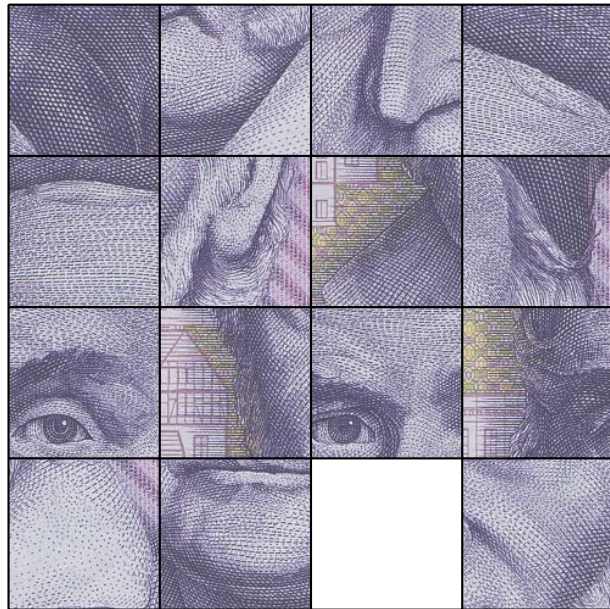


Lineare Algebra

Benjamin Sambale
Leibniz Universität Hannover

Version: 7. März 2026



Inhaltsverzeichnis

Einleitung	5
Vorwort	5
Motivation	5
Notation	6
Konventionen	8
Lineare Algebra I	11
1 Aussagenlogik und Mengenlehre	11
1.1 Aussagen	11
1.2 Mengen	13
1.3 Vollständige Induktion	16
2 Kartesische Produkte und Funktionen	17
2.1 Paare und Tupel	17
2.2 Injektive und surjektive Funktionen	19
3 Körper und Vektorräume	23
3.1 Gruppen und Körper	23
3.2 Vektorräume und Unterräume	25
4 Basen und Dimension	29
4.1 Lineare Unabhängigkeit und Erzeugendensysteme	29
4.2 Charakterisierung und Existenz von Basen	31
4.3 Dimension	33
5 Matrizen	35
5.1 Der Matrizen-Vektorraum	35
5.2 Matrizenmultiplikation	37
5.3 Der Rang einer Matrix	39
6 Der Gauß-Algorithmus	41
6.1 Gleichungssysteme	41
6.2 Elementare Zeilenoperationen	43
6.3 Anwendungen	45
7 Lineare Abbildungen	50
7.1 Definitionen und Beispiele	50
7.2 Darstellungsmatrizen	54
7.3 Dualräume	59

8	Eigenwerte und Eigenvektoren	63
8.1	Definitionen und Beispiele	63
8.2	Diagonalisierbarkeit	64
9	Determinanten	67
9.1	Rekursive Definition	67
9.2	Eigenschaften	70
9.3	Laplace-Entwicklung	72
9.4	Die Leibniz-Formel	75
	Aufgaben	80
	Lineare Algebra II	86
10	Polynome	86
10.1	Der Vektorraum der Polynome	86
10.2	Nullstellen	89
10.3	Charakteristische Polynome	91
10.4	Minimalpolynome	95
11	Euklidische Geometrie	99
11.1	Skalarprodukte	99
11.2	Orthonormalbasen	102
11.3	Symmetrische und orthogonale Abbildungen	104
11.4	Komplexe Zahlen	107
11.5	Der Hauptsatz	110
12	Bilinearformen	113
12.1	Gram-Matrizen	113
12.2	Sylvesters Trägheitssatz	116
12.3	Positiv definite Matrizen	120
13	Unitäre Räume	125
13.1	Sesquilinearformen	125
13.2	Adjungierte Abbildungen	126
13.3	Der Spektralsatz	129
14	Die Jordan-Normalform	134
14.1	Haupträume	134
14.2	Jordanblöcke	138
14.3	Anwendungen	142
15	Die Frobenius-Normalform	145
15.1	Irreduzible Polynome	145
15.2	Begleitmatrizen	148
15.3	Zentralisatoren	153
15.4	Zerfällungskörper	155
16	Die Jordan-Chevalley-Zerlegung	157
16.1	Der chinesische Restsatz	157

16.2 Separable und halbeinfache Abbildungen	160
16.3 Verallgemeinerte Jordanblöcke	161
Aufgaben	165
Lineare Algebra III	171
17 Numerische Verfahren	171
17.1 Effiziente Arithmetik	171
17.2 Die Konditionszahl	175
17.3 Stabile Varianten des Gauß-Algorithmus	179
17.4 Iterative Verfahren	184
17.5 Matrixnormen	188
17.6 Eigenwertberechnung	191
17.7 Orthonormalisierung	196
18 Analytische Aspekte	199
18.1 Eigenwertabschätzungen	199
18.2 Der Spektralradius	200
18.3 Die Exponentialfunktion einer Matrix	202
18.4 Nicht-negative Matrizen	205
18.5 Der Page-Rang	210
19 Lineare Optimierung	212
19.1 Lineare Programme	212
19.2 Konvexe Mengen	214
19.3 Der Simplex-Algorithmus	217
20 Gitter und quadratische Formen	222
20.1 Gitter	222
20.2 Die Minimal-Norm	224
20.3 Ganzzahlige Matrizen	227
20.4 Der LLL-Algorithmus	231
20.5 Quadratische Formen	236
20.6 Sukzessive Minima	241
Aufgaben	243
Anhang	249
Stichwortverzeichnis	255

Einleitung

Vorwort

Dieses Skript ist eine Erweiterung meiner Vorlesung Lineare Algebra A&B im Wintersemester 2020/21 und Sommersemester 2021 an der Leibniz Universität Hannover. Während sich die Vorlesung hauptsächlich an Studierende der Informatik gewendet hat, ist das vorliegende Skript an Mathematiker gerichtet. Der dritte Teil wurde Ende 2025 ergänzt und beschäftigt sich hauptsächlich praktischen Anwendungen. Einige der präsentierten Sätze (u. a. von Fillmore, Mirsky, Mazur-Ulam, Schur-Horn und Frobenius) findet man nur schwer in der Standard-Literatur. Außerdem wird die Frobenius-Normalform in voller Allgemeinheit behandelt. Die Übungsaufgaben beschränken sich auf theoretische Aspekte und sollten durch praktische Beispiele ergänzt werden. Ich danke Annika Bartelt und Gereon Koßmann für Fehlerhinweise (weitere Hinweise sind willkommen). Das folgende Buch deckt in etwa die Themen von Teil I und II ab:

Hoffman, Kunze, *Linear algebra*, 2nd edition, Prentice-Hall, New Jersey, 1971

Motivation

Sie haben zu verschiedenen Zeitpunkten $t_1 = 1, t_2 = 2, \dots$ durch physikalische Experimente Messdaten $d_1 = -2, d_2 = 3, \dots$ gewonnen. Aus theoretischen Überlegungen sei bekannt, dass diese Daten einem Gesetz folgen, das heißt, es gibt eine Funktion f mit $f(t_i) = d_i$ für $i = 1, 2, \dots$. Dabei hängt f (linear) von unbekanntem Parametern x_1, x_2, \dots ab, zum Beispiel $f(t) = t^2 x_1 - t x_2 + x_3$. Die Bestimmung dieser Parameter auf Grundlage der Messdaten führt auf ein lineares Gleichungssystem:

$$\begin{aligned}x_1 - x_2 + x_3 &= -2 \\4x_1 - 2x_2 + x_3 &= 3 \\&\vdots\end{aligned}\tag{S}$$

Wir beantworten unter anderem folgende Fragen:

- Wann ist das System (S) lösbar? (Satz 6.4)
- Wie viele Lösungen gibt es? (Bemerkung 6.7(b))
- Welche Struktur hat die Lösungsmenge? (Satz 6.6)
- Wie berechnet man alle Lösungen in der Praxis? (Satz 6.15)

Die entwickelten Methoden (Vektorräume, Matrizen und lineare Abbildungen) haben zahlreiche Anwendungen in anderen Gebieten:

- Bildverarbeitung: Wie erkennt man Gesichter auf Fotos?

- Suchmaschinen: Nach welchen Kriterien bewertet Google Internetseiten? (Abschnitt 18.5)
- Codierungstheorie: Wie erkennt und korrigiert man Fehler bei der Übertragung digitaler Daten?
- Kryptografie: Wie verschlüsselt man Daten resistent gegen Angriffe mit Quantencomputern? (Bemerkung 20.11)
- Elektrotechnik: Wie berechnet man Widerstände in Schaltkreisen?
- Meteorologie: Wie sagt man das Wetter von Morgen voraus?
- Stochastik: Mit welcher Wahrscheinlichkeit gelangt man nach einer Irrfahrt zum Ziel?
- Künstliche Intelligenz: Wie werden Large Language Models trainiert?

Notation

bzgl.	bezüglich
d. h.	das heißt
ggf.	gegebenenfalls
o. B. d. A.	ohne Beschränkung der Allgemeinheit
u. ä.	und ähnliche
vgl.	vergleiche
w, f	wahr, falsch
$\wedge, \vee, \neg, \Rightarrow, \Leftrightarrow, \exists, \forall$	logische Ausdrücke
$:=, :\Leftrightarrow$	linke Seite wird durch rechte Seite definiert
\square	Beweisende
\emptyset	leere Menge
$\mathcal{P}(M)$	Potenzmenge von M
$\cup, \dot{\cup}, \cap, \setminus$	Vereinigung (disjunkt), Durchschnitt, Differenz von Mengen
$ A $	Mächtigkeit von A (Anzahl der Elemente)
$(a, b), (x_1, \dots, x_n)$	Paar, n -Tupel
$A_1 \times \dots \times A_n$	kartesisches Produkt von Mengen A_1, \dots, A_n
$[a]$	Äquivalenzklasse von $a \in A$
$A \rightarrow B, a \mapsto b$	Abbildung von A nach B
$\text{Abb}(A, B)$	Menge aller Abbildungen $A \rightarrow B$
$\mathbb{N}, \mathbb{N}_0, \mathbb{Z}$	natürliche Zahlen (ohne und mit 0), ganze Zahlen
$\mathbb{Q}, \mathbb{R}, \mathbb{R}_{>0}, \mathbb{R}_{\geq 0}, \mathbb{C}$	rationale, reelle (positive, nicht-negative) und komplexe Zahlen
\mathbb{F}_2	Körper mit zwei Elementen
K^\times	$= K \setminus \{0\}$ (multiplikative Gruppe)
$f _A, f^{-1}, f \circ g$	Einschränkung, Umkehrfunktion und Komposition von Funktionen
f^*	duale oder zu f adjungierte Abbildung
$f(A), f^{-1}(B), \text{Ker}(f)$	Bild, Urbild, Kern von f
id, id_V	Identität (auf V)
$0_K, 0_V, 1_G$	Nullelement, Nullvektor, neutrales Element in G
δ_{ij}	Kronecker-Delta
$\sum_{i=1}^n \lambda_i v_i$	Linearkombination
$U \leq V, U < V, V/U$	Unterraum (echter), Faktorraum
$H \leq G, H < G$	Untergruppe (echte) von G
$U \cong V$	isomorphe Vektorräume
V^*, V^{**}	Dualraum, Bidualraum

