Technische Universität Braunschweig SS 22
Institut für Analysis und Algebra
Prof. Dr. Volker Bach, Dr. Konstantin Merz

# Tutoriumsblatt zur Vorlesung Lineare Algebra 2

Hilfreiche Literatur: [Beu94, Fis05, MK08, Dym13, Bos14, Bos20, SP08]

## 1 1. Tutoriumsblatt (28. April 2022)

### 1.1 Euklidsche Ringe

**Definition 1.1.** 1. Ein kommutativer Ring $(R, +, \cdot)$ mit Einselement heißt *nullteilerfrei* oder *Integritätsbereich*, wenn für alle $a, b \in R$ mit $a \cdot b = 0$ folgt, dass $a = 0$ oder $b = 0$ ist.

2. Für einen Integritätsbereich $(R, +, \cdot)$ heißt eine Abbildung $\rho : \mathbb{R} \setminus \{0\}$ *Ringnorm*, wenn für alle $a, b \in R$ gilt, dass $\rho(ab) \geq \rho(a)$.

3. Ein Integritätsbereich $(R, +, \cdot)$ mit Ringnorm $\rho$ heißt *euklidscher Ring*, wenn der Divisionsalgorithmus gilt: für alle $a, b \in R$ mit $b \neq 0$ gibt es $q, r \in R$, sodass $a = qb + r$ und $\rho(r) < \rho(b)$.

*Bemerkung* 1.2. Der Restklassenring $\mathbb{Z}_9$ ist nicht nullteilerfrei, denn $[0] = [9] = [3 \cdot 3] = [3] \cdot [3]$.

Hier sind einige weitere Begriffe und Resultate; siehe beispielsweise Schulze–Pillot [SP08, Kapitel 0–2].

**Definition 1.3.** 1. Sei $R$ ein kommutativer Ring mit Einselement. Ein Element $a \in R$ heißt *Einheit*, wenn es $a' \in R$ gibt, sodass $aa' = 1$. Die Menge aller Einheiten wird mit $R^\times$ bezeichnet.

2. Sei $R$ ein Integritätsbereich. Dann heißen $a, b \in R$ *(zueinander) assoziiert*, wenn es eine Einheit $\varepsilon \in R^\times$ gibt, sodass $b = a\varepsilon$.

*Bemerkung* 1.4. 1. Ein Integritätsbereich, in dem $R^\times = R \setminus \{0\}$ ist, ist ein Körper.

2. Die Menge aller Einheiten in $\mathbb{Z}$ ist $\mathbb{Z}^\times = \{\pm 1\}$.

3. Die Einheiten im Ring $C(\mathbb{R} : \mathbb{C})$ aller stetigen Funktionen sind alle Funktionen ohne Nullstellen.

Sind $R$ ein Integritätsbereich und $a, b \in R$, so sagen wir, dass $b$ durch $a$ geteilt wird (oder $a$ teilt $b$), wenn es $c \in R$ gibt, sodass $b = ac$. Man schreibt $a|b$.

**Lemma 1.5.** *Sei $R$ ein Integritätsbereich. Dann gelten folgende Aussagen.*

1. *Zwei Elemente $a, b \in R$ sind genau dann assoziiert, wenn $a|b$ und $b|a$ gelten.*

2. *In $R = \mathbb{Z}$ sind $a, b \in \mathbb{Z}$ genau dann assoziiert, wenn $|a| = |b|$.*

3. *Zwei Polynome $f, g \in R[X]$ sind genau dann assoziiert, wenn sie durch Multiplikation mit einer Einheit von $R$ auseinander hervorgehen, wenn es also $\varepsilon \in R$ gibt, sodass $f = \varepsilon g$.*

4. *Ist $R = K$ ein Körper, so sind zwei Polynome $f, g \in K[X]$ genau dann assoziiert, wenn sie skalare Vielfache voneinander sind.*

5. *Für $R = K$ sind alle $a, b \in K \setminus \{0\}$ zueinander assoziiert.*

**Beispiel 1.6.** 1. $(\mathbb{Z}, +, \cdot)$ ist ein euklidscher Ring mit $\rho(x) = |x|$ (Betrag) für alle $x \in \mathbb{Z}$.

2. $(K[X], +, \cdot)$ ist ein euklidscher Ring mit $\rho(f) = \deg(f)$ für alle $f \in K[X]$.

3. $(K, +, \cdot)$ ist ein euklidscher Ring mit $\rho(a) = 1$ für alle $a \in K$.

4. Ein weiteres Beispiel für einen Integritätsbereich liefert

$$\mathbb{Z}[\sqrt{d}] := \{a + \sqrt{d} \cdot b : \; a, b \in \mathbb{Z}\} \subseteq \mathbb{C} \tag{1.1}$$

für beliebiges $d \in \mathbb{Z}$. Offenbar ist $\mathbb{Z}[\sqrt{d}]$ ein Teilring des Körpers $\mathbb{C}$. (Besonders interessant ist die Situation, wenn $d$ kein Quadrat ist.) In jedem Fall ist $\mathbb{Z}[\sqrt{d}]$ ein Integritätsbereich. (Dies folgt aus der Tatsache, dass jeder Körper ein Integritätsbereich ist und, dass Teilringe von Körpern wieder nullteilerfrei sein müssen.)

Für $d = -1$ erhält man die *gaußschen Zahlen*. Sie bilden sogar einen euklidschen Ring mit Ringnorm $\rho(a+ib) := a^2 + b^2$. In der Tat gilt der Divisionsalgorithmus: seien dazu $a, b \in \mathbb{Z}$ und $c, d \in \mathbb{Z} \setminus \{0\}$. Definiere $\mathbb{C} \ni x + iy := \frac{a+ib}{c+id}$. Dann finden wir nächste Nachbarn $p, q \in \mathbb{Z}$ von $x, y$ so, dass $|x-p| \vee |y-q| \leq 1/2$. (Wir erinnern an die Notation $a \vee b := \max\{a, b\}$ und $a \wedge b := \min\{a, b\}$.) Dann gilt $|x + iy - (p + iq)|^2 = |x-p|^2 + |y-q|^2 \leq 1/2 < 1$. Multipliziert man dies mit $|c + id|$, erhält man

$$|\underbrace{(a+ib) - (c+id)(p+iq)}_{=:r+is}|^2 < |c+id|^2 , \tag{1.2}$$

was $\rho(r + is) < \rho(c + id)$ zeigt, wie gewünscht.

## 1.2 Ggt und euklidscher Algorithmus

**Definition 1.7.** 1. Sei $R$ ein Integritätsbereich und $a, b \in R$. Dann heißt $g \in R \setminus \{0\}$ *gemeinsamer Teiler* von $a, b$, wenn $g|a \wedge g|b$. Dieser gemeinsame Teiler heißt *größter gemeinsamer Teiler* von $a, b$, wenn für alle weiteren gemeinsamen Teiler $\tilde{g}$ gilt, dass $\tilde{g}|g$. Manchmal schreibt man (im Gegensatz zur Vorlesung!) $g = \mathrm{ggT}(a, b)$.

2. Ein Element $m \in R$ heißt *gemeinsames Vielfaches* von $a, b$, wenn $a|m \wedge b|m$. Es heißt *kleinstes gemeinsames Vielfaches*, wenn für alle weiteren gemeinsamen Vielfachen $\tilde{m} \in R$ gilt, dass $m|\tilde{m}$. Wir schreiben $m = \mathrm{kgV}(a, b)$.

**Lemma 1.8.** *Sei $R$ ein Integritätsbereich und $a, b \in R$.*

1. *Wenn $\mathrm{ggT}(a, b)$ und $\mathrm{kgV}(a, b)$ existieren, so sind sie, bis auf Assoziiertheit, eindeutig bestimmt. In diesem Fall sind $\mathrm{ggT}(a, b) \cdot \mathrm{kgV}(a, b)$ und $a \cdot b$ assoziiert.*

2. *Für $d \in R$ ist $d \cdot \mathrm{ggT}(a, b) = \mathrm{ggT}(da, db)$.*

3. *Existiert $\mathrm{ggT}(a, b)$, so existiert auch $\mathrm{kgV}(a, b)$.*

In beliebigen Integritätsbereichen existieren ggT (und somit kgV) nicht. Wir führen (auch im Hinblick auf die spätere Diskussion) folgende zwei Begriffe ein, die in Integritätsbereichen im Allgemeinen nicht äquivalent sind.

**Definition 1.9.** Sei $R$ ein Integritätsbereich.

1. Ein Element $0 \neq p \in R \setminus R^\times$ heißt *Primelement*, wenn gilt:

$$\forall\, a, b \in R \text{ mit } p|ab \Rightarrow p|a \vee p|b\,. \tag{1.3}$$

2. Ein Element $0 \neq a \in R \setminus R^\times$ heißt *unzerlegbar/irreduzibel*, wenn gilt:

$$\text{Ist } a = bc \text{ mit } b, c \in R\,, \text{ dann ist } b \in R^\times \vee c \in R^\times\,. \tag{1.4}$$

**Beispiel 1.10.** Wir betrachten $\mathbb{Z}[\sqrt{-5}]$. Man überzeugt sich, dass $2, 3, 1 \pm \sqrt{-5}$ unzerlegbar sind. Wir betrachten nun die gemeinsamen Teiler von 6 und $3(1 + \sqrt{-5})$. Die Teiler von 6 sind $\{\pm 1, \pm 2, \pm 3, \pm(1 \pm \sqrt{-5})\}$ (da $(1 + \sqrt{-5})(1 - \sqrt{-5}) = 1 + 5 = 6$). Die Teiler von $3(1 + \sqrt{-5})$ sind lediglich $\{\pm 1, \pm 3, \pm(1 \pm \sqrt{-5})\}$. Da jedoch 3 und $1 + \sqrt{-5}$ unzerlegbar sind, teilt weder das eine noch das andere Element jeweils das andere. Somit gibt es keinen größten gemeinsamen Teiler von 6 und $3(1 + \sqrt{5})$ in $\mathbb{Z}[\sqrt{-5}]$.

In der Vorlesung hatten wir bereits gesehen, dass ggT und kgV in euklidschen Ringen schon existieren. Definiere dazu

$$\mathcal{M}(a,b) := \{\alpha a + \beta b \neq 0 : \alpha, \beta \in R\}. \tag{1.5}$$

**Lemma 1.11** (Lemma XI.4)**.** *Sei $R$ ein euklidscher Ring und $a, b \in R \setminus \{0\}$ und $\mathrm{ggT}(a,b)$ so definiert wie in der Vorlesung. Dann ist $\mathcal{M} \neq \emptyset$ und es gilt*

$$\mathrm{ggT}(a,b) = \{g \in \mathcal{M}(a,b) : \rho(g) = \min_{\tilde{g} \in \mathcal{M}(a,b)} \rho(\tilde{g})\}. \tag{1.6}$$

**Lemma 1.12** (Lemma XI.5)**.** *Seien $R$ ein Integritätsbereich und $a, b, c \in R$. Wenn $\mathrm{ggT}(a,b)$ existiert, dann gilt $\mathrm{ggT}(a,b) = \mathrm{ggT}(a+bc, b)$.*

Mit diesen beiden Lemmas kann der euklidsche Algorithmus hergeleitet werden.

**Satz 1.13.** *Seien $R$ ein euklidscher Ring und $a_1, a_2 \in R$ mit $\rho(a_1) > \rho(a_2)$. Dann gibt es $n \in \mathbb{N} \setminus \{1\}$ und $a_3, a_4, ..., a_{n+1} \in R$ und $q_1, ..., q_{n-1} \in R$ mit*

$$\begin{aligned} a_j &= q_j a_{j+1} + a_{j+2}, \quad 1 \leq j \leq n-1 \quad \textit{mit } \rho(a_{j+2}) < \rho(a_{j+1}), \quad 1 \leq j < n-2 \\ a_{n+1} &= 0 \end{aligned} \tag{1.7}$$

*und es gilt*

$$\begin{aligned} a_n &= \mathrm{ggT}(a_1, a_2) = \mathrm{ggT}(q_1 a_2 + a_3, a_2) = \mathrm{ggT}(a_3, a_2) = \mathrm{ggT}(a_3, q_2 a_3 + a_4) = \mathrm{ggT}(a_3, a_4) \\ &= ... = \mathrm{ggT}(a_{n-1}, a_n) = \mathrm{ggT}(q_n a_n + a_{n+1}, a_n) = \mathrm{ggT}(a_{n+1}, a_n) = \mathrm{ggT}(0, a_n). \end{aligned} \tag{1.8}$$

*Da $\rho(a_{j+2}) < \rho(a_{j+1})$ bricht der Algorithmus nach höchstens $\rho(a_2)$ vielen Schritten ab.*

**Beispiel 1.14.**

1. In $\mathbb{Z}$ ist beispielsweise $\mathrm{ggT}(221, 247) = \mathrm{ggT}(221, 247-221) = \mathrm{ggT}(221, 26) = \mathrm{ggT}(221 - 8 \cdot 26, 26) = \mathrm{ggT}(13, 26) = \{\pm 13\}$.

2. In $\mathbb{Q}[X]$ ist

$$\begin{aligned} &\mathrm{ggT}(x^3 - 3x + 2, x^3 + 4x^2 + 5x + 2) = \mathrm{ggT}(x^3 - 3x + 2, 4x^2 + 8x) \\ &\quad = \mathrm{ggT}(-2x^2 - 3x + 2, 4x^2 + 8x) = \mathrm{ggT}(x + 2, 4x^2 + 8x) = (x+2) \cdot \mathcal{Q} \setminus \{0\}. \end{aligned}$$

## 1.3 Fundamentalsatz der Arithmetik

Wörtlich besagt der Fundamentalsatz der Arithmetik

> *Jedes $n \in \mathbb{N}$ lässt sich eindeutig in ein Produkt von Primzahlen, bzw. nach Zusammenfassung gleicher Faktoren, in ein Produkt von Primzahlpotenzen zerlegen.*

Dieser Satz kann nicht nur in $\mathbb{N}$, sondern auch in bestimmten Integritätsbereichen formuliert werden. Zunächst aber kommen wir kurz zurück auf die Begriffe *prim* und *unzerlegbar*.

**Lemma 1.15.** *Sei $R$ ein Integritätsbereich. Dann ist jedes Primelement in $R$ auch unzerlegbar.*

*Beweis.* Hausaufgabe. □

**Beispiel 1.16.** In einem Polynomring $K[X]$ über einem Körper $K$ sind die unzerlegbaren Elemente genau die Polynome, die sich nicht als Produkt von zwei Polynomen *echt* kleineren Grades zerlegen lassen; siehe auch (XI.37).

Interessanterweise gilt die Umkehrung des obigen Lemmas nicht, sprich es gibt unzerlegbare Elemente, die nicht prim sind.

**Beispiel 1.17.** Betrachte das Element 2 im Integritätsbereich $\mathbb{Z}[\sqrt{10}]$. Offenbar ist 2 unzerlegbar und 2 teilt die Zahl $6 = 16 - 10 = (4 + \sqrt{10})(4 - \sqrt{10})$. Jedoch teilt 2 weder $4 + \sqrt{10}$ noch $4 - \sqrt{10}$. Somit ist 2 nicht Prim in $\mathbb{Z}[\sqrt{10}]$.

Wir überspringen den Begriff des *Hauptidealrings* hier und bemerken lediglich, dass euklidsche Ringe Beispiele für Hauptidealringe sind. Für Details, siehe z.B. [MK08, SP08].

**Satz 1.18** (Fundamentalsatz der Arithmetik). *Sei $R$ ein Hauptidealring.*

1. *Die Primelemente in $R$ sind genau die unzerlegbaren Elemente.*
2. *Jedes $0 \neq a \in R \backslash R^\times$ hat (bis auf Reihenfolge und Assoziiertheit) eine eindeutige Zerlegung*
$$a = q_1 ... q_s \tag{1.9}$$
*in ein Produkt nicht notwendig verschiedener Primfaktoren. Fasst man hierin assoziierte Faktoren zusammen, so erhält man eine Zerlegung*
$$a = \varepsilon p_1^{\nu_1} ... p_r^{\nu_r} \tag{1.10}$$
*mit $\varepsilon \in R^\times$, $\nu_1, ..., \nu_r \in \mathbb{N}$ und paarweise nicht zueinander assoziierter Primelemente $p_1, ..., p_r$.*

## 1.4 Unendlichkeit der Primzahlenmenge

**Satz 1.19** (Euklid). *In $\mathbb{N}$ gibt es unendlich viele Primzahlen.*

*Beweis.* Sei $M = \{p_1, ..., p_r\} \subseteq \mathbb{N}$ eine beliebige, endliche Menge von Primzahlen. Dann ist $p_1 ... p_r + 1$ durch keine der $p_1, ..., p_r$ teilbar. Nach dem Fundamentalsatz der Arithmetik muss es daher eine weitere Primzahl $p \notin M$ geben, die $p_1 ... p_r + 1$ teilt. □

**Beispiel 1.20.** $p_1 = 2$, $p_2 = p_1 + 1 = 3$, $p_3 = p_1 p_2 + 1 = 7$, $p_4 = p_1 p_2 p_3 + 1 = 43$, ...

Eine erste Quantifizierung liefert folgender Satz.

**Satz 1.21** (Primzahlsatz von Hadamard und Valée–Poussin). *Für $x \in (0, \infty)$, sei $\pi(x) := \#\{p \leq x : p \text{ Primzahl}\}$. Dann gilt*
$$\lim_{x \to \infty} \frac{\pi(x)}{x / \log(x)} = 1 \, . \tag{1.11}$$

Die Richtigkeit der folgenden Vermutung ist äquivalent zur *Riemannschen Vermutung*, eine der sieben "Millenium Problems" des Clay-Instituts.

**Vermutung 1.22.** *Es gilt*
$$\pi(x) = \mathrm{li}(x) + \mathcal{O}_{x \to \infty}(x^{1/2+\varepsilon}) \tag{1.12}$$
*mit dem Dilogarithmus*
$$\mathrm{li}(x) = \int_2^x \frac{dt}{\log t} = \frac{x}{\log x} + \mathcal{O}_{x \to \infty}\left( \frac{x}{(\log x)^2} \right) . \tag{1.13}$$

Weitere offene Fragen:

1. Gibt es unendlich viele "Primzahlzwillinge", also Paare $(p, p')$ von Primzahlen $p$ und $p'$ mit $|p - p'| = 2$?
2. *Goldbach-Vermutung*: Jede gerade Zahl $> 2$ ist die Summe zweier Primzahlen. (Christian Goldbach an Leonhard Euler im Jahre 1742)

   Eine Abschwächung, die *ternäre Goldbach-Vermutung*, wurde (wohl?) 2013 von Harald Helfgott [Hel13] bewiesen: jede ungerade Zahl $> 5$ ist die Summe dreier Primzahlen. Siehe auch **Wiki**

**Satz 1.23** (Green–Tao [GT08]). *Sei $A$ eine Teilmenge der Primzahlen, die*
$$\limsup_{N \to \infty} \frac{|A \cap [1, N]|}{\pi(N)} > 0 \tag{1.14}$$
*erfüllt. Für beliebiges $k \in \mathbb{N}$ beinhaltet die Menge $A$ dann unendlich viele arithmetische Progressionen der Länge $k$. Insbesondere beinhaltet die Folge der Primzahlen beliebig lange arithmetische Progressionen.*

# 2 2. Tutoriumsblatt (3. Mai 2022)

## 2.1 Minimalpolynom und Satz von Cayley–Hamilton

Wir starten mit dem Minimalpolynom und dem Satz von Cayley–Hamilton aus der Vorlesung.

# X. Das Minimalpolynom und der Satz von Cayley-Hamilton

In diesem Kapitel betrachten wir einen Vektorraum $X$ endlicher Dimension $N = \dim(X) < \infty$ über einem Körper $\mathbb{F}$ und einen Endomorphismus $\Phi \in \mathcal{L}(X)$. Wir untersuchen Polynome $f \in \mathbb{F}[x]$ in $\Phi$, d.h. für $f = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \ldots + \alpha_K x^K$ definieren wir

$$f(\Phi) \ := \ \alpha_0 + \alpha_1 \Phi + \alpha_2 \Phi^2 + \ldots + \alpha_K \Phi^K, \tag{X.1}$$

wobei $\Phi^k = \Phi \circ \Phi \circ \cdots \circ \Phi$ die $k$-fache Iteration der Abbildung $\Phi$ notiert.

## X.1. Das Minimalpolynom

Als erstes stellen wir fest, dass es Polynome $f \in \mathbb{F}[x] \setminus \{0\}$ gibt, sodass $f(\Phi) = 0$ die Nullabbildung ist. Der Vektorraum $\mathcal{L}(X)$ der Endomorphismen auf $X$ hat nämlich die Dimension

$$\dim \big[\mathcal{L}(X)\big] \ := \ \big[\dim(X)\big]^2 \ = \ N^2. \tag{X.2}$$

Daher ist die $(N^2+1)$-elementige Menge

$$\big\{ \mathbb{1}, \ \Phi, \ \Phi^2, \ldots, \ \Phi^{(N^2)} \big\} \ \subseteq \ \mathcal{L}(X) \tag{X.3}$$

linear abhängig, und es gibt eine nichttriviale Linearkombination mit

$$f(\Phi) \ = \ \alpha_0 + \alpha_1 \Phi + \alpha_2 \Phi^2 + \ldots + \alpha_{N^2} \Phi^{(N^2)} \ = \ 0. \tag{X.4}$$

**Definition X.1.** Seien $X$ ein $\mathbb{F}$-Vektorraum der Dimension $\dim(X) = N \in \mathbb{N}$, $\Phi \in \mathcal{L}(X) \setminus \{0\}$ ein Endomorphismus auf $X$ und

$$\mathcal{F} \ := \ \big\{ f \in \mathbb{F}[x] \setminus \{0\} \ \big| \ f(\Phi) = 0 \big\}. \tag{X.5}$$

Ist $\mu \in \mathcal{F}$ von kleinstem Grad und normiert,

$$\forall f \in \mathcal{F}: \quad \deg(f) \geq \deg(\mu) \ , \qquad \mu \ = \ \sum_{m=0}^{M-1} \beta_m x^m + x^M, \tag{X.6}$$

so heißt $\mu$ **Minimalpolynom (von $\Phi$)**.

**Bemerkungen und Beispiele.**

- Wegen (X.4) ist $\mathcal{F} \neq \emptyset$ und $\deg(\mu) \leq N^2$.

**Lemma X.2.** *Sind $\mu \in \mathcal{F}$ ein Minimalpolynom von $\Phi \in \mathcal{L}(X) \setminus \{0\}$ und $f \in \mathcal{F}$, so wird $f$ von $\mu$ geteilt. Insbesondere ist das Minimalpolynom von $\Phi$ eindeutig.*

*Beweis.* Seien $\mu, f \in \mathcal{F}$ und $\mu$ ein Minimalpolynom von $\Phi$. Weil $\mathbb{F}[x]$ euklidsch ist, gibt es $q, r \in \mathbb{F}[x]$ mit $r = 0$ oder $\deg(r) < \deg(\mu)$ so, dass $f = q \cdot \mu + r$. Damit ist

$$r(\Phi) \ = \ f(\Phi) - q(\Phi) \cdot \mu(\Phi) \ = \ 0. \tag{X.7}$$

Wäre $r \neq 0$, so wäre $r \in \mathcal{F}$, was wegen der Minimalität des Grads $\deg(\mu) > \deg(r)$ unmöglich ist. Also ist $r = 0$ bzw. $f = q \cdot \mu$.
Die Eindeutigkeit des Minimalpolynoms folgt nun leicht aus $\mu | \tilde{\mu}$, für alle Minimalpolynome $\mu, \tilde{\mu} \in \mathcal{F}$. □

## X.2. Der Satz von Cayley-Hamilton

Der nun folgende Satz von Cayley-Hamilton behauptet, dass das charakteristische Polynom von $\Phi$ ein Element von $\mathcal{F}$ ist. Zu seiner Formulierung erinnern wir an die Definition des charakteristischen Polynoms

$$\chi(\lambda) \ = \ \det\big[A - \lambda \cdot \mathbb{1}\big] \tag{X.8}$$

eines Endomorphismus $\Phi \in \mathcal{L}(X)$, wobei $A$ (irgend-)eine Matrixdarstellung von $\Phi$ ist.

**Satz X.3** (Cayley-Hamilton). *Sind $X$ ein $\mathbb{F}$-Vektorraum der Dimension $\dim(X) = N \in \mathbb{N}$, $\Phi \in \mathcal{L}(X)$ ein Endomorphismus auf $X$ und $\chi \in \mathbb{F}[x]$ das charakteristische Polynom, so ist $\chi \in \mathcal{F}$, d.h.*

$$\chi(\Phi) \ = \ 0. \tag{X.9}$$

*Beweis.* Es reicht, $\chi[A] = 0$ für eine Matrixdarstellung $A \in \mathfrak{M}_{N \times N}(\mathbb{F})$ von $\Phi \in \mathcal{L}(X)$ zu zeigen. Nach Satz VI.7 ist

$$G \cdot G_{\mathrm{minor}}^T \ = \ \det[G] \cdot \mathbb{1}, \tag{X.10}$$

für jede $N \times N$-Matrix $G \in \mathfrak{M}_{N \times N}(\mathbb{F})$, wobei $G_{\mathrm{minor}}$ die zugehörige Minorenmatrix ist. Deren Matrixelemente sind gerade die Determinanten der $(N-1) \times (N-1)$-Matrizen, die durch Streichung einer Spalte und einer Zeile aus $G$ hervorgehen; genauer gilt mit $G = \big(g_{i,j}\big)_{i,j=1}^N$, dass

$$\big(G_{\mathrm{minor}}\big)_{i,j} \ = \ (-1)^{i+j} \cdot \det\left[\begin{pmatrix} g_{1,1} & \cdots & g_{1,j-1} & g_{1,j+1} & \cdots & g_{1,N} \\ \vdots & & \vdots & \vdots & & \vdots \\ g_{i-1,1} & \cdots & g_{i-1,j-1} & g_{i-1,j+1} & \cdots & g_{i-1,N} \\ g_{i+1,1} & \cdots & g_{i+1,j-1} & g_{i+1,j+1} & \cdots & g_{i+1,N} \\ \vdots & & \vdots & \vdots & & \vdots \\ g_{N,1} & \cdots & g_{N,j-1} & g_{N,j+1} & \cdots & g_{N,N} \end{pmatrix}\right]. \tag{X.11}$$

Dabei wird ein Diagonalelement aus $G$ gestrichen, falls $i = j$, und für $i \neq j$ werden sogar zwei Diagonalelemente aus $G$ gestrichen. Jedenfalls wird mindestens ein Diagonalelement aus $G$ gestrichen, und deshalb ist

$$(A - \lambda \cdot \mathbb{1})^T_{\text{minor}} \;=\; B_0 + B_1\lambda + B_2\lambda^2 + \ldots + B_{N-1}\lambda^{N-1}, \tag{X.12}$$

für geeignet gewählte Matrizen $B_0, B_1, \ldots, B_{N-1} \in \mathfrak{M}_{N\times N}(\mathbb{F})$. Nach (X.10) gilt mit $\chi(\lambda) = \alpha_0 + \alpha_1\lambda + \ldots + \alpha_N\lambda^N$ also, dass

$$\begin{aligned} \alpha_0 \cdot \mathbb{1} \,+\, \alpha_1\lambda \cdot \mathbb{1} \,+ \ldots + \alpha_N\lambda^N \cdot \mathbb{1} \;&=\; \chi(\lambda)\cdot \mathbb{1} \;=\; \det\big[A - \lambda\cdot\mathbb{1}\big]\cdot \mathbb{1} \\ &= \big(A - \lambda\cdot\mathbb{1}\big)\cdot\big(B_0 + B_1\lambda + B_2\lambda^2 + \ldots + B_{N-1}\lambda^{N-1}\big) \\ &= AB_0 + \lambda(AB_1 - B_0) + \lambda^2(AB_2 - B_1) + \ldots + \lambda^{N-1}(AB_{N-1} - B_{N-2}) - \lambda^N B_{N-1}. \end{aligned} \tag{X.13}$$

Durch Koeffizientenvergleich gewinnen wir daraus, dass

$$\begin{aligned} && \alpha_0\cdot\mathbb{1} &= A\,B_0, && \text{(X.14)}\\ \alpha_1\cdot\mathbb{1} = A\,B_1 - B_0 &\;\Rightarrow& \alpha_1\cdot A &= A^2\,B_1 - A\,B_0, && \text{(X.15)}\\ \alpha_2\cdot\mathbb{1} = A\,B_2 - B_1 &\;\Rightarrow& \alpha_2\cdot A^2 &= A^3\,B_2 - A^2\,B_1, && \text{(X.16)}\\ && &\vdots && \\ \alpha_{N-1}\cdot\mathbb{1} = A\,B_{N-1} - B_{N-2} &\;\Rightarrow& \alpha_{N-1}\cdot A^{N-1} &= A^N\,B_{N-1} - A^{N-1}\,B_{N-2}, && \text{(X.17)}\\ \alpha_N\cdot\mathbb{1} = -B_{N-1} &\;\Rightarrow& \alpha_N\cdot A^N &= -\,A^N\,B_{N-1}. && \text{(X.18)} \end{aligned}$$

Die Behauptung erhalten wir nun durch Addition der rechten Gleichungen in (X.14)–(X.18),

$$\begin{aligned} \chi(A) \;&= \alpha_0\cdot\mathbb{1} \,+\, \alpha_1 A \,+ \ldots + \alpha_N A^N \\ &= AB_0 + \big(A^2B_1 - AB_0\big) + \big(A^3B_2 - A^2B_1\big) + \\ &\qquad \ldots + \big(A^N B_{N-1} - A^{N-1}B_{N-2}\big) - A^N B_{N-1} \\ &= 0. \end{aligned} \tag{X.19}$$

□

**Korollar X.4.** *Ist $\mu \in \mathbb{F}[x]$ das Minimalpolynom eines Endomorphismus' $\Phi \in \mathcal{L}(X) \setminus \{0\}$ eines $\mathbb{F}$-Vektorraums $X$ der Dimension $\dim(X) = N \in \mathbb{N}$, so teilt $\mu$ das charakteristische Polynom $\chi \in \mathbb{F}[x]$ von $\Phi$ und es gilt*

$$\deg(\mu) \;\leq\; N. \tag{X.20}$$

**Bemerkungen und Beispiele.**

- Sind die Eigenwerte $\lambda_1, \ldots, \lambda_N \in \mathbb{F}$ von $\Phi$ -also die Nullstellen des charakteristischen Polynoms- paarweise verschieden, d.h. gilt $\lambda_i \neq \lambda_j$ falls $i \neq j$, so ist

$$\chi(\lambda) \;=\; (\lambda_1 - \lambda)(\lambda_2 - \lambda)\cdots(\lambda_N - \lambda). \tag{X.21}$$

Außerdem ist $\Phi$ diagonalisierbar, denn die zugehörigen Eigenvektoren $\vec{v}_1, \ldots, \vec{v}_N \in X \setminus \{\vec{0}\}$, $\Phi\vec{v}_j = \lambda_j \vec{v}_j$, bilden eine Basis.

