Analysis II (FS 2017): BANACHRÄUME UND LINEARE OPERATOREN

Dietmar A. Salamon
ETH-Zürich

28. Februar 2017

1 Äquivalente Normen

Dieser erste Abschnitt ruft einige Definitionen und Resultate in Erinnerung, die aus der Analysis I Vorlesung bekannt sind. Sei X ein reeller Vektorraum. Eine **Norm** auf X ist eine Abbildung $X \to \mathbb{R} : x \mapsto \|x\|$ mit folgenden Eigenschaften.

(i) Für alle $x \in X$ gilt $\|x\| \geq 0$ und $\|x\| = 0 \iff x = 0$.

(ii) Für alle $x \in X$ und alle $\lambda \in \mathbb{R}$ gilt $\|\lambda x\| = |\lambda| \, \|x\|$.

(iii) Für alle $x, y \in X$ gilt $\|x + y\| \leq \|x\| + \|y\|$.

Ein **normierter Vektorraum** ist ein Paar $(X, \|\cdot\|)$ welches aus einem reellen Vektorraum X und einer Normfunktion $\|\cdot\|$ auf X besteht. Eine wichtige Klasse normierter Vektorräume sind solche bei denen die Norm von einem inneren Produkt erzeugt wird. Ein **inneres Produkt** auf X ist eine symmetrische bilineare Abbildung $X \times X \to \mathbb{R} : (x, y) \mapsto \langle x, y\rangle$, welche die Bedingung $\langle x, x\rangle > 0$ für alle $x \in X \setminus \{0\}$ erfüllt. Ist solch ein inneres Produkt gegeben, so definiert die Formel

$$\|x\| := \sqrt{\langle x, x\rangle} \qquad \text{für } x \in X \tag{1}$$

eine Norm und es gilt die Cauchy–Schwarz-Ungleichung $|\langle x, y\rangle| \leq \|x\| \, \|y\|$ für alle $x, y \in X$.

Sei $(X, \|\cdot\|)$ ein normierter reeller Vektorraum. Dann definiert die Formel

$$d(x, y) := \|x - y\| \qquad \text{für } x, y \in X \tag{2}$$

eine Abstandsfunktion $d : X \times X \to \mathbb{R}$, so dass X damit zu einem metrischen Raum wird. Der normierte Vektorraum $(X, \|\cdot\|)$ heisst **Banachraum** wenn der zugehörige metrische Raum (X, d) **vollständig** ist, das heisst wenn jede Cauchy-Folge in X konvergiert. (Es sei daran erinnert, dass eine Folge $(x_n)_{n\in\mathbb{N}}$ in X eine **Cauchy-Folge** ist, wenn es zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ gibt so dass für alle $n, m \in \mathbb{N}$ gilt: $n, m \geq n_0 \implies d(x_n, x_m) < \varepsilon$.) Ein **(reeller) Hilbertraum** ist ein Paar $(X, \langle\cdot,\cdot\rangle)$ welches aus einem reellen Vektorraum X und einem inneren Product $\langle\cdot,\cdot\rangle$ auf X besteht, so dass der zugehörige metrische Raum (X, d) mit der durch (1) und (2) definierten Abstandsfunktion vollständig ist. Also ist jeder Hilbertraum auch ein Banachraum mit der durch (1) definierten Normfunktion.

Es sei nun $(X, \|\cdot\|)$ ein normierter Vektorraum. Ist $x \in X$ und $r > 0$, so wird die Menge

$$B_r(x) := B_r(x; \|\cdot\|) := B_r(x; X, \|\cdot\|) := \left\{ y \in X \,\middle|\, \|x - y\| < r \right\}$$

der **Ball vom Radius r mit Mittelpunkt x** genannt. Eine Teilmenge $U \subset X$ heisst **offen**, wenn es für jedes Element $x \in U$ ein $\varepsilon > 0$ gibt so dass $B_\varepsilon(x) \subset U$ ist. Die Gesamtheit der offenen Mengen wird mit

$$\mathscr{U}(X, \|\cdot\|) := \left\{ U \subset X \,\middle|\, U \text{ ist offen bezüglich } \|\cdot\| \right\} \tag{3}$$

bezeichnet.

Beispiel 1.1. 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

$$\|x\|_p := \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}, \qquad x \in \mathbb{R}^n, \tag{4}$$

eine Norm auf dem $\mathbb{R}^n$. Die Dreiecksungleichung für diese Norm heisst *Minkowski-Ungleichung*. Im Fall $p > 1$ verwendet ihr Beweis die *Hölder-Ungleichung* $\langle x, y\rangle = \sum_{i=1}^n x_i y_i \leq \|x\|_p \|y\|_q$ für $x, y \in \mathbb{R}^n$ und $1/p + 1/q = 1$. Für $p = \infty$ definiert die Formel

$$\|x\|_\infty := \max\{|x_1|, \dots, |x_n|\}, \qquad x \in \mathbb{R}^n, \tag{5}$$

eine Norm auf $\mathbb{R}^n$.

Beispiel 1.2. 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 sie die Bedingung

$$x^T Ax = \sum_{i,j=1}^{n} x_i a_{ij} x_j > 0 \qquad \text{für alle } x \in \mathbb{R}^n \setminus \{0\}$$

erfüllt. Jede positiv definite $(n \times n)$-Matrix A definiert ein inneres Product

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

auf dem $\mathbb{R}^n$. Die zugehörige Norm eines Vektors $x \in \mathbb{R}^n$ ist durch die Formel $\|x\|_A := \sqrt{x^T Ax}$ gegeben.

Beispiel 1.3. Sei (M, d) ein metrischer Raum. Denn ist der Raum

$$\mathrm{BC}(M) := \left\{ f : M \to \mathbb{R} \,\middle|\, f \text{ is stetig und beschränkt} \right\}$$

mit der durch

$$\|f\|_\infty := \sup_{p\in M} |f(p)| \qquad \text{für } f \in \mathrm{BC}(M) \tag{6}$$

definierten Norm ein Banachraum. Die Vollständigkeit folgt aus der Tatsache, dass der Grenzwert einer gleichmässig konvergierenden Folge stetiger Funktionen selbst wieder stetig ist, und dass gleichmässige Konvergenz äquivalent ist zur Konvergenz bezüglich der Supremumsnorm. Ist $(M.d)$ kompakt (das heisst jede Folge in M besitzt eine konvergente Teilfolge) so ist jede stetige Funktion $f : M \to \mathbb{R}$ notwendigerweise beschränkt. In diesem Fall stimmt also der Raum $\mathrm{BC}(M)$ mit dem Raum $\mathrm{C}(M)$ aller stetigen reellwertigen Funktionen auf M überein, und der Raum $\mathrm{C}(M)$ ist mit der Supremumsnorm (6) ein Banachraum.

Beispiel 1.4. Sei $I = [a, b]$ ein kompaktes Interval reeller Zahlen (mit $a < b$). Für $1 \le p < \infty$ definiert die Formel

$$\|f\|_p := \left(\int_a^b |f(t)|^p \, dt \right)^{1/p} \qquad \text{für } f \in \mathrm{C}(I) \tag{7}$$

eine Norm auf dem Raum $\mathrm{C}(I)$ aller stetigen reellwertigen Funktionen auf I. Der Raum $\mathrm{C}(I)$ ist mit der Norm (7) nicht vollständig (Übung) und ist daher kein Banachraum. Im Fall $p = 2$ wird die Norm (7) durch das innere Produkt

$$\langle f, g\rangle := \int_a^b f(t)g(t)\, dt \qquad \text{für } f, g \in \mathrm{C}(I) \tag{8}$$

erzeugt.

Beispiel 1.5. Sei $I = [a, b]$ ein kompaktes Interval reeller Zahlen (mit $a < b$) und sei $\ell \geq 0$ eine nichtnegative ganze Zahl. Dann ist der reelle Vektorraum

$$\mathrm{C}^\ell(I) := \left\{ f : I \to \mathbb{R} \,\middle|\, f \text{ ist } \ell \text{ mal stetig differenzierbar} \right\}$$

mir der durch

$$\|f\|_{\mathrm{C}^\ell} := \max_{k=0,\ldots,\ell} \sup_{a \leq t \leq b} \left| f^{(k)}(t) \right|$$

für $f \in \mathrm{C}^\ell(I)$ definierten Norm ein Banachraum. Für $\ell = 0$ ist dies der Raum $\mathrm{C}(I)$ der stetigen Funktionen $f : I \to \mathbb{R}$ mit der Supremumsnorm wie in Beispiel 1.3.

Definition 1.6. *Seien $\|\cdot\|$ und $\|\cdot\|'$ zwei Normfunktionen auf einem reellen Vektorraum X. Die Normfunktion $\|\cdot\|$ heisst* **äquivalent** *zu $\|\cdot\|'$ (Schreibweise $\|\cdot\| \sim \|\cdot\|'$) wenn eine Konstante $c > 0$ existiert, so dass*

$$\frac{1}{c} \|x\| \leq \|x\|' \leq c \|x\| \qquad \text{für alle } x \in X. \tag{9}$$

Bemerkung 1.7. (i) Sei X ein reeller Vektorraum. Dann definiert die *"Äquivalenz von Normen"* eine Äquivalenzrelation auf der Menge aller Normen auf X. Das heisst, für alle Normen $\|\cdot\|$, $\|\cdot\|'$, $\|\cdot\|''$ auf X gilt
- $\|\cdot\| \sim \|\cdot\|$,
- $\|\cdot\| \sim \|\cdot\|' \iff \|\cdot\|' \sim \|\cdot\|$,
- $\|\cdot\| \sim \|\cdot\|'$, $\|\cdot\|' \sim \|\cdot\|'' \implies \|\cdot\| \sim \|\cdot\|''$.