U^0, U_0	duale Komplemente
$U \oplus W, U \times W$	direkte Summe/Produkt von U und W
e_1, \dots, e_n	Standardbasis von K^n
b_1^*, \dots, b_n^*	duale Basis
$\langle S \rangle, \langle s_1, \dots, s_n \rangle$	Spann von S
$\dim_K V = \dim V$	Dimension von V über K
${}_B[v]$	Koordinatendarstellung von v bzgl. der Basis B
$\text{Hom}(V, W)$	Vektorraum aller linearen Abbildungen $V \rightarrow W$
$\text{End}(V)$	$= \text{Hom}(V, V)$
$K^{n \times m}$	Vektorraum der $n \times m$ -Matrizen über K
$\text{GL}(V), \text{GL}(n, K)$	allgemeine lineare Gruppe
$\text{SL}(n, K)$	spezielle lineare Gruppe
$\text{O}(V), \text{O}(n, K)$	orthogonale Gruppe
$\text{SO}(n, K)$	spezielle orthogonale Gruppe
$\text{U}(V), \text{U}(n, \mathbb{C})$	unitäre Gruppe
$\text{SU}(n, \mathbb{C})$	spezielle unitäre Gruppe
$\text{Aff}(V)$	affine Gruppe
$0_{n \times m}, 0_n, 1_n$	Nullmatrix, Einheitsmatrix
E_{st}	Standardmatrix mit 1 an Position (s, t)
$A \sim B$	A zeilen-äquivalent zu B
$A \approx B$	A ähnlich zu B
$(A b)$	erweiterte Koeffizientenmatrix
A_{st}	$= (a_{ij} : i \neq s, j \neq t)$ (Streichen von Zeile/Spalte)
A^{-1}, A^+	Inverse und Pseudoinverse von A
A^t, A^{-t}	Transponierte und Transponiert-Inverse von A
\hat{A}, \tilde{A}	Zeilenstufenform, komplementäre Matrix von A
\bar{A}, A^*	komplex-konjugierte, adjungierte Matrix von A
$\text{rk}(A), \text{tr}(A), \det(A)$	Rang, Spur, Determinante von A
$\text{vol}(D)$	Volumen (Jordan-Maß) von $A \subseteq \mathbb{R}^n$
$C[f]_B, [f], C\Delta_B$	Darstellungsmatrix, Basiswechselmatrix
$E_\lambda(f), H_\lambda(f)$	Eigenraum, Hauptraum zum Eigenwert λ von f
S_n	symmetrische Gruppe vom Grad n
A_n	alternierende Gruppe vom Grad n
$\text{sgn}(\sigma), P_\sigma$	Signum, Permutationsmatrix von $\sigma \in S_n$
$K[X]$	Vektorraum der Polynome mit Koeffizienten in K
$K(X)$	Körper der rationalen Funktionen
α'	Ableitung von $\alpha \in K[X]$
$\deg(\alpha)$	Grad von $\alpha \in K[X]$
$\alpha \mid \beta$	α teilt β
$\alpha \equiv \beta \pmod{\delta}$	$\delta \mid \alpha - \beta$
$\chi_A, \chi_f, \mu_A, \mu_f$	charakteristisches Polynom, Minimalpolynom von A, f
μ_v	Minimalpolynom des zyklischen Unterraums
$[v, w], v $	Skalarprodukt, euklidische Norm von v
$v \perp w$	v und w sind orthogonal, d. h. $[v, w] = 0$
π	Länge des Halbkreisbogens mit Radius 1
$\cos \varphi, \sin \varphi$	Kosinus, Sinus von φ
S^\perp	orthogonales Komplement von $S \subseteq V$
$v \times w$	Kreuzprodukt von v und w
$D(\varphi), S(\varphi)$	Drehung, Spiegelung in \mathbb{R}^2

$D_{st}(\phi)$	Givens-Rotation
S_v	Spiegelung an der Hyperebene v^\perp
$\operatorname{Re}(z), \operatorname{Im}(z), \bar{z}$	Realteil, Imaginärteil, komplexe Konjugation von z
i	imaginäre Einheit
$\operatorname{Bil}(V)$	Vektorraum der Bilinearformen auf V
$B[\beta]_B$	Gram-Matrix einer Bilinearform β
$\operatorname{ind}(\beta), \operatorname{ind}(A)$	Index von $\beta \in \operatorname{Bil}(V), A \in K^{n \times n}$
$J_n(\lambda)$	Jordanblock für $\lambda \in K$ der Größe $n \times n$
$J_k(\gamma)$	verallgemeinerter Jordanblock für $\gamma \in K[X]$
$B(\alpha)$	Begleitmatrix von $\alpha \in K[X]$
$C(f), C(A)$	Zentralisator von $f \in \operatorname{End}(V), A \in K^{n \times n}$
W_n	n -te Fourier-Matrix
H_n	n -te Hilbert-Matrix
\mathcal{F}_n	diskrete Fourier-Transformation
$\kappa(A)$	Konditionszahl von A
$ A $	Frobenius-Norm von A
$\ v\ _p, \ A\ _p$	p -(Matrix-)Norm
$\ v\ _{\max}, \ A\ _{\max}$	Maximum-Norm, Zeilensummen-Norm
$\lambda_k(A)$	Der k . größte Eigenwert von $A = A^*$
$\rho(A)$	Spektralradius von A
$\exp(A)$	Exponentialfunktion von A
A_+	nicht-negative Matrix $(a_{ij})_{ij}$
L^*	zu L duales lineares Programm
$\operatorname{con}(\Delta)$	Konvexe Hülle von $\Delta \subseteq \mathbb{R}^n$
$\operatorname{supp}(x)$	Träger von $x \in \mathbb{R}^n$
$\lfloor x \rfloor, \lceil x \rceil$	abgerundet und gerundeter Wert von $x \in \mathbb{R}$
$\operatorname{disc}(\Delta)$	Diskriminante des Gitters Δ
Δ^*	zu Δ duales Gitter
E_8	Ganzes Gitter in \mathbb{R}^8 mit vollem Rang
$\Delta \perp \Lambda$	orthogonale Zerlegung von Gittern
γ_n	n -te Hermite-Konstante
gT, ggT	(größter) gemeinsame(r) Teiler
$\det(q)$	Determinante der quadratischen Form q
$\mu_i(q)$	Sukzessives Minimum von q

Konventionen

- Eigennamen sind bei erster Verwendung in KAPITÄLCHEN geschrieben.
- K ist stets ein Körper, V ein endlich-dimensionaler K -Vektorraum, Unterräume heißen meist U, W, V_1 etc.
- Mengen und Matrizen werden mit lateinischen Großbuchstaben bezeichnet (A, B, \dots, M, \dots).
- Elemente von Mengen werden mit Kleinbuchstaben bezeichnet, Vektoren mit u, v, w , natürliche Zahlen mit n, m, k, l , Abbildungen mit f, g, h etc.
- Für Mengen von Mengen benutzt man oft „geschwungene“ Buchstaben (\mathcal{M}, \mathcal{P})

- Für Polynome, Skalare (Körper Elemente im Kontext von Vektorräumen) und Bilinearformen verwenden wir griechische Buchstaben. Die gebräuchlichsten sind:

α	β	γ, Γ	δ, Δ	ϵ, ε	ζ	η	$\theta, \vartheta, \Theta$	λ, Λ	μ
alpha	beta	gamma	delta	epsilon	zeta	eta	theta	lambda	my
ν	ξ	π, Π	ρ, ϱ	σ, Σ	τ	φ, ϕ, Φ	χ	ψ, Ψ	ω, Ω
ny	xi	pi	rho	sigma	tau	phi	chi	psi	omega

Lineare Algebra I

1 Aussagenlogik und Mengenlehre

1.1 Aussagen

Bemerkung 1.1. Die Sprache der Mathematik basiert auf logischen Prinzipien, die man letztlich als gegeben hinnehmen muss. Alle „höheren“ mathematischen Objekte lassen sich auf mengentheoretische Konstrukte zurückführen. Wir behandeln diese Themen hier nur so weit, wie sie zum Verständnis der linearen Algebra benötigt werden. Mehr Informationen findet man in meinem Skript zur Logik und Mengenlehre.

Definition 1.2.

- Eine *Aussage* A ist ein deutscher Satz, der entweder den *Wahrheitswert wahr* (**w**) oder *falsch* (**f**) annimmt. Man sagt dann A *gilt* bzw. A *gilt nicht*.
- Für Aussagen A und B sind auch $\neg A$ (*nicht A*), $A \wedge B$ (*A und B*), $A \vee B$ (*A oder B*), $A \Rightarrow B$ (*A impliziert B*) und $A \Leftrightarrow B$ (*A genau dann wenn B*) Aussagen mit folgenden Wahrheitswerten:

A	B	$\neg A$	$A \wedge B$	$A \vee B$	$A \Rightarrow B$	$A \Leftrightarrow B$
w	w	f	w	w	w	w
w	f	f	f	w	f	f
f	w	w	f	w	w	f
f	f	w	f	f	w	w

- Zwei Aussagen A und B nennt man *äquivalent*, falls $A \Leftrightarrow B$ wahr ist, d. h. wenn A und B den gleichen Wahrheitswert haben.
- Ein *Prädikat* ist eine Eigenschaft $A = A(x)$, die erst durch Einsetzen einer Variablen x zu einer Aussage wird. Ggf. sind $\forall x : A(x)$ (*für alle x gilt $A(x)$*) und $\exists x : A(x)$ (*es existiert ein x , sodass $A(x)$ gilt*) Aussagen.

Beispiel 1.3. Folgende Sätze sind Aussagen (selbst wenn wir den Wahrheitswert nicht kennen):

- Alle blauen Katzen können fliegen (**w**).
- $1 + 1 = 3$ (**f**).
- Jede gerade Zahl größer als 2 ist die Summe zweier Primzahlen (?).¹

Keine Aussagen dagegen sind:

- Sei $\epsilon > 0$ (*Annahme*).
- $a^2 + b^2 = c^2$ (*Gleichung*).

¹GOLDBACHS Vermutung

- Dieser Satz ist falsch (*Paradoxon*).

Aus dem Prädikat $x > 0$ kann man die wahre Aussage $\forall x > 4 : x > 0$ bilden.

Bemerkung 1.4.

- (a) Im Gegensatz zum alltäglichen Sprachgebrauch unterscheidet sich das mathematische *oder* vom *entweder oder*. Das heißt, die Aussage $\mathbf{w} \vee \mathbf{w}$ ist wahr. Wir führen kein eigenständiges Symbol für *entweder oder* ein.² Unterscheiden Sie außerdem die Formulierungen „Es gibt ein...“ und „Es gibt genau ein...“.
- (b) Die Wahrheit der Aussage $\mathbf{f} \Rightarrow \mathbf{f}$ irritiert viele Anfänger (siehe Beispiel 1.3). Interpretation: Wenn die Voraussetzung nicht erfüllt ist, ist auch nichts zu zeigen. Man unterscheide außerdem die Aussage $A \Rightarrow B$ von ihrer *Umkehrung* $B \Rightarrow A$.³
- (c) Für Aussagen A_1, \dots, A_n definiert man $A_1 \wedge \dots \wedge A_n$ durch $\forall i : A_i$ und $A_1 \vee \dots \vee A_n$ durch $\exists i : A_i$.
- (d) Um den Wahrheitswert einer Aussage A zu bestimmen, führt man *Äquivalenzumformungen* durch, d. h. man ersetzt A durch eine äquivalente Aussage. Dafür sind folgende Schlussregeln nützlich.⁴

Lemma 1.5. *Seien A, B und C Aussagen. Dann gilt:*

- (a) *Die folgenden Aussagen sind äquivalent zu A :*

$$\neg\neg A, \quad A \wedge \mathbf{w}, \quad A \vee \mathbf{f}, \quad A \wedge A, \quad A \vee A, \quad \mathbf{w} \Rightarrow A$$

- (b) $A \wedge B$ und $B \wedge A$ sind äquivalent sowie $A \vee B$ und $B \vee A$ (Kommutativgesetz).
- (c) Es gilt $A \vee \neg A$ (Satz vom ausgeschlossenen Dritten) und $\neg(A \wedge \neg A)$ (Satz vom Widerspruch).
- (d) $A \wedge (B \vee C)$ und $(A \wedge B) \vee (A \wedge C)$ sind äquivalent sowie $A \vee (B \wedge C)$ und $(A \vee B) \wedge (A \vee C)$ (Distributivgesetz).
- (e) $\neg(A \wedge B)$ und $\neg A \vee \neg B$ sind äquivalent sowie $\neg(A \vee B)$ und $\neg A \wedge \neg B$ (DE MORGANSche Regeln).
- (f) $A \Rightarrow B$, $\neg A \vee B$ und $(\neg B) \Rightarrow (\neg A)$ sind äquivalent (Kontraposition).
- (g) $(A \Rightarrow B) \wedge (B \Rightarrow C)$ impliziert $A \Rightarrow C$ (Transitivität).
- (h) Aus $A \wedge (A \Rightarrow B)$ folgt B (Modus ponens).
- (i) $A \Leftrightarrow B$ ist äquivalent zu $(A \Rightarrow B) \wedge (B \Rightarrow A)$.

Beweis. Alle Behauptungen lassen sich leicht durch Wahrheitstabellen verifizieren. Bei drei Variablen muss man dafür $2^3 = 8$ Fälle unterscheiden (wer einen schnelleren Weg findet, kann eine Million Dollar verdienen⁵). Alternativ kann man einige der Behauptungen aus bereits bewiesenen ableiten. So folgt die zweite De Morgansche Regel aus der ersten:

$$\neg(A \vee B) \stackrel{(a)}{\iff} \neg((\neg\neg A) \vee (\neg\neg B)) \iff \neg(\neg(\neg A \wedge \neg B)) \stackrel{(a)}{\iff} (\neg A \wedge \neg B). \quad \square^6$$

²In der Informatik spricht man von XOR.