Seien nun $n \in \mathbb{Z}_1^N$ und $f \in \mathcal{F}$, also $f(\Phi) = 0$ und $\deg(f) \geq 1$ (letzteres wegen $\Phi \neq 0$). Dann gibt es $g \in \mathbb{F}[x]$ und $\alpha \in \mathbb{F}$ so, dass

$$f(\lambda) \;=\; g(\lambda) \cdot (\lambda_n - \lambda) + \alpha. \tag{X.22}$$

Damit ist

$$\alpha\, \vec{v}_n \;=\; f(\Phi)\vec{v}_n - g(\Phi) \cdot (\lambda_n - \Phi)\vec{v}_n \;=\; \vec{0} \tag{X.23}$$

und wegen $\vec{v}_n \neq \vec{0}$ ist $\alpha = 0$, d.h. $\lambda_n - \lambda$ teilt $f$.

Da dies für jedes $n \in \mathbb{Z}_1^N$ gilt, folgt, dass $\chi$ jedes Polynom in $\mathcal{F}$ teilt – auch das Minimalpolynom. Deshalb ist in diesem Fall

$$\mu(\lambda) \;=\; (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_N) \;=\; (-1)^N \cdot \chi(\lambda). \tag{X.24}$$

- Es kann aber auch das umgekehrte Extrem auftreten: Ist $\Phi\vec{x} = \vec{x}$ die Identitätsabbildung, so gelten offensichtlich

$$\chi(\lambda) \;=\; (1 - \lambda)^N \,, \qquad \mu(\lambda) \;=\; 1 - \lambda. \tag{X.25}$$

- Die konkrete Berechnung des Minimalpolynoms behandeln wir in den Übungen.

## X.3. Invariante Unterräume und irreduzible Polynome

**Definition X.5.** Sei $\Phi \in \mathcal{L}(X)$ ein Endomorphismus eines $\mathbb{F}$-Vektorraums $X$. Ein Unterraum $U \subseteq X$ heißt **invariant unter $\Phi$** oder einfach **$\Phi$-invariant**

$$:\Leftrightarrow \qquad \Phi(U) \;\subseteq\; U. \tag{X.26}$$

**Bemerkungen und Beispiele.**

- Ist $\vec{v} \in X \setminus \{\vec{0}\}$ ein Eigenvektor von $\Phi$ zum Eigenwert $\lambda \in \mathbb{F}$, $\Phi\vec{v} = \lambda\vec{v}$, so ist $\mathbb{F} \cdot \vec{v} \subseteq X$ offensichtlich $\Phi$-invariant.

- Allgemeiner ist für $f \in \mathbb{F}[x]$ auch $\mathrm{Ker}[f(\Phi)] \subseteq X$ $\Phi$-invariant, denn für $f, g \in \mathbb{F}[x]$ kommutieren die Endomorphismen $f(\Phi)$ und $g(\Phi) \in \mathcal{L}(X)$, da mit $f = \alpha_0 + \alpha_1 x + \ldots + \alpha_M x^M$ und $g = \beta_0 + \beta_1 x + \ldots + \beta_L x^L$

$$f(\Phi) \cdot g(\Phi) \;=\; \sum_{k=1}^{K} \sum_{\ell=1}^{L} \alpha_k \beta_\ell \, \Phi^{k+\ell} \;=\; g(\Phi) \cdot f(\Phi). \tag{X.27}$$

Für $\vec{x} \in \mathrm{Ker}[f(\Phi)]$ ist also $f(\Phi)(\Phi\vec{x}) = \Phi[f(\Phi)\vec{x}] = 0$, d.h. auch $\Phi\vec{x} \in \mathrm{Ker}[f(\Phi)]$.

### 2.1.1 Anwendung: Analytische Funktionen von Matrizen

Seien $A \in \mathbb{C}^{n\times n}$ und $f : \mathbb{C} \to \mathbb{C}$ eine analytische Funktion mit Laurent-Reihe

$$f(x) = \sum_{k\geq 0} a_k x^k \,. \tag{2.1}$$

Sei $q \in \mathbb{C}[X]$ ein Quotientenpolynom und $r \in \mathbb{C}[X]$ das Restpolynom in der Polynomdivision

$$f(x) = q(x)\chi_A(x) + r(x)\,, \tag{2.2}$$

wobei $\deg(r) < \deg(\chi_A) = n$. Mit Cayley–Hamilton können wir $f(A)$ vereinfachen und erhalten

$$f(A) = q(A)\chi_A(A) + r(A) = r(A)\,. \tag{2.3}$$

Um diese Konsequenz auszunutzen, bestimmen wir nun die Koeffizienten von $r(x)$. Angenommen, das Restpolynom sei von der Gestalt

$$r(x) = \sum_{k=0}^{n-1} c_k x^k \,. \tag{2.4}$$

Wenn $\{\lambda_j\}_{j=1}^n \subseteq \mathbb{C}$ die Eigenwerte von $A$ bezeichnen, so folgt aus $\chi_A(\lambda_j) = 0$, dass

$$f(\lambda_j) = r(\lambda_j) = \sum_{k=0}^{n-1} c_k \lambda_j^k\,, \quad j \in \{1, 2, ..., n\}\,. \tag{2.5}$$

Wenn alle Eigenwerte einfach sind, erhält man $n$ Gleichungen, wodurch die Koeffizienten $c_0, ..., c_{n-1}$ eindeutig bestimmt werden können. Sollte der Eigenwert $\lambda_j$ algebraische Vielfachheit $m > 1$ haben, so kann obiges Gleichungssystem nicht mehr eindeutig gelöst werden, da es $m$ gleiche (also insbesondere linear abhängige) Gleichungen gibt. Jedoch verschwinden gleichzeitig die ersten $m - 1$ Ableitungen von $\chi_A(x)$ bei $x = \lambda_j$. Dies führt auf $m - 1$ weitere linear unabhängige Gleichungen

$$\frac{\mathrm{d}^\ell f(x)}{\mathrm{d}x^\ell}\Big|_{x=\lambda_j} = \frac{\mathrm{d}^\ell r(x)}{\mathrm{d}x^\ell}\Big|_{x=\lambda_j}\,, \quad j = 1, 2, ..., m-1\,. \tag{2.6}$$

So erhält man wieder $n$ Gleichungen für die Koeffizienten $c_0, ..., c_{n-1}$ des Restpolynoms $r(x)$.

**Beispiel 2.1.** Sei $A = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$. Unser Ziel ist es, $f(A) = \mathrm{e}^{tA}$ zu finden. Wir haben $\chi_A(x) = (x-1)(x-3)$, sprich die Eigenwerte sind $\lambda_1 = 1$ und $\lambda_2 = 3$. Sei $r(x) = c_0 + c_1 x$. Mit obigem Verfahren sind $c_0, c_1$ durch die Gleichungen

$$\mathrm{e}^{t\lambda_1} = \mathrm{e}^t = c_0 + c_1 \cdot \lambda_1 = c_0 + c_1$$

und

$$\mathrm{e}^{t\lambda_2} = \mathrm{e}^{3t} = c_0 + c_1 \cdot \lambda_2 = c_0 + 3c_1$$

bestimmt. Daraus folgt $c_0 = (3\mathrm{e}^t - \mathrm{e}^{3t})/2$ und $c_1 = (\mathrm{e}^{3t} - \mathrm{e}^t)/2$. Somit ist

$$\mathrm{e}^{tA} = c_0 \mathbf{1}_2 + c_1 A = \begin{pmatrix} \mathrm{e}^t & \mathrm{e}^{3t} - \mathrm{e}^t \\ 0 & \mathrm{e}^{3t} \end{pmatrix}\,. \tag{2.7}$$

Ähnlich gelangt man zu Polynomdarstellungen von $\sin(tA)$, was als Übung überlassen ist. (Wir erinnern uns an $\sin(tA) = (\mathrm{e}^{itA} - \mathrm{e}^{-itA})/(2i)$.)

**Beispiel 2.2.** Letztes Semester hatten wir kurz die Erzeuger von Rotationsmatrizen betrachtet. Sei nun $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Dann findet man mit obigem Verfahren in der Tat

$$\mathrm{e}^{tJ} = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix} = \cos(t) \cdot \mathbf{1}_2 + \sin(t) \cdot J\,. \tag{2.8}$$

### 2.1.2 Weitere Resultate

Wir notieren nun noch einige weitere Resultate.

**Satz 2.3.** *Ähnliche Matrizen in $K^{n\times n}$ haben dasselbe Minimalpolynom.*

*Beweis.* Seien $A, B \in K^{n\times n}$ und $S \in \mathrm{GL}(n, K)$ so, dass $B = S^{-1}AS$. Seien außerdem $a_0, a_1, ..., a_{m-1} \in K$ so, dass $A^m + a_{m-1}A^{m-1} + ... + a_1 A + a_0 = 0$. (Dies trifft insbesondere für das Minimalpolynom zu.) Dann gilt

$$\begin{aligned} B^m + a_{m-1}B^{m-1} + ... + a_1 B + a_0 &= (S^{-1}AS)^m + a_{m-1}(S^{-1}AS)^{m-1} + ... a_1 S^{-1}AS + a_0 \\ &= S^{-1}\left(A^m + a_{m-1}A^{m-1} + ... + a_1 A + a_0\right) S = 0\,, \end{aligned}$$

was die Behauptung zeigt. □

**Satz 2.4.** *Seien $n \in \mathbb{N}$ und $V$ ein $n$-dimensionaler $K$-Vektorraum und $f \in \mathrm{End}(V)$ ein Endomorphismus mit Minimalpolynom $\mu_f$. Für $r \in K$ gilt*

$$r \textit{ ist ein Eigenwert von } f \Leftrightarrow \mu_f(r) = 0 \tag{2.9}$$

*Beweis.* $\Rightarrow$: Sei $0 \neq v \in V$ der zu $r$ gehörende Eigenvektor, sprich $f(v) = rv$ und sei

$$\mu_f = X^m + c_{m-1}X^{m-1} + ... + c_1 X + c_0\,.$$

Dann gilt

$$0 = \mu_f(f)v = f(v)^m + c_{m-1}f(v)^{m-1} + ... + c_1 f(v) + c_0 v = \mu_f(r)v\,.$$

Da $v \neq 0$, muss $\mu_f(r) = 0$ sein, wie behauptet.

$\Leftarrow$: Dies folgt aus der Tatsache, dass $\mu_f$ das charakteristische Polynom $\chi_f(\lambda) = \prod_{j=1}^n(\lambda_j - \lambda)$ teilt. (Hierbei sind $\lambda_j$ die Eigenwerte von $f$ gezählt gemäß algebraischer Vielfachheit.) □

**Satz 2.5.** *Seien $n \in \mathbb{N}$ und $V$ ein $n$-dimensionaler $K$-Vektorraum und $f \in \mathrm{End}(V)$ ein Endomorphismus mit Minimalpolynom $\mu_f$. Dann ist $f$ genau dann diagonalisierbar, wenn sich $\mu_f$ als $\mu_f = (X - a_1)...(X - a_k)$ mit paarweise verschiedenen $a_1, ..., a_k \in K$ schreiben lässt.*

*Beweis.* $\Rightarrow$: Angenommen $f$ sei diagonalisierbar. Dann gibt es eine Basis in $V$, sodass $f$ die Matrixdarstellung $A = \mathrm{diag}(a_1, ..., a_1, a_2, ..., a_2, ..., a_k, ..., a_k)$ hat, wobei die $(a_j)_{j=1}^k$ jeweils gemäß geometrischer Vielfachheit gezählt werden. Diese sind nach Satz 2.4 Nullstellen des Minimalpolynoms $\mu_f$, sprich $\mu_f = (X - a_1)...(X - a_k) \cdot \varphi$ für ein $\varphi \in K[X]$. Die Behauptung folgt aus $(A - a_1\mathbf{1}_n)(A - a_2\mathbf{1}_n)...(A - a_k\mathbf{1}_n) = 0$.

$\Leftarrow$: Angenommen $\mu = (X - a_1)...(X - a_k)$ mit paarweise verschiedenen $a_1, ..., a_k \in K$. Für $i = 1, ..., k$ hat das Polynom

$$\varphi_i := \prod_{j\neq i} \frac{X - a_j}{a_i - a_j}$$

den Grad $k - 1$ und es gilt

$$\varphi_i(a_t) = \begin{cases} 0 & \text{für } t \neq i\,, \\ 1 & \text{für } t = i\,. \end{cases}$$

Daher hat das Polynom $\varphi := \varphi_1 + ... + \varphi_k$ höchstens Grad $k-1$ und es gilt $\varphi(a_1) = ... = \varphi(a_k) = 1$. Da Polynome $n$-ten Grades höchstens $n$ Nullstellen haben können, muss also $\varphi \equiv 1$ ein konstantes Polynom sein. Somit ist

$$\mathrm{id}_V = \sum_{i=1}^k \prod_{j\neq i} \frac{f - a_j \mathrm{id}_V}{a_i - a_j}\,.$$

Für $v \in V$ gilt also

$$v = \sum_{i=1}^k \underbrace{\prod_{j\neq i}\left(\frac{f - a_j\mathrm{id}_V}{a_i - a_j}\right)(v)}_{=:v_i}\,.$$

Dabei ist jeweils

$$(f - a_i \mathrm{id}_V)(v_i) = \left( \prod_{j \neq i} \frac{1}{a_i - a_j} \right) \underbrace{(f - a_1 \mathrm{id}_V) ... (f - a_k \mathrm{id}_V)}_{=0} = 0\,,$$

sprich $v_i \in \ker(f - a_i \mathrm{id}_V)$ für alle $i = 1, ..., k$. Das bedeutet, dass $V$ die Summe der Eigenräume zu den Eigenwerten $a_1, ..., a_k$ von $f$ ist. Folglich ist $f$ diagonalisierbar. □

Wir erinnern daran, dass ein Körper $K$ algebraisch abgeschlossen heißt, wenn jedes nichtkonstante Polynom mit Koeffizienten in $K$ eine Nullstelle hat. Aus dem Fundamentalsatz der Algebra folgt, dass $\mathbb{C}$ algebraisch abgeschlossen ist. Dagegen ist $\mathbb{R}$ nicht algebraisch abgeschlossen: beispielsweise hat die Gleichung $x^2 + 1 = 0$ keine reelle Lösung.

**Satz 2.6.** *Seien $K$ algebraisch abgeschlossen, $n \in \{2, 3\}$, $A, B \in K^{n \times n}$. Dann sind $A$ und $B$ genau dann ähnlich zueinander, wenn sie dasselbe charakteristische Polynom und Minimalpolynom haben.*

*Beweis.* $\Rightarrow$: Folgt aus Satz 2.3 (gleiches Minimalpolynom) und der Tatsache, dass $\det(A - \lambda) = \det(S^{-1}(A - \lambda)S)$ ist (gleiches charakteristisches Polynom).

$\Leftarrow$: Wir unterscheiden zwischen $n = 2$ und $n = 3$.

1. $\underline{n = 2}$: Die Matrix $A$ habe das charakteristische Polynom $\chi_A = (X - a)(X - b)$ und das Minimalpolynom $\mu_A$. Falls $a \neq b$, so ist $A$ ähnlich zu $\mathrm{diag}(a, b)$. Falls $a = b$, so ist $\chi_A = (X - a)^2$. Falls $\mu_A = (X - a)^2$, so ist $A$ ähnlich zu $\begin{pmatrix} a & 0 \\ 1 & a \end{pmatrix}$. Falls $\mu_A = (X - a)$, so ist $A$ ähnlich zu $\mathrm{diag}(a, a)$.

2. $\underline{n = 3}$: Die Matrix $A$ habe das charakteristische Polynom $\chi_A = (X - a)(X - b)(X - c)$ und das Minimalpolynom $\mu_A$. Im Fall $|\{a, b, c\}| = 3$ ist $A$ ähnlich zu $\mathrm{diag}(a, b, c)$. Sei nun $\chi_A = (X - a)^2(X - b)$ mit $a \neq b$. Im Fall $\mu_A = (X - a)^2(X - b)$ ist $A$ ähnlich zu $\begin{pmatrix} a & 0 & 0 \\ 1 & a & 0 \\ 0 & 0 & b \end{pmatrix}$. Im Fall $\mu_A = (X - a)(X - b)$ ist $A$ ähnlich zu $\begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{pmatrix}$. Angenommen schließlich $\chi_A = (X - a)^3$. Im Fall $\mu_A = (X - a)^3$ ist $A$ ähnlich zu $\begin{pmatrix} a & 0 & 0 \\ 1 & a & 0 \\ 0 & 1 & a \end{pmatrix}$. Im Fall $\mu_A = (X - a)^2$ ist $A$ ähnlich zu $\begin{pmatrix} a & 0 & 0 \\ 1 & a & 0 \\ 0 & 0 & a \end{pmatrix}$. Im Fall $\mu_A = (X - a)$ ist $A$ ähnlich zu $\begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{pmatrix}$.

□

## 2.2 Aufgaben

**Aufgabe 2.1.** Seien $A_1, A_2, A_3 \in \mathbb{R}^{3 \times 3}$ gegeben durch

$$A_1 = \begin{pmatrix} -2 & 1 & 1 \\ -6 & 3 & 3 \\ 2 & -1 & -1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 2 & 4 & -2 \\ 0 & -4 & 3 \\ -3 & -7 & 4 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

Ermitteln Sie die Minimalpolynome der Matrizen $A_1, A_2, A_3$.

**Aufgabe 2.2.** Geben Sie jeweils eine $\mathbb{R}^{3 \times 3}$-Matrix an, welche die angegebenen charakteristischen Polynome und Minimalpolynome besitzt.

1. $\chi_A = -X^3$ und $\mu_A \in \{X, X^2, X^3\}$.
2. $\chi_B = (X-1)^2(X-3)$ und $\mu_B \in \{(X-1)(X-3), (X-1)^2(X-3)\}$.

**Aufgabe 2.3.** Beweise oder widerlege: Je zwei $n \times n$-Matrizen über $K$ mit demselben Minimalpolynom sind ähnlich.

Die folgende Aufgabe eignet sich für später, wenn wir die Jordansche Normalform besprechen.

**Aufgabe 2.4.** Beweise oder widerlege:

1. Je zwei $n \times n$-Matrizen über $K$ mit demselben Minimalpolynom und demselben charakteristischen Polynom sind ähnlich.
2. Je zwei $n \times n$-Matrizen über $K$ mit demselben demselben charakteristischen Polynom, welches außerdem keine mehrfachen irreduziblen Faktoren hat, sind ähnlich. (Zur Erinnerung: Ein Polynom $f \in K[X]$ heißt irreduzibel genau dann, wenn alle Teiler von $f$ entweder Grad Null oder $\deg(f)$ haben.)

## 2.3 Lösungen

**Lösung 2.1.** Die charakteristischen Polynome der Matrizen sind $\chi_{A_1} = -X^3$, $\chi_{A_2} = -(X-2)(X-1)(X+1)$ und $\chi_{A_3} = -(X-1)(X-2)^2$. Offenbar ist schon $\mu_{A_2} = \chi_{A_2}$ (siehe Vorlesung nach Korollar XII.4). Um $\mu_{A_1}$ zu bestimmen, beobachten wir, dass zwar $A_1 \neq 0$, aber $A_1^2 = 0$ ist. Somit ist $\mu_{A_1} = X^2$. Schließlich bestimmen wir $\mu_{A_3} \in \{(X-1)(X-2), (X-1)(X-2)^2\}$. Man findet $(A_3 - 1)(A_3 - 2) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$, womit $\mu_{A_3} = (X-1)(X-2)^2$ sein muss.

**Lösung 2.2.**

1. Ansatz: Mache die zu Null gehörenden Eigenräume immer kleiner durch Vergrößern des Ranges.

   $A = 0$ hat $\chi_A = -X^3$ und $\mu_A = X$.

   Die Matrix $A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ hat wieder $\chi_A = -X^3$, aber $\mu_A = X^2$.

   Die Matrix $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ hat wieder $\chi_A = -X^3$, aber $\mu_A = X^3$.

2. Nach Satz 2.5 ist $B$ im Falle $\mu_B = (X-1)(X-3)$ diagonalisierbar. Wir können also bspw. $B = \mathrm{diag}(1,1,3)$ wählen. Durch Verkleinern des zu 1 gehörenden Eigenraums, wie $B = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ erreichen wir $\mu_B = (X-1)^2(X-3) = \chi_B$.

**Lösung 2.3.** Die komplexen Matrizen

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

besitzen beide das Minimalpolynom $(X-1)(X-2)$. Da aber $\det(A) = 2 \neq \det(B) = 4$, können sie nicht ähnlich zueinander sein.

**Lösung 2.4.**

1. Die komplexen Matrizen

$$A = \begin{pmatrix} 1 & 1 & & & \\ & 1 & 1 & & \\ & & 1 & & \\ & & & 1 & \\ & & & & 1 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 1 & 1 & & & \\ & 1 & 1 & & \\ & & 1 & & \\ & & & 1 & 1 \\ & & & & 1 \end{pmatrix}$$

haben beide charakteristisches Polynom $X^5$ und Minimalpolynom $X^3$. Da sie aber unterschiedliche Jordan-Normalformen haben, können sie nicht ähnlich zueinander sein.

2. Sei $A$ eine Matrix mit einem charakteristischen Polynom, welches keine mehrfachen irreduziblen Faktoren besitzt. In der Jordanschen Normalform von $A$ ist der Jordanblock zu jedem irreduziblen Faktor $p(X)$ von $\chi_A(X)$ gleich der Begleitmatrix von $p(X)$. Somit ist die Jordansche Normalform von $A$ durch $\chi_A(X)$ eindeutig bestimmt. Insbesondere haben also zwei Matrizen von dieser Form dieselbe Jordansche Normalform, sind also ähnlich. Also gilt die Aussage.

## 3 3. Tutoriumsblatt (12. Mai 2022)

### 3.1 Aufgaben

**Aufgabe 3.1.** Seien $A_1, A_2, A_3 \in \mathbb{R}^{3\times 3}$ gegeben durch

$$A_1 = \begin{pmatrix} 2 & 4 & -2 \\ 0 & -4 & 3 \\ -3 & -7 & 4 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

(a) Ermitteln Sie die Minimalpolynome der Matrizen $A_1, A_2, A_3$.

Hinweise zur Bestimmung von $\mu_{A_3}$:

(a) Erinnern Sie sich an die Determinantenformel

$$\det{}_{K^{2n\times 2n}} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det{}_{K^n}(AD - CB)$$

für $K^{2n\times 2n}$ Blockmatrizen mit $A, B, C, D \in K^{n\times n}$. Diese wurde bspw. in Aufgabe 9.2.(b) des Tutoriumblatts vom WS 21/22 bewiesen.

(b) Erinnern Sie sich an Satz XII.9 und Korollar XI.12 der Vorlesung: Das Minimalpolynom eines Endomorphismus über einem $n$-dimensionalen $K$-Vektorraum zerfällt bekanntermaßen eindeutig in ein Produkt von normierten, irreduziblen und paarweise teilerfremden Polynome

$$\mu(x) = \mu_1^{s_1}(x) \cdot \mu_2^{s_2}(x) \cdot \ldots \cdot \mu_K^{s_K}(x)$$

mit $s_1 + \ldots + s_K \leq n$.

(b) Bestimmen Sie, ob $A_1, A_2, A_3$ diagonalisierbar sind.

(c) Bestimmen Sie, ob $A_1, A_2, A_3$ invertierbar sind.

(d) In der Hausaufgabe werden Sie zeigen, dass für $A \in \mathrm{GL}(n, K)$ gilt, dass

$$A^{-1} = -\frac{1}{a_0}\left(A^{m-1} + a_{m-1}A^{m-2} + \ldots + a_1\right).$$

Hierbei sind $a_0, \ldots, a_{m-1}$ die Koeffizienten des Minimalpolynoms $\mu_A(x) = x^m + a_{m-1}x^{m-1} + \ldots + a_1 x + a_0$ mit $m \leq n$.

Sofern $A_j$ invertierbar ist, bestimmen Sie damit die Matrizen $A_j^{-1}$.

**Aufgabe 3.2.** Sei $K$ ein Körper und definiere $K[X]_{\leq n} := \{f \in K[X] : \deg(f) \leq n\}$. Zusammen mit der üblichen Addition und skalaren Multiplikation ist $K[X]_{\leq n}$ ein $K$-Vektorraum. Dieser Vektorraum kann bspw. mit der Monombasis $B := \{1, X, X^2, \ldots, X^n\}$ ausgestattet werden.

Für $K = \mathbb{Q}$ betrachten wir die Abbildungen

$$\varphi_1 : \mathbb{Q}[X]_{\leq 2} \to \mathbb{Q}[X]_{\leq 2}, \qquad \sum_{j=0}^{2} a_j x^j \mapsto \sum_{j=1}^{2} j a_j x^{j-1} \quad \text{(formale Ableitung)}$$

$$\varphi_2 : \mathbb{Q}[X]_{\leq 2} \to \mathbb{Q}[X]_{\leq 2}, \qquad f \mapsto f(0) + [f(1) - f(-1)]x + [f(1) + f(-1)]x^2.$$

(a) Bestimmen Sie die Matrixdarstellungen der $\varphi_j$ bzgl. der Monombasis $B$.

(b) Berechnen Sie die charakteristischen Polynome $\chi_{\varphi_j}$.

(c) Berechnen Sie die Minimalpolynome $\mu_{\varphi_j}$.

(d) Entscheiden Sie auf Basis von (c), ob die $\varphi_j$ diagonalisierbar sind.

Bonus: Geben Sie eine Basis $B'$ von zugehörigen Eigenvektoren an, in der die Matrixdarstellung von $\varphi_j$ diagonal ist.

(e) Entscheiden Sie auf Basis von (b), ob $\varphi_j$ invertierbar sind. Berechnen Sie wie in (d) von Aufgabe 3.1 die Inverse $\varphi_j^{-1}$, sofern sie existiert.

## 3.2 Lösungen

**Lösung 3.1.** (a) Die charakteristischen Polynome der Matrizen sind $\chi_{A_1} = -(X-2)(X-1)(X+1)$ und $\chi_{A_2} = -(X-1)(X-2)^2$. Schließlich berechnen wir $\chi_{A_3}$ und erhalten

$$\chi_{A_3}(x) = \det \begin{pmatrix} 1-x & 0 & 0 & 0 \\ 0 & 1-x & 0 & 0 \\ 0 & 0 & -x & -1 \\ 0 & 0 & 1 & -x \end{pmatrix} = \det \left[ \begin{pmatrix} 1-x & 0 \\ 0 & 1-x \end{pmatrix} \cdot \begin{pmatrix} -x & -1 \\ 1 & -x \end{pmatrix} \right]$$

$$= \det \begin{pmatrix} 1-x & 0 \\ 0 & 1-x \end{pmatrix} \cdot \det \begin{pmatrix} -x & -1 \\ 1 & -x \end{pmatrix} = (x-1)^2(x^2+1)\,.$$

Wir bestimmen nun die Minimalpolynome. Offenbar ist schon

$$\mu_{A_1} = \chi_{A_1}\,,$$

siehe auch die Bemerkungen nach Korollar XII.4 der Vorlesung. Weiter haben wir $\mu_{A_2} \in \{(X-1)(X-2), (X-1)(X-2)^2\}$. Man findet $(A_2-1)(A_2-2) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$, womit

$$\mu_{A_2} = (X-1)(X-2)^2\,.$$

Schließlich bestimmen wir $\mu_{A_3}$. Dazu beobachten wir zunächst, dass $(X^2+1) \in \mathbb{R}[X]$ irreduzibel ist. (Wir erinnern, dass $f \in K[X]$ per Definition genau dann irreduzibel ist, wenn für alle $h \in K[X]$, die $f$ teilen gilt, dass $\deg[h(x)] = 0$, oder $\deg[h(x)] = \deg[f(x)]$.) Seien also $g, h \in \mathbb{R}[X]$ so, dass $X^2+1 = g \cdot h$. Offenbar ist $2 = \deg(g) + \deg(h)$. Per Definition von Irreduzibilität genügt es zu zeigen, dass $\deg(g) = \deg(h) = 1$ nicht wahr sein kann. In dem Fall wären $g$ und $h$ Linearfaktoren, d.h. $g$ und $h$ haben insbesondere reelle Nullstellen. Das würde aber bedeuten, dass auch $g \cdot h = X^2+1$ reelle Nullstellen hätte, was aber offenbar nicht wahr ist. Somit ist $X^2+1$ irreduzibel.