(ii) Seien $\|\cdot\|$ und $\|\cdot\|'$ zwei äquivalente Normen auf einem reellen Vektorraum X. Dann erzeugen die beiden Normen dieselben offenen Mengen. (Übung; siehe auch Beispiel 2.4).

(iii) Seien $\|\cdot\|$ und $\|\cdot\|'$ zwei äquivalente Normen auf einem reellen Vektorraum X. Da die offenen Teilmengen von X bezüglich der beiden Normen dieselben sind, folgt, dass alle *topologischen Aussagen in X* (also Aussagen die sich mit Hilfe offener Mengen formulieren lassen, zum Beispiel Konvergenz, Stetigkeit, Abgeschlossenheit, Kompaktheit) nicht davon abhängen, welche der beiden äquivalenten Normen wir dazu verwenden. Ist zum Beispiel $A \subset X$ kompakt (beziehungsweise abgeschlossen) bezüglich der Norm $\|\cdot\|$ so ist A auch kompakt (beziehungsweise abgeschlossen) bezüglich der Norm $\|\cdot\|'$. Weiterhin gilt, dass eine Folge $(x_k)_{k \in \mathbb{N}}$ in X genau dann bezüglich der Norm $\|\cdot\|$ eine Cauchy-Folge ist, beziehungsweise konvergiert, wenn sie bezüglich der Norm $\|\cdot\|'$ eine Cauchy-Folge ist, beziehungsweise konvergiert. Daher ist $(X, \|\cdot\|)$ genau dann ein Banachraum wenn $(X, \|\cdot\|')$ ein Banachraum ist.

Sei $I = [a, b] \subset \mathbb{R}$ ein kompaktes Intervall mit $a < b$. Dann ist die Menge $U := \{f \in \mathrm{C}(I) \mid \sup_I |f| < 1\}$ offen bezüglich der Supremumsnorm $\|\cdot\|_\infty$ in Beispiel 1.3 aber nicht bezüglich der Norm $\|\cdot\|_p$ in Beispiel 1.4 für $1 \le p < \infty$. Also sind die Normen $\|\cdot\|_\infty$ und $\|\cdot\|_p$ auf dem Raum $\mathrm{C}(I)$ nicht äquivalent. Dies hängt damit zusammen, dass der Vektorraum der stetigen Funktionen auf einem nichttrivialen kompakten Intervall unendlichdimensional ist. Der nächste Satz zeigt, dass auf endlichdimensionalen Vektorräumen je zwei Normen äquivalent sind.

Satz 1.8. *Sei X ein endlichdimensionaler reeller Vektorraum. Dann sind je zwei Normen auf X äquivalent.*

Beweis. Der Beweis hat vier Schritte. Sei $n := \dim X \in \mathbb{N}$. Wir bezeichnen mit $\|\cdot\|_2$ die Euklidische Norm auf dem $\mathbb{R}^n$ in (4) mit $p = 2$.

Schritt 1. *Sei $\|\cdot\|$ eine beliebige Norm auf dem $\mathbb{R}^n$. Dann existiert eine Konstante $c > 0$ so dass jedes $x \in \mathbb{R}^n$ die Ungleichung $\|x\| \le c \|x\|_2$ erfüllt.*

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 $c := (\sum_{i=1}^n \|e_i\|^2)^{1/2} > 0$. Dann gilt $x = \sum_{i=1}^n x_i e_i$ für jeden Vektor $x \in \mathbb{R}^n$ und daher

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

Hier folgt die erste Ungleichung aus der Dreiecksungleichung für $\|\cdot\|$ und die zweite aus der Cauchy–Schwarz-Ungleichung für die Euklidische Norm.

Schritt 2. *Sei $\|\cdot\|$ wie in Schritt 1. Dann existiert eine Konstante $\delta > 0$ so dass jedes $x \in \mathbb{R}^n$ die Ungleichung $\delta \|x\|_2 \le \|x\|$ erfüllt.*

Die Einheitssphäre $S^{n-1} := \{x \in \mathbb{R}^n \mid \|x\|_2 = 1\}$ ist mit der Abstandsfunktion $d(x, y) := \|x - y\|_2$ für $x, y \in S^{n-1}$ ein kompakter metrischer Raum nach dem Satz von Heine–Borel. Wir definieren die Funktion $f : S^{n-1} \to \mathbb{R}$ durch $f(x) := \|x\|$ für $x \in S^{n-1}$. Nach Schritt 1 gilt $|f(x) - f(y)| = |\|x\| - \|y\|| \le \|x - y\| \le c \|x - y\|_2$ für alle $x, y \in S^{n-1}$ und daher ist f stetig. Da S^{n-1} kompakt ist, existiert ein $x_0 \in S^{n-1}$ mit $\inf_{S^{n-1}} f = f(x_0) =: \delta > 0$. Für $x \in \mathbb{R}^n \setminus \{0\}$ gilt daher

$$\delta \le f\left(\|x\|_2^{-1} x\right) = \left\|\|x\|_2^{-1} x\right\| = \|x\|_2^{-1} \|x\| .$$

Hieraus folgt die Behauptung von Schritt 2.

Schritt 3. *Je zwei Normen auf dem $\mathbb{R}^n$ sind äquivalent.*

Nach Schritt 1 und Schritt 2 ist jede Norm auf dem $\mathbb{R}^n$ äquivalent zur Euklidischen Norm. Daher folgt Schritt 3 aus Teil (i) von Bemerkung 1.7.

Schritt 4. *Je zwei Normen auf X sind äquivalent.*

Wähle eine Basis $e_1, \dots, e_n$ von X. Dann ist die Abbildung

$$\Phi : \mathbb{R}^n \to X, \qquad \Phi\xi := \sum_{i=1}^{n} \xi_i e_i \quad \text{für } \xi = (\xi_1, \dots \xi_n) \in \mathbb{R}^n,$$

linear und bijektiv, und ist daher ein Vektorraum-Isomorphismus. Sind $\|\cdot\|$ und $\|\cdot\|'$ zwei Normen auf X, so definieren die Formeln $\|\xi\|_\Phi := \|\Phi\xi\|$ und $\|\xi\|'_\Phi := \|\Phi\xi\|'$ für $\xi \in \mathbb{R}^n$ zwei Normen auf dem $\mathbb{R}^n$. Diese sind nach Schritt 3 äquivalent und daher sind auch die Normen $\|\cdot\|$ und $\|\cdot\|'$ auf X äquivalent. Damit ist Satz 1.8 bewiesen. □

Satz 1.8 hat einige wichtige Konsequenzen.

Korollar 1.9. *Jeder endlichdimensionaler normierter Vektorraum ist ein Banachraum.*

Beweis. In der Analysis I Vorlesung wurde gezeigt, dass der $\mathbb{R}^n$ mit der Euklidischen Norm ein Banachraum ist. Nach Satz 1.8 ist daher der $\mathbb{R}^n$ mit jeder beliebigen Norm ein Banachraum. Da jeder endlichdimensionale normierte Vektorraum zum $\mathbb{R}^n$ isomorph ist für ein $n \in \mathbb{N}_0$, folgt daraus die Behauptung. □

Korollar 1.10. *Eine Teilmenge eines endlichdimensionalen normierten Vektorraumes ist genau dann kompakt wenn sie abgeschlossen und beschränkt ist.*

Beweis. Nach dem Satz von Heine–Borel gilt dies für den $\mathbb{R}^n$ mit der Euklidischen Norm. Nach Satz 1.8 sind aber je zwei Normen auf dem $\mathbb{R}^n$ äquivalent und daher sind alle drei Eigenschaften (Kompaktheit, Abgeschlossenheit, Beschränktheit) einer Teilmenge des $\mathbb{R}^n$ unabhängig von der Wahl der Norm. Also gilt die Behauptung für jede beliebige Norm auf dem $\mathbb{R}^n$. Da jeder endlichdimensionale reelle Vektorraum zum $\mathbb{R}^n$ für ein $n \in \mathbb{N}_0$ isomorph ist, folgt daraus die Behauptung im allgemeinen. □

Nach Korollar 1.10 ist die Einheitssphäre $S := \{x \in X \mid \|x\| = 1\}$ in jedem endlichdimensionalen normierten Vektorraum kompakt. Die folgende Übung zeigt, dass diese Eigenschaft genau die endlichdimensionalen normierten Vektorräume charakterisiert.

Übung 1.11. Sei $(X, \|\cdot\|)$ ein normierter reeller Vektorraum, so dass die Einheitssphähre $S := \{x \in X \mid \|x\| = 1\}$ kompakt ist. Dann ist X endlichdimensional. **Hinweis:** Beweisen Sie folgendes.

(i) Jeder endlichdimensionale lineare Unterraum von X ist eine abgeschlossene Teilmenge von X.

(ii) Sei $Y \subset X$ eine abgeschlossene Teilmenge. Dann gilt für jedes $x \in X \setminus Y$ die Ungleichung $d(x, Y) := \inf_{y \in Y} \|x - y\| > 0$.

