung

$$x'' = -\frac{gR^2}{(x + R)^2}, \quad x \geq 0.$$

Integration dieser Gleichung liefert *zwei* Integrationskonstanten. Wir benötigen daher zwei Anfangswerte. Typischerweise werden der Anfangsort x und die Anfangsgeschwindigkeit x' vorge-schrieben:

$$x(t_0) = x_0, \quad x'(t_0) = x_1,$$

was aus physikalischer Sicht auch sinnvoll ist, um die Dynamik von $x(t)$ vollständig zu beschreiben.

Die obige Differentialgleichung *zweiter* Ordnung können wir in ein System von Differentialgleichungen *erster* Ordnung umwandeln. Seien hierzu $y_1 = x$, $y_2 = x'$. Dann folgt $y_1' = x' = y_2$ und $y_2' = x'' = -gR^2(y_1 + R)^{-2}$ und

$$\frac{d}{dt} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = f(y_1, y_2) := \begin{pmatrix} y_2 \\ -gR^2(y_1 + R)^{-2} \end{pmatrix},$$

also eine Differentialgleichung vom Typ $y' = f(y)$ mit $y \in \mathbb{R}^2$. □

Die Vorgehensweise des obigen Beispiels lässt sich verallgemeinern. Sei die Differentialgleichung n -ter Ordnung

$$x^{(n)} = F(x^{(n-1)}, \dots, x', x, t), \quad t \in \mathbb{R}, \tag{4.9}$$

mit Anfangswerten

$$x(t_0) = x_0, \quad x'(t_0) = x_1, \dots, \quad x^{(n-1)}(t_0) = x_{n-1} \tag{4.10}$$

gegeben. Definieren wir $y = (y_n, \dots, y_1)$ mit $y_1 = x$, $y_2 = x'$, \dots , $y_n = x^{(n-1)}$, so ist (4.9)-(4.10) äquivalent zu

$$y' = f(y, t), \quad t \in \mathbb{R}, \quad y(t_0) = y_0,$$

wobei

$$f(y,t) = \begin{pmatrix} y_2 \\ \vdots \\ y_n \\ F(y_n, \dots, y_1, t) \end{pmatrix}, \quad y_0 = \begin{pmatrix} x_0 \\ \vdots \\ x_{n-1} \end{pmatrix}.$$

Aus dem Satz von Picard-Lindelöf können wir auf die Lösbarkeit von (4.9) schließen: Ist $f = f(y,t)$ (lokal) lipschitzstetig in y gleichmäßig in t , so existiert eine eindeutige (lokale) Lösung.

Beispiel 4.23. Wir setzen Beispiel 4.22 fort:

$$y' = F(y) = \begin{pmatrix} y_2 \\ -gR^2(y_1 + R)^{-2} \end{pmatrix}, \quad y(0) = \begin{pmatrix} x_0 \\ x_1 \end{pmatrix}.$$

Physikalisch ist es sinnvoll, $y_1 \geq 0$ und $y_2 \in \mathbb{R}$ vorauszusetzen. Wir behaupten, dass F global lipschitzstetig ist; dies impliziert die Existenz einer globalen Lösung $y : [0, \infty) \rightarrow [0, \infty) \times \mathbb{R}$. Dazu rechnen wir

$$\begin{aligned} |F(y) - F(z)|^2 &= (y_2 - z_2)^2 + \left(\frac{gR^2}{(y_1 + R)^2} - \frac{gR^2}{(z_1 + R)^2} \right)^2 \\ &= (y_2 - z_2)^2 + g^2 R^4 \left(\frac{(z_1 + R)^2 - (y_1 + R)^2}{(y_1 + R)^2 (z_1 + R)^2} \right)^2 \\ &= (y_2 - z_2)^2 + g^2 R^4 \underbrace{\left(\frac{y_1 + z_1 + 2R}{(y_1 + R)^2 (z_1 + R)^2} (z_1 - y_1) \right)^2}_{\leq C_1(R)} \\ &\leq C_2(R) ((y_1 - z_1)^2 + (y_2 - z_2)^2) \\ &= C_2(R) |y - z|^2, \end{aligned}$$

wobei $C_i(R)$ positive Konstanten seien, die nur von R abhängen. Die Behauptung folgt mit $L = \sqrt{C_2(R)}$. \square

4.3 Lineare Differentialgleichungen

In diesem Abschnitt lösen wir lineare Differentialgleichungen der Form

$$x' = A(t)x + b(t), \quad t \in \mathbb{R}, \quad x(t_0) = x_0,$$

wobei $A(t) \in \mathbb{R}^{n \times n}$ die Koeffizientenmatrix und $b(t) \in \mathbb{R}^n$ die Inhomogenität seien. Ist $b = 0$, so nennen wir diese Gleichung *homogen*, ansonsten *inhomogen*. Sind A und b stetig und beschränkt auf \mathbb{R} , so ist $f(x,t) = A(t)x + b(t)$ global lipschitzstetig in x gleichmäßig in t , und das obige Anfangswertproblem besitzt eine eindeutige Lösung $x : \mathbb{R} \rightarrow \mathbb{R}^n$, die für alle $t \in \mathbb{R}$ definiert ist. In diesem Abschnitt stellen wir einige Methoden vor, mit denen das Anfangswertproblem (explizit) gelöst werden kann.

Zunächst betrachten wir homogene Systeme:

$$x' = A(t)x, \quad t \in \mathbb{R}. \quad (4.11)$$

Wir möchten den Raum aller Lösungen charakterisieren. Sind x_1 und x_2 zwei Lösungen, so ist klar, dass dann auch $x_1 + x_2$ und λx_1 für alle $\lambda \in \mathbb{R}$ Lösungen sind. Der Lösungsraum bildet also einen Vektorraum.

Lemma 4.24. *Die Menge aller Lösungen von (4.11) ist ein Unterraum im Vektorraum aller stetig differenzierbaren Funktionen $x : \mathbb{R} \rightarrow \mathbb{R}^n$.*

Sind $x_1, \dots, x_n : I \rightarrow \mathbb{R}^n$ Funktionen mit $x_j = (x_{j1}, \dots, x_{jn})^\top$, so nennen wir die Matrix

$$W(t) = (x_1(t), \dots, x_n(t)) \in \mathbb{R}^{n \times n}$$

die *Wronski-Matrix* und $\det W(t)$ die *Wronski-Determinante*. Mit Hilfe dieser Determinante können wir die lineare Unabhängigkeit von (x_1, \dots, x_n) nachweisen. Lineare Unabhängigkeit von (x_1, \dots, x_n) bedeutet, dass aus

$$\lambda_1 x_1(t) + \dots + \lambda_n x_n(t) = 0 \quad \text{für alle } t \in I \quad (4.12)$$

folgt $\lambda_1 = \dots = \lambda_n = 0$.

Proposition 4.25. *Wenn $\det W(t_0) \neq 0$ für ein $t_0 \in I$, dann sind (x_1, \dots, x_n) linear unabhängig.*

Beweis. Schreiben wir die Gleichung (4.12) komponentenweise

$$\begin{aligned} \lambda_1 x_{11}(t_0) + \dots + \lambda_n x_{n1}(t_0) &= 0, \\ &\vdots \\ \lambda_1 x_{1n}(t_0) + \dots + \lambda_n x_{nn}(t_0) &= 0, \end{aligned}$$

so erhalten wir eine Linearkombination der Spaltenvektoren der Wronski-Matrix in t_0 zum Nullvektor. Wegen $\det W(t_0) \neq 0$ müssen die Spaltenvektoren aber linear unabhängig sein, woraus $\lambda_1 = \dots = \lambda_n = 0$ folgt. \square

Bemerkung 4.26. Die Umkehrung der obigen Proposition ist i.Allg. falsch: Aus $\det W(t) = 0$ für alle $t \in I$ folgt *nicht* die lineare Abhängigkeit von (x_1, \dots, x_n) . Ein Gegenbeispiel ist gegeben durch

$$x_1(t) = \begin{cases} (0, 0)^\top & : t \leq 0 \\ (t^2, 2t)^\top & : t > 0, \end{cases} \quad x_2(t) = \begin{cases} (t^2, 2t)^\top & : t \leq 0 \\ (0, 0)^\top & : t > 0. \end{cases}$$

Es gilt für $t \leq 0$

$$\det W(t) = \det \begin{pmatrix} 0 & t^2 \\ 0 & 2t \end{pmatrix} = 0$$

und analog $\det W(t) = 0$ für $t > 0$. Andererseits folgt aus

$$\lambda_1 x_1(t) + \lambda_2 x_2(t) = 0$$

für $t = 1$, dass $\lambda_1 = 0$, und für $t = -1$, dass $\lambda_2 = 0$. Also ist (x_1, x_2) linear unabhängig. \square

Wie viele linear unabhängige Lösungen besitzt das homogene System (4.11)? Die Antwort liefert der folgende Satz.

Satz 4.27. Sei $A : \mathbb{R} \rightarrow \mathbb{R}^{n \times n}$ stetig. Dann besitzt das homogene System (4.11) genau n linear unabhängige Lösungen auf \mathbb{R} .

Beweis. Sei x_k ($k = 1, \dots, n$) die eindeutige Lösung von (4.11) mit Anfangswert $x_k(t_0) = e_k$, wobei e_k den k -ten Einheitsvektor im \mathbb{R}^n bezeichne. Da die Einheitsvektoren linear unabhängig sind, gilt dies auch für die Lösungen (x_1, \dots, x_n) . Um zu zeigen, dass es genau n linear unabhängige Lösungen gibt, sei $U \subset C^1(\mathbb{R}; \mathbb{R}^n)$ die Menge aller Lösungen von (4.11). Wegen der Linearität der Differentialgleichung ist U ein Unterraum von $C^0(\mathbb{R}; \mathbb{R}^n)$. Die Abbildung

$$F : U \rightarrow \mathbb{R}^n, \quad F(x) = (x_1(t_0), \dots, x_n(t_0)), \quad x = (x_1, \dots, x_n) \in U,$$

ist aus demselben Grund linear. Aus der Eindeutigkeit der Lösungen von (4.11) folgt die Injektivität von F : Gilt nämlich $F(x) = 0$ für ein $x \in U$, so löst x die Differentialgleichung mit verschwindenden Anfangswerten. Dieses Anfangswertproblem besitzt die eindeutige Lösung $x = 0$. Außerdem ist F surjektiv, was sofort aus der Lösbarkeit des entsprechenden Anfangswertproblems folgt. Damit ist F eine bijektive Funktion und $\dim(U) = \dim(\mathbb{R}^n) = n$. \square

Definition 4.28 (Fundamentalsystem). Eine Basis (x_1, \dots, x_n) des Lösungsraums des homogenen Systems (4.11) heißt ein Fundamentalsystem. Schreiben wir die Basis als $X = (x_1, \dots, x_n)$, so nennen wir X eine Fundamentalmatrix.

Beispiel 4.29. Wir wollen das System

$$x' = A(t)x, \quad t \in [1, \infty), \quad A(t) = \frac{1}{2} \begin{pmatrix} -1/t & 1/t^2 \\ 1 & 1/t \end{pmatrix},$$

lösen. Wir vermuten, dass die Lösungen aus Funktionen der Form t^α bestehen und machen den Ansatz $x(t) = (at^\alpha, bt^\beta)$ mit gewissen reellen Konstanten a , α , b und β . Setzen wir den Ansatz in die Differentialgleichung ein, so erhalten wir

$$\begin{pmatrix} a\alpha t^{\alpha-1} \\ b\beta t^{\beta-1} \end{pmatrix} = x' = A(t)x = \frac{1}{2} \begin{pmatrix} -at^{\alpha-1} + bt^{\beta-2} \\ at^\alpha + bt^{\beta-1} \end{pmatrix}.$$

Damit die Exponenten in den t -Termen übereinstimmen, sollte $\alpha = \beta - 1$ gelten. Dann können wir die erste Gleichung durch $t^{\alpha-1} \neq 0$ dividieren und die zweite Gleichung durch t^α :

$$a\alpha = \frac{1}{2}(-a + b), \quad b\beta = \frac{1}{2}(a + b)$$

oder als lineares Gleichungssystem in a und b :

$$\begin{pmatrix} 1 + 2\alpha & -1 \\ -1 & 1 + 2\alpha \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Ist die Determinante der Koeffizientenmatrix ungleich null, so gibt es nur die triviale Lösung $a = b = 0$. Anderenfalls folgt $(1 + 2\alpha)^2 - 1 = 0$ oder äquivalent $4\alpha(\alpha + 1) = 0$. Dies impliziert $\alpha = 0$ oder $\alpha = -1$. Im ersten Fall ist $\beta = \alpha + 1 = 1$ und $a = b$, im zweiten Fall $\beta = 0$ und $a = -b$. Wir erhalten die Lösungen

$$x_1(t) = a \begin{pmatrix} 1 \\ t \end{pmatrix}, \quad x_2(t) = a \begin{pmatrix} 1/t \\ -1 \end{pmatrix}.$$

Damit ist

$$X(t) = \begin{pmatrix} 1 & 1/t \\ t & -1 \end{pmatrix}$$

eine Fundamentalmatrix. □

Leider gibt es keine allgemeine Lösungsmethode für homogene Systeme der Form (4.11). Den Fall eines Systems mit konstanten Koeffizienten behandeln wir im folgenden Abschnitt. Wir kommen zu inhomogenen Anfangswertproblemen.

Satz 4.30 (Inhomogene lineare Systeme). *Seien $A : [t_0, t_1] \rightarrow \mathbb{R}^{n \times n}$ stetig, $b \in L^1(t_0, t_1)$ und $X \in C^0([t_0, t_1]; \mathbb{R}^{n \times n})$ eine Fundamentalmatrix von (4.11). Dann besitzt das Anfangswertproblem*

$$x' = A(t)x + b(t), \quad t \geq t_0, \quad x(t_0) = x_0, \quad (4.13)$$

die eindeutig bestimmte Lösung

$$x(t) = \phi(t, t_0)x_0 + \int_{t_0}^t \phi(t, s)b(s)ds, \quad t \in [t_0, t_1],$$

wobei $\phi(t, s) = X(t)X^{-1}(s)$ die Übergangsmatrix heißt.

Die obige Lösungsformel hat eine gewisse Ähnlichkeit mit der entsprechenden Formel für lineare skalare Differentialgleichungen (siehe Proposition 4.9) und kann mit der Methode der Variation der Konstanten hergeleitet werden.

Beweis. Zuerst bemerken wir, dass die Übergangsmatrix wohldefiniert ist, da $X(t)$ nach Konstruktion invertierbar ist. Außerdem wissen wir bereits, dass (4.13) eindeutig lösbar ist und dass $x(t_0) = \phi(t_0, t_0)x_0 = x_0$. Es bleibt lediglich die Lösungsformel zu verifizieren. Dazu differenzieren wir die Übergangsmatrix:

$$\frac{\partial \phi}{\partial t}(t, s) = X'(t)X^{-1}(s) = A(t)X(t)X^{-1}(s) = A(t)\phi(t, s).$$

Insbesondere gilt $\phi(t, t) = X(t)X(t)^{-1} = I$. Es folgt

$$\begin{aligned} x'(t) &= A(t)\phi(t, t_0)x_0 + \phi(t, t)b(t) + \int_{t_0}^t A(t)\phi(t, s)b(s)ds \\ &= A(t)\left(\phi(t, t_0)x_0 + \int_{t_0}^t \phi(t, s)b(s)ds\right) + b(t) = A(t)x(t) + b(t), \end{aligned}$$

also die Behauptung. □

Beispiel 4.31. Wir setzen Beispiel 4.29 fort und betrachten $x' = A(t)x + b(t)$, $x(1) = (1, 0)^\top$, mit

$$A(t) = \frac{1}{2} \begin{pmatrix} -1/t & 1/t^2 \\ 1 & 1/t \end{pmatrix}, \quad b(t) = \begin{pmatrix} 0 \\ t \end{pmatrix}, \quad t \geq 1.$$

Die Übergangsmatrix lautet

$$\phi(t, s) = \begin{pmatrix} 1 & -1/t \\ t & 1 \end{pmatrix} \cdot \frac{1}{2} \begin{pmatrix} 1 & 1/s \\ -s & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1+s/t & 1/s-1/t \\ t-s & 1+t/s \end{pmatrix}.$$

Damit erhalten wir mit

$$\begin{aligned} x(t) &= \frac{1}{2} \begin{pmatrix} 1+1/t \\ t-1 \end{pmatrix} + \frac{1}{2} \int_1^t \begin{pmatrix} 1+s/t & 1/s-1/t \\ t-s & 1+t/s \end{pmatrix} \begin{pmatrix} 0 \\ s \end{pmatrix} ds \\ &= \frac{1}{2} \begin{pmatrix} 1+1/t \\ t-1 \end{pmatrix} + \frac{1}{2} \int_1^t \begin{pmatrix} 1-s/t \\ s+t \end{pmatrix} ds \\ &= \frac{1}{2} \begin{pmatrix} 1+1/t \\ t-1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} t/2-1+1/(2t) \\ 3t^2/2-t-1/2 \end{pmatrix} \\ &= \frac{1}{4} \begin{pmatrix} t+3/t \\ 3(t^2-1) \end{pmatrix} \end{aligned}$$

die Lösung des obigen Anfangswertproblems. □

Mit den bisher vorgestellten Methoden können wir auch lineare Differentialgleichungen höherer Ordnung lösen. Dies sind Differentialgleichungen der Form

$$x^{(n)} + a_{n-1}(t)x^{(n-1)} + \dots + a_1(t)x' + a_0(t)x = f(t), \quad t \in \mathbb{R}. \quad (4.14)$$

Hierbei sind $a_0, \dots, a_{n-1}, f: \mathbb{R} \rightarrow \mathbb{R}$ stetige Funktionen. Schreiben wir nämlich

$$y_1 = x, \quad y_2 = x', \dots, \quad y_n = x^{(n-1)},$$

so können wir (4.14) äquivalent umformulieren als

$$\begin{aligned} y_1' &= y_2, \\ y_2' &= y_3, \\ &\vdots \\ y_{n-1}' &= y_n, \\ y_n' &= -a_{n-1}(t)y_n - \dots - a_0(t)y_1 + f(t) \end{aligned}$$

oder mit dem Vektor $y = (y_1, \dots, y_n)$

$$y' = A(t)y + b(t), \quad t \in \mathbb{R},$$

wobei

$$A(t) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & \vdots & & \ddots & \\ 0 & 0 & 0 & & 1 \\ -a_0(t) & -a_1(t) & -a_2(t) & \cdots & -a_{n-1}(t) \end{pmatrix}, \quad b(t) = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ f(t) \end{pmatrix}.$$

Aus den oben geführten Ergebnissen folgt sofort der folgende Satz.

Satz 4.32 (Lineare Differentialgleichungen n -ter Ordnung). *Seien $a_{n-1}, \dots, a_0, f : \mathbb{R} \rightarrow \mathbb{R}$ stetige Funktionen. Dann besitzt das Anfangswertproblem (4.14) und*

$$x(t_0) = x_0, \quad x'(t_0) = x_1, \quad \dots, \quad x^{(n-1)}(t_0) = x_{n-1}$$

eine eindeutige Lösung auf \mathbb{R} . Das homogene Problem (4.14) mit $b = 0$ besitzt genau n linear unabhängige Lösungen (y_1, \dots, y_n) mit der Eigenschaft

$$\det \begin{pmatrix} y_1(t) & \cdots & y_n(t) \\ y_1'(t) & \cdots & y_n'(t) \\ \vdots & & \vdots \\ y_1^{(n-1)}(t) & \cdots & y_n^{(n-1)}(t) \end{pmatrix} \neq 0 \quad \text{für alle } t \in \mathbb{R}.$$

Beispiel 4.33. Wir betrachten einen LC-Schaltkreis, in dem ein Kondensator und eine Spule in Reihe geschaltet sind. Der Strom genügt der Differentialgleichung

$$I'' + \omega^2 I = \frac{U'(t)}{L}, \quad t > 0, \quad I(0) = I_0, \quad I'(0) = I_1.$$

Ist die angelegte Spannung konstant, $U'(t) = 0$, so müssen wir nur die homogene Gleichung lösen, die wir mit $x = (I, I')$ auch als das System

$$x' = Ax, \quad t > 0, \quad x(0) = x_0,$$

schreiben können, wobei

$$A = \begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix}, \quad x_0 = \begin{pmatrix} I_0 \\ I_1 \end{pmatrix},$$

und $\omega > 0$. Das Fundamentalsystem können wir leicht erraten, denn $I(t) = \sin(\omega t)$ und $I(t) = \cos(\omega t)$ sind zwei linear unabhängige Lösungen:

$$x_1(t) = \begin{pmatrix} \sin(\omega t) \\ \omega \cos(\omega t) \end{pmatrix}, \quad x_2(t) = \begin{pmatrix} \cos(\omega t) \\ -\omega \sin(\omega t) \end{pmatrix}.$$

Wir prüfen die Wronski-Determinante nach:

$$\det(x_1, x_2) = \det \begin{pmatrix} \sin(\omega t) & \cos(\omega t) \\ \omega \cos(\omega t) & -\omega \sin(\omega t) \end{pmatrix} = -\omega \sin^2(\omega t) - \omega \cos^2(\omega t) = -\omega \neq 0$$

für alle $t \in \mathbb{R}$. Also sind (x_1, x_2) wirklich linear unabhängig und ein Fundamentalsystem für das obige System. \square

4.4 Lineare Differentialgleichungen mit konstanten Koeffizienten

Wir untersuchen lineare Differentialgleichungen mit konstanten Koeffizienten:

$$x' = Ax + b(t), \quad t \in \mathbb{R}, \quad x(t_0) = x_0,$$

wobei $A \in \mathbb{R}^{n \times n}$ eine Matrix und $b : \mathbb{R} \rightarrow \mathbb{R}$ eine stetige Funktion seien. Wir betrachten zuerst die homogene Situation $b = 0$.