Es verbleibt zu untersuchen, ob $\mu_{A_3}$ gleich $(x-1)^2(x^2+1)$ oder $(x-1)^1(x^2+1)$ ist. Man rechnet

$$(A_3^2+1)(A_3-1) = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix} = 0\,,$$

womit wir

$$\mu_{A_3}(x) = (x-1)(x^2+1)$$

gefunden haben.

(b) Wir verwenden Satz 2.5.

- $A_1$ diagonalisierbar, da $\mu_{A_1}$ in Linearfaktoren mit einfachen Nullstellen zerfällt.
- Zwar zerfällt auch $\mu_{A_2}$ in Linearfaktoren, doch mindestens einer dieser Linearfaktoren hat eine doppelte Nullstelle. Somit ist $A_2$ nicht diagonalisierbar.
- $A_3$ ist nicht diagonalisierbar, da $\mu_{A_3}$ nicht in Linearfaktoren zerfällt.

(c) Alle $A_j$ sind invertierbar, denn $\chi_{A_j}(x=0) = \det(A_j) \neq 0$.

(d) Für ein Minimalpolynom $\mu_A(x) = x^m + a_{m-1}x^{m-1} + ... + a_1 x + a_0$ schreiben wir $p_A(x) = -\frac{1}{a_0}\left(x^{m-1} + a_{m-1}x^{m-2} + ... + a_1\right)$. Somit erhalten wir

$$\begin{aligned} \mu_{A_1}(x) &= x^3 - 2x^2 - x + 2 &\Rightarrow\quad p_{A_1}(x) &= -\frac{1}{2}x^2 + x + \frac{1}{2} \\ \mu_{A_2}(x) &= x^3 - 5x^2 + 8x - 4 &\Rightarrow\quad p_{A_2}(x) &= \frac{1}{4}x^2 - \frac{5}{4}x + 2 \\ \mu_{A_3}(x) &= x^3 - x^2 + x - 1 &\Rightarrow\quad p_{A_3}(x) &= x^2 - x + 1 \end{aligned}$$

Einsetzen liefert

$$A_1^{-1} = \begin{pmatrix} -5/2 & 1 & -2 \\ 9/2 & -1 & 3 \\ 6 & -1 & 4 \end{pmatrix}, \quad A_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & -1/4 \\ 0 & 0 & 1/2 \end{pmatrix}, \quad A_3^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}.$$

**Lösung 3.2.** (a) Es sind $\varphi_1(1) = 0$, $\varphi_1(x) = 1$ und $\varphi_1(x^2) = 2x$ sowie $\varphi_2(1) = 1 + 2x^2$, $\varphi_2(x) = 2x$ und $\varphi_2(x^2) = 2x^2$. Somit sind

$$M_B(\varphi_1) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \qquad \text{und} \qquad M_B(\varphi_2) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix}$$

Matrixdarstellungen der $\varphi_j$ in der Monombasis $B$.

(b) Man rechnet

$$\chi_{\varphi_1}(x) = x^3,$$
$$\chi_{\varphi_2}(x) = (x-1)(x-2)^2.$$

(c) Mögliche Minimalpolynome $\mu_{\varphi_1}$ sind in der Menge $\{x^j : j = 1, 2, 3\}$ enthalten. Man rechnet, dass sowohl $M_B(\varphi_1) \neq 0$ als auch $M_B(\varphi_1)^2 \neq 0$. Somit ist $\mu_{\varphi_1}(x) = x^3$.

Mögliche Minimalpolynome $\mu_{\varphi_2}$ sind in der Menge $\{(x-1)(x-2), (x-1)(x-2)^2\}$ enthalten. Man rechnet, dass $(M_B(\varphi_2) - 1) \cdot (M_B(\varphi_2) - 2) = 0$. Somit ist $\mu_{\varphi_2}(x) = (x-1)(x-2)$.

(d) Zwar zerfällt $\mu_{\varphi_1}$ in Linearfaktoren, doch sind die Potenzen nicht alle eins. Im Gegenteil, $\mu_{\varphi_1}$ hat eine dreifache Nullstelle. Somit ist $\mu_{\varphi_1}$ nicht diagonalisierbar über $\mathbb{Q}$.

Das Minimalpolynom $\mu_{\varphi_2}$ zerfällt in Linearfaktoren mit einfacher Potenz. Somit ist $\varphi_2$ diagonalisierbar. Die Eigenwerte können sofort zu $\lambda_1 = 1$ und $\lambda_2 = 2$ abgelesen werden. Der (bis auf Normierung eindeutige) zu $\lambda_1 = 1$ gehörende Eigenvektor ist $w_1 = (-1, 0, 2)^T$. Die zu $\lambda_2 = 2$ gehörenden Eigenvektoren sind $w_2 = (0, 1, 0)^T$ und $w_3 = (0, 0, 1)^T$. Sei $S := (w_1, w_2, w_3) \in \mathbb{Q}^{3\times3}$. Dann ist $S^{-1} = S$ und in der Basis $B' = \{w_1, w_2, w_3\}$ ist somit

$$M_{B'}(\varphi_2) = S^{-1} M_B(\varphi_2) S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

(e) Da $\chi_{\varphi_1}(x = 0) = \det(\varphi_1) = 0$, ist $\varphi_1$ nicht invertierbar.

Da $\det(\varphi_2) \neq 0$, ist $\varphi_2$ invertierbar. Gemäß Teil (d) von Aufgabe 3.1 und $\mu_{\varphi_2}(x) = x^2-3x+2$, ist die Inverse durch das Polynom $-\frac{1}{2}x + \frac{3}{2}$, also

$$M_B(\varphi_2)^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ -1 & 0 & 1/2 \end{pmatrix}$$

gegeben.

# 4 4. Tutoriumsblatt (17. Mai 2022)

## 4.1 Diagonalisierbarkeit

Seien $K$ ein Körper, $V$ ein $n$-dimensionaler $K$-Vektorraum und $F \in \mathrm{End}(V)$ ein Endomorphismus mit Eigenwerten $\{\lambda_j\}_{j=1}^k \subseteq K$ und $k \leq n$. Der zu $\lambda \in K$ gehörende *Eigenraum* ist

$$\mathrm{Eig}(F;\lambda) := \{v \in V : Fv = \lambda v\} = \ker(F - \lambda)\,. \tag{4.1}$$

Offenbar ist $\mathrm{Eig}(F;\lambda) \setminus \{0\}$ die Menge der zu $\lambda$ gehörenden Eigenvektoren.

**Lemma 4.1.** *Seien $v_1, ..., v_m$ Eigenvektoren zu paarweise verschiedenen Eigenwerten $\lambda_1, ..., \lambda_m$. Dann sind die $v_1, ..., v_m$ linear unabhängig. Insbesondere ist $m \leq \dim(V) = n$.*

*Beweis.* Wir beweisen die Aussage per Induktion über $m$. Der Fall $m = 1$ ist wegen $v_1 \neq 0$ klar. Angenommen, die Aussage sei für $m - 1$ wahr und es gelte

$$\sum_{j=1}^m \alpha_j v_j = 0\,.$$

Darauf können wir einerseits $F$ anwenden und andererseits mit $\lambda_1$ multiplizieren. Dies ergibt die Gleichungen

$$\sum_{j=1}^m \alpha_j \lambda_j v_j = 0$$
$$\lambda_1 \sum_{j=1}^m \alpha_j v_j = 0\,.$$

Zieht man diese voneinander ab, erhält man

$$\sum_{j=2}^m \alpha_j (\lambda_j - \lambda_1) v_j = 0\,.$$

Per Induktionsannahme sind die $v_2, ..., v_m$ linear unabhängig. Daraus folgt, dass $\alpha_j(\lambda_j - \lambda_1) = 0$ für alle $j = 2, ..., m$. Da die Eigenwerte paarweise verschieden sind, müssen $\alpha_j = 0$ für alle $j = 2, ..., m$ sein. Setzt man dies in $\sum_{j=1}^m \alpha_j v_j = 0$ ein, folgt auch $\alpha_1 = 0$, also $\alpha_1 = ... = \alpha_m = 0$. Somit sind die $v_1, ..., v_m$ linear unabhängig voneinander. □

Aus dem Lemma folgt unmittelbar

**Korollar 4.2.** *Sind $\lambda_1, \lambda_2 \in K$ verschieden, so ist $\mathrm{Eig}(F;\lambda_1) \cap \mathrm{Eig}(F;\lambda_2) = \{0\}$.*

Wir erinnern daran, dass ein Endomorphismus genau dann diagonalisierbar ist, wenn die Eigenvektoren von $F$ eine Basis von $V$ bilden.

**Satz 4.3.** *Sei $F \in \mathrm{End}(V)$. Dann gelten folgende Aussagen.*

1. *Ist $F$ diagonalisierbar mit nicht notwendigerweise paarweise verschiedenen Eigenwerten $\lambda_1, ..., \lambda_n$, so zerfällt $\chi_F(x) = \prod_{j=1}^n (\lambda_j - x)$ in Linearfaktoren.*

2. *Ist $\chi_F(x) = \prod_{j=1}^n (\lambda_j - x)$ mit paarweise verschiedenen $\lambda_1, ..., \lambda_n \in K$, so ist $F$ diagonalisierbar.*

*Beweis.* 1. Folgt unmittelbar aus der Invarianz der Determinante unter Ähnlichkeitstransformationen.

2. Folgt aus der Tatsache, dass Eigenvektoren zu verschiedenen Eigenwerten linear unabhängig sind.

□

Oft sind die Eigenwerte von $F \in \mathrm{End}(V)$ entartet. Genauer haben wir folgende Notation.

**Definition 4.4.** Sei $F \in \mathrm{End}(V)$ mit charakteristischem Polynom $\chi_F(x) = \prod_{j=1}^k (\lambda_j - x)^{r_j}$ und paarweise verschiedenen $\{\lambda_j\}_{j=1}^k$ und $r_j \in \mathbb{N}$.

1. Die Zahl $m_g(\lambda_j) := \dim\ker(F - \lambda_j)$ heißt *geometrische Entartung/Vielfachheit* des Eigenwerts $\lambda_j$.

2. Die Zahl $m_a(\lambda_j) := r_j$ heißt *algebraische Entartung/Vielfachheit* des Eigenwerts $\lambda_j$.

**Korollar 4.5.** *Sei $F \in \mathrm{End}(V)$ und $\lambda$ ein Eigenwert. Dann gilt $1 \le m_g(\lambda) \le m_a(\lambda)$.*

*Beweis.* Sei $v_1, ..., v_s$ eine Basis von $\mathrm{Eig}(F, \lambda)$. Dann ist $1 \le m_g(\lambda)$ offenbar. Wir ergänzen $v_1, ..., v_s$ nun zu einer Basis

$$B := \{v_1, ..., v_s, v_{s+1}, ..., v_n\}$$

von $V$. Sei $A \in K^{n\times n}$ die Matrixdarstellung von $F$ in der Basis $B$. Dann haben wir offenbar

$$A = \begin{pmatrix} \lambda\mathbf{1}_s & * \\ 0 & A' \end{pmatrix}$$

für ein $A' \in K^{(n-s)\times(n-s)}$. Aus dem Kästchensatz (Satz VIII.1) folgt $\chi_F(x) = (\lambda - x)^s \cdot \chi_{A'}(x)$ und daher insbesondere $m_g(\lambda) = s \le m_a(\lambda)$. $\square$

**Satz 4.6.** *Sei $F \in \mathrm{End}(V)$. Dann sind folgende Aussagen äquivalent.*

*1. $F$ ist diagonalisierbar.*

*2. (a) Das charakteristische Polynom $\chi_F$ zerfällt in Linearfaktoren und*

*(b) $m_g(\lambda) = m_a(\lambda)$ für alle Eigenwerte $\lambda \in K$ von $F$.*

*3. Sind $\lambda_1, ..., \lambda_k$ die paarweise verschiedenen Eigenwerte von $F$, so gilt*

$$V = \bigoplus_{j=1}^k \mathrm{Eig}(F, \lambda_k) \tag{4.2}$$

Zum Beweis des Satzes erinnern wir an die Definition der direkten Summe. Es gilt

$$V = \bigoplus_{j=1}^k W_j$$

genau dann, wenn

1. $V = \sum_{j=1}^k W_j$ und

2. von Null verschiedene Vektoren $w_j \in W_j$ $(j = 1, ..., k)$ linear unabhängig voneinander sind.

**Wichtig:** Für $k > 2$ darf die zweite Bedingung *nicht* durch

$$W_1 \cap ... \cap W_k = \{0\} \quad \text{oder} \quad W_j \cap W_\ell = \{0\}\,, \quad \forall j \ne \ell$$

ersetzt werden!

Weiter gilt folgendes Lemma. (Siehe z.B. Satz 1.6.4 in [Fis05].)

**Lemma 4.7.** *Für Untervektorräume $W_1, ..., W_k$ eines endlichdimensionalen Vektorraums $V$ sind folgende Aussagen äquivalent.*

*1. $V = \bigoplus_{j=1}^k W_j$.*

2. *Ist für jedes $j \in \{1, ..., k\}$ eine Basis $\{v_j^{(1)}, ..., v_j^{(r_j)}\}$ von $W_j$ gegeben, so ist*

$$\{v_1^{(1)}, ..., v_1^{(r_1)}, ..., v_k^{(1)}, ..., v_k^{(r_k)}\}$$

*eine Basis von $V$.*

3. $V = \sum_{j=1}^k W_k$ *und* $\dim(V) = \sum_{j=1}^k \dim(W_j)$.

*Beweis von Satz 4.6.* (1) $\Rightarrow$ (2): Sei $F$ diagonalisierbar. Wir ordnen die Basis von $V$ der zugehörigen Eigenvektoren entsprechend den verschiedenen Eigenwerten $\lambda_1, ..., \lambda_k$, sprich, wir betrachten für $j = 1, ..., k$ die Basis

$$\{v_j^{(1)}, ..., v_j^{(s_j)}\} \quad \text{von} \quad \mathrm{Eig}(F, \lambda_j)\,.$$

Insbesondere gilt $\sum_{j=1}^k s_j = n$. Da $F$ diagonalisierbar ist, zerfällt $\chi_F(x) = \prod_{j=1}^k (\lambda_j - x)^{r_j}$ mit $\sum_{j=1}^k r_j = n$ (Satz 4.3) und $r_j \geq s_j$ (Korollar 4.5) für alle $j = 1, ..., k$. Daraus und $n = \sum_{j=1}^k s_j = \sum_{j=1}^k r_j$ folgt aber, dass $s_j = r_j$ sein muss.

(2) $\Rightarrow$ (3): Sei $W := \sum_{j=1}^k \mathrm{Eig}(F, \lambda_j)$. Nach Lemma 4.1 und der Definition der direkten Summe folgt $W = \bigoplus_{j=1}^k \mathrm{Eig}(F, \lambda_j)$. Deshalb folgt aus (1) $\Leftrightarrow$ (3) in Lemma 4.7, dass $\dim(W) = \sum_{j=1}^k \dim(\mathrm{Eig}(F, \lambda_j)) = \sum_{j=1}^k r_j = n$. Somit muss $W = V$ sein.

(3) $\Rightarrow$ (1): Für jedes $j = 1, ..., k$ sei $B_j := \{v_j^{(1)}, ..., v_j^{(s_j)}\}$ eine Basis von $\mathrm{Eig}(F, \lambda_j)$. Dank (1) $\Leftrightarrow$ (2) in Lemma 4.7 folgt, dass $B := \{v_1^{(1)}, ..., v_1^{(s_1)}, ..., v_k^{(1)}, ..., v_k^{(s_k)}\}$ eine Basis von $V$ ist. Da sie nach Definition aus Eigenvektoren von $F$ besteht, ist $F$ (per Definition) diagonalisierbar. Insbesondere ist $s_j = r_j$ für alle $j = 1, ..., k$ und wir haben

$$M_B(F) = \begin{pmatrix} \lambda_1 \mathbf{1}_{r_1} & 0 & 0 & \cdots \\ 0 & \lambda_2 \mathbf{1}_{r_2} & \ddots & \ddots \\ \vdots & \ddots & \ddots & 0 \\ \vdots & \ddots & 0 & \lambda_k \mathbf{1}_{r_k} \end{pmatrix}$$

□

*Bemerkung* 4.8. Sei $K$ ein algebraisch abgeschlossener Körper, sprich jedes nicht-konstante Polynom mit Koeffizienten in $K$ hat eine Nullstelle. Aus dem Fundamentalsatz der Algebra folgt, dass $\mathbb{C}$ algebraisch abgeschlossen ist. Dann sehen wir, dass $F \in \mathrm{End}(V)$ genau dann diagonalisierbar ist, wenn

$$\sum_{j=1}^k m_g(\lambda_j) = n\,.$$

*Bemerkung* 4.9. Rezept zur Diagonalisierung von $F \in \mathrm{End}(V)$:

1. Berechne $\chi_F(x)$. Falls $\chi_F$ nicht in Linearfaktoren zerfällt, so kann $F$ nicht diagonalisiert werden.

2. Bestimme eine Basis von $\mathrm{Eig}(F, \lambda)$ für alle Eigenwerte $\lambda \in K$ von $F$. Wenn $m_g(\lambda) = m_a(\lambda)$ für alle $\lambda$ gilt, so ist $F$ diagonalisierbar.

**Beispiel 4.10.** Betrachte

$$A = \begin{pmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix} \in \mathbb{R}^{3\times 3}$$

mit $\chi_A(x) = (1-x)^2(1+x)$. Also sind $\lambda_1 = 1$ und $\lambda_2 = -1$ die Eigenwerte. Man findet, dass $\{(1,0,1)^T, (0,1,1)^T\}$ eine Basis von $\mathrm{Eig}(A, 1)$ und $\{(1,3,2)^T\}$ eine Basis von $\mathrm{Eig}(A, -1)$ ist.

Somit ist $B = \{(1,0,1)^T, (0,1,1)^T, (1,3,2)^T\}$ eine Basis von $\mathbb{R}^3$ bestehend aus Eigenvektoren von $A$. Schreibt man

$$S = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 3 \\ 1 & 1 & 2 \end{pmatrix}$$

so findet man in der Tat $\mathrm{diag}(1,1,-1) = S^{-1}AS$.

## 4.2 Trigonalisierbarkeit

In Satz 4.6 haben wir gesehen, dass ein Endomorphismus $F \in \mathrm{End}(V)$ genau dann diagonalisierbar ist, wenn

1. $\chi_F(x)$ in Linearfaktoren zerfällt und

2. die algebraischen und geometrischen Vielfachheiten aller Eigenwerte übereinstimmen.

In diesem Abschnitt untersuchen wir den Fall, in dem zwar (1), jedoch nicht (2) erfüllt ist. Wir werden sehen, dass wir in diesem Fall $F$ immerhin noch auf obere Dreiecksform bringen können. Dieses Verfahren bezeichnet man als Trigonalisierung.

Wir erinnern zunächst an den Begriff des *invarianten Unterraums*. Ist $F \in \mathrm{End}(V)$ ein Endomorphismus, so heißt ein Unterraum $W \subseteq V$ invariant, falls $F(W) \subseteq W$.

**Beispiel 4.11.**
- Die trivialen Unterräume $W = \{0\}$ oder $W = V$ sind für alle Endomorphismen invariant.

- Sei $F$ diagonalisierbar mit Eigenwerten $\lambda_1, ..., \lambda_k \in K$. Dann sind $\mathrm{Eig}(F, \lambda_j)$ invariante, $m_g(\lambda_j) = m_a(\lambda_j)$-dimensionale Unterräume.

**Lemma 4.12.** *Ist $W \subseteq V$ ein $F$-invarianter Unterraum, so ist $\chi_{F|W}$ ein Teiler von $\chi_F$. (Hierbei meint $F|W$ die Einschränkung von $F$ auf $W$.)*

*Beweis.* Sei $B'$ eine Basis von $W$. Dann ergänzen wir $B'$ zu einer Basis $B$ von $V$. Dann ist

$$M_B(F) = \begin{pmatrix} M_{B'}(F|W) & * \\ 0 & A \end{pmatrix}$$

für ein $A \in K^{(n-\dim(W))\times(n-\dim(W)}$ und es gilt $\chi_F(x) = \chi_{F|W}(x)\chi_A(x)$ nach dem Kästchensatz. □

**Definition 4.13.** Unter einer *Fahne* $(V_r)_{r=0}^n$ in einem $n$-dimensionalen Vektorraum $V$ versteht man eine Kette

$$\{0\} = V_0 \subseteq V_1 \subseteq \cdots \subseteq V_n = V \tag{4.3}$$

von Untervektorräumen $V_r$ mit $\dim(V_r) = r$.

Ist $F \in \mathrm{End}(V)$, so heißt die Fahne $(V_r)_{r=1}^n$ genau dann $F$-invariant, wenn $F(V_r) \subseteq V_r$ für alle $r = 0, 1, ..., n$.

*Bemerkung* 4.14. Man kann sich $V_0$ als Befestigungspunkt, $V_1$ als Stange, $V_2$ als Tuch, etc. der Fahne vorstellen.

Mit Hilfe jeder Basis eines Vektorraums kann man viele Fahnen konstruieren. Aber aus der Existenz einer $F$-invarianten Fahne folgt insbesondere aus $F(V_1) \subseteq V_1$, dass es einen Eigenvektor geben muss.

**Lemma 4.15.** *Für $F \in \mathrm{End}(V)$ sind folgende Aussagen äquivalent.*

1. *Es gibt eine $F$-invariante Fahne in $V$.*

2. *Es gibt eine Basis $B$ von $V$, sodass $M_B(F)$ eine obere Dreiecksmatrix ist.*

*Ist eine (und damit beide) der Aussagen erfüllt, so heißt $F$ trigonalisierbar.*

*Beweis.* Die Beziehung zwischen Fahne $(V_r)_{r=1}^n$ und Basis $B = (v_1, ..., v_n)$ ist durch $V_r = \operatorname{span}(v_1, ..., v_r)$ für alle $r = 1, ..., n$ geregelt. $\square$

**Definition 4.16.** Eine Matrix $A \in K^{n\times n}$ heißt trigonalisierbar, wenn es $S \in \mathrm{GL}(n, K)$ gibt, sodass $S^{-1}AS$ eine obere Dreiecksmatrix ist.

Folgender Satz ist das Hauptergebnis dieses Abschnitts.

**Satz 4.17** (Trigonalisierungssatz)**.** *Für einen Endomorphismus $F \in \mathrm{End}(V)$ eines $n$-dimensionalen $K$-Vektorraums sind folgende Aussagen äquivalent.*

1. *$F$ ist trigonalisierbar.*
2. *Das charakteristische Polynom $\chi_F(x) = \prod_{j=1}^n (\lambda_j - x)$ zerfällt in Linearfaktoren mit nicht notwendigerweise verschiedenen $\lambda_1, ..., \lambda_n \in K$.*

**Korollar 4.18.** *Ist $K$ algebraisch abgeschlossen (wie z.B. im Falle $K = \mathbb{C}$), so ist jeder Endomorphismus $F \in \mathrm{End}(V)$ in einem endlichdimensionalen Vektorraum $V$ trigonalisierbar.*

*Beweis von Satz 4.17.* (1) $\Rightarrow$ (2): Folgt aus dem Kästchensatz.

(2) $\Rightarrow$ (1): Wir induzieren über $n$. Für $n = 0, 1$ ist nichts zu zeigen. Ist $n \geq 2$, so wähle man einen Eigenvektor $v_1 \in V \setminus \{0\}$ zum Eigenwert $\lambda_1 \in K$. Diesen ergänzen wir zu einer Basis $B = (v_1, w_2, ..., w_n)$ von $V$. Dann ist

$$V = V_1 \oplus W \quad \text{mit} \quad V_1 := \operatorname{span}(v_1) \quad \text{und} \quad W := \operatorname{span}(w_2, ..., w_n)$$

und wir haben

$$M_B(F) = \left(\begin{array}{c|ccc} \lambda_1 & a_{12} & \cdots & a_{1n} \\ \hline 0 & & & \\ \vdots & & \tilde{A} & \\ 0 & & & \end{array}\right)$$

mit $(\tilde{A})_{\ell,k} := a_{\ell,k}$ und $\ell, k \in \{2, ..., n\}$. Dann ist $V_1$ offenbar $F$-invariant. Jedoch ist $W$ wegen der Einträge $a_{12}, ..., a_{1n}$ noch nicht $F$-invariant. Für $w \in W$ können wir $F(w) = H(w) + G(w)$ mit

$$H(w_j) = a_{1j}v_1 \quad \text{und} \quad G(w_j) = a_{2j}w_2 + ... + a_{nj}w_n\,, \quad j \in \{2, ..., n\}$$

zerlegen, wobei offenbar $H : W \to V_1$ und $G : W \to W$ linear abbilden. Für die charakteristischen Polynome gilt

$$\chi_F(x) = (\lambda_1 - x) \cdot \chi_G(x)\,, \quad \text{also} \quad \chi_G(x) = \prod_{j=2}^n (\lambda_j - x)\,.$$

Wendet man die Induktionsvoraussetzung auf $W$ an, so erhalten wir eine $G$-invariante Fahne

$$\{0\} = W_0 \subseteq W_1 \subseteq \cdots \subseteq W_{n-1} = W\,.$$

Wir behaupten nun, dass durch $V_r := V_1 + W_{r-1}$ eine $F$-invariante Fahne gegeben ist. Erstens sind die Einbettungen $V_{\ell-1} \subseteq V_\ell$ mit $\ell \in \{1, ..., n-1\}$ offenbar. Zweitens gilt für $w \in W_{r-1}$ in der Tat

$$F(\mu v_1 + w) = \lambda_1\mu_1 v_1 + H(w) + G(w) \subseteq V_1 + W_{r-1}\,,$$

da $H(w) \in V_1$ und $G(w) \in W_{r-1}$ für $w \in W_{r-1}$. Dies schließt den Beweis. $\square$

In [Fis05, Abschnitt 4.4.5] findet man ein explizites Rezept zur Trigonalisierung. Dieses besprechen wir hier jedoch nicht, sondern gehen sofort zur verbesserten *Trigonalisierungsmethode von Jordan*. Dazu benötigen wir zwei vorbereitende Schritte: die *Hauptraumzerlegung* und das *Maßschneidern einer Basis für nilpotente Matrizen.*

### 4.3 Hauptraumzerlegung

Satz 4.17 besagt, dass Endomorphismen in $K$-Vektorräumen über algebraisch abgeschlossenen Körpern immer trigonalisierbar sind, und, dass die Diagonaleinträge durch die Eigenwerte, also die Nullstellen der Linearfaktoren des charakteristischen Polynoms gegeben sind. Der Satz besagt jedoch nichts über die Einträge oberhalb der Diagonalen. Diese Einträge sind Gegenstand der folgenden Untersuchung.

Wir beginnen mit folgenden Beobachtungen zu nilpotenten Endomorphismen. Wir erinnern, dass $F \in \mathrm{End}(V)$ in einem $n$-dimensionalen $K$-Vektorraum nilpotent heißt, wenn es $k \in \mathbb{N}$ gibt, sodass $F^k = 0$.

**Satz 4.19.** *Sei $F \in \mathrm{End}(V)$. Dann sind folgende Aussagen äquivalent.*

1. *$F$ ist nilpotent.*
2. *$F^d = 0$ für ein $d \in \{1, ...n\}$.*
3. *$\chi_F(x) = (-x)^n$.*
4. *Es gibt eine Basis $B$ von $V$, sodass*

$$M_B(F) = \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix}$$

.

*Beweis.* (1) $\Rightarrow$ (3): Angenommen es gibt $k \in \mathbb{N}$, sodass $F^k = 0$. Dann hat $F$ Null als einzigen Eigenwert. (Sei dazu $\lambda \in K$ ein Eigenwert mit Eigenvektor $v \in V\setminus\{0\}$. Dann gilt $0 = F^k v = \lambda^k v$. Da $v \neq 0$ ist, muss $\lambda = 0$ sein.) Das bedeutet, dass $\chi_F(x) = (-x)^n$.

(3) $\Rightarrow$ (2): Wegen Cayley–Hamilton gilt $F^n = 0$. Somit gibt es $d \in \{1, ..., n\}$, sodass $F^d = 0$.

(3) $\Rightarrow$ (4): Folgt aus der Tatsache, dass Null der einzige Eigenwert von $F$ ist und dem Trigonalisierungssatz 4.17.

(4) $\Rightarrow$ (1): Nach dem Kästchensatz folgt $\chi_{M_B(F)}(x) = (-x)^n = \chi_F(x)$. Aus Cayley–Hamilton folgt $F^n = 0$. □

Im Folgenden sei $K$ stets ein beliebiger Körper und $V$ ein $n$-dimensionaler $K$-Vektorraum. Sei $F \in \mathrm{End}(V)$ mit

$$\chi_F(x) = (\lambda_1 - x)^{r_1} \cdots (\lambda_k - x)^{r_k}$$

und paarweise verschiedenen $\lambda_1, ..., \lambda_k$ und $r_1, ..., r_k \in \mathbb{N}$. Sind $m_g(\lambda_j) = r_j$ für alle $j = 1, ..., k$, so ist $F$ diagonalisierbar mit $V = \bigoplus_{j=1}^k \mathrm{Eig}(F, \lambda_j)$. Falls der Eigenraum $\mathrm{Eig}(F, \lambda_j)$ zu klein sein sollte, also, wenn $m_g(\lambda_j) < m_a(\lambda_j)$, so kann man ihn mit der Beobachtung

$$\mathrm{Eig}(F, \lambda_j) = \ker(F - \lambda_j) \subseteq \ker((F - \lambda_j)^{m_a(\lambda_j)}) =: \mathrm{Hau}(F, \lambda_j)$$

vergrößern.