³Man könnte auch $A \Leftarrow B$ schreiben.

⁴Ein Lemma ist ein Hilfssatz mit wenig eigener Bedeutung.

⁵Das SAT-Problem der theoretischen Informatik ist NP-vollständig. Eines der sieben Millenniumsprobleme fragt, ob $P = NP$.

⁶Diese Box markiert das Ende eines Beweises.

Bemerkung 1.6.

- (a) Die De Morganschen Regeln lassen sich allgemeiner in der Form $(\neg\forall x : A(x)) \Leftrightarrow (\exists x : (\neg A(x)))$ und $(\neg\exists x : A(x)) \Leftrightarrow (\forall x : (\neg A(x)))$ für Prädikate formulieren.
- (b) Lemma 1.5 zeigt, dass man allein mit den Symbolen \neg und \wedge alle weiteren Terme ausdrücken kann. Zu Gunsten der Lesbarkeit sollte man jedoch alle Symbole sparsam einsetzen.

1.2 Mengen

Definition 1.7 (CANTOR). Eine *Menge* M ist eine Zusammenfassung von bestimmten wohlunterschiedenen Objekten x unserer Anschauung oder unseres Denkens zu einem Ganzen.⁷ Man sagt dann: x ist ein *Element* von M und schreibt $x \in M$ sowie $M = \{x : x \in M\}$ (bzw. $x \notin M$ für $\neg(x \in M)$). Die Anzahl $|M|$ der Elemente von M heißt *Kardinalität* oder *Mächtigkeit* von M . Im Fall $|M| < \infty$ heißt M *endlich* und anderenfalls *unendlich*.

Bemerkung 1.8.

- (a) Definition 1.7 ist ungenau, denn sie lässt Mengen zu, die zu logischen Widersprüchen führen. Sei beispielsweise

$$M := \{x : x \notin x\} \qquad \text{(RUSSELLsche Antinomie)}^8$$

Die Aussage $M \in M$ kann dann weder wahr noch falsch sein. In der modernen Mathematik verhindert man solche Widersprüche durch Einführung eines *Axiomensystems*, d. h. man gibt den Wahrheitswert von möglichst wenigen „elementaren“ Aussagen (Axiomen) vor. Weit verbreitet ist das ZERMELO-FRAENKEL-System. Eines seiner Axiome besagt:

Mengen sind genau dann gleich, wenn sie die gleichen Elemente enthalten.

Dies impliziert, dass die Elemente einer Menge keine feste Reihenfolge haben. Es gilt also $\{2, 1, 1, 2, 2\} = \{1, 2\}$.

- (b) In manchen Situationen benötigt man zusätzlich das sogenannte *Auswahlaxiom* (siehe Beispiel 2.3). Es wird von den meisten Mathematikern anerkannt, obwohl es die Konstruktion kontraintuitiver Mengen zulässt: Das BANACH-TARSKI-*Paradoxon* besagt beispielsweise, dass man eine Kugel vom Volumen 1 in fünf Teile zerlegen kann, die anders zusammengesetzt zwei Kugeln vom Volumen 1 ergeben.
- (c) Nach GÖDELS zweitem *Unvollständigkeitssatz* ist es unmöglich zu beweisen, dass die Zermelo-Fraenkel-Axiome keine Widersprüche liefern.⁹ Ist dies tatsächlich der Fall (wovon die meisten Mathematiker ausgehen), so besagt Gödels erster Unvollständigkeitssatz, dass es Aussagen gibt, deren Wahrheitswert sich nicht bestimmen lässt.¹⁰ Das bekannteste Beispiel hierfür ist die *Kontinuumshypothese* (siehe Bemerkung 2.11(f)).

⁷Cantors Wortlaut

⁸Das Symbol $:=$ besagt, dass die linke Seite durch die rechte Seite festgelegt wird.

⁹Siehe Skript zur Logik und Mengenlehre

¹⁰Die Beweisidee besteht darin die Aussage „Dieser Satz ist nicht beweisbar.“ zu formalisieren.

Definition 1.9.

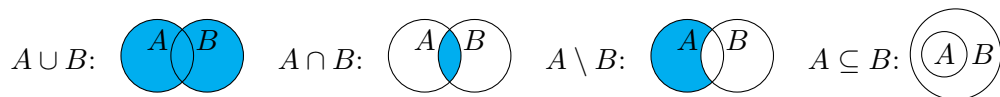
(a) Für Mengen A und B sei

$$\begin{aligned} \emptyset &:= \{\} && (\text{leere Menge}), \\ A \cup B &:= \{x : x \in A \vee x \in B\} && (\text{Vereinigung}), \\ A \cap B &:= \{x : x \in A \wedge x \in B\} && (\text{Durchschnitt}),^{11} \\ A \setminus B &:= \{x : x \in A \wedge x \notin B\} && (\text{Differenz}).^{12} \end{aligned}$$

- (b) Im Fall $A \cup B = B$ ist A eine *Teilmenge* von B . Man schreibt dann $A \subseteq B$ oder $A \subsetneq B$, falls zusätzlich $A \neq B$ (man spricht dann von einer *echten* Teilmenge¹³). Ist A keine Teilmenge von B , so schreibt man $A \not\subseteq B$.
- (c) Man nennt A und B *disjunkt*, falls $A \cap B = \emptyset$. Ggf. nennt man $A \dot{\cup} B := A \cup B$ eine *disjunkte Vereinigung*.

Bemerkung 1.10.

(a) Beziehungen zwischen Mengen lassen sich durch VENN-Diagramme veranschaulichen:



Achtung: Sind mehr als drei Mengen im Spiel, so kann die allgemeine Situation nicht mehr durch Kreise dargestellt werden.¹⁴

(b) Vereinigung und Durchschnitt von beliebig vielen Mengen A_i (wobei i aus einer Indexmenge I stammt) lassen sich wie folgt definieren:

$$\bigcup_{i \in I} A_i := \{x : \exists i \in I : x \in A_i\}, \quad \bigcap_{i \in I} A_i := \{x : \forall i \in I : x \in A_i\}.$$

Ist A die disjunkte Vereinigung von Mengen A_i , so spricht man von einer *Partition* von A .

(c) Um die Gleichheit von Mengen $A = B$ zu beweisen, ist es oft einfacher die äquivalente Aussage $(A \subseteq B) \wedge (B \subseteq A)$ zu zeigen.

Beispiel 1.11.

- (a) Die Menge der *natürlichen Zahlen* $\mathbb{N} := \{1, 2, 3, \dots\}$. Wir setzen $\mathbb{N}_0 := \{0, 1, 2, \dots\} = \mathbb{N} \cup \{0\}$.¹⁵
Achtung: Bei manchen Autoren ist $0 \in \mathbb{N}$.
- (b) Die Menge der *ganzen Zahlen* $\mathbb{Z} := \{\dots, -2, -1, 0, 1, 2, \dots\}$. Es gilt $\mathbb{N} = \{n \in \mathbb{Z} : n > 0\}$. Die ganzen Zahlen der Form $2n$ (bzw. $2n + 1$) mit $n \in \mathbb{Z}$ heißen *gerade* (bzw. *ungerade*).
- (c) Die Menge der *rationalen Zahlen* $\mathbb{Q} := \{\frac{a}{b} : a, b \in \mathbb{Z}, b \neq 0\}$.

¹¹Man beachte die Ähnlichkeit der Symbole \cup und \vee sowie \cap und \wedge .

¹²In manchen Büchern schreibt man $A - B$ anstatt $A \setminus B$.

¹³Das Symbol \subset wird in der Literatur leider nicht einheitlich benutzt.

¹⁴siehe <https://de.wikipedia.org/wiki/Mengendiagramm>

¹⁵Streng axiomatisch definiert man $0 := \emptyset$, $1 := \{\emptyset\}$ und allgemein $n + 1 := n \cup \{n\}$.

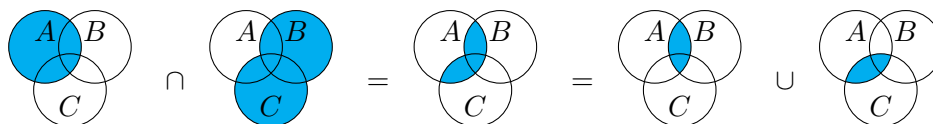
- (d) Die Menge der *reellen Zahlen* \mathbb{R} besteht aus allen *Dezimalbrüchen* wie $2 = 2.0$, $\frac{1}{3} = 0.33\dots$, $\sqrt{2} = 1.4142\dots$ oder $\pi = 3.1415\dots$ (die Dezimalbruchentwicklung kann abbrechend, periodisch oder unperiodisch sein). In der Analysis definiert man reelle Zahlen als Grenzwerte von rationalen CAUCHY-Folgen. Im Folgenden setzen wir die üblichen Regeln für die Grundrechenarten voraus. Auch diese lassen sich streng axiomatisch einführen.
- (e) Es gilt $\mathbb{N} \subsetneq \mathbb{N}_0 \subsetneq \mathbb{Z} \subsetneq \mathbb{Q} \subsetneq \mathbb{R}$. Die Behauptung $\mathbb{Q} \neq \mathbb{R}$ zeigen wir indirekt. Annahme: $\mathbb{Q} = \mathbb{R}$. Dann ist $\sqrt{2} \in \mathbb{Q}$ und es existieren $a, b \in \mathbb{Z}$ mit $\sqrt{2} = \frac{a}{b}$ und $b \neq 0$. Ohne Beschränkung der Allgemeinheit (kurz: o. B. d. A.) können wir annehmen, dass a und b teilerfremd sind (anderenfalls kann man $\frac{a}{b}$ kürzen). Umstellen ergibt $2b^2 = a^2$. Insbesondere ist a^2 gerade. Da das Quadrat einer ungeraden Zahl ungerade ist ($(2n+1)^2 = 2(2n^2+2n)+1$), ist a gerade, sagen wir $a = 2c$. Es folgt $b^2 = 2c^2$. Mit dem gleichen Argument ist nun auch b gerade. Also ist 2 ein gemeinsamer Teiler von a und b . Dieser Widerspruch zeigt, dass die Annahme falsch war. Also ist $\mathbb{Q} \neq \mathbb{R}$.
- (f) Die Elemente einer Menge können durchaus selbst Mengen sein. In solchen Fällen benutzt man oft geschwungene Buchstaben. Zum Beispiel besteht $\mathcal{M} := \{\{1, 2\}, \{1, 3\}, \{2, 3\}\}$ aus allen 2-elementigen Teilmengen von $\{1, 2, 3\}$.

Lemma 1.12. Für Mengen A, B und C gilt:

- (a) $A \cap B \subseteq A \subseteq A \cup B$.
- (b) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ und $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ (Distributivgesetz).
- (c) $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$ und $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$ (De Morgansche Regeln).
- (d) $|A \cup B| + |A \cap B| = |A| + |B|$ und $|A \dot{\cup} B| = |A| + |B|$.

Beweis.

- (a) Folgt direkt aus der Definition.
- (b) Wir beweisen nur die erste Gleichheit (beweisen Sie die zweite selbst):



- (c) Diesmal benutzen wir Lemma 1.5 (für die erste Gleichung):

$$\begin{aligned} x \in A \setminus (B \cup C) &\iff (x \in A \wedge (x \notin B \cup C)) \iff (x \in A \wedge (x \notin B \wedge x \notin C)) \\ &\iff ((x \in A \wedge x \notin B) \wedge (x \in A \wedge x \notin C)) \iff x \in (A \setminus B) \cap (A \setminus C). \end{aligned}$$

- (d) Ist A oder B unendlich, so auch $A \cup B$ und die Behauptung gilt, wenn man $\infty + n = \infty$ für $n \in \mathbb{N}_0 \cup \{\infty\}$ interpretiert. Seien nun A und B endlich, sagen wir $A \cap B = \{x_1, \dots, x_s\}$, $A = \{x_1, \dots, x_s, a_1, \dots, a_t\}$ und $B = \{x_1, \dots, x_s, b_1, \dots, b_u\}$. Dann gilt

$$|A \cup B| + |A \cap B| = s + t + u + s = |A| + |B|.$$

Sind A und B disjunkt, so gilt $|A \cap B| = |\emptyset| = 0$ und die zweite Behauptung folgt. \square

Definition 1.13. Die *Potenzmenge* $\mathcal{P}(M)$ einer Menge M ist die Menge aller Teilmengen von M , d. h.

$$\mathcal{P}(M) := \{N : N \subseteq M\}.$$

1.3 Vollständige Induktion

Satz 1.14 (Prinzip der vollständigen Induktion). Sei $A(n)$ ein Prädikat für $n \in \mathbb{N}$ mit den Eigenschaften:

- Induktionsanfang: $A(1)$ gilt.
- Induktionsschritt: $\forall n \in \mathbb{N} : (A(n) \implies A(n+1))$.