Ist A eine Zahl, können wir die Lösung des homogenen Problems $x' = Ax$, $x(t_0) = x_0$, sofort hinschreiben: $x(t) = e^{A(t-t_0)}x_0$. Diese Schreibweise können wir erstaunlicherweise auf den Matrizenfall übertragen. Hierzu müssen wir definieren, was wir unter e^A verstehen. Ähnlich wie in der Analysis definieren wir wie folgt.

Definition 4.34 (Matrix-Exponentialfunktion). Sei $A \in \mathbb{R}^{n \times n}$. Wir definieren die Matrix-Exponentialfunktion von A durch

$$e^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k.$$

Es ist zu klären, ob durch diese Definition eine konvergente Reihe erklärt ist. Sei die Matrixnorm wie bei linearen Operatoren über

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

mit einer beliebigen Norm $\|x\|$ auf \mathbb{R}^n gegeben. Die Norm darf beliebig sein, da auf endlich-dimensionalen Räumen alle Normen äquivalent sind. Dann gilt $\|Ax\| \leq \|A\| \|x\|$ für alle $x \in \mathbb{R}^n$ nach Definition sowie wegen $\|ABx\| \leq \|A\| \|Bx\| \leq \|A\| \|B\| \|x\|$ auch die *Submultiplikationseigenschaft*

$$\|AB\| \leq \|A\| \|B\| \quad \text{für alle } A, B \in \mathbb{R}^{n \times n}. \quad (4.15)$$

Dann folgt aus der Definition der Matrix-Exponentialfunktion

$$\|e^A\| \leq \sum_{k=0}^{\infty} \frac{1}{k!} \|A^k\| \leq \sum_{k=0}^{\infty} \frac{1}{k!} \|A\|^k \leq e^{\|A\|}, \quad (4.16)$$

also ist e^A für alle Matrizen A konvergent.

Satz 4.35 (Homogene lineare Systeme). Seien $A \in \mathbb{R}^{n \times n}$, $t_0 \in \mathbb{R}$ und $x_0 \in \mathbb{R}^n$. Dann ist $x(t) = e^{A(t-t_0)}x_0$, $t \in \mathbb{R}$, die eindeutige Lösung des homogenen Anfangswertproblems

$$x' = Ax, \quad x(t_0) = x_0. \quad (4.17)$$

Beweis. Die Reihe $e^{A(t-t_0)}$ konvergiert wegen (4.16) gleichmäßig in t , sofern t in einem beschränkten Intervall liegt. Daher dürfen wir Differentiation nach t und die Summe vertauschen:

$$\begin{aligned} x'(t) &= \frac{d}{dt} e^{A(t-t_0)}x_0 = \sum_{k=0}^{\infty} \frac{1}{k!} \frac{d}{dt} (t-t_0)^k A^k x_0 = \sum_{k=1}^{\infty} \frac{1}{(k-1)!} (t-t_0)^{k-1} A^k x_0 \\ &= \sum_{j=0}^{\infty} \frac{1}{j!} (t-t_0)^j A(A^j x_0) = A e^{A(t-t_0)}x_0 = Ax(t). \end{aligned}$$

Außerdem gilt $x(t_0) = e^0 x_0 = Ix_0 = x_0$. □

Gemäß den Definitionen aus dem vorigen Abschnitt ist $X(t) = e^{A(t-t_0)}$ eine Fundamentalmatrix und $\phi(t, s) = X(t)X(s)^{-1} = e^{A(t-t_0)}e^{-A(s-t_0)} = e^{A(t-s)}$ die Übergangsmatrix. Die letzte Gleichheit ist eine Konsequenz aus folgendem Resultat

$$e^{At}e^{As} = e^{A(t+s)} = e^{As}e^{At} \quad \text{für alle } s, t \in \mathbb{R}$$

(Übungsaufgabe).

Wie können wir e^{At} ausrechnen? Dazu erinnern wir an einige Ergebnisse aus der linearen Algebra: Jede Matrix $A \in \mathbb{C}^{n \times n}$ kann mit Hilfe einer Ähnlichkeitstransformation auf die *Jordansche Normalenform* transformiert werden, d.h., es gibt eine invertierbare Matrix $S \in \mathbb{C}^{n \times n}$ und Matrizen J_1, \dots, J_p , so dass

$$A = SJS^{-1} \quad \text{mit } J = \text{diag}(J_1, \dots, J_p), \quad J_k = \lambda_k I + N_k, \quad k = 1, \dots, p,$$

und die Matrizen N_k sind gegeben durch

$$N_k = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & & & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \in \mathbb{R}^{k \times k}.$$

Hierbei bezeichnet $\text{diag}(J_1, \dots, J_p)$ eine Block-Diagonalmatrix. Damit können wir e^{At} vereinfachen:

$$\begin{aligned} e^{At} &= \sum_{k=0}^{\infty} \frac{t^k}{k!} (SJS^{-1})^k = \sum_{k=0}^{\infty} \frac{t^k}{k!} SJ^k S^{-1} = S \sum_{k=0}^{\infty} \frac{t^k}{k!} \text{diag}(J_1^k, \dots, J_p^k) S^{-1} \\ &= S \text{diag}(e^{J_1 t}, \dots, e^{J_p t}) S^{-1}. \end{aligned} \quad (4.18)$$

Wegen

$$e^{J_k t} = e^{(\lambda_k I + N_k)t} = e^{\lambda_k t} e^{N_k t}$$

genügt es also, $e^{N_k t}$ auszurechnen. Die Matrix N_k ist nilpotent: Eine Rechnung zeigt, dass $N_k^k = 0$. Dies bedeutet, dass die Reihe $e^{N_k t}$ sich zu einer Summe der ersten k Summanden reduziert. Da in dieser Summe Potenzen von t auftreten, schließen wir, dass $e^{N_k t}$ Terme $c_j t^j$ mit $c_j \geq 0$ und $j \in \{0, \dots, k-1\}$ enthält. Insgesamt enthalten die Elemente der Matrix e^{At} also Ausdrücke der Form $c_j t^j e^{\lambda_k t}$, wobei λ_k komplex sein kann.

In Rechnungen ist es zuweilen einfacher, direkt einen Lösungsansatz der Form $x(t) = e^{\lambda t} v$ mit $\lambda \in \mathbb{C}$ und $v \in \mathbb{C}^n$ zu machen. Setzen wir diesen Ansatz in das homogene System ein, so folgt

$$\lambda e^{\lambda t} v = x' = Ax = e^{\lambda t} Av$$

oder $Av = \lambda v$. Der Ansatz führt also zum Erfolg, wenn λ ein Eigenwert von A mit zugehörigem Eigenvektor v ist. Ist der Eigenraum mehrdimensional mit der Basis (v_1, \dots, v_q) aus Eigenvektoren, so sind $e^{\lambda t} v_1, \dots, e^{\lambda t} v_q$ Lösungen des homogenen Systems. Etwas schwieriger wird es,

wenn die geometrische Vielfachheit (Dimension des Eigenraums) kleiner ist als die algebraische Vielfachheit (ein k -facher Eigenwert), da wir dann Hauptvektoren v benötigen. In der Matrix-Exponentialfunktion ist dieser Fall mit den nilpotenten Matrizen dargestellt. Wir geben unten nur ein einfaches Beispiel und verweisen für allgemeine Resultate auf die Literatur.

Beispiel 4.36. (1) Betrachte das homogene System $x' = Ax$ mit

$$A = \begin{pmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{pmatrix}.$$

Wir berechnen die Eigenwerte aus den Nullstellen der charakteristischen Gleichung:

$$0 = \det(A - \lambda I) = -\lambda^3 - \lambda^2 + 21\lambda + 45 = -(\lambda - 5)(\lambda + 3)^2,$$

also $\lambda_1 = 5$ und $\lambda_2 = -3$. Die Eigenvektoren zum Eigenwert $\lambda_1 = 5$ erhalten wir aus dem Gleichungssystem $(A - 5I)x = 0$ mit dem eindimensionalen Lösungsraum $\text{span}(1, 2, -1)^\top$. Der Lösungsraum zum Gleichungssystem $(A + 3I)x = 0$ ist zweidimensional, und wir wählen die beiden linear unabhängigen Vektoren $(-2, 1, 0)^\top$ und $(3, 0, 1)^\top$. Daher lautet die allgemeine Lösung von $x' = Ax$

$$x(t) = c_1 e^{-3t} \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} + c_2 e^{-3t} \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} + c_3 e^{5t} \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}.$$

(2) Betrachte $x' = Ax$ mit

$$A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

Aus der charakteristischen Gleichung $\lambda^2 - 2\lambda + 2 = 0$ erhalten wir die Eigenwerte $1 \pm i$ mit zugehörigen Eigenvektoren $(\pm i, 1)^\top$. Die allgemeine Lösung von $x' = Ax$ lautet

$$x(t) = C_1 e^{(1+i)t} \begin{pmatrix} i \\ 1 \end{pmatrix} + C_2 e^{(1-i)t} \begin{pmatrix} -i \\ 1 \end{pmatrix} \quad \text{mit } C_1, C_2 \in \mathbb{C}.$$

Da die Matrix A reell ist, würden wir gerne reellwertige Lösungen erhalten. Dies erreichen wir, indem wir den Real- und Imaginärteil einer der beiden Lösungsanteile berechnen. Wegen

$$e^{(1+i)t} \begin{pmatrix} i \\ 1 \end{pmatrix} = e^t (\cos t + i \sin t) \begin{pmatrix} i \\ 1 \end{pmatrix} = e^t \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} + i e^t \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$$

erhalten wir die allgemeine reelle Lösung

$$x(t) = c_1 \text{Re} e^{(1+i)t} \begin{pmatrix} i \\ 1 \end{pmatrix} + c_2 \text{Im} e^{(1+i)t} \begin{pmatrix} i \\ 1 \end{pmatrix} = c_1 e^t \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} + c_2 e^t \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}, \quad c_1, c_2 \in \mathbb{R}.$$

Der andere Lösungsanteil $e^{(1-i)t}(-i, 1)^\top$ ist lediglich das komplex Konjugierte des ersten Lösungsanteils und liefert daher keine neuen reellen Lösungen.

(3) Als letztes Beispiel untersuchen wir

$$x' = Ax = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix} x, \quad x(0) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Die charakteristische Gleichung $(1 + \lambda)^2 = 0$ führt auf den doppelten Eigenwert $\lambda = -1$. Der Eigenraum ist nur eindimensional, und wir wählen den Eigenvektor $v_1 = (1, 0)^\top$. Eine Lösung lautet also $x_1(t) = (e^{-t}, 0)^\top$. Um eine zweite linear unabhängige Lösung zu finden, können wir einen Hauptvektor bestimmen oder im Hinblick auf die Jordansche Normalenform einen Ansatz machen. Die obige Matrix lässt sich zerlegen in

$$\begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \lambda I + N_1.$$

Da $N_1^2 = 0$, enthält die Matrix $e^{N_1 t}$ nur Polynome in t höchstens der Ordnung eins. Wir machen daher den Ansatz

$$x(t) = e^{\lambda t} \begin{pmatrix} a + bt \\ c + dt \end{pmatrix}, \quad \lambda = -1.$$

Setzen wir ihn in die Differentialgleichung ein und dividieren durch e^{-t} , so folgt

$$\begin{pmatrix} -(a + bt) + b \\ -(c + dt) + d \end{pmatrix} = \begin{pmatrix} -(a + bt) + (c + dt) \\ -(c + dt) \end{pmatrix}$$

oder $b = c + dt$ und $d = 0$. Dies impliziert $b = c$, und wir erhalten

$$x(t) = e^{-t} \begin{pmatrix} a + bt \\ b \end{pmatrix}, \quad a, b \in \mathbb{R}.$$

Den Lösungsanteil $e^{-t}(a, 0)^\top$ haben wir schon oben gefunden. Daher ist (mit $b = 1$) $x_2(t) = e^{-t}(t, 1)^\top$ die gesuchte zweite Lösung. Ein Fundamentalsystem lautet

$$e^{At} = (x_1(t), x_2(t)) = e^{-t} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}.$$

Die Lösung des Anfangswertproblems ist dann gegeben durch $x(t) = e^{At}(1, 1)^\top = e^{-t}(1 + t, 1)^\top$. \square

Die Darstellung der Lösung mit Hilfe der Matrix-Exponentialfunktion erlaubt es, eine Aussage über das Langzeitverhalten der Lösungen zu machen.

Satz 4.37 (Langzeitverhalten für homogene lineare Systeme). *Sei $\|\cdot\|$ eine Norm auf $\mathbb{C}^{n \times n}$, die die Submultiplikationseigenschaft (4.15) erfüllt und sei $A \in \mathbb{R}^{n \times n}$ eine Matrix mit Eigenwerten $\lambda_1, \dots, \lambda_p \in \mathbb{C}$. Wenn $\operatorname{Re}(\lambda_k) < 0$ für alle $k = 1, \dots, p$, dann existiert zu jedem $\varepsilon > 0$ ein $K = K(\varepsilon)$ mit*

$$\|e^{At}\| \leq K e^{(-\gamma + \varepsilon)t}, \quad t > 0,$$

wobei $\gamma = -\max\{\operatorname{Re}(\lambda_k) : k = 1, \dots, p\} > 0$. Insbesondere gilt für jede Lösung von $x' = Ax$, $x(t_0) = x_0$,

$$\|x(t)\| \rightarrow 0 \quad \text{für } t \rightarrow \infty.$$

Beweis. Aus (4.18) und (4.15) folgt

$$\|e^{At}\| = \|S \operatorname{diag}(e^{J_1 t}, \dots, e^{J_p t}) S^{-1}\| \leq \|S\| \|\operatorname{diag}(e^{J_1 t}, \dots, e^{J_p t})\| \|S^{-1}\| \leq C \|\operatorname{diag}(e^{J_1 t}, \dots, e^{J_p t})\|.$$

Wir wählen eine Norm, für die $\|\operatorname{diag}(A, B)\|^2 \leq \|A\|^2 + \|B\|^2$ und $\|I\| = 1$ erfüllt ist; die folgenden Abschätzungen gelten für jede Norm, da alle Normen auf $\mathbb{C}^{n \times n}$ äquivalent sind. Wegen $J_k = \lambda_k I + N_k$ und der Definition von γ ergibt sich

$$\|e^{At}\|^2 \leq C^2 \sum_{k=1}^p \|e^{J_k t}\|^2 = C^2 \sum_{k=1}^p |e^{\lambda_k t}|^2 \|e^{N_k t}\|^2 \leq C^2 e^{-2\gamma t} \sum_{k=1}^p \|e^{N_k t}\|^2.$$

Die Koeffizienten der Matrix $e^{N_k t}$ sind Polynome in t , so dass deren Norm durch ein Polynom in t abgeschätzt werden kann. Zu jedem $\varepsilon > 0$ können wir dieses Polynom durch $C_0^2(\varepsilon) e^{2\varepsilon t}$ für ein geeignetes $C_0^2(\varepsilon) > 0$ abschätzen. Also folgt

$$\|e^{At}\|^2 \leq C^2 C_0^2(\varepsilon) e^{2(-\gamma+\varepsilon)t}.$$

Wählen wir nun $0 < \varepsilon < \gamma$, konvergiert $e^{(-\gamma+\varepsilon)t}$ gegen null für $t \rightarrow \infty$. Also ist

$$\|x(t)\| \leq \|e^{A(t-t_0)}\| \|x_0\| \leq C C_0(\varepsilon) \|x_0\| e^{(-\gamma+\varepsilon)(t-t_0)} \rightarrow 0 \quad \text{für } t \rightarrow \infty,$$

was die zweite Behauptung zeigt. □

Aus Satz 4.30 folgt eine Lösungsformel für inhomogene Systeme.

Satz 4.38 (Inhomogene lineare Systeme). *Seien $A \in \mathbb{R}^{n \times n}$, $b \in L^1(\mathbb{R}; \mathbb{R}^n)$ und $x_0 \in \mathbb{R}^n$. Dann besitzt*

$$x' = Ax + b(t), \quad t \in \mathbb{R}, \quad x(t_0) = x_0,$$

die eindeutige Lösung

$$x(t) = e^{A(t-t_0)} x_0 + \int_{t_0}^t e^{A(t-s)} b(s) ds, \quad t \in \mathbb{R}.$$

Beweis. Der Satz folgt aus der Lösungsformel in Satz 4.30 und der Darstellung $\phi(t, s) = e^{A(t-s)}$ für die Übergangsmatrix. □

Beispiel 4.39. Als Beispiel lösen wir das inhomogene Anfangswertproblem

$$x' = Ax + b(t) = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix} x + \begin{pmatrix} \sin t \\ \sin t \end{pmatrix}, \quad x(0) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Wir haben bereits in Beispiel 4.36 das homogene Problem gelöst mit

$$e^{At} = e^{-t} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}.$$

Mit der Lösungsformel aus Satz 4.38 erhalten wir

$$\begin{aligned}
 x(t) &= e^{-t} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \int_0^t e^{-(t-s)} \begin{pmatrix} 1 & t-s \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \sin s \\ \sin s \end{pmatrix} ds \\
 &= \begin{pmatrix} (1+t)e^{-t} \\ e^{-t} \end{pmatrix} + \int_0^t \begin{pmatrix} e^{-(t-s)}(1+t-s) \sin s \\ e^{-(t-s)} \sin s \end{pmatrix} ds \\
 &= \begin{pmatrix} (1+t)e^{-t} \\ e^{-t} \end{pmatrix} + \begin{pmatrix} (1+\frac{1}{2}t)e^{-t} + \frac{1}{2} \sin t - \cos t \\ \frac{1}{2}(e^{-t} + \sin t - \cos t) \end{pmatrix} \\
 &= \begin{pmatrix} (2+\frac{3}{2}t)e^{-t} + \frac{1}{2} \sin t - \cos t \\ \frac{1}{2}(3e^{-t} + \sin t - \cos t) \end{pmatrix}.
 \end{aligned}$$

Die homogene Lösung konvergiert gegen null im Langzeitgrenzwert $t \rightarrow \infty$. Dies gilt nicht für den inhomogenen Lösungsanteil, da die rechte Seite mit den Sinustermen nichtverschwindende angeregte Kräfte darstellt. Nach hinreichend langer Zeit wird die Lösung jedoch durch Sinus- und Kosinusterme dominiert, und zwar mit derselben Frequenz wie die anregende Kraft. \square

Abschließend behandeln wir kurz den Fall linearer Gleichungen höherer Ordnung. Diese können grundsätzlich auf den Fall linearer Systeme zurückgeführt werden, wie wir am Ende von Abschnitt 4.3 ausgeführt haben. Wir wissen bereits, wie die Lösungen aussehen: Sie enthalten Terme der Form $e^{\lambda t}$ bzw., wenn Hauptvekturlösungen erforderlich sind, $t^k e^{\lambda t}$ mit $\lambda \in \mathbb{C}$. Es bietet sich also an, bei *homogenen* Gleichungen

$$x^{(n)} + a_{n-1}x^{(n-1)} + \dots + a_1x' + a_0x = 0, \quad t \in \mathbb{R},$$

den Ansatz $e^{\lambda t}$ zu probieren. Dies führt auf

$$(\lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_1\lambda + a_0)e^{\lambda t} = 0$$

oder, nach Division durch $e^{\lambda t}$,

$$\lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_1\lambda + a_0 = 0.$$

Wenn diese Gleichung n verschiedene Lösungen besitzt, erhalten wir eine Lösungsbasis mit den Funktionen $e^{\lambda t} = e^{\alpha t}(\cos(\beta t) + i \sin(\beta t))$, wobei $\lambda = \alpha + i\beta$, die wir in Real- und Imaginärteile zerlegen können, um eine reelle Lösungsbasis zu gewinnen. Im Falle eines k -fachen Eigenwerts λ können wir die linear unabhängigen Funktionen $e^{\lambda t}, t e^{\lambda t}, \dots, t^{k-1} e^{\lambda t}$ verwenden.

