

Analysis II

Banachräume und lineare Operatoren

Frühjahrssemester 2009

Dietmar A. Salamon
ETH-Zürich

18. Februar 2009

1 Äquivalente Normen

Zur Erinnerung: Sei V ein reeller Vektorraum. Eine **Norm** auf V ist eine Abbildung $V \rightarrow \mathbb{R} : v \mapsto \|v\|$ mit folgenden Eigenschaften

- (i) Für alle $v \in V$ gilt $\|v\| \geq 0$ und $\|v\| = 0 \iff v = 0$.
- (ii) Für alle $v \in V$ und alle $\lambda \in \mathbb{R}$ gilt $\|\lambda v\| = |\lambda| \|v\|$.
- (iii) Für alle $v, w \in V$ gilt $\|v + w\| \leq \|v\| + \|w\|$.

Wenn die Abbildung $V \times V \rightarrow \mathbb{R} : (v, w) \mapsto \langle v, w \rangle$ ein inneres Produkt ist so definiert die Formel

$$\|v\| := \sqrt{\langle v, v \rangle}$$

eine Norm auf V und es gilt die Cauchy-Schwarz-Ungleichung

$$|\langle v, w \rangle| \leq \|v\| \|w\|$$

für alle $v, w \in V$. Jeder normierte reelle Vektorraum $(V, \|\cdot\|)$ besitzt eine Metrik $d : V \times V \rightarrow \mathbb{R}$ die durch $d(v, w) := \|v - w\|$ definiert ist. Ein normierter Vektorraum $(V, \|\cdot\|)$ heißt **Banachraum** wenn der zugehörige metrische Raum (V, d) **vollständig** ist, d.h. jede Cauchy-Folge in V konvergiert. (Eine Folge $(v_n)_{n \in \mathbb{N}}$ in V heißt **Cauchy-Folge** wenn es zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ gibt so daß für alle $n, m \in \mathbb{N}$ gilt: $n, m \geq n_0 \implies \|v_n - v_m\| < \varepsilon$.) Eine Teilmenge $U \subset V$ heißt **offen** wenn es für jedes Element $x \in U$ ein $\varepsilon > 0$ gibt so daß

$$B_\varepsilon(x; V, \|\cdot\|) := \{y \in V \mid \|x - y\| < \varepsilon\} \subset U.$$

Eine Abbildung $f : X \rightarrow Y$ zwischen zwei normierten Vektorräumen X und Y ist genau dann stetig wenn, für jede offene Teilmenge $U \subset Y$, auch die Urbildmenge $f^{-1}(U) := \{x \in X \mid f(x) \in U\}$ offen in X ist. (Das gleiche Kriterium für Stetigkeit gilt auch für Abbildungen zwischen beliebigen metrischen Räumen.)

Beispiele. Sei \mathbb{R}^n der euklidische Raum aller Vektoren $x = (x_1, \dots, x_n)$ mit $x_i \in \mathbb{R}$. Für $1 \leq p < \infty$ definiert die Formel

$$\|x\|_p := \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}, \quad x \in \mathbb{R}^n,$$

eine Norm auf \mathbb{R}^n . Für $p = \infty$ definiert die Formel

$$\|x\|_\infty := \max\{|x_1|, \dots, |x_n|\}, \quad x \in \mathbb{R}^n,$$

eine Norm auf \mathbb{R}^n . Eine symmetrische Matrix $A = A^T \in \mathbb{R}^{n \times n}$, mit Koeffizienten $a_{ij} = a_{ji}$ für $i, j = 1, \dots, n$, heißt **positiv definit** wenn

$$x^T A x = \sum_{i,j=1}^n x_i a_{ij} x_j > 0 \quad \forall x \in \mathbb{R}^n \setminus \{0\}.$$

Jede positiv definite $(n \times n)$ -Matrix A definiert ein inneres Product auf \mathbb{R}^n :

$$\langle x, y \rangle_A := x^T A y = \sum_{i,j=1}^n x_i a_{ij} y_j.$$

Die zugehörige Norm eines Vektors $x \in \mathbb{R}^n$ ist gegeben durch

$$\|x\|_A := \sqrt{x^T A x}.$$

Definition. Zwei Normen $\|\cdot\|$ und $\|\cdot\|'$ auf einem reellen Vektorraum V heissen **äquivalent** wenn es eine Konstante $c > 0$ gibt so daß

$$\frac{1}{c} \|v\| \leq \|v\|' \leq c \|v\|, \quad \forall v \in V.$$

Bemerkung. Zwei äquivalente Normen $\|\cdot\|$ und $\|\cdot\|'$ erzeugen dieselben offenen Mengen. Um das zu zeigen wählen wir $c > 0$ wie in der obigen Definition. Ist $U \subset V$ offen bezüglich $\|\cdot\|$ und $x \in U$ so gibt es ein $\varepsilon > 0$ so daß $B_\varepsilon(x; \|\cdot\|) \subset U$. Sei $\delta := \varepsilon/c$. Dann gilt für alle $y \in V$:

$$\begin{aligned} y \in B_\delta(x; \|\cdot\|') &\implies \|x - y\|' < \delta \\ &\implies \|x - y\| \leq c \|x - y\|' < c\delta = \varepsilon \\ &\implies y \in B_\varepsilon(x; \|\cdot\|) \\ &\implies y \in U. \end{aligned}$$

Das heißt: für jedes $x \in U$ haben wir ein $\delta > 0$ gefunden so daß $B_\delta(x; \|\cdot\|') \subset U$; also ist U auch offen bzgl $\|\cdot\|'$. Daß die offenen Teilmengen von V bzgl der beiden Normen dieselben sind, folgt daß alle *topologischen Aussagen in V* (also Aussagen die sich mit Hilfe offener Mengen formulieren lassen, zB Konvergenz, Stetigkeit, Abgeschlossenheit, Kompaktheit) nicht davon abhängen welche der beiden äquivalenten Normen wir dazu verwenden.

Satz 1. Je zwei Normen auf dem euklidischen Raum \mathbb{R}^n sind äquivalent.

Beweis. Sei $\|\cdot\|$ eine (beliebige) Norm auf \mathbb{R}^n . Es genügt zu beweisen, daß $\|\cdot\|$ äquivalent zu $\|\cdot\|_2$ ist. Wir zeigen dies in zwei Schritten.

Schritt 1. Es gibt ein $c > 0$ so daß $\|x\| \leq c \|x\|_2$ für alle $x \in \mathbb{R}^n$.

Die Vektoren $e_i := (0, \dots, 0, 1, 0, \dots, 0)$ (mit der 1 an der i ten Stelle) für $i = 1, \dots, n$ bilden die Standardbasis des \mathbb{R}^n . Sei

$$c := \sqrt{\sum_{i=1}^n \|e_i\|^2} > 0.$$

Dann gilt $x = \sum_{i=1}^n x_i e_i$ für jeden Vektor $x \in \mathbb{R}^n$ und daher

$$\|x\| \leq \sum_{i=1}^n |x_i| \|e_i\| \leq \sqrt{\sum_{i=1}^n |x_i|^2} \sqrt{\sum_{i=1}^n \|e_i\|^2} = c \|x\|_2.$$

Hier folgt die erste Ungleichung aus der Dreiecksungleichung für $\|\cdot\|$ und die zweite Ungleichung aus der Cauchy–Schwarz-Ungleichung für die Euklidische Norm auf \mathbb{R}^n .

Schritt 2. Es gibt ein $\delta > 0$ so daß $\delta \|x\|_2 \leq \|x\|$ für alle $x \in \mathbb{R}^n$.