Dann gilt $A(n)$ für alle $n \in \mathbb{N}$.

Beweis. Beweis durch Widerspruch: Gilt $A(n)$ nicht für alle $n \in \mathbb{N}$, so gibt es ein kleinstes n mit $\neg A(n)$. Nach dem Induktionsanfang ist $n \neq 1$. Nach Wahl von n gilt $A(n-1)$. Nach dem Induktionsschritt gilt $A(n-1) \implies A(n)$. Also gilt $A(n)$ nach Modus ponens. Widerspruch. \square

Bemerkung 1.15. Man verwendet oft Varianten der vollständigen Induktion. Zum Beispiel:

- Induktionsanfang: $A(1) \wedge A(2)$ gilt.
- Induktionsschritt: $\forall n \in \mathbb{N} : ((A(n) \wedge A(n+1)) \implies A(n+2))$.

Beispiel 1.16. Wir beweisen $(1 + 2 + \dots + n)^2 = 1^3 + 2^3 + \dots + n^3$ für alle $n \in \mathbb{N}$.

Induktionsanfang: Für $n = 1$ gilt $1^2 = 1 = 1^3$.

Induktionsvoraussetzung: Es gelte bereits $(1 + 2 + \dots + n)^2 = 1^3 + 2^3 + \dots + n^3$ (*).

Induktionsschritt: Wir müssen die Behauptung für $n + 1$ beweisen. Zunächst eine Nebenrechnung:

$$\begin{aligned} 2(1 + 2 + \dots + n) &= (1 + 2 + \dots + n) + (n + (n-1) + \dots + 1) \\ &= (1 + n) + (2 + n-1) + \dots + (n+1) = n(n+1) \end{aligned}$$

(das hat GAUSS als 9-Jähriger erkannt¹⁶). Nach der binomischen Formel gilt nun

$$\begin{aligned} ((1 + 2 + \dots + n) + (n+1))^2 &= (1 + 2 + \dots + n)^2 + 2(1 + 2 + \dots + n)(n+1) + (n+1)^2 \\ &\stackrel{(*)}{=} 1^3 + 2^3 + \dots + n^3 + n(n+1)(n+1) + (n+1)^2 \\ &= 1^3 + 2^3 + \dots + n^3 + (n+1)^3. \end{aligned} \quad \square$$

¹⁶siehe Wikipedia

2 Kartesische Produkte und Funktionen

2.1 Paare und Tupel

Bemerkung 2.1. Nach Bemerkung 1.8 sind die Elemente einer Menge ungeordnet. Wir führen eine geordnete Variante ein.

Definition 2.2.

- Seien A und B Mengen. Das *kartesische Produkt* von A und B ist die Menge $A \times B$ bestehend aus allen (*geordneten*) *Paaren*¹ (a, b) mit $a \in A, b \in B$, sodass gilt

$$(a, b) = (a', b') \iff (a = a' \wedge b = b').$$

Es gilt $|A \times B| = |A||B|$, sofern man die Regeln $\infty \cdot 0 = 0$ und $\infty \cdot n = \infty$ für $n \in \mathbb{N} \cup \{\infty\}$ benutzt.

- Analog definiert man *Tripel* (a, b, c) und n -*Tupel* (a_1, \dots, a_n) für $n \geq 2$. Für Mengen A_1, \dots, A_n setzt man

$$A_1 \times \dots \times A_n := \{(a_1, \dots, a_n) : a_1 \in A_1, \dots, a_n \in A_n\}.$$

Gilt $A := A_1 = \dots = A_n$, so benutzt man die Abkürzung $A^n := A_1 \times \dots \times A_n$.

- Kartesische Produkte lassen sich auch für beliebige Familien von Mengen definieren. Sei I eine Indexmenge und $(A_i : i \in I)$ eine Familie von Mengen. Man definiert $\times_{i \in I} A_i := \{(a_i)_{i \in I} : \forall i \in I : a_i \in A_i\}$.

Beispiel 2.3.

(a) Das kartesische Produkt $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ besteht aus allen Koordinaten in der 2-dimensionalen Ebene.

(b) Es gilt

$$\{1, 2\} \times \{2, 3, 4\} = \{(1, 2), (1, 3), (1, 4), (2, 2), (2, 3), (2, 4)\}.$$

(c) Das kartesische Produkt $\times_{i \in \mathbb{N}} \mathbb{R}$ ist die Menge aller reellen Folgen aus der Analysis.

(d) Sei $(A_i : i \in I)$ eine beliebige Familie von nicht-leeren Mengen. Das bereits erwähnte *Auswahlaxiom* besagt $\times_{i \in I} A_i \neq \emptyset$, d. h. man kann aus jeder Menge A_i *gleichzeitig* ein Element auswählen.

Definition 2.4. Eine *Relation* auf einer nicht-leeren Menge A ist eine Teilmenge $R \subseteq A \times A$. Üblicherweise wählt man ein Symbol, zum Beispiel \sim , und schreibt $a \sim b$, falls $(a, b) \in R$. Man nennt R

- *reflexiv*, falls $\forall a \in A : a \sim a$.
- *symmetrisch*, falls $\forall a, b \in A : (a \sim b \Rightarrow b \sim a)$.

¹Die formale Definition von Paaren kann man auf Mengen zurückführen: $(a, b) := \{\{a\}, \{a, b\}\}$.

- *antisymmetrisch*, falls $\forall a, b \in A : ((a \sim b \wedge b \sim a) \Rightarrow a = b)$.²
- *transitiv*, falls $\forall a, b, c \in A : ((a \sim b \wedge b \sim c) \Rightarrow a \sim c)$.
- *Äquivalenzrelation*, falls R reflexiv, symmetrisch und transitiv ist.
- (*partielle*) *Ordnungsrelation*, falls R reflexiv, antisymmetrisch und transitiv ist.

Ist R eine Äquivalenzrelation auf der Menge A und $a \in A$, so nennt man $[a] := \{b \in A : a \sim b\} \subseteq A$ die *Äquivalenzklasse* von a .

Beispiel 2.5.

- Die *triviale* Relation $R = A \times A$ ist eine (uninteressante) Äquivalenzrelation.
- Die Gleichheitsrelation $\{(a, a) : a \in A\}$ mit dem Symbol $=$ ist die „kleinste“ reflexive Relation auf A . Trivialerweise handelt es sich um eine Äquivalenzrelation.³ Man kann viele weitere Äquivalenzrelationen auf die Gleichheit zurückführen. Sei beispielsweise A die Menge aller Menschen und $a \sim b$ falls $a, b \in A$ im gleichen Land leben. Die Äquivalenzklassen entsprechen dann den Ländern.
- Man kann durch einfache Beispiele zeigen, dass die Eigenschaften reflexiv, symmetrisch und transitiv unabhängig voneinander sind. Zum Beispiel ist die Relation

$$\{(1, 1), (2, 2), (3, 3), (1, 2), (2, 1), (2, 3), (3, 2)\}$$

auf $A = \{1, 2, 3\}$ reflexiv und symmetrisch, aber nicht transitiv.

- Auf \mathbb{R} ist die Kleinergleichrelation \leq eine Ordnungsrelation. Sie hat zusätzlich die Eigenschaft, dass je zwei Zahlen a und b in Relation stehen, d. h. es gilt $a \leq b$ oder $b \leq a$ (man spricht von einer *totalen* Ordnungsrelation).
- Auf der Potenzmenge jeder Menge A ist die Inklusionsrelation \subseteq eine Ordnungsrelation. Im Fall $A = \mathbb{N}$ stehen $\{1\}$ und $\{2\}$ nicht in Relation ($\{1\} \not\subseteq \{2\} \not\subseteq \{1\}$). Im Gegensatz zu \leq ist \subseteq also nicht total.

Lemma 2.6. *Sei R eine Äquivalenzrelation auf einer Menge A . Dann existiert eine Teilmenge $T \subseteq A$, sodass die Äquivalenzklassen $[t]$ mit $t \in T$ eine Partition von A bilden, d. h. $A = \dot{\bigcup}_{t \in T} [t]$.*

Beweis. Sei \sim das Symbol von R . Seien $a, b \in A$ und $c \in [a] \cap [b]$. Dann gilt $a \sim c$ und $b \sim c$. Da \sim symmetrisch ist, gilt $c \sim b$. Da \sim transitiv ist, gilt $a \sim b$. Für jedes $d \in [b]$ gilt also $a \sim b \sim d$ und $a \sim d$. Dies zeigt $[b] \subseteq [a]$ und analog erhält man $[a] \subseteq [b]$. Es folgt $[a] = [b]$. Somit sind je zwei Äquivalenzklassen entweder gleich oder disjunkt. Die Existenz von T folgt nun aus dem Auswahlaxiom. \square

Bemerkung 2.7.

- In der Situation von Lemma 2.6 nennt man T ein *Repräsentantensystem* für die Äquivalenzklassen.
- Ist $A = \dot{\bigcup}_{i \in I} A_i$ eine Partition von A , so definiert

$$a \sim b :\iff \exists i \in I : a, b \in A_i$$

eine Äquivalenzrelation auf A (nachrechnen). Daher entsprechen sich Partitionen und Äquivalenzrelationen.

²Achtung: Es gibt auch die stärkere Eigenschaft *asymmetrisch*: $\forall a, b \in A : (a \sim b \Rightarrow b \not\sim a)$.

³Für symmetrische Relationen sollte man „symmetrische“ Symbole wählen. Die Umkehrung ist leider nicht immer gegeben, z. B. das Symbol $|$ für die antisymmetrische Teilbarkeitsrelation auf \mathbb{N} .

Beispiel 2.8. Für die Äquivalenzklasse auf der Menge aller Menschen aus Beispiel 2.5 bilden die Präsidenten jedes Landes ein Repräsentantensystem.

2.2 Injektive und surjektive Funktionen

Definition 2.9.

- Seien A und B Mengen. Eine *Funktion* oder *Abbildung* f von A nach B ist eine Vorschrift, die jedem $a \in A$ genau ein $f(a) \in B$ zuordnet.⁴ Man schreibt dann⁵

$$f: A \rightarrow B, \quad a \mapsto f(a).$$

Die Menge aller Abbildungen $A \rightarrow B$ bezeichnen wir mit $\text{Abb}(A, B)$.

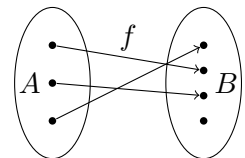
- Man nennt A den *Definitionsbereich* und B den *Wertebereich* von f . Außerdem ist $f(a)$ das *Bild* von a unter f und $f(A) := \{f(a) : a \in A\} \subseteq B$ ist das *Bild* von f . Für $B' \subseteq B$ ist

$$f^{-1}(B') := \{a \in A : f(a) \in B'\} \subseteq A$$

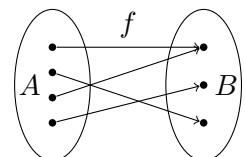
das *Urbild* von B' unter f .

- Man nennt $f: A \rightarrow B$

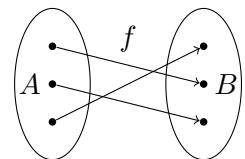
- *injektiv*, falls $\forall a, a' \in A : (f(a) = f(a') \implies a = a')$.



- *surjektiv*, falls $\forall b \in B : \exists a \in A : f(a) = b$, d. h. $f(A) = B$.



- *bijektiv* (oder *Bijektion*), falls f injektiv und surjektiv ist.
Ggf. nennt man A und B *gleichmächtig*.