Beispiel 4.40. (1) Die Differentialgleichung

$$x'' + 4x' + 4x = 0, \quad t \in \mathbb{R},$$

besitzt die charakteristische Gleichung $\lambda^2 + 4\lambda + 4 = (\lambda + 2)^2 = 0$ mit einem doppelten Eigenwert $\lambda = -2$. Die allgemeine Lösung lautet

$$x(t) = c_1 e^{-2t} + c_2 t e^{-2t}, \quad c_1, c_2 \in \mathbb{R}.$$

(2) Die charakteristische Gleichung von

$$x'' + 6x' + 10x = 0, \quad t \in \mathbb{R},$$

besitzt die Lösungen $\lambda_{1/2} = -3 \pm i$. Die allgemeine komplexe Lösung lautet also

$$x(t) = C_1 e^{(-3+i)t} + C_2 e^{(-3-i)t} = C_1 e^{-3t} (\cos t + i \sin t) + C_2 e^{-3t} (\cos t - i \sin t), \quad C_1, C_2 \in \mathbb{C}.$$

Die allgemeine reelle Lösung ist die Kombination aus dem Real- und Imaginärteil des ersten Summanden (der zweite Summand liefert keine neue Information):

$$x(t) = c_1 e^{-3t} \cos t + c_2 e^{-3t} \sin t, \quad c_1, c_2 \in \mathbb{R}.$$

□

4.5 Stabilität linearer autonomer Systeme

In den vorigen beiden Abschnitten haben wir Lösungsmethoden für *lineare* Differentialgleichungen und ihre Systeme vorgestellt. Leider gibt es für *nichtlineare* Differentialgleichungen keine vergleichbar allgemeinen Lösungstechniken. Andererseits können diese Probleme häufig näherungsweise mit Hilfe linearer Gleichungen gelöst werden. Ein Grund dafür ist die Tatsache, dass sich viele Systeme in den Anwendungen in der Nähe des Gleichgewichts befinden. Sei nämlich $x(t)$ eine Lösung von $x' = f(x)$. Unter einem *Gleichgewicht* verstehen wir eine konstante Lösung \bar{x} mit $f(\bar{x}) = 0$. Befindet sich $x(t)$ in der Nähe von \bar{x} , so folgt nach Taylor-Entwicklung von f :

$$x' = f(x) = \underbrace{f(\bar{x})}_{=0} + \underbrace{f'(\bar{x})}_{=:A}(x - \bar{x}) + O(|x - \bar{x}|^2) = A(x - \bar{x}) + O(|x - \bar{x}|^2).$$

Führen wir die neue Variable $y = x - \bar{x}$ ein, so ist diese Gleichung von der Form

$$y' = Ay + O(|y|^2),$$

wobei die rechte Seite $O(|y|^2)$ "klein" ist, falls $x(t)$ hinreichend nahe an \bar{x} liegt.

In vielen Anwendungen ist man daran interessiert zu verstehen, ob eine Lösung einer Differentialgleichung gegen ein Gleichgewicht konvergiert oder wie man das System in die Gleichgewichtslage bringen kann. In diesem Abschnitt untersuchen wir die *Stabilität* dieser Gleichgewichtslagen für lineare (linearisierte) und nichtlineare Systeme. Wir betrachten nur autonome Gleichungen der Form

$$x' = f(x), \quad t \geq 0, \quad x(0) = \xi. \quad (4.19)$$

Wir bemerken, dass wir ohne Einschränkung $t_0 = 0$ setzen können. Anderenfalls (wenn $x(t_0) = \xi$) erfüllt die Funktion $\tilde{x}(t) = x(t + t_0)$ das Anfangswertproblem $\tilde{x}' = f(\tilde{x})$, $\tilde{x}(0) = x(t_0) = \xi$. Wir nehmen an, dass $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine lokal lipschitzstetige Funktion ist und dass es ein $\bar{x} \in \mathbb{R}^n$ mit $f(\bar{x}) = 0$ gibt. Unter der ersten Bedingung liefert der Satz 4.15 von Picard-Lindelöf lokale Lösungen $x(\cdot, \xi): [0, b) \rightarrow \mathbb{R}^n$ für ein $b > 0$. Die zweite Bedingung besagt, dass \bar{x} ein Gleichgewicht von (4.19) ist. Insbesondere ist $x(\cdot, \bar{x}) = \bar{x}$ eine konstante Lösung von (4.19). Wir führen die folgenden Begriffe ein.

Definition 4.41 (Stabilität). Seien $x : I \rightarrow \mathbb{R}^n$ eine Lösung von (4.19) und $\bar{x} \in \mathbb{R}^n$ mit $f(\bar{x}) = 0$.

(i) Wir nennen in diesem Fall $\bar{x} \in \mathbb{R}^n$ ein Gleichgewicht (oder Gleichgewichtspunkt) von (4.19).

(ii) Ein Gleichgewicht \bar{x} heißt stabil (im Sinne von Lyapunov), wenn

$$\forall \varepsilon > 0 : \exists \delta > 0 : \forall \xi \in B_\delta(\bar{x}) : I = [0, \infty) \text{ und } \forall t \geq 0 : x(t, \xi) \in B_\varepsilon(\bar{x}).$$

(iii) Ein Gleichgewicht \bar{x} heißt attraktiv, wenn

$$\exists \delta > 0 : \forall \xi \in B_\delta(\bar{x}) : I = [0, \infty) \text{ und } \lim_{t \rightarrow \infty} x(t, \xi) = \bar{x}.$$

(iv) Ein Gleichgewicht heißt asymptotisch stabil, wenn es stabil und attraktiv ist.

(v) Ein Gleichgewicht heißt instabil, wenn es nicht stabil ist.

Die Bedingung $I = [0, \infty)$ stellt lediglich sicher, dass die Lösung für alle Zeiten $t \geq 0$ existiert. Ein Gleichgewicht ist also nur dann stabil, wenn die entsprechende Lösung für alle Zeiten existiert. Stabilität eines Gleichgewichts bedeutet, dass die Lösung für alle Zeiten in einer Kugel um das Gleichgewicht bleibt, wenn die Anfangswerte hinreichend nahe am Gleichgewicht gewählt werden. Es gibt attraktive, aber nicht stabile Gleichgewichte. Ein Beispiel ist eine Lösungstrajektorie, die sich für kleine Zeiten "weit" vom Gleichgewicht entfernt, aber für große Zeiten gegen das Gleichgewicht konvergiert.

Beispiel 4.42. (1) Die Lösung des Anfangswertproblems $x' = x(1-x)$, $x(0) = \xi$, lautet

$$x(t, \xi) = \frac{\xi}{\xi + (1-\xi)\exp(-t)}, \quad t \geq 0.$$

Es gibt zwei Gleichgewichte, $\bar{x}_1 = 0$ und $\bar{x}_2 = 1$. Das Gleichgewicht $\bar{x}_2 = 1$ ist stabil, denn ist $\varepsilon \in (0, 1)$ gegeben und wählen wir $\delta > 0$ mit $\delta/(1-\delta) < \varepsilon$, so folgt für alle $|\xi - 1| < \delta$ wegen $|\xi| \geq 1 - \delta$

$$|x(t, \xi) - 1| = \frac{|1-\xi|e^{-t}}{|\xi + (1-\xi)e^{-t}|} \leq \frac{|1-\xi|e^{-t}}{|\xi|} \leq \frac{\delta}{1-\delta} < \varepsilon.$$

Außerdem gilt $x(t, \xi) \rightarrow 1$ für $t \rightarrow \infty$. Also ist \bar{x}_2 attraktiv und damit asymptotisch stabil. Wir behaupten, dass das Gleichgewicht $\bar{x}_1 = 0$ instabil ist. Wir müssen zeigen:

$$\exists \varepsilon > 0 : \forall \delta > 0 : \exists \xi \in B_\delta(\bar{x}_1) : \exists t_0 \geq 0 : x(t_0, \xi) \notin B_\varepsilon(\bar{x}_1).$$

Sei $\varepsilon = 1/4$ und sei $\delta \in (0, 1)$ beliebig. Sei weiter $\xi = \delta/2$. Wir wählen $t_0 = -\ln \delta$. Dann folgt

$$|x(t_0, \xi)| = \frac{\delta}{\delta + (2-\delta)\exp(-t_0)} = \frac{\delta}{\delta + (2-\delta)\delta} \geq \frac{\delta}{\delta + 2\delta} = \frac{1}{3} > \varepsilon.$$

Also ist \bar{x}_1 instabil.

- (2) Die Schwingung eines Pendels wird durch die folgende Differentialgleichung beschrieben (siehe auch Beispiel 4.14):

$$x'' + \sin x = 0, \quad t \in \mathbb{R}, x \in [0, 2\pi).$$

Setzen wir $y_1 = x$ und $y_2 = x'$, so können wir diese Gleichung als System $y' = f(y)$ mit $y = (y_1, y_2)$ und

$$f(y) = \begin{pmatrix} y_2 \\ -\sin y_1 \end{pmatrix}$$

formulieren. Die Gleichgewichte sind $\bar{y}_1 = (0, 0)$ und $\bar{y}_2 = (\pi, 0)$. Intuitiv erwarten wir, dass \bar{y}_1 stabil ist, da für kleine Auslenkungen das Pendel um das Gleichgewicht schwingt. Das Gleichgewicht \bar{y}_2 wird nicht stabil sein, da dieser Punkt einem um π ausgelenkten Pendel entspricht, das bei einer kleinen Störung nach unten kippt und um $(0, 0)$ zu pendeln beginnt. Wir setzen diese Betrachtungen in Beispiel 4.47 fort.

- (3) Die Pendelgleichung kann vermittels $\sin x \approx x$ für kleine Auslenkungen x durch die Gleichung $x'' + x = 0$ approximiert werden. Das entsprechende System lautet $y' = f(y)$ mit

$$f(y) = \begin{pmatrix} y_2 \\ -y_1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

Das lineare System mit Anfangswert $y(0) = \xi = (\xi_1, \xi_2)^\top$ besitzt die Lösung

$$y(t, \xi) = \begin{pmatrix} x \\ x' \end{pmatrix} = \xi_1 \begin{pmatrix} \cos t \\ -\sin t \end{pmatrix} + \xi_2 \begin{pmatrix} \sin t \\ \cos t \end{pmatrix}.$$

Das einzige Gleichgewicht ist klarerweise $\bar{y} = (0, 0)^\top$. Wir wollen untersuchen, ob es stabil oder sogar asymptotisch stabil ist. Sei $\varepsilon > 0$. Mit $\delta = \varepsilon$ folgt für alle $|\xi| < \delta$ (wir wählen die euklidische Norm) wegen $\sin^2 t + \cos^2 t = 1$

$$|y(t, \xi)|^2 = (\xi_1 \cos t + \xi_2 \sin t)^2 + (-\xi_1 \sin t + \xi_2 \cos t)^2 = \xi_1^2 + \xi_2^2 = |\xi|^2 < \delta^2 = \varepsilon^2.$$

Also ist das Gleichgewicht stabil. Es ist *nicht* asymptotisch stabil, weil $y(t, \xi)$ nicht gegen null konvergiert. Das Beispiel zeigt, dass stabile Gleichgewichte nicht notwendigerweise attraktiv sein müssen.

- (4) Die Lorenz-Gleichungen

$$x' = \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = \begin{pmatrix} -\rho(x_1 - x_2) \\ rx_1 - x_2 - x_1x_3 \\ x_1x_2 - bx_3 \end{pmatrix} =: f(x)$$

mit positiven Parametern ρ , r und b beschreiben die Strömung einer Flüssigkeit zwischen zwei horizontalen Platten, von denen die untere beheizt wird (siehe Abbildung 4.9). Die Variablen haben folgende Bedeutung:

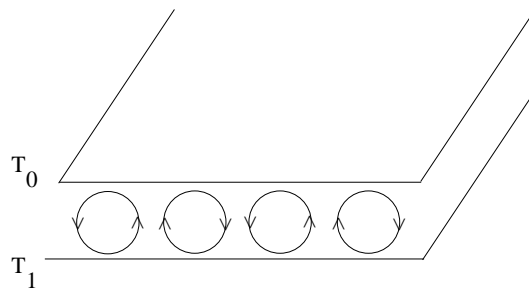


Abbildung 4.9: Die Flüssigkeit zwischen zwei parallelen Platten wird von unten beheizt. Für genügend große Temperaturdifferenzen $T_1 - T_0$ bildet sich eine periodische Strömung, die bei sehr großen Temperaturdifferenzen in eine chaotische Strömung übergeht. Dazwischen bilden sich sogenannte *Bénard-Zellen* aus.

- x_1 ist proportional zum Betrag der Konvektionsgeschwindigkeit,
- x_2 ist proportional zur Temperaturdifferenz zwischen aufsteigender und abfallender Strömung,
- x_3 ist proportional zur Abweichung vom linearen vertikalen Temperaturprofil.

Der Parameter ρ ist proportional zur Viskosität der Flüssigkeit, b ist ein Maß für die Geometrie der Bénard-Zellen, und r hängt von der Temperaturdifferenz $T_1 - T_0$ zwischen den beiden Platten ab.

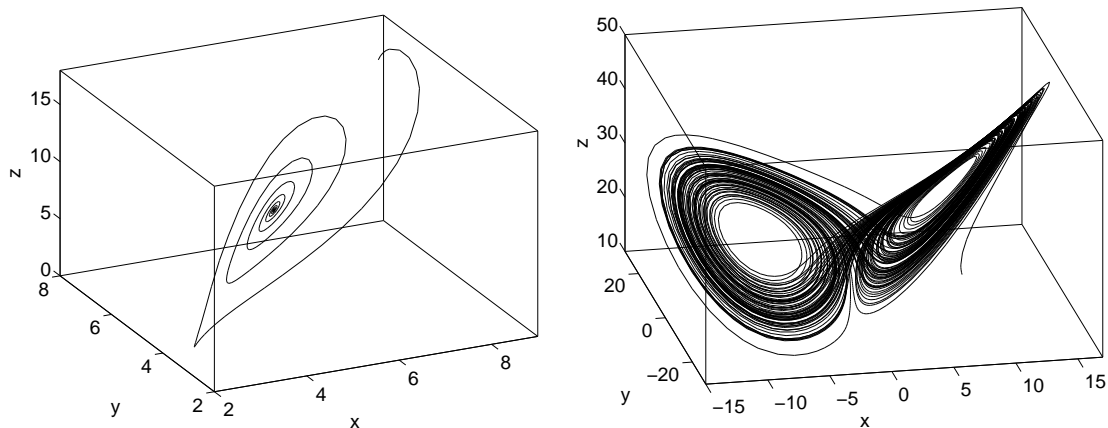


Abbildung 4.10: Trajektorien der Lösung der Lorenz-Gleichungen für $\rho = 10$, $b = 8/3$, $r = 10$ (links) und $r = 28$ (rechts).

Die Lorenz-Gleichungen haben folgende Gleichgewichte, falls $r > 1$: Setzen wir $f(x) = 0$, so folgt aus der ersten Gleichung $x_1 = x_2$, aus der zweiten Gleichung $(r - 1 - x_3)x_2 = 0$ und aus der dritten Gleichung $x_1^2 = bx_3$. Falls $x_2 = 0$, folgt auch $x_1 = x_3 = 0$. Wenn $r - 1 - x_3 =$

0, so erhalten wir $x_3 = r - 1$, $x_1 = x_2 = \pm \sqrt{b(r-1)}$. Dies ergibt drei Gleichgewichte:

$$\bar{x}_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad \bar{x}_2 = \begin{pmatrix} \sqrt{b(r-1)} \\ \sqrt{b(r-1)} \\ r-1 \end{pmatrix}, \quad \bar{x}_3 = \begin{pmatrix} -\sqrt{b(r-1)} \\ -\sqrt{b(r-1)} \\ r-1 \end{pmatrix}.$$

Die Dynamik der Lösungen der Lorenz-Gleichungen ist recht komplex: Für hinreichend kleine Parameter $r > 1$ konvergiert $x(t)$ gegen ein Gleichgewicht, während für genügend große $r > 1$ sich ein "chaotisches" Verhalten zeigt (siehe Abbildung 4.10). \square

Wir beginnen unsere Untersuchung der Stabilität mit homogenen linearen Systemen. Bei solchen Systemen ist $\bar{x} = 0$ immer ein Gleichgewicht.

Satz 4.43 (Stabilität homogener linearer Systeme). *Seien $A \in \mathbb{R}^{n \times n}$ und $\bar{x} = 0$ ein Gleichgewicht von*

$$x' = Ax, \quad t \geq 0. \quad (4.20)$$

Dann sind folgende Aussagen äquivalent:

- (i) Das Gleichgewicht $\bar{x} = 0$ ist asymptotisch stabil.
- (ii) Das Gleichgewicht $\bar{x} = 0$ ist attraktiv.
- (iii) Es gilt für alle Eigenwerte λ von A : $\operatorname{Re}(\lambda) < 0$.

Beweis. Aus der Definition folgt die Implikation (i) \Rightarrow (ii). Es gelte (ii), d.h., sei $\bar{x} = 0$ attraktiv. Angenommen, es gäbe einen Eigenwert $\lambda = \alpha + i\beta$ von A mit $\alpha \geq 0$. Sei v der zugehörige Eigenvektor. Dann ist

$$x(t) = \operatorname{Re}(e^{\lambda t} v) = e^{\alpha t} \operatorname{Re}(v), \quad t \geq 0,$$

wegen $x' = \operatorname{Re}(e^{\lambda t} \lambda v) = \operatorname{Re}(e^{\lambda t} Av) = Ax$ eine Lösung von (4.20), deren Grenzwert $t \rightarrow \infty$ offensichtlich *nicht* gleich null ist, Widerspruch. Dies beweist (ii) \Rightarrow (iii). Gilt (iii), so erhalten wir aus Satz 4.37 die Abschätzung

$$\|e^{At}\| \leq Ke^{(-\gamma+\varepsilon)t},$$

wobei $\gamma = -\max\{\operatorname{Re}(\lambda) : \lambda \text{ ist Eigenwert von } A\} > 0$. Wählen wir $0 < \varepsilon < \gamma$, so folgt für eine Lösung $x(t, \xi)$ von (4.20), $x(0) = \xi$,

$$\|x(t, \xi)\| = \|e^{At} \xi\| \leq Ke^{(-\gamma+\varepsilon)t} \|\xi\| \leq K \|\xi\|, \quad t \geq 0. \quad (4.21)$$

Wähle weiter $\delta = \varepsilon/K$. Wir erhalten für alle $\|\xi\| < \delta$

$$\|x(t, \xi)\| \leq K \|\xi\| < K\delta = \varepsilon.$$

Dies impliziert die Stabilität von $\bar{x} = 0$, also in Verbindung mit Satz 4.37 die Implikation (iii) \Rightarrow (i). \square

Der obige Beweis impliziert noch mehr: Ist nämlich das Gleichgewicht $\bar{x} = 0$ asymptotisch stabil, also $\operatorname{Re}(\lambda) < 0$, so folgt für $\varepsilon = \gamma/2$

$$\|x(t, \xi)\| \leq Ke^{-\gamma t/2} \|\xi\|, \quad t \geq 0.$$

Die Lösung konvergiert also gegen null exponentiell schnell. Diese Eigenschaft bezeichnen wir als *exponentielle Stabilität*. Für homogene lineare Systeme stimmt also die asymptotische und die exponentielle Stabilität überein.

Beispiel 4.44. In diesem Beispiel klassifizieren wir ebene Systeme der Form

$$x_1' = ax_1 + bx_2, \quad x_2' = cx_1 + dx_2$$

mit reellen Koeffizienten a, b, c und d . Wir betrachten nur das Gleichgewicht $\bar{x} = (0, 0)$. Die charakteristische Gleichung der entsprechenden Matrix ist

$$\lambda^2 - (a+d)\lambda + ad - bc = 0.$$

Die Nullstellen lauten mit den Abkürzungen $\Delta = ad - bc$ und $s = a + d$

$$\lambda_{1/2} = \frac{1}{2} \left(s \pm \sqrt{s^2 - 4\Delta} \right).$$

Wir unterscheiden verschiedene Fälle:

- $s^2 - 4\Delta > 0$: Beide Eigenwerte sind reell. Falls $\Delta < 0$, haben die beiden Eigenwerte verschiedene Vorzeichen. Man nennt das Gleichgewicht $\bar{x} = (0, 0)$ einen *Sattel*. Falls $\Delta > 0$, erhalten wir positive Eigenwerte, wenn $s > 0$, und negative Eigenwerte, wenn $s < 0$. Im ersteren Fall ist das Gleichgewicht instabil, im letzteren Fall stabil. Falls $\Delta = 0$, dann ist \bar{x} Teil einer Linie von Gleichgewichten, die instabil sind, wenn $s > 0$, und stabil, wenn $s < 0$.
- $s^2 - 4\Delta < 0$: Die Eigenwerte sind komplex. Falls $s > 0$, haben die Eigenwerte einen positiven Realteil. Die Lösungen sind von der Form $x_1(t) = c_1 e^{\alpha t} \cos(\beta t)$, $x_2(t) = c_2 e^{\alpha t} \sin(\beta t)$ für gewisse reelle Zahlen α und β . Es handelt sich um einen *instabilen Strudel*. Falls $s < 0$, so haben die Eigenwerte einen negativen Realteil und das Gleichgewicht ist ein *stabiler Strudel*. Falls $s = 0$, sind die Eigenwerte rein imaginär, die Lösungskurven sind Ellipsen, und man nennt das Gleichgewicht ein *Zentrum*.
- $s^2 - 4\Delta = 0$: Falls $s > 0$, erhalten wir einen positiven doppelten Eigenwert. Die Lösung hat in der Jordan-Normalform die Gestalt $x_1(t) = c_1 e^{\lambda t} + c_2 t e^{\lambda t}$, $x_2(t) = c_2 e^{\lambda t}$. Das Gleichgewicht ist instabil. Wenn $s < 0$, so ist der doppelte Eigenwert negativ, und das Gleichgewicht ist stabil.

Diese Szenarien sind veranschaulicht in Abbildung 4.11. □

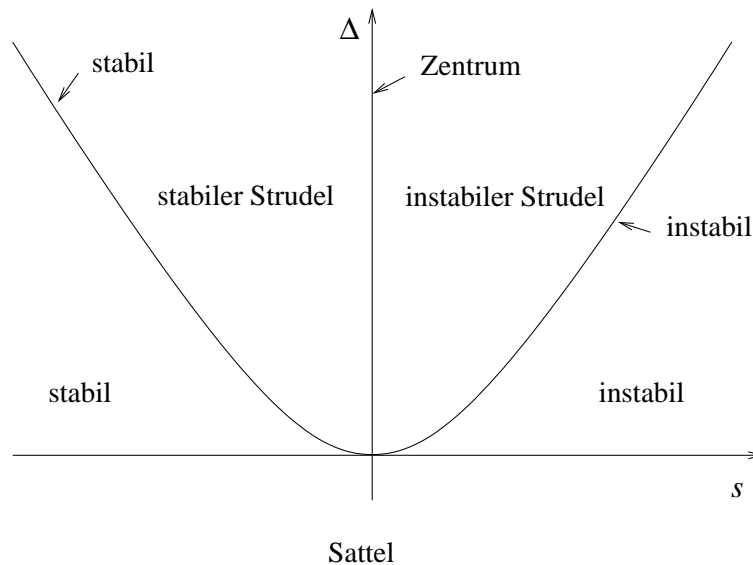


Abbildung 4.11: Klassifikation ebener linearer Systeme. Die Kurve beschreibt die Parabel $s^2 - 4\Delta = 0$.