Die Teilmenge

$$S^{n-1} := \{x \in \mathbb{R}^n \mid \|x\|_2 = 1\}$$

ist abgeschlossen und beschränkt und daher, nach dem Satz von Heine–Borel, kompakt. Wir definieren die Funktion $f : S^{n-1} \rightarrow \mathbb{R}$ durch

$$f(x) := \|x\|.$$

Nach Schritt 1 ist diese Funktion Lipschitz-stetig, denn es gilt

$$|f(x) - f(y)| = |\|x\| - \|y\|| \leq \|x - y\| \leq c \|x - y\|_2$$

für alle $x, y \in S^{n-1}$. Daher ist f stetig. Nach einem Satz aus Analysis I nimmt jede stetige Funktion auf einem kompakten metrischen Raum ihr Minimum an (siehe [3, Kap. V, Satz 7, 6.11.2008]). Es gibt also einen Punkt $x_0 \in S^{n-1}$ so daß

$$f(x_0) \leq f(x) \quad \forall x \in S^{n-1}$$

Sei $\delta := f(x_0) = \|x_0\| > 0$ und $x \in \mathbb{R}^n$ irgendein von Null verschiedener Vektor. Dann gilt $\|x\|_2^{-1} x \in S^{n-1}$ und daher

$$\delta \leq f\left(\frac{x}{\|x\|_2}\right) = \left\|\frac{x}{\|x\|_2}\right\| = \frac{\|x\|}{\|x\|_2}.$$

Hieraus folgt die Behauptung von Schritt 2. □

Definition. Eine Teilmenge $K \subset \mathbb{R}^n$ heißt **konvex** wenn für alle $x, y \in \mathbb{R}^n$ und alle $t \in \mathbb{R}$ folgendes gilt:

$$x, y \in K, \quad 0 \leq t \leq 1 \quad \implies \quad (1-t)x + ty \in K.$$

Übung. Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ eine stetige Funktion. Die offene Teilmenge

$$U := \{(x, y) \in \mathbb{R}^2 \mid y > f(x)\}$$

ist genau dann konvex wenn die Funktion f konvex ist

Übung. Sei $U \subset \mathbb{R}^n$ eine offene, beschränkte, konvexe Teilmenge so daß

$$x \in U \iff -x \in U.$$

(Nach Satz 1 hängt der Begriff “beschränkte Teilmenge von \mathbb{R}^n ” nicht ab von der Norm welche für die Definition verwendet wird.) Dann definiert die Formel

$$\|x\| := \inf \{ \lambda > 0 \mid \lambda^{-1}x \in U \}$$

eine Norm auf \mathbb{R}^n und es gilt

$$U = \{x \in \mathbb{R}^n \mid \|x\| < 1\}.$$

Übung. Auf dem Raum $\mathcal{C}([0, 1])$ aller stetigen Funktionen $f : [0, 1] \rightarrow \mathbb{R}$ hatten wir die Normen

$$\|f\|_\infty := \sup_{0 \leq t \leq 1} |f(t)|, \quad \|f\|_p := \left(\int_0^1 |f(t)|^p \right)^{1/p}$$

für $1 \leq p < \infty$ definiert. Zeigen Sie, daß die Normen $\|\cdot\|_p$ und $\|\cdot\|_q$ für $p \neq q$ nicht äquivalent sind. Für welche p und q gibt es eine Konstante $c = c(p, q) > 0$ so daß $\|f\|_p \leq c \|f\|_q$ für alle $f \in \mathcal{C}([0, 1])$?

2 Die Operator-Norm

Definition. Seien X, Y normierte reelle Vektorräume. Ein linearer Operator¹ $T : X \rightarrow Y$ heißt **beschränkt**, wenn es ein $c > 0$ gibt so daß

$$\|Tx\|_Y \leq c \|x\|_X \quad \forall x \in X.$$

¹Der Begriff “Operator” wird in der mathematischen Literatur oft gleichbedeutend mit dem Begriff “Abbildung” verwendet. Die Herkunft dieser Wortwahl liegt darin begründet, daß viele der wichtigsten motivierenden Beispiele *Differentialoperatoren* oder *Integraloperatoren* aus der mathematischen Physik sind, für deren Studium die Theorie der linearen Abbildungen entwickelt wurde. Dies führt hin zum Gebiet der “*Funktionalanalysis*” welches wir in dieser Vorlesung nur am Rande berühren.

Die kleinste solche Konstante c heißt **Norm** (bzw **Operatornorm**) von T und wird mit

$$\|T\| := \sup_{x \neq 0} \frac{\|Tx\|_Y}{\|x\|_X}$$

bezeichnet. (Hier ist das Supremum über alle von Null verschiedenen Vektoren $x \in X$ zu verstehen.) Die Menge der beschränkten linearen Operatoren von X nach Y bezeichnen wir mit $\mathcal{L}(X, Y)$.

Beispiel. Seien $\|\cdot\|_1$ und $\|\cdot\|_2$ zwei Normen auf einem reellen Vektorraum V . Diese beiden Normen sind genau dann äquivalent wenn die beiden linearen Abbildungen $\text{id} : (V, \|\cdot\|_1) \rightarrow (V, \|\cdot\|_2)$ und $\text{id} : (V, \|\cdot\|_2) \rightarrow (V, \|\cdot\|_1)$ beschränkt sind. **Übung:** Zeigen Sie mit Hilfe von Lemma 2, daß die durch $\|\cdot\|_1$ und $\|\cdot\|_2$ bestimmten offenen Teilmengen von V übereinstimmen. Zeigen Sie mit Hilfe von Satz 1 und dem folgenden Lemma 2, daß die Topologie auf dem \mathbb{R}^n (und damit auf jedem endlichdimensionalen Vektorraum) unabhängig von der Wahl der Norm ist.

Lemma 2. *Seien X, Y normierte Vektorräume und $T : X \rightarrow Y$ ein linearer Operator. Dann sind folgende Aussagen äquivalent.*

- (i) T ist beschränkt.
- (ii) T ist stetig.
- (iii) T ist stetig and der Stelle 0.

Beweis. “(i) \implies (ii)”. Wenn T beschränkt ist, so ist T auch Lipschitz-stetig:

$$\|Tx_0 - Tx_1\|_Y = \|T(x_0 - x_1)\|_Y \leq \|T\| \|x_0 - x_1\|_X.$$

Also ist T stetig.

“(ii) \implies (iii)”. Dies folgt sofort aus der Definition der Stetigkeit.

“(iii) \implies (i)”. Ist T stetig an der Stelle 0, so folgt aus der Definition der Stetigkeit mit $\varepsilon = 1$, daß es ein $\delta > 0$ gibt so daß für alle $x \in X$ folgendes gilt:

$$\|x\|_X \leq \delta \quad \implies \quad \|Tx\|_Y \leq 1.$$

Nun sei $x \in X$ von Null verschieden. Dann gilt

$$\left\| \frac{\delta}{\|x\|_X} x \right\|_X = \delta$$

und daher

$$1 \geq \left\| \frac{\delta}{\|x\|_X} Tx \right\|_Y = \frac{\delta}{\|x\|_X} \|Tx\|_Y.$$

Hieraus folgt

$$\|Tx\|_Y \leq \frac{1}{\delta} \|x\|_X$$

für alle $x \in X$. Also ist T beschränkt. □

Lemma 3. $\mathcal{L}(X, Y)$ ist ein normierter Vektorraum.

Beweis. Folgende drei Aussagen sind zu beweisen.