**Definition 4.20.** Sei $F \in \mathrm{End}(V)$ mit einem Eigenwert $\lambda \in K$ algebraischer Vielfachheit $r \in \mathbb{N}$. Dann bezeichnet man mit

$$\mathrm{Hau}(F, \lambda) := \ker((F - \lambda)^r) = \{v \in V : (F - \lambda)v = 0 \text{ für ein } r \in \mathbb{N}\} \tag{4.4}$$

den dazugehörigen *Hauptraum* oder *verallgemeinerten Eigenraum.*

Für trigonalisierbare Matrizen haben wir dann folgendes wichtige Resultat, das die Eigenraumzerlegung verallgemeinert.

**Satz 4.21** (Hauptraumzerlegung). *Sei $F \in \mathrm{End}(V)$ und $\chi_F(x) = \prod_{j=1}^k (\lambda_j - x)^{r_j}$ mit paarweise verschiedenen $\lambda_1, ..., \lambda_k \in K$ und $r_1, ..., r_k \in \mathbb{N}$. Wir bezeichnen mit $V_j := \mathrm{Hau}(F, \lambda_j)$ den zu $\lambda_j$ gehörenden Hauptraum. Dann gelten folgende Aussagen.*

1. *$F(V_j) \subseteq V_j$ und $\dim(V_j) = r_j$ für alle $j = 1, ..., k$.*

2. *$V = \bigoplus_{j=1}^k V_j$.*

3. *$F$ hat eine Zerlegung $F = F_D + F_N$ mit*

   (a) *$F_D$ ist diagonalisierbar,*

   (b) *$F_N$ ist nilpotent und*

   (c) *$F_D F_N = F_N F_D$.*

*Beweis.* Siehe [Fis05, Abschnitte 4.6.1–4.6.3]. □

Übersetzt man dies in die Sprache der Matrizen und verwendet, dass jede nilpotente Matrix in eine strikte obere Dreiecksmatrix transformiert werden kann, so erhält man

**Korollar 4.22.** *Sei $A \in K^{n\times n}$ mit $\chi_A(x) = \prod_{j=1}^k (\lambda_j - x)^{r_j}$ mit paarweise verschiedenen $\lambda_1, ..., \lambda_k \in K$ und $r_1, ..., r_k \in \mathbb{N}$. Dann gibt es $S \in \mathrm{GL}(n, K)$, sodass*

$$S^{-1}AS = \begin{pmatrix} \boxed{\lambda_1 \mathbf{1}_{r_1} + N_1} & & \\ & \ddots & \\ & & \boxed{\lambda_k \mathbf{1}_{r_k} + N_k} \end{pmatrix} =: \tilde{A}\,,$$

*wobei für $j = 1, ..., k$*

$$\lambda_j \mathbf{1}_{r_j} + N_j = \begin{pmatrix} \lambda_j & & * \\ & \ddots & \\ & & \lambda_j \end{pmatrix} \in K^{r_j \times r_j}\,.$$

*Insbesondere ist $\tilde{A} = D + N$, wobei $D = \mathrm{diag}(\underbrace{\lambda_1, ..., \lambda_1}_{r_1 \text{ mal}}, \underbrace{\lambda_2, ..., \lambda_2}_{r_2 \text{ mal}}, ..., \underbrace{\lambda_k, ..., \lambda_k}_{r_k \text{ mal}})$ und $N$ nilpotent ist mit $DN = ND$.*

**Beispiel 4.23.** Sei $A = \begin{pmatrix} 25 & 34 & 18 \\ -14 & -19 & -10 \\ -4 & -6 & -1 \end{pmatrix}$ mit $\chi_A(x) = (t-1)^2(t-3)$. Die Eigenwerte $\lambda_1 = 1$ und $\lambda_2 = 3$ sind $r_1 = 2$- bzw. $r_2 = 1$-fach algebraisch entartet. Man findet (siehe die Rechnungen in der darauffolgenden Bemerkung), dass $\mathrm{rang}(A - \lambda_1) = 2 > 1 = \mathrm{rang}((A - \lambda_1)^2)$, was bedeutet, dass $\mathrm{Eig}(A, \lambda_1) = \ker(A - \lambda_1) \subsetneq \ker((A - \lambda_1)^2) = \mathrm{Hau}(A, \lambda_1)$. Insbesondere ist $A$ nicht diagonalisierbar. Man findet, dass $\{(2, 0, -1)^T, (0, 2, -1)^T\}$ eine Basis von $\mathrm{Hau}(A, \lambda_1)$ und $\{(-7, 4, 1)^T\}$ eine Basis von $\mathrm{Hau}(A, \lambda_2) = \mathrm{Eig}(A, \lambda_2)$ bereitstellt. Schreibt man

$$S = \begin{pmatrix} 2 & 0 & -7 \\ 0 & 2 & 4 \\ -1 & -1 & 1 \end{pmatrix},$$

so findet man

$$S^{-1}AS = \begin{pmatrix} 16 & 25 & 0 \\ -9 & -14 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

mit

$$\begin{pmatrix} 16 & 25 \\ -9 & -14 \end{pmatrix} \equiv 1 \cdot \mathbf{1}_2 + N \quad \text{und } N^2 = 0\,.$$

Schließlich transformieren wir noch die Basis von $\mathrm{Hau}(A, \lambda_1)$ so, dass einer der Basisvektoren zu einem Eigenvektor $v$ zu $\lambda_1$ wird. Dieser hat die Form

$$v = \alpha(2, 0, -1)^T + \beta(0, 2, -1)^T\,.$$

Die Bedingung $Av = \lambda_1 v$ ergibt $3\alpha + 5\beta = 0 \Leftrightarrow \alpha = -5\beta/3$, sprich wir können $\beta = -3/2$ und somit $v = (5, -3, -1)^T$ wählen. Damit ist der Basisvektor $(0, 2, -1)^T$ eliminiert. Schreibt man die neuen Basisvektoren in

$$T = \begin{pmatrix} 5 & 2 & -7 \\ -3 & 0 & 4 \\ -1 & -1 & 1 \end{pmatrix}$$

so findet man schließlich

$$T^{-1}AT = \begin{pmatrix} 1 & 6 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix} .$$

<u>Alternativ:</u> Die Matrix $\begin{pmatrix} 16 & 25 \\ -9 & -14 \end{pmatrix}$ können wir auch mit dem Verfahren aus Abschnitt 6 trigonalisieren. Ein zu eins gehörender Eigenvektor ist $(-5, 3)^T$. Schreiben wir also

$$T_1 = \begin{pmatrix} -5 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} ,$$

so erhalten wir

$$T_1^{-1} S^{-1} A S T_1 = \begin{pmatrix} 1 & -5 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix} .$$

**Nebenrechnungen zu Beispiel 4.23:**

Wir rechnen noch $\operatorname{rang}(A - \lambda_1) = 2 > 1 = \operatorname{rang}((A - \lambda_1)^2)$ sowie $\ker((A - \lambda_1)^2)$ nach. Wir haben

$$A - \lambda_1 = \begin{pmatrix} 24 & 34 & 18 \\ -14 & -20 & -10 \\ -4 & -6 & -2 \end{pmatrix}$$

und sehen, dass

$$24 = 14a + 4b \quad 34 = 20a + 6b \quad 18 = 10a + 2b$$

z.B. durch $a = 2$ und $b = -1$ gelöst wird. Das heißt, dass die erste Zeile multipliziert mit minus eins gleich zwei mal der zweiten minus einmal der dritten Zeile ist. Somit ist $\operatorname{rang}(A - 1) = 2$.

Weiter haben wir

$$(A - \lambda_1)^2 = \begin{pmatrix} 28 & 28 & 56 \\ -16 & -16 & -32 \\ -4 & -4 & -8 \end{pmatrix} .$$

Da die ersten beiden Spalten identisch sind und die dritte Spalte gleich zwei mal der ersten Spalte ist, folgt $\operatorname{rang}((A - 1)^2) = 1$.

Schließlich bestimmen wir den zweidimensionalen $\ker((A - \lambda_1)^2)$. Da alle Spalten linear abhängig voneinander sind, genügt es

$$x + y + 2z = 0$$

zu lösen. Dies ist ein Gleichungssystem mit einer Gleichung für drei Variablen, d.h. wir haben zwei freie Parameter. Wir können z.B. $y = 0$ und $z = -1$ setzen, dann erhalten wir $x = 2$. Dies liefert die erste Lösung $(2, 0, -1)^T$. Wählen wir $x = 0$ und $z = -1$, erhalten wir $y = 2$, was auf die zweite Lösung $(0, 2, -1)^T$ führt, welche von der ersten linear unabhängig ist.

## 4.4 Jordanisierung nilpotenter Matrizen

Wir maßschneidern nun Basen zu nilpotenten Matrizen. Dazu führen wir den Jordan-Block

$$J_k = \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix} \in K^{k\times k}.$$

Offenbar gilt $J_k^k = 0$. Beachte $J_1 = (0)$.

**Satz 4.24.** *Sei $V$ ein $r$-dimensionaler $K$-Vektorraum, $G \in \mathrm{End}(V)$ nilpotent und $d := \min\{\ell : G^\ell = 0\}$. Dann gibt es eindeutig bestimmte $s_1, ..., s_d \in \mathbb{N}$ mit*

$$d \cdot s_d + (d-1)s_{d-1} + ... + s_1 = r = \dim(V)$$

*und eine Basis $B$ von $V$, sodass*

$$M_B(G) = \mathrm{diag}(\underbrace{J_d, ..., J_d}_{s_d \text{ mal}}, \underbrace{J_{d-1}, ..., J_{d-1}}_{s_{d-1} \text{ mal}}, ..., \underbrace{J_1, ..., J_1}_{s_1 \text{ mal}}). \tag{4.5}$$

*Beweis.* Siehe [Fis05, Abschnitt 4.6.5]. Das Verfahren wird auch noch in der Vorlesung behandelt. □

**Beispiel 4.25.** Sei $B = \begin{pmatrix} 0 & 1 & 3 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$. Dann sind $B^2 = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ und $B^3 = 0$, sprich $d = 3$ in obiger Notation. Für $K^{(\ell)} := \ker(B^\ell)$ gilt

$$\{0\} = K^{(0)} \subsetneq K^{(1)} = \mathrm{span}(e_1) \subsetneq K^{(2)} = \mathrm{span}(e_1, e_2) \subsetneq K^{(3)} = \mathbb{R}^3.$$

Ziel: Schrittweise $V$ in eine direkte Summe zerlegen. Wir wählen zunächst ein $W_d = W_3$ in $V = \mathbb{R}^3$ mit

$$\mathbb{R}^3 = V = K^{(3)} = K^{(2)} \oplus W_3.$$

In diesem Beispiel können wir $\boxed{W_3 = \mathrm{span}(e_3)}$ wählen. In der Notation von Satz 4.24 ist $s_3 = 1$. (Damit ist sofort klar, dass $s_2 = s_1 = 1$ sein müssen.)

Um $K^{(2)}$ weiter zu zerlegen, verwenden wir die Tatsache $B^{-1}K^{(\ell-1)} = K^{(\ell)}$ für $1 \le \ell \le d$. Insbesondere ist damit $BK^{(\ell)} \subseteq K^{(\ell-1)}$. Das bedeutet aber wegen $K^{(3)} = K^{(2)} \oplus W_3$, dass $BW_3 \subseteq K^{(2)}$ und (wegen $K^{(1)} \subsetneq K^{(2)}$), dass $BW_3 \cap K^{(1)} = \{0\}$. Somit gibt es eine Zerlegung

$$\mathbb{R}^3 = K^{(2)} \oplus W_3 = K^{(1)} \oplus W_2 \oplus W_3 \quad \text{mit } B(W_3) \subseteq W_2.$$

Eine Basis von $W_2$ erhalten wir durch Vektoren in $BW_3$ und möglicherweise ergänzt durch $s_2$ weitere Vektoren. Da $\dim(W_2) = 1$ ist $\boxed{W_2 = \mathrm{span}\{Be_3\} = \mathrm{span}\{(3,2,0)^T\}}$ der richtige Basisvektor von $W_2$; insbesondere is $s_2 = 0$, da nicht mit anderen Vektoren ergänzt werden muss. Schließlich haben wir die Zerlegung

$$\mathbb{R}^3 = \{0\} \oplus W_1 \oplus W_2 \oplus W_3.$$

Eine Basis von $W_1$ erhalten wir durch Vektoren in $B^2W_3$, ergänzt durch Vektoren in $BW_2$ und möglicherweise ergänzt durch weitere $s_1$ Vektoren. Da $\dim W_1$, ist $\boxed{W_1 = \mathrm{span}\{B^2e_3\} = \mathrm{span}\{2e_1\}}$ der korrekte Basisvektor von $W_1$; insbesondere ist $s_1 = 0$. Schreibt man die Basisvektoren $W_1$, $W_2$ und $W_3$ in dieser Reihenfolge in

$$T = \begin{pmatrix} 2 & 3 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

so finden wir

$$T^{-1}BT = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

## 4.5 Jordan-Normalform

Durch Kombination der Hauptraumzerlegung und der Maßschneiderung von Basen für nilpotente Matrizen finden wir

**Satz 4.26** (Jordan-Normalform)**.** *Sei $F \in \operatorname{End}(V)$ so, dass $\chi_F(x) = \prod_{j=1}^{k} (\lambda_j - x)^{r_j}$ mit paarweise verschiedenen $\lambda_1, ..., \lambda_k \in K$. Dann gibt es eine Basis $B$ von $V$, sodass*

$$M_B(F) = \begin{pmatrix} \boxed{\lambda_1 \mathbf{1}_{r_1} + N_1} & & 0 \\ & \ddots & \\ 0 & & \boxed{\lambda_k \mathbf{1}_{r_k} + N_k} \end{pmatrix},$$

*wobei $N_j$ in der Normalform (4.5) geschrieben ist. Ausgeschrieben bedeutet das*

$$\lambda_j \mathbf{1}_{r_j} + N_j = \begin{pmatrix} \lambda_j & 1 & & & & & & & & & \\ & \ddots & \ddots & & & & & & & & \\ & & \ddots & 1 & & & & & & & \\ & & & \lambda_j & 0 & & & & & & \\ & & & & \lambda_j & 1 & & & & & \\ & & & & & \ddots & \ddots & & & & \\ & & & & & & \ddots & 1 & & & \\ & & & & & & & \lambda_j & 0 & & \\ & & & & & & & & \ddots & \ddots & \\ & & & & & & & & & \ddots & 0 \\ & & & & & & & & & & \lambda_j \end{pmatrix}$$

Die Nuller rechts neben den $\lambda_j$'s deuten an, wo ein Jordanblock zu Ende ist.

*Beweis.* Für $j = 1, ..., k$ setzen wir

$$V_j := \operatorname{Hau}(F, \lambda_j) \quad \text{und} \quad G_j := (F - \lambda_j)|_{V_j} .$$

Wendet man Satz 4.24 auf $G_j$ an, so erhält man eine Basis $B_j$ von $V_j$. Aus den gefundenen Basen $B_1, ..., B_k$ baut man sich die gesuchte Basis $B$ von $V$. □

Wir schließen mit zwei Beobachtungen. Die Teilmatrix, die zum Eigenwert $\lambda_j$ gehört, besteht aus $s_1^{(j)} + ... + s_{d_j}^{(j)}$ vielen Jordanblöcken, deren Superdiagonale aus 0, 1, ..., bzw. $d_j - 1$ vielen Einsern besteht. Der oberste und größte Jordanblock in $\lambda_j \mathbf{1}_{r_j} + N_j$ hat die Größe $d_j$ mit

$$1 \leq d_j \leq \min\{\ell \in \mathbb{N} : N_j^\ell = 0\} \leq r_j .$$

Die Zahl $d_j$ ist genau die Potenz, zu der der Linearfaktor $(x - \lambda_j)$ im Minimalpolynom vorkommt. Zweitens ist die Zahl der Jordanblöcke gerade die geometrische Multiplizität.

**Satz 4.27.** *Sei $F \in \operatorname{End}(V)$, dessen charakteristisches Polynom in Linearfaktoren $\chi_F(x) = \prod_{j=1}^{k} (\lambda_j - x)^{r_k}$ mit paarweise verschiedenen $\lambda_1, ..., \lambda_k \in K$ zerfällt. Dann gelten folgende Aussagen.*

1. *Das Minimalpolynom ist $\mu_F(x) = \prod_{j=1}^{k} (x - \lambda_j)^{d_j}$, wobei $d_j$ die Größe des größten zu $\lambda_j$ gehörenden Jordanblocks bezeichnet.*

2. *Die geometrische Multiplizität von $\lambda_j$ ist gleich der Zahl der zu $\lambda_j$ gehörenden Jordan-Blöcke, also $s_{d_j} + ... + s_1$.*

*Beweis.* Siehe [Fis05, Aufgabe 8 in Abschnitt 4.6]. □

Die zweite Aussage in Satz 4.27 folgt aus der Tatsache $m_g(\lambda \mathbf{1}_k + J_k, \lambda) = 1$ für einen $k \times k$ Jordan-Block und folgender Beobachtung.

**Satz 4.28.** *Sei $F \in \operatorname{End}(V)$ und $V = V_1 \oplus V_2$ so, dass $F(V_j) \subseteq V_j$. (Das heißt es gibt eine Basis in der $F$ als Blockdiagonalmatrix dargestellt werden kann.) Dann sind die algebraischen bzw. geometrischen Multiplizitäten eines Eigenwerts $\lambda \in K$ von $F$ gleich der Summe der algebraischen bzw. geometrischen Multiplizitäten des Eigenwerts $\lambda$ von $F|_{V_1}$ und $F|_{V_2}$, sprich*

$$m_a(F,\lambda) = m_a(F|_{V_1},\lambda) + m_a(F|_{V_2},\Lambda) \quad \textit{und} \quad m_g(F,\lambda) = m_g(F|_{V_1},\lambda) + m_g(F|_{V_2},\Lambda)\,.$$

In der Matrixsprache bedeutet das, dass die algebraischen bzw. geometrischen Vielfachheiten der Eigenwerte der Blockdiagonalmatrix $A = \operatorname{diag}(B,C)$ gleich der Summe der algebraischen bzw. geometrischen Vielfachheiten der Eigenwerte von $B$ und $C$ sind, wobei $A \in K^{\dim(V)\times\dim(V)}$, $B \in K^{\dim(V_1)\times\dim(V_1)}$ und $C \in K^{\dim(V_2)\times\dim(V_2)}$.

*Beweis.* Für die algebraischen Multiplizitäten folgt dies bspw. aus der Matrixdarstellung und dem Kästchensatz.

Nun zu den geometrischen Multiplizitäten. Seien $E_\lambda := \operatorname{Eig}(F,\lambda)$ sowie $F_j := F|_{V_j}$ für $j = 1,2$. Dann ist $E_\lambda \cap V_j$ der zu $\lambda$ gehörende Eigenraum von $F_j$. Offenbar ist $E_\lambda = (E_\lambda \cap V_1) \oplus (E_\lambda \cap V_2)$. Die Inklusion "$\supseteq$" ist klar; die Inklusion "$\subseteq$" kann wie folgt gezeigt werden. Sei $x \in E_\lambda$. Da $V = V_1 \oplus V_2$, können wir $x = x_1 + x_2$ für $x_j \in V_j$ zerlegen. Da $F(x) = \lambda x$, gilt

$$F_1(x_1) + F_2(x_2) = F(x_1 + x_2) = \lambda(x_1 + x_2) = \lambda x_1 + \lambda x_2\,.$$

Da $F_j(x_j) \in V_j$ und $V = V_1 \oplus V_2$, erhalten wir die Implikation

$$F_1(x_1) + F_2(x_2) = \lambda x_1 + \lambda x_2 \Rightarrow F_j(x_j) = \lambda x_j \quad \text{für } j = 1,2\,.$$

Das zeigt aber gerade, dass $x \in (E_\lambda \cap V_1) \oplus (E_\lambda \cap V_2)$ und somit $E_\lambda = (E_\lambda \cap V_1) \oplus (E_\lambda \cap V_2)$ wie behauptet. Daraus und Lemma 4.7 folgt schließlich

$$\dim(E_\lambda \cap V_1) + \dim(E_\lambda \cap V_2) = \dim((E_\lambda \cap V_1) \oplus (E_\lambda \cap V_2)) = \dim(E_\lambda)\,.$$

Die linke Seite ist gerade die Summe der geometrischen Multiplizitäten von $\lambda$ bzgl. der Blöcke $F_1$ und $F_2$, und die rechte Seite ist gerade geometrische Multiplizität von $\lambda$ bzgl. $F$. Dies schließt den Beweis. $\square$

**Beispiel 4.29.** Betrachte $A = \begin{pmatrix} 3 & 4 & 3 \\ -1 & 0 & -1 \\ 1 & 2 & 3 \end{pmatrix}$ mit $\chi_A(x) = (2-x)^3$. Wir haben es somit mit einem einzigen Hauptraum zu tun.

Definiere $B = A - 2 = \begin{pmatrix} 1 & 4 & 3 \\ -1 & -2 & -1 \\ 1 & 2 & 1 \end{pmatrix}$ mit $B^2 = (A-2)^2 = \begin{pmatrix} 0 & 2 & 2 \\ 0 & -2 & -2 \\ 0 & 2 & 2 \end{pmatrix}$ und $B^3 = 0$ (Cayley–Hamilton). Es sind

$$K^{(1)} = \ker(B) = \operatorname{span}\{(1,-1,1)^T\} \quad \text{und} \quad K^{(2)} = \ker(B^2) = \operatorname{span}\{(1,-1,1)^T, (0,-1,1)^T\}.$$

Wie in Beispiel 4.25 haben wir die Zerlegungen

$$\mathbb{R}^3 = K^{(2)} \oplus W_3 = K^{(1)} \oplus W_2 \oplus W_1 = \{0\} \oplus W_1 \oplus W_2 \oplus W_3$$

mit $\boxed{W_3 = \operatorname{span}\{e_3\}}$ (womit $s_3 = 1$ und daher wieder $s_1 = s_2 = 0$ wegen $r = 3 = d \cdot s_d + \ldots = 3s_3 + 2s_2 + s_1$). Als Nächstes sind $\boxed{W_2 = \operatorname{span}\{Be_3\} = \operatorname{span}\{(3,-1,1)^T\}}$ und schließlich $\boxed{W_1 = \operatorname{span}\{B^2 e_3\} = \operatorname{span}\{(2,-2,2)^T\}}$. Schreibt man diese Basisvektoren von $W_1$, $W_2$ und $W_3$ in dieser Reihenfolge in

$$T = \begin{pmatrix} 2 & 3 & 0 \\ -2 & -1 & 0 \\ 2 & 1 & 1 \end{pmatrix}$$

erhält man

$$T^{-1}AT = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} .$$

*Bemerkung* 4.30. In Beispiel 4.29 hatte $A$ nur einen Eigenwert. Hätte $A$ mehrere Eigenwerte, so müsste zunächst eine Hauptraumzerlegung durchgeführt werden. Anschließend müsste obiges Verfahren für jeden Hauptraum individuell durchgeführt werden. Da der gesamte Raum die direkte Summe der Haupträume ist, ergeben die unabhängig voneinander konstruierten Basen der Haupträume zusammen eine Basis der gewünschten Art des gesamten Raumes.

# 5 5. Tutoriumsblatt (24. Mai und 02. Juni 2022)

Im letzten Tutoriumsblatt hatten wir

$$J_k = \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix} \in K^{k\times k}.$$

eingeführt. Um mit der Notation der Vorlesung konsistent zu sein, bezeichnen wir diese Matrix nun mit $\tilde{J}_k$ und redefinieren

$$J_k = \begin{pmatrix} 0 & & & 0 \\ 1 & 0 & & 0 \\ & \ddots & \ddots & \\ 0 & \ddots & 1 & 0 \end{pmatrix} \in K^{k\times k}.$$

Zur Interpretation (und Verifikation) der Ergebnisse der folgenden Aufgabe:

- Die Zahl der zu einem Eigenwert gehörenden Jordanblöcke ist gleich der geometrischen Multiplizität des Eigenwerts in Frage. (Satz 4.27)
- Die Größe des größten Jordanblocks ist gleich dem Exponenten des entsprechenden Linearfaktors im Minimalpolynom. (Satz 4.27)
- Die Summe der Zahlen aller Jordanblöcke einer Größe $d_j$ multipliziert mit ihrer Häufigkeit ist gleich dem Exponenten des entsprechenden Linearfaktors im charakteristischen Polynom. (Satz 4.24)

*Bemerkung* 5.1 (Algorithmus zum Jordanisieren). Sei $X$ ein $\mathbb{F}$-Vektorraum mit $N = \dim(X)$ und $\Phi \in \mathrm{End}(X)$ nilpotent mit $\mu_\Phi(x) = x^M$ für ein $M \in \{1, ..., N\}$.

1. Bestimme die $K^{(m)} := \ker(\Phi^m)$ für $m = 1, ..., M$. Offenbar ist $K^{(M)} = X$.
2. Für $m = 1, ..., M$ wähle $\mathcal{Y}^{(m)}$ so, dass
$$\begin{aligned} \mathcal{Y}^{(1)} &\subseteq K^{(1)} \\ \mathcal{Y}^{(1)} \cup \mathcal{Y}^{(2)} &\subseteq K^{(2)} \\ &\vdots \\ \bigcup_{m=1}^{M} \mathcal{Y}^{(m)} &\subseteq K^{(M)} = X \end{aligned}$$
Basen sind.
3. Konstruiere Mengen $\mathcal{W}^{(m)}$ so, dass
$$\Phi^\ell \mathcal{W}^{(M)} \cup \Phi^{\ell-1}\mathcal{W}^{(M-1)} \cup ... \cup \mathcal{W}^{(M-\ell)} \subseteq \mathrm{span}(\mathcal{Y}^{(M-\ell)}), \quad \ell \in \{0, 1, ..., M-1\}$$
Basen sind. Man beginnt mit der Konstruktion von $\mathcal{W}^{(M)}$ durch Setzen von $\mathcal{W}^{(M)} = \mathcal{Y}^{(M)}$ und fährt sukzessive mit der Konstruktion von $\mathcal{W}^{(M-1)}$, $\mathcal{W}^{(M-2)}$, ..., $\mathcal{W}^{(1)}$ fort.
4. Die in $\mathcal{W}^{(1)}$, $\mathcal{W}^{(2)}$, $\Phi\mathcal{W}^{(2)}$, ..., $\mathcal{W}^{(M)}$, $\Phi\mathcal{W}^{(M)}$, ..., $\Phi^{M-1}\mathcal{W}^{(M)}$ enthaltenenen Vektoren werden von rechts nach links in die Basistransformationsmatrix eingetragen. (Ganz rechts stehen also die Vektoren in $\Phi^{M-1}\mathcal{W}^{(M)}$, dann kommen die Vekoren in $\Phi^{M-2}\mathcal{W}^{(M)}$ usw. bis man ganz links die verbleibenden Vektoren in $\mathcal{W}^{(1)}$ einträgt. Dabei werden linear abhängige Vektoren natürlich *nicht* eingetragen.

## 5.1 Aufgaben

**Aufgabe 5.1.** Bestimmen Sie die Jordansche Normalformen der komplexwertigen Matrizen

$$A = \begin{pmatrix} 0 & 1 & 2 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & -1 & 1 & -2 \\ 1 & -1 & -1 & 1 & -1 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 & 1 & 1 & -2 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & -1 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}, \qquad C = \begin{pmatrix} 2 & 2 & 3 & -1 & 0 & -1 \\ 0 & 3 & 1 & 0 & 0 & -1 \\ 1 & 0 & 2 & 1 & -2 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ -1 & 1 & 2 & -1 & 4 & -1 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{pmatrix}.$$

**Aufgabe 5.2.**

## 5.2 Lösungen

**Lösung 5.1.** 1. Wir rechnen zunächst die charakteristischen Polynome und Minimalpolynome aus und erhalten

$$\begin{aligned} \chi_A(x) = \chi_B(x) &= -x^5\,, \quad \chi_C(x) = (3-x)^4(2-x)^2 \\ \mu_A(x) = \mu_B(x) &= x^3\,, \quad \mu_C(x) = (x-3)^3(x-2)^2\,. \end{aligned}$$

Offenbar sind $A$ und $B$ nilpotent, d.h. wir können sofort mit der "Jordanisierung" beginnen. Bei $C$ müssen wir zunächst die Hauptraumzerlegung durchführen.

2. Zur Hauptraumzerlegung von $C$ bestimmen wir

$$\begin{aligned} \ker(C-3) &= \operatorname{span}(\{(2,1,0,0,1,0)^T, (-1,0,0,1,0,0)^T\}) \\ \ker((C-3)^2) &= \operatorname{span}(\{(2,1,0,0,1,0)^T, (-1,0,0,1,0,0)^T, (1,0,1,0,0,0)^T\}) \\ \ker((C-3)^3) &= \operatorname{span}(\{(1,0,0,0,1,0)^T, (-1,0,0,1,0,0)^T, (1,0,1,0,0,0)^T, (1,1,0,0,0,0)^T\}) \\ \ker((C-3)^4) &= \operatorname{span}(\{(1,0,0,0,1,0)^T, (-1,0,0,1,0,0)^T, (1,0,1,0,0,0)^T, (1,1,0,0,0,0)^T\}) \\ &= \ker((C-3)^3) \end{aligned}$$

und

$$\begin{aligned} \ker(C-2) &= (\{(2,0,0,0,1,0)^T\}) \\ \ker((C-2)^2) &= (\{(0,0,1,0,0,1)^T, (2,0,0,0,1,0)^T\})\,. \end{aligned}$$

Schreiben wir die Vektoren aus $\ker((C-3)^4)$ und $\ker((C-2)^2)$ in die Matrix

$$S = \begin{pmatrix} 1 & -1 & 1 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix},$$

so erhalten wir

$$S^{-1}CS = \begin{pmatrix} 3 & 0 & 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 \\ -1 & 0 & 3 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{pmatrix},$$

wobei der obere $4 \times 4$-Block als $3\mathbf{1}_4 + N_1$ mit $N_1 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$ und der untere $2 \times 2$-Block als $2\mathbf{1}_2 + N_2$ mit $N_2 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ geschrieben werden kann.