- Die *Einschränkung* von $f: A \rightarrow B$ auf eine Teilmenge $A' \subseteq A$ ist die Funktion

$$f|_{A'}: A' \rightarrow B, \quad a \mapsto f(a).$$

Für eine weitere Funktion $g: B \rightarrow C$ nennt man die Abbildung

$$g \circ f: A \rightarrow C, \quad a \mapsto g(f(a))$$

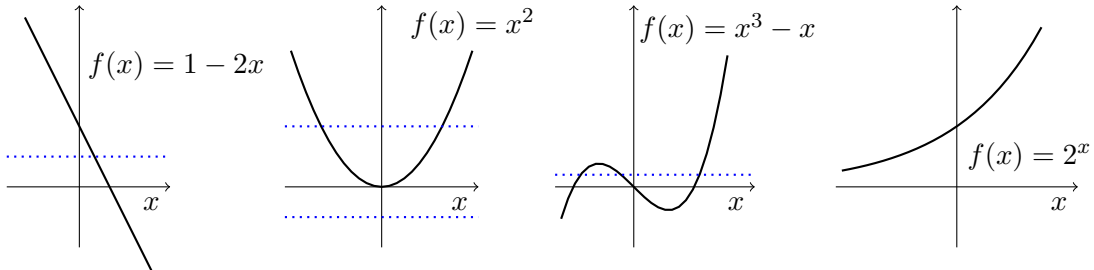
die *Komposition* (oder *Hintereinanderausführung*, *Verkettung*) von f und g .

⁴Formal: Eine Funktion ist eine Teilmenge $f \subseteq A \times B$, sodass für jedes $a \in A$ genau ein $b \in B$ mit $(a, b) \in f$ existiert.

⁵Man beachte die unterschiedlichen Pfeile \rightarrow und \mapsto .

Beispiel 2.10.

- (a) Für jede Menge A und $B \subseteq A$ ist $f: B \rightarrow A, b \mapsto b$ eine injektive Funktion, die man *Inklusionsabbildung* nennt. Im Fall $B = A$ ist f sogar bijektiv und man nennt $f = \text{id}_A$ die *Identität* auf A .
- (b) Abbildungen $f: \mathbb{R} \rightarrow \mathbb{R}$ lassen sich grafisch darstellen:



Injektiv (bzw. surjektiv) bedeutet, dass der Graph von f jede horizontale Gerade höchstens (bzw. mindestens) einmal schneidet. Wir lesen ab:

Funktion	injektiv	surjektiv	bijektiv
$f(x) = 1 - 2x$	✓	✓	✓
$f(x) = x^2$	✗	✗	✗
$f(x) = x^3 - x$	✗	✓	✗
$f(x) = 2^x$	✓	✗	✗

Bemerkung 2.11.

- (a) Sind A und B endliche Mengen, so gilt $|\text{Abb}(A, B)| = |B|^{|A|}$, denn für $f: A \rightarrow B$ und jedes $a \in A$ hat man $|B|$ Möglichkeiten $f(a) \in B$ zu wählen.
- (b) Achtung: Injektiv ist nicht das Gegenteil von surjektiv (ein häufiger Anfängerfehler)!
- (c) Man kann jede Funktion $f: A \rightarrow B$ surjektiv machen, indem man auf den Wertebereich auf das Bild einschränkt: $f: A \rightarrow f(A)$.
- (d) Für $f: A \rightarrow B$ gilt

$$\begin{aligned}
 f \text{ injektiv} &\implies |A| = |f(A)| \leq |B| \\
 f \text{ surjektiv} &\implies |B| = |f(A)| \leq |A| \\
 f \text{ bijektiv} &\implies |A| = |B|
 \end{aligned}$$

(wobei $\infty \leq \infty$).

- (e) Zwei endliche Mengen A und B sind genau dann gleichmächtig, wenn $|A| = |B|$. In diesem Fall sind die Eigenschaften injektiv, surjektiv und bijektiv nach (d) äquivalent. Für unendliche Mengen ist dies im Allgemeinen falsch (Beispiel 2.10).
- (f) Obwohl \mathbb{N} nur „halb so viele“ Zahlen wie \mathbb{Z} enthält, sind \mathbb{N} und \mathbb{Z} durch die Bijektion

$$\mathbb{N} \rightarrow \mathbb{Z}, \quad n \mapsto \begin{cases} \frac{n-1}{2} & \text{falls } n \text{ ungerade,} \\ -\frac{n}{2} & \text{falls } n \text{ gerade} \end{cases}$$

gleichmächtig. Eine Menge, die gleichmächtig zu \mathbb{N} ist, nennt man *abzählbar*⁶. Nach Cantors *Diagonalisierungsargumenten* ist \mathbb{Q} abzählbar, aber \mathbb{R} nicht, d. h. \mathbb{R} ist *überabzählbar*. Die (nicht beweisbare) *Kontinuumshypothese* besagt, dass jede unendliche Teilmenge von \mathbb{R} entweder zu \mathbb{N} oder zu \mathbb{R} gleichmächtig ist. Der folgende Satz zeigt, dass es beliebig „große“ Mengen gibt (*Kardinalzahlen*), die man in der Praxis aber selten antrifft.

Satz 2.12 (CANTOR). *Jede Menge M ist „kleiner“ als ihre Potenzmenge, d. h. es existiert eine injektive Abbildung $M \rightarrow \mathcal{P}(M)$, aber keine Bijektion. Ist M endlich, so gilt $|\mathcal{P}(M)| = 2^{|M|}$.*

Beweis. Die Abbildung $M \rightarrow \mathcal{P}(M)$, $a \mapsto \{a\}$ ist sicher injektiv. Nehmen wir an es existiert eine Bijektion $f: M \rightarrow \mathcal{P}(M)$. Sei

$$A := \{x \in M : x \notin f(x)\} \in \mathcal{P}(M).$$

Dann existiert ein $a \in M$ mit $f(a) = A$. Es folgt der Widerspruch $a \in A = f(a) \iff a \notin f(a)$. Für die zweite Behauptung sei $M = \{x_1, \dots, x_n\}$. Dann ist die Abbildung

$$\mathcal{P}(M) \rightarrow \{0, 1\}^n, \quad A \mapsto (a_1, \dots, a_n)$$

mit $a_i = 1 \iff x_i \in A$ eine Bijektion. Also ist $|\mathcal{P}(M)| = |\{0, 1\}^n| = 2^n = 2^{|M|}$. □

Lemma 2.13. *Seien $f: A \rightarrow B$, $g: B \rightarrow C$, $h: C \rightarrow D$ Funktionen. Dann gilt:*

- (a) $(h \circ g) \circ f = h \circ (g \circ f)$ (Assoziativgesetz).
- (b) Sind f und g injektiv, so auch $g \circ f$.
- (c) Sind f und g surjektiv, so auch $g \circ f$.
- (d) Ist $g \circ f$ injektiv, so auch f .
- (e) Ist $g \circ f$ surjektiv, so auch g .
- (f) Genau dann ist f bijektiv, wenn eine Funktion $g: B \rightarrow A$ mit $g \circ f = \text{id}_A$ und $f \circ g = \text{id}_B$ existiert. Ggf. ist g eindeutig bestimmt und man nennt $f^{-1} := g$ die Umkehrfunktion von f .

Beweis.

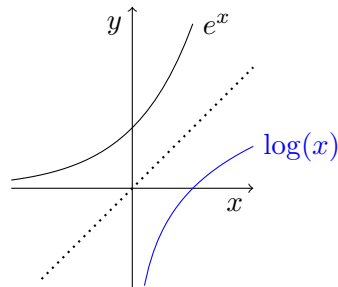
- (a) Für $a \in A$ ist $((h \circ g) \circ f)(a) = (h \circ g)(f(a)) = h(g(f(a))) = h((g \circ f)(a)) = (h \circ (g \circ f))(a)$.
- (b) Für $a, a' \in A$ mit $(g \circ f)(a) = (g \circ f)(a')$ gilt $g(f(a)) = g(f(a'))$, also $f(a) = f(a')$ und $a = a'$.
- (c) Es gilt $(g \circ f)(A) = g(f(A)) = g(B) = C$.
- (d) Sei $f(a) = f(a')$ für $a, a' \in A$. Dann ist $(g \circ f)(a) = g(f(a)) = g(f(a')) = (g \circ f)(a')$. Da $g \circ f$ injektiv ist, folgt $a = a'$.
- (e) Es gilt $C = (g \circ f)(A) = g(f(A)) \subseteq g(B) \subseteq C$, also $g(B) = C$.
- (f) Ist $g \circ f = \text{id}_A$ und $f \circ g = \text{id}_B$, so ist f injektiv nach (d) und surjektiv nach (e), also auch bijektiv. Sei umgekehrt f bijektiv. Für jedes $b \in B$ existiert dann genau ein $g(b) \in A$ mit $f(g(b)) = b$. Daher ist $g: B \rightarrow A$ die einzige Abbildung mit $f \circ g = \text{id}_B$. Aus $f(a) = f(g(f(a)))$ folgt $g(f(a)) = a$ für alle $a \in A$, da f injektiv ist. Dies zeigt $g \circ f = \text{id}_A$. □

⁶In manchen Büchern zählen die endlichen Mengen auch zu den abzählbaren Mengen.

Bemerkung 2.14. Verwechseln Sie die Umkehrfunktion nicht mit dem Urbild. Der Zusammenhang beider Konzepte ist $f^{-1}(\{b\}) = \{f^{-1}(b)\}$ für jede Bijektion $f: A \rightarrow B$ und $b \in B$.

Beispiel 2.15.

- (a) Die Abbildung $f: \mathbb{Q} \rightarrow \mathbb{Q}$, $x \mapsto 2x + 1$ ist eine Bijektion mit Umkehrabbildung $f^{-1}: \mathbb{Q} \rightarrow \mathbb{Q}$, $x \mapsto \frac{x-1}{2}$ (nachrechnen).
- (b) Die Umkehrabbildung der *Exponentialfunktion* $\exp: \mathbb{R} \rightarrow \mathbb{R}_{>0}$, $x \mapsto e^x$ ist der *natürliche Logarithmus* $\log: \mathbb{R}_{>0} \rightarrow \mathbb{R}$. Man erhält den Graphen von \log durch Spiegelung an der Geraden $y = x$:



Man beachte, dass die reine Existenz der Umkehrfunktion noch lange keine konkrete Formel für $f^{-1}(x)$ liefert. Diesen Umstand macht man sich in der Kryptographie zunutze (*Einwegfunktion*).

3 Körper und Vektorräume

3.1 Gruppen und Körper

Bemerkung 3.1. In fast allen Anwendungen der linearen Algebra wird nur von den vier Grundrechenarten (Addition, Subtraktion, Multiplikation und Division) Gebrauch gemacht. Damit man nicht jede Aussage für jeden Zahlbereich ($\mathbb{Q}, \mathbb{R}, \dots$) neu beweisen muss, ersetzt man Zahlbereiche durch abstrakte *Gruppen* (mit einer Operation) und *Körper* (mit zwei Operationen). Zur Beschreibung von Lösungsmengen von Gleichungssystemen führt man *Vektorräume* ein. Beachten Sie, dass dies lediglich Modelle zur Untersuchung linearer Probleme sind, die sich im Laufe der Zeit bewährt haben (so wie metrische Räume in der Analysis oder das Bohrsche Atommodell in der Chemie).

Definition 3.2. Eine *Verknüpfung* \cdot auf einer Menge G ist eine Abbildung $G \times G \rightarrow G$, $(x, y) \mapsto x \cdot y$. Man nennt das Paar (G, \cdot) (oder auch nur G) eine *Gruppe*, falls

- $\forall x, y, z \in G : (x \cdot y) \cdot z = x \cdot (y \cdot z)$ (*Assoziativgesetz*),
- $\exists e \in G : (\forall x \in G : e \cdot x = x = x \cdot e)$ (*neutrales Element*),
- $\forall x \in G : (\exists y \in G : y \cdot x = e = x \cdot y)$ (*inverses Element*).

Gilt zusätzlich

- $\forall x, y \in G : x \cdot y = y \cdot x$ (*Kommutativgesetz*),

so heißt G *abelsch*.¹

Bemerkung 3.3. Sei G eine Gruppe mit neutralem Element e .

- (a) Aus Bequemlichkeit schreiben wir oft xy anstatt $x \cdot y$.
- (b) Ist auch $e' \in G$ ein neutrales Element, so gilt $e' = e' \cdot e = e$. Also ist e eindeutig bestimmt und wir schreiben oft $e = 1_G = 1$ oder $e = 0_G = 0$, falls die Verknüpfung $+$ ist.
- (c) Seien $y, y' \in G$ invers zu $x \in G$. Dann ist

$$y' = y'e = y'(xy) = (y'x)y = ey = y.$$

Somit hat x genau ein Inverses und wir schreiben $y = x^{-1}$ oder $y = -x$, falls die Verknüpfung $+$ ist. Im letzten Fall schreiben wir $x - y := x + (-y)$ für beliebige $x, y \in G$.²

- (d) Für $x, y \in G$ ist $\boxed{(x^{-1})^{-1} = x}$ und $\boxed{(xy)^{-1} = y^{-1}x^{-1}}$ (Achtung Reihenfolge!).