(iii) Sei $Y \subsetneq X$ ein abgeschlossener linearer Unterraum von X der nicht gleich dem ganzen Raum X ist. Dann existiert ein Vektor $x \in X$ so dass $\|x\| = 1$ und $d(x, Y) \geq 1/2$ ist. (Sei $x_0 \in X \setminus Y$. Dann existiert nach (ii) ein $y_0 \in Y$ so dass $\|x_0 - y_0\| \leq 2d(x_0, Y)$ ist. Nun sei $x := \|x_0 - y_0\|^{-1}(x_0 - y_0)$.)

(iv) Ist X unendlichdimensional, so existiert eine Folge $(x_n)_{n \in \mathbb{N}}$ in S so dass $\|x_n - x_m\| \geq 1/2$ ist für alle $m, n \in \mathbb{N}$ mit $m \neq n$.

Übung 1.12. Sei X ein endlichdimensionaler reeller Vektorraum. Eine Teilmenge $K \subset X$ heisst **konvex** wenn sie, für alle $x, y \in X$ und alle $t \in \mathbb{R}$, die Bedingung

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

erfüllt.

(i) Sei $f : X \to \mathbb{R}$ eine Funktion. Die offene Teilmenge

$$U := \left\{(x, s) \in X \times \mathbb{R} \,\middle|\, s > f(x)\right\}$$

ist genau dann konvex, wenn die Funktion f konvex ist.

(ii) Sei $U \subset X$ eine offene, beschränkte, konvexe Teilmenge so dass für alle $x \in X$ folgendes gilt:

$$x \in U \qquad \Longleftrightarrow \qquad -x \in U.$$

(Nach Satz 1.8 hängen die Begriffe *"offen"* und *"beschränkt"* für Teilmengen von X nicht von der Norm ab, welche für die Definitionen verwendet wird.) Dann definiert die Formel

$$\|x\| := \inf\left\{\lambda > 0 \,|\, \lambda^{-1}x \in U\right\} \qquad \text{für } x \in X$$

eine Normfunktion auf X und es gilt $U = \{x \in X \mid \|x\| < 1\}$.

Übung 1.13. Auf dem Raum $\mathrm{C}([0,1])$ betrachten wir die Normen $\|\cdot\|_p$ in Beispiel 1.4 für $1 \leq p \leq \infty$. Für welche p und q gibt es eine Konstante $c > 0$ so dass die Ungleichung $\|f\|_p \leq c\,\|f\|_q$ für alle $f \in \mathrm{C}([0,1])$ gilt? Zeigen Sie, dass die Normen $\|\cdot\|_p$ und $\|\cdot\|_q$ für $p \neq q$ nicht äquivalent sind.

2 Die Operatornorm

Seien $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ normierte reelle Vektorräume.

Definition 2.1. *Ein linearer Operator $A : X \to Y$ heisst* **beschränkt***, wenn es eine Konstante $c \geq 0$ gibt so dass die Ungleichung*

$$\|Ax\|_Y \leq c \|x\|_X \tag{10}$$

für alle $x \in X$ gilt. Die kleinste solche Konstante $c \geq 0$ heisst **Norm** *(beziehungsweise* **Operatornorm***) von A und wird mit*

$$\|A\| := \sup_{x \neq 0} \frac{\|Ax\|_Y}{\|x\|_X} \tag{11}$$

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 wird mit $\mathcal{L}(X, Y)$ bezeichnet.

Der 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.

Korollar 2.2. *Sei X endlichdimensional. Dann ist jeder lineare Operator $A : X \to Y$ beschränkt.*

Beweis. Wir definieren eine Funktion

$$X \to \mathbb{R} : x \mapsto \|x\|_A := \|x\|_X + \|Ax\|_Y . \tag{12}$$

Da A linear ist, ist die Funktion (12) eine Norm. Da X endlichdimensional ist, existiert nach Satz 1.8 eine Konstante $c \geq 1$ so dass für jedes $x \in X$ die Ungleichung $\|x\|_A \leq c \|x\|_X$ gilt. Daraus folgt unmittelbar die gewünschte Ungleichung $\|Ax\|_Y \leq c \|x\|_X$ für alle $x \in X$. $\square$

Satz 2.3. *Sei $A : X \to Y$ ein linearer Operator. Dann sind folgende Aussagen äquivalent.*

(i) *A ist beschränkt.*

(ii) *A ist stetig.*

(iii) *A ist stetig an der Stelle $x = 0$.*

Beweis. Wir zeigen, dass (ii) aus (i) folgt. Wenn A beschränkt ist, so ist A Lipschitz-stetig mit der Lipschitz-Konstanten $\|A\|$, denn es gilt

$$\begin{aligned} d_Y(Ax_0, Ax_1) &= \|Ax_0 - Ax_1\|_Y = \|A(x_0 - x_1)\|_Y \\ &\leq \|A\| \, \|x_0 - x_1\|_X = \|A\| \, d_X(x_0, x_1) \end{aligned}$$

für alle $x_0, x_1 \in X$; damit ist A auch stetig. Dass (iii) aus (ii) folgt, ergibt sich unmittelbar aus der Definition der Stetigkeit.

Wir zeigen, dass (i) aus (iii) folgt. Sei also A an der Stelle $x = 0$ stetig. Dann folgt aus der Definition der Stetigkeit mit $\varepsilon = 1$, dass ein $\delta > 0$ existiert, so dass jedes $x \in X$ mit $\|x\|_X \leq \delta$ die Ungleichung $\|Ax\|_Y \leq 1$ erfüllt. Sei nun $x \in X \setminus \{0\}$. Dann gilt

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

und daher

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

Hieraus folgt $\|Ax\|_Y \leq \delta^{-1} \|x\|_X$ für alle $x \in X$. Also ist A beschränkt. □

Beispiel 2.4. Seien $\|\cdot\|$ und $\|\cdot\|'$ zwei Normen auf einem rellen Vektorraum X. Wir betrachten die Aussage

$$\exists c > 0 \, \forall x \in X : \|x\|' \leq c \|x\| . \tag{13}$$

Diese Bedingung gilt genau dann, wenn der lineare Operator

$$\mathrm{id} : (X, \|\cdot\|) \to (X, \|\cdot\|') \tag{14}$$

beschränkt ist. Nach Satz 2.3 heisst das, dass die Abbildung (14) stetig ist. Das wiederum heisst, dass das Urbild jeder offenen Menge in $(X, \|\cdot\|')$ unter der Abbildung (14) offen in $(X, \|\cdot\|)$ ist. Dies bedeutet

$$\mathscr{U}(X, \|\cdot\|') \subset \mathscr{U}(X, \|\cdot\|). \tag{15}$$

Also haben wir gezeigt, dass die Aussagen (13) und (15) äquivalent sind. Daraus folgt die Aussage

$$\|\cdot\| \sim \|\cdot\|' \qquad \Longleftrightarrow \qquad \mathscr{U}(X, \|\cdot\|) = \mathscr{U}(X, \|\cdot\|')$$

für je zwei Normen auf einem reellen Vektorraum X (siehe auch Teil (ii) von Bemerkung 1.7). Nach Satz 1.8 heisst das wiederum, dass die Gesamtheit der offenen Mengen für einen endlichdimensionalen reellen Vektorraum X unabhängig von der Wahl der Norm ist. Damit besitzt jeder endlichdimensionale reelle Vektorraum eine *natürliche Topologie.*

Lemma 2.5. *$\mathcal{L}(X, Y)$ ist ein normierter Vektorraum.*

Beweis. Folgende drei Aussagen sind zu beweisen.
(a) Für alle $A \in \mathcal{L}(X, Y)$ gilt $\|A\| = 0 \implies A = 0$.
(b) Für alle $A \in \mathcal{L}(X, Y)$ und $\lambda \in \mathbb{R}$ gilt $\lambda A \in \mathcal{L}(X, Y)$ und $\|\lambda A\| = |\lambda| \, \|A\|$.
(c) Für alle $A, B \in \mathcal{L}(X, Y)$ gilt $A+B \in \mathcal{L}(X, Y)$ und $\|A + B\| \leq \|A\|+\|B\|$.
Die Aussagen (a) und (b) folgen unmittelbar aus den Definitionen. Wir beweisen (c). Seien $A, B : X \to Y$ beschränkte lineare Operatoren. Dann gilt für alle $x \in X$ folgende Ungleichung

$$\begin{aligned}
\|(A+B)x\|_Y &= \|Ax + Bx\|_Y \\
&\leq \|Ax\|_Y + \|Bx\|_Y \\
&\leq \|A\| \, \|x\|_X + \|B\| \, \|x\|_X \\
&= (\|A\| + \|B\|) \, \|x\|_X \, .
\end{aligned}$$

Also ist $A + B$ beschränkt und für $x \in X \setminus \{0\}$ gilt $\frac{\|(A+B)x\|_Y}{\|x\|_X} \leq \|A\| + \|B\|$. Damit folgt die Dreiecksungleichung $\|A + B\| \leq \|A\|+\|B\|$ aus der Definition der Operatornorm. □

Satz 2.6. *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 $(A_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in $\mathcal{L}(X, Y)$. Dann zeigt die Ungleichung

$$\|A_n x - A_m x\|_Y \leq \|A_n - A_m\| \, \|x\|_X \, ,$$

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

$$Ax := \lim_{n\to\infty} A_n x$$

für $x \in X$. Wir werden beweisen, dass A ein beschränkter linearer Operator ist und dass die Folge A_n in der Operatornorm gegen A konvergiert.