4.6 Stabilität nichtlinearer autonomer Systeme

In diesem Abschnitt stellen wir zwei Techniken vor, um die Stabilität nichtlinearer Systeme zu untersuchen: die Eigenwertmethode, die auf den Ergebnissen des vorigen Abschnitts aufbaut, und die Methode von Lyapunov, die eine gewisse Monotonieeigenschaft ausnutzt.

Satz 4.45 (Stabilität linearer Systeme mit Störung). Sei $A \in \mathbb{R}^{n \times n}$ eine Matrix, deren Eigenwerte λ in der linken Halbebene $\{\operatorname{Re}(\lambda) < 0\}$ liegen, und sei $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine lokal lipschitzstetige Funktion mit

$$\lim_{|x| \rightarrow 0} \frac{|g(x)|}{|x|} = 0.$$

Dann ist das Gleichgewicht $\bar{x} = 0$ der Differentialgleichung

$$x' = Ax + g(x), \quad t \geq 0, \quad (4.22)$$

asymptotisch stabil.

Beweis. Aus Satz 4.43 und der Bemerkung nach dem Beweis dieses Satzes folgt, dass das Gleichgewicht der linearen Gleichung exponentiell stabil ist, d.h., es existieren $K > 0$ und $\lambda > 0$, so dass für alle $t \geq 0$

$$\|e^{At}\| \leq Ke^{-\lambda t}.$$

Die Voraussetzung an g impliziert, dass ein $\delta > 0$ existiert, so dass für alle $|x| < \delta$

$$|g(x)| \leq K^{-1}\lambda|x|.$$

Da g lokal Lipschitzstetig ist, existiert nach dem Satz von Picard-Lindelöf eine eindeutige Lösung $x(\cdot, \xi) : [0, T) \rightarrow \mathbb{R}^n$ von (4.22) und $x(0) = \xi$ mit $|\xi| < \delta/K$. Für diese Lösung gilt nach Satz 4.38 (definiere $b(t) = g(x(t))$) für alle $0 \leq t < T$

$$\begin{aligned} |x(t, \xi)| &= \left| e^{At} \xi + \int_0^t e^{A(t-s)} g(x(s, \xi)) ds \right| \leq \underbrace{\|e^{At}\|}_{\leq Ke^{-2\lambda t}} |\xi| + \int_0^t \underbrace{\|e^{A(t-s)}\|}_{\leq Ke^{-2\lambda(t-s)}} \underbrace{|g(x(s, \xi))|}_{\leq K^{-1}\lambda |x(s, \xi)|} ds \\ &\leq Ke^{-2\lambda t} |\xi| + \lambda \int_0^t e^{-2\lambda(t-s)} |x(s, \xi)| ds. \end{aligned}$$

Setzen wir $y(t) = e^{2\lambda t} |x(t, \xi)|$, ergibt dies

$$y(t) \leq K|\xi| + \lambda \int_0^t y(s) ds, \quad 0 \leq t < T.$$

Nach dem Lemma 4.17 von Gronwall (mit $h(t) = \lambda$ und $a = 0$) folgt

$$y(t) \leq K|\xi| e^{\lambda t} \quad \text{für } 0 \leq t < T,$$

also für die ursprüngliche Variable $x(t, \xi)$

$$|x(t, \xi)| \leq K|\xi| e^{-\lambda t}, \quad 0 \leq t < T.$$

Daraus schließen wir zum einen, dass die Lösung auf $[0, \infty)$ fortgesetzt werden kann; die Abschätzung gilt dann auch für $T = \infty$. Zum anderen folgern wir die Stabilität und Attraktivität aus dieser Ungleichung (siehe den Beweis von Satz 4.43 für einen analogen Beweis). \square

Der obige Satz kann unter bestimmten Bedingungen auf allgemeine nichtlineare autonome Systeme angewendet werden.

Korollar 4.46 (Stabilität nichtlinearer Systeme). *Seien $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ stetig differenzierbar und $\bar{x} \in \mathbb{R}^n$ mit $f(\bar{x}) = 0$. Ist \bar{x} asymptotisch stabil für die linearisierte Gleichung*

$$x' = f'(\bar{x})x, \quad t \geq 0,$$

dann ist \bar{x} auch asymptotisch stabil für die nichtlineare Gleichung

$$x' = f(x), \quad t \geq 0.$$

Beweis. Ohne Einschränkung können wir $\bar{x} = 0$ setzen. Da f stetig differenzierbar ist, ergibt eine Taylor-Entwicklung $f(x) = Ax + g(x)$ mit $A = f'(0)$. Die Funktion g erfüllt die Eigenschaften $g(0) = 0$ und $g(x) = o(|x|)$, also $|g(x)|/|x| \rightarrow 0$ für $|x| \rightarrow 0$. Der vorige Satz ist also anwendbar, und die Behauptung folgt. \square

Der Satz sagt aus, dass die asymptotische Stabilität des linearisierten Systems die asymptotische Stabilität des nichtlinearen Systems impliziert. Dies gilt *nicht* für die Stabilität. Die Linearisierung der Gleichung $x' = x^3$, $x(0) = x_0$, um das Gleichgewicht $\bar{x} = 0$ liefert die Gleichung $x' = 3\bar{x}^2 = 0$, $x(0) = x_0$, mit der stabilen konstanten Lösung $x(t) = x_0$. Die nichtlineare Gleichung besitzt jedoch die Lösung $x(t) = x_0(1 - 2x_0^2 t)^{-1/2}$, die für $t \rightarrow 1/(2x_0^2)$ unbeschränkt wird. Wir folgern: Aus der reinen Stabilität der linearisierten Gleichung können wir *nicht* auf die Stabilität der nichtlinearen Gleichung schließen.

Beispiel 4.47. (1) Wir erläutern die obigen Ergebnisse an der nichtlinearen Pendelgleichung mit Dämpfung $d > 0$:

$$x'' + 2dx' + \sin x = 0, \quad t \geq 0, \quad x \in [0, 2\pi).$$

Das entsprechende System lautet $y' = f(y)$, $y = (y_1, y_2) \in [0, 2\pi) \times \mathbb{R}$ mit

$$f(y) = \begin{pmatrix} y_2 \\ -2dy_2 - \sin y_1 \end{pmatrix}.$$

Es gibt zwei Gleichgewichte $\bar{y}_1 = (0, 0)$ und $\bar{y}_2 = (\pi, 0)$. Die Linearisierung in \bar{y}_1 lautet

$$A = f'(0, 0) = \begin{pmatrix} 0 & 1 \\ -1 & -2d \end{pmatrix}.$$

Die Realteile der Eigenwerte $\lambda_{1/2} = -d \pm \sqrt{d^2 - 1}$ sind negativ, also ist \bar{y}_1 asymptotisch stabil. Die Linearisierung im zweiten Gleichgewicht \bar{y}_2 führt auf die Matrix

$$A = f'(\pi, 0) = \begin{pmatrix} 0 & 1 \\ 1 & -2d \end{pmatrix}$$

mit den Eigenwerten $\lambda_{1/2} = -d \pm \sqrt{d^2 + 1}$. Wegen $\lambda_1 > 0$ ist der obige Satz nicht anwendbar. In Beispiel 4.42 haben wir bereits begründet, warum dieses Gleichgewicht aus physikalischer Sicht nicht stabil sein kann.

(2) Wir betrachten wie in Beispiel 4.42 die Lorenz-Gleichungen

$$x' = f(x) = \begin{pmatrix} -\rho(x_1 - x_2) \\ rx_1 - x_2 - x_1x_3 \\ x_1x_2 - bx_3 \end{pmatrix}.$$

Sei $r > 1$. Wir haben gezeigt, dass dieses System drei Gleichgewichte besitzt. Die Linearisierung im ersten Gleichgewicht $\bar{y} = (0, 0, 0)$ lautet

$$A = f'(\bar{y}) = \begin{pmatrix} -\rho & \rho & 0 \\ r & -1 & 0 \\ 0 & 0 & -b \end{pmatrix}.$$

Die charakteristische Gleichung

$$(\lambda + b)(\lambda^2 + (\rho + 1)\lambda - \rho(r - 1)) = 0$$

besitzt zwei negative Nullstellen und eine positive Nullstelle:

$$\lambda_1 = -b, \quad \lambda_{2/3} = -\frac{1}{2}(\rho + 1) \pm \sqrt{\frac{1}{4}(\rho + 1)^2 + \rho(r - 1)},$$

also ist der obige Satz nicht anwendbar. Die Linearisierung im zweiten und dritten Gleichgewicht ist

$$A = f'(\bar{y}) = \begin{pmatrix} -\rho & \rho & 0 \\ 1 & -1 & \mp \sqrt{b(r-1)} \\ \pm \sqrt{b(r-1)} & \pm \sqrt{b(r-1)} & -b \end{pmatrix}$$

mit der charakteristischen Gleichung

$$\lambda^3 + (\rho + b + 1)\lambda^2 + b(\rho + r)\lambda + 2b\rho(r - 1) = 0. \quad (4.23)$$

Zwar kann man die Nullstellen dieses Polynoms explizit ermitteln, aber die entsprechenden Formeln sind sehr kompliziert. Uns genügt die Information, unter welchen Bedingungen die Nullstellen einen negativen Realteil besitzen. Hierfür benutzen wir das folgende Kriterium:

Routh-Hurwitz-Kriterium: Das Polynom $p(x) = x^3 + ax^2 + bx + c$ mit reellen Koeffizienten besitzt genau dann nur Nullstellen mit negativem Realteil, wenn (i) alle Koeffizienten positiv sind und (ii) $ab > c$.

Der Beweis dieses Kriteriums ist nicht schwer: Ein Polynom dritter Ordnung hat genau drei komplexe Nullstellen (und zwar drei reelle oder zwei konjugiert komplexe und eine reelle), es lässt sich also zerlegen in

$$p(x) = (x + \alpha)(x^2 + \beta x + \gamma) = x^3 + (\alpha + \beta)x^2 + (\alpha\beta + \gamma)x + \alpha\gamma.$$

mit $\alpha, \beta, \gamma \in \mathbb{R}$. Ein Koeffizientenvergleich ergibt

$$a = \alpha + \beta, \quad b = \alpha\beta + \gamma, \quad c = \alpha\gamma.$$

Es gelte $a, b, c > 0$ und $ab - c > 0$. Wegen

$$0 < ab - c = \alpha^2\beta + \alpha\beta^2 + \beta\gamma = \beta(\alpha^2 + b)$$

folgt $\beta > 0$. Daher ist $\alpha b = \alpha\gamma + \alpha^2\beta = c + \alpha^2\beta > 0$ und damit auch $\alpha > 0$. Schließlich folgt $\gamma = c/\alpha > 0$. Die Nullstellen von p lauten $x_1 = -\alpha < 0$ und $x_{2/3} = (-\beta \pm \sqrt{\beta^2 - 4\gamma})/2$, und wegen $\beta, \gamma > 0$ folgt $\operatorname{Re}(x_{2/3}) < 0$. Seien umgekehrt die Realteile der

Nullstellen von $p(x)$ negativ. Dies bedeutet $\alpha > 0$; wegen $x_{2/3}$ gilt $\beta > 0$ (ansonsten wäre $\operatorname{Re}(x_{2/3}) \geq 0$); und es ist $\gamma > 0$ (ansonsten wäre $x_2 \geq 0$). Daraus folgt sofort $a, b, c > 0$ sowie $ab - c = \beta(\alpha^2 + b) > 0$.

Wir kommen zur charakteristischen Gleichung (4.23) zurück. Wegen $r > 1$ sind alle Koeffizienten positiv. Damit die Realteile der Nullstellen negativ sind, muss nach dem Routh-Horwith-Kriterium

$$0 < (\rho + b + 1)b(\rho + r) - 2b\rho(r - 1) = b((1 + b - \rho)r + \rho(\rho + b + 3))$$

gelten. Aus physikalischen Gründen ist $b > 0$. Falls $\rho \leq b + 1$, sind alle $r > 0$ erlaubt. Anderenfalls erhalten wir eine Schranke an r :

$$\rho > b + 1, \quad r < r_c := \frac{\rho(\rho + b + 3)}{\rho - b - 1}.$$

Wenn die Parameter diese Werte annehmen, sind die zweiten und dritten Gleichgewichte asymptotisch stabil. Im Beispiel 4.42 haben wir die Werte $\rho = 10$ und $b = 8/3$ verwendet. Wir erhalten also asymptotische Stabilität für alle $r < r_c = 470/19 = 24,7368\dots$. Diese Situation entspricht der linken Abbildung 4.10. Die obigen Resultate erlauben keine Aussage für die Parameterwahl in der rechten Abbildung. \square

Es gibt nichtlineare Systeme, deren Eigenwerte der linearisierten Matrix am Gleichgewicht einen *nichtnegativen* Realteil haben und deren Gleichgewicht dennoch stabil (sogar asymptotisch stabil) sein kann. Eine alternative Methode zum Nachweis der Stabilität wurde von Lyapunov entwickelt, die wir im Folgenden vorstellen. Zunächst ein illustratives Beispiel.

Beispiel 4.48. Wir betrachten das System

$$x' = f(x) = \begin{pmatrix} -x_1^3 - 2x_2 \\ x_1^3 - x_2 \end{pmatrix}.$$

Das einzige Gleichgewicht ist $\bar{x} = (0, 0)$. Die linearisierte Matrix lautet

$$f'(0) = \begin{pmatrix} -3x_1^2 & -2 \\ 3x_1^2 & -1 \end{pmatrix} \Big|_{x=0} = \begin{pmatrix} 0 & -2 \\ 0 & -1 \end{pmatrix}$$

mit Eigenwerten $\lambda = 0$ und $\lambda = -1$. Die obige Theorie ist also nicht anwendbar. Um dennoch Stabilität nachzuweisen, lautet die Idee, das Funktional $V(x) = x_1^4 + 4x_2^2$ entlang einer Trajektorie $x(t)$ nach t abzuleiten:

$$\begin{aligned} \frac{d}{dt}V(x(t)) &= 4x_1^3(t)x_1'(t) + 8x_2(t)x_2'(t) \\ &= 4x_1^3(t)(-x_1(t)^3 - 2x_2(t)) + 8x_2(t)(x_1(t)^3 - x_2(t)) = -4x_1^6(t) - 8x_2^2(t) \leq 0. \end{aligned}$$

Daraus folgt

$$0 \leq V(x_1(t), x_2(t)) = V(x_1(0), x_2(0)) - \int_0^t (4x_1^6(s) + 8x_2^2(s)) ds. \quad (4.24)$$

Wählen wir die Anfangswerte in einer kleinen Kugel um das Gleichgewicht, so bleibt die Lösung in dieser Kugel. Insbesondere ist das Gleichgewicht stabil. Wir behaupten, dass das Gleichgewicht sogar attraktiv ist, d.h., dass $x(t) \rightarrow 0$ für $t \rightarrow \infty$ gilt. Dies können wir wie folgt einsehen: Die Funktion $t \mapsto V(x_1(t), x_2(t))$ ist monoton fallend und durch null nach unten beschränkt. Daher existiert der Grenzwert $a = \lim_{t \rightarrow \infty} V(x_1(t), x_2(t))$. Falls $a > 0$, erhalten wir wegen der Monotonie von $t \mapsto V(x_1(t), x_2(t))$

$$0 < a \leq V(x_1(t), x_2(t)) \leq V(x_1(0), x_2(0)), \quad t \geq 0.$$

Sei $m = \min\{4x_1^6 + 8x_2^2 : x_1, x_2 \in \mathbb{R}, a \leq V(x_1, x_2) \leq V(x_1(0), x_2(0))\}$. Es ist $m > 0$, da andernfalls das Minimum an $x_1 = x_2 = 0$ angenommen wird, was zu dem Widerspruch $0 < a \leq V(0, 0) = 0$ führt. Wir erhalten aus (4.24)

$$0 \leq V(x_1(t), x_2(t)) \leq V(x_1(0), x_2(0)) - m \int_0^t dt = V(x_1(0), x_2(0)) - mt.$$

Für hinreichend großes $t > 0$ folgt $0 \leq V(x_1(t), x_2(t)) < 0$, Widerspruch. Daher muss $a = 0$ gelten. Wir haben gezeigt, dass das Gleichgewicht asymptotisch stabil ist. \square

Wesentlich im obigen Beispiel ist die Existenz eines Funktional, das gewisse Monotonieeigenschaften besitzt. Dies wird in der folgenden Definition präzisiert. Betrachte hierfür weiterhin die Differentialgleichung $x' = f(x)$.

Definition 4.49 (Lyapunov-Funktion). Seien $\bar{x} \in \mathbb{R}^n$ und $U \subset \mathbb{R}^n$ eine Umgebung von \bar{x} . Die Funktion $V \in C^1(U)$ heißt Lyapunov-Funktion, wenn

- (i) $V(\bar{x}) = 0$ und $V(x) > 0$ für alle $x \in U$, $x \neq \bar{x}$;
- (ii) $\nabla V(x) \cdot f(x) \leq 0$ für alle $x \in U$.

Die letzte Bedingung impliziert entlang einer Lösungstrajektorie:

$$\frac{dV}{dt}(x(t)) = \nabla V(x(t)) \cdot x'(t) = \nabla V(x(t)) \cdot f(x(t)) \leq 0,$$

d.h., die Lyapunov-Funktion ist monoton fallend entlang der Trajektorie.

Beispiel 4.50. Die nichtlineare Pendelgleichung mit Dämpfung $x'' + 2dx' + \sin x = 0$ besitzt die Lyapunov-Funktion $V(x, x') = (x')^2/2 - \cos x + 1$, $(x, x') \in U = [-\pi, \pi] \times \mathbb{R}$. Schreiben wir nämlich die Differentialgleichung als ein System $y' = f(y)$ mit $y = (x, x')$ und $f(y) = (x', -2dx' - \sin x)$, so folgt

$$\nabla V(y) \cdot f(y) = \begin{pmatrix} \sin x \\ x' \end{pmatrix} \cdot \begin{pmatrix} x' \\ -2dx' - \sin x \end{pmatrix} = -2d(x')^2 \leq 0.$$

Außerdem gilt $V(0, 0) = 0$ und $V(x, x') > 0$ für alle $(x, x') \in U \setminus \{(0, 0)\}$. Physikalisch können wir V als die Gesamtenergie interpretieren, und zwar als die Summe aus kinetischer Energie $(x')^2/2$ und potentieller Energie $1 - \cos x$. \square

Die Existenz einer Lyapunov-Funktion impliziert bereits die Stabilität des Gleichgewichts und, unter geeigneten Voraussetzungen, dessen Attraktivität.

Satz 4.51 (Stabilitätssatz von Lyapunov). *Seien \bar{x} ein Gleichgewicht von $x' = f(x)$, f eine stetige Funktion, $U \subset \mathbb{R}^n$ eine Umgebung von \bar{x} und $V \in C^1(U)$ eine Lyapunov-Funktion. Dann gilt:*

- (i) *Das Gleichgewicht \bar{x} ist stabil.*
- (ii) *Das Gleichgewicht \bar{x} ist attraktiv genau dann, wenn eine Umgebung W von \bar{x} existiert, so dass \bar{x} die einzige Lösung von $x' = f(x)$ ist mit $x(0) \in W$ und $(dV/dt)(x(t)) = 0$ für $t \in [0, \infty)$.*

Beweis. Wir können wieder ohne Einschränkung annehmen, dass $\bar{x} = 0$ (ansonsten definiere $y = x - \bar{x}$ und $\tilde{f}(y) = f(y + \bar{x})$).

(i) Sei $r > 0$ mit $B_{2r}(0) \subset U$ und definiere

$$\alpha = \min_{|x|=r} V(x) > 0, \quad U_\alpha = \{x \in U : V(x) < \alpha\} \cap \overline{B_r(0)}$$

(siehe Abbildung 4.12). Nach Definition der Lyapunov-Funktion ist $\alpha > 0$, und $U_\alpha \neq \emptyset$ (denn $0 \in U_\alpha$) ist eine Umgebung von null. Sei nun $x : [0, t^*) \rightarrow \mathbb{R}^n$ eine Lösung von $x' = f(x)$, $x(0) = \xi \in U_\alpha$. Dann folgt $V(\xi) < \alpha$ nach Konstruktion von U_α und $(dV/dt)(x(t)) \leq 0$ für $t \in [0, t^*)$ nach Definition von V .

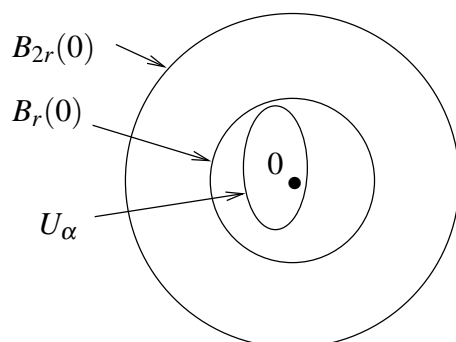


Abbildung 4.12: Illustration für den Beweis des Stabilitätssatzes.