¹Benannt nach N. ABEL.

²In nicht-abelschen Gruppen ist die Schreibweise $\frac{x}{y}$ problematisch, denn es könnte sowohl xy^{-1} als auch $y^{-1}x$ gemeint sein.

Beispiel 3.4.

- (a) Wegen $e \in G$ ist eine Gruppe niemals leer. Andererseits gibt es die *triviale* Gruppe $G = \{e\}$.
- (b) Nach den üblichen Rechenregeln sind $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$ und $(\mathbb{R}, +)$ abelsche Gruppen mit neutralem Element 0. Andererseits ist $(\mathbb{Z}, -)$ *keine* Gruppe, denn das Assoziativgesetz ist verletzt:

$$(1 - 2) - 3 = -4 \neq 2 = 1 - (2 - 3).$$

Ebenso besitzt $(\mathbb{N}, +)$ kein neutrales Element und in $(\mathbb{N}_0, +)$ hat nicht jedes Element ein Inverses (z. B. $-1 \notin \mathbb{N}_0$).

- (c) Offenbar sind $(\mathbb{Q} \setminus \{0\}, \cdot)$ und $(\mathbb{R} \setminus \{0\}, \cdot)$ abelsche Gruppen mit neutralem Element 1, aber nicht $(\mathbb{Z} \setminus \{0\}, \cdot)$, denn $2^{-1} = \frac{1}{2} \notin \mathbb{Z}$.
- (d) Für Gruppen G_1, \dots, G_n ist auch $G_1 \times \dots \times G_n$ eine Gruppe mit

$$(x_1, \dots, x_n) \cdot (y_1, \dots, y_n) := (x_1 y_1, \dots, x_n y_n)$$

für $(x_1, \dots, x_n), (y_1, \dots, y_n) \in G_1 \times \dots \times G_n$ (Aufgabe I.8). Das neutrale Element ist $(1_{G_1}, \dots, 1_{G_n})$. Man spricht dann vom *direkten Produkt* von G_1, \dots, G_n (anstelle vom kartesischen Produkt).

Definition 3.5. Ein *Körper* ist eine Menge K mit Verknüpfungen $+$ und \cdot , sodass folgende Eigenschaften gelten:

- $(K, +)$ ist eine abelsche Gruppe mit neutralem Element 0.
- $(K \setminus \{0\}, \cdot)$ ist eine abelsche Gruppe mit neutralem Element 1. Man setzt $K^\times := K \setminus \{0\}$.
- $\forall x, y, z \in K : x \cdot (y + z) = (x \cdot y) + (x \cdot z)$ (*Distributivgesetz*).

Bemerkung 3.6. Im Folgenden sei K stets ein Körper.

- (a) Durch die Vereinbarung „Punktrechnung vor Strichrechnung“ sparen wir Klammern ein. Zum Beispiel sei $xy + z := (x \cdot y) + z$ für $x, y, z \in K$.
- (b) Für alle $x \in K$ gilt $x \cdot 0 = 0 = 0 \cdot x$, denn $x0 = x(0 + 0) = x0 + x0$. Es folgt $(-x)y = -(xy)$ für $x, y \in K$.
- (c) Für $x, y, z \in K$ und $z \neq 0$ gilt die *Kürzungsregel* $xz = yz \implies x = y$, denn

$$x = x \cdot 1 = x(zz^{-1}) = (xz)z^{-1} = (yz)z^{-1} = \dots = y.$$

Beispiel 3.7.

- (a) Nach den gewohnten Rechenregeln sind \mathbb{Q} und \mathbb{R} Körper. Es gibt außerdem unendlich viele Körper „zwischen“ \mathbb{Q} und \mathbb{R} (vgl. Aufgabe I.15). Andererseits ist $(\mathbb{Z}, +, \cdot)$ kein Körper, da $(\mathbb{Z} \setminus \{0\}, \cdot)$ keine Gruppe ist.
- (b) Jeder Körper besitzt mindestens die beiden Elemente 0 und 1. Tatsächlich ist $\mathbb{F}_2 = \{0, 1\}$ bereits ein Körper, wenn man $1 + 1 := 0$ definiert. Die Verknüpfungstabellen sind dadurch vollständig bestimmt:

$+$	0	1	\cdot	0	1
0	0	1	0	0	0
1	1	0	1	0	1

Auf Computern werden alle Rechnungen in \mathbb{F}_2 durchgeführt, indem man 0 und 1 als *Bits* interpretiert. In der Algebra³ konstruiert man für jede Primzahlpotenz q ein Körper mit genau q Elementen (vgl. Aufgabe I.9).

3.2 Vektorräume und Unterräume

Definition 3.8. Ein *Vektorraum* V über einem Körper K (kurz: K -Vektorraum) ist eine abelsche Gruppe bzgl. $+$ zusammen mit einer *Skalarmultiplikation* $K \times V \rightarrow V$, $(\lambda, v) \mapsto \lambda \cdot v$ mit folgenden Eigenschaften:

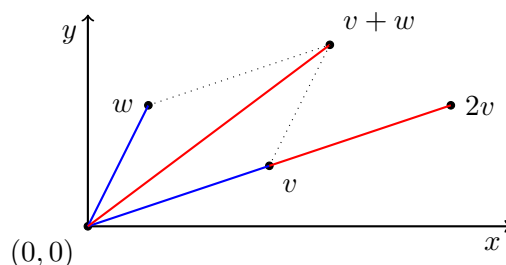
- $\forall v \in V : 1 \cdot v = v$,
- $\forall v, w \in V, \lambda \in K : \lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w$,
- $\forall v \in V, \lambda, \mu \in K : (\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v$,
- $\forall v \in V, \lambda, \mu \in K : (\lambda \cdot \mu) \cdot v = \lambda \cdot (\mu \cdot v)$.

Die Elemente von V heißen *Vektoren* und die Elemente in K *Skalare* (in diesem Kontext). Das neutrale Element 0_V in V nennt man den *Nullvektor*.

Bemerkung 3.9. Man beachte, dass $+$ sowohl die Addition in K als auch in V bezeichnet. Ebenso steht \cdot für die Multiplikation in K und für die Skalarmultiplikation (das ist ungenau, aber durchaus üblich). Wir werden in beiden Fällen das Symbol \cdot oft einsparen. Falls Missverständnisse ausgeschlossen sind, schreiben wir auch 0 anstatt 0_V . Sie müssen im Zweifel in der Lage sein zu entscheiden, ob das Nullelement in K oder V gemeint ist.

Beispiel 3.10.

- (a) Der *Nullraum* $V = \{0_V\}$ mit der Skalarmultiplikation $\lambda \cdot 0_V := 0_V$ für alle $\lambda \in K$.
- (b) Für K -Vektorräume V_1, \dots, V_n ist auch das direkte Produkt (bzgl. $+$) $V_1 \times \dots \times V_n$ ein Vektorraum mit komponentenweiser Skalarmultiplikation: $\lambda(v_1, \dots, v_n) := (\lambda v_1, \dots, \lambda v_n)$ für $v_i \in V_i$ und $\lambda \in K$ (nachrechnen).
- (c) Offenbar ist K selbst ein Vektorraum, in dem die Skalarmultiplikation mit der gewöhnlichen Multiplikation übereinstimmt. Nach (b) ist auch K^n für $n \geq 1$ ein Vektorraum. In \mathbb{R}^2 lassen sich Vektoraddition und Skalarmultiplikation geometrisch deuten:



³siehe Algebra-Skript

- (d) Sind v_1, \dots, v_n Vektoren aus V und $\lambda_1, \dots, \lambda_n \in K$, so liegt auch die *Linearkombination* $\lambda_1 v_1 + \dots + \lambda_n v_n$ in V (Nachweis durch Induktion nach n). Man benutzt dafür das Summenzeichen:

$$\sum_{i=1}^n \lambda_i v_i := \lambda_1 v_1 + \dots + \lambda_n v_n.$$

Sind v_1, \dots, v_n paarweise verschieden (d.h. $v_i \neq v_j$ für $i \neq j$)⁴ und mindestens ein $\lambda_i \neq 0$, so nennt man die Linearkombination *nicht-trivial*. Manchmal tritt die *leere Summe* ohne Summanden auf. Diese wird stets als 0_V interpretiert. Zum Beispiel $\sum_{i=1}^0 v_i = 0$. Sei auch $v_i = \mu_{i1} w_{i1} + \mu_{i2} w_{i2} + \dots + \mu_{im} w_{im}$ eine Linearkombination für $i = 1, \dots, n$.⁵ Dann erhält man eine *Doppelsumme*:

$$\sum_{i=1}^n v_i = \sum_{i=1}^n \sum_{j=1}^m \mu_{ij} w_{ij}.$$

Da $(V, +)$ abelsch ist, darf man die Summanden beliebig umordnen und somit die Summenzeichen vertauschen:

$$\sum_{i=1}^n \sum_{j=1}^m \mu_{ij} w_{ij} = \sum_{j=1}^m \sum_{i=1}^n \mu_{ij} w_{ij}.$$

Ein Vorzug der Algebra gegenüber der Analysis ist, dass alle Summen endlich sind und man keine Konvergenzbetrachtungen anstellen muss.

Satz 3.11. Sei $A \neq \emptyset$ eine Menge und V ein K -Vektorraum. Dann ist $\text{Abb}(A, V)$ mit den folgenden Verknüpfungen ein K -Vektorraum:

$$\begin{aligned} (f + g)(a) &:= f(a) + g(a) & (f, g \in \text{Abb}(M, V), a \in M) \\ (\lambda f)(a) &:= \lambda f(a) & (\lambda \in K) \end{aligned}$$

Beweis. Offenbar liegen $f + g$ und λf in $\text{Abb}(A, V)$. Die triviale Abbildung $f(a) = 0$ für alle $a \in A$ ist das neutrale Element bzgl. $+$. Für $f: A \rightarrow V$ ist $-f: A \rightarrow V, a \mapsto -f(a)$ invers zu f bzgl. $+$. Für $f, g, h: A \rightarrow V$ und $a \in A$ ist

$$\begin{aligned} ((f + g) + h)(a) &= (f + g)(a) + h(a) = (f(a) + g(a)) + h(a) = f(a) + (g(a) + h(a)) \\ &= f(a) + (g + h)(a) = (f + (g + h))(a). \end{aligned}$$

Daher ist $+$ assoziativ. Auf die gleiche Weise übertragen sich die verbleibenden Vektorraumaxiome von V nach $\text{Abb}(A, V)$. \square

Definition 3.12.

- Eine Teilmenge H einer Gruppe G heißt *Untergruppe*, falls H mit der eingeschränkten Verknüpfung selbst eine Gruppe ist, d. h.
 - $1_G \in H$,
 - $\forall g, h \in H : gh \in H$,
 - $\forall h \in H : h^{-1} \in H$.

⁴„paarweise verschieden“ ist stärker als die Formulierung „nicht alle sind gleich“

⁵Im Zweifel sollte man doppelte Indizes durch Komma trennen $\mu_{i,1}$.

Wir schreiben ggf. $H \leq G$. Im Fall $H \neq G$ nennt man H eine *echte* Untergruppe und schreibt $H < G$.

- Eine Teilmenge U eines Vektorraums V heißt *Unterraum*, falls U mit den eingeschränkten Verknüpfungen selbst einen Vektorraum ist, d. h.
 - $(U, +)$ ist eine Untergruppe von $(V, +)$,
 - $\forall v \in U, \lambda \in K : \lambda v \in U$.

Wir schreiben dann $U \leq V$ wie bei Untergruppen. Im Fall $U \neq V$ nennt man U einen *echten* Unterraum und schreibt $U < V$.

Bemerkung 3.13.

- (a) Die Bedingungen garantieren, dass H unter Multiplikation bzw. U unter Addition und Skalarmultiplikation *abgeschlossen* ist. Somit sind die Verknüpfungen auf H bzw. U *wohldefiniert*. Die verbleibenden Gruppenaxiome bzw. Vektorraumaxiome muss man nicht prüfen, da sie bereits in der größeren Menge G bzw. V gelten.
- (b) Im Folgenden beschränken wir uns auf die Untersuchung von Unterräumen. Die meisten Aussagen gelten sinngemäß auch für (abelsche) Gruppen.
- (c) Für Vektorräume kann man die Bedingungen wie folgt zusammenfassen: Eine *nichtleere* Teilmenge $U \subseteq V$ ist genau dann ein Unterraum, wenn für alle $u, v \in U$ und $\lambda \in K$ gilt: $\lambda u + v \in U$ (Aufgabe I.10).