- (1) Für alle $T \in \mathcal{L}(X, Y)$ gilt $\|T\| = 0 \implies T = 0$.
- (2) Für alle $T \in \mathcal{L}(X, Y)$ und $\lambda \in \mathbb{R}$ gilt $\lambda T \in \mathcal{L}(X, Y)$ und $\|\lambda T\| = |\lambda| \|T\|$.
- (3) Für alle $S, T \in \mathcal{L}(X, Y)$ gilt $S + T \in \mathcal{L}(X, Y)$ und $\|S + T\| \leq \|S\| + \|T\|$.

Die Aussagen (1) und (2) folgen unmittelbar aus den Definitionen. Wir beweisen (3). Seien $S, T : X \rightarrow Y$ beschränkte lineare Operatoren. Dann gilt für alle $x \in X$:

$$\begin{aligned} \|(S + T)x\|_Y &= \|Sx + Tx\|_Y \\ &\leq \|Sx\|_Y + \|Tx\|_Y \\ &\leq \|S\| \|x\|_X + \|T\| \|x\|_X \\ &= (\|S\| + \|T\|) \|x\|_X. \end{aligned}$$

Also ist $S + T$ beschränkt und für $x \in X \setminus \{0\}$ gilt

$$\frac{\|(S + T)x\|_Y}{\|x\|_X} \leq \|S\| + \|T\|.$$

Damit folgt die Dreiecksungleichung

$$\|S + T\| \leq \|S\| + \|T\|$$

aus der Definition der Operatornorm. □

Lemma 4. Ist Y ein Banachraum so ist auch $\mathcal{L}(X, Y)$ ein Banachraum.

Beweis. Es ist zu zeigen, daß $\mathcal{L}(X, Y)$ vollständig ist. Sei also $(T_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge in $\mathcal{L}(X, Y)$. Dann zeigt die Ungleichung

$$\|T_n x - T_m x\|_Y \leq \|T_n - T_m\| \|x\|_X,$$

daß die Folge $(T_n x)_{n \in \mathbb{N}}$ für jedes $x \in X$ eine Cauchy-Folge in Y ist. Da Y vollständig ist, konvergiert diese Folge und wir definieren $T : X \rightarrow Y$ durch

$$Tx := \lim_{n \rightarrow \infty} T_n x$$

für $x \in X$. Wir werden beweisen, daß T ein beschränkter linearer Operator ist und daß die Folge T_n in der Operatornorm gegen T konvergiert.

T ist linear. Für $x_0, x_1 \in X$ gilt

$$T(x_0 + x_1) = \lim_{n \rightarrow \infty} T_n(x_0 + x_1) = \lim_{n \rightarrow \infty} T_n x_0 + \lim_{n \rightarrow \infty} T_n x_1 = Tx_0 + Tx_1.$$

Genauso zeigt man, daß $T(\lambda x) = \lambda Tx$ ist für alle $\lambda \in \mathbb{R}$ und $x \in X$.

T ist beschränkt. Zunaechst folgt aus der Ungleichung

$$|\|T_n\| - \|T_m\|| \leq \|T_n - T_m\|,$$

daß die Folge $(\|T_n\|)_{n \in \mathbb{N}}$ der Normen eine Cauchy-Folge reeller Zahlen ist und daher konvergiert. Sei

$$c := \lim_{n \rightarrow \infty} \|T_n\|.$$

Dann erfüllt jedes $x \in X$ die Ungleichung

$$\|Tx\|_Y = \left\| \lim_{n \rightarrow \infty} T_n x \right\|_Y = \lim_{n \rightarrow \infty} \|T_n x\|_Y \leq \lim_{n \rightarrow \infty} \|T_n\| \|x\|_X = c \|x\|_X.$$

Daher ist T beschränkt und $\|T\| \leq c$.

T_n konvergiert gegen T . Sei $\varepsilon > 0$. Da T_n eine Cauchy-Folge ist, gibt es ein $n_0 \in \mathbb{N}$ so daß für alle $n, m \in \mathbb{N}$ gilt:

$$m > n \geq n_0 \quad \implies \quad \|T_n - T_m\| < \varepsilon.$$

Hieraus folgt die Ungleichung $\|T_n x - T_m x\|_Y < \varepsilon \|x\|_X$ für alle $x \in X$ und $m > n \geq n_0$ und daher

$$\|T_n x - Tx\|_Y = \lim_{m \rightarrow \infty} \|T_n x - T_m x\|_Y \leq \varepsilon \|x\|_X$$

für alle $x \in X$ und $n \geq n_0$. Teilen wir durch $\|x\|_X$ im Fall $x \neq 0$ so ergibt sich

$$\|T_n - T\| = \sup_{x \neq 0} \frac{\|T_n x - Tx\|_Y}{\|x\|_X} \leq \varepsilon$$

für alle $n \geq n_0$. Damit ist das Lemma bewiesen. \square

Der Dualraum. Ein interessanter Spezialfall tritt auf mit $Y = \mathbb{R}$. In diesem Fall nennen wir

$$X^* := \mathcal{L}(X, \mathbb{R})$$

den **Dualraum** des normierten Vektorraumes X . Lemma 4 zeigt, daß dieser Dualraum immer vollständig ist, auch dann wenn X selbst nicht vollständig ist.

Übung. Sei X ein endlichdimensionaler normierter Vektorraum. Beweisen Sie:

- (a) X ist vollständig.
- (b) Eine Teilmenge $K \subset X$ ist genau dann kompakt wenn sie abgeschlossen und beschränkt ist.
- (c) Jede lineare Abbildung $\phi : X \rightarrow \mathbb{R}$ ist beschränkt.
- (d) Der Dualraum $X^* = \mathcal{L}(X, \mathbb{R})$ ist endlichdimensional und $\dim X^* = \dim X$.

Übung. Seien $1 \leq p, q \leq \infty$ mit $1/p + 1/q = 1$. Konstruieren Sie einen Isomorphismus

$$\mathbb{R}^n \rightarrow (\mathbb{R}^n)^* : y \mapsto \phi_y$$

so daß

$$\|y\|_q = \sup_{x \neq 0} \frac{|\phi_y(x)|}{\|x\|_p}$$

für alle $y \in \mathbb{R}^n$.

3 Banachalgebren

Ein zweiter Spezialfall von Lemma 4 ergibt sich mit $Y = X$. In diesem Fall setzen wir voraus, daß X vollständig ist, so daß dann auch der Raum

$$\mathcal{L}(X) := \mathcal{L}(X, X)$$

ein Banachraum ist. Dieser Raum besitzt eine Produktstruktur $(S, T) \mapsto S \circ T$, die durch die Komposition gegeben ist. Diese Produktstruktur ist im allgemeinen nicht kommutativ, und es gibt auch nicht für jeden von Null verschiedenen Operator $T \in \mathcal{L}(X)$ ein inverses Element; also ist $\mathcal{L}(X)$ kein Körper; dieser Raum erfüllt aber alle anderen Eigenschaften eines Körpers; einen solchen Raum nennt man eine Algebra. Der Begriff einer Banachalgebra ist nützlich für das Rechnen mit inversen Operatoren und Exponentialmatrizen.

Definition. Eine **reelle Algebra** besteht aus einem reellen Vektorraum \mathcal{A} , einer bilinearen Abbildung²

$$\mathcal{A} \times \mathcal{A} \rightarrow \mathcal{A} : (x, y) \mapsto xy,$$

(das **Produkt** genannt) und einem Element $\mathbb{1} \in \mathcal{A}$ (das **1-Element** genannt) welche folgende Axiome erfüllen.

- (i) Das Produkt ist assoziativ: $(xy)z = x(yz)$ für alle $x, y, z \in \mathcal{A}$.
- (ii) Für jedes $x \in \mathcal{A}$ gilt $\mathbb{1}x = x\mathbb{1} = x$.