Wir behaupten, dass $x(t)$ für alle $t \in [0, t^*)$ in der Kugel $B_r(0)$ bleibt. Gäbe es nämlich ein $t_1 \in (0, t^*)$ mit $|x(t_1)| = r$, führt

$$V(\xi) < \underbrace{\alpha \leq V(x(t_1))}_{\text{Definition von } \alpha} \quad \text{und} \quad \underbrace{V(x(t_1)) \leq V(x(0)) = V(\xi)}_{\text{Monotonie von } V}$$

auf einen Widerspruch. Die Lösung $x(t)$ bleibt also beschränkt und existiert daher für alle $t \geq 0$. Das Lösungsintervall lautet also $I = [0, \infty)$. Die eben bewiesene Aussage impliziert auch die Stabilität von $\bar{x} = 0$, da aus $|x(0)| = |\xi| \leq r$ die Abschätzung $|x(t)| < r$ folgt.

(ii) Sei $\bar{x} = 0$ attraktiv. Dann existiert eine Umgebung W_1 von \bar{x} , so dass für alle $\xi \in W_1$ gilt $\lim_{t \rightarrow \infty} x(t, \xi) = \bar{x} = 0$. Die Menge $W = U_\alpha \cap W_1$ ist auch eine Umgebung von $\bar{x} = 0$, denn $0 \in U_\alpha$. Sei $x : [0, \infty) \rightarrow W$ eine Lösung mit $x(0) = \xi \in W$ und $(dV/dt)(x(t)) = 0$ für alle $t \in [0, \infty)$. Insbesondere ist $V(x(t)) = V(x(0)) = V(\xi)$. Wir müssen nachweisen, dass $x = 0$. Teil (i) des Beweises zeigt, dass der Wertebereich von x wirklich in W liegt. Dann folgt wegen der Stetigkeit von V

$$V(\xi) = \lim_{t \rightarrow \infty} V(x(t)) = V\left(\lim_{t \rightarrow \infty} x(t)\right) = V(0) = 0.$$

Dies impliziert $\xi = 0$. Aus $V(x(t)) = V(\xi) = 0$ folgt nach Definition von V , dass $x = 0$, was zu zeigen war.

Um die andere Richtung zu beweisen, sei W gegeben und setze $W_2 = U_\alpha \cap W$. Sei x eine Lösung von $x' = f(x)$ mit $x(0) \in W_2$. Teil (i) des Beweises zeigt, dass $x : [0, \infty) \rightarrow B_r(0)$ gilt und $x(t)$ insbesondere beschränkt ist. Daher enthält jede Folge $(x(t_n))$ mit $t_n \rightarrow \infty$ ($n \rightarrow \infty$) eine konvergente Teilfolge $(x(t_{n_k}))$, d.h. $\lim_{k \rightarrow \infty} x(t_{n_k}) = x_1$. Wir müssen zeigen, dass $x_1 = 0$.

Angenommen, es ist $x_1 \neq 0$. Aus den Eigenschaften der Lyapunov-Funktion folgt, dass die Folge $(V(x(t_{n_k})))$ monoton fällt und $\lim_{k \rightarrow \infty} V(x(t_{n_k})) = V(x_1)$ gilt. Insbesondere ist $V(x_1) < \alpha$. Sei nun y die Lösung von $x' = f(x)$, $x(0) = x_1$. Aus Teil (i) des Beweises folgt $y : [0, \infty) \rightarrow B_r(0)$ und $V(y(t)) \leq V(x_1)$ für $t \in [0, \infty)$. Wegen $x_1 \neq 0$ und $V(0) = 0$ existiert ein $\tau > 0$ mit $V(y(\tau)) < V(x_1)$. Aus der Autonomie der Differentialgleichung schließen wir, dass die Funktion $x_k : [0, \infty) \rightarrow \mathbb{R}^n$ mit $x_k(t) := x(t + t_{n_k})$ eine Lösung von $x' = f(x)$ mit Anfangswert $x_k(0) = x(t_{n_k})$ ist. Mit dem Satz 4.18 über die stetige Abhängigkeit von den Anfangswerten schließen wir aus $\lim_{k \rightarrow \infty} x(t_{n_k}) = x_1$ und damit $\lim_{k \rightarrow \infty} (x_k(0) - y(0)) = \lim_{k \rightarrow \infty} (x(t_{n_k}) - x_1) = 0$:

$$\lim_{k \rightarrow \infty} (x(t + t_{n_k}) - y(t)) = 0$$

oder

$$\lim_{k \rightarrow \infty} x_k(t) = \lim_{k \rightarrow \infty} x(t + t_{n_k}) = y(t), \quad t \in [0, \tau].$$

Wir erhalten

$$\lim_{k \rightarrow \infty} V(x(\tau + t_{n_k})) = V(y(\tau)) < V(x_1).$$

Andererseits existiert wegen der Monotonie von $(V(x(t_{n_k})))$ und der Eigenschaft $x(t_{n_k}) \rightarrow x_1$ ($k \rightarrow \infty$) zu jedem $k \in \mathbb{N}$ ein $m \in \mathbb{N}$ mit

$$V(x(\tau + t_{n_k})) \geq V(x(t_{n_m})) \geq V(x_1),$$

Widerspruch. Also ist $x_1 = 0$. Die obige Argumentation gilt für jede Teilfolge von (t_n) , aus der wir eine konvergente Teilfolge $(t_{n'})$ extrahieren können mit $\lim_{n' \rightarrow \infty} x(t_{n'}) = 0$. Also konvergiert die gesamte Folge (t_n) , und weil (t_n) beliebig ist, folgt $\lim_{t \rightarrow \infty} x(t) = 0$. Daher ist das Gleichgewicht $\bar{x} = 0$ attraktiv, und der Satz ist bewiesen. \square

Wir können den Stabilitätssatz verschärfen, wenn wir eine zusätzliche Voraussetzung an die Lapunov-Funktion machen.

Korollar 4.52. Seien \bar{x} ein Gleichgewicht von $x' = f(x)$, $U \subset \mathbb{R}^n$ eine Umgebung von \bar{x} und $V \in C^1(U)$ eine Lyapunov-Funktion mit

$$\nabla V(x) \cdot f(x) < 0 \quad \text{für alle } x \in U \setminus \{\bar{x}\}.$$

Dann ist \bar{x} asymptotisch stabil.

Beweis. Seien $\bar{x} = 0$ und $x: [0, \infty) \rightarrow \mathbb{R}^n$ eine Lösung von $x' = f(x)$, $x(0) \neq \bar{x}$. Nach Voraussetzung ist

$$\frac{dV}{dt}(x(t)) = \nabla V(x(t)) \cdot x'(t) < 0,$$

sofern $x(t) \neq \bar{x}$. Die einzige Lösung x , die $(dV/dt)(x(t)) = 0$ erfüllt, ist also $x \equiv \bar{x} = 0$. Nach Satz 4.51 ist $\bar{x} = 0$ stabil und attraktiv, also asymptotisch stabil. \square

Beispiel 4.53. (1) In Beispiel 4.48 haben wir das System

$$x' = f(x) = \begin{pmatrix} -x_1^3 - 2x_2 \\ x_1^3 - x_2 \end{pmatrix}$$

untersucht und bereits die Funktion $V(x) = x_1^4 + 4x_2^2$ eingeführt. Die Abschätzung

$$\frac{dV}{dt}(x(t)) = -4(x_1^6(t) + 2x_2^2(t))$$

zeigt, dass V eine Lyapunov-Funktion ist, die die Voraussetzungen von Korollar 4.52 erfüllt ($\bar{x} = (0, 0)$ ist das einzige Gleichgewicht). Also ist $\bar{x} = 0$ asymptotisch stabil. Dies liefert einen alternativen Beweis der asymptotischen Stabilität.

(2) In Abschnitt 3.2 haben wir gezeigt, dass der Strom durch einen RLC-Schaltkreis ohne anregende Spannung der linearen Differentialgleichung

$$I'' + 2dI' + \omega^2 I = 0, \quad t \geq 0, \quad I(0) = I_0, \quad I'(0) = I_1,$$

genügt. Setzen wir $x_1 = I$ und $x_2 = I'$, entspricht dies dem System $x = f(x)$, $x(0) = (I_0, I_1)$ mit

$$f(x) = \begin{pmatrix} 0 & 1 \\ -\omega^2 & -2d \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Natürlich können wir für dieses lineare System die Stabilitätsaussagen aus dem vorigen Abschnitt verwenden. Wir verwenden hier zur Illustration die Methode von Lyapunov. Eine Lyapunov-Funktion ist gegeben durch

$$V(x) = \omega^2 x_1^2 + \frac{1}{2} x_2^2, \quad x \in \mathbb{R}^2,$$

denn

$$\nabla V(x) \cdot f(x) = \begin{pmatrix} 2\omega^2 x_1 \\ x_2 \end{pmatrix} \begin{pmatrix} x_2 \\ -\omega^2 x_1 - 2d x_2 \end{pmatrix} = -d x_2^2.$$

Die Voraussetzungen von Korollar 4.52 sind zwar nicht erfüllt (denn $\nabla V(1,0) \cdot f(1,0) = 0$), aber Satz 4.51 ist anwendbar, denn aus $0 = \nabla V(x(t)) \cdot f(x(t)) = -dx_2(t)^2$ folgt $x_2(t) = 0$ und wegen $\omega^2 x_1 = -x_2' - 2dx_2 = 0$ auch $x_1(t) = 0$. Wir folgern aus Satz 4.51, dass $\bar{x} = 0$ ein asymptotisch stabiles Gleichgewicht ist. \square

Die Hauptschwierigkeit bei der Methode von Lyapunov ist die Bestimmung einer Lyapunov-Funktion V . Hierfür gibt es keine allgemeinen Regeln; meistens benötigt man Hintergrundwissen über die hinter der Differentialgleichung liegenden Anwendung. In vielen Fällen ist V etwa durch die Energie des Systems gegeben. Ein Beispiel ist die Pendelgleichung aus Beispiel 4.50, bei der V als die Summe aus der kinetischen und potentiellen Energie angesetzt werden kann. Ein anderes Beispiel sind *Gradientensysteme*

$$x' = f(x) := -\nabla V(x), \quad t \in \mathbb{R},$$

wobei $V \in C^1(\mathbb{R}^n)$. Wir setzen voraus: $V(\bar{x}) = 0$, $V(x) > 0$ für $x \neq \bar{x}$ und $\nabla V(\bar{x}) = 0$. In diesem Fall folgt

$$\nabla V(x) \cdot f(x) = -|\nabla V(x)|^2 \leq 0,$$

also ist V eine Lyapunov-Funktion, und wir folgern aus Satz 4.51, dass \bar{x} ein stabiles Gleichgewicht ist.

4.7 Räuber-Beute-Modelle

In diesem Abschnitt untersuchen wir das qualitative Verhalten und die Stabilität der Gleichgewichte von Räuber-Beute-Modellen. Schule

Ein Räuber-Beute-Modell mit unbegrenztem Wachstum. Wir beginnen mit den Lotka-Volterra-Gleichungen

$$\begin{aligned} x_1' &= (a - bx_2)x_1, \\ x_2' &= (-c + dx_1)x_2, \quad t \geq 0. \end{aligned}$$

Wir erinnern, dass x_1 die Beutepopulation und x_2 die Räuberpopulation beschreibt und dass $a, b, c, d > 0$ Konstanten sind. Wir formulieren die Gleichungen als das System $x' = f(x)$ mit

$$f(x) = \begin{pmatrix} (a - bx_2)x_1 \\ (-c + dx_1)x_2 \end{pmatrix},$$

wobei $x = (x_1, x_2)$. Es gibt zwei Gleichgewichte: $\bar{x} = (0, 0)$ und $\bar{x} = (c/d, a/b)$. Um die Stabilität der Gleichgewichte zu überprüfen, betrachten wir die Linearisierung:

$$f'(0,0) = \begin{pmatrix} a - bx_2 & -bx_1 \\ dx_2 & -c + dx_1 \end{pmatrix} \Big|_{(x_1, x_2) = (0,0)} = \begin{pmatrix} a & 0 \\ 0 & -c \end{pmatrix}, \quad f'\left(\frac{c}{d}, \frac{a}{b}\right) = \begin{pmatrix} 0 & -\frac{bc}{d} \\ \frac{ad}{b} & 0 \end{pmatrix}.$$

Die Eigenwerte der ersten Matrix sind $\lambda_1 = a > 0$ und $\lambda_2 = -c < 0$. Mit der Notation von Beispiel 4.44 ist $s = a - c$ (die Spur der Matrix) und $\Delta = -ac$ (die Determinante der Matrix),

also $s^2 - 4\Delta = (a + c)^2 > 0$. Es liegt also ein Sattel vor. Was bedeutet dies für die Populationen? Starten wir mit dem Anfangswert $(x_1, 0)$ (keine Räuber), so ist $x(t) = (x_1 e^{at}, 0)$ die Lösung des Systems; die Beutepopulation vermehrt sich exponentiell. Starten wir andererseits mit dem Anfangswert $(0, x_2)$ (keine Beute), so ist $x(t) = (0, x_2 e^{-ct})$ eine Lösung, die exponentiell schnell gegen das Gleichgewicht konvergiert. Die Eigenwerte der zweiten Matrix sind rein imaginär. Das zweite Gleichgewicht ist daher stabil (die Lösungen "oszillieren" um das Gleichgewicht), aber wir haben keine exponentielle Stabilität.

Da in beiden Fällen keine exponentielle Stabilität vorliegt, können wir die Stabilitätssätze, die von der (exponentiellen) Stabilität des linearen Systems auf die (exponentielle) Stabilität des nichtlinearen Systems schließen, nicht anwenden. Mehr Informationen können wir durch die Methode von Lyapunov gewinnen. Um eine Lyapunov-Funktion zu bestimmen, rechnen wir zunächst formal:

$$\frac{dx_1}{dx_2} = \frac{dx_1/dt}{dx_2/dt} = \frac{(a - bx_2)x_1}{(-c + dx_1)x_2}.$$

Behandeln wir den Differentialquotienten wie einen Bruch, erhalten wir (formal)

$$\int_{c/d}^{x_1(t)} \frac{-c + dx_1}{x_1} dx_1 = \int_{a/b}^{x_2(t)} \frac{a - bx_2}{x_2} dx_2,$$

also nach Integration

$$-c \ln x_1(t) + dx_1(t) + c \ln \frac{c}{d} - c = a \ln x_2 - bx_2 - a \ln \frac{a}{b} + a.$$

Wir setzen daher an:

$$V(x) = -c \left(\ln x_1 - \ln \frac{c}{d} + 1 \right) - a \left(\ln x_2 - \ln \frac{a}{b} + 1 \right) + dx_1 + bx_2.$$

Wegen $V(c/d, a/b) = 0$ und

$$\nabla V(x) \cdot f(x) = \begin{pmatrix} -\frac{c}{x_1} + d \\ -\frac{a}{x_2} + b \end{pmatrix} \cdot \begin{pmatrix} (a - bx_2)x_1 \\ (-c + dx_1)x_2 \end{pmatrix} = 0 \quad (4.25)$$

ist V eine Lyapunov-Funktion. (Den Nachweis, dass $V(x) > 0$ für alle $x \neq (c/d, a/b)$ ist, überlassen wir den LeserInnen.) Nach Satz 4.51 ist $\bar{x} = (c/d, a/b)$ ein stabiles Gleichgewicht.

Wir erhalten noch mehr: Aus (4.25) folgt, dass $(dV/dt)(x(t)) = 0$, also $V(x(t)) = V(x_0)$ für alle $t \geq 0$ konstant ist. Starten wir in x_0 , so läuft die Trajektorie entlang der Höhenlinie $\{x : V(x) = V(x_0)\}$ (siehe Abbildung 4.13). Man kann zeigen, dass diese Höhenlinie eine geschlossene Kurve um das Gleichgewicht $(c/d, a/b)$ ist. Dieses Resultat können wir wie folgt interpretieren: Die Beutepopulation vermehrt sich so lange, bis genügend Räuber herangewachsen sind, die die Beutetiere reduzieren. Ist die Beutepopulation zu klein geworden, haben die Räuber zu wenig Nahrung, und ihre Population schrumpft wieder. Diese Reduktion erlaubt es den Beutetieren, sich nahezu ungehindert zu vermehren, und der Kreislauf beginnt von vorne.

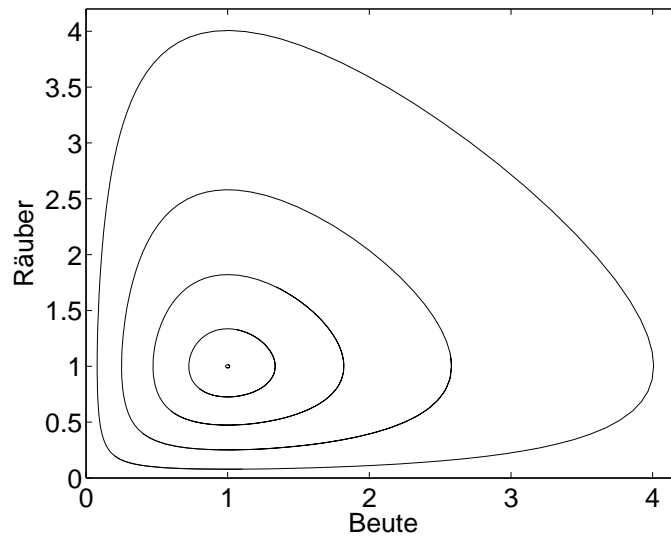


Abbildung 4.13: Phasenportrait für das Lotka-Volterra-Modell mit unbegrenztem Wachstum ($a = b = c = d = 1$). Das stabile Gleichgewicht befindet sich im Punkt $(1, 1)$.

Ein Räuber-Beute-Modell mit begrenztem Wachstum. Im Lotka-Volterra-Modell wird für die Beutepopulation bei Abwesenheit von Räubern unbegrenztes Wachstum vorausgesetzt, was nicht sehr realistisch ist. Wir ändern daher das Modell ab:

$$\begin{aligned}x_1' &= (a - bx_2 - ex_1)x_1, \\x_2' &= (-c + dx_1 - fx_2)x_2, \quad t \geq 0,\end{aligned}$$

wobei alle Koeffizienten positiv seien. Ohne Räuber vermehrt sich die Beutepopulation gemäß der logistischen Gleichung $x_1' = (a - ex_1)x_1$, was ungehindertes Wachstum verhindert. In diesem Fall konvergiert die Beutepopulation (bei positivem Anfangswert) gegen den asymptotischen Wert a/e . Die Koeffizienten e und f modellieren den *intraspezifischen Wettbewerb*.

Wir berechnen die Gleichgewichte. Falls $x_1 = 0$, folgt $x_2 = 0$ oder $x_2 = -c/f$. Falls $x_2 = 0$, erhalten wir $x_1 = 0$ oder $x_1 = a/e$. Dies ergibt die Gleichgewichte

$$\bar{x}_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \bar{x}_2 = \begin{pmatrix} a/e \\ 0 \end{pmatrix}.$$

Das Gleichgewicht $\bar{x} = (0, -c/f)^\top$ verwerfen wir, da wir nur an nichtnegativen Lösungen interessiert sind. Der letzte Fall ist durch $a - bx_2 - ex_1 = 0$ und $-c + dx_1 - fx_2 = 0$ gegeben, woraus

$$\bar{x}_3 = \frac{1}{bd + ef} \begin{pmatrix} af + bc \\ ad - ce \end{pmatrix}$$

folgt. Als nächstes betrachten wir die linearisierte Gleichung. Es ist

$$f'(x) = \begin{pmatrix} a - bx_2 - 2ex_1 & -bx_1 \\ dx_2 & -c + dx_1 - 2fx_2 \end{pmatrix}.$$

Das erste Gleichgewicht ist wie oben ein Sattel. Für das zweite Gleichgewicht erhalten wir

$$f'(\bar{x}_2) = \begin{pmatrix} -a & -\frac{ab}{e} \\ 0 & -c + \frac{ad}{e} \end{pmatrix}$$

mit den Eigenwerten $\lambda_1 = -a < 0$ und $\lambda_2 = -c + ad/e$. Falls $a/e < c/d$, so ist das zweite Gleichgewicht gemäß den Resultaten der vorigen Abschnitte asymptotisch stabil (und zwar auch für das nichtlineare System). Da a/e gerade die asymptotische Anzahl der Beutetiere ist, bedeutet die Bedingung $a/e < c/d$, dass die Sterberate der Räuber zu groß ist, um durch das vorhandene Nahrungsangebot kompensiert zu werden; die Räuber sterben aus. (Dies gilt auch, wenn $a/e = c/d$.)