Wir haben $\mu_{N_1}(x) = x^3$ und $\mu_{N_2}(x) = x^2$. In der Notation der Vorlesung ist der Block $N_2$ bereits in der gewünschten Jordan-Normalform $N_2 = J_2$; in diesem Fall ist die zugehörige Transformationsmatrix zur Jordanisierung also einfach gleich $\mathbf{1}_2$. Wenn wir mit $\tilde{J}_k$-Matrizen arbeiten wollten, könnte $N_2$ mit der Ähnlichkeitstransformation

$$\tilde{T}_{N_2} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \tilde{T}_{N_2}^{-1}$$

in $\tilde{T}_{N_2}^{-1} N_2 \tilde{T}_{N_2} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \tilde{J}_2$ übergeführt werden.

Es verbleibt, $N_1$ zu Jordanisieren. Dazu bestimmen wir

$$\begin{aligned} \ker(N_1) &= \operatorname{span}(\{(1,0,0,1)^T, (0,1,0,0)^T\}), \\ \ker(N_1^2) &= \operatorname{span}(\{(1,0,0,1)^T, (0,1,0,0)^T, (0,0,1,0)^T\}) \,. \end{aligned}$$

Damit können wir (in der Notation der Vorlesung)

$$\begin{aligned} \mathcal{Y}_{N_1}^{(1)} &= \{(1,0,0,1)^T, (0,1,0,0)^T\}, \\ \mathcal{Y}_{N_1}^{(2)} &= \{(0,0,1,0)^T\}, \\ \mathcal{Y}_{N_1}^{(3)} &= \{(0,0,0,1)^T\} \end{aligned}$$

wählen, sodass $\mathcal{Y}_{N_1}^{(1)} \cup \mathcal{Y}_{N_1}^{(2)} \cup \mathcal{Y}_{N_1}^{(3)}$ eine Basis von $\mathbb{R}^4$ bildet. Wir wählen nun $\mathcal{W}_{N_1}^{(3)} = \mathcal{Y}_{N_1}^{(3)}$ und rechnen

$$\begin{aligned} N_1 \operatorname{span}(\mathcal{W}_{N_1}^{(3)}) &= \operatorname{span}(\{(0,0,1,0)^T\}), \\ N_1^2 \operatorname{span}(\mathcal{W}_{N_1}^{(3)}) &= \operatorname{span}(\{(1,0,0,1)^T\}), \end{aligned}$$

d.h. wir können $\mathcal{W}_{N_1}^{(2)} = \{0\}$ und $\mathcal{W}_{N_1}^{(1)} = \{(0,1,0,0)^T\}$ wählen. Schreiben wir die Vektoren $(\mathcal{W}_{N_1}^{(1)}, \mathcal{W}_{N_1}^{(3)}, N_1\mathcal{W}_{N_1}^{(3)}, N_1^2\mathcal{W}_{N_1}^{(3)})$ in die Matrix

$$T_{N_1} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix},$$

so erhalten wir

$$T_{N_1}^{-1} N_1 T_{N_1} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} = 0_1 \oplus J_3.$$

Schreiben wir die Ähnlichkeitstransformationen für die $N_1$ und $N_2$ in die Matrix

$$T_N = \operatorname{diag}(T_{N_1}, T_{N_2}) = T_{N_1} \oplus T_{N_2} = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

so erhalten wir die Jordan-Normalform

$$T_N^{-1} S^{-1} C S T_N = \begin{pmatrix} 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 1 & 3 & 0 & 0 & 0 \\ 0 & 0 & 1 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{pmatrix} = (3\mathbf{1}_1) \oplus (3\mathbf{1}_3 + J_3) \oplus \left(2\mathbf{1}_2 + \tilde{J}_2\right).$$

Alternativ: Schreiben wir die Vektoren $(N_1^2\mathcal{W}_{N_1}^{(3)}, N_1\mathcal{W}_{N_1}^{(3)}, \mathcal{W}_{N_1}^{(3)}, \mathcal{W}_{N_1}^{(1)})$ in die Matrix

$$\tilde{T}_{N_1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix},$$

so erhalten wir

$$\tilde{T}_{N_1}^{-1} N_1 \tilde{T}_{N_1} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \tilde{J}_3 \oplus 0_1.$$

Schreiben wir die Ähnlichkeitstransformationen für die $N_1$ und $N_2$ in die Matrix

$$\tilde{T}_N = \operatorname{diag}(\tilde{T}_{N_1}, \tilde{T}_{N_2}) = \tilde{T}_{N_1} \oplus \tilde{T}_{N_2} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix},$$

so erhalten wir die Jordan-Normalform

$$\tilde{T}_N^{-1} S^{-1} C S \tilde{T}_N = \begin{pmatrix} 3 & 1 & 0 & 0 & 0 & 0 \\ 0 & 3 & 1 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{pmatrix} = \left(3\mathbf{1}_3 + \tilde{J}_3\right) \oplus (3\mathbf{1}_1) \oplus \left(2\mathbf{1}_2 + \tilde{J}_2\right).$$

3. Wir Jordanisieren jetzt die Matrix $A$. Wir beginnen mit der Bestimmung der Kerne von Potenzen von $A$ und erhalten

$$\begin{aligned} \ker(A) &= \operatorname{span}(\{(2,1,0,0,1)^T, (-1,0,0,1,0)^T\}) \\ \ker(A^2) &= \operatorname{span}(\{(2,0,0,0,1)^T, (-1,0,0,1,0)^T, (1,0,1,0,0)^T, (0,1,0,0,0)^T\}) \\ &= \ker(A) \oplus \operatorname{span}(\{(1,0,1,0,0)^T, (0,1,0,0,0)^T\}). \end{aligned}$$

Somit sind (in der Notation der Vorlesung)

$$\begin{aligned} \mathcal{Y}_A^{(1)} &= \{(2,1,0,0,1)^T, (-1,0,0,1,0)^T\}, \\ \mathcal{Y}_A^{(2)} &= \{(1,0,1,0,0)^T, (0,1,0,0,0)^T\}, \\ \mathcal{Y}_A^{(3)} &= \{(0,0,0,1,0)^T\} \end{aligned}$$

und $\mathcal{Y}_A^{(1)} \cup \mathcal{Y}_A^{(2)} \cup \mathcal{Y}_A^{(3)}$ bilden eine Basis von $\mathbb{R}^5$. Wir wählen nun $\mathcal{W}_A^{(3)} = \mathcal{Y}_A^{(3)}$ und rechnen

$$\begin{aligned} A \operatorname{span}(\mathcal{W}_A^{(3)}) &= \operatorname{span}(\{(0,0,1,1,0)^T\}), \\ A^2 \operatorname{span}(\mathcal{W}_A^{(3)}) &= \operatorname{span}(\{(2,1,0,0,1)^T\}). \end{aligned}$$

Da $\mathcal{Y}_A^{(2)}$ nicht durch Linearkombinationen von Vektoren in $A\mathcal{W}_A^{(3)}$ aufgespannt wird, wählen wir $\mathcal{W}_A^{(2)} = \mathcal{Y}_A^{(2)}$. Wir haben

$$A \operatorname{span}(\mathcal{W}_A^{(2)}) = \operatorname{span}(\{(2,1,0,0,1)^T, (1,0,0,-1,0)^T\}) = \operatorname{span}\{(1,0,0,-1,0)^T\} \cup A^2 \operatorname{span}(\mathcal{W}_A^{(3)}).$$

Die Vektoren von $\mathcal{Y}_A^{(1)}$, sprich $(2,1,0,0,1)^T, (-1,0,0,1,0)^T$ liegen in $A \operatorname{span}(\mathcal{W}_A^{(2)})$, weshalb wir $\mathcal{W}_A^{(1)} = \{0\}$ wählen können.

Sammeln wir die Basisvektoren $(\mathcal{W}_A^{(2)}, A\mathcal{W}_A^{(2)}, \mathcal{W}_A^{(3)}, A\mathcal{W}_A^{(3)}, A^2\mathcal{W}_A^{(3)})$ (wir füllen die Matrix dabei von rechts, beginnend mit Vektoren in $A^2\mathcal{W}_A^{(3)}$, nach links, endend mit Vektoren in $A\mathcal{W}_A^{(2)}$ auf; beobachte, dass $(1,0,1,0,0)^T \in \mathcal{W}_A^{(2)}$ als Linearkombination der Vektoren in $(1,0,0,-1,0)^T \in A\mathcal{W}_A^{(2)}$ und $(0,0,1,1,0)^T \in A\mathcal{W}_A^{(3)}$ geschrieben werden kann und somit nicht in die Matrix eingetragen wird!) in

$$T_A = \begin{pmatrix} 0 & 1 & 0 & 0 & 2 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

so erhalten wir

$$T_A^{-1} A T_A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} = J_2 \oplus J_3.$$

<u>Alternativ:</u> Sammeln wir die Basisvektoren $(A^2\mathcal{W}_A^{(3)}, A\mathcal{W}_A^{(3)}, \mathcal{W}_A^{(3)}, A\mathcal{W}_A^{(2)}, \mathcal{W}_A^{(2)})$ (jetzt füllen wir von links nach rechts auf) in

$$\tilde{T}_A = \begin{pmatrix} 2 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & -1 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix},$$

so erhalten wir

$$\tilde{T}_A^{-1} A \tilde{T}_A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = \tilde{J}_3 \oplus \tilde{J}_2.$$

4. Schließlich Jordanisieren wir $B$. Wir rechnen

$$\begin{aligned} \ker(B) &= \operatorname{span}(\{(2,0,0,0,1)^T, (-1,0,0,1,0)^T, (0,1,0,0,0)^T\}) \\ \ker(B^2) &= \operatorname{span}(\{(2,0,0,0,1)^T, (-1,0,0,1,0)^T, (0,1,0,0,0)^T, (1,0,1,0,0)^T\}) \\ &= \ker(B) \oplus \operatorname{span}(\{(1,0,1,0,0)^T\}) \end{aligned}$$

Somit sind (in der Notation der Vorlesung)

$$\begin{aligned} \mathcal{Y}_B^{(1)} &= \{(2,0,0,0,1)^T, (-1,0,0,1,0)^T, (0,1,0,0,0)^T\} \\ \mathcal{Y}_B^{(2)} &= \{(1,0,1,0,0)^T\} \\ \mathcal{Y}_B^{(3)} &= \{(0,0,0,1,0)^T\} \end{aligned}$$

und $\mathcal{Y}_B^{(1)} \cup \mathcal{Y}_B^{(2)} \cup \mathcal{Y}_B^{(3)}$ bilden eine Basis von $\mathbb{R}^5$. Wir wählen nun $\mathcal{W}_B^{(3)} = \mathcal{Y}_B^{(3)}$ und rechnen

$$\begin{aligned} B \operatorname{span}(\mathcal{W}_B^{(3)}) &= \operatorname{span}(\{(1,0,1,0,0)^T\}) \\ B^2 \operatorname{span}(\mathcal{W}_B^{(3)}) &= \operatorname{span}(\{(2,1,0,0,1)^T\}) \end{aligned}$$

Da $\mathcal{Y}_B^{(2)} = B \operatorname{span} \mathcal{W}_B^{(3)}$, wählen wir $\mathcal{W}_B^{(2)} = \{0\}$. Wir wählen nun noch $\mathcal{W}_B^{(1)} \subseteq \mathcal{Y}_B^{(1)}$ so, dass $B^2\mathcal{W}_B^{(3)} \cup B\mathcal{W}_B^{(2)} \cup \mathcal{W}_B^{(1)} = B^2\mathcal{W}_B^{(3)} \cup \mathcal{W}_B^{(1)}$ eine Basis von $\operatorname{span}(\mathcal{Y}_B^{(1)})$ bildet. Eine mögliche Wahl ist $\mathcal{W}_B^{(1)} = \{(2,0,0,0,1)^T, (-1,0,0,1,0)^T\}$.

Sammeln wir die Basisvektoren $(\mathcal{W}_B^{(1)}, \mathcal{W}_B^{(3)}, B\mathcal{W}_B^{(3)}, B^2\mathcal{W}_B^{(3)})$ (wir füllen die Matrix dabei wieder von rechts nach links auf; beobachte, dass die Reihenfolge, in der wir bspw. Vektoren in $\mathcal{W}_B^{(1)}$ eintragen keine Rolle spielt) in

$$T_B = \begin{pmatrix} -1 & 2 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix},$$

so erhalten wir

$$T_B^{-1} B T_B = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} = 0_2 \oplus J_3$$

<u>Alternativ:</u> Sammeln wir die Basisvektoren $(B^2\mathcal{W}_B^{(3)}, B\mathcal{W}_B^{(3)}, \mathcal{W}_B^{(3)}, \mathcal{W}_B^{(1)})$ (jetzt füllen wir von links nach rechts auf) in

$$\tilde{T}_B = \begin{pmatrix} 2 & 1 & 0 & 2 & -1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{pmatrix},$$

so erhalten wir

$$\tilde{T}_B^{-1} B \tilde{T}_B = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = \tilde{J}_3 \oplus 0_2$$

# 6 6. Tutoriumsblatt (02. Juni 2022)

Die Jordansche Normalform ist sehr sensitiv gegenüber Störungen, wenn die zugrundeliegende Matrix mehrfache Eigenwerte hat oder ihre Eigenwerte nahe beinander liegen. Betrachten wir beispielsweise

$$A_\varepsilon = \begin{pmatrix} 1 & 1 \\ \varepsilon & 1 \end{pmatrix},$$

so ist $A_0 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ bereits in Jordanscher Normalform. Jedoch ist die Jordan-Normalform von $A_\varepsilon$ für $\varepsilon \neq 0$ durch $\begin{pmatrix} 1+\sqrt{\varepsilon} & 0 \\ 0 & 1-\sqrt{\varepsilon} \end{pmatrix}$ gegeben.

Wir geben nun ein Verfahren an, welches numerisch stabil durch den **QR-Algorithmus** (welcher nicht die Nullstellen des charakteristischen Polynoms im Voraus braucht) umgesetzt werden kann. Wir halten uns an [Fis05, Abschnitt 4.4.5]. Für ein Beispiel dieses Verfahrens blicke man zurück auf Beispiel 4.23.

## 6.1 Algorithmus zur Trigonalisierung von Matrizen über algebraisch abgeschlossenen Körpern

Sei $A \in K^{n\times n}$ mit

$$\chi_A(x) = \prod_{j=1}^{n} (\lambda_j - x)$$

mit nicht notwendigerweise verschiedenen Eigenwerten $\lambda_j \in K$. Wir suchen eine Matrix $S \in GL(n,K)$, sodass

$$D = S^{-1}AS$$

eine obere Dreiecksmatrix ist. Der Beweis des Trigonalisierungssatzes 4.17 gibt dann folgendes Verfahren.

*Schritt 1:* Betrachte $W_1 = K^n$ mit Basis $B_1 = \mathcal{K}$ und Endomorphismus $A_1 = A$. Zum Eigenwert $\lambda_1$ berechnet man einen Eigenvektor $v_1 \in K^n$. Nach dem Austauschlemma bestimmt man ein $j_1 \in \{1, ..., n\}$ so, dass

$$B_2 = (v_1, e_1, e_2, ..., \hat{e}_{j_1}, ..., e_n)$$

wieder eine Basis ist, wobei die Notation $\hat{e}_{j_1}$ meint, dass der Vektor $e_{j_1}$ ausgelassen wird. Wir betrachten dann die Transformationsmatrix

$$S_1 = T_{B_1}^{B_2}$$

mit der Basis $B_2$ als Spaltenvektoren. Dann ist

$$A_2 = S_1^{-1} A_1 S_1 = \left(\begin{array}{c|ccc} \lambda_1 & * & \cdots & * \\ \hline 0 & & & \\ \vdots & & A_2' & \\ 0 & & & \end{array}\right).$$

*Schritt 2:* Wir betrachten $W_2$ mit der Basis $B_2' = (e_1, ..., \hat{e}_{j_1}, ..., e_n)$ und dem Endomorphismus $A_2'$ aus dem vorigen Schritt. Wir haben

$$\chi_{A_2'}(x) = \prod_{j=2}^{n} (\lambda_2 - x).$$

Wir berechnen einen zu $\lambda_2$ gehörenden Eigenwert $v_2 \in W_2$ und wählen $j_2 \neq j_1$ so, dass

$$B_3' = (v_2, e_1, ..., \hat{e}_{j_1}, ..., \hat{e}_{j_2}, ..., e_n)$$

eine Basis von $W_2$ ist. Somit ist

$$B_3 = (v_1, v_2, e_1, ..., \hat{e}_{j_1}, ..., \hat{e}_{j_2}, ..., e_n)$$

eine Basis von $K^n$. Mit der Transformationsmatrix

$$S_2 = T_{B_1}^{B_3}$$

mit der Basis $B_3$ als Spaltenvektoren ist

$$A_3 = S_2^{-1} A_1 S_2 = \left(\begin{array}{cc|ccc} \lambda_1 & * & \cdots & \cdots & * \\ 0 & \lambda_2 & * & \cdots & * \\ \hline \vdots & 0 & & & \\ \vdots & \vdots & & A_3' & \\ 0 & 0 & & & \end{array}\right).$$

Bei der Berechnung von $S_2$ kann man aussnutzen, dass

$$S_2 = T_{B_1}^{B_3} = T_{B_1}^{B_2} T_{B_2}^{B_3} \quad \text{mit } T_{B_2}^{B_3} = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & T_{B_2'}^{B_3'} & \\ 0 & & & \end{array}\right).$$

Spätestens im $(n-1)$-ten Schritt erhält man eine obere Dreiecksmatrix $A_n$, denn $A_n'$ ist eine $1 \times 1$-Matrix. Also ist

$$D = A_n = S_{n-1}^{-1} A_1 S_{n-1}$$

eine obere Dreiecksmatrix.

*Bemerkung* 6.1. Die im ersten Schritt berechnete erste Zeile in $A_2$ kann sich in jedem der folgenden Schritt verändern, analog für die anderen Zeilen der Matrizen $A_j$ $(j = 3, ..., n-1)$.

## 6.2 Aufgaben

**Aufgabe 6.1.** Seien $K = \mathbb{R}$, $n = 3$ und $A = \begin{pmatrix} 3 & 4 & 3 \\ -1 & 0 & -1 \\ 1 & 2 & 3 \end{pmatrix}$. Bestimmen Sie $D$ mit dem Trigonalisierungsverfahren des vorigen Abschnitts.

## 6.3 Lösungen

**Lösung 6.1.** Es ist $\chi_A(x) = (2-x)^3$, also $\lambda_1 = \lambda_2 = \lambda_3 = 2$. Da $\operatorname{rang}(A-2) = 2$, ist $\dim\ker(A-2) = 1$, also ist $A$ nicht diagonalisierbar. Die Iteration des Verfahrens läuft wie folgt.

*Schritt 1:* Sei $v_1 = (1, -1, 1)$, $j_1 = 1$ und $B_2 = (v_1, e_2, e_3)$. Dann ist

$$S_1 = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

und somit

$$A_2 = S_1^{-1} A S_1 = \left(\begin{array}{c|cc} 2 & 4 & 3 \\ \hline 0 & 4 & 2 \\ 0 & -2 & 0 \end{array}\right).$$

*Schritt 2:* Es ist $W_2 = \operatorname{span}(e_2, e_3)$. Wähle nun $v_2 = e_2 - e_3$ und $j_2 = 2$. Stecken wir diese Vektoren in

$$S_2 = (v_1, v_2, e_3) \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix},$$

so erhalten wir

$$A_3 = S_2^{-1} A S_2 = \left(\begin{array}{cc|c} 2 & 1 & 3 \\ 0 & 2 & 2 \\ \hline 0 & 0 & 2 \end{array}\right).$$

# 7 7. Tutoriumsblatt (16. Juni 2022)

## 7.1 Min-Max-Prinzip

Das Variations- oder Min-Max-Prinzip von **Courant–Fischer oder Rayleigh–Ritz** charakterisiert die Eigenwerte selbstadjungierter Matrizen und stellt eine mächtige Methode zur analytischen und numerischen Abschätzung der Eigenwerte dar.

**Satz 7.1.** *Sei $A \in \mathbb{C}^{n\times n}$ selbstadjungiert mit Eigenwerten $\lambda_1 \leq \lambda_2 \leq ... \leq \lambda_n$. Dann gelten*

$$\lambda_k = \min_{\substack{U\subseteq\mathbb{C}^n\\ \dim U=k}} \max_{\substack{\varphi\in U\\ \|\varphi\|=1}} \langle\varphi, A\varphi\rangle, \tag{7.1}$$

$$= \max_{\substack{U\subseteq\mathbb{C}^n\\ \dim U=n-k+1}} \min_{\substack{\varphi\in U\\ \|\varphi\|=1}} \langle\varphi, A\varphi\rangle, \tag{7.2}$$

$$= \max_{\psi_1,...,\psi_{k-1}} \min_{\varphi\in[\psi_1,...,\psi_{k-1}]^\perp,\, \|\varphi\|=1} \langle\varphi, A\varphi\rangle, \tag{7.3}$$

*wobei $[\psi_1, ..., \psi_\ell]^\perp := \{\varphi : \langle\psi_j, \varphi\rangle = 0,\ j = 1, ..., \ell\}$. Insbesondere sind*

$$\lambda_1 = \min_{\|\varphi\|=1} \langle\varphi, A\varphi\rangle \quad \textit{und} \quad \lambda_n = \max_{\|\varphi\|=1} \langle\varphi, A\varphi\rangle. \tag{7.4}$$

*Bemerkung* 7.2. Die Bedingung, dass $A$ selbstadjungiert ist, ist notwendig, wie folgendes Beispiel zeigt. Sei $N = \begin{pmatrix} 0 & 1\\ 0 & 0\end{pmatrix}$. Dann wissen wir, dass die Eigenwerte $\lambda_1 = \lambda_2 = 0$ sind. Jedoch ist $\max \frac{\langle\varphi,N\varphi\rangle}{\|\varphi\|^2} = \frac{1}{2}$. (Dies ist aus $N(a, b)^T = (b, 0)^T$ und $\max \frac{ab}{|a|^2+|b|^2} = \frac{1}{2}$ ersichtlich.)

*Beweis von Satz 7.1.* Wir definieren den Rayleigh-Quotient durch

$$R_A(v) = \frac{\langle v, Av\rangle}{\|v\|^2}, \quad v \in \mathbb{C}^n \setminus \{0\}.$$

Seien $u_1, ..., u_n$ die zu $\lambda_1, ..., \lambda_n$ gehörenden Eigenvektoren von $A$. Wir beweisen zunächst (7.1). Sei dazu $U$ ein $k$-dimensionaler Unterraum von $\mathbb{C}^n$. Dann ist $U\cap\text{span}\{u_k, ..., u_n\} \neq \emptyset$, denn sonst wäre $\dim(\text{span}(U \cup \{u_k, ..., u_n\})) = k + (n - k + 1) = n + 1$, was widersprüchlich ist. Sei also $0 \neq v \in U\cap\text{span}\{u_k, ..., u_n\}$ durch $v = \sum_{j=k}^n \alpha_j u_j \in U$ gegeben. Dann ist der Rayleigh-Quotient

$$R_A(v) = \frac{\sum_{j=k}^n |\alpha_j|^2\lambda_j}{\sum_{j=k}^n |\alpha_j|^2} \geq \lambda_k, \quad \forall v \in U.$$

Daher ist

$$\max\{R_A(x) : x \in U\} \geq \lambda_k \quad \forall U \subseteq \mathbb{C}^n \quad \text{mit } \dim(U) = k$$

und somit

$$\min_{\substack{U\subseteq\mathbb{C}^n\\ \dim(U)=k}} \max\{R_A(x) : x \in U\} \geq \lambda_k.$$

Dies zeigt "$\leq$". Andererseits gilt durch Setzen von $U = \text{span}\{u_1, ..., u_k\}$, dass

$$\min_{\substack{U\subseteq\mathbb{C}^n\\ \dim(U)=k}} \max\{R_A(x) : x \in U\} \leq \max\{R_A(x) : x \in \text{span}\{u_1, ..., u_k\}\} \leq \lambda_k,$$

was die umgekehrte Ungleichung "$\geq$" zeigt.

Wir zeigen nun (7.2). Dazu setzen wir $A' := -A$, welche die Eigenwerte $\lambda'_k$ besitzen solle. Dann gilt nach (7.1), dass

$$\lambda'_k = -\lambda_{n-k+1} = \min_{\dim U=k} \max_{\varphi\in U} R_{A'}(\varphi) = -\max_{\dim U=k} \min_{\varphi\in U} R_A(\varphi).$$

Durch Ersetzen von $k \mapsto n - k + 1$ folgt daher

$$\lambda_k = \max_{\dim U=n-k+1} \min_{\varphi\in U} R_A(\varphi).$$

Dies zeigt (7.2).

Schließlich folgt (7.3) aus (7.2) und der Beobachtung, dass $[\psi_1, ..., \psi_{k-1}]^\perp$ ein $(n - k + 1)$-dimensionaler Unterraum von $\mathbb{C}^n$ ist. $\square$

## 7.2 Aufgaben

**Aufgabe 7.1.** Bestimmen Sie die Singulärwertzerlegungen von $A = \begin{pmatrix} 2 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & -2 & 0 \end{pmatrix} \in \mathbb{C}^{3\times 3}$ und $B = \begin{pmatrix} \frac{3}{\sqrt{2}} & -\frac{3}{\sqrt{2}} & 0 \\ 0 & 0 & 0 \\ \sqrt{2} & \sqrt{2} & 0 \end{pmatrix} \in \mathbb{C}^{3\times 3}$.

Folgende Aufgabe zeigt, dass $AB$ und $BA$ dieselben von Null verschiedenen Eigenwerte haben. (Tatsächlich stimmen auch die Multiplizitäten überein, siehe bspw. Sakai [Sak71] oder Deift [Dei78]. Weiter gilt sogar

$$\lambda(AB+\lambda)^{-1} + A(BA+\lambda)^{-1}B = 1$$

im Sinne, dass, wenn $0 \neq \lambda \in \rho(-BA)$, dann ist $\lambda \in \rho(-AB)$ und $\lambda^{-1}(1 - A(BA+\lambda)^{-1}B)$ ergibt $(AB+\lambda)^{-1}$.)

**Aufgabe 7.2.** Seien $A, B \in K^{n\times n}$.

1. Überzeugen Sie sich, dass $\begin{pmatrix} \mathbf{1}_n & -A \\ 0 & \mathbf{1}_n \end{pmatrix}$ die zu $\begin{pmatrix} \mathbf{1}_n & A \\ 0 & \mathbf{1}_n \end{pmatrix}$ inverse Matrix ist.

2. Zeigen Sie

$$\begin{pmatrix} \mathbf{1}_n & A \\ 0 & \mathbf{1}_n \end{pmatrix}^{-1} \begin{pmatrix} AB & 0 \\ B & 0 \end{pmatrix} \begin{pmatrix} \mathbf{1}_n & A \\ 0 & \mathbf{1}_n \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ B & BA \end{pmatrix}.$$

3. Folgern Sie, dass $AB$ und $BA$ dieselben Eigenwerte haben. (Erinnern Sie sich an die Determinantenformel $\det_{2n} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det_n(AD - CB)$ für Blockmatrizen $A, B, C, D \in K^{n\times n}$.)

Die folgende Aufgabe zeigt, dass für zwei Matrizen $A, B \in K^{n\times n}$, die Produkte $AB$ und $BA$ zwar dieselben von Null verschiedenen Eigenwerte haben, die Singulärwerte von $AB$ und $BA$ nicht unbedingt übereinstimmen müssen.

**Aufgabe 7.3.** Seien $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ und $B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. Bestimmen Sie die Eigen- und Singulärwerte von $AB$ und $BA$.

In den folgenden Aufgaben werden Sie einige einfache Abschätzungen für die Singulärwerte von Produkten und Summen von Matrizen kennenlernen. (Mehr Details z.B. in Bhatia [Bha97].) Folgende Charakterisierung der Singulärwerte (die aus dem Variations- / Min-Max-Prinzip von **Courant–Fischer**, siehe Satz 7.1, durch Ersetzen von $A$ durch $-A^*A$ folgt), dürfen Sie als gegeben betrachten.

**Satz 7.3.** *Sei $A \in K^{n\times n}$ mit Singulärwerten $\rho_k(A)$ in nicht-aufsteigender Folge geordnet. Dann gilt*

$$\rho_k(A) = \min_{\varphi_1,\dots,\varphi_{k-1}} \max_{\psi \in [\varphi_1,\dots,\varphi_{k-1}]^\perp} \frac{\|A\psi\|}{\|\psi\|}, \quad k \in \mathbb{N}.$$

**Aufgabe 7.4.** Seien $A, B \in K^{n\times n}$ mit Singulärwerten $\rho_k(\cdot)$. Zeigen Sie

$$\max\{\rho_k(AB), \rho_k(BA)\} \leq \|B\|\rho_k(A), \quad k \in \mathbb{N}.$$

Hinweise:

- Min-Max-Prinzip (Satz 7.3).

- $\rho_k(A) = \rho_k(A^*)$ als Konsequenz der Singulärwertzerlegung.
- $\|B\| = \|B^*\|$.

Die Abschätzung der letzten Aufgabe lässt sich verallgemeinern. Man erhält folgende Ungleichung von Fan [Fan51],

$$\rho_{k+\ell+1}(AB) \leq \rho_{k+1}(A)\rho_{\ell+1}(B), \quad k,\ell \in \mathbb{N}_0.$$

Fan zeigte außerdem die Abschätzung der folgenden Aufgabe.