Eine **reelle normierte Algebra** besteht aus einer reellen Algebra \mathcal{A} und einer Norm $\mathcal{A} \rightarrow [0, \infty) : x \mapsto \|x\|$ welche die Produkt-Ungleichung

$$\|xy\| \leq \|x\| \|y\| \tag{1}$$

für alle $x, y \in \mathcal{A}$ erfüllt. Eine **reelle Banachalgebra** ist eine reelle normierte Algebra in der jede Cauchy-Folge konvergiert (das heißt, als normierter Vektorraum betrachtet ist \mathcal{A} vollständig, also ein Banachraum).

Beispiel 1. Der Raum $\mathbb{R}^{n \times n}$ der quadratischen Matrizen ist eine Banachalgebra mit jeder Norm welche der Bedingung (1) genügt. Eine solche Norm auf dem Raum $\mathbb{R}^{n \times n}$ wird **Matrixnorm** genannt. Zum Beispiel kann dies die zu einer beliebigen Norm auf dem \mathbb{R}^n gehörige Operatornorm auf $\mathbb{R}^{n \times n}$ sein. Ein anderes Beispiel wäre die euklidische Norm die sich durch das Identifizieren vom $\mathbb{R}^{n \times n}$ mit dem euklidischen Raum \mathbb{R}^{n^2} ergibt; in diesem Fall erhält man

$$\|A\|_2 = \sqrt{\sum_{i,j=1}^n a_{ij}^2}$$

für $A = (a_{ij}) \in \mathbb{R}^{n \times n}$. **Übung:** Diese Norm erfüllt die Bedingung (1) sowie $\|Ax\|_2 \leq \|A\|_2 \|x\|_2$ für jede Matrix $A \in \mathbb{R}^{n \times n}$ und jeden Vektor $x \in \mathbb{R}^n$.

² "Bilinear" heißt: die Abbildungen $\mathcal{A} \rightarrow \mathcal{A} : x \mapsto xy$ und $\mathcal{A} \rightarrow \mathcal{A} : x \mapsto yx$ sind beide linear für jedes $y \in \mathcal{A}$.

Beispiel 2. Ist X ein Banachraum so ist $\mathcal{L}(X)$ mit der Operatornorm eine Banachalgebra, nach Lemma 4. **Übung:** Die Operatornorm auf $\mathcal{L}(X)$ erfüllt die Bedingung (1).

Beispiel 3. Der Raum $\mathcal{C}([0, 1])$ der stetigen, reellwertigen Funktionen auf dem Intervall $[0, 1]$ ist eine Banachalgebra mit der Supremumsnorm. Gleiches gilt für den Raum $\mathcal{C}(X)$ der beschränkten, stetigen und reellwertigen Funktionen auf einem beliebigen metrischen Raum (X, d) .

Beispiel 4. Eine Folge $x = (x_0, x_1, x_2, \dots)$ reeller Zahlen heißt **absolut summierbar** wenn die Reihe $\sum_{k=0}^{\infty} x_k$ absolut konvergiert. Die Norm einer solchen absolut summierbaren Folge ist definiert durch

$$\|x\| := \sum_{k=0}^{\infty} |x_k|.$$

Wir bezeichnen mit ℓ^1 den Raum aller absolut summierbaren Folgen $(x_k)_{k \in \mathbb{N}}$ reeller Zahlen. Der Vektorraum ℓ^1 ist mit dieser Norm ein Banachraum. Die **Faltung** zweier Folgen $x, y \in \ell^1$ ist definiert durch

$$(x * y)_k := \sum_{i=0}^k x_i y_{k-i},$$

Diese Folge ist ebenfalls absolut summierbar und es gilt

$$\|x * y\| \leq \|x\| \|y\|. \quad (2)$$

(Dies folgt aus dem Beweis von Satz 9 in [3, Kap. IV, 29.10.2008].) Damit ist ℓ^1 eine Banachalgebra. **Übung:** Beweisen Sie, daß ℓ^1 ein Banachraum ist. Zeigen Sie daß das Faltungsprodukt bilinear und assoziativ ist.

4 Potenzreihen in Banachalgebren

Die Begriffe von Konvergenz und Stetigkeit hatten wir für beliebige metrische Räume eingeführt, sie gelten also insbesondere für Funktionen und Folgen mit Werten in einem Banachraum oder einer Banachalgebra. Aufgrund der Vektorraumstruktur lässt sich der Begriff der Konvergenz einer Reihe auch sofort auf Reihen mit Werten in Banachräumen übertragen: Ist $(X, \|\cdot\|)$ ein Banachraum und $(x_k)_{k \in \mathbb{N}}$ eine Folge in X , so nennen wir die Reihe $\sum_{k=1}^{\infty} x_k$ **konvergent**, wenn die Folge der **Partialsommen**

$$s_n := \sum_{k=1}^n x_k$$

konvergiert; in dem Fall bezeichnen wir den Limes mit

$$\sum_{k=1}^{\infty} x_k := \lim_{n \rightarrow \infty} s_n.$$

Wir benutzen also die gleiche Bezeichnungsweise wie für Reihen reeller oder komplexer Zahlen, d.h. der Ausdruck “ $\sum_{k=1}^{\infty} x_k$ ” hat zwei Bedeutungen: er steht zum einen für die Folge der Partialsummen $(s_n)_{n \in \mathbb{N}}$ und zum anderen für deren Grenzwert, falls dieser existiert. Wir nennen die Reihe $\sum_{k=1}^{\infty} x_k$ **absolut konvergent**, wenn die Reihe $\sum_{k=1}^{\infty} \|x_k\|$ konvergiert.

Lemma 5. Sei $(X, \|\cdot\|)$ ein Banachraum und $(x_k)_{k \in \mathbb{N}}$ eine Folge in X . Konvergiert die Reihe $\sum_{k=1}^{\infty} x_k$ absolut, so konvergiert sie.

Beweis. Sei $s_n := \sum_{k=1}^n x_k$ die Folge der Partialsummen und $\varepsilon > 0$. Da die Reihe $\sum_{k=1}^{\infty} x_k$ absolut konvergiert, konvergiert die Reihe $\sum_{k=1}^{\infty} \|x_k\|$ in \mathbb{R} . Das heißt, die Folge der Partialsummen

$$\sigma_n := \sum_{k=1}^n \|x_k\|$$

der Reihe der Normen konvergiert und ist daher eine Cauchy-Folge. Also gibt es eine natürliche Zahl $n_0 \in \mathbb{N}$ so daß für alle $m, n \in \mathbb{N}$ folgendes gilt:

$$m > n \geq n_0 \quad \implies \quad \sum_{k=n+1}^m \|x_k\| = \sigma_m - \sigma_n < \varepsilon.$$

Es folgt nun aus der Dreiecksungleichung, daß

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

für $m > n \geq n_0$. Da $\varepsilon > 0$ beliebig gewählt war, haben wir gezeigt, daß $(s_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge in X ist. Da X ein Banachraum (und somit vollständig) ist, konvergiert diese Folge, und das heißt genau, daß die Reihe $\sum_{k=1}^{\infty} x_k$ konvergiert, wie behauptet. \square

Sei \mathcal{A} eine reelle Banachalgebra und

$$P(z) := \sum_{k=0}^{\infty} a_k z^k \tag{3}$$

eine Potenzreihe mit reellen Koeffizienten $a_k \in \mathbb{R}$ und Konvergenzradius

$$\rho := \frac{1}{\limsup_{n \rightarrow \infty} |a_n|^{1/n}}.$$

Wir nehmen an, dieser Konvergenzradius sei positiv (oder sogar unendlich). Wir wollen nun die Variable z durch ein Element der Banachalgebra \mathcal{A} ersetzen. Formal macht dies Sinn, da ja in einer Banachalgebra die Produkte x^k erklärt sind. Wir bezeichnen die daraus resultierende Reihe in \mathcal{A} mit

$$P_{\mathcal{A}}(x) := \sum_{k=0}^{\infty} a_k x^k. \tag{4}$$

Man kann nun die Frage stellen, für welche $x \in \mathcal{A}$ diese Reihe konvergiert.