Das dritte Gleichgewicht liegt genau dann im ersten Quadranten, wenn $ad - ce > 0$ bzw. $a/e > c/d$. Die Eigenwerte von $f'(\bar{x}_3)$ lauten

$$\lambda_{1/2} = -\frac{1}{2}(eA + fB) \pm \sqrt{\frac{1}{4}(eA + fB)^2 - (bd + ef)AB},$$

wobei $A = cf + bc > 0$ und $B = ad - ce > 0$. Die Realteile sind also negativ, und nach Satz 4.45 ist \bar{x}_3 asymptotisch stabil. Selbst wenn wir eine große Räuberpopulation zu Beginn aussetzen, wird die Anzahl der Beutetiere nach hinreichend langer Zeit $(af + bc)/(ef + bd)$ betragen. Dieser Fall wird *langfristige Koexistenz* genannt. Wir fassen zusammen (siehe Abbildung 4.14):

$$\begin{aligned} \frac{a}{e} \leq \frac{c}{d} : & \quad x_1(t) \rightarrow \frac{a}{e}, \quad x_2(t) \rightarrow 0 \quad \text{für } t \rightarrow \infty, \\ \frac{a}{e} > \frac{c}{d} : & \quad x_1(t) \rightarrow \frac{af + bc}{ef + bd}, \quad x_2(t) \rightarrow \frac{ad - ce}{ef + bd} \quad \text{für } t \rightarrow \infty. \end{aligned}$$

4.8 Diskrete Dynamik

Populationen bestehen aus endlich vielen Organismen und ändern sich in diskreten Zeitsprüngen. Bei einer großen Anzahl von Organismen ist es zweckmäßig, die Populationszahl $x(t)$ kontinuierlich in der Zeit t zu betrachten. Ein anderer Ansatz lautet, eine Dynamik mit äquidistanten Zeitschritten $\tau > 0$ zu definieren. Sei x_n die Populationszahl zur Zeit $t_n = n\tau$ ($n \in \mathbb{N}$, $\tau > 0$) und $a > 0$ die Geburtenrate pro Zeiteinheit. In der Zeit τ wächst die Population um den Anteil τax_n , so dass die Population zur Zeit $t_{n+1} = t_n + \tau$

$$x_{n+1} = x_n + \tau ax_n = (1 + \tau a)x_n \tag{4.26}$$

lautet. Bei dieser Gleichung handelt es sich um eine *Rekursionsgleichung* oder *Differenzgleichung*. In diesem Abschnitt analysieren wir die diskrete Dynamik von Gleichungen der Form

$$x_{n+1} = f(x_n, x_{n-1}, \dots, x_{n-k}).$$

Inwiefern unterscheiden sich Differenzgleichungen von Differentialgleichungen? Einerseits können wir den Differentialquotienten dx/dt durch den Differenzenquotienten $(x(t_{n+1}) - x(t_n))/\tau$ approximieren. Beispielsweise können wir $x' = ax$ durch

$$\frac{1}{\tau}(x(t_{n+1}) - x(t_n)) = ax(t_n)$$

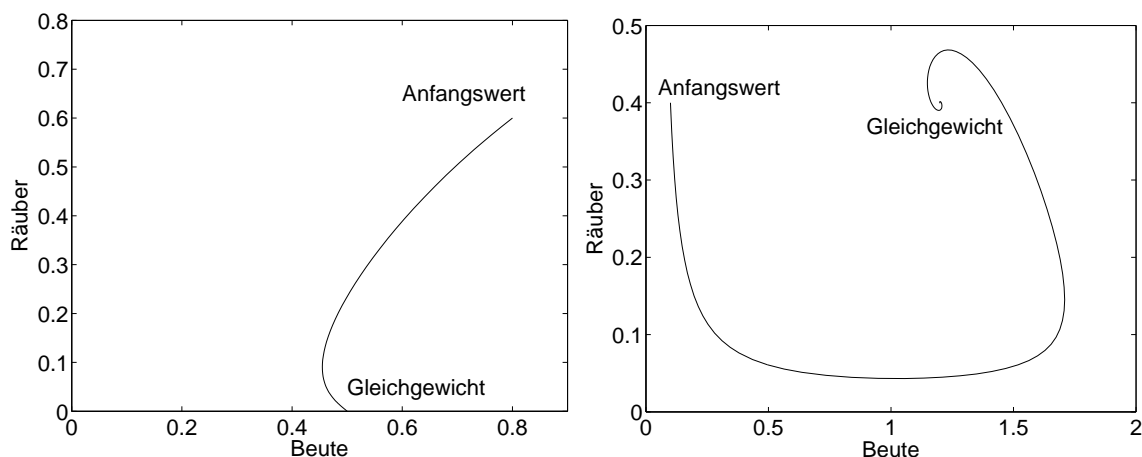


Abbildung 4.14: Phasenportrait für das Lotka-Volterra-Modell mit begrenztem Wachstum ($a = b = c = d = 1$). Links: Die Räuberpopulation stirbt aus, das Gleichgewicht ist $\bar{x} = (1/2, 0)$ ($e = f = 2$). Rechts: Die Populationen konvergieren gegen das Gleichgewicht $\bar{x} = (6/5, 2/5)$ ($e = f = 1/2$).

ersetzen, also, mit der Abkürzung $x_n = x(t_n)$,

$$x_{n+1} - x_n = \tau a x_n,$$

was äquivalent zu (4.26) ist. Andererseits werden wir in diesem Abschnitt zeigen, dass die Dynamik von Differenzgleichungen für hinreichend große $\tau > 0$ qualitativ sehr verschieden von der kontinuierlichen Dynamik sein kann. Ein Charakteristikum der diskreten Dynamik ist, dass recht einfache nichtlineare Differenzgleichungen über eine erstaunlich komplexe Dynamik verfügen und dass die Lösung (x_n) mit Computerhilfe einfach zu berechnen und zu visualisieren ist. Des Weiteren sind für viele Resultate keine Kenntnisse der Differentialrechnung erforderlich, so dass die Theorie auch für den fortgeschrittenen Schulunterricht zugänglich ist.

Schule

Wir beginnen mit linearen Differenzgleichungen k -ter Ordnung

$$x_n + a_1 x_{n-1} + \dots + a_k x_{n-k} = 0, \quad n \geq k, \quad (4.27)$$

wobei $a_1, \dots, a_k \in \mathbb{R}$ und $x_0, \dots, x_{k-1} \in \mathbb{R}$ gegeben sind. Durch diese Rekursionsvorschrift ist eine eindeutige Lösung definiert. Wir wollen eine explizite Formel für (x_n) finden.

Beispiel 4.54. Die Vermehrung von Kaninchenpaaren kann approximativ durch die *Fibonacci-Folge*

$$x_n = x_{n-1} + x_{n-2}, \quad x_0 = 0, \quad x_1 = 1,$$

modelliert werden. Hierbei beschreibt n die Anzahl der Monate. In dem Modell wird angenommen, dass jedes neugeborene Paar ab dem zweiten Lebensmonat geschlechtsreif wird und dann pro Monat ein weiteres Paar wirft. Die Kaninchen sterben nicht. Jeden Monat kommt zu der Paarzahl, die im letzten Monat gelebt haben, eine Anzahl neugeborener Paare hinzu, die der Paarzahl

entspricht, die im vorletzten Monat gelebt haben:

$$\begin{aligned}
 x_1 &= \underbrace{1}_{\text{Kind}}, & x_2 &= \underbrace{1}_{\text{erwachsen}}, \\
 x_3 &= \underbrace{1}_{\text{erwachsen}} + \underbrace{1}_{\text{Kind}} = 2, & x_4 &= \underbrace{2}_{\text{erwachsen}} + \underbrace{1}_{\text{Kind}} = 3, \\
 x_5 &= \underbrace{3}_{\text{erwachsen}} + \underbrace{2}_{\text{Kind}} = 5, & x_6 &= \underbrace{5}_{\text{erwachsen}} + \underbrace{3}_{\text{Kind}} = 8.
 \end{aligned}$$

Die Reihe wächst sehr schnell:

$$x_7 = 13, \quad x_8 = 21, \quad x_9 = 34, \quad x_{10} = 55, \quad x_{11} = 89, \dots$$

Um die Rechnung zu vereinfachen, wäre es praktisch, eine explizite Formel zur Berechnung von x_n zu haben. Wir leiten eine derartige Formel weiter unten her. \square

Im Schulunterricht ist es vermutlich ausreichend, nur den Fall $k = 2$ zu behandeln. Hierfür vereinfachen sich die folgenden Resultate und Beweise erheblich.

Schule

Eine Differenzgleichung der Form $x_n = f(x_{n-1})$ können wir als eine Approximation einer Differentialgleichung erster Ordnung interpretieren. Analog können wir Differenzgleichungen der Form (4.27) als eine Näherung einer Differentialgleichung k -ter Ordnung ansehen. Für diese Differentialgleichungen haben wir den Lösungsansatz $e^{\lambda t}$ gemacht, der auf die charakteristische Gleichung in λ geführt hat. Diese Analogie motiviert die folgende Vorgehensweise.

Definition 4.55 (Unabhängigkeit von Lösungen). *Wir nennen die Lösungen $(x_n^{(1)}), \dots, (x_n^{(m)})$ von (4.27) mit Anfangswerten x_0, \dots, x_{k-1} unabhängig, wenn es für alle x_0, \dots, x_{k-1} Koeffizienten $a_1, \dots, a_m \in \mathbb{R}$ gibt, so dass gilt:*

$$a_1 x_j^{(1)} + \dots + a_m x_j^{(m)} = x_j, \quad j = 0, \dots, k-1.$$

Falls $m = k$, nennen wir $(x_n^{(1)}), \dots, (x_n^{(m)})$ ein Fundamentalsystem.

Satz 4.56. *Die charakteristische Gleichung $\lambda^k + a_1 \lambda^{k-1} + \dots + a_k = 0$ besitze k paarweise verschiedene Nullstellen $\lambda_1, \dots, \lambda_k \in \mathbb{R} \setminus \{0\}$. Dann ist durch*

$$x_n^{(j)} = \lambda_j^n, \quad n \in \mathbb{N}, \quad j = 1, \dots, k,$$

ein Fundamentalsystem von (4.27) definiert. Die allgemeine Lösung von (4.27) ist gegeben durch $x_n = \sum_{j=1}^k c_j x_n^{(j)}$, wobei die Koeffizienten c_j aus den Anfangswerten x_0, \dots, x_{k-1} berechnet werden können.

Beweis. 1. Wir weisen zuerst nach, dass durch $x_n = \lambda_j^n$ eine Lösung von (4.27) gegeben ist. Dies folgt aus

$$x_n + a_1 x_{n-1} + \dots + a_k x_{n-k} = \lambda^{n-k} (\lambda^k + a_1 \lambda^{k-1} + \dots + a_k) = 0,$$

denn λ_j ist eine Nullstelle.

2. Um zu zeigen, dass durch die Lösungen ein Fundamentalsystem definiert ist, ist für gegebene x_0, \dots, x_{k-1} das Gleichungssystem

$$\begin{aligned} a_1 \lambda_1^0 + a_2 \lambda_2^0 + \dots + a_k \lambda_k^0 &= x_0, \\ &\vdots \\ a_1 \lambda_1^{k-1} + a_2 \lambda_2^{k-1} + \dots + a_k \lambda_k^{k-1} &= x_{k-1} \end{aligned} \quad (4.28)$$

zu lösen. Definiere hierfür die *Lagrange-Polynome*

$$L_j(\lambda) = \frac{\prod_{m \neq j} (\lambda - \lambda_m)}{\prod_{m \neq j} (\lambda_j - \lambda_m)}, \quad j = 1, \dots, k.$$

Sie besitzen die Eigenschaft $L_j(\lambda_m) = \delta_{jm}$ und können geschrieben werden als

$$L_j(\lambda) = \sum_{m=0}^{k-1} c_{j,m} \lambda^m, \quad j = 1, \dots, k, \quad (4.29)$$

mit gewissen Koeffizienten $c_{j,m}$. Ferner gilt

$$\sum_{j=1}^k \lambda_j^p L_j(\lambda_m) = \sum_{j=1}^k \lambda_j^p \delta_{jm} = \lambda_m^p, \quad m = 1, \dots, k, \quad p = 0, \dots, k-1.$$

Ein Polynom $(k-1)$ -ter Ordnung ist durch die Vorgabe von k Werten eindeutig bestimmt. Da L_j ein Polynom $(k-1)$ -ter Ordnung ist, gilt die obige Gleichung sogar für alle $\lambda \in \mathbb{R}$:

$$\sum_{j=1}^k \lambda_j^p L_j(\lambda) = \lambda^p.$$

Setzen wir die Formulierung (4.29) in diese Gleichung ein, erhalten wir

$$\sum_{m=0}^{k-1} \left(\sum_{j=1}^k \lambda_j^p c_{j,m} \right) \lambda^m = \lambda^p \quad \text{für alle } \lambda \in \mathbb{R}.$$

Auf beiden Seiten stehen Polynome, die für alle λ übereinstimmen; also müssen ihre Koeffizienten gleich sein. Dies impliziert

$$\begin{aligned} m = p : \quad & \sum_{j=1}^k \lambda_j^p c_{j,m} = 1, \\ m \neq p : \quad & \sum_{j=1}^k \lambda_j^p c_{j,m} = 0. \end{aligned} \quad (4.30)$$

3. Wir behaupten, dass durch $a_j = c_{j,k-1}x_{k-1} + \dots + c_{j,0}x_0$, $j = 1, \dots, k$, eine Lösung von (4.28) gegeben ist, was die Behauptung des Satzes zeigt. Sei $p = 0, \dots, k-1$. Dann folgt aus der Definition von a_j und (4.30)

$$\sum_{j=1}^k a_j \lambda_j^p = \sum_{j=1}^k \left(\sum_{m=0}^{k-1} c_{j,m} x_m \right) \lambda_j^p = \sum_{m=0}^{k-1} \left(x_m \sum_{j=1}^k \lambda_j^p c_{j,m} \right) = \sum_{m=0}^{k-1} x_m \delta_{mp} = x_p,$$

also löst a_j in der Tat (4.28). □

Beispiel 4.57. Wir betrachten die Fibonacci-Gleichung

$$x_n - x_{n-1} - x_{n-2} = 0, \quad n \geq 2.$$

Dies ist von der Form (4.27) mit $k = 2$ und $a_1 = a_2 = -1$. Die charakteristische Gleichung $\lambda^2 - \lambda - 1 = 0$ besitzt die Nullstellen $\lambda_{1/2} = (1 \pm \sqrt{5})/2$. Wir erhalten also die unabhängigen Lösungen

$$x_n^{(1)} = \frac{1}{2^n} (1 + \sqrt{5})^n, \quad x_n^{(2)} = \frac{1}{2^n} (1 - \sqrt{5})^n.$$

Die allgemeine Lösung lautet somit

$$x_n = \frac{1}{2^n} (c_1 (1 + \sqrt{5})^n + c_2 (1 - \sqrt{5})^n), \quad n \in \mathbb{N}.$$

Die Koeffizienten c_1 und c_2 bestimmen wir aus den Anfangswerten

$$0 = x_0 = c_1 + c_2, \quad 1 = x_1 = \frac{1}{2} (c_1 (1 + \sqrt{5}) + c_2 (1 - \sqrt{5})),$$

also $c_1 = -c_2 = 1/\sqrt{5}$. Die Fibonacci-Folge kann also explizit über

$$x_n = \frac{1}{2^n \sqrt{5}} ((1 + \sqrt{5})^n - (1 - \sqrt{5})^n), \quad n \in \mathbb{N},$$

berechnet werden. Erstaunlicherweise liefert die Differenz der beiden irrationalen Zahlen stets ein ganzzahliges Ergebnis. □

Besitzt die charakteristische Gleichung mehrfache Nullstellen, so können wir in Analogie zu den Differentialgleichungen n -ter Ordnung Lösungen der Form $p(n)\lambda^k$ ansetzen, wobei $p(n)$ ein Polynom in n ist. Genauer gilt:

Proposition 4.58. Sei λ eine m -fache Nullstelle der charakteristischen Gleichung $P(\lambda) = \lambda^k + a_1 \lambda^{k-1} + \dots + a_k = 0$. Dann sind

$$x_n^{(0)} = n \lambda^{n-1}, \quad x_n^{(1)} = n(n-1) \lambda^{n-2}, \dots, \quad x_n^{(m-1)} = n(n-1) \dots (n-m+2) \lambda^{n-m+1}$$

unabhängige Lösungen von (4.27).

Beweis. Da λ eine m -fache Nullstelle ist, gilt $P^{(j)}(\lambda) = 0$ für alle $j = 0, \dots, m-1$. Dies ergibt

$$\begin{aligned} 0 &= (\lambda^{n-k}P(\lambda))^{(j)} = (\lambda^n + a_1\lambda^{n-1} + \dots + a_k\lambda^{n-k})^{(j)} \\ &= n(n-1)\cdots(n-j+1)\lambda^{n-j} + a_1(n-1)\cdots(n-j)\lambda^{n-1-j} \\ &\quad + \dots + a_k(n-k)\cdots(n-k-j+1)\lambda^{n-k-j} \\ &= x_n^{(j)} + a_1x_{n-1}^{(j)} + \dots + a_kx_{n-k}^{(j)}. \end{aligned}$$

Also ist $x_n^{(j)}$ eine Lösung von (4.27). Den Nachweis, dass die Lösungen unabhängig sind, überlassen wir den LeserInnen. \square

Beispiel 4.59. Wir betrachten die Gleichung

$$x_n - x_{n-1} + \frac{1}{4}x_{n-2} = 0, \quad x_0 = 0, \quad x_1 = 1.$$

Die charakteristische Gleichung $0 = \lambda^2 - \lambda + \frac{1}{4} = (\lambda - \frac{1}{2})^2$ besitzt die doppelte Nullstelle $\lambda = \frac{1}{2}$. Die allgemeine Lösung lautet daher

$$x_n = c_1 \frac{n}{2^{n-1}} + c_2 \frac{n(n-1)}{2^{n-2}}, \quad n \geq 2.$$

Aus $1 = x_1 = c_1$ und $1 = x_2 = c_1 + 2c_2$ folgt $c_1 = 1$ und $c_2 = 0$, also $x_n = n2^{-(n-1)}$, $n \geq 1$. Insbesondere gilt $x_n \rightarrow 0$ für $n \rightarrow \infty$. \square

Wir wenden uns nun nichtlinearen Differenzgleichungen der Form

$$x_{n+1} = f(x_n), \quad n \in \mathbb{N}, \quad x_0 = \xi, \tag{4.31}$$

zu. Die Existenz von Lösungen ist kein Problem; durch die Rekursion wird ja eine Lösung automatisch definiert. Wir bezeichnen eine Lösung mit $x_n(\xi)$. Uns interessiert das Langzeitverhalten und die Stabilität von Gleichgewichten.

Definition 4.60 (Stabilität). Für die Differenzgleichung (4.31) definieren wir:

(i) Ein Punkt $\bar{x} \in \mathbb{R}$ heißt Gleichgewicht von (4.31), wenn $f(\bar{x}) = \bar{x}$.

(ii) Ein Gleichgewicht \bar{x} heißt stabil, wenn

$$\forall \varepsilon > 0 : \exists \delta > 0 : \forall \xi \in B_\delta(\bar{x}) : \forall n \in \mathbb{N} : x_n(\xi) \in B_\varepsilon(\bar{x}).$$

(iii) Ein Gleichgewicht \bar{x} heißt asymptotisch stabil, wenn es stabil ist und wenn

$$\exists \delta > 0 : \forall \xi \in B_\delta(\bar{x}) : \lim_{n \rightarrow \infty} x_n(\xi) = \bar{x}.$$

Beispiel 4.61. Wir behaupten, dass das Gleichgewicht $\bar{x} = 0$ für $x_{n+1} = \sin x_n$ asymptotisch stabil ist. Sei dazu $0 < \varepsilon < 1$ und wähle $\delta = \varepsilon$. Sei weiter $|\xi| < \delta$. Wir bemerken, dass für alle $x \in \mathbb{R}$ gilt $|\sin x| \leq |x|$. Dann folgt $|x_1(\xi)| = |\sin \xi| \leq |\xi| < \delta = \varepsilon$, $|x_2(\xi)| = |\sin x_1(\xi)| \leq |x_1(\xi)| < \varepsilon$, also nach endlich vielen Schritten

$$|x_{n+1}(\xi)| = |\sin x_n(\xi)| \leq |x_n(\xi)| < \varepsilon.$$

Es bleibt die Konvergenz von $(x_n(\xi))$ zu zeigen. Seien $\delta = 1$ und $|\xi| < \delta$. Die obige Abschätzung zeigt, dass $(|x_n(\xi)|)$ monoton fallend ist. Starten wir mit einem positiven Anfangswert $0 < \xi < 1$, so gilt $0 \leq x_{n+1}(\xi) \leq x_n(\xi) \leq 1$, also ist auch $(x_n(\xi))$ monoton fallend. Falls $-1 < \xi < 0$, so folgt $-1 \leq x_n(\xi) \leq x_{n+1}(\xi) \leq 0$. Folglich ist $(x_n(\xi))$ eine monotone, beschränkte Folge und damit konvergent. Sei $x = \lim_{n \rightarrow \infty} x_n(\xi) \in [-1, 1]$. Der Grenzwert in der Rekursionsgleichung liefert $x = \sin x$, also muss $x = 0$ gelten. Wir haben gezeigt, dass $\bar{x} = 0$ asymptotisch stabil ist. \square

Ein einfaches Kriterium für asymptotische Stabilität ist wie folgt.

Satz 4.62. Seien $f : \mathbb{R} \rightarrow \mathbb{R}$ stetig differenzierbar und \bar{x} ein Gleichgewicht von (4.31). Falls $|f'(\bar{x})| < 1$, so ist \bar{x} asymptotisch stabil.