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

$$A(x_0 + x_1) = \lim_{n\to\infty} A_n(x_0 + x_1) = \lim_{n\to\infty} A_n x_0 + \lim_{n\to\infty} A_n x_1 = Ax_0 + Ax_1.$$

Genauso zeigt man, dass $A(\lambda x) = \lambda Ax$ ist für alle $\lambda \in \mathbb{R}$ und alle $x \in X$.

A **ist beschränkt.** Zunächst folgt aus der Ungleichung

$$|\|A_n\| - \|A_m\|| \leq \|A_n - A_m\|,$$

dass die Folge $(\|A_n\|)_{n\in\mathbb{N}}$ eine Cauchy-Folge reeller Zahlen ist und daher konvergiert. Sei $c := \lim_{n\to\infty} \|A_n\|$. Dann erfüllt jedes $x \in X$ die Ungleichung

$$\|Ax\|_Y = \left\|\lim_{n\to\infty} A_n x\right\|_Y = \lim_{n\to\infty} \|A_n x\|_Y \leq \lim_{n\to\infty} \|A_n\| \|x\|_X = c \|x\|_X .$$

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

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

$$m > n \geq n_0 \qquad \Longrightarrow \qquad \|A_n - A_m\| < \varepsilon.$$

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

$$\|A_n x - Ax\|_Y = \lim_{m\to\infty} \|A_n x - A_m x\|_Y \leq \varepsilon \|x\|_X$$

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

$$\|A_n - A\| = \sup_{x\neq 0} \frac{\|A_n x - Ax\|_Y}{\|x\|_X} \leq \varepsilon$$

für alle $n \in \mathbb{N}$ mit $n \geq n_0$. Damit ist Satz 2.6 bewiesen. □

Ein wichtiger Spezialfall ist $Y = \mathbb{R}$. Der Banachraum $X^* := \mathcal{L}(X, \mathbb{R})$ heisst **Dualraum** des normierten Vektorraumes X. Satz 2.6 zeigt, dass dieser Dualraum stets vollständig ist, auch wenn X nicht vollständig ist.

Übung 2.7. (i) Sei X ein endlichdimensionaler normierter Vektorraum. Dann ist $X^* = \mathcal{L}(X, \mathbb{R})$ endlichdimensional und $\dim X^* = \dim X$.

(ii) Sei $n \in \mathbb{N}$ und seien $1 \leq p, q \leq \infty$ mit $1/p + 1/q = 1$. Für $y \in \mathbb{R}^n$ definieren wir die lineare Abbildung $\phi_y : \mathbb{R}^n \to \mathbb{R}$ durch $\phi_y(x) := \sum_{i=1}^n x_i y_i = \langle x, y\rangle$ für $x \in \mathbb{R}^n$. Dann ist die Abbildung $\mathbb{R}^n \to (\mathbb{R}^n)^* : y \mapsto \phi_y$ ein Vektorraum-Isomorphismus und es gilt

$$\|y\|_q = \sup_{x\neq 0} \frac{|\langle x, y\rangle|}{\|x\|_p}$$

für alle $y \in \mathbb{R}^n$.

3 Banachalgebren

Sei $(X, \|\cdot\|)$ ein Banachraum. Dann ist auch der Raum

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

aller beschränkten linearen Operatoren von X nach X ein Banachraum (siehe Satz 2.6). 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 invereses 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 3.1. *Eine* **reelle Algebra** *besteht aus einem reellen Vektorraum $\mathcal{A}$, einer Abbildung*

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

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

(i) *Das Produkt ist assoziativ, das heisst für alle $x, y, z \in \mathcal{A}$ gilt*

$$(xy)z = x(yz).$$

(ii) *Das Produkt ist bilinear, das heisst für alle $x, y, z \in \mathcal{A}$ und alle $\lambda \in \mathbb{R}$ gilt*

$$x(y + z) = xy + xz, \qquad (x + y)z = xz + yz, \qquad x(\lambda z) = \lambda(xz) = (\lambda x)z.$$

(iii) *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 $\mathcal{A} \to [0, \infty) : x \mapsto \|x\|$ welche die Produkt-Ungleichung*

$$\|xy\| \leq \|x\| \, \|y\| \tag{16}$$

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).*

Man beachte, dass das 1-Element in einer reellen Algebra $\mathcal{A}$ durch das Produkt eindeutig bestimmt ist. Ist nämlich $\mathbb{1}' \in \mathcal{A}$ ein weiteres Element welches die Bedingung $\mathbb{1}'x = x\mathbb{1}' = \mathbb{1}'$ für alle $x \in \mathcal{A}$ erfüllt, so gilt $\mathbb{1}' = \mathbb{1}'\mathbb{1} = \mathbb{1}$. Eine Algebra unterscheidet sich von einem Körper dadurch, dass sie auch eine (mit dem Produkt kompatible) Vektorraumstruktur besitzt, dass das Produkt nicht kommutativ sein muss, und dass nicht jedes von Null verschiedene Element von $\mathcal{A}$ invertierbar sein muss.

Definition 3.2. *Sei $\mathcal{A}$ eine reelle Algebra. Ein Element $x \in \mathcal{A}$ heisst* **invertierbar***, wenn ein Element $y \in \mathcal{A}$ existiert, welches die Bedingung*

$$xy = yx = \mathbb{1} \tag{17}$$

erfüllt. Existiert ein solches Element, so ist es durch x eindeutig bestimmt (denn jedes weitere Element $y' \in \mathcal{A}$, welches der Bedingung $y'x = \mathbb{1}$ genügt, erfüllt die Gleichung $y' = y'\mathbb{1} = y'(xy) = (y'x)y = \mathbb{1}y = y$). Ist $x \in \mathcal{A}$ invertierbar so wird das eindeutige Element $y \in \mathcal{A}$, das die Bedingung (17) *erfüllt, das* **inverse Element von** *x genannt und mit $x^{-1} := y$ bezeichnet. Die Gruppe der invertierbaren Elemente von $\mathcal{A}$ bezeichnen wir mit*

$$\mathcal{G} := \left\{ x \in \mathcal{A} \,\middle|\, \textit{es gibt ein } y \in \mathcal{A} \textit{ so dass } xy = yx = \mathbb{1} \right\}.$$

Lemma 3.3. *Sei $\mathcal{A}$ eine reelle Algebra mit* 1*-Element $\mathbb{1}$. Dann gilt folgendes.*

(i) $\mathbb{1} \in \mathcal{G}$ *und* $\mathbb{1}^{-1} = \mathbb{1}$.

(ii) *Ist $x \in \mathcal{G}$ so ist $x^{-1} \in \mathcal{G}$ und es gilt $(x^{-1})^{-1} = x$.*

(iii) *Sind $x, y \in \mathcal{G}$ so ist $xy \in \mathcal{G}$ und es gilt $(xy)^{-1} = y^{-1}x^{-1}$.*

Beweis. Diese Aussagen folgen unmittelbar aus der Definition. □

Beispiel 3.4. Die reellen und die komplex Zahlen bilden Banachalgebren mit dem üblichen Produkt und der Betragsfunktion als Norm. Ein ähnliches Beispiel ist durch die Quaternionen gegeben. Die **Quaternionen** sind durch eine geeignete Produktstruktur auf dem reellen Vektorraum $\mathbb{H} \cong \mathbb{R}^4$ definiert. Wir schreiben die Elemente von $\mathbb{H}$ als formale Summen $x = x_0 + \mathbf{i}x_1 + \mathbf{j}x_2 + \mathbf{k}x_3$ mit $x_i \in \mathbb{R}$ und definieren die Produktstruktur so, dass die Gleichungen

$$\mathbf{jk} = -\mathbf{kj} = \mathbf{i}, \quad \mathbf{ki} = -\mathbf{ik} = \mathbf{j}, \quad \mathbf{ij} = -\mathbf{ji} = \mathbf{k}, \quad \mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1$$

gelten und 1 das 1-Element ist. Die Norm auf $\mathbb{H}$ ist $|x| := \sqrt{\sum_{i=0}^{3} x_i^2}$ und es gilt $|xy| = |x||y|$ für alle $x, y \in \mathbb{H}$. Eine bilineare Abbildung, die ebenfalls diese Eigenschaft hat aber nicht mehr assoziativ ist, gibt es noch in der Dimension acht (die Oktonionen), aber in keiner Dimension ausser $1, 2, 4, 8$.

Beispiel 3.5. Ist X ein Banachraum so ist $\mathcal{L}(X)$ mit der Operatornorm, der Komposition, und der Identität eine Banachalgebra. Die Vollständigkeit folgt aus Satz 2.6. **Übung:** Die Operatornorm auf $\mathcal{L}(X)$ erfüllt die Bedingung (16).