**Aufgabe 7.5.** Seien $A, B \in K^{n\times n}$ mit Singulärwerten $\rho_k(\cdot)$ in nicht-aufsteigender Reihenfolge angeordnet. Zeigen Sie

$$\rho_{k+\ell+1}(A+B) \leq \rho_{k+1}(A) + \rho_{\ell+1}(B), \quad k,\ell \in \mathbb{N}_0.$$

Hinweise:

- Betrachten Sie die Abbildung $Q_k(\varphi_1, ..., \varphi_k; A+B) := \max\{\|A\psi\| : \|\psi\| = 1, \psi \in [\varphi_1, ..., \varphi_k]^\perp\}$.
- Schätzen Sie $Q_k(\varphi_1, ..., \varphi_k; A+B)$ mit Hilfe von $\|(A+B)\psi\| \leq \|A\psi\| + \|B\psi\|$ geeignet ab.
- Verwenden Sie das Min-Max-Prinzip (Satz 7.3).

## 7.3 Lösungen

**Lösung 7.1.** (1) Wir bestimmen zunächst die Singulärwerte $\rho_j(A)$ (mit $j = 1,2,3$) von $A$, sprich die Wurzeln der Eigenwerte von $A^*A = \begin{pmatrix} 8 & 2 & 0 \\ 2 & 5 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

Wir rechnen $\chi_{A^*A}(x) = (8-x)(5-x)(-x) - (-x)\cdot 4 = -x^3 + 13x^2 - 36x = -x(x-9)(x-4)$ und erhalten somit $\rho_1 = 3$, $\rho_2 = 2$ und $\rho_3 = 0$. Zugehörige orthonormierte Eigenvektoren sind

$$x_1 = 5^{-1/2}\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \quad x_2 = 5^{-1/2}\begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}, \quad x_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Setzt man (gemäß Notation der Vorleung) $y_j = \rho_j(A)^{-1}Ax_j$, erhält man

$$y_1 = 3^{-1}\cdot 5^{-1/2}\begin{pmatrix} 4 \\ 5 \\ -2 \end{pmatrix}, \quad y_2 = 2^{-1}\cdot 5^{-1/2}\begin{pmatrix} -2 \\ 0 \\ -4 \end{pmatrix}, \quad y_3 = 3^{-1}\begin{pmatrix} 2 \\ -2 \\ -1 \end{pmatrix},$$

wobei wir $y_3$ durch Ergänzen von $y_1$ und $y_2$ zu einer Basis von $\mathbb{R}^3$ mit $y_3 \perp y_1$ und $y_3 \perp y_2$ gewonnen haben. In jedem Fall erhalten wir

$$A = \sum_{j=1}^{3} \rho_j(A)|y_j\rangle\langle x_j| = 3|y_1\rangle\langle x_1| + 2|y_2\rangle\langle x_2|.$$

Alternativ: wir schreiben

$$U = (y_1\ y_2\ y_3) = \begin{pmatrix} 3^{-1}\cdot 5^{-1/2}\cdot 4 & 2^{-1}\cdot 5^{-1/2}\cdot(-2) & 2/3 \\ 3^{-1}\cdot 5^{-1/2}\cdot 5 & 0 & -2/3 \\ 3^{-1}\cdot 5^{-1/2}\cdot(-2) & 2^{-1}\cdot 5^{-1/2}\cdot(-4) & -1/3 \end{pmatrix}$$

$$V = (x_1\ x_2\ x_3) = \begin{pmatrix} 2\cdot 5^{-1/2} & -5^{-1/2} & 0 \\ 5^{-1/2} & 2\cdot 5^{-1/2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

und erhalten

$$A = U\,\mathrm{diag}(3,2,0)V^*.$$

(2) Wir bestimmen zunächst die Singulärwerte $\rho_j(B)$ (mit $j = 1, 2, 3$) von $B$, sprich die Wurzeln der Eigenwerte von $B^*B = \begin{pmatrix} \frac{13}{2} & -\frac{5}{2} & 0 \\ -\frac{5}{2} & \frac{13}{2} & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

Wir rechnen $\chi_{B^*B}(x) = (8-x)(5-x)(-x) - (-x)\cdot 4 = -x^3 + 13x^2 - 36x = -x(x-9)(x-4)$ (was dasselbe Ergebnis wie zuvor ist) und erhalten somit $\rho_1(B) = 3$, $\rho_2(B) = 2$ und $\rho_3(B) = 0$. Zugehörige orthonormierte Eigenvektoren sind

$$x_1 = 2^{-1/2}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \quad x_2 = 2^{-1/2}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad x_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Setzt man (gemäß Notation der Vorleung) $y_j = \rho_j(B)^{-1}Ax_j$, erhält man

$$y_1 = 3^{-1}\cdot 2^{-1/2}\begin{pmatrix} -3\sqrt{2} \\ 0 \\ 0 \end{pmatrix} = -e_1, \quad y_2 = 2^{-1}\cdot 2^{-1/2}\begin{pmatrix} 0 \\ 0 \\ 2\sqrt{2} \end{pmatrix} = e_3, \quad y_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = e_2,$$

wobei wir $y_3$ durch Ergänzen von $y_1$ und $y_2$ zu einer Basis von $\mathbb{R}^3$ mit $y_3 \perp y_1$ und $y_3 \perp y_2$ gewonnen haben. In jedem Fall erhalten wir

$$B = \sum_{j=1}^{3} \rho_j(B)|y_j\rangle\langle x_j| = 3|y_1\rangle\langle x_1| + 2|y_2\rangle\langle x_2|.$$

**Lösung 7.2.** Folgt durch Nachrechnen. Für den dritten Punkt rechnen wir mit $S = \begin{pmatrix} \mathbf{1}_n & A \\ 0 & \mathbf{1}_n \end{pmatrix}$

$$\begin{aligned} -\lambda \det(AB - \lambda\mathbf{1}_n) &= \det\begin{pmatrix} AB - \lambda\mathbf{1}_n & 0 \\ B & -\lambda\mathbf{1}_n \end{pmatrix} = \det\left(\begin{pmatrix} AB & 0 \\ B & 0 \end{pmatrix} - \lambda\mathbf{1}_n\right) \\ &= \det\left(S\left[\begin{pmatrix} AB & 0 \\ B & 0 \end{pmatrix} - \lambda\mathbf{1}_n\right]S^{-1}\right) = \det\left(\begin{pmatrix} 0 & 0 \\ B & BA \end{pmatrix} - \lambda\mathbf{1}_n\right) \\ &= -\lambda\det(BA - \lambda\mathbf{1}_n). \end{aligned}$$

Für $\lambda \neq 0$ folgt also $\det(AB - \lambda\mathbf{1}_n) = \det(BA - \lambda\mathbf{1}_n)$ wie behauptet.

**Lösung 7.3.** Wir haben $AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ und $BA = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Da $\chi_{BA}(x) = -x^2$, sind $\sigma(AB) = \sigma(BA) = \{0\}$. Da $(BA)^*(BA) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, sind die Singulärwerte von $BA$ gerade $\rho_1(BA) = 1$ und $\rho_2(BA) = 0$.

**Lösung 7.4.** Aus dem Min-Max-Prinzip (Satz 7.3) sowie $\|BA\psi\| \le \|B\|\|A\psi\|$ folgt $\rho_k(BA) \le \|B\|\rho_k(A)$. Diese Ungleichung impliziert $\rho_k(AB) = \rho_k((AB)^*) = \rho_k(B^*A^*) \le \|B^*\|\rho_k(A^*) = \|B\|\rho_k(A)$.

**Lösung 7.5.** Sei $Q_k(\varphi_1, ..., \varphi_k; A) := \max\{\|A\psi\| : \|\psi\| = 1,\, \psi \in [\varphi_1, ..., \varphi_k]^\perp\}$. Dann folgt aus $\|(A+B)\psi\| \le \|A\psi\| + \|B\psi\|$, dass

$$\begin{aligned} Q_{k+\ell}(\varphi_1, ..., \varphi_{k+\ell}; A+B) &\le Q_{k+\ell}(\varphi_1, ..., \varphi_{k+\ell}; A) + Q_{k+\ell}(\varphi_1, ..., \varphi_{k+\ell}; B) \\ &\le Q_k(\varphi_1, ..., \varphi_k; A) + Q_\ell(\varphi_{k+1}, ..., \varphi_{k+\ell}; B). \end{aligned}$$

Die Behauptung folgt durch Minimierung über die $\varphi_1, ..., \varphi_{k+\ell}$ sowie dem Min-Max-Prinzip (Satz 7.3).

# 8 8. Tutoriumsblatt (21. Juni 2022)

## 8.1 Aufgaben

**Aufgabe 8.1.** Sei $\mathcal{P}$ der Vektorraum aller reellen Polynome $p : \mathbb{R} \to \mathbb{R}$. Für $p \in \mathcal{P}$, $p(x) = \sum_{k=0}^n a_k x^k$ setze $\|p\| := \sum_{k=0}^n |a_k|$, welche eine Norm auf $\mathcal{P}$ definiert (Hausaufgabe).

Untersuchen Sie, ob der Operator $T : \mathcal{P} \to \mathcal{P}$, definiert durch $Tp := p'$ (Ableitung, punktweise durch $(Tp)(x) = p'(x)$ definiert) beschränkt ist. Gerechnen Sie ggf. $\|T\|_{\mathrm{Op}}$.

**Aufgabe 8.2.** Sei $\mathcal{H}$ ein (nicht notwendigerweise endlichdimensionaler) Hilbertraum und $\{e_1, e_2, ...\}$ ein Orthonormalsystem in $\mathcal{H}$. Zeigen Sie für alle $x \in \mathcal{H}$ die *Besselsche Ungleichung*

$$\sum_{j\geq 1} |\langle e_j, x\rangle|^2 \leq \|x\|^2.$$

**Aufgabe 8.3.** Auf $\mathbb{C}^{n\times n}$ sei die Operatornorm durch $\|A\|_{\mathrm{op}} := \sup_{x\in\mathbb{C}^n} \frac{\|Ax\|_2}{\|x\|_2}$ gegeben. Zeigen Sie, dass $\|A\|_{\mathrm{op}} = \rho_1(A)$, wobei $\rho_1(A)$, der größte Singulärwert von $A$ ist.

**Aufgabe 8.4.** Auf $\mathbb{C}^{n\times n}$ seien die Operatornormen $\|A\|_{\mathrm{op},p} := \sup_{x\in\mathbb{C}^n} \frac{\|Ax\|_p}{\|x\|_p}$ mit $p \in \{1, 2, \infty\}$ gegeben. Zeigen Sie, dass die Operatornorm $\|\cdot\|_{\mathrm{op},p}$ nicht durch ein Skalarprodukt induziert wird.

Hinweis: Erinnern Sie sich daran, dass für alle Normen $\|\cdot\|$, die durch ein Skalarprodukt $\langle\cdot,\cdot\rangle$ induziert werden, die Parallelogramgleichung $\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2$ gilt, siehe Aufgabe K.10.4 des letzten Semesters. Kombinieren Sie dies mit der vorigen Aufgabe sowie der Hausaufgabe 8.2, welche $\|A\|_{\mathrm{op},1} = \max_{1\leq j\leq n} \sum_{i=1}^n |A_{ij}|$ (*Spaltensummennorm*) und $\|A\|_{\mathrm{op},\infty} = \max_{1\leq i\leq n} \sum_{j=1}^n |A_{ij}|$ (*Zeilensummennorm*) besagt.

## 8.2 Lösungen

**Lösung 8.1.** Wir argumentieren per Widerspruch. Angenommen $T$ wäre beschränkt, sprich es gäbe $c > 0$, sodass $\|Tp\| \leq c\|p\|$ für alle $p \in \mathcal{P}$. Sei dann $p(x) = x^n$ mit $n \in \mathbb{N}$ so, dass $n > c$. Dann sind $\|p\| = 1$ und $p'(x) = nx^{n-1}$, also $\|Tp\| = n = n\|p\| > c\|p\|$. Dies widerspricht der Annahme, dass $T$ beschränkt ist.

**Lösung 8.2.** Sei $n \in \mathbb{N}$ beliebig. Dann gilt

$$\begin{aligned} 0 \leq \|x - \sum_{j=1}^n \langle e_j, x\rangle e_j\|^2 &= \|x\|^2 - 2\sum_{j=1}^n \mathrm{Re}\langle\langle e_j, x\rangle e_j, x\rangle + \sum_{j=1}^n |\langle e_j, x\rangle|^2 \\ &= \|x\|^2 - 2\sum_{j=1}^n |\langle e_j, x\rangle|^2 + \sum_{j=1}^n |\langle e_j, x\rangle|^2 = \|x\|^2 - \sum_{j=1}^n |\langle e_j, x\rangle|^2. \end{aligned}$$

Da die Ungleichung für beliebiges $n \in \mathbb{N}$ gilt, folgt die Behauptung.

**Lösung 8.3.** Sei

$$A = \sum_{j=1}^n \rho_j(A)|y_j\rangle\langle x_j|$$

die Singulärwertzerlegung von $A$, wobei $\{x_j\}_{j=1}^n$ und $\{y_j\}_{j=1}^n$ Orthonormalsysteme in $\mathbb{C}^n$ sind. Dann folgt mit der Besselungleichung

$$\begin{aligned} \|A\psi\|_2^2 &= \sum_{j,\ell} \rho_j(A)\rho_\ell(A)\langle x_j, \psi\rangle\overline{\langle x_\ell, \psi\rangle}\langle y_\ell, y_j\rangle = \sum_j \rho_j(A)^2 \sum_j |\langle x_j, \psi\rangle|^2 \\ &\leq \rho_1(A)^2 \sum_j |\langle x_j, \psi\rangle|^2 \leq \rho_1(A)^2\|\psi\|^2, \end{aligned}$$

also $\|A\|_{\mathrm{op}} \leq \rho_1(A)$. Andererseits gilt (mit derselben Rechnung) auch die umgekehrte Ungleichung, denn $\|A\|_{\mathrm{op}} \geq \|Ax_1\|_2 = \rho_1(A)$.

**Lösung 8.4.** Seien $A, B \in \mathbb{C}^{2\times 2}$ durch

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

gegeben. Durch die offenbaren Singulärwertzerlegungen der Matrizen $A$, $B$, $A+B$ und $A-B$ sowie Hausaufgabe 8.2 sieht man, dass $\|A\|_{\mathrm{op},p} = \|B\|_{\mathrm{op},p} = \|A+B\|_{\mathrm{op},p} = \|A-B\|_{\mathrm{op},p} = 1$ gelten. Damit ist die Parallelogramgleichung verletzt, sprich $\|\cdot\|_{\mathrm{op},p}$ wird nicht durch ein Skalarprodukt induziert.

# 9 9. Tutoriumsblatt (30. Juni 2022)

## 9.1 Aufgaben

**Aufgabe 9.1.** Seien $K \in \{\mathbb{R}, \mathbb{C}\}$ und $p \in [1, \infty)$. Dann ist

$$\ell^p := \{x = (x_j)_{j\in\mathbb{N}} \in K^{\mathbb{N}} : \sum_{j\in\mathbb{N}} |x_j|^p < \infty\}$$

ein Vektorraum und $\|x\|_p := (\sum_{j\in\mathbb{N}} |x_j|^p)^{1/p}$ definiert eine Norm auf $\ell^p$. Insbesondere ist also $(\ell^p, \|\cdot\|_p)$ ein Banachraum.

*Hinweis:* Erinnern Sie sich an die Minkowski-Ungleichung $\|x+y\|_p \leq \|x\|_p + \|y\|_p$ für $x, y \in \ell^p$ (siehe bspw. Aufgabe 11.3 des Tutoriumblatts des letzten Semesters).

Sei $X \neq \emptyset$ eine Menge.

**Definition 9.1.** Eine *Metrik auf* $X$ ist eine Abbildung $d : X \times X \to [0, \infty)$ mit den Eigenschaften

1. (Definitheit) $\forall\, x, y \in X$ gilt $d(x, y) = 0$ genau dann, wenn $x = y$
2. (Symmetrie) $\forall\, x, y \in X$ gilt $d(x, y) = d(y, x)$
3. (Dreiecksungleichung) $\forall\, x, y, z \in X$ gilt $d(x, y) \leq d(x, z) + d(z, y)$.

Wenn $d$ eine Metrik auf $X$ ist, so heißt $(X, d)$ *metrischer Raum.*

Wenn $(X, \|\cdot\|)$ ein normierter Raum ist, so ist $(X, d)$ mit $d(x, y) := \|x - y\|$ offenbar ein metrischer Raum.

**Definition 9.2.** 1. Eine Teilmenge $Q \subseteq X$ heißt genau dann *offen*, wenn es für alle $x_0 \in Q$ einen Radius $r > 0$ gibt, sodass $B(x_0, r) = \{x \in X : d(x, x_0) < r\} \subseteq Q$.

2. Eine Teilmenge $A \subseteq X$ heißt genau dann *abgeschlossen*, wenn $X \setminus A$ offen ist.

*Bemerkung* 9.3. 1. Beliebige Vereinigungen oder endliche Schnitte offener Mengen sind wieder offen.

2. Beliebige Schnitte oder endliche Vereinigungen abgeschlossener Mengen sind wieder abgeschlossen.

**Definition 9.4.** Sei $(M, d)$ ein metrischer Raum.

1. Eine Teilmenge $U \subseteq X$ heißt genau dann *Umgebung eines Punkts* $x_0 \in X$, wenn es $r > 0$ gibt, sodass $B(x_0, r) \subseteq U$.
2. Sei $M \subseteq X$. Dass *Innere von* $M$ ist definiert durch

$$\text{int}(M) := \{x \in X : M \text{ ist eine Umgebung von } x\}.$$

3. Der *Abschluß von* $M \subseteq X$ ist definiert durch

$$\overline{M} := \{x \in X : U \cap M \neq \emptyset \text{ für alle Umgebungen } U \text{ von } x\}.$$

4. Der *Rand von* $M \subseteq X$ ist definiert durch

$$\partial M = \{x \in X : U \cap M \neq \emptyset \text{ and } U \cap (X \setminus M) \neq \emptyset \text{ für jede Umgebung } U \text{ von } x\}.$$

5. Ein Punkt $x \in M$ heißt *innerer Punkt von* $M$, wenn $x \in \text{int}(M)$ und *Randpunkt von* $M$, wenn $x \in \partial M$.

*Bemerkung* 9.5. Seien $(X, d)$ ein metrischer Raum und $M, M_1, M_2 \subseteq X$.

1. $M$ ist genau dann offen, wenn $M = \operatorname{int}(M)$.
2. $M$ ist genau dann abgeschlossen, wenn $M = \overline{M}$. (In Satz 9.8 zeigen wir das später noch auf andere Art, ausgehend von der Definition, dass $\overline{M}$ die Menge aller Kontaktpunkte[1] von $M$ ist.)
3. Es ist $M \subseteq \overline{M}$. Weiter folgt aus $M_1 \subseteq M_2$, dass $\overline{M_1} \subseteq \overline{M_2}$.
4. $\partial M = \overline{M} \cap \overline{X \setminus M}$. Insbesondere ist $\partial M$ abgeschlossen.
5. Es ist $\overline{M} = \operatorname{int}(M) \dot\cup \partial M$ und $X = \operatorname{int}(M) \dot\cup \partial M \dot\cup \operatorname{int}(X \setminus M)$. (Hierbei bezeichnet $A \dot\cup B$ die disjunkte Vereinigung zweier Mengen $A$ und $B$.)

**Behauptung 9.6.** *Seien $M \subseteq X$ und $x \in X$. Dann ist $x \in \overline{M}$ genau dann, wenn es eine Folge $(x_n)_{n\in\mathbb{N}} \subseteq M$ gibt, sodass $x_n \to x$.*

*Beweis.* "$\Rightarrow$": Sei $x \in \overline{M}$. Für jedes $n$ finden wir $x_n \in B(x, 1/n) \cap M$. Dann ist $(x_n)_{n\in\mathbb{N}} \subseteq M$ eine Folge mit $d(x_n, x) \leq 1/n \to 0$.

"$\Leftarrow$": Sei $x \in X$ durch die Folge $(x_n)_{n\in\mathbb{N}} \subseteq M$ approximiert und $r > 0$ ein beliebiger Radius. Dann gibt es $n \in \mathbb{N}$, sodass $d(x_n, x) < r$. Insbesondere ist $x_n \in B(x, r) \cap M$. Das zeigt $x \in \overline{M}$. □

**Behauptung 9.7.** *Sei $M \subseteq X$. Dann gilt $\overline{\overline{M}} = \overline{M}$.*

*Beweis.* $\overline{M} \subseteq \overline{\overline{M}}$ ist klar, wir zeigen also die umgekehrte Inklusion. Seien $f \in \overline{\overline{M}}$ und $\varepsilon > 0$. Dann gibt es $g \in \overline{M}$, sodass $\|f - g\| < \varepsilon$. Außerdem gibt es $h \in M$, sodass $\|g - h\| < \varepsilon$. Insgesamt gilt also $\|f - h\| \leq 2\varepsilon$. Das bedeutet aber $f \in \overline{M}$, sprich $\overline{\overline{M}} \subseteq \overline{M}$ wie behauptet. □

**Satz 9.8.** *Sei $(X, d)$ ein metrischer Raum und $M \subseteq X$. Dann ist $M$ genau dann abgeschlossen, wenn $M = \overline{M}$.*

*Beweis.* Angenommen $M = \overline{M}$. Wir zeigen, dass $\overline{M}$ abgeschlossen, also $X \setminus \overline{M}$ offen ist. Sei also $f \in X \setminus \overline{M}$. Da $\overline{M} = \overline{\overline{M}}$, gilt auch $f \in X \setminus \overline{\overline{M}}$, sprich $f$ ist kein Kontaktpunkt von $\overline{M}$, d.h. es gibt ein $\varepsilon > 0$, sodass für alle $g \in \overline{M}$ gilt, dass $\|f - g\| > \varepsilon$. Oder anders gesprochen, es gibt $\varepsilon > 0$, sodass $B_f(\varepsilon) \cap \overline{M} = \emptyset$. Das bedeutet aber $B_f(\varepsilon) \subseteq (X \setminus \overline{M})$, sprich $X \setminus \overline{M}$ ist offen, wie behauptet.

Sei nun $M$ abgeschlossen, sprich $X \setminus M$ ist offen. Wegen $M \subseteq \overline{M}$ genügt es $\overline{M} \subseteq M$ zu zeigen. Da $X \setminus M$ offen ist, gibt es für alle $f \in X \setminus M$ ein $\varepsilon > 0$, sodass $B_f(\varepsilon)$ ganz in $X \setminus M$ enthalten ist, sprich für das $B_f(\varepsilon) \cap M = \emptyset$ gilt. Das bedeutet aber, dass $f \notin \overline{M}$. Damit ist $X \setminus M \subseteq X \setminus \overline{M}$, also $\overline{M} \subseteq M$ gezeigt. Dies schließt den Beweis. □

**Lemma 9.9.** *Sei $(X, d)$ ein vollständiger metrischer Raum und $\emptyset \neq M \subseteq X$. Dann ist $(M, d_{M\times M})$ genau dann vollständig, wenn $M$ in $X$ abgeschlossen ist.*

*Beweis.* "$\Rightarrow$": Sei $M$ vollständig und $(x_n)_{n\in\mathbb{N}} \subseteq M$ eine Folge mit $x_n \to x \in X$. Dann ist $(x_n)_{n\in\mathbb{N}}$ insbesondere eine Cauchy-Folge, sprich $x_n$ konvergiert gegen ein $\tilde{x} \in M$. Das bedeutet $x = \tilde{x} \in M$. Also ist $M$ abgeschlossen.

"$\Leftarrow$": Sei $M$ abgeschlossen in $X$ und $(x_n)_{n\in\mathbb{N}} \subseteq M$ eine Cauchyfolge in $M$. Da $X$ vollständig ist, gibt es $x \in X$, sodass $x_n \to x$. Da $M$ abgeschlossen ist, folgt $x \in M$, was bedeutet, dass $(x_n)_{n\in\mathbb{N}}$ einen Grenzwert in $M$ hat, sprich $M$ ist vollständig. □

**Aufgabe 9.2.** Sei

$$\ell^\infty := \{x = (x_j)_{j\in\mathbb{N}} \in K^{\mathbb{N}} : \sup_{j\in\mathbb{N}} |x_j| < \infty\}$$

ausgestattet mit $\|x\|_\infty := \sup_{j\in\mathbb{N}} |x_j|$. Mit denselben Argumenten wie aus Aufgabe 9.1 folgt, dass $\ell^\infty$ ein Banachraum ist. Sei nun

$$c_0 := \{x = (x_j)_{j\in\mathbb{N}} \in \ell^\infty : \lim_{j\to\infty} x_j = 0\}$$

[1] $f \in X$ heißt genau dann Kontaktpunkt $M$, wenn es für alle $\varepsilon > 0$ ein $g \in M$ gibt, sodass $\|f - g\| < \varepsilon$. Äquivalent: $f \in X$ heißt genau dann Kontaktpunkt von $M$, wenn es $\varepsilon > 0$ gibt, sodass $B_f(\varepsilon) \cap M \neq \emptyset$.

ausgestattet mit $\|\cdot\|_\infty$. Zeigen Sie, dass $c_0$ ein abgeschlossener Teilraum von $\ell^\infty$ und (wegen Lemma 9.9) ein Banachraum ist.

**Aufgabe 9.3.** Seien $L^\infty(\mathbb{R}:\mathbb{C})$ der Raum aller beschränkten Funktionen $f:\mathbb{R}\to\mathbb{C}$ und $C^0(\mathbb{R}:\mathbb{C})$ der Raum aller stetigen Funktionen $f:\mathbb{R}\to\mathbb{C}$ jeweils ausgestattet mit der Supremumsnorm $\|f\|_\infty = \sup_{x\in\mathbb{R}}|f(x)|$. Zeigen Sie folgende Aussagen.

(a) Sei $(f_n)_{n\in\mathbb{N}}\subseteq L^\infty(\mathbb{R})$ eine Cauchy-Folge. Dann hat $(f_n)_{n\in\mathbb{N}}$ einen punktweisen Grenzwert $f\in L^\infty(\mathbb{R})$, sprich es gibt $f\in L^\infty$, sodass für alle $x\in\mathbb{R}$ gilt, dass $f_n(x)\to f(x)$.

(b) Sei $(f_n)_{n\in\mathbb{N}}\subseteq L^\infty(\mathbb{R})$ die obige Cauchy-Folge mit punktweisem Grenzwert $f$. Dann stimmt der punktweise Grenzwert $f$ mit dem Grenzwert in $\|\cdot\|_\infty$ überein, sprich es gilt $\|f_n-f\|_\infty\to 0$ für $n\to\infty$. Insbesondere ist $L^\infty$ vollständig.

(c) $C^0(\mathbb{R})$ ist vollständig. (Hinweis: Zeigen Sie, dass der gleichmäßige Grenzwert stetiger Funktionen ist wieder stetig: d.h., falls $(f_n)_{n\in\mathbb{N}}\subseteq C^0$ und $\|f_n-f\|_\infty\to 0$, so ist auch $f$ stetig.)

**Satz 9.10.** *Es gibt punktweise konvergente Funktionenfolgen $(f_n)_{n\in\mathbb{N}}\subseteq C([0,1]:\mathbb{C})$ deren punktweiser Grenzwert nicht stetig ist.*

*Beweis.* **1. Beispiel:** Sei $\mathcal{C}\subseteq[0,1]$ eine abgeschlossene Menge mit

$$\operatorname{int}(\mathcal{C}) = \{x\in X : \mathcal{C} \text{ ist eine Umgebung von } x\} = \emptyset,$$

wie z.B. die **Cantormenge**, siehe auch diesen **Überblick** von [DMRV06]. Definiere nun

$$f_n(x) = \max\{0, 1-n\cdot\operatorname{dist}(x,\mathcal{C})\}$$

mit $\operatorname{dist}(x,\mathcal{C}) = \inf_{y\in\mathcal{C}}|x-y|$. Wir behaupten, dass $f_n$ stetig ist. Dies würde unmittelbar aus der Stetigkeit von $x\mapsto\operatorname{dist}(x,\mathcal{C})$ folgen. Die Distanzfunktion ist aber in der Tat stetig, was aus

$$|\operatorname{dist}(x_1,\mathcal{C})-\operatorname{dist}(x_2,\mathcal{C})|\le|x_1-x_2|$$

und der Dreiecksungleichung ersichtlich ist.

Wenn $x\in\mathcal{C}$, dann ist $f_n(x)=1$ für alle $n\in\mathbb{N}$. Wenn $x\notin\mathcal{C}$ und $n$ hinreichend groß ist, so ist $f_n(x)=0$. Also konvergiert $f_n(x)\to\mathbf{1}_{\mathcal{C}}(x)$ punktweise gegen die charakteristische Funktion von $\mathcal{C}$. Diese Funktion ist jedoch an allen Punkten in $\mathcal{C}$ unstetig.