Korollar. Die Reihe (4) konvergiert für jedes $x \in \mathcal{A}$ mit $\|x\| < \rho$.

Beweis. Es folgt aus der Bedingung (1) in der Definition einer Banachalgebra, daß $\|a_k x^k\| \leq |a_k| \|x\|^k$. Da $\|x\| < \rho$ ist, konvergiert die Reihe $\sum_{k=0}^{\infty} |a_k| \|x\|^k$. Nach dem Majorantenkriterium konvergiert daher auch die Reihe $\sum_{k=0}^{\infty} \|a_k x^k\|$. Dies heißt aber genau, daß die Reihe $\sum_{k=0}^{\infty} a_k x^k$ in \mathcal{A} absolut konvergiert und somit, nach Lemma 5, konvergiert. \square

Wir haben also gezeigt, daß die Formel (4) eine Abbildung

$$P_{\mathcal{A}} : \mathcal{B}_{\rho} := \{x \in \mathcal{A} \mid \|x\| < \rho\} \rightarrow \mathcal{A} \quad (5)$$

definiert, wobei ρ der Konvergenzradius von P ist. Das nächste Lemma zeigt, daß diese Abbildung stetig ist. Zunächst notieren wir noch, daß sich der Begriff der Differenzierbarkeit, und der der Ableitung, auf Funktionen mit Werten in einem Banachraum (Wort für Wort) übertragen lässt.

Definition. Sei $I \subset \mathbb{R}$ ein Intervall und $(X, \|\cdot\|)$ ein Banachraum. Eine Funktion $f : I \rightarrow X$ heißt **differenzierbar an der Stelle $t \in I$ mit Ableitung $a =: f'(t) \in X$** wenn gilt: $\forall \varepsilon > 0 \exists \delta > 0 \forall h \in \mathbb{R}$:

$$0 < |h| < \delta, \quad t + h \in I \quad \implies \quad \frac{\|f(t+h) - f(t) - ha\|}{|h|} < \varepsilon.$$

Äquivalenterweise heißt das, daß der Differenzenquotient $h^{-1}(f(t+h) - f(t))$ in X gegen a konvergiert (für $h \rightarrow 0$). Mit anderen Worten

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

falls dieser Limes existiert, und wenn er existiert so nennen wir f differenzierbar an der Stelle t .

Lemma 6. Sei \mathcal{A} eine reelle Banachalgebra und (3) eine Potenzreihe mit reellen Koeffizienten und Konvergenzradius $\rho > 0$. Dann gilt folgendes

(i) Für jedes $r < \rho$ ist die Restriktion der Abbildung (5) auf die Teilmenge $\overline{\mathcal{B}}_r := \{x \in \mathcal{A} \mid \|x\| \leq r\}$ Lipschitz-stetig; es gilt

$$\|P_{\mathcal{A}}(x) - P_{\mathcal{A}}(y)\| \leq c_r \|x - y\|, \quad c_r := \sum_{k=1}^{\infty} k |a_k| r^{k-1},$$

für $x, y \in \mathcal{A}$ mit $\|x\|, \|y\| \leq r$.

(ii) Für jedes $x \in \mathcal{A}$ ist die Funktion $t \mapsto P_{\mathcal{A}}(tx)$ auf dem offenen Intervall $(-\rho/\|x\|, \rho/\|x\|)$ differenzierbar und es gilt

$$\frac{d}{dt} P_{\mathcal{A}}(tx) = x Q_{\mathcal{A}}(tx), \quad Q := P'. \quad (6)$$

Beweis. Wir beweisen (i). Es gilt

$$x^k - y^k = \sum_{j=0}^{k-1} x^{k-1-j} (x-y) y^j$$

für jedes $k \in \mathbb{N}$ und damit

$$\|x\| \leq r, \|y\| \leq r \quad \implies \quad \|x^k - y^k\| \leq kr^{k-1} \|x - y\|.$$

Hieraus folgt

$$\|P_{\mathcal{A}}(x) - P_{\mathcal{A}}(y)\| \leq \sum_{k=0}^{\infty} |a_k| \|x^k - y^k\| \leq \sum_{k=0}^{\infty} k |a_k| r^{k-1} \|x - y\| = c_r \|x - y\|$$

für $x, y \in \mathcal{A}$ mit $\|x\|, \|y\| \leq r$. Damit ist (i) bewiesen.

Zum Beweis von (ii) beachten wir, daß $Q_{\mathcal{A}}(x) = \sum_{k=1}^{\infty} k a_k x^{k-1}$ und daher

$$\begin{aligned} & \|P_{\mathcal{A}}((t+h)x) - P_{\mathcal{A}}(tx) - hxQ_{\mathcal{A}}(tx)\| \\ &= \left\| \sum_{k=2}^{\infty} ((t+h)^k - t^k - kht^{k-1}) a_k x^k \right\| \\ &\leq \sum_{k=2}^{\infty} |(t+h)^k - t^k - kht^{k-1}| |a_k| \|x\|^k \\ &\leq \sum_{k=2}^{\infty} \sum_{j=2}^k \binom{k}{j} |h|^j |t|^{k-j} |a_k| \|x\|^k \\ &= \sum_{k=2}^{\infty} \left((|t|+|h|)^k - |t|^k - k|h||t|^{k-1} \right) |a_k| \|x\|^k \\ &= |p(|t|+|h|) - p(|t|) - |h|\dot{p}(|t|)|, \end{aligned}$$

wobei p die folgende Potenzreihe ist:

$$p(\lambda) := \sum_{k=0}^{\infty} |a_k| \|x\|^k \lambda^k.$$

Diese Potenzreihe hat Konvergenzradius $\rho/\|x\|$. Daher gibt es für $|t| < \rho/\|x\|$ zwei Konstanten $\delta > 0$ und $c > 0$ so daß für alle $h \in \mathbb{R}$ folgendes gilt

$$|h| < \delta \quad \implies \quad |p(|t|+|h|) - p(|t|) - |h|\dot{p}(|t|)| \leq c|h|^2.$$

Damit gilt auch

$$\|P_{\mathcal{A}}((t+h)x) - P_{\mathcal{A}}(tx) - hxQ_{\mathcal{A}}(tx)\| \leq c|h|^2$$

für jedes $h \in \mathbb{R}$ mit $|h| < \delta$. Damit haben wir gezeigt, daß

$$\lim_{h \rightarrow 0} \left\| \frac{P_{\mathcal{A}}((t+h)x) - P_{\mathcal{A}}(tx)}{h} - xQ_{\mathcal{A}}(tx) \right\| = 0$$

und daraus folgt (ii). □

Betrachten wir den Spezialfall $\mathcal{A} := \mathbb{R}^{n \times n}$ mit einer beliebigen Matrixnorm $\|\cdot\|$. Dann folgt aus Lemma 6, daß die Matrix-Reihe $P(A) := \sum_{k=0}^{\infty} a_k A^k$ für jede Matrix $A \in \mathbb{R}^{n \times n}$ mit $\|A\| < \rho$ in dieser Norm konvergiert. Während die Schlußfolgerung (Konvergenz) nach Satz 1 unabhängig von der Wahl der Norm ist, benötigen wir bei der Voraussetzung (dh. $\|A\| < \rho$) eine Matrixnorm.