Beispiel 3.6. Der Raum $\mathbb{R}^{n\times n}$ der quadratischen Matrizen ist eine Banachalgebra mit jeder Norm welche der Bedingung (16) 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 die Identifiktion $\mathbb{R}^{n\times n} \cong \mathbb{R}^{n^2}$ ergibt; in diesem Fall erhält man

$$\|A\|_2 = \left(\sum_{i,j=1}^{n} a_{ij}^2\right)^{1/2}$$

für $A = (a_{ij}) \in \mathbb{R}^{n\times n}$. **Übung:** Diese Norm erfüllt die Bedingung (16) sowie $\|Ax\|_2 \le \|A\|_2 \|x\|_2$ für jede Matrix $A \in \mathbb{R}^{n\times n}$ und jeden Vektor $x \in \mathbb{R}^n$.

Beispiel 3.7. Sei (X, d) ein kompakter metrischer Raum. Dann ist $\mathrm{C}(X)$ mit der Supremumsnorm eine Banachalgebra (siehe Beispiel 1.3).

Beispiel 3.8. 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 $(x_k)_{k\in\mathbb{N}_0}$ reeller Zahlen. Der Vektorraum ℓ^1 ist mit dieser Norm ein Banachraum. Die **Faltung** zweier Folgen $x, y \in \ell^1$ ist definiert durch

$$(x * y)_k := \sum_{i=0}^{k} x_i y_{k-i} \qquad \text{für } k \in \mathbb{N}_0.$$

Diese Folge ist ebenfalls absolut summierbar und es gilt $\|x * y\| \le \|x\| \|y\|$. (Dies folgt aus dem grossen Umordnungssatz in [2].) Damit ist ℓ^1 eine Banachalgebra. **Übung:** ℓ^1 ist vollständig.

Übung 3.9. In jeder Banachalgebra gilt $\|\mathbb{1}\| \ge 1$. Für welche der oben genannten Banachalgebren ist das Produkt kommutativ? Ist $\mathcal{A}$ eine Banachalgebra so ist $\mathcal{A}^2$ mit der Normfunktion $\mathcal{A}^2 \to \mathbb{R} : (x, y) \mapsto \|x\| + \|y\|$ ein Banachraum und die Produktabbildung $\mathcal{A}^2 \to \mathcal{A} : (x, y) \mapsto xy$ ist stetig.

4 Potenzreihen in Banachalgebren

Die Begriffe von Konvergenz und Stetigkeit wurden in der Analysis I Vorlesung 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 auf Reihen mit Werten in Banachräumen übertragen. Ist $(X, \|\cdot\|)$ ein Banachraum und $(x_k)_{k\in\mathbb{N}_0}$ eine Folge in X, so nennen wir die Reihe $\sum_{k=0}^{\infty} x_k$ **konvergent**, wenn die Folge der **Partialsummen**

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

konvergiert; in dem Fall bezeichnen wir den Grenzwert mit

$$\sum_{k=0}^{\infty} x_k := \lim_{n\to\infty} s_n.$$

Wir benutzen also die gleiche Bezeichnungsweise wie für Reihen reeller oder komplexer Zahlen, das heisst der Ausdruck “$\sum_{k=0}^{\infty} x_k$” hat zwei Bedeutungen: er steht zum einen für die Folge der Partialsummen $(s_n)_{n\in\mathbb{N}_0}$ und zum anderen für deren Grenzwert, falls dieser existiert. Wir nennen die Reihe $\sum_{k=0}^{\infty} x_k$ **absolut konvergent**, wenn die Reihe $\sum_{k=0}^{\infty} \|x_k\|$ konvergiert. In diesem Fall nennen wir auch die Folge $(x_k)_{k\in\mathbb{N}_0}$ **absolut summierbar**.

Satz 4.1. *Sei $(X, \|\cdot\|)$ ein Banachraum und sei $(x_k)_{k\in\mathbb{N}_0}$ eine Folge in X. Konvergiert die Reihe $\sum_{k=0}^{\infty} x_k$ absolut, so konvergiert sie.*

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

$$n > m \geq n_0 \qquad \Longrightarrow \qquad \sigma_n - \sigma_m = \|x_{m+1}\| + \cdots + \|x_n\| < \varepsilon.$$

Es folgt aus der Dreiecksungleichung, dass für alle $m, n \in \mathbb{N}$ mit $n > m \geq n_0$ die folgende Ungleichung gilt

$$\|s_n - s_m\| = \|x_{m+1} + \cdots + x_n\| \leq \|x_{m+1}\| + \cdots + \|x_n\| < \varepsilon.$$

Damit ist gezeigt, dass $(s_n)_{n\in\mathbb{N}_0}$ eine Cauchy-Folge in X ist. De X ein Banachraum (und somit vollständig) ist, konvergiert diese Folge, und das heisst, dass die Reihe $\sum_{k=0}^{\infty} x_k$ konvergiert. Damit ist Satz 4.1 bewiesen. $\square$

Sei $\mathcal{A}$ eine reelle Banachalgebra und

$$P(z) := \sum_{k=0}^{\infty} a_k z^k \tag{18}$$

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

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

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. Insbesondere definieren wir $x^0 := \mathbb{1}$ für jedes $x \in \mathcal{A}$. Die daraus resultierende Reihe in $\mathcal{A}$ ist

$$P_{\mathcal{A}}(x) := \sum_{k=0}^{\infty} a_k x^k. \tag{19}$$

Man kann nun die Frage stellen, für welche $x \in \mathcal{A}$ die Reihe (19) konvergiert.

Korollar 4.2. *Die Reihe (19) konvergiert für jedes $x \in \mathcal{A}$ mit $\|x\| < \rho$.*

Beweis. Sei $x \in \mathcal{A}$ mit $\|x\| < \rho$. Es folgt aus der Bedingung (16) in der Definition einer Banachalgebra, dass die Ungleichung

$$\left\|a_k x^k\right\| \le |a_k| \, \|x\|^k$$

für jedes $k \in \mathbb{N}$ gilt. Da $\|x\| < \rho$ ist, konvergiert die Reihe $\sum_{k=0}^{\infty} |a_k| \, \|x\|^k$. Nach dem Majorantenkriterium konvergiert auch die Reihe $\sum_{k=0}^{\infty} \left\|a_k x^k\right\|$. Dies heisst aber genau, dass die Reihe $\sum_{k=0}^{\infty} a_k x^k$ in $\mathcal{A}$ absolut konvergiert und somit, nach Satz 4.1, konvergiert. $\square$

Wir haben also gezeigt, dass die Formel (19) eine Abbildung

$$P_{\mathcal{A}} : \mathcal{B}_\rho := \{x \in \mathcal{A} \mid \|x\| < \rho\} \to \mathcal{A} \tag{20}$$

definiert, wobei ρ der Konvergenzradius von P ist. Das nächste Lemma zeigt, dass diese Abbildung stetig ist. Zunächst bemerken 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 4.3. *Sei $I \subset \mathbb{R}$ ein Intervall und $(X, \|\cdot\|)$ ein Banachraum. Eine Funktion $f : I \to X$ heisst* **differenzierbar an der Stelle** $t \in I$ *mit* **Ableitung** $a =: f'(t) \in X$ *wenn gilt:* $\forall \varepsilon > 0\ \exists \delta > 0\ \forall h \in \mathbb{R}$*:*

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

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

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

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

Satz 4.4. *Sei $\mathcal{A}$ eine reelle Banachalgebra und sei* (18) *eine Potenzreihe mit reellen Koeffizienten und Konvergenzradius $\rho > 0$. Dann gilt folgendes*

(i) *Für jedes $r < \rho$ ist die Restriktion der Abbildung (20) auf die Teilmenge $\overline{\mathcal{B}}_r := \{x \in \mathcal{A} \,|\, \|x\| \le r\}$ Lipschitz-stetig; es gilt*

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

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

(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*

$$\frac{d}{dt} P_{\mathcal{A}}(tx) = x Q_{\mathcal{A}}(tx), \qquad Q(z) := P'(z) := \sum_{k=1}^{\infty} k a_k z^{k-1}, \tag{21}$$

für alle $t \in \mathbb{R}$ mit $|t|\,\|x\| < \rho$.

Beweis. Wir beweisen (i). Es gilt $x^k - y^k = \sum_{j=1}^{k} x^{k-j}(x-y)y^{j-1}$ für alle $k \in \mathbb{N}$ und alle $x, y \in \mathcal{A}$, und daher

$$\|x\| \le r, \ \|y\| \le r \qquad \Longrightarrow \qquad \left\|x^k - y^k\right\| \le k r^{k-1} \|x - y\|.$$

Hieraus folgt

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

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

Zum Beweis von (ii) beachten wir, dass $Q_{\mathcal{A}}(x) = \sum_{k=1}^{\infty} ka_k x^{k-1}$ ist. Daher gilt

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

Hier bezeichnet p die Potenzreihe

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

mit dem Konvergenzradius $\rho / \|x\|$. Sei $t \in \mathbb{R}$ mit $|t| < \rho / \|x\|$. Dann existieren zwei Konstanten $\delta > 0$ und $c > 0$ so dass $|t| + \delta < \rho / \|x\|$ ist und jedes $h \in \mathbb{R}$ mir $|h| < \delta$ die Ungleichung

$$|p(|t| + |h|) - p(|t|) - |h| \, \dot{p}(|t|)| \leq c \, |h|^2$$

erfüllt. Daraus folgt

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

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

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

ist und daraus folgt (ii). Damit ist Satz 4.4 bewiesen. $\square$

Betrachten wir den Spezialfall $\mathcal{A} := \mathbb{R}^{n \times n}$ mit einer Matrixnorm $\|\cdot\|$. Dann folgt aus Satz 4.4, 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 1.8 unabhängig von der Wahl der Norm ist, benötigen wir bei der Voraussetzung ($\|A\| < \rho$) eine Matrixnorm.