Beweis. 1. Wir zeigen zuerst die Stabilität von \bar{x} . Sei $\varepsilon > 0$. Wegen der Stetigkeit von f' existiert ein $0 < \delta < \varepsilon$, so dass $|f'(x)| \leq 1$ für alle $|x - \bar{x}| < \delta$. Sei $|\xi - \bar{x}| < \delta$. Wir behaupten, dass $|x_n(\xi) - \bar{x}| < \delta < \varepsilon$ für alle $n \in \mathbb{N}$. Wir zeigen dies durch vollständige Induktion. Der Induktionsanfang $|x_0(\xi) - \bar{x}| = |\xi - \bar{x}| < \delta < \varepsilon$ ist erfüllt. Nach dem Mittelwertsatz existiert ein z zwischen $x_{n-1}(\xi)$ und \bar{x} , so dass

$$|x_n(\xi) - \bar{x}| = |f(x_{n-1}(\xi)) - f(\bar{x})| \leq |f'(z)| |x_{n-1}(\xi) - \bar{x}|.$$

Wegen der Induktionsannahme ist $|x_{n-1}(\xi) - \bar{x}| < \delta$ erfüllt. Dann gilt auch $|z - \bar{x}| < \delta$ und damit $|f'(z)| \leq 1$. Dies ergibt

$$|x_n(\xi) - \bar{x}| \leq |x_{n-1}(\xi) - \bar{x}| < \delta < \varepsilon,$$

also die Induktionsbehauptung.

2. Es bleibt zu zeigen, dass $(x_n(\xi))$ gegen \bar{x} konvergiert. Es existiert ein $\delta > 0$, so dass für alle $|x - \bar{x}| < \delta$ gilt $|f'(x)| \leq q < 1$ für ein $0 < q < 1$. Sei $|\xi - \bar{x}| < \delta$. Nach dem Mittelwertsatz existiert ein z_1 zwischen ξ und \bar{x} , das $|z_1 - \bar{x}| < \delta$ und damit $|f'(z_1)| \leq 1$ erfüllt, so dass

$$|x_1(\xi) - \bar{x}| = |f(x_0(\xi)) - f(\bar{x})| = |f'(z_1)(\xi - \bar{x})| \leq |\xi - \bar{x}| < \delta.$$

Mit vollständiger Induktion können wir so zeigen, dass $|x_n(\xi) - \bar{x}| < \delta$ für alle $n \in \mathbb{N}$. Sei nun $\varepsilon > 0$ gegeben und wähle $N \in \mathbb{N}$ so, dass $q^N \delta < \varepsilon$. Dann folgt für alle $n \in \mathbb{N}$ mit dem Mittelwertsatz

$$|x_{n+N}(\xi) - \bar{x}| = |f'(z_2)(x_{n+N-1}(\xi) - \bar{x})|.$$

Die Zahl z_2 liegt zwischen x_{n+N-1} und \bar{x} und erfüllt $|z_2 - \bar{x}| < \delta$ sowie $|f'(z_2)| \leq q$. Dies impliziert

$$|x_{n+N}(\xi) - \bar{x}| \leq q |x_{n+N-1}(\xi) - \bar{x}|.$$

Mit vollständiger Induktion können wir schließen, dass

$$|x_{n+N}(\xi) - \bar{x}| \leq q^N |x_n(\xi) - \bar{x}| < q^N \delta < \varepsilon.$$

Da $\varepsilon > 0$ beliebig war, folgern wir, dass $x_n(\xi) \rightarrow \bar{x}$ für $n \rightarrow \infty$. □

Das vorige Beispiel $x_{n+1} = \sin x_n$ zeigt, dass die Bedingung $|f'(\bar{x})| < 1$ nicht notwendig für die asymptotische Stabilität ist, denn die Funktion $f(x) = \sin x$ erfüllt nur die Gleichung $f'(\bar{x}) = f'(0) = \cos 0 = 1$.

Beispiel 4.63. Wir betrachten die logistische Gleichung

Schule

$$x_{n+1} = rx_n(1 - x_n), \quad n \in \mathbb{N}, \quad x_0 \geq 0,$$

mit $r > 0$. Es gibt zwei Gleichgewichte: $\bar{x}_1 = 0$ und $\bar{x}_2 = 1 - 1/r$. Die Ableitung von f lautet $f'(x) = r(1 - 2x)$. Wegen $f'(0) = r$ ist das Gleichgewicht \bar{x}_1 nach Satz 4.62 asymptotisch stabil, falls $r < 1$. Man kann zeigen, dass \bar{x}_1 für $r > 1$ instabil wird. Für das zweite Gleichgewicht gilt $|f'(1 - 1/r)| = |2 - r| < 1$ genau dann, falls $1 < r < 3$. Also ist \bar{x}_2 für alle $1 < r < 3$ asymptotisch stabil. Dieses Gleichgewicht wird für $r > 3$ instabil.

An der Stelle $r = 3$ geschieht etwas Überraschendes. Betrachten wir nämlich die Rekursion $x_{n+1} = f(f(x_{n-1}))$, also

$$x_{n+1} = f(f(x_{n-1})) = r^2 x_{n-1} (1 - x_{n-1}) (rx_{n-1}^2 - rx_{n-1} + 1).$$

Die Gleichgewichte sind Lösungen der Gleichung

$$x = r^2 x(1 - x)(rx^2 - rx + 1) \quad \text{oder} \quad x(rx - r + 1)(r^2 x^2 - r(r + 1)x + r + 1) = 0.$$

Wir erhalten die bereits bekannten Gleichgewichte $\bar{x}_1 = 0$ und $\bar{x}_2 = 1 - 1/r$ sowie (wegen $r \geq 3$)

$$\bar{x}_{3/4} = \frac{1}{2r} (r + 1 \pm \sqrt{(r + 1)(r - 3)}).$$

Diese Gleichgewichte entsprechen *periodischen* Lösungen der ursprünglichen Rekursionsgleichung, nämlich solche mit der Periode zwei. Wir nennen sie *2-periodische Fixpunkte* für die logistische Gleichung. Das Gleichgewicht \bar{x}_2 verzweigt sich bei $r_0 = 3$ in zwei 2-periodische Fixpunkte (siehe Abbildung 4.15). Man kann zeigen, dass sich beide Fixpunkte wiederum verzweigen in periodische Fixpunkte, und zwar ab $r_1 = 1 + \sqrt{6}$. Dies ergibt vier 4-periodische Fixpunkte für die logistische Gleichung. Diesen Vorgang nennt man *Periodenverdopplung*. Es gibt eine Folge (r_n) , bei denen die 2^n -periodischen Lösungen sich in 2^{n+1} -periodische Lösungen verzweigen. Interessanterweise konvergiert (r_n) gegen den Grenzwert $r_\infty = 3,5699\dots$ (siehe Abbildung 4.16). Außerdem gilt

$$\lim_{n \rightarrow \infty} \frac{r_n - r_{n-1}}{r_{n+1} - r_n} = \delta = 4,6692\dots$$

Diese Konstante ist universell für Periodenverdopplungen und wird die *Feigenbaumkonstante* genannt. Für Parameter $r > r_\infty$ wird die Dynamik chaotisch, d.h., die Lösungen scheinen irregulär zu verlaufen und hängen sehr sensitiv von der Wahl der Anfangswerte ab. Dieses Verhalten wird in der Chaostheorie genauer untersucht. □

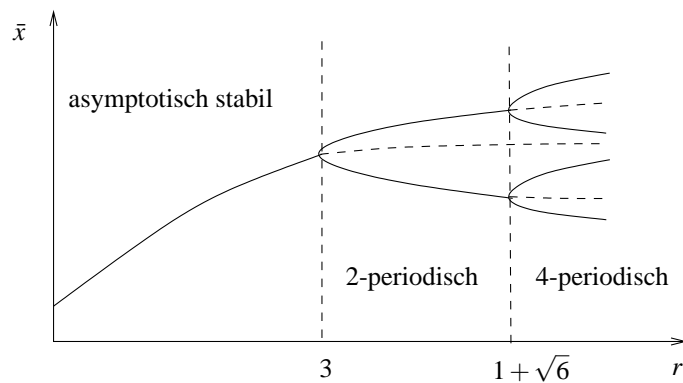


Abbildung 4.15: Schematische Darstellung der Fixpunkte der logistischen Gleichung in Abhängigkeit des Parameters $r < 4$. Die gestrichelten Linien stellen die instabilen periodischen Fixpunkte dar.

5 Partielle Differentialgleichungen erster Ordnung

Ziel dieses Kapitels ist die Modellierung und Analysis des Verkehrsflusses auf einer einspurigen Straße mit partiellen Differentialgleichungen. Eine partielle Differentialgleichung ist eine Gleichung, in der partielle Ableitungen auftreten und bei denen die abzuleitende Funktion die Unbekannte ist. Obwohl partielle Ableitungen in der Regel nicht zum Schulstoff gehören, können die Differentialgleichungen teilweise recht elementar, nämlich graphisch, gelöst werden, so dass dieses Thema in der Schule denkbar ist.

Schule

5.1 Die Burgers-Gleichung

Wir beschreiben die Fahrzeuge nicht individuell, sondern betrachten die Fahrzeugdichte $n(x, t)$, also die Anzahl der Fahrzeuge pro Kilometer, am Ort x zur Zeit t . Die Anzahl der Fahrzeuge, die sich im Intervall $[a, b]$ befinden, lautet

$$\int_a^b n(x, t) dx.$$

Wir wollen eine Gleichung für die zeitliche Entwicklung der Fahrzeugdichte herleiten. Sei $v(x, t)$ die mittlere Geschwindigkeit der Fahrzeuge in x zur Zeit t . Die Anzahl der Fahrzeuge, die den Punkt x zur Zeit t passieren, lautet pro Längeneinheit $n(x, t)v(x, t)$. Die Anzahl der Fahrzeuge in $[a, b]$ ändert sich zeitlich je nach den in $[a, b]$ eintreffenden oder hinausfahrenden Fahrzeugen:

$$\frac{d}{dt} \int_a^b n(x, t) dx = n(a, t)v(a, t) - n(b, t)v(b, t).$$

Wir nehmen für die Herleitung an, dass die Funktionen hinreichend regulär sind. Dann können wir Differentiation und Integration vertauschen und über $t \in [t_1, t_2]$ integrieren:

$$\int_{t_1}^{t_2} \int_a^b n_t(x, t) dx dt = \int_{t_1}^{t_2} (n(a, t)v(a, t) - n(b, t)v(b, t)) dt = - \int_{t_1}^{t_2} \int_a^b (nv)_x(x, t) dx dt,$$

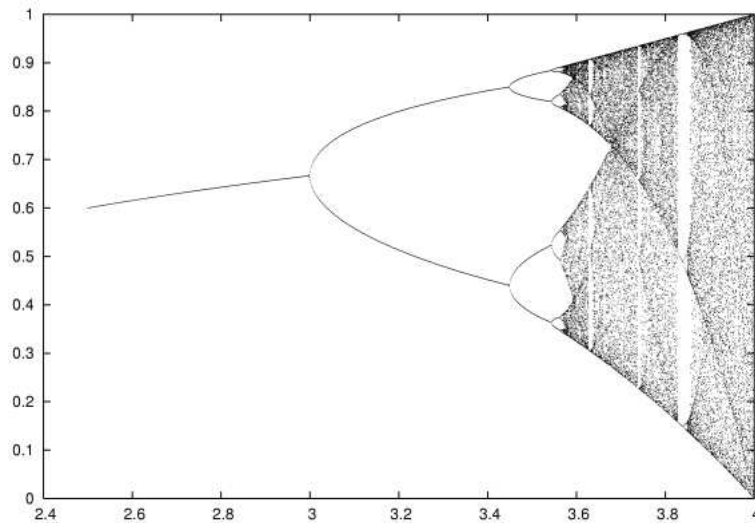


Abbildung 4.16: Numerische Lösung der 2^n -periodischen Fixpunkte der logistischen Gleichung in Abhängigkeit des Parameters $r < 4$.

wobei die Indizes für die entsprechende partielle Ableitung stehen. Für alle $a, b \in \mathbb{R}$ und $t_1, t_2 > 0$ gilt also

$$\int_{t_1}^{t_2} \int_a^b (n_t + (nv)_x)(x, t) dx dt = 0.$$

Dies ist aber bei stetigen Funktionen nur möglich, wenn der Integrand verschwindet. (Anderenfalls ist der Integrand in einer Umgebung eines Punktes von null verschieden, und Integration über diese Umgebung lieferte einen von null verschiedenen Wert, Widerspruch.) Wir erhalten also die partielle Differentialgleichung

$$n_t + (nv)_x = 0, \quad x \in \mathbb{R}, t > 0.$$

Wie bei den gewöhnlichen Differentialgleichungen schreiben wir eine Anfangsbedingung für $t = 0$ vor:

$$n(x, 0) = n_0(x), \quad x \in \mathbb{R}.$$

Wir benötigen eine Gleichung für die Geschwindigkeit. Wir nehmen an, dass sie von der Fahrzeugdichte abhängt: Wenn die Straße frei ist ($n = 0$), werden die Autos mit der maximal zulässigen Geschwindigkeit fahren, $v = v_{\max}$. Ist die Straße voll, werden die Fahrzeuge langsamer und stoppen ($v = 0$) in einem Stau, in dem die Fahrzeugdichte maximal ist ($n = n_{\max}$). Ein einfaches Modell, das diese Eigenschaften erfüllt, ist die lineare Beziehung

$$v(n) = v_{\max} \left(1 - \frac{n}{n_{\max}} \right), \quad 0 \leq n \leq n_{\max}.$$

Um die Gleichung zu vereinfachen, nehmen wir an, dass $n_{\max} = 1$ und $v_{\max} = 1$, und führen die Variablentransformation $u = 1 - 2n$ durch. Dann ist

$$u_t = -2n_t = 2(nv(n))_x = 2(n(1-n))_x = 2(1-2n)n_x = -uu_x = -\left(\frac{u^2}{2}\right)_x.$$

Unsere partielle Differentialgleichung lautet also

$$u_t + \left(\frac{u^2}{2}\right)_x = 0, \quad x \in \mathbb{R}, t > 0, \quad u(\cdot, 0) = u_0, \quad (5.1)$$

wobei $u_0 = 1 - 2n_0$. Sie wird die (*reibungslose*) *Burgers-Gleichung* genannt. Eine leere Straße entspricht dem Wert $u = 1$, eine volle Straße dem Wert $u = -1$.

Beispiel 5.1. Wir lösen diese Gleichung für spezielle Anfangswerte.

(1) Wir betrachten eine Anfangsdichte, die in Fahrtrichtung abnimmt:

$$u_0(x) = \begin{cases} -1 & : x < -1 \\ x & : -1 \leq x \leq 1 \\ 1 & : x > 1. \end{cases} \quad (5.2)$$

In der linken Hälfte $x > 1$ ist die Straße voll ($u = -1$), in der rechten Hälfte $x > 1$ ist sie leer ($u = 1$), und dazwischen nimmt die Fahrzeugdichte linear ab. Wir behaupten, dass die Funktion

$$u(x, t) = \begin{cases} -1 & : x < -1 - t \\ x/(1+t) & : -1 - t \leq x \leq 1 + t \\ 1 & : x > 1 + t, \end{cases} \quad x \in \mathbb{R}, t > 0,$$

die Gleichung (5.1) bis auf die Geraden $x = -1 - t$ und $x = 1 + t$ löst (siehe Abbildung 5.17 links). Seien $t \geq 0$ und $x \neq -1 - t, x \neq 1 + t$. Dann ist

$$u_t = \begin{cases} 0 & : x < -1 - t \\ -x/(1+t)^2 & : -1 - t < x < 1 + t, \\ 0 & : x > 1 + t \end{cases}, \quad u_x = \begin{cases} 0 & : x < -1 - t \\ 1/(1+t) & : -1 - t < x < 1 + t \\ 0 & : x > 1 + t, \end{cases}$$

woraus sofort $u_t = -uu_x$ folgt. Das Ergebnis können wir folgendermaßen interpretieren. Die Fahrzeugfront setzt sich in den Bereich $x > 1$ fort, während der Stau in $x < -1$ kleiner wird. Da sich der Verkehr im Bereich $(-1 - t, 1 + t)$ entspannt, nennt man diese Situation eine *Verdünnungswelle*.

(2) Sei

$$u_0(x) = \begin{cases} 1 & : x < -1 \\ -x & : -1 \leq x \leq 1 \\ -1 & : x > 1. \end{cases} \quad (5.3)$$

Die Fahrzeugdichte nimmt also in Fahrtrichtung zu. Anfangs ist die Straße im linken Bereich $x < -1$ leer, während der rechte Bereich $x > 1$ voller Fahrzeuge ist. Im Bereich $(-1, 1)$ variiert die Fahrzeugdichte. Eine Lösung von (5.1) lautet

$$u(x, t) = \begin{cases} 1 & : x < -1 + t \\ -x/(1-t) & : -1 + t \leq x \leq 1 - t \\ 0 & : x > 1 - t. \end{cases}$$

Dies können wir wie oben nachrechnen. Beispielsweise ist für $-1+t < x < 1-t$

$$u_t(x,t) = -\frac{x}{(1-t)^2}, \quad uu_x = \frac{-x}{1-t} \frac{-1}{1-t} = \frac{-x}{(1-t)^2}.$$

Diese Lösung kann nur für $t < 1$ existieren! Die Fahrzeuge im Intervall $[-1, 1]$ fahren nach rechts, bis sie auf das Stauende ($u = -1$) treffen. Das Stauende verschiebt sich dadurch nach links und erreicht zur Zeit $t = 1$ den Punkt $x = 0$. Wir nennen diese Situation einen *Schock*. \square

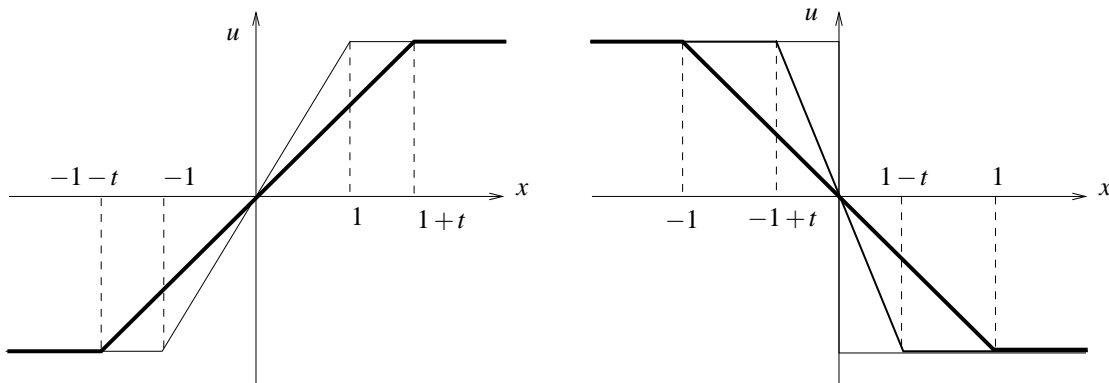


Abbildung 5.17: Links: Verdünnungswelle mit $u(x,0)$ (dünne Linie) und $u(x,t)$ (dicke Linie). Rechts: Ein Schock bildet sich zur Zeit $t = 1$ mit $u(x,0)$ (dicke Linie), $u(x,t)$ (mittelstarke Linie) und $u(x,1)$ (dünne Linie).

Das obige Beispiel wirft einige Fragen auf:

- Inwiefern sind die obigen Funktionen Lösungen von (5.1)?
- Was geschieht für Zeiten $t > 1$? Physikalisch sollte eine Lösung für alle Zeiten existieren. Wie können wir dies mathematisch formulieren?

Wir beantworten diese Fragen im folgenden Abschnitt.

Bemerkung 5.2. Die Differentialgleichung

$$n_t + (nv)_x = 0$$

ist auch gültig für die Evolution von Gasen oder Flüssigkeiten. Sie wird komplementiert durch Gleichungen für die Geschwindigkeit v und Energie E :

$$\begin{aligned} (nv)_t + (nv^2 + p(n,E))_x &= 0, \\ (nE)_t + (nEv + p(n,E)v)_x &= 0, \end{aligned}$$

wobei der Druck $p(n,E)$ eine Funktion der Teilchendichte n und Energie E und die Energie die Summe aus der internen und kinetischen Energie ist, $E = e + v^2/2$. Die Variablen sind also n , v und e . Diese Gleichungen heißen *Euler-Gleichungen* und sind grundlegend für die Beschreibung von kompressiblen Gasen oder Flüssigkeiten ohne viskose Effekte. \square

5.2 Methode der Charakteristiken

Die Burgers-Gleichung aus dem vorigen Abschnitt ist vom Typ

$$u_t + f(u)_x = 0 \quad x \in \mathbb{R}, t > 0, \quad u(\cdot, 0) = u_0, \quad (5.4)$$

wobei $f : \mathbb{R} \rightarrow \mathbb{R}$ eine Funktion sei. Wir wollen diese Gleichung mit der sogenannten *Methode der Charakteristiken* lösen. Hierfür müssen wir zuerst definieren, was wir unter einer Charakteristik verstehen.

Definition 5.3 (Charakteristik). Sei $u : [0, T] \rightarrow \mathbb{R}$ eine differenzierbare Lösung von (5.4). Die Lösungen $x : [0, T] \rightarrow \mathbb{R}$ des Anfangswertproblems

$$x'(t) = f'(u(x(t), t)), \quad t > 0, \quad x(0) = x_0,$$

wobei $x_0 \in \mathbb{R}$, heißen Charakteristiken von (5.4).

Die Haupteigenschaft von Charakteristiken $x(t)$ ist, dass die Lösung u konstant entlang $x(t)$ ist:

$$\frac{d}{dt}u(x(t), t) = u_t(x(t), t) + u_x(x(t), t)x'(t) = u_t(x(t), t) + u_x(x(t), t)f'(u(x(t), t)) = 0,$$

also $u(x(t), t) = \text{const.}$ für alle $t > 0$.