5 Die geometrische Reihe

Definition. Sei \mathcal{A} eine Banachalgebra. Ein Element $x \in \mathcal{A}$ heißt **invertierbar**, wenn es ein Element $y \in \mathcal{A}$ gibt so daß $xy = yx = \mathbb{1}$. Ist x invertierbar, so ist dieses Element y durch x eindeutig bestimmt; es heißt **Inverse** von x und wird mit x^{-1} bezeichnet. Die Teilmenge der invertierbaren Elemente bildet eine multiplikative Gruppe

$$\mathcal{A}^* := \{x \in \mathcal{A} \mid x \text{ ist invertierbar}\}.$$

Beispiel. Wir betrachten den Spezialfall

$$\mathcal{A} = \mathbb{R}^{n \times n}.$$

Für eine Matrix $A \in \mathbb{R}^{n \times n}$ bezeichnen wir mit $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ die durch A induzierte lineare Abbildung. Die Matrix A ist invertierbar (im obigen Sinne als Element von \mathcal{A}) genau dann wenn die lineare Abbildung T_A bijektiv ist und die inverse Matrix A^{-1} ist die, welche die Umkehrabbildung $T_{A^{-1}} = T_A^{-1}$ induziert. Aus der *Linearen Algebra* wissen wir, daß die invertierbaren Matrizen gerade die mit Determinante ungleich Null sind:

$$\mathcal{A}^* = \{A \in \mathbb{R}^{n \times n} \mid \det(A) \neq 0\} =: \text{GL}(n, \mathbb{R}).$$

Übung. Sind $x, y \in \mathcal{A}^*$ so ist auch $xy \in \mathcal{A}^*$ und $(xy)^{-1} = y^{-1}x^{-1}$.

Geometrische Reihe. Wir betrachten die Potenzreihe

$$G(z) := \sum_{k=0}^{\infty} z^k$$

mit Konvergenzradius $\rho = 1$. Es gilt $G(z) = (1 - z)^{-1}$ für $z \in \mathbb{C}$ mit $|z| < 1$. Diese Formel gilt auch in einer beliebigen Banachalgebra \mathcal{A} . Mit anderen Worten, ist $x \in \mathcal{A}$ mit $\|x\| < 1$ so ist $\mathbb{1} - x \in \mathcal{A}^*$ und

$$(\mathbb{1} - x)^{-1} = \sum_{k=0}^{\infty} x^k. \tag{7}$$

Dies folgt aus den Formeln

$$(1 - x)G_{\mathcal{A}}(x) = G_{\mathcal{A}}(x)(1 - x) = \sum_{k=0}^{\infty} x^k - \sum_{k=0}^{\infty} x^{k+1} = \mathbb{1}$$

mit $G_{\mathcal{A}}(x) := \sum_{k=0}^{\infty} x^k$ für $\|x\| < 1$. Die Gleichung (7) ist besonders nützlich für das Invertieren einer Matrix.

Satz 7. Sei \mathcal{A} eine Banachalgebra. Dann ist die Menge $\mathcal{A}^* \subset \mathcal{A}$ offen und die Abbildung $\mathcal{A}^* \rightarrow \mathcal{A}^* : x \mapsto x^{-1}$ ist stetig.

Beweis. Sei $a \in \mathcal{A}^*$ und $x \in \mathcal{A}$ mit

$$\|x - a\| < \varepsilon := \frac{1}{\|a^{-1}\|}.$$

Dann gilt

$$\|\mathbb{1} - a^{-1}x\| \leq \|a^{-1}\| \|a - x\| < \|a^{-1}\| \varepsilon = 1.$$

Hieraus folgt $a^{-1}x \in \mathcal{A}^*$ und damit auch $x \in \mathcal{A}^*$ und

$$x^{-1} = (a^{-1}x)^{-1}a^{-1} = \sum_{k=0}^{\infty} (\mathbb{1} - a^{-1}x)^k a^{-1}.$$

Damit erhalten wir

$$\begin{aligned} \|x^{-1} - a^{-1}\| &= \left\| \sum_{k=1}^{\infty} (\mathbb{1} - a^{-1}x)^k a^{-1} \right\| \\ &\leq \|a^{-1}\| \sum_{k=1}^{\infty} \|\mathbb{1} - a^{-1}x\|^k \\ &= \|a^{-1}\| \frac{\|\mathbb{1} - a^{-1}x\|}{1 - \|\mathbb{1} - a^{-1}x\|} \\ &\leq \frac{\|a^{-1}\|^2 \|x - a\|}{1 - \|a^{-1}\| \|x - a\|} \end{aligned}$$

für jedes $x \in \mathcal{A}$ mit $\|x - a\| < \varepsilon$. Damit ist die Inversen-Abbildung stetig an der Stelle a . Da $a \in \mathcal{A}^*$ beliebig gewählt war, ist die Inversen-Abbildung stetig. \square

6 Die Exponentialabbildung

Sei \mathcal{A} eine Banachalgebra. Die Potenzreihe $\exp(z) := \sum_{k=0}^{\infty} z^k/k!$ hat den Konvergenzradius $\rho = \infty$ und induziert damit eine Abbildung $\mathcal{A} \rightarrow \mathcal{A} : x \mapsto e^x$ die auf ganz \mathcal{A} definiert ist; und zwar durch die Formel

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

Man zeigt wie im Fall $\mathcal{A} = \mathbb{C}$, daß sich die Exponentialabbildung in der Form

$$e^x = \lim_{n \rightarrow \infty} \left(\mathbb{1} + \frac{x}{n} \right)^n$$

ausdrücken lässt (siehe [3, Kap. V, Lemma 8, 10.11.2008]). Ebenfalls wie im Fall $\mathcal{A} = \mathbb{C}$ (siehe [3, Kap. IV, Satz 8, 29.10.2008]) zeigt man für $x, y \in \mathcal{A}$:

$$xy = yx \quad \implies \quad e^{x+y} = e^x e^y.$$

Damit ist das Element e^x stets invertierbar mit inversem Element $(e^x)^{-1} = e^{-x}$. Es folgt, daß die Abbildung

$$\mathbb{R} \rightarrow \mathcal{A}^* : t \mapsto e^{tx} \quad (8)$$

für jedes $x \in \mathcal{A}$ ein Gruppenhomomorphismus ist. Nach Lemma 6 ist die Abbildung (8) differenzierbar und es gilt

$$\frac{d}{dt} e^{tx} = x e^{tx}. \quad (9)$$

Übung (Komposition von Potenzreihen). Seien

$$P = \sum_{k=1}^{\infty} a_k x^k, \quad Q = \sum_{\ell=0}^{\infty} b_\ell y^\ell$$

zwei Potenzreihen mit reellen Koeffizienten und positiven Konvergenzradien ρ_P und ρ_Q . Wähle eine positive Zahl $\rho < \rho_P$ so daß $\sum_{k=1}^{\infty} |a_k| \rho^k < \rho_Q$. Dann gilt $|P(z)| < \rho_Q$ für jedes $z \in \mathbb{C}$ mit $|z| \leq \rho$. Wir bezeichnen mit $Q \circ P$ die durch $Q \circ P(z) := Q(P(z))$ für $|z| \leq \rho$ definierte Potenzreihe. Ihr Konvergenzradius ist grösser als ρ . Beweisen Sie die Formel

$$(Q \circ P)_{\mathcal{A}}(x) = Q_{\mathcal{A}}(P_{\mathcal{A}}(x))$$

für jedes $x \in \mathcal{A}$ mit $\|x\| \leq \rho$. Folgern Sie daraus, daß die durch

$$\log(\mathbb{1} - x) := - \sum_{k=1}^{\infty} \frac{1}{k} x^k$$

definierte Logarithmusabbildung $\log : \{y \in \mathcal{A} \mid \|\mathbb{1} - y\| < 1\} \rightarrow \mathcal{A}$ die Gleichung

$$\exp(\log(\mathbb{1} - x)) = \mathbb{1} - x$$

für $x \in \mathcal{A}$ mit $\|x\| < 1$ erfüllt. **Hinweis:** Siehe [3, Kap. VI, Satz 9, 27.11.2008] oder Übung 1.7.18 in [2, Seite 44].