5 Die geometrische Reihe

Die **geometrische Reihe** hat den Konvergenzradius $\rho = 1$ und stellt die Funktion

$$\sum_{k=0}^{\infty} z^k = \frac{1}{1-z} \qquad \text{für } z \in \mathbb{C} \text{ mit } |z| < 1$$

dar. Nach Satz 4.4 ergibt sich daraus folgendes Resultat für Banachalgebren.

Satz 5.1. *Sei $\mathcal{A}$ eine reelle Banachalgebra mit $\|\mathbb{1}\| = 1$ und sei $\mathcal{G} \subset \mathcal{A}$ die Gruppe der invertierbaren Elemente. Dann gilt folgendes.*

(i) *Sei $x \in \mathcal{A}$ mit $\|x\| < 1$. Dann ist $\mathbb{1} - x \in \mathcal{G}$ und*

$$(\mathbb{1} - x)^{-1} = \sum_{k=0}^{\infty} x^k. \tag{22}$$

(ii) *Sei $x \in \mathcal{G}$ und $y \in \mathcal{A}$ so dass*

$$\|x - y\| \left\|x^{-1}\right\| < 1. \tag{23}$$

Dann ist $y \in \mathcal{G}$ und

$$y^{-1} = x^{-1} \sum_{k=0}^{\infty} \left(\mathbb{1} - yx^{-1}\right)^k, \qquad \left\|x^{-1} - y^{-1}\right\| \le \frac{\|x^{-1}\|^2 \|x - y\|}{1 - \|x^{-1}\| \|x - y\|}. \tag{24}$$

(iii) *Die Gruppe $\mathcal{G}$ der invertierbaren Elemente in $\mathcal{A}$ ist eine offene Teilmenge von $\mathcal{A}$ und die Abbildung $\mathcal{G} \to \mathcal{G} : x \mapsto x^{-1}$ ist stetig.*

Beweis. Sei $x \in \mathcal{A}$ mit $\|x\| < 1$. Dann gilt

$$\sum_{k=0}^{\infty} \left\|x^k\right\| \le \sum_{k=0}^{\infty} \|x\|^k = \frac{1}{1 - \|x\|} < \infty.$$

Daher konvergiert Reihe $\sum_{k=0}^{\infty} x^k$ nach Satz 4.1. Für alle $n \in \mathbb{N}$ gilt

$$(\mathbb{1} - x) \sum_{k=0}^{n} x^k = \sum_{k=0}^{n} x^k (\mathbb{1} - x) = \sum_{k=0}^{n} x^k - \sum_{k=0}^{n} x^{k+1} = \mathbb{1} - x^{n+1}$$

und daraus folgt $(\mathbb{1} - x) \sum_{k=0}^{\infty} x^k = \sum_{k=0}^{\infty} x^k (\mathbb{1} - x) = \mathbb{1}$. Also ist $\mathbb{1} - x \in \mathcal{G}$ und $(\mathbb{1} - x)^{-1} = \sum_{k=0}^{\infty} x^k$ und damit ist (i) bewiesen.

Nun seien $x \in \mathcal{G}$ und $y \in \mathcal{A}$ gegeben mit $\|x - y\| \|x^{-1}\| < 1$. Dann gilt

$$\left\|\mathbb{1} - yx^{-1}\right\| = \left\|(x-y)x^{-1}\right\| \leq \|x - y\| \left\|x^{-1}\right\| < 1.$$

Nach Teil (i) folgt daraus $yx^{-1} \in \mathcal{G}$ und $(yx^{-1})^{-1} = \sum_{k=0}^{\infty}(\mathbb{1} - yx^{-1})^k$. Daher ist $y \in \mathcal{G}$ und es gilt

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

und

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

Damit ist Teil (ii) bewiesen und Teil (iii) folgt unmittelbar aus (ii). Damit ist Satz 5.1 bewiesen. □

Beispiel 5.2. Wir betrachten den Fall $\mathcal{A} = \mathbb{R}^{n\times n}$. Jede Matrix $A \in \mathbb{R}^{n\times n}$ induziert eine lineare Abbldung $T_A : \mathbb{R}^n \to \mathbb{R}^n$, die durch $T_A x := Ax$ für $x \in \mathbb{R}^n$ definiert ist. Nach Definition ist eine Matrix $A \in \mathbb{R}^{n\times n}$ genau dann ein invertierbares Element von $\mathcal{A}$, wenn eine Matrix $B \in \mathbb{R}^{n\times n}$ existiert, welche die Gleichungen $AB = BA = \mathbb{1}$ erfüllt. Solch eine Matrix B existiert genau dann wenn die lineare Abbildung T_A bijektiv ist. In diesem Fall induziert die inverse Matrix $A^{-1} = B$ die Umkehrabbildung $T_A^{-1} = T_{A^{-1}}$. Aus der *Linearen Algebra* wissen wir, dass die invertierbaren Matrizen gerade die mit Determinante ungleich Null sind (siehe [3]). Daher ist die Gruppe der invertierbaren Elemente in $\mathcal{A} = \mathbb{R}^{n\times n}$ durch

$$\mathrm{GL}(n, \mathbb{R}) := \left\{A \in \mathbb{R}^{n\times n} \mid \det(A) \neq 0\right\}$$

gegeben. Diese ist die **allgemeine lineare Gruppe**. Nach Satz 5.1 ist dies eine offene Teilmenge von $\mathbb{R}^{n\times n}$ und die Abbildung

$$\mathrm{GL}(n, \mathbb{R}) \to \mathrm{GL}(n, \mathbb{R}) : A \mapsto A^{-1}$$

ist stetig

6 Die Exponentialabbildung

Die Potenzreihe $\exp(z) := \sum_{k=0}^{\infty} z^k/k!$ hat den Konvergenzradius $\rho = \infty$ und induziert damit eine Exponentialabbildung auf jeder Banachalgebra.

Definition 6.1. *Sei $\mathcal{A}$ eine reelle Banachalgebra. Die* **Exponentialabbildung auf** $\mathcal{A}$ *ist die durch*

$$\exp(x) := e^x := \sum_{k=0}^{\infty} \frac{x^k}{k!} \qquad \text{für } x \in \mathcal{A} \tag{25}$$

definierte Abbildung

$$\exp : \mathcal{A} \to \mathcal{A}.$$

Nach Satz 4.4 konvergiert die Reihe in (25) absolut für jedes $x \in \mathcal{A}$.

Satz 6.2. *Sei $\mathcal{A}$ eine reelle Banachalgebra. Dann ist die Exponentialabbildung $\exp : \mathcal{A} \to \mathcal{A}$ in (25) stetig und hat folgende Eigenschaften.*

(i) *Für jedes $x \in \mathcal{A}$ und jede Folge $(x_n)_{n\in\mathbb{N}}$ in $\mathcal{A}$ gilt*

$$\lim_{n\to\infty} x_n = x \qquad \Longrightarrow \qquad \exp(x) = \lim_{n\to\infty} \left(\mathbb{1} + \frac{x_n}{n}\right)^n. \tag{26}$$

(ii) *Es gilt $\exp(0) = \mathbb{1}$ und, für alle $x, y \in \mathcal{A}$,*

$$xy = yx \qquad \Longrightarrow \qquad \exp(x+y) = \exp(x)\exp(y). \tag{27}$$

(iii) *Für jedes $x \in \mathcal{A}$ gilt*

$$\exp(x) \in \mathcal{G}, \qquad \exp(x)^{-1} = \exp(-x). \tag{28}$$

(iv) *Für jedes $x \in \mathcal{A}$ und jedes $y \in \mathcal{G}$ gilt*

$$\exp(yxy^{-1}) = y\exp(x)y^{-1}. \tag{29}$$

(v) *Für jedes $x \in \mathcal{A}$ ist die Funktion $\mathbb{R} \to \mathcal{A} : t \mapsto \exp(tx)$ stetig differenzierbar und ihre Ableitung ist*

$$\frac{d}{dt}\exp(tx) = x\exp(tx). \tag{30}$$

Beweis. Die Stetigkeit der Exponentialabbildung folgt aus Satz 4.4 und die Aussage (i) beweist man wie im Fall $\mathcal{A} = \mathbb{R}$ (siehe [2]). Hier ist noch einmal das Argument zum Beweis von (i). Sei $\varepsilon > 0$ gegeben und sei $N \in \mathbb{N}$ so gewählt, dass $\|x_n - x\| \leq 1$ ist für jede natürliche Zahl $n \geq N$ und

$$\sum_{k=N}^{\infty} \frac{(\|x\|+1)^k}{k!} < \frac{\varepsilon}{3}. \tag{31}$$

Dann gilt $\lim_{n\to\infty} \binom{n}{k} \frac{1}{n^k} = \frac{1}{k!}$ für jedes $k \in \{0, 1, \ldots, N-1\}$. Daher existiert eine ganze Zahl $n_0 \geq N$ so dass für alle $n \in \mathbb{N}$ folgendes gilt:

$$n \geq n_0 \qquad \Longrightarrow \qquad \sum_{k=0}^{N-1} \left\| \binom{n}{k} \frac{x_n^k}{n^k} - \frac{x^k}{k!} \right\| < \frac{\varepsilon}{3}. \tag{32}$$