Beispiel 3.14.

- (a) Jeder Vektorraum V besitzt die Unterräume $\{0_V\}$ und V .
- (b) Aus $U \leq W \leq V$ folgt $U \leq V$. Aus $U, W \leq V$ und $U \subseteq W$ folgt sicher auch $U \leq W$.
- (c) Der Durchschnitt von beliebig vielen Unterräumen ist wieder ein Unterraum (nachrechnen).
- (d) Wir beweisen $U := \{(x, 0) : x \in \mathbb{R}\} \leq \mathbb{R}^2$ mit Hilfe von Bemerkung 3.13: Wegen $(0, 0) \in U$ ist $U \neq \emptyset$. Für $(x_1, 0), (x_2, 0) \in U$ und $\lambda \in \mathbb{R}$ gilt

$$\lambda(x_1, 0) + (x_2, 0) = (\lambda x_1, 0) + (x_2, 0) = (\lambda x_1 + x_2, 0) \in U.$$

Geometrisch entspricht U der x -Achse in der Ebene. Analog ist die xy -Ebene

$$U := \{(x, y, 0) \in \mathbb{R}^3 : x, y \in \mathbb{R}\}$$

ein Unterraum von \mathbb{R}^3 .

- (e) Die Teilmenge $U := \{(x, x^2) : x \in \mathbb{Q}\}$ von \mathbb{Q}^2 ist *kein* Unterraum, denn $(1, 1) \in U$, aber $2 \cdot (1, 1) = (2, 2) \notin U$. Wir zeigen in Bemerkung 7.19, dass sich jeder Unterraum durch *lineare* Gleichungen beschreiben lässt.
- (f) Offenbar sind $U := \{(0, 0), (1, 0)\}$ und $W := \{(0, 0), (0, 1)\}$ Unterräume von \mathbb{F}_2^2 , aber nicht $U \cup W$ (Warum?)

Lemma 3.15. Sei V ein K -Vektorraum und $U \leq V$. Für $v \in V$ sei $v + U := \{v + u : u \in U\} \subseteq V$. Dann wird $V/U := \{v + U : v \in V\}$ mit

$$\begin{aligned}(v + U) + (w + U) &:= (v + w) + U & (v, w \in V), \\ \lambda(v + U) &:= \lambda v + U & (\lambda \in K)\end{aligned}$$

zu einem K -Vektorraum.

Beweis. Sei $\bar{v} := v + U$ für $v \in V$. Für $\bar{v} = \bar{v}'$ und $\bar{w} = \bar{w}'$ gilt

$$\overline{v + w} = v + w + U = v + w' + U = w' + v + U = w' + v' + U = \overline{v' + w'}.$$

Dies zeigt, dass die Addition auf V/U wohldefiniert ist. Das neutrale Element ist $0 + U = U$. Für $\lambda \in K$ gilt analog

$$\overline{\lambda v} = \lambda v + U = \lambda v + \lambda U = \lambda(v + U) = \lambda(v' + U) = \lambda v' + U = \overline{\lambda v'}.$$

Also ist auch die Skalarmultiplikation wohldefiniert. Die Vektorraumaxiome für V/U folgen unmittelbar aus den Axiomen für V . \square

Bemerkung 3.16. Man nennt V/U den *Faktorraum* von V nach U . Die Mengen $v + U$ werden manchmal als *affine Räume* bezeichnet (Aufgabe III.25). Sie sind die Äquivalenzklassen der Relation $v \sim w : \iff v - w \in U$.

4 Basen und Dimension

4.1 Lineare Unabhängigkeit und Erzeugendensysteme

Bemerkung 4.1. Um unendlich große Vektorräume vergleichen zu können, führen wir die Dimension als feinere Kenngröße ein. Es wird sich zeigen, dass Vektorräume allein durch ihre Dimension weitestgehend bestimmt sind (Satz 7.10).

Definition 4.2. Sei V ein Vektorraum.

- (a) Für $S \subseteq V$ sei $\langle S \rangle \subseteq V$ die Menge aller Linearkombinationen von Elementen aus S . Man nennt $\langle S \rangle$ den *Spann* von S .¹ Im Fall $S = \{s_1, \dots, s_n\}$ schreiben wir auch $\langle s_1, \dots, s_n \rangle$ anstatt $\langle S \rangle$ (d. h. wir sparen die Mengenklammern ein).
- (b) Für Unterräume $U, W \leq V$ sei

$$U + W := \{u + w : u \in U, w \in W\} \subseteq V$$

die (MINKOWSKI-)Summe von U und W . Im Fall $U \cap W = \{0\}$ nennt man die Summe *direkt* und schreibt $U \oplus W$ anstatt $U + W$.²

Lemma 4.3. Sei V ein Vektorraum, $S \subseteq V$ und $U, W \leq V$. Dann sind $\langle S \rangle$ und $U + W$ Unterräume von V .

Beweis. Offenbar ist 0 eine Linearkombination von Elementen aus S , d. h. $0 \in \langle S \rangle$ (im Fall $S = \emptyset$ wähle man die leere Summe). Addition und Skalarmultiplikation von Linearkombinationen sind wieder Linearkombinationen. Dies zeigt $\langle S \rangle \leq V$. Wegen $0 \in U \cap W$ ist $0 = 0 + 0 \in U + W$. Seien $u_1 + w_1, u_2 + w_2 \in U + W$ und $\lambda \in K$. Dann gilt

$$\lambda(u_1 + w_1) + (u_2 + w_2) = \underbrace{(\lambda u_1 + u_2)}_{\in U} + \underbrace{(\lambda w_1 + w_2)}_{\in W} \in U + W.$$

Also ist auch $U + W \leq V$. □

Beispiel 4.4.

- (a) Es gilt $\langle \emptyset \rangle = \{0\}$, denn die leere Summe ist die einzige Linearkombination aus \emptyset .
- (b) Für $U \leq W \leq V$ gilt $U + W = W$ und $\langle U \rangle = U = U \oplus \{0\}$.
- (c) Für $s_1, \dots, s_n \in V$ gilt $\langle s_i \rangle = \{\lambda s_i : \lambda \in K\} =: Ks_i$ und $\langle s_1, \dots, s_n \rangle = Ks_1 + \dots + Ks_n$. Insbesondere ist $\mathbb{R}^2 = \mathbb{R}(1, 0) \oplus \mathbb{R}(0, 1)$.

¹In manchen Büchern schreibt man $\text{Span}(S)$ anstatt $\langle S \rangle$.

²Dies ersetzt die (disjunkte) Vereinigung, siehe Aufgabe I.12.

Definition 4.5. Eine Teilmenge S eines Vektorraums V heißt

- *Erzeugendensystem*, falls $\langle S \rangle = V$. Im Fall $|S| < \infty$ nennt man V *endlich erzeugt*.
- *linear abhängig*, falls 0_V eine nicht-triviale Linearkombination von Elementen aus S ist.
- *linear unabhängig*, falls nicht linear abhängig, d. h. für paarweise verschiedene Elemente $s_1, \dots, s_n \in S$ und $\lambda_1, \dots, \lambda_n \in K$ gilt:

$$\sum_{i=1}^n \lambda_i s_i = 0 \quad \implies \quad \lambda_1 = \dots = \lambda_n = 0.$$

- *Basis*, falls S ein linear unabhängiges Erzeugendensystem ist.

Bemerkung 4.6. Da Basen Mengen sind, besitzen ihre Elemente keine feste Anordnung. Tatsächlich hängen aber viele Sätze von der Reihenfolge der Basiselemente ab. Wir führen daher folgende Sprechweise ein: Vektoren s_1, \dots, s_n heißen linear unabhängig (bzw. bilden eine Basis), falls sie paarweise verschieden sind und $\{s_1, \dots, s_n\}$ linear unabhängig (bzw. eine Basis) ist.

Beispiel 4.7.

- Die leere Menge ist stets linear unabhängig und bildet eine Basis des Nullraums.
- Wegen $1_K \cdot 0_V = 0_V$ ist der Nullvektor niemals Bestandteil einer linear unabhängigen Menge. Ein einzelner Vektor $v \neq 0$ ist hingegen stets linear unabhängig, denn aus $\lambda v = 0$ mit $\lambda \in K^\times$ folgt der Widerspruch

$$v = 1v = (\lambda^{-1}\lambda)v = \lambda^{-1}(\lambda v) = \lambda^{-1}0 = 0.$$

- Vektoren $v, w \in V \setminus \{0\}$ sind genau dann linear abhängig, wenn $Kv = Kw$, d. h. v ist ein skalares Vielfache von w und umgekehrt.
- Jede Teilmenge einer linear unabhängigen Menge ist linear unabhängig.
- Für $n \geq 1$ seien

$$\begin{aligned} e_1 &:= (1, 0, \dots, 0), \\ e_2 &:= (0, 1, 0, \dots, 0), \\ &\vdots \\ e_n &:= (0, \dots, 0, 1) \end{aligned}$$

Vektoren aus K^n . Da sich jeder Vektor $v = (v_1, \dots, v_n) \in K^n$ in der Form $v = \sum_{i=1}^n v_i e_i$ schreiben lässt, ist $\{e_1, \dots, e_n\}$ ein Erzeugendensystem von K^n . Aus $v = 0 \iff v_1 = \dots = v_n = 0$ folgt die lineare Unabhängigkeit von $\{e_1, \dots, e_n\}$. Man nennt e_1, \dots, e_n die *Standardbasis* von K^n .

4.2 Charakterisierung und Existenz von Basen

Satz 4.8. Sei b_1, \dots, b_n eine Basis eines K -Vektorraums V . Dann lässt sich jedes $v \in V$ eindeutig in der Form $v = \sum_{i=1}^n \lambda_i b_i$ mit $\lambda_1, \dots, \lambda_n \in K$ schreiben. Insbesondere ist die Abbildung

$$B[\cdot]: V \rightarrow K^n, \quad v \mapsto B[v] := (\lambda_1, \dots, \lambda_n)$$

eine Bijektion.

Beweis. Wegen $V = \langle b_1, \dots, b_n \rangle$ ist jedes $v \in V$ eine Linearkombination der angegebenen Form. Seien $\lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_n \in K$ mit

$$v = \sum_{i=1}^n \lambda_i b_i = \sum_{i=1}^n \mu_i b_i.$$

Dann ist $0 = v - v = \sum_{i=1}^n (\lambda_i - \mu_i) b_i$. Da $\{b_1, \dots, b_n\}$ linear unabhängig ist, folgt $\lambda_i = \mu_i$ für $i = 1, \dots, n$. \square

Definition 4.9. In der Situation von Satz 4.8 nennt man $B[v]$ die *Koordinatendarstellung* von v bzgl. B .

Lemma 4.10. Für einen Vektorraum V und $B \subseteq V$ sind äquivalent:

- (1) B ist eine Basis von V .
- (2) B ist ein minimales Erzeugendensystem, d. h. für alle $b \in B$ ist $B \setminus \{b\}$ kein Erzeugendensystem.
- (3) B ist maximal linear unabhängig, d. h. für alle $v \in V \setminus B$ ist $B \cup \{v\}$ linear abhängig.

Beweis. Wir führen einen Ringbeweis.³

(1) \Rightarrow (2): Sei B eine Basis, also insbesondere ein Erzeugendensystem von V . Nehmen wir an, dass auch $B \setminus \{b\}$ für ein $b \in B$ ein Erzeugendensystem ist. Dann existieren $\lambda_1, \dots, \lambda_n \in K$ und $b_1, \dots, b_n \in B \setminus \{b\}$ mit $b = \sum_{i=1}^n \lambda_i b_i$. Wegen $-b + \sum_{i=1}^n \lambda_i b_i = 0$ wäre B dann linear abhängig. Widerspruch.

(2) \Rightarrow (3): Sei B ein minimales Erzeugendensystem. Sei $\sum_{i=1}^n \lambda_i b_i = 0$ für $\lambda_1, \dots, \lambda_n \in K$ und paarweise verschiedene $b_1, \dots, b_n \in B$. Ist $\lambda_i \neq 0$ für ein i , so gilt

$$b_i = -\lambda_i^{-1} \sum_{j \neq i} \lambda_j b_j = \sum_{j \neq i} (-\lambda_i^{-1} \lambda_j) b_j \in \langle B \setminus \{b_i\} \rangle.$$

Dann wäre aber auch $B \setminus \{b_i\}$ ein Erzeugendensystem. Also ist $\lambda_1 = \dots = \lambda_n = 0$ und B ist linear unabhängig. Sei nun $v \in V \setminus B$. Wegen $\langle B \rangle = V$ existieren $\lambda_1, \dots, \lambda_n \in K$ und $b_1, \dots, b_n \in B$ mit $v = \sum_{i=1}^n \lambda_i b_i$ und $-v + \sum_{i=1}^n \lambda_i b_i = 0$. Insbesondere ist $B \cup \{v\}$ linear abhängig.