**2. Beispiel:** Nach [GO03, S. 77–78] gibt es eine Funktionenfolge $(f_n)_{n\in\mathbb{N}}\subseteq C(\mathbb{R})$, die gegen

$$f(x) = \begin{cases} q^{-1}, & x=\frac{p}{q} \text{ in gekürzter Form mit } p\in\mathbb{Z} \text{ und } q\in\mathbb{N} \\ 0, & x\in\mathbb{R}\setminus\mathbb{Q} \end{cases}$$

konvergiert. Diese Funktion ähnelt stark der Dirichlet-Funktion

$$\mathbf{1}_{\mathbb{Q}}(x)\begin{cases} 1, & x\in\mathbb{Q} \\ 0, & x\notin\mathbb{Q} \end{cases},$$

die man in der Analysis kennenlernt. Diese Funktion ist nirgends stetig:

(a) Für $x\in\mathbb{Q}$ ist $\mathbf{1}_{\mathbb{Q}}(x)=1$. Dann gibt es $\varepsilon>0$, sodass es kein $\delta(\varepsilon)>0$ gibt, sodass für $y\in\mathbb{R}$ mit $|x-y|$ gilt, dass $|\mathbf{1}_{\mathbb{Q}}(x)-\mathbf{1}_{\mathbb{Q}}(y)|<\varepsilon$. Z.B. kann man $\varepsilon=1/2$ wählen. Da $\mathbb{R}\setminus\mathbb{Q}$ dicht in $\mathbb{R}$ liegt, gibt es für alle $\delta$ ein $y\in\mathbb{R}\setminus\mathbb{Q}$, sodass $|x-y|<\delta$, aber $f(y)=0$ ist mindestens $1/2$ weg von $f(x)=1$.

(b) Für $x\in\mathbb{R}\setminus\mathbb{Q}$ ist $f(x)=0$. Dann gibt es wegen der Dichtheit von $\mathbb{Q}$ in $\mathbb{R}$ für alle $\delta$ ein $y\in\mathbb{Q}$ mit $|x-y|<\delta$, aber $|f(x)-f(y)|>1/2$.

□

**Satz 9.11.** *Es gibt Funktionenfolgen $(f_n)_{n\in\mathbb{N}}\subseteq C([0,1]:\mathbb{C})$ deren punktweise und gleichmäßige Grenzwerte nicht übereinstimmen.*

*Beweis.* Die Funktionenfolge

$$f_n : \mathbb{R} \to \mathbb{R}\,, \qquad x \mapsto f_n(x) := \frac{x^2 + nx}{n}$$

konvergiert punktweise gegen $f : \mathbb{R} \to \mathbb{R}$ definiert durch $f(x) := x$ für $x \in \mathbb{R}$. Dazu sei $x \in \mathbb{R}$ beliebig. Dann gilt

$$\lim_{n\to\infty} |f_n(x) - f(x)| = \lim_{n\to\infty} |\frac{x^2}{n} + x - x| = \lim_{n\to\infty} \frac{x^2}{n} = 0\,.$$

Also konvergiert $f_n(x) \to x$ punktweise für $n \to \infty$. Jedoch ist $f$ nicht der gleichmäßige Grenzwert, denn die rechte Seite von

$$\sup_{x\in\mathbb{R}} |f_n(x) - f(x)| = \sup_{x\in\mathbb{R}} \frac{x^2}{n} \geq \frac{n^2}{n} = n$$

konvergiert nicht gegen Null für $n \to \infty$. □

## 9.2 Lösungen

**Lösung 9.1.** Offensichtlich sind $0 \in \ell^p$ sowie für $\alpha \in K$ und $x \in \ell^p$ auch $\alpha x \in \ell^p$ mit $\|\alpha x\|_p = |\alpha|\|x\|_p$. Die Dreiecksungleichung ist gerade die Minkowski-Ungleichung.

Es verbleibt die Vollständigkeit zu zeigen. Sei also $(x^{(n)})_{n\in\mathbb{N}}$ eine Cauchy-Folge in $\ell^p$. Dann ist insbesondere die Folge der Komponenten $(x_j^{(n)})_{n\in\mathbb{N}}$ eine Cauchy-Folge in $K$ für alle $j \in \mathbb{N}$ und der Grenzwert $x_j^{(0)} := \lim_{n\to\infty} x_j^{(n)}$ existiert in $K$. Somit ist $x^{(0)} := (x_j^{(0)})_{j\in\mathbb{N}}$ ein guter Kandidat für den Grenzwert der Folge $(x^{(n)})_{n\in\mathbb{N}}$.

Wir zeigen nun, dass $x^{(0)} \in \ell^p$ und $\|x^{(0)} - x^{(n)}\|_p \to 0$. Da $(x^{(n)})_{n\in\mathbb{N}}$ Cauchy ist, gibt es für alle $\varepsilon > 0$ ein $n_0 = n_0(\varepsilon) \in \mathbb{N}$, sodass $\|x^{(n)} - x^{(m)}\|_p \leq \varepsilon$ für alle $n, m \geq n_0$. Daraus folgt für alle $n > n_0$ und *fixierten* $N \in \mathbb{N}$, dass

$$\left(\sum_{j=1}^{N} |x_j^{(n)} - x_j^{(0)}|^p\right)^{1/p} = \lim_{m\to\infty} \left(\sum_{j=1}^{N} |x_j^{(n)} - x_j^{(m)}|^p\right)^{1/p} \leq \|x^{(n)} - x^{(m)}\|_p \leq \varepsilon.$$

Lässt man $N \to \infty$ gehen, folgt $\|x^{(n)} - x^{(0)}\|_p < \infty$ für alle $n \geq n_0$. Insbesondere folgt, dass $x^{(0)} = x^{(0)} - x^{(n_0)} + x^{(n_0)} \in \ell^p$ und $\|x^{(n)} - x^{(0)}\|_p \to 0$ für $\varepsilon \to 0$, also $n \to \infty$.

**Lösung 9.2.** Sei $(x^{(n)})_{n\in\mathbb{N}} \subseteq c_0$ eine Folge mit $x^{(n)} \to x^{(0)} \in \ell^\infty$. Es genügt zu zeigen, dass $x^{(0)} \in c_0$. Sei dazu $\varepsilon > 0$. Dann gibt es $n \in \mathbb{N}$, sodass $\|x^{(n)} - x^{(0)}\|_\infty < \varepsilon$. Außerdem gibt es $j_0 \in \mathbb{N}$, sodass $|x_j^{(n)}| < \varepsilon$ für alle $j \geq j_0$, da $x^{(n)} \in c_0$. Daraus folgt für alle $j \geq j_0$, dass

$$|x_j^{(0)}| \leq |x_j^{(0)} - x_j^{(n)} + x_j^{(n)}| \leq 2\varepsilon\,,$$

sprich $x_j^{(0)} \to 0$ für $j \to \infty$, also $x^{(0)} \in c_0$.

**Lösung 9.3.**

1. Sei $(f_n)_{n\in\mathbb{N}} \subseteq L^\infty$ Cauchy. Dann ist für festes $x \in \mathbb{R}$ auch $(f_n(x))_{n\in\mathbb{N}} \subseteq \mathbb{C}$ Cauchy. Da $\mathbb{C}$ vollständig ist, ist der Grenzwert $\lim_{n\to\infty} f_n(x) =: f(x) \in \mathbb{C}$. Insbesondere ist $f \in L^\infty$. Um das zu sehen, sei $N \in \mathbb{N}$ so, dass für alle $n, m \geq N$ gilt, dass $\|f_n - f_m\|_\infty < 1$. Dann gilt $|f(x) - f_N(x)| = |\lim_n f_n(x) - f_N(x)| = \lim_n |f_n(x) - f_N(x)| \leq 1$ für alle $x$. Insbesondere gilt dann für alle $x$, dass

   $$|f(x)| \leq |f(x) - f_N(x) + f_N(x)| \leq \|f - f_N\|_\infty + \|f_N\|_\infty \leq 1 + \|f_N\|_\infty < \infty,$$

   womit $f \in L^\infty$ gezeigt ist.

2. Seien $\varepsilon > 0$ und $N = N(\varepsilon) \in \mathbb{N}$ so, dass für alle $n, m \geq N$ gilt, dass $\|f_n - f_m\|_\infty < \varepsilon$. Dann gilt für alle $n \geq N$, dass

$$|f(x) - f_n(x)| = |\lim_m f_m(x) - f_n(x)| = \lim_m |f_m(x) - f_n(x)| \leq \varepsilon$$

für alle $x$ und somit $\|f - f_n\|_\infty \to 0$ für $n \to \infty$.

3. Folgt aus dem vorigen Punkt, dem uniform limit theorem **(Wiki)** sowie Lemma 9.9. Dieses beweisen wir jetzt. Für gegebenes $\varepsilon > 0$ sei $N = N(\varepsilon) \in \mathbb{N}$ so, dass für alle $m, n \geq N$ gilt, dass $\|f_m - f_n\|_\infty < \varepsilon$. Weiter wissen wir wegen der Stetigkeit von $f_N$, dass es $\delta = \delta(N)$ gibt, sodass für alle $x_1, x_2$ mit $|x_1 - x_2| < \delta$ gilt, dass $|f_N(x_1) - f_N(x_2)| < \varepsilon$. Dann gilt für alle $x_1, x_2$ mit $|x_1 - x_2| < \delta$, dass

$$\begin{aligned}|f(x_1) - f(x_2)| &\leq |f(x_1) - f_N(x_1)| + |f_N(x_1) - f_N(x_2)| + |f_N(x_2) - f(x_2)| \\ &= \lim_{m\to\infty} |f_m(x_1) - f_N(x_1)| + |f_N(x_1) - f_N(x_2)| + \lim_{m\to\infty} |f_N(x_2) - f_m(x_2)| \\ &\leq 3\varepsilon.\end{aligned}$$

Damit ist die Stetigkeit von $f$ gezeigt.

# 10 10. Tutoriumsblatt (07. und 14. Juli 2022)

## 10.1 Hinweise zum 10. Übungsblatt (07. Juli 2022)

Bei Aufgabe 10.2 sollte man sich erinnern, dass alle Normen auf endlichdimensionalen Vektorräumen (wie dem Raum der $N \times N$-Matrizen) äquivalent sind. Z.B. ist die Standard-Operatornorm $\|A\|_{2\to 2} = \sup_{x\neq 0} \|Ax\|_2/\|x\|_2$ äquivalent zur Zeilen- oder Spaltensummennorm, die im 8. Blatt diskutiert wurde.

Nun zu Hinweisen für Aufgabe 10.1.

**Satz 10.1.** *Seien $(\mathcal{H}, \langle\cdot|\cdot\rangle)$ ein Hilbertraum sowie $M_1$ und $M_2$ orthogonale Teilräume von $\mathcal{H}$. Dann ist $M_1 \oplus M_2$ genau dann abgeschlossen, wenn $M_1$ und $M_2$ abgeschlossen sind.*

*Tipps für den Beweis.* "$\Rightarrow$": Zu zeigen ist, dass $M_1$ abgeschlossen ist, sprich $\overline{M_1} \subseteq M_1$. Dazu verwendet man die Charakterisierung aus Lemma 9.6 abgeschlossener Mengen. Wir haben $f \in \overline{M_1}$ genau dann, wenn es eine Folge $(f_n)_{n\in\mathbb{N}} \subseteq M_1$ gibt, sodass $f_n \to f$. Jetzt sollte man $M_1 \subseteq M_1 \oplus M_2$ und damit $\overline{M_1} \subseteq \overline{M_1 \oplus M_2} = M_1 \oplus M_2$ ausnutzen.

"$\Leftarrow$": Es ist zu zeigen, dass $\overline{M_1 \oplus M_2} \subseteq M_1 \oplus M_2$. Für $f \in \overline{M_1 \oplus M_2}$ gibt es eine Folge $(f_n)_{n\in\mathbb{N}} = (f_{1,n} + f_{2,n})_{n\in\mathbb{N}} \subseteq M_1 + M_2$, sodass $f_{1,n} + f_{2,n} \to f$. Man ist fertig, wenn man zeigen kann, dass $f_{j,n}$ (mit $j = 1, 2$) Grenzwerte in $M_j$ haben. Hierfür nutzt man die Vollständigkeit der $M_j$ sowie Orthogonalität ($\|f_1 + f_2\|^2 = \|f_1\|^2 + \|f_2\|^2$ für $f_j \in M_j$) aus. □

Hier noch ein Hinweis zur Wichtigkeit der Orthogonalität bei Aufgabe 10.1.(a).

**Satz 10.2.** *Sei $\mathcal{H}$ ein Hilbertraum. Dann gibt es abgeschlossene, aber nicht zueinander orthogonale $M_1, M_2 \subseteq \mathcal{H}$, sodass $M_1 \cap M_2 = \{0\}$, aber $M_1 + M_2 = M_1 \dot{+} M_2$ nicht abgeschlossen ist.*

*Beweis.* Seien $\{e_j\}_{j\in\mathbb{N}}$ und $\{f_j\}_{j\in\mathbb{N}}$ zwei zueiander orthogonale Folgen orthonormaler Vektoren in $\mathcal{H}$. Seien dann

$$M_1 = \overline{\operatorname{span}\{e_j : j \in \mathbb{N}\}} \quad \text{und} \quad M_2 = \overline{\operatorname{span}\{e_j + \frac{1}{j} f_j : j \in \mathbb{N}\}}.$$

Dann sind $M_1, M_2$ per Definition abgeschlossen und $M_1 \cap M_2 = \{0\}$. Jedoch ist $M_1$ nicht orthogonal auf $M_2$. Zudem ist die direkte Summe $M_1 \dot{+} M_2$ nicht abgeschlossen. In der Tat gilt zwar $j^{-1} f_j \in M_1 + M_2$ für alle $j \in \mathbb{N}$, doch die Folge der Partialsummen

$$\sum_{j=1}^{N} \frac{1}{j} f_j$$

hat keinen Grenzwert in $M_1 + M_2$. (Beobachte, dass $\|\sum_{j\geq 1} j^{-1} f_j\|^2 = \sum_{j\geq 1} j^{-2} < \infty$.) Wenn es doch so wäre, so könnten wir $\sum_{j\geq 1} \frac{1}{j} f_j = m_1 + m_2$ für gewisse $m_k \in M_k$ $(k = 1, 2)$ schreiben. Das würde aber bedeuten, dass notwendigerweise $m_2 = \sum_{j\geq 1}(e_j + j^{-1} f_j)$ sein muss. Da diese Reihe jedoch nicht konvergiert, kann $\sum_{j\geq 1} \frac{1}{j} f_j$ nicht in $M_1 + M_2$ sein. □

## 10.2 Resolventen (14. Juli 2022)

**Satz 10.3** (Neumann)**.** *Sei $K$ ein Körper, $(X, \|\cdot\|)$ ein $K$-Banachrraum und $S, T \in \mathcal{B}(X)$. Angenommen $T$ sei bijektiv und so, dass $\|ST^{-1}\| < 1$. Dann ist auch $S + T$ invertierbar und die Reihen*

$$\sum_{n\geq 0} (-1)^n T^{-1} (ST^{-1})^n = \sum_{n\geq 0} (-1)^n (T^{-1}S)^n T^{-1} = (S + T)^{-1}$$

*konvergieren in Operatornorm.*

Der folgende Beweis funktioniert streng genommen nur auf *endlichdimenionalen Hilberträumen*, da wir das Konzept von *abgeschlossenen Operatoren* noch nicht kennen. Wir machen auf die entsprechende Stelle im Beweis aufmerksam. Man kann die Behauptung auch aus Satz XV.12 erkennen, denn

$$(S+T)^{-1} = [(ST^{-1}+1)T]^{-1} = T^{-1}(ST^{-1}+1) = T^{-1}(1-(-ST^{-1})) = T^{-1}\sum_{n\geq 0}(-ST^{-1})^n.$$

*Beweis.* Wir zeigen zuerst, dass $S+T$ injektiv ist. Anschließend untersuchen wir die Konvergenz der Reihen. Dann zeigen wir, dass die Reihen konvergieren und das Grenzelement ein Rechts- und Linksinverses zu $S+T$ ist. Diese Tatsache zeigt die Surjektivität und damit die Gleichheit von $(S+T)^{-1}$ und den Reihen.

(1) Injektivität: Es gilt

$$\|(T+S)x\| \geq \|Tx\| - \|Sx\| = \|Tx\| - \|ST^{-1}Tx\| \geq \|Tx\|(1-\|ST^{-1}\|) > 0.$$

Somit ist $\ker(T+S) = \{0\}$, also ist $S+T$ injektiv.

(2) Konvergenz der Reihen: Für $N \in \mathbb{N}$ definiere die Partialsumme

$$A_N := \sum_{n=0}^{N}(-1)^n T^{-1}(ST^{-1})^n = \sum_{n=0}^{N}(-1)^n (T^{-1}S)^n T^{-1}.$$

Wir zeigen, dass $(A_N)_{N\in\mathbb{N}} \subseteq \mathcal{B}(X)$ eine Cauchyfolge ist. Da $\mathcal{B}(X)$ ein Banachraum ist, folgt daraus, dass $(A_N)_{N\in\mathbb{N}} \to A$ konvergiert für ein $A \in \mathcal{B}(X)$. Seien also $M, N \in \mathbb{N}$ hinreichend groß und $M > N$. Dann gilt

$$\begin{aligned}\|A_M - A_N\| &= \|\sum_{n=N+1}^{M}(-1)^n T^{-1}(ST^{-1})^n\| \\ &\leq \|T^{-1}\|\|ST^{-1}\|^{N+1}\sum_{n=0}^{M-N-1}\|ST^{-1}\|^n \\ &\leq \underbrace{\|T^{-1}\|}_{=\text{const}}\underbrace{\|ST^{-1}\|}_{<1}{}^{N+1}\cdot\sum_{n\geq 0}\underbrace{\|ST^{-1}\|}_{<1}{}^n\end{aligned}$$

und die rechte Seite konvergiert gegen Null, wenn $N, M \to \infty$. Wir haben also

$$A_N \to \sum_{n\geq 0}(-1)^n T^{-1}(ST^{-1})^n \in \mathcal{B}(X)$$

in Operatornorm gezeigt.

(3) Rechts- und Linksinverses: Wir zeigen nur, dass $\lim_{N\to\infty} A_N$ ein Rechtsinverses zu $S+T$ ist. Dass es auch ein Linksinverses ist, folgt analog. Wir haben

$$(S+T)A_N = (1+ST^{-1})\sum_{n=0}^{N}(-1)^n(ST^{-1})^n = 1 + (-1)^N(ST^{-1})^{N+1}.$$

Daraus folgt $\operatorname{ran}(A_N) \subseteq \mathcal{D}(S+T)$ und außerdem konvergiert $(S+T)A_N \to 1$. Für $\psi \in X$ erhalten wir sowohl $A_N\psi \to A\psi$ als auch $(S+T)A_N\psi \to \psi$. Jetzt verwenden wir, dass $S+T$ ein abgeschlossener Operator ist; dies folgt aus der Tatsache, dass $T$ abgeschlossen ist (notwendig, damit $T^{-1}$ sinnvoll definiert werden kann; wenn $T$ nicht abgeschlossen wäre, so wäre $\rho(T) = \emptyset$) und, dass $S$ relativ $T$-beschränkt mit $T$-Schranke $< 1$ ist. Aus der Abgeschlossenheit von $S+T$ folgt $A\psi \in \mathcal{D}(S+T)$ und, dass $(S+T)A\psi = \psi$ für alle $\psi \in X$. Somit ist $(S+T)A = 1$ wie behauptet.

(4) Surjektivität: Aus dem letzten Argument folgt insbesondere, dass $S+T$ surjektiv ist. Zusammen mit der Injektivität folgt, dass $S+T$ bijektiv ist und, dass die Resolvente durch $A = \sum_{n\geq 0}(-1)^n T^{-1}(ST^{-1})^n$ gegeben ist.

$\square$

In den Hausaufgaben zeigen Sie die folgenden Identitäten.

**Satz 10.4** (Resolventenformeln)**.** *Für lineare Operatoren $T, S$ in einem Banachraum gelten*

$$\begin{aligned} T^{-1} - S^{-1} = T^{-1}(S-T)S^{-1} &= S^{-1}(S-T)T^{-1}\,, \quad \textit{falls } 0 \in \rho(T)\cap\rho(S)\,, \\ (T-z)^{-1} - (S-z)^{-1} &= (T-z)^{-1}(S-T)(S-z)^{-1}\,, \quad \textit{falls } z \in \rho(T)\cap\rho(S)\,, \\ (T-z)^{-1} - (T-w)^{-1} &= (z-w)(T-w)^{-1}(T-z)^{-1}\,, \quad \textit{falls } w, z \in \rho(T)\,. \end{aligned}$$

*Beweisskizze.* Schreibe $T^{-1} = T^{-1}SS^{-1}$ und forme so um, dass die gewünschte Behauptung entsteht. $\square$

Die Formeln gelten auch für *unbeschränkte Operatoren* sofern die Definitionsbereiche von $T$ und $S$ passend sind. (Beispielsweise gilt die erste Identität der ersten Zeile, wenn $\mathcal{D}(S) \subseteq \mathcal{D}(T)$. Die zweite Identität gilt, wenn $\mathcal{D}(T) \subseteq \mathcal{D}(S)$. Beide Identitäten gelten, wenn $\mathcal{D}(T) = \mathcal{D}(S)$.) Beachten Sie, dass $T$ und $S$ im Allgemeinen nicht kommutieren, d.h. die Reihenfolge der Operatoren auf beiden Seiten der Gleichungen ist wichtig. Beispielsweise ist im Allgemeinen $T^{-1} - S^{-1} \neq T^{-1}S^{-1}(S-T)$.

**Satz 10.5.** *Sei $(X, \|\cdot\|)$ ein komplexer Banachraum und $T \in \mathcal{B}(X)$. Dann gelten folgende Aussagen.*

*(a) Die Resolventenmenge $\rho(T) \subseteq \mathbb{C}$ ist offen, sprich für alle $w \in \rho(T)$ gibt es $r_w > 0$, sodass auch $\{z \in \mathbb{C} : |z-w| < r_w\} \subseteq \rho(T)$ gilt. (Dementsprechend ist $\sigma(T) = \mathbb{C}\setminus\rho(T)$ abgeschlossen.)*

*(b) Für $w \in \rho(T)$ gilt $\|(T-w)^{-1}\| \geq [\mathrm{dist}(w, \sigma(T))]^{-1}$.*

*(c) Sei $w \in \rho(T)$ und $r_w := \|(T-w)^{-1}\|^{-1}$. Dann konvergiert die Reihe*

$$\sum_{n\geq 0}(z-w)^n[(T-w)^{-1}]^{n+1} = (T-z)^{-1}$$

*in Operatornorm für alle $z \in \mathbb{C}$ mit $|z-w| < r_w$. (Das heißt die Abbildung $\rho(T) \ni z \mapsto (T-z)^{-1} \in \mathcal{B}(X)$ ist komplex analytisch.)*

*(d) Seien $z \in \rho(T)$ und $(z_n)_{n\in\mathbb{N}} \subseteq \rho(T)$ eine Folge mit $z_n \to z$. Dann gilt $\|(T-z_n)^{-1} - (T-z)^{-1}\| \to 0$ für $n \to \infty$. (Das heißt die Abbildung $\rho(T) \ni z \mapsto (T-z)^{-1} \in \mathcal{B}(X)$ ist normstetig.)*

*(e) Es gilt $\sigma(T) \subseteq \{z \in \mathbb{C} : |z| \leq \|T\|\}$.*

*Beweisskizze.* (a) Sei $w \in \rho(T)$ sowie $r_w := \|(T-w)^{-1}\|^{-1} > 0$. Zeige dann, dass $B_w(r_w) \subseteq \rho(T)$, indem man für $z \in B_w(r_w)$ zeigt, dass $(T-z)^{-1}$ als Neumann-Reihe geschrieben werden kann.

(b) Sei $w \in \rho(T)$ und $r_w := \|(T-w)^{-1}\|^{-1} > 0$. Verwende dann $B_w(r_w) \subseteq \rho(T)$, um die Aussage zu folgern.

(c) Für $w \in \rho(T)$ und $r_w := \|(T-w)^{-1}\|^{-1}$ ist $w-z$ mit $z \in B_w(r_w)$ eine kleine Störung im Sinne, dass $\|(w-z)(T-w)^{-1}\| < 1$. Das bedeutet, dass auch $T-w+w-z = T-z$ invertierbar ist. Schreibe nun $(T-z)^{-1}$ als Neumann-Reihe.

(d) Für $z \in \rho(T)$ sei wieder $r_z := \|(T-z)^{-1}\|^{-1} > 0$. Betrachte alle Folgenglieder $z_n$ von $(z_n)_{n\in\mathbb{C}}$ mit $z_n \in B_z(r_z)$, dass auch $T-z_n$ invertierbar ist. Vergleiche dann die zu $z$ und $z_n$ gehörenden Neumann-Reihen in Operatornorm.

(e) Sei $z \in \mathbb{C}\setminus\{0\}$ so, dass $|z| > \|T\|$. Wir behaupten, dass $z \in \rho(T)$. Beobachte dazu $\|T/z\| < 1$ und verwende dies, um $(T-z)^{-1}$ als Neumannreihe aufzuschreiben.

$\square$

## 10.3 Einige Resultate zum Spektrum bestimmter linearer Operatoren

Weitaus mehr Resultate findet man in klassischen Büchern wie [RS72, RS78, Tes14, Wei00, Wei03].

**Satz 10.6** (Einfacher Spektralabbildungssatz). *Sei $(X, \|\cdot\|)$ ein komplexer Banachraum und $T : \mathcal{D}(T) \to X$ injektiv mit $\operatorname{ran}(T) \subseteq X$ dicht. Dann gilt $\sigma(T^{-1}) \setminus \{0\} = \{\lambda^{-1} : \lambda \in \sigma(T) \setminus \{0\}\}$. Insbesondere gilt $T\psi = z\psi$ für $z \in \mathbb{C} \setminus \{0\}$ und $\psi \in X \setminus \{0\}$ genau dann, wenn $T^{-1}\psi = z^{-1}\psi$.*

*Beweisskizze.* Angenommen $z \in \rho(T) \setminus \{0\}$. Wir zeigen dann, dass $z^{-1} \in \rho(T^{-1})$. Sobald wir dies gezeigt haben, folgt – durch Vertauschen der Rollen von $T$ und $T^{-1}$ bzw. $z$ und $z^{-1}$ – die Aussage $(z^{-1} \in \rho(T^{-1}) \setminus \{0\}) \Rightarrow (z \in \rho(T))$. Damit wäre die Behauptung gezeigt.

Sei also $z \in \rho(T) \setminus \{0\}$. Wir behaupten dann, dass $T^{-1} - z^{-1}$ beschränkt invertierbar ist. Tut man so, als wäre $T$ eine komplexe Zahl, kommt man auf $-zT(T-z)^{-1}$ als Resolvente von $T^{-1} - z^{-1}$, denn

$$\frac{1}{\frac{1}{T} - \frac{1}{z}} = \frac{1}{\frac{z-T}{Tz}} = -\frac{Tz}{T-z}\,.$$

Dies Nachzurechnen ist Teil der Hausaufgabe. $\square$

Wir erinnern an die Definition (oder Charakterisierung) unitärer Operatoren.

**Satz 10.7** (Definition und Charakterisierung unitärer Operatoren). *Sei $(\mathcal{H}, \langle\cdot,\cdot\rangle)$ ein komplexer Hilbertraum und $U \in \mathcal{B}(\mathcal{H})$. Dann heißt $U$ unitär genau dann, wenn eine (und damit alle) der folgenden Aussagen gelten.*

1. *Es gilt $UU^* = U^*U = \mathbf{1}_{\mathcal{H}}$. (Typischerweise als Definition angesehen.)*
2. *$U$ ist surjektiv und für alle $x, y \in \mathcal{H}$ gilt $\langle Ux, Uy\rangle = \langle x, y\rangle$.*
3. *$U$ ist surjektiv und isometrisch, d.h. $\|Ux\| = \|x\|$ für alle $x \in \mathcal{H}$.*
4. *$\operatorname{ran}(U)$ ist dicht in $\mathcal{H}$ und für alle $x, y \in \mathcal{H}$ gilt $\langle Ux, Uy\rangle = \langle x, y\rangle$.*

*Beweis.* (1) $\Rightarrow$ (2) $\Rightarrow$ (3) sind klar. Aus der Polarisationsidentität ($\langle x, y\rangle = 4^{-1}[\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2]$) folgt (3) $\Rightarrow$ (2).

Um (2) $\Rightarrow$ (1) zu sehen, bemerkt man, dass $\langle Ux, Uy\rangle = \langle x, y\rangle$ für alle $x, y \in \mathcal{H}$ impliziert, dass $U^*U = \mathbf{1}_{\mathcal{H}}$ (d.h. $U$ ist isometrisch). Insbesondere ist $U^*$ eine Links-Inverse von $U$. Das bedeutet $U$ ist injektiv. Zusammen mit der geforderten Surjektivität von $U$ bedeutet das, dass $U$ bijektiv mit Inverser $U^*$ ist.

Die Implikation (2) $\Rightarrow$ (4) ist offenbar. Die Implikation (4) $\Rightarrow$ (2) folgt aus der Tatsache, dass Isometrien von oben *und unten* beschränkt, sowie der Annahme, dass $\operatorname{ran}(U)$ dicht in $\mathcal{H}$ ist, siehe auch Lemma 10.8 unten. $\square$

**Lemma 10.8.** *Sei $X$ ein Banachraum und $T \in \mathcal{B}(X)$. Dann ist $T$ genau dann beschränkt invertierbar, wenn $T$ nach unten beschränkt ist — im Sinne $\|Tx\| \gtrsim \|x\|$ — und, dass $\operatorname{ran}(T) \subseteq X$ dicht ist.*

*Beweis.* "$\Rightarrow$": Da $T$ invertierbar ist, ist $\operatorname{ran}(T) = X$ und es gibt $T^{-1} : \mathcal{D}(T) \to X$ (mit $\mathcal{D}(T) = X$ oBdA, da $T$ beschränkt ist), sodass $T^{-1}T = \mathbf{1}_X$. Daraus folgt $\|x\| = \|T^{-1}Tx\| \leq \|T^{-1}\|\|Tx\|$, also $\|Tx\| \geq \frac{\|x\|}{\|T^{-1}\|}$, womit gezeigt ist, dass $T$ von unten beschränkt ist.