Sei nun $n \in \mathbb{N}$ mit $n \geq n_0$. Dann ist $n \geq N$ und es gilt

$$\begin{aligned}
\left\| \left(\mathbb{1} + \frac{x_n}{n} \right) - e^x \right\| &= \left\| \sum_{k=0}^{n} \binom{n}{k} \frac{x_n^k}{n^k} - \sum_{k=0}^{\infty} \frac{x^k}{k!} \right\| \\
&\leq \sum_{k=0}^{N-1} \left\| \binom{n}{k} \frac{x_n^k}{n^k} - \frac{x^k}{k!} \right\| + \sum_{k=N}^{n} \binom{n}{k} \frac{\|x_n\|^k}{n^k} + \sum_{k=N}^{\infty} \frac{\|x\|^k}{k!} \\
&\leq \sum_{k=0}^{N-1} \left\| \binom{n}{k} \frac{x_n^k}{n^k} - \frac{x^k}{k!} \right\| + 2 \sum_{k=N}^{\infty} \frac{(\|x\|+1)^k}{k!} \\
&< \varepsilon.
\end{aligned}$$

Hier folgt der vorletzte Schritt aus den Ungleichungen $\|x_n\| \leq \|x\| + 1$ für $n \geq N$ und $\binom{n}{k} \frac{1}{n^k} \leq \frac{1}{k!}$ für $n \in \mathbb{N}$ mit $n \geq k$. Der letzte Schritt folgt aus (31) und (32). Damit ist (i) bewiesen.

Die Aussage (ii) beweist man ebenfalls genau wie im Fall $\mathcal{A} = \mathbb{R}$ (siehe [2]). Man kann diese Aussage jedoch auch direkt aus (i) herleiten, denn wenn $x, y \in \mathcal{A}$ kommutieren so dass $xy = yx$ ist, dann gilt

$$\left(\mathbb{1} + \frac{x}{n} \right)^n \left(\mathbb{1} + \frac{y}{n} \right)^n = \left(\mathbb{1} + \frac{x + y + xy/n}{n} \right)^n$$

für jedes $n \in \mathbb{N}$. Beim Grenzübergang $n \to \infty$ ergibt sich daraus nach (i) die gewünschte Gleichung $\exp(x)\exp(y) = \exp(x+y)$.

Teil (iii) folgt unmittelbar aus Teil (ii) mit $y = -x$, Teil (iv) folgt aus der Gleichung $(yxy^{-1})^k = yx^ky^{-1}$ für $y \in \mathcal{G}$, und Teil (v) folgt aus Satz 4.4. Damit ist Satz 6.2 bewiesen. $\square$

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

$$\begin{aligned}
\dot{x}_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n, \qquad & x_1(0) &= x_{01}, \\
\dot{x}_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n, \qquad & x_2(0) &= x_{02}, \\
&\vdots \\
\dot{x}_n &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n, \qquad & x_n(0) &= x_{0n}.
\end{aligned} \tag{33}$$

Hier sind die $a_{ij} \in \mathbb{R}$ konstante reelle Koeffizienten und, für jedes i, ist die Variable x_i eine stetig differenzierbare Abbildung $x_i : \mathbb{R} \to \mathbb{R}$. Wir stellen uns das Argument dieser Abbildung als Zeitvariable t vor und bezeichnen mit $\dot{x}_i = dx_i/dt$ die Ableitung der Funktion x_i nach der Zeit t. In der physikalischen Interpretation bezeichnet das n-Tupel

$$x(t) := (x_1(t), \ldots, x_n(t)) \in \mathbb{R}^n$$

den Ortsvektor eines Teilchens (oder mehrerer Teilchen gleichzeitig) zum Zeitpunkt t und die Ableitung $\dot{x}(t) = (\dot{x}_1(t), \ldots, \dot{x}_n(t)) \in \mathbb{R}^n$ den Geschwindigkeitsvektor zum Zeitpunkt t. Fasst man die konstanten Koeffizienten zu einer Matrix $A = (a_{ij})_{i,j=1}^n \in \mathbb{R}^{n\times n}$ zusammen so lässt sich das Differentialgleichungssystem (33) auch in der Kurzform

$$\dot{x} = Ax, \qquad x(0) = x_0, \tag{34}$$

schreiben. Hier bezeichnet $x_0 := (x_{01}, \ldots, x_{0n}) \in \mathbb{R}^n$ den **Anfangsvektor** (oder den Ortsvektor zum Zeitpunkt $t = 0$). Es folgt nun aus der Gleichung (30), dass die eindeutige Lösung der Differentialgleichung (34) durch die Formel

$$x(t) := e^{At}x_0 \tag{35}$$

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 $\mathcal{A} = \mathbb{R}^{n\times n}$ (siehe Satz 6.2).

Beispiel 7.1. Für
$$\Lambda := \operatorname{diag}(\lambda_1, \dots, \lambda_n)$$
gilt $\Lambda^k = \operatorname{diag}(\lambda_1^k, \dots, \lambda_n^k)$ und damit
$$e^\Lambda = \sum_{k=0}^\infty \frac{\Lambda^k}{k!} = \operatorname{diag}(e^{\lambda_1}, \dots, e^{\lambda_n}).$$

Beispiel 7.2. Für $\lambda \in \mathbb{R}$ gilt
$$A = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} \qquad \Longrightarrow \qquad e^{At} = e^{\lambda t} \begin{pmatrix} 1 & t & t^2/2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix}.$$

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 7.3. Für
$$H := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$
gilt $H^{2k} = (-1)^k \mathbb{1}$ und damit
$$\begin{aligned} e^{Ht} &= \left(1 - \frac{t^2}{2!} + \frac{t^4}{4!} \mp \cdots \right) \mathbb{1} + \left(t - \frac{t^3}{3!} + \frac{t^5}{5!} \mp \cdots \right) H \\ &= \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}. \end{aligned}$$

(Siehe [2] für die Potenzreihendarstellungen von Cosinus und Sinus.) Da die Matrizen $a\mathbb{1}$ und bH kommutieren, gilt für die Matrix
$$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$
die Gleichung
$$e^{tA} = e^{at} \begin{pmatrix} \cos(bt) & -\sin(bt) \\ \sin(bt) & \cos(bt) \end{pmatrix}$$
für alle $t \in \mathbb{R}$.

8 Komposition von Potenzreihen

Seien $(a_k)_{k\geq 1}$ und $(b_\ell)_{\ell\geq 0}$ zwei Folgen reeller Zahlen, so dass die Potenzreihen

$$P(z) := \sum_{k=1}^{\infty} a_k z^k, \qquad Q(z) := \sum_{\ell=1}^{\infty} b_\ell z^\ell$$

positive Konvergenzradien haben, das heisst

$$\rho_P := \frac{1}{\limsup_{k\to\infty} |a_k|^{1/k}} > 0, \qquad \rho_Q := \frac{1}{\limsup_{\ell\to\infty} |b_\ell|^{1/\ell}} > 0.$$

Insbesondere gilt $P(0) = 0$. Sei $0 < \rho \leq \rho_P$ so gewählt, dass

$$z \in \mathbb{C}, \quad |z| < \rho \qquad \Longrightarrow \qquad \sum_{k=1}^{\infty} |a_k||z|^k < \rho_Q. \tag{36}$$

Ist $z \in \mathbb{C}$ mit Betrag $|z| < \rho \leq \rho_P$ so gilt $|P(z)| < \rho_Q$ nach (36). Damit is die Komposition $Q \circ P : \{z \in \mathbb{C} \,|\, |z| < \rho\} \to \mathbb{C}$, wohldefiniert. Der folgende Satz besagt, dass sich diese Komposition als Potenzreihe darstellen lässt.

Satz 8.1. *Für $k \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$ sei*

$$c_k := \sum_{\ell=0}^{k} b_\ell \ell! \sum_{\substack{\sum_i m_i = \ell \\ \sum_i i m_i = k}} \prod_{\substack{i\in\mathbb{N} \\ m_i > 0}} \frac{a_i^{m_i}}{m_i!}. \tag{37}$$

Die zweite Summe in (37) ist über alle Folgen $(m_i)_{i\in\mathbb{N}}$ nichtnegativer ganzer Zahlen zu verstehen, die die Gleichungen $\sum_i i m_i = k$ und $\sum_i m_i = \ell$ erfüllen. Da $m_i \leq k$ ist für alle i und $m_i = 0$ ist für $i > k$, ist diese Summe endlich und verschwindet für $\ell > k$. Sei

$$R(z) := \sum_{k=0}^{\infty} c_k z^k \tag{38}$$

die Potenzreihe mit den Koeffizienten c_k. Dann gilt folgendes.