(3) \Rightarrow (1): Sei B maximal linear unabhängig. Wir müssen $\langle B \rangle = V$ zeigen. Sei $v \in V$. Im Fall $v \in B$ ist $v \in \langle B \rangle$. Sei also $v \notin B$. Dann ist $B \cup \{v\}$ linear abhängig. Also existieren $\lambda_1, \dots, \lambda_n \in K^\times$, $\mu \in K$, $b_1, \dots, b_n \in B$ mit $\mu v + \sum_{i=1}^n \lambda_i b_i = 0$. Da B linear unabhängig ist, muss $\mu \neq 0$ gelten. Dies liefert

$$v = -\mu^{-1} \sum_{i=1}^n \lambda_i b_i = \sum_{i=1}^n (-\mu^{-1} \lambda_i) b_i \in \langle B \rangle.$$

Insgesamt ist $V = \langle B \rangle$. \square

³Ein *Zirkelschluss* hingegen ist eine fehlerhafte Argumentation, bei der die Behauptung bereits vorausgesetzt wird.

Satz 4.11 (Basisergänzungssatz). *Sei V ein Vektorraum mit einem endlichen Erzeugendensystem $E \subseteq V$. Dann lässt sich jede linear unabhängige Menge $U \subseteq V$ durch Hinzunahme von Elementen aus E zu einer Basis von V ergänzen.*

Beweis. Sei $E = \{s_1, \dots, s_n\}$. Im Fall $E \subseteq \langle U \rangle$ ist $V = \langle E \rangle \subseteq \langle U \rangle$, d. h. U ist bereits eine Basis. Sei also $E \not\subseteq \langle U \rangle$ und o. B. d. A. $s_1 \notin \langle U \rangle$. Wie üblich ist dann $U_1 := U \cup \{s_1\}$ linear unabhängig. Wir können nun das Argument mit U_1 anstelle von U wiederholen. Im Fall $E \subseteq \langle U_1 \rangle$ ist U_1 eine Basis und anderenfalls können wir $s_2 \notin \langle U_1 \rangle$ annehmen. Dann ist $U_2 := U_1 \cup \{s_2\}$ linear unabhängig usw. Da E endlich ist, erhält man nach endlich vielen Schritten eine Basis von V . \square

Beispiel 4.12. Die linear unabhängige Menge $U := \{(1, 2, 0), (2, 1, 0)\} \subseteq \mathbb{R}^3$ lässt sich mit dem Standardbasisvektor e_3 zu einer Basis ergänzen (aber nicht mit e_1 oder e_2).

Satz 4.13 (STEINITZER Austauschatz). *Sei V ein Vektorraum mit Erzeugendensystem E . Für jede linear unabhängige Teilmenge $U \subseteq V$ gilt $|U| \leq |E|$.*

Beweis. O. B. d. A. sei E endlich, sagen wir $E = \{s_1, \dots, s_n\}$. Seien $u_1, \dots, u_m \in U$ paarweise verschieden. Wir müssen $m \leq n$ zeigen. Da U linear unabhängig ist, gilt $0 \neq u_1 = \sum_{i=1}^n \lambda_i s_i$, wobei nicht alle $\lambda_1, \dots, \lambda_n \in K$ verschwinden. Sei also o. B. d. A. $\lambda_1 \neq 0$ und daher

$$s_1 = \lambda_1^{-1} u_1 + \sum_{i=2}^n (-\lambda_1^{-1} \lambda_i) s_i \in \langle u_1, s_2, \dots, s_n \rangle.$$

Folglich ist auch $\{u_1, s_2, \dots, s_n\}$ ein Erzeugendensystem mit n Elementen (wir haben s_1 gegen u_1 ausgetauscht). Schreibe nun $u_2 = \mu_1 u_1 + \sum_{i=2}^n \mu_i s_i$ mit $\mu_1, \dots, \mu_n \in K$. Wegen $u_2 \notin \langle u_1 \rangle$ muss mindestens ein μ_i mit $i \geq 2$ ungleich 0 sein. Sagen wir $\mu_2 \neq 0$. Wegen

$$s_2 = -\mu_2^{-1} \mu_1 u_1 + \mu_2^{-1} u_2 - \sum_{i=3}^n \mu_2^{-1} \mu_i s_i \in \langle u_1, u_2, s_3, \dots, s_n \rangle$$

kann man s_2 auf die gleiche Weise gegen u_2 austauschen. Wiederholt man diesen Prozess, so erhält man schließlich das Erzeugendensystem $\{u_1, \dots, u_m, s_{m+1}, \dots, s_n\}$ von V . Insbesondere ist $m \leq n$. \square

Beispiel 4.14. Die Menge $\{(1, 2, 3, 4), (-1, 4, 0, 2), (0, 5, 2, 1), (0, 0, -7, 1), (-3, 4, 1, 0)\} \subseteq \mathbb{R}^4$ muss linear abhängig sein, da $\{e_1, e_2, e_3, e_4\}$ ein Erzeugendensystem von \mathbb{R}^4 ist (beachten Sie, dass man nichts rechnen muss).

Satz 4.15. *Jeder endlich erzeugte Vektorraum besitzt eine endliche Basis und je zwei Basen sind gleichmächtig.*

Beweis. Sei V ein Vektorraum mit endlichem Erzeugendensystem E . Nach dem Basisergänzungssatz kann man die linear unabhängige Menge \emptyset mit Elementen aus E zu einer Basis B von V ergänzen. Insbesondere ist $|B| \leq |E| < \infty$. Sei auch C eine Basis von V . Nach dem Austauschatz gilt $|C| \leq |B| \leq |C|$, also $|C| = |B|$. Da B und C endlich sind, müssen sie nach Bemerkung 2.11(e) gleichmächtig sein. \square

Folgerung 4.16. *Jeder Unterraum U eines endlich erzeugten Vektorraums V ist endlich erzeugt und besitzt ein Komplement $W \leq V$, d. h. es gilt $V = U \oplus W$.*

Beweis. Sei B eine Basis von V und $S \subseteq U$ linear unabhängig. Nach dem Austauschatz gilt $|S| \leq |B| < \infty$. Insbesondere besitzt U eine maximal linear unabhängige Teilmenge C . Nach Lemma 4.10 ist C eine (endliche) Basis von U . Also ist U endlich erzeugt. Nach dem Basisergänzungssatz lässt sich C zu einer Basis D von V ergänzen. Die zweite Behauptung folgt dann mit $W := \langle D \setminus C \rangle$. \square

Bemerkung 4.17.

- (a) In der Situation von Folgerung 4.16 ist W im Allgemeinen nicht eindeutig bestimmt. Zum Beispiel gilt

$$\mathbb{R}^2 = \mathbb{R}(1, 0) \oplus \mathbb{R}(0, 1) = \mathbb{R}(1, 0) \oplus \mathbb{R}(1, 1).$$

- (b) Mit dem Auswahlaxiom (genauer mit ZORNs Lemma) kann man zeigen, dass jeder Vektorraum eine (möglicherweise unendliche) Basis besitzt und je zwei Basen gleichmächtig sind. Zum Beispiel hat \mathbb{R} als \mathbb{Q} -Vektorraum unendliche Basen, von denen man keine explizit angeben kann. Wir überlegen uns in Satz 6.12 wie man Basen von endlich-erzeugten Vektorräumen effizient berechnet.

4.3 Dimension

Definition 4.18. Sei B eine Basis eines endlich erzeugten K -Vektorraums V . Dann nennt man

$$d = \dim_K V = \dim V := |B| \in \mathbb{N}_0$$

die *Dimension* von V . Nach Satz 4.15 hängt d nicht von der Wahl von B ab. Anstatt „endlich erzeugt“ kann man nun *endlich-dimensional* oder genauer *d-dimensional* sagen.

Beispiel 4.19.

- (a) Für jeden Körper K und $n \geq 1$ hat K^n Dimension n (wähle die Standardbasis). Der Unterraum $U := \{(x, x) \in K^2 : x \in K\} \leq K^2$ ist 1-dimensional mit Basis $\{(1, 1)\}$.
- (b) Im \mathbb{R}^3 beschreibt $\{(x, y, 0) : x, y \in \mathbb{R}\}$ eine 2-dimensionale Ebene. Allgemeiner nennt man einen $(d - 1)$ -dimensionalen Unterraum eines d -dimensionalen Raums eine *Hyperebene*.
- (c) Sei V ein d -dimensionaler \mathbb{F}_2 -Vektorraum. Die Koordinatendarstellung bzgl. einer Basis zeigt

$$|V| = |\mathbb{F}_2^d| = |\mathbb{F}_2 \times \dots \times \mathbb{F}_2| = 2^d.$$

Bemerkung 4.20.

- (a) Aus den obigen Sätzen folgen einige nützliche Fakten:
- Für $U \leq V$ gilt $\dim U \leq \dim V$ mit Gleichheit genau dann, wenn $U = V$ (ergänze eine Basis von U zu einer Basis von V).
 - Jedes Erzeugendensystem E von V enthält eine Basis von V (reduziere zu einem minimalen Erzeugendensystem). Insbesondere ist $|E| \geq \dim V$.
 - $d + 1$ Vektoren eines d -dimensionalen Vektorraums sind linear abhängig.
- (b) In der linearen Algebra stehen endlich-dimensionale Vektorräume im Vordergrund, während unendlich-dimensionale Vektorräume Gegenstand der *Funktionalanalysis* sind.

(c) Die folgende Formel entspricht der Gleichung $|A \cup B| = |A| + |B| - |A \cap B|$ für endliche Mengen A und B (Lemma 1.12).

Satz 4.21 (Dimensionsformel). Für Unterräume U und W eines endlich-dimensionalen Vektorraums V gilt

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W).$$

Ist die Summe direkt, so gilt $\dim(U \oplus W) = \dim U + \dim W$.

Beweis. Sei $\{b_1, \dots, b_n\}$ eine Basis von $U \cap W$. Wir ergänzen zu einer Basis $\{b_1, \dots, b_n, c_1, \dots, c_s\}$ von U und einer Basis $\{b_1, \dots, b_n, d_1, \dots, d_t\}$ von W . Da $U + W$ aus den Elementen der Form $u + w$ mit $u \in U$ und $w \in W$ besteht, wird $U + W$ von $b_1, \dots, b_n, c_1, \dots, c_s, d_1, \dots, d_t$ erzeugt.

Für die lineare Unabhängigkeit seien $\lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_s, \rho_1, \dots, \rho_t \in K$ mit

$$\underbrace{\sum_{i=1}^n \lambda_i b_i}_{=:v} + \underbrace{\sum_{j=1}^s \mu_j c_j}_{=:u} + \underbrace{\sum_{k=1}^t \rho_k d_k}_{=:w} = 0.$$

Dann ist $v + u = -w \in U \cap W$. Also lässt sich $v + u$ als Linearkombination von b_1, \dots, b_n ausdrücken. Andererseits ist die Darstellung von $v + u$ bzgl. der Basis $\{b_1, \dots, b_n, c_1, \dots, c_s\}$ eindeutig nach Satz 4.8. Dies zeigt $\mu_1 = \dots = \mu_s = 0$. Nun ist $v + w = 0$ eine Linearkombination der Basis $\{b_1, \dots, b_n, d_1, \dots, d_t\}$. Dies geht nur falls $\lambda_1 = \dots = \lambda_n = \rho_1 = \dots = \rho_t = 0$. Daher ist $\{b_1, \dots, b_n, c_1, \dots, c_s, d_1, \dots, d_t\}$ linear unabhängig und folglich eine Basis von $U + W$. Man erhält

$$\dim(U + W) = n + s + t = (n + s) + (n + t) - n = \dim U + \dim W - \dim(U \cap W).$$

Ist die Summe direkt, so gilt $U \cap W = \{0\}$ und die zweite Behauptung folgt. □

Beispiel 4.22. Sei

$$U := \langle (1, 1, 0), (0, 2, 1) \rangle \leq \mathbb{R}^3,$$

$$W := \langle (1, 1, 1) \rangle \leq \mathbb{R}^3.$$

Offenbar