Beispiel 5.4. Sei $f(u) = u^2/2$. Dann ist $f'(u) = u$ und die Charakteristiken sind Lösungen von

$$x'(t) = u(x(t), t), \quad t > 0, \quad x(0) = x_0.$$

Da $u(x(t), t)$ konstant ist für alle $t > 0$ (und gleich $u(x(0), 0) = u_0(x_0)$), ist $x(t)$ also eine Gerade:

$$x(t) = u(x(t), t)t + x_0 = u_0(x_0)t + x_0.$$

In der (t, x) -Ebene besitzt diese Gerade die Steigung $u_0(x_0)$. Es ist jedoch üblich, die Geraden in der (x, t) -Ebene zu betrachten; in diesem Fall besitzt die Charakteristik die Steigung $1/u_0(x_0)$. Charakteristiken bieten eine alternative Möglichkeit, Lösungen zu illustrieren. Beachte, dass eine Charakteristik *nicht* den Weg eines Fahrzeugs darstellt; vielmehr beschreibt sie die Entwicklung der Dichtewerte. Wir betrachten die beiden Situationen aus Beispiel 5.1.

- (1) Verdünnungswelle: Sei der Anfangswert durch (5.2) gegeben. Die Charakteristiken durch $x_0 < -1$ besitzen die Steigung $1/u_0(x_0) = -1$, sind also nach links geneigt (siehe Abbildung 5.18 links). Die Charakteristiken durch $x_0 > 1$ haben die Steigung $1/u_0(x_0) = 1$ und sind nach rechts geneigt. Im Bereich $x_0 \in (-1, 1)$ variiert die Steigung; dort ist die Verdünnungswelle. Im Punkt $x_0 = 0$ ist die Steigung unendlich, d.h., die Charakteristik durch null ist vertikal, ansonsten ist die Steigung durch $1/u_0(x_0) = (1+t)/x_0$ gegeben.

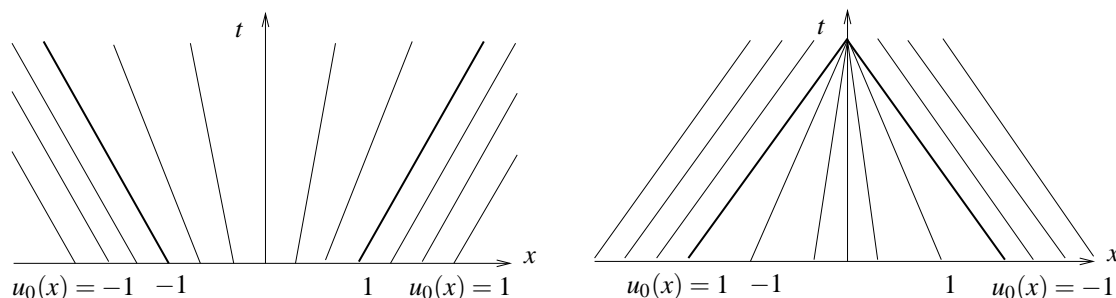


Abbildung 5.18: Links: Charakteristiken im Falle einer Verdünnungswelle. Rechts: Zur Zeit $t = 1$ treffen sich die Charakteristiken im Punkt $x = 0$ und bilden einen Schock.

- (2) Schock: Sei u_0 durch (5.3) gegeben. Die Charakteristiken durch $x_0 < -1$ (bzw. $x_0 > 1$) haben die Steigung 1 (bzw. -1); für $x_0 \in (-1, 1)$ ist die Steigung gleich $1/u_0(x_0) = -(1-t)/x_0$ (siehe Abbildung 5.18 rechts). Wiederum ist sie unendlich, falls $x_0 = 0$. Zur Zeit $t = 1$ treffen sich die Charakteristiken, die in $x_0 \in (-1, 1)$ starten, im Punkt $x_0 = 0$. Entlang jeder Charakteristik ist die Lösung konstant; also muss die Lösung u zur Zeit $t = 1$ im Punkt $x = 0$ unstetig sein. \square

Wie können wir *unstetige* Lösungen von (5.4) definieren? Die Idee lautet, die Differentialgleichung zu integrieren. Sei $C_0^1(\mathbb{R}^2)$ die Menge aller stetig differenzierbaren Funktionen, die außerhalb einer kompakten Menge verschwinden, und sei $\phi \in C_0^1(\mathbb{R}^2)$. Wir multiplizieren (5.4) mit ϕ und integrieren über \mathbb{R}^2 :

$$0 = \int_0^\infty \int_{\mathbb{R}} (u_t + f(u)_x) \phi \, dx \, dt = - \int_0^\infty \int_{\mathbb{R}} (u \phi_t + f(u) \phi_x) \, dx \, dt - \int_{\mathbb{R}} u(x, 0) \phi(x, 0) \, dx.$$

Im letzten Schritt haben wir partiell integriert. Die Integrale auf der rechten Seite existieren für alle *integrierbaren* Funktionen u ; Differenzierbarkeit für u ist nicht erforderlich. Dies motiviert die folgende Definition.

Definition 5.5 (Schwache Lösung). *Eine Funktion $u : \mathbb{R} \times [0, T) \rightarrow \mathbb{R}$ heißt eine schwache Lösung von (5.4), wenn*

$$\int_0^\infty \int_{\mathbb{R}} (u \phi_t + f(u) \phi_x) \, dx \, dt = - \int_{\mathbb{R}} u(x, 0) \phi(x, 0) \, dx \quad \text{für alle } \phi \in C_0^1(\mathbb{R}^2)$$

erfüllt ist und insbesondere die auftretenden Integrale definiert sind.

Die Lösung in Beispiel 5.4 (2) zur Zeit $t = 1$ lautet

$$u(x, 1) = \begin{cases} 1 & : x < 0 \\ -1 & : x > 0. \end{cases}$$

Wir verwenden diese Funktion als Anfangswert von (5.4). Wie lautet die Lösung dieses Anfangswertproblems? Links von der Unstetigkeitsstelle ist die Straße leer, rechts davon stehen

die Fahrzeuge im Stau. Was geschieht für $t > 1$? Nichts. Es kommen ja von links keine Fahrzeuge nach, und der Stau erstreckt sich rechts bis ins Unendliche. Interessanter ist die folgende Situation:

$$u(x, 1) = \begin{cases} 0 & : x < 0 \\ -1 & : x > 0. \end{cases}$$

Dies bedeutet, dass von links stets Fahrzeuge nachkommen und auf das Stauende treffen. Dadurch verschiebt sich das Stauende nach links. Wir erwarten also, dass die Lösung zur Zeit $t > 1$ die Form

$$u(x, t) = \begin{cases} 0 & : x < st \\ -1 & : x > st \end{cases}$$

hat, wobei s die Geschwindigkeit ist, mit der sich das Stauende nach links ausbreitet. Die Frage ist, welchen Wert s hat. Die Antwort wird im folgenden Satz gegeben.

Satz 5.6. Sei

$$u_0(x) = \begin{cases} u_\ell & : x < 0 \\ u_r & : x \geq 0. \end{cases} \quad (5.5)$$

Dann ist die Funktion

$$u(x, t) = \begin{cases} u_\ell & : x < st \\ u_r & : x \geq st \end{cases} \quad (5.6)$$

eine schwache Lösung von (5.4) und (5.5) genau dann, wenn

$$s = \frac{f(u_\ell) - f(u_r)}{u_\ell - u_r}.$$

Die Zahl s wird die *Schockgeschwindigkeit* genannt. Sie gibt an, wie schnell sich der Schock (Unstetigkeitsstelle) zeitlich ändert. Die obige Beziehung für s wird die *Rankine-Hugoniot-Bedingung* genannt.

Beweis. Sei $\phi \in C_0^1(\mathbb{R}^2)$. Da $u = \text{const.}$ außer auf der Geraden $x = st$, können wir wie folgt rechnen:

$$\begin{aligned} \int_0^\infty \int_{\mathbb{R}} u \phi_t dx dt &= \int_0^\infty \left(\int_{-\infty}^{st} u \phi_t dx + \int_{st}^\infty u \phi_t dx \right) dt \\ &= \int_0^\infty \left(\frac{\partial}{\partial t} \int_{-\infty}^{st} u \phi dx - \underbrace{s u(st-0, t)}_{=u_\ell} \phi(st, t) + \frac{\partial}{\partial t} \int_{st}^\infty u \phi dx + \underbrace{s u(st+0, t)}_{=u_r} \phi(st, t) \right) dt \\ &= \int_0^\infty \frac{\partial}{\partial t} \int_{\mathbb{R}} u \phi dx dt - s(u_\ell - u_r) \int_0^\infty \phi(st, t) dt \\ &= \int_{\mathbb{R}} u \phi dx dt \Big|_0^\infty - s(u_\ell - u_r) \int_0^\infty \phi(st, t) dt \\ &= - \int_{\mathbb{R}} u(x, 0) \phi(x, 0) dx - s(u_\ell - u_r) \int_0^\infty \phi(st, t) dt. \end{aligned}$$

Hier haben wir die Eigenschaft verwendet, dass $\phi = 0$ außerhalb einer kompakten Menge. Wir integrieren nun partiell:

$$\begin{aligned} \int_0^\infty f(u)\phi_x dx dt &= \int_0^\infty \left(- \int_{-\infty}^{st} \underbrace{f(u)_x}_{=0} \phi dx + \underbrace{f(u(st-0,t))}_{=f(u_\ell)} \phi(st,t) \right. \\ &\quad \left. - \int_{st}^\infty \underbrace{f(u)_x}_{=0} \phi dx - \underbrace{f(u(st+0,t))}_{=f(u_r)} \phi(st,t) \right) dt \\ &= (f(u_\ell) - f(u_r)) \int_0^\infty \phi(st,t) dt. \end{aligned}$$

Addieren wir diese Ausdrücke, so erhalten wir

$$\int_0^\infty \int_{\mathbb{R}} (u\phi_t + f(u)\phi_x) dx dt + \int_{\mathbb{R}} u(x,0)\phi(x,0) dx = (f(u_\ell) - f(u_r) - s(u_\ell - u_r)) \int_0^\infty \phi(st,t) dt.$$

Damit die rechte Seite verschwindet, müssen wir s wie im Satz wählen. Dies zeigt, dass u eine schwache Lösung von (5.4) ist. \square

Beispiel 5.7. Wir setzen unser Beispiel für den Verkehrsfluss mit $f(u) = u^2/2$ fort.

(1) Wir betrachten wie in den vorigen Beispielen den Anfangswert

$$u_0(x) = \begin{cases} 1 & : x < 0 \\ -1 & : x > 0. \end{cases}$$

Wie wir bereits oben vermutet haben, ist eine schwache Lösung durch (5.6) mit $s = \frac{1}{2}(u_\ell + u_r) = 0$ gegeben. Das Stauende verändert sich nicht.

(2) Bei dem Anfangswert

$$u_0(x) = \begin{cases} 0 & : x < 0 \\ -1 & : x > 0 \end{cases}$$

kommen von links Fahrzeuge nach. Die Schockgeschwindigkeit $s = \frac{1}{2}(0 - 1) = -\frac{1}{2}$ ist negativ; das Stauende wandert also nach links. Wie erwartet dehnt sich der Stau nach links aus. \square

Der Begriff der schwachen Lösung erlaubt eine sinnvolle Beschreibung unstetiger Lösungen. Sie hat allerdings den Nachteil, dass Verdünnungswellen nicht mehr eindeutig definiert werden können. Sei nämlich der Anfangswert durch

$$u_0(x) = \begin{cases} 0 & : x < 0 \\ 1 & : x > 0 \end{cases} \quad (5.7)$$

gegeben. Dann existieren *zwei* schwache Lösungen; siehe Abbildung 5.19. (Man kann zeigen, dass beide Funktionen wirklich schwache Lösungen sind.) Welche der beiden Lösungen ist die physikalisch sinnvolle?

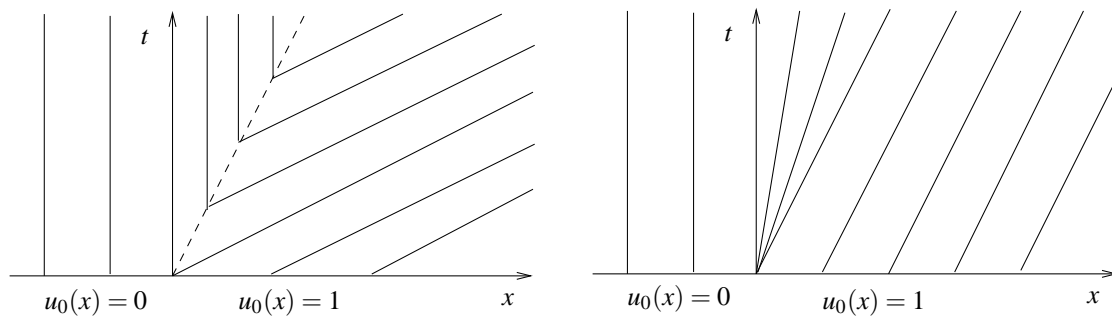


Abbildung 5.19: Zwei schwache Lösungen u_1 (links) und u_2 (rechts) mit dem Anfangswert (5.7).

Zu Beginn ist die Straße in $x < 0$ moderat befahren und leer in $x > 0$. Die Lösung u_1 bedeutet, dass die Fahrzeuge in $x < st$ mit der Geschwindigkeit $v(u_\ell) = 1 - n = 1 - \frac{1}{2}(1 - u_\ell) = \frac{1}{2}$ fahren, während die Fahrzeuge in $x > st$ sich mit der maximalen Geschwindigkeit $v(u_r) = v(1) = 1$ bewegen. An der Geraden $x = st$ ist eine Unstetigkeitsstelle. Dies ist nicht sehr realistisch. Es ist realistischer anzunehmen, dass die Fahrer vor dem Anfangsschock in $x = 0$ so schnell wie möglich fahren, da ja vor ihnen eine leere Straße liegt. Nach einer gewissen Zeit gibt es Fahrzeuge, die mit einer Geschwindigkeit fahren, die zwischen $v(0) = \frac{1}{2}$ und $v(1) = 1$ liegt. Diese Situation entspricht gerade der Funktion u_2 . Wir würden also gerne diese Funktion als die korrekte Lösung zulassen. Wir suchen also eine mathematische Bedingung, die genau diese Lösung auswählt. Eine Antwort ist die *Entropiebedingung*.

Definition 5.8 (Entropiebedingung von Oleinik). *Eine schwache Lösung von (5.4) und (5.5) erfüllt die Entropiebedingung von Oleinik genau dann, wenn entlang jeder Unstetigkeitskurve $x = \psi(t)$*

$$\frac{f(u_\ell(t)) - f(v)}{u_\ell(t) - v} \geq \psi'(t) \geq \frac{f(u_r(t)) - f(v)}{u_r(t) - v}$$

für alle $t > 0$ und alle v zwischen $u_\ell(t)$ und $u_r(t)$ gilt, wobei

$$u_\ell(t) = \lim_{x \rightarrow \psi(t)-0} u(x,t), \quad u_r(t) = \lim_{x \rightarrow \psi(t)+0} u(x,t).$$

Wenn die schwache Lösung stetig ist, ist die Entropiebedingung stets erfüllt, da dann $u_\ell(t) = u_r(t)$. Insbesondere erfüllt u_2 die Entropiebedingung. Für die im obigen Beispiel definierte Funktion u_1 gilt $\psi(t) = \frac{1}{2}t$ sowie $u_\ell(t) = 0$, $u_r(t) = 1$ und damit

$$\frac{u_\ell(t)^2 - v^2}{2(u_\ell(t) - v)} = \frac{v}{2} < \psi'(t) = \frac{1}{2} < \frac{u_r(t)^2 - v^2}{2(u_r(t) - v)} = \frac{1+v}{2}$$

für alle $0 < v < 1$. Also erfüllt u_1 die Entropiebedingung *nicht*.

Beispiel 5.9. Wir wollen die obigen Ergebnisse verwenden, um einen Stau vor einer Ampelanlage zu beschreiben. Wir nehmen an, dass sich zur Zeit $t = 0$ auf der Straße in $x < 0$ Fahrzeuge mit

der konstanten Fahrzeugdichte $\bar{n} = \frac{1}{2}(1 - \bar{u})$ befinden. An der Stelle $x = 0$ sei eine rote Ampel. Hinter der Ampel ($x > 0$) sei die Straße leer, d.h. $u = 1$. Im Laufe der Zeit wird sich ein Stau vor der Ampel bilden. Zur Zeit $t = \omega$ schalte die Ampel auf grün, und die Fahrzeuge fahren von links nach rechts. Wir wollen wissen, was mit dem Stauende zur Zeit $t > \omega$ geschieht.

Rotphase ($0 < t < \omega$): Wir lösen die Burgers-Gleichung im Intervall $(-\infty, 0)$ mit dem Anfangswert $u_0(x) = \bar{u}$, $x < 0$, und dem Randwert $u(0, t) = -1$. Dies modelliert die rote Ampel. Es bildet sich ein Stau, also ist die Lösung dieses Problems gegeben durch

$$u(x, t) = \begin{cases} \bar{u} & : x < st \\ -1 & : x > st, \end{cases}$$

wobei $x < 0$, $0 < t < \omega$, und die Schockgeschwindigkeit ist gegeben durch $s = (u_\ell + u_r)/2 = (\bar{u} - 1)/2$ (siehe Abbildung 5.20). Hinter der Ampel gibt es keine Fahrzeuge: $u(x, t) = 1$ für $x > 0$, $0 < t < \omega$.

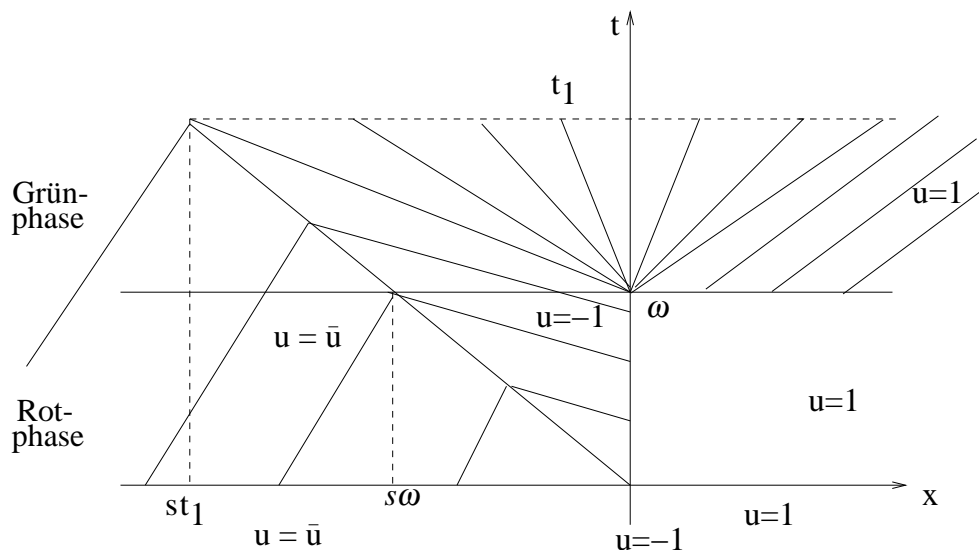


Abbildung 5.20: Charakteristiken im Ampelproblem.

Grünphase ($t \geq \omega$): Wir lösen die Burgers-Gleichung mit Anfangswert $u_0(x) = u(x, \omega)$. Wir erwarten, dass sich entlang $\psi(t) = st$ ein Schock bildet (durch nachfolgende Fahrzeuge, die auf das noch nicht aufgelöste Stauende treffen). Andererseits wird sich wegen der von rot auf grün schaltenden Ampel eine Verdünnungswelle an $x = 0$ bilden. Die Lösung ist also gegeben durch

$$u(x, t) = \begin{cases} \bar{u} & : x < st \\ -1 & : st \leq x < \omega - t \\ x/(t - \omega) & : \omega - t \leq x < t - \omega \\ 1 & : x > t - \omega, \end{cases}$$

wobei $x \in \mathbb{R}$, $t > \omega$. Diese Lösung macht Sinn, solange $st < \omega - t$ oder $t < t_1 := \omega/(s+1) = 2\omega/(\bar{u}+1)$.

Für $t > t_1$ wird die Situation komplizierter und wir erwähnen nur das Ergebnis. Aus der Entropiebedingung können wir die Unstetigkeitslinie berechnen (die jetzt nicht mehr eine Gerade ist):

$$\psi(t) = \bar{u}(t - \omega) - \sqrt{\omega(t - \omega)(1 - \bar{u}^2)}, \quad t \geq t_1.$$

Falls $\bar{u} \leq 0$ (dies entspricht einem relativ dichten Verkehr zu Beginn), so folgt $\psi(t) \rightarrow -\infty$ für $t \rightarrow \infty$. Dies bedeutet, dass sich das Stauende, das durch die Rotphase entstanden ist, weiter nach links ausbreitet, obwohl die Ampel bereits auf grün geschaltet hat. Zwar wird der Unstetigkeitsprung an $\psi(t)$ immer kleiner; dennoch beobachten die Fahrer den Schock eine lange Zeit nach der Rotphase. Dies entspricht der Erfahrung: Eine kurze Verkehrsstörung (Baustelle, rote Ampel, Unfall o.a.) verursacht eine Dichteunstetigkeit, die noch lange nach Beseitigung der Störung beobachtet werden kann. Auf der Website <http://www.juengel.at.vu>, Rubrik Simulations, sind Simulationen des obigen Ampelproblems verfügbar, wobei die Parameter variabel eingestellt werden können. \square