(i) *Der Konvergenzradius $\rho_R := (\limsup_{k\to\infty} |c_k|^{1/k})^{-1}$ der Potenzreihe (38) ist grösser oder gleich ρ und $R(z) = Q(P(z))$ für alle $z \in \mathbb{C}$ mit $|z| < \rho$.*

(ii) *Sei $\mathcal{A}$ eine reelle Banachalgebra. Dann erfüllt jedes $x \in \mathcal{A}$ mit $\|x\| < \rho$ die Ungleichung $\|P_{\mathcal{A}}(x)\| < \rho_Q$ und die Gleichung*

$$R_{\mathcal{A}}(x) = Q_{\mathcal{A}}(P_{\mathcal{A}}(x)). \tag{39}$$

Beweis. Für $k \in \mathbb{N}_0$ und $N \in \mathbb{N}$ seien die reellen Koeffizienten $c_{k,N}$ durch

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

gegeben. Dann gilt $c_{k,N} = c_k$ für $k \le N$ und $c_{k,N} = 0$ für $k > N^2$ (weil jedes N-Tupel $(m_1, \ldots, m_N) \in \mathbb{N}_0^N$ mit $\sum_i m_i = \ell \le N$ die Ungleichung $k = \sum_i im_i \le N \sum_i m_i = N\ell \le N^2$ erfüllt). Insbesondere ist die Funktion

$$R_N(z) := \sum_{k=0}^{\infty} c_{k,N} z^k$$

ein Polynom. Für $N \in \mathbb{N}$ seien die Polynome P_N und Q_N durch

$$P_N(z) := \sum_{k=1}^{N} a_k z^k, \qquad Q_N(z) := \sum_{\ell=0}^{N} b_\ell z^\ell$$

für $z \in \mathbb{C}$ definiert. Dann gilt für alle $z \in \mathbb{C}$ die Gleichung

$$\begin{aligned}
R_N(z) &= \sum_{k=0}^{\infty} c_{k,N} z^k \\
&= \sum_{k=0}^{\infty} \sum_{\ell=0}^{N} b_\ell \ell! \sum_{\substack{m_1+\cdots+m_N=\ell \\ \sum_{i=1}^N im_i=k}} \frac{a_1^{m_1} a_2^{m_2} \cdots a_N^{m_N}}{m_1! \cdots m_N!} z^k \\
&= \sum_{\ell=0}^{N} b_\ell \sum_{m_1+\cdots+m_N=\ell} \frac{\ell!}{m_1! \cdots m_N!} (a_1 z)^{m_1} (a_2 z^2)^{m_2} \cdots (a_N z^N)^{m_N} \\
&= \sum_{\ell=0}^{N} b_\ell \left(a_1 z + a_2 z^2 + \cdots + a_N z^N\right)^\ell \\
&= Q_N(P_N(z)).
\end{aligned}$$

Sei nun $0 < r < \rho$. Dann gilt $s := \sum_{k=1}^\infty |a_k| r^k < \rho_Q$. Sei ausserdem $z \in \mathbb{C}$ mit $|z| \le r$. Dann konvergiert $P_N(z)$ gegen $P(z)$ und es ist $|P_N(z)| \le s$ für alle $N \in \mathbb{N}$. Da Q_N auf dem Ball $\{z \in \mathbb{C} \,|\, |z| \le s\}$ gleichmässig gegen Q konvergiert, folgt daraus, dass $Q_N(P_N(z))$ gegen $Q(P(z))$ konvergiert, und daher gilt

$$\lim_{N \to \infty} R_N(z) = Q(P(z)) \qquad \text{für alle } z \in \mathbb{C} \text{ mit } |z| \le r. \tag{40}$$

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

$$|c_k| \le \bar{c}_k := \sum_{\ell=0}^{k} |b_\ell|\, \ell! \sum_{\substack{\sum_i m_i=\ell \\ \sum_i i m_i = k}} \prod_{\substack{i\in\mathbb{N} \\ m_i>0}} \frac{a_i^{m_i}}{m_i!},$$

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

Da $\bar{c}_{k,N} \le \bar{c}_k$ ist für alle k, N sowie $\bar{c}_{k,N} = \bar{c}_k$ ist für alle $k \le N$ und $\bar{c}_{k,N} = 0$ ist für alle $k > N^2$, folgt daraus, für jedes $N' \in \mathbb{N}$,

$$N' \le \sqrt{N} \quad \Longrightarrow \quad \sum_{k=0}^{\infty} \bar{c}_{k,N'} r^k \le \sum_{k=0}^{N} \bar{c}_k r^k \le \sum_{k=0}^{\infty} \bar{c}_{k,N} r^k. \tag{41}$$

Hier konvergieren die beiden äusseren Terme gegen

$$\sum_{\ell=0}^{\infty} |b_\ell| \left(\sum_{k=1}^{\infty} |a_k|\, r^k \right)^{\ell} = \sum_{\ell=0}^{\infty} |b_\ell|\, s^\ell$$

für $N, N' \to \infty$, nach dem gleichen Argument wie oben. Also konvergiert die Differenz $\sum_{k>N} \bar{c}_{k,N} r^k$ der letzten beiden Terme in (41) gegen Null, und für $z \in \mathbb{C}$ mit $|z| \le r$ gilt daher

$$\left| R_N(z) - \sum_{k=0}^{N} c_k z^k \right| = \left| \sum_{k=N+1}^{\infty} c_{k,N} z^k \right| \le \sum_{k=N+1}^{\infty} \bar{c}_{k,N} r^k \stackrel{N\to\infty}{\longrightarrow} 0. \tag{42}$$

Es folgt aus (40) und (42), dass die Gleichung

$$R(z) = \sum_{k=0}^{\infty} c_k z^k = \lim_{N\to\infty} \sum_{k=0}^{N} c_k z^k = \lim_{N\to\infty} R_N(z) = Q(P(z))$$

für jedes $z \in \mathbb{C}$ mit $|z| \le r$ gilt. Da r als beliebige reelle Zahl $0 < r < \rho$ gewählt war, folgt daraus, dass der Konvergenzradius der Potenzreihe R die Ungleichung $\rho_R \ge \rho$ erfüllt und dass $R(z) = Q(P(z))$ ist für jedes $z \in \mathbb{C}$ mit $|z| < \rho$. Damit ist (i) bewiesen. Teil (ii) folgt aus dem gleichen Argument, indem man $z \in \mathbb{C}$ durch $x \in \mathcal{A}$ und den Betrag einer komplexen Zahl durch die Norm in $\mathcal{A}$ ersetzt. Damit ist Satz 8.1 bewiesen. $\square$

Beispiel 8.2 (Logarithmus). Die Reihe

$$L(z) := \sum_{k=1}^{\infty} \frac{(-1)^{k-1} z^k}{k} = z - \frac{z^2}{2} + \frac{z^3}{3} - \frac{z^4}{4} \pm \cdots$$

hat den Konvergenzradius $\rho_L = 1$ und stellt die Funktion

$$L(z) = \log(1+z)$$

für $z \in \mathbb{C}$ mit $|z| < 1$ dar. Das heisst, L erfüllt die Gleichung

$$\exp(L(z)) = 1 + z \qquad \text{für alle } z \in \mathbb{C} \text{ mit } |z| < 1.$$

(Siehe [1, Seite 44, Übung 1.7.18].) Nach Teil (i) von Satz 8.1 folgt daraus zunächst, dass die Koeffizienten $c_0 = c_1 = 1$ und $c_k = 0$ für $k \geq 2$ der Potenzreihe $1 + z$ durch die Formel (37) mit $a_i = (-1)^{i-1}/i$ und $b_\ell = 1/\ell!$ gegeben sind. Das heisst, für alle $k \in \mathbb{N}_0$ gilt

$$\sum_{\ell=0}^{k} (-1)^{k-\ell} \sum_{\substack{\sum_i m_i = \ell \\ \sum_i i m_i = k}} \prod_{\substack{i \in \mathbb{N} \\ m_i > 0}} \frac{1}{m_i! i^{m_i}} = \begin{cases} 1, \text{ falls } k = 0, 1, \\ 0, \text{ falls } k \geq 2. \end{cases} \tag{43}$$

Weiterhin folgt aus Teil (ii) von Satz 8.1 dass, wenn $\mathcal{A}$ eine Banachalgebra ist, jedes $x \in \mathcal{A}$ mit $\|x\| < 1$ die Gleichung

$$\exp\left(\sum_{k=1}^{\infty} \frac{(-1)^{k-1} x^k}{k} \right) = \mathbb{1} + x \tag{44}$$

erfüllt. Insbesondere gilt die Gleichung (44) für jede Matrix $x \in \mathbb{R}^{n \times n}$ mit $\|x\| < 1$, wobei $\|\cdot\|$ eine beliebige Matrixnorm auf $\mathbb{R}^{n \times n}$ ist.

Schlussbemerkungen

Für den Satz über die Existenz und Eindeutigkeit von Lösungen gewöhnlicher Differentialgleichungen sei an die Vorlesung Analysis I erinnert. Dieser Satz wurde dort in viel grösserer Allgemeinheit (auch für nichtlineare Differentialgleichungen) bewiesen. Formeln wie in den Beispielen im Abschnitt 7 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, das heisst 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**, einem wichtigen 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.

Literatur

[1] K. Königsberger, *Analysis 2*, 5. Auflage, Springer Verlag, 2003.

[2] D.A. Salamon, *Summierbare Familien*, ETHZ, Herbstsemester 2016.
`http://www.math.ethz.ch/~salamon/PREPRINTS/ana1-sum.pdf`

[3] D.A. Salamon, *Die Determinante*, ETHZ, Frühjahrssemester 2017.
`http://www.math.ethz.ch/~salamon/PREPRINTS/ana2-det.pdf`