"$\Leftarrow$": Wenn $T$ von unten beschränkt ist, so ist $T$ injektiv. Andernfalls gäbe es $x \in \mathcal{D}(T)\setminus\{0\}$, sodass $Tx = 0$. Doch dann würde $0 = \|Tx\| \gtrsim \|x\| > 0$ gelten, was widersprüchlich ist. Es verbleibt zu zeigen, dass $T$ surjektiv ist, sprich, dass $\operatorname{ran}(T) = X$ gilt. Dass $\operatorname{ran}(T)$ dicht in $X$ liegt ist, ist gleichbedeutend mit $X = \overline{\operatorname{ran}(T)}$. Es genügt also zu zeigen, dass $\operatorname{ran}(T) = \overline{\operatorname{ran}(T)}$ gilt, also $\operatorname{ran}(T)$ abgeschlossen ist. Wegen Lemma 9.9 genügt es zu zeigen, dass $\operatorname{ran}(T)$ vollständig ist. Sei also $(y_n)_{n\in\mathbb{N}} \subseteq \operatorname{ran}(T)$ eine Cauchyfolge, die wir als $(Tx_n)_{n\in\mathbb{N}}$ mit $(x_n)_{n\in\mathbb{N}} \subseteq \mathcal{D}(T)$ schreiben. Aus der unteren Beschränktheit von $T$ folgt $\|(x_n - x_m)\| \lesssim \|T(x_n - x_m)\|$; ergo ist auch $(x_n)_{n\in\mathbb{N}} \subseteq \mathcal{D}(T) = X$ eine Cauchyfolge. Sei $x \in X$ der zugehörige Grenzwert $x = \lim_{n\to\infty} x_n$. Da $T$ beschränkt (also stetig, sh. Lemma XV.4) ist, gilt $\lim_{n\to\infty} y_n = \lim_{n\to\infty}(Tx_n) = T(\lim_{n\to\infty} x_n) = Tx \in \operatorname{ran}(T)$, das heißt die Cauchyfolge $(y_n)_{n\in\mathbb{N}} \subseteq \operatorname{ran}(T)$ mit der wir gestartet sind, konvergiert gegen ein Element in $\operatorname{ran}(T)$. Also ist $\operatorname{ran}(T)$ vollständig (bzw. abgeschlossen), was zu zeigen war. Wir haben also die Bijektivität von $T$ gezeigt. Zusammen mit der Beschränktheit von $T$ und dem Satz vom abgeschlossenen Graphen (siehe auch **Wiki**), folgt, dass $T$ beschränkt invertierbar ist. Dies schließt den Beweis. □

*Bemerkung* 10.9. Operatoren mit dichtem Bild sind nicht notwendigerweise invertierbar. Betrachte $A : \ell^2 \to \ell^2$ definiert durch $\ell^2 \ni x = (x_n)_{n\in\mathbb{N}} \mapsto Ax := (2^{-n}x_n)_{n\in\mathbb{N}}$. Dann ist die Menge aller endlichen (bzw. kompakt getragenen) Folgen $Y = \{y \in \mathbb{C}^{\mathbb{N}} : \exists\, m \in \mathbb{N}, \text{ sodass } y = (y_1, ..., y_m, 0, ...)\}$ in $\operatorname{ran}(A)$ enthalten. Außerdem ist $Y \subseteq \ell^2$ dicht. (So wie $C_c^\infty(\mathbb{R}^d)$ dicht in $L^2(\mathbb{R}^d)$ liegt.) Jedoch gibt es kein $x \in \ell^2$, sodass beispielsweise $Ax = (1/n)_{n\in\mathbb{N}} \in \ell^2$.

Der hier betrachtete Operator $A$ ist *nicht* von unten beschränkt. Dazu betrachtet man für $N \in \mathbb{N}$ die Beispielsfolge

$$(x_n^{(N)})_{n\in\mathbb{N}} := \left(\frac{1}{n}\mathbf{1}_{n>N}\right)_{n\in\mathbb{N}} .$$

Dann sind $\|x_n^{(N)}\|_{\ell_n^2(\mathbb{N})} \sim N^{-1/2}$ aber $\|2^{-n} \cdot x_n^{(N)}\|_{\ell_n^2(\mathbb{N})} \sim 2^{-N}$.

**Satz 10.10** (Spektrum adjungierter Operatoren)**.** *Sei $T \in \mathcal{B}(\mathcal{H})$. Dann gilt*

$$\rho(T^*) = \{\overline{z} \in \mathbb{C} : z \in \rho(T)\} \quad \textit{und} \quad \sigma(T^*) = \{\overline{z} \in \mathbb{C} : z \in \sigma(T)\}$$

*Beweis.* Es genügt die Aussage über die Resolventenmenge zu beweisen.

"$\supseteq$": Wenn $z \in \rho(T)$ ist, dann ist $T - z$ beschränkt invertierbar, also existiert $(T-z)^{-1} \in \mathcal{B}(\mathcal{H})$. Insbesondere existiert $[(T-z)^{-1}]^* = (T^* - \overline{z})^{-1} \in \mathcal{B}(\mathcal{H})$. Damit ist $\overline{z} \in \rho(T^*)$ wie behauptet.

"$\subseteq$": Da $T = (T^*)^*$, folgt aus dem eben Bewiesenem (für $T^*$ statt $T$), dass

$$\{\overline{z} \in \mathbb{C} : z \in \rho(T^*)\} \subseteq \rho((T^*)^*) = \rho(T) .$$

Das heißt, für $z \in \rho(T^*)$ ist $\overline{z} \in \rho(T)$ wie behauptet. □

**Satz 10.11** (Spektrum unitärer Operatoren)**.** *Sei $(\mathcal{H}, \langle\cdot,\cdot\rangle)$ ein komplexer Hilbertraum und $U : \mathcal{H} \to \mathcal{H}$ unitär. Dann gelten folgende Aussagen.*

*(a) Es gilt $\sigma(U) \subseteq \{z \in \mathbb{C} : |z| = 1\}$. Weiter haben alle Eigenwerte von $U$ Absolutbetrag gleich Eins.*

*(b) Seien $\lambda_1, \lambda_2 \in \sigma(U)$ mit $\lambda_1 \neq \lambda_2$ zwei verschiedene Eigenwerte mit zugehörigen Eigenvektoren $\psi_1, \psi_2$. Dann ist $\langle\psi_1|\psi_2\rangle = 0$.*

*Beweisskizze.* (a) Dass alle Eigenwerte Absolutbetrag gleich Eins haben folgt durch Betrachten von $\|Ux\|$ für beliebige Eigenvektoren $x$ von $U$ (Hausaufgabe). Wir zeigen noch $\sigma(U) \subseteq \{z \in \mathbb{C} : |z| = 1\}$. Aus (e) in Satz 10.5 folgt, dass $\sigma(U) \subseteq \{z \in \mathbb{C} : |z| \leq 1\}$. Da auch $U^{-1}$ unitär ist, folgt aus Satz 10.6, dass $\sigma(U) = \sigma((U^{-1})^{-1}) = \{z^{-1} : z \in \sigma(U^{-1})\} \subseteq \{z \in \mathbb{C} : |z| \geq 1\}$. Daraus folgt $\sigma(U) \subseteq \{z \in \mathbb{C} : |z| = 1\}$.

(b) Aus $U^* = U^{-1}$ sowie $U^*\psi_j (= U^{-1}\psi_j = \lambda_j^{-1}\psi_j) = \overline{\lambda_j}\psi_j$ (Sätze 10.6 und 10.10) folgt

$$(\lambda_1 - \lambda_2)\langle\psi_1, \psi_2\rangle = \langle U^*\psi_1, \psi_2\rangle - \langle\psi_1, U\psi_2\rangle = 0 ,$$

wie behauptet.

□

## 10.4 Spektralradius

**Behauptung 10.12.** *Sei $A \in \mathbb{C}^{n\times n}$. Dann gilt $r(A) < 1$ genau dann, wenn*

$$\lim_{k\to\infty} A^k = 0\,.$$

*Wenn andernfalls $r(A) > 1$, dann gilt $\lim_{k\to\infty} \|A^k\| = \infty$ (für jede Matrixnorm).*

*Beweis.* "$\Leftarrow$": Angenommen $\lim_{k\to\infty} A^k = 0$. Sei $(v, \lambda)$ ein Eigenvektor-Eigenwert-Paar von $A$. Aus $A^k v = \lambda^k v$ folgt

$$0 = (\lim_{k\to\infty} A^k)v = \lim_{k\to\infty} (A^k v) = \lim_{k\to\infty} \lambda^k v = v \lim_{k\to\infty} \lambda^k\,.$$

Da $v \neq 0$, folgt $\lim_{k\to\infty} \lambda^k = 0$, woraus $|\lambda| < 1$ für alle Eigenwerte von $A$ und damit $r(A) < 1$ folgt.

"$\Rightarrow$": Aus der Jordan-Normalform wissen wir, dass es invertierbares $S \in \mathbb{C}^{n\times n}$ und block-diagonales $J \in \mathbb{C}^{n\times n}$ gibt, sodass $A = SJS^{-1}$ mit

$$J = \begin{pmatrix} J_{m_1}(\lambda_1) & 0 & 0 & \cdots & 0 \\ 0 & J_{m_2}(\lambda_2) & 0 & \cdots & 0 \\ \vdots & \cdots & \ddots & \cdots & \vdots \\ 0 & \cdots & 0 & J_{m_{s-1}}(\lambda_{s-1}) & 0 \\ 0 & \cdots & \cdots & 0 & J_{m_s}(\lambda_s) \end{pmatrix}, \quad 1 \le s \le n\,,$$

wobei

$$J_{m_i}(\lambda_i) = \begin{pmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_i & 1 \\ 0 & 0 & \cdots & 0 & \lambda_i \end{pmatrix} \in \mathbb{C}^{m_i\times m_i}, \quad 1 \le i \le s\,.$$

Wir haben bereits bewiesen, dass $A^k = SJ^kS^{-1}$. Da $J$ blockdiagonal ist, gilt

$$J^k = \begin{pmatrix} J_{m_1}(\lambda_1)^k & 0 & 0 & \cdots & 0 \\ 0 & J_{m_2}(\lambda_2)^k & 0 & \cdots & 0 \\ \vdots & \cdots & \ddots & \cdots & \vdots \\ 0 & \cdots & 0 & J_{m_{s-1}}(\lambda_{s-1})^k & 0 \\ 0 & \cdots & \cdots & 0 & J_{m_s}(\lambda_s)^k \end{pmatrix}.$$

Durch Induktion und binomische Formel weist man

$$J_{m_i}(\lambda_i)^k = \begin{pmatrix} \lambda_i^k & \binom{k}{1}\lambda_i^{k-1} & \binom{k}{2}\lambda_i^{k-2} & \cdots & \binom{k}{m_i-1}\lambda_i^{k-m_i+1} \\ 0 & \lambda_i^k & \binom{k}{1}\lambda_i^{k-1} & \cdots & \binom{k}{m_i-2}\lambda_i^{k-m_i+2} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_i^k & \binom{k}{1}\lambda_i^{k-1} \\ 0 & 0 & \cdots & 0 & \lambda_i^k \end{pmatrix}$$

nach. Wenn $r(A) < 1$, dann sind alle $|\lambda_i| < 1$, woraus

$$\lim_{k\to\infty} J_{m_i}(\lambda_i)^k = 0 \Rightarrow \lim_{k\to\infty} J^k = 0 \Rightarrow \lim_{k\to\infty} A^k = 0\,.$$

Ist andererseits $r(A) > 1$, so gibt es mindestens einen Jordanblock $J_{m_i}$ für den $\|J_{m_i}^k\|$ mit $k$ wächst. Dies schließt den Beweis. $\square$

Die Aussagen des folgenden Satzes gelten auch im unendlichdimensionalen Fall.

**Satz 10.13.** *Sei $(X, \|\cdot\|)$ ein endlichdimensionaler normierter $\mathbb{C}$-Vektorraum, $A \in \mathcal{B}(X)$ und*

$$r(A) := \sup\{|\lambda| : \lambda \in \sigma(A)\}$$

*der Spektralradius von $A$. Sei $\|\cdot\|$ eine beliebige Norm auf $X$ und $\|\cdot\|_{\mathrm{op}}$ die zugehörige Operatornorm. Dann gelten folgende Aussagen.*

*(a) Es gilt $\|A\|_{\mathrm{op}} \geq r(A)$.*

*(b) Es gilt die Gelfandsche Formel $\lim_{n\to\infty}(\|A^n\|_{\mathrm{op}})^{1/n} = r(A)$.*

*(c) Sind speziell $(X, \langle\cdot,\cdot\rangle)$ ein $\mathbb{C}$-Hilbertraum und $\|x\| = \sqrt{\langle x, x\rangle}$, $x \in X$, die durch das Skalarprodukt auf $X$ induzierte Norm sowie $A$ normal, d.h. $AA^* = A^*A$, so gilt $\|A\|_{\mathrm{op}} = r(A)$.*

*Beweis.* (a) Für alle $k \in \mathbb{N}$ und jedes Eigenvektor-Eigenwert-Paar $(v, \lambda)$ von $A$ gilt (dank der Submultiplikativität der Operatornorm)

$$|\lambda|^k\|v\| = \|\lambda^k v\| = \|A^k v\| \leq \|A^k\|\|v\|\,.$$

Da $v \neq 0$, ist

$$|\lambda|^k \leq \|A^k\| \Rightarrow |\lambda| \leq \|A^k\|^{1/k}\,.$$

Die Aussage folgt dann, indem man $k = 1$ setzt und $\lambda$ als den betragsmäßig gröss ten Eigenwert von $A$ wählt.

(b) Nun zum Beweis von Gelfands Formel mit Hilfe von Proposition 10.12.

**Satz 10.14.** *Sei $A \in \mathbb{C}^{n\times n}$ und $r(A) = \max\{|\lambda| : \lambda \in \sigma(A)\}$. Es gilt $r(A) = \lim_{k\to\infty} \|A^k\|^{1/k}$ in jeder Matrixnorm $\|\cdot\|$.*

*Beweis.* Für alle $\varepsilon > 0$ definiere

$$A_\pm := \frac{1}{r(A) \pm \varepsilon} A\,.$$

Dann folgt

$$r(A_\pm) = \frac{r(A)}{r(A) \pm \varepsilon} \quad \text{und} \quad r(A_+) < 1 < r(A_-)\,.$$

Wir wenden nun Behauptung 10.12 auf $A_+$ an und sehen, dass $\lim_{k\to\infty} A_+^k = 0$. Das bedeutet, dass es $N_+ = N_+(\varepsilon) \in \mathbb{N}$ gibt, sodass für alle $k \geq N_+$ wir haben, dass $\|A_+^k\| < 1$. Per Definition von $A_+$ folgt

$$\|A^k\|^{1/k} < r(A) + \varepsilon\,.$$

Analog folgt aus Behauptung 10.12 für $A_-$, dass es $N_- = N_-(\varepsilon) \in \mathbb{N}$ gibt, sodass für alle $k \geq N_-$ gilt, dass $\|A_-^k\| > 1$. Daraus folgt

$$\|A^k\|^{1/k} > r(A) - \varepsilon\,.$$

Wir kombinieren diese beiden Aussagen nun. Dann gibt es für alle $\varepsilon > 0$ ein $N = N(\varepsilon) \in \mathbb{N}$, wie beispielsweise $N := 1 + \max\{N_+(\varepsilon), N_-(\varepsilon)\}$, sodass für alle $k \geq N$ gilt, dass

$$r(A) - \varepsilon < \|A^k\|^{1/k} < r(A) + \varepsilon\,.$$

Dies ist gleichbedeutend mit

$$\lim_{k\to\infty} \|A^k\|^{1/k} = r(A)\,,$$

was den Beweis von Theorem 10.14 schließt. □

(c) Wir verwenden Gelfands Formel aus Teil (b). Zunächst folgt aus der Normalität von $A$ die metrische Gleichheit

$$\|A\psi\|^2 = \langle A\psi, A\psi\rangle = \langle \psi, A^*A\psi\rangle = \langle \psi, AA^*\psi\rangle = \|A^*\psi\|^2 .$$

(Beobachte, dass dies stärker als die allgemeingültige Gleichheit $\|A\| = \|A^*\|$ ist.) Dann folgt aus $\|A\|^2 = \|A^*A\|$, dass

$$\|A\|^2 = \|A^*A\| = \max_{\|x\|=1} \|A^* \underbrace{Ax}_{\equiv\psi} \| = \max \|AAx\| = \|A^2\| .$$

Durch Iteration folgt daraus $\|A^{2n}\| = \|A\|^{2n}$ und mit Gelfands Formel

$$r(A) = \lim_{n\to\infty} \|A^n\|^{1/n} = \lim_{n\to\infty} \|A^{2n}\|^{1/(2n)} = \|A\| .$$

□

**Beispiel 10.15** (Input-Output-Analyse nach V. Leontieff)**.** (a) Eine Volkswirtschaft besitze die Industrien $X_1, ..., X_n$, die gewisse Outputs erzeugen. Um einen Output im Wert von 1 EUR zu erzeugen, benötigt Industrie $X_j$ Inputs der Industrien $X_i$ im Wert von $a_{ij}$ EUR für $i = 1, ..., n$. Dabei ist vernünftigerweise

$$0 \le a_{ij} \quad \text{für } i, j = 1, ..., n \quad \text{und} \quad \sum_{i=1}^{n} a_{ij} < 1 \quad \text{für } j = 1, ..., n \tag{10.1}$$

anzunehmen. Produziert nun jede Industrie $X_i$ einen Output im Wert von $x_i$ EUR, so stehen für Konsumenten nur noch die Outputs $x_i - \sum_{j=1}^{n} a_{ij}x_j$ zur Verfügung. Das Problem besteht darin, genau soviel zu produzieren, dass eine gegebene Nachfrage $d = (d_1, ..., d_n)$ befriedigt werden kann.

(b) Dazu schreibt man $x = (x_1, ..., x_n)$ für den Produktionsvektor und führt die Matrix $A := (a_{ij}) \in \mathbb{R}^{n\times n}$ ein. Zu lösen ist dann die Gleichung $x - Ax = d$ oder $(1 - A)x = d$.

(c) Nun folgt aus (10.1) sofort $\|A\|_{1\to1} = \max_{1\le j\le n} \sum_{i=1}^{n} |A_{ij}| < 1$ (Spaltensummennorm). Also konvergiert die Neumannsche Reihe $(1 - A)^{-1} = \sum_{n\ge0} A^n$. Da alle Matrixeinträge von $A$ nicht-negativ sind, sind auch alle Matrixeinträge von $(1 - A)^{-1}$ nicht-negativ.
<u>Beispiel:</u> Für die Industrien Kohle, Stahl und Elektrizität hat man die Matrix

$$A = \begin{pmatrix} 0 & 0,15 & 0,43 \\ 0,02 & 0,03 & 0,20 \\ 0,01 & 0,08 & 0,05 \end{pmatrix} .$$

Es gilt $\|A\|_{\infty\to\infty} = \max_{1\le i\le n} \sum_{j=1}^{n} |A_{ij}| = 0,58$ (Zeilensummennorm), $\|A\|_{1\to1} = 0,68$ und $\|A\|_{\mathrm{HS}} = (\sum_{j=1}^{3} \rho(A)^2)^{1/2} = 0,51$ (Hilbert–Schmidt-Norm). Mit den Eigenwerten $0,204$, $-0,043$ und $-0,081$ ergibt sich für den Spektralradius $r(A) = 0,204$. Weiter ist

$$(1 - A)^{-1} = \begin{pmatrix} 1,00892 & 0,19711 & 0,49817 \\ 0,02340 & 1,05372 & 0,23243 \\ 0,01259 & 0,09081 & 1,07745 \end{pmatrix} .$$

## 11 11. Tutoriumsblatt (28. Juli 2022)

Tipps zur Vorbereitung auf die Klausur:

- Lernen Sie den Inhalt der Vorlesung nicht einfach auswendig. Investieren Sie Zeit darin, die Aussagen der einzelnen Lemmata und Sätze zu verstehen. Stellen Sie sich dazu folgende Fragen: *Warum ist dieser oder jener Satz wichtig? Wo und wie kann man ihn anwenden?*

- Machen Sie sich mit den Voraussetzungen und Konsequenzen (oder Äquivalenzen) der Lemmata und Sätze vertraut. Stellen Sie in den Beweisen fest, wo welche Annahme wie stark gebraucht wird. Wie ändert sich die Aussage, wenn Sie eine Annahme abschwächen oder weglassen? Welche weiteren Annahmen wären natürlich zu fordern, wenn man die Aussage verstärken wollte?

- Entwickeln Sie ein Verständnis dafür, wie die einzelnen Lemmata und Sätze miteinander in Beziehung stehen. Welche Lemmata wurden aus welchen Gründen für den Beweis eines Satzes gebraucht?

- Zur Illustration der Aussagen der Sätze gehen Sie gründlich durch die einfachen Beispiele in den Vorlesungsnotizen sowie die Übungs- und Tutoriumsaufgaben durch. Machen Sie sich damit vertraut, für welche Art von Problem welche Methode angewandt wurde. Durchdringen Sie, warum diese oder jene Methode verwendet wurde. Gibt es alternative Beweismethoden?

Folgende Punkte sollten Sie unter Anderem aus der Vorlesung mitgenommen haben:

(1) Euklidsche Ringe und euklidscher Algorithmus.

  (a) Definition euklidscher Ring (Definition XI.1, Lemma XI.2).
  (b) Definition und Charakterisierung von $\mathrm{ggT}(a, b)$ und $\mathrm{kgV}(a, b)$ (Definition XI.3, Lemmata XI.4–5, Übungsblatt 2).
  (c) Euklidscher Algorithmus (S. 130–131 im Skript oder Satz 1.13 im Tutoriumsblatt).
  (d) Beispiel in $\mathbb{Z}$ oder $\mathbb{Z}_p$ mit $p$ Prim (Übungsblatt 1).
  (e) Definition von Prim und Zusammenhang mit Unzerlegbarkeit in euklidschen Ringen (Übungsblatt 1).

(2) Polynome.

  (a) Begriffe zu Polynomen. Normiert, irreduzibel, teilerfremd (Definition XI.6).
  (b) Einfache Aussagen zur Teilung von Polynomen (Lemmata XI.7–8, Korollar XI.9).
  (c) Analogon der Primfaktorzerlegung für Polynome (Satz XI.10, Korollar XI.12).
  (d) Euklidscher Algorithmus in $\mathbb{F}[X]$ mit $\mathbb{F} \in \{\mathbb{Z}, \mathbb{Z}_p, \mathbb{Q}\}$ (Übungsblätter 1–2).

(3) Minimalpolynom und Cayley–Hamilton.

  (a) Polynome von Endomorphismen auf endlichdimensionalen Vektorräumen.
  (b) Definition Minimalpolynom. Existenz, Eindeutigkeit und einfach obere Schranke für den Grad des Minimalpolynoms. (Definition XII.1, Bemerkung, Lemma XII.2).
  (c) Cayley–Hamilton (Satz XII.3), verbesserte Schranke an Grad des Minimalpolynoms. Berechnung des Minimalpolynoms anhand von Beispielen (Übungsblätter 3–4).
  (d) Zusammenhang zwischen charakteristischem Polynom und Diagonalisierbarkeit (S. 138–139 im Skript oder Satz 4.6 im Tutoriumsblatt).
  (e) Zusammenhang zwischen Minimalpolynom und den Eigenwerten sowie der Diagonalisierbarkeit (Sätze 2.4–2.5 im Tutoriumsblatt).

(f) *Optional:* Anwendung von Cayley–Hamilton zur Berechnung analytischer Funktionen von Matrizen (Abschnitt 2.1.1 im Tutoriumsblatt).

(4) Jordansche Normalform.

(a) Eindeutige Zerlegung des Minimalpolynoms in Potenzen normierter, irreduzibler, paarweise teilerfremder Polynome gemäß Korollar XI.12 $\Rightarrow$ Hauptraumzerlegung (Satz XII.9 oder Satz 4.21 im Tutoriumsblatt. Übungsblatt 4).

(b) Charakterisierung nilpotenter Matrizen (Satz 4.19 im Tutoriumsblatt).

(c) Jordanisierung nilpotenter Matrizen (Satz XIII.2 oder Satz 4.24 im Tutoriumsblatt)

(d) Jordansche Normalform für Matrizen über algebraisch abgeschlossenen Körpern, wie bspw. $\mathbb{C}$. (Satz XIII.3 oder Satz 4.26 im Tutoriumsblatt. Übungsblätter 4–6.).

(e) Zusammenhang zwischen Größe der Jordanblöcke und algebraischer Vielfachheit (Satz XIII.3 oder Satz 4.26 im Tutoriumsblatt).

(f) *Optional:* Zusammenhang zwischen Größe der Jordanblöcke und geometrischer Vielfachheit (Satz 4.27 im Tutoriumsblatt).

(5) Singulärwertzerlegung.

(a) Bra-Ket-Notation von Dirac (Definition XIV.1).

(b) Darstellung der Einheitsmatrix, selbstadjungierter und allgemeiner Endomorphismen in Bra-Ket-Notation (Lemma XIV.2).

(c) Singulärwerte und Singulärwertzerlegung (Satz XIV.3, Übungsblatt 7).

(d) $\ker(\Phi) = \ker(\Phi^*\Phi)$ und $\ker(\Phi^*) = \ker(\Phi\Phi^*)$ für Endomorphismen $\Phi$ in komplexen Vektorräumen (S. 156 im Skript).

(e) *Optional:* $\sigma(AB) = \sigma(BA)$, aber Singulärwerte von $AB$ stimmen nicht unbedingt mit denen von $BA$ überein. Min-Max-Prinzip. Einfache Schranken für Singulärwerte (Aufgaben 7.2–7.5 im Tutoriumsblatt).

(6) Hilberträume und Operatoren.

(a) Definitionen Vollständigkeit, Offenheit, Abgeschlossenheit, Banachraum, Hilbertraum (Definition XV.1 und anschließende Bemerkung, oder 9. Tutoriumsblatt).

(b) Definitionen Operatornorm, beschränkte lineare Operatoren und Dualraum (Definition XV.2, Lemma XV.3).

(c) Eigenschaften und Charakterisierungen Operatornorm und Beispiele (Lemmata XV.5, XV.4, XV.6, Satz XV.13, Übungsblätter 8–9).

(d) Orthogonale Komplemente (Satz XV.8, Korollar XV.9, Übungsblatt 10)

(e) Rieszscher Darstellungssatz und adjungierter Operator (Satz XV.10, Korollar XV.11)

(f) Resolvente, Resolventenmenge, Spektrum, Neumann-Reihe (Abschnitt XV.3 im Skript und Übungsblätter 11–12).

(g) *Optional:* Spektralradius (Satz 10.13 im Tutoriumsblatt).

# Literatur


[Beu94] Albrecht Beutelspacher. *Lineare Algebra.* Vieweg, 1994.

[Bha97] Rajendra Bhatia. *Matrix Analysis*, volume 169 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 1997.

[Bos14] S. Bosch. *Lineare Algebra.* Springer-Lehrbuch. Springer Berlin Heidelberg, 2014.

[Bos20] Siegfried Bosch. *Algebra.* Springer Berlin Heidelberg, 2020.

[Dei78] P. A. Deift. Applications of a commutation formula. *Duke Math. J.*, 45(2):267–310, 1978.

[DMRV06] O. Dovgoshey, O. Martio, V. Ryazanov, and M. Vuorinen. The Cantor function. *Expo. Math.*, 24(1):1–37, 2006.

[Dym13] Harry Dym. *Linear Algebra in Action*, volume 78 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, second edition, 2013.

[Fan51] Ky Fan. Maximum properties and inequalities for the eigenvalues of completely continuous operators. *Proc. Nat. Acad. Sci., U. S. A.*, 37:760–766, 1951.

[Fis05] G. Fischer. *Lineare Algebra.* Vieweg-Studium : Grundkurs Mathematik. Vieweg, 2005.

[GO03] Bernard R. Gelbaum and John M. H. Olmsted. *Counterexamples in Analysis.* Dover Publications, Inc., Mineola, NY, 2003. Corrected reprint of the second (1965) edition.

[GT08] Ben Green and Terence Tao. The primes contain arbitrarily long arithmetic progressions. *Ann. of Math. (2)*, 167(2):481–547, 2008.

[Hel13] H. A. Helfgott. The ternary Goldbach conjecture is true. *arXiv e-prints*, page arXiv:1312.7748, December 2013.

[MK08] Gerhard Michler and H.-J. Kowalsky. *Lineare Algebra.* De Gruyter, 2008.

[RS72] Michael Reed and Barry Simon. *Methods of Modern Mathematical Physics. I. Functional Analysis.* Academic Press, New York, 1972.

[RS78] Michael Reed and Barry Simon. *Methods of Modern Mathematical Physics*, volume 4: Analysis of Operators. Academic Press, New York, 1 edition, 1978.

[Sak71] Shôichirô Sakai. *$C^*$-Algebras and $W^*$-Algebras.* Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 60. Springer-Verlag, New York-Heidelberg, 1971.

[SP08] R. Schulze-Pillot. *Einführung in Algebra und Zahlentheorie.* Springer-Lehrbuch. Springer Berlin Heidelberg, 2008.

[Tes14] Gerald Teschl. *Mathematical Methods in Quantum Mechanics*, volume 157 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, second edition, 2014. With applications to Schrödinger operators.

[Wei00] Joachim Weidmann. *Lineare Operatoren in Hilberträumen. Teil 1.* Mathematische Leitfäden. [Mathematical Textbooks]. B. G. Teubner, Stuttgart, 2000. Grundlagen. [Foundations].

[Wei03] Joachim Weidmann. *Lineare Operatoren in Hilberträumen. Teil II.* Mathematische Leitfäden. [Mathematical Textbooks]. B. G. Teubner, Stuttgart, 2003. Anwendungen. [Applications].