7 Lineare Differentialgleichungssysteme

Die Exponentialabbildung auf der Banachalgebra $\mathcal{A} = \mathbb{R}^{n \times n}$ aller $(n \times n)$ -Matrizen spielt eine wichtige Rolle bei der Lösung eines Systems von n linearen Differentialgleichungen erster Ordnung in n Variablen. Ein solches Differentialgleichungssystem hat die Form

$$\begin{aligned} \dot{x}_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n, & x_1(0) &= x_{01}, \\ \dot{x}_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n, & x_2(0) &= x_{02}, \\ & \vdots & & \\ \dot{x}_n &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n, & x_n(0) &= x_{0n}, \end{aligned} \quad (10)$$

wobei die $a_{ij} \in \mathbb{R}$ konstante reelle Koeffizienten sind und, für jedes i , die Variable x_i eine stetig differenzierbare Abbildung $x_i : \mathbb{R} \rightarrow \mathbb{R}$ bezeichnen soll. Wir stellen uns das Argument dieser Abbildung als Zeitvariable t vor und bezeichnen mit

$$\dot{x}_i = \frac{dx_i}{dt}$$

die Ableitung der Funktion x_i nach der Zeit t . In der physikalischen Interpretation bezeichnet das n -Tupel $x(t) := (x_1(t), \dots, x_n(t)) \in \mathbb{R}^n$ den Ortsvektor eines Teilchens (oder mehrerer Teilchen gleichzeitig) zum Zeitpunkt t und die Ableitung $\dot{x}(t) = (\dot{x}_1(t), \dots, \dot{x}_n(t)) \in \mathbb{R}^n$ den Geschwindigkeitsvektor zum Zeitpunkt t . Fassen wir die konstanten Koeffizienten zu einer Matrix

$$A = (a_{ij})_{i,j=1}^n \in \mathbb{R}^{n \times n}$$

zusammen so können wir das Differentialgleichungssystem (10) auch in der folgenden Kurzform schreiben

$$\dot{x} = Ax, \quad x(0) = x_0. \quad (11)$$

Hier bezeichnet $x_0 := (x_{01}, \dots, x_{0n}) \in \mathbb{R}^n$ den **Anfangsvektor** (oder den Ortsvektor zum Zeitpunkt $t = 0$). Es folgt nun aus der Gleichung (9), daß die eindeutige Lösung der Differentialgleichung (11) durch die Formel

$$x(t) := e^{At} x_0 \quad (12)$$

gegeben ist, wobei

$$e^{At} = \sum_{k=0}^{\infty} \frac{A^k t^k}{k!} \in \mathbb{R}^{n \times n}$$

die Exponentialmatrix bezeichnet. Die Konvergenz dieser Reihe folgt aus der Konvergenz der Exponentialabbildung für beliebige Banachalgebren \mathcal{A} im Spezialfall $\mathcal{A} = \mathbb{R}^{n \times n}$.

Beispiel 1. Für $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ gilt $\Lambda^k = \text{diag}(\lambda_1^k, \dots, \lambda_n^k)$ und damit

$$e^{\Lambda} = \sum_{k=0}^{\infty} \frac{\Lambda^k}{k!} = \text{diag}(e^{\lambda_1}, \dots, e^{\lambda_n}).$$

Beispiel 2. Für $\lambda \in \mathbb{R}$ gilt

$$A = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} \quad \implies \quad e^{At} = e^{\lambda t} \begin{pmatrix} 1 & t & t^2/2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix}.$$

Übung: Verwenden Sie diese Formel zur Bestimmung der Lösungen eines geeigneten Systems von drei Differentialgleichungen erster Ordnung in drei Variablen. Verallgemeinern Sie die Formel in diesem Beispiel auf vier und fünf Variablen, dann auf Matrizen beliebiger Größe.

Beispiel 3. Für

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

gilt $H^{2k} = (-1)^k \mathbb{1}$ und damit

$$\begin{aligned} e^{tH} &= \left(1 - \frac{t^2}{2!} + \frac{t^4}{4!} \mp \dots \right) \mathbb{1} + \left(t - \frac{t^3}{3!} + \frac{t^5}{5!} \mp \dots \right) H \\ &= \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}. \end{aligned}$$

(Siehe [3, Kap. V, 12.11.2008] für die Potenzreihendarstellungen von Cosinus und Sinus.) Da die Matrizen $a\mathbb{1}$ und bH kommutieren gilt

$$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \implies e^{tA} = e^{at} \begin{pmatrix} \cos(bt) & -\sin(bt) \\ \sin(bt) & \cos(bt) \end{pmatrix}.$$

Schlußbemerkungen

Für den Satz über die Existenz und Eindeutigkeit von Lösungen gewöhnlicher Differentialgleichungen sei an [3, Kap. VIII, Satz 1, 12.12.2008] erinnert. Dieser Satz wurde in viel grösserer Allgemeinheit (auch für nichtlineare Differentialgleichungen) bewiesen. Formeln wie in den obigen Beispielen erlauben es uns im Prinzip, ein lineares Differentialgleichungssystem beliebiger Ordnung explizit zu lösen (indem wir es zunächst in ein Differentialgleichungssystem erster Ordnung verwandeln, anschliessend die Eigenwert-Zerlegung, dh die Jordan'sche Normalform, der Matrix A bestimmen, und schliesslich die Exponentialmatrix der Jordan'schen Normalform ausrechnen). Bei nichtlinearen Differentialgleichungen ist es aber nur in den seltensten Fällen möglich, explizite Lösungen hinzuschreiben, und man kann oft bestenfalls hoffen, etwas über das *qualitative Verhalten* der Lösungen aussagen zu können. Dies führt hin zur qualitativen Theorie **dynamischer Systeme** - einem ganz eigenen und spannenden Gebiet der mathematischen Forschung.

Ein gänzlich anderes Thema, das hier kurz angeschnitten wurde, ist das der Banachräume und der linearen Operatoren. Diese Begriffe stehen am Anfang der **Funktionalanalysis**, ein wichtiges Gebiet der Mathematik, das erst im 5. Semester vertieft wird, und in weiten Teilen sowohl der Mathematik (Partielle Differentialgleichungen bis hin zur Topologie und Geometrie) als auch der Physik (zum Beispiel Quantenmechanik) eine zentrale Rolle spielt.

Schliesslich haben wir hier einen kurzen Einblick in die Banachalgebren erhalten, die im Schnittpunkt zwischen Analysis, Topologie und Algebra stehen, und auch ein ganz eigenes Gebiet der mathematischen Forschung sind.

A Komposition von Potenzreihen

Gegeben seien zwei Folgen $(a_k)_{k \geq 1}$ und $(b_\ell)_{\ell \geq 0}$ reeller Zahlen. Wir betrachten die Potenzreihen

$$f(x) := \sum_{k=1}^{\infty} a_k x^k, \quad g(y) := \sum_{\ell=1}^{\infty} b_\ell y^\ell$$

und nehmen an, daß ihre Konvergenzradien positiv sind:

$$\rho_f := \frac{1}{\limsup_{k \rightarrow \infty} |a_k|^{1/k}} > 0, \quad \rho_g := \frac{1}{\limsup_{\ell \rightarrow \infty} |b_\ell|^{1/\ell}} > 0.$$

Für $k = 0, 1, 2, 3 \dots$ definieren wir

$$c_k := \sum_{\ell \geq 0} b_\ell \sum_{\substack{\sum_i m_i = \ell \\ \sum_i i m_i = k}} \frac{\ell!}{\prod_i m_i!} \prod_i a_i^{m_i}.$$

Die zweite Summe ist zu verstehen über alle endlichen Folgen (m_1, m_2, m_3, \dots) nichtnegativer ganzer Zahlen so daß $\sum_i i m_i = k$ und $\sum_i m_i = \ell$ ist. Da jedes m_i kleiner oder gleich k ist und jedes m_i mit $i > k$ verschwindet, ist diese Summe endlich und verschwindet für alle bis auf endlich viele ℓ . Damit ist c_k wohldefiniert. Der nächste Satz zeigt, daß die Potenzreihe

$$h(x) := \sum_{k=0}^{\infty} c_k x^k$$

einen positiven Konvergenzradius hat und in einer hinreichend kleinen Umgebung des Ursprungs mit der Komposition $g \circ f$ übereinstimmt.

Satz 8. Sei ρ eine reelle Zahl so daß

$$0 < \rho < \rho_f, \quad \sum_{k=1}^{\infty} |a_k| \rho^k < \rho_g.$$

Dann gilt

$$\rho_h := \frac{1}{\limsup_{k \rightarrow \infty} |c_k|^{1/k}} > \rho$$

und, für jede reelle Zahl x ,

$$|x| \leq \rho \quad \implies \quad |f(x)| < \rho_g, \quad g(f(x)) = h(x).$$

Proof. Für jedes Paar ganzer Zahlen $k, N \geq 0$ definieren wir

$$c_{k,N} := \sum_{\ell=0}^N b_\ell \sum_{\substack{m_1 + \dots + m_N = \ell \\ \sum_{i=1}^N i m_i = k}} \frac{\ell!}{m_1! \dots m_N!} a_1^{m_1} a_2^{m_2} \dots a_N^{m_N}.$$

Dann gilt $c_{k,N} = c_k$ für $k \leq N$ und $c_{k,N} = 0$ für $k > N^2$. Insbesondere ist $h_N(x) := \sum_{k \geq 0} c_{k,N} x^k$ ein Polynom.

Definieren wir

$$f_N(x) := \sum_{k=1}^N a_k x^k, \quad g_N(y) := \sum_{\ell=0}^N b_\ell y^\ell,$$

so erhalten wir, für jede reelle Zahl x ,

$$\begin{aligned} h_N(x) &= \sum_{k \geq 0} c_{k,N} x^k \\ &= \sum_{k \geq 0} \sum_{\ell=0}^N b_\ell \sum_{\substack{m_1 + \dots + m_N = \ell \\ \sum_{i=1}^N i m_i = k}} \frac{n!}{m_1! \dots m_N!} a_1^{m_1} a_2^{m_2} \dots a_N^{m_N} x^k \\ &= \sum_{\ell=0}^N b_\ell \sum_{m_1 + \dots + m_N = \ell} \frac{n!}{m_1! \dots m_N!} (a_1 x)^{m_1} (a_2 x^2)^{m_2} \dots (a_N x^N)^{m_N} \\ &= \sum_{\ell=0}^N b_\ell (a_1 x + a_2 x^2 + \dots + a_N x^N)^\ell \\ &= g_N(f_N(x)). \end{aligned}$$

Sei nun $r > 0$ so gewählt daß $\sum_{k=1}^{\infty} |a_k| \rho^k < r < \rho_g$ und sei $x \in [-\rho, \rho]$. Dann konvergiert $f_N(x)$ gegen $f(x) \in (-r, r)$. Da g_N auf dem Intervall $[-r, r]$ gleichmäßig gegen g konvergiert, folgern wir daß $g_N(f_N(x))$ gegen $g(f(x))$ konvergiert. Damit haben wir gezeigt, daß

$$\lim_{N \rightarrow \infty} h_N(x) = g(f(x)) \quad \forall x \in [-\rho, \rho]. \quad (13)$$

Ersetzen wir die Koeffizienten a_k und b_ℓ durch ihre Beträge, so erhalten wir

$$\begin{aligned} |c_k| &\leq \bar{c}_k := \sum_{\ell \geq 0} |b_\ell| \sum_{\substack{\sum_i m_i = \ell \\ \sum_i i m_i = k}} \frac{\ell!}{\prod_i m_i!} \prod_i |a_i|^{m_i}, \\ |c_{k,N}| &\leq \bar{c}_{k,N} := \sum_{n=0}^N |b_n| \sum_{\substack{m_1 + \dots + m_N = n \\ \sum_{i=1}^N i m_i = k}} \frac{n!}{m_1! \dots m_N!} |a_1|^{m_1} |a_2|^{m_2} \dots |a_N|^{m_N}. \end{aligned}$$

Da $\bar{c}_{k,N} \leq \bar{c}_k$ für alle k, N sowie $\bar{c}_{k,N} = \bar{c}_k$ für $k \leq N$ und $\bar{c}_{k,N} = 0$ für $k > N^2$, folgt

$$N' \leq \sqrt{N} \quad \implies \quad \sum_{k \geq 0} \bar{c}_{k,N'} \rho^k \leq \sum_{k=0}^N \bar{c}_k \rho^k \leq \sum_{k \geq 0} \bar{c}_{k,N} \rho^k.$$

Hier konvergieren die beiden äußeren Terme gegen $\sum_{\ell \geq 0} |b_\ell| (\sum_{k \geq 1} |a_k| \rho^k)^\ell$ für $N, N' \rightarrow \infty$, nach dem gleichen Argument wie oben. Also konvergiert die Differenz $\sum_{k > N} \bar{c}_{k,N} \rho^k$ der letzten beiden Terme gegen Null, und für $|x| \leq \rho$ gilt

$$\left| h_N(x) - \sum_{k=0}^N c_k x^k \right| = \left| \sum_{k > N} c_{k,N} x^k \right| \leq \sum_{k > N} \bar{c}_{k,N} \rho^k \xrightarrow{N \rightarrow \infty} 0. \quad (14)$$

Es folgt aus (13) und (14), daß

$$\sum_{k \geq 0} c_k x^k = \lim_{N \rightarrow \infty} \sum_{k=0}^N c_k x^k = \lim_{N \rightarrow \infty} h_N(x) = g(f(x))$$

für jedes $x \in \mathbb{R}$ mit $|x| \leq \rho$. Insbesondere gilt $\rho_h \geq \rho$, da andernfalls die Reihe $\sum_{k \geq 0} c_k x^k$ für $|x| = \rho$ divergieren würde. Da wir ρ ein wenig vergrößern können, ohne die Voraussetzungen des Satzes zu verletzen, gilt sogar $\rho_h > \rho$ und damit ist der Satz bewiesen. \square

References

- [1] C. Blatter, *Analysis Zwei*, ETHZ, Sommersemester 2004.
- [2] K. Königsberger, *Analysis 2*, 5. Auflage, Springer Verlag, 2003.
- [3] D. Salamon, *Analysis I*, Vorlesung an der ETHZ im Herbstsemester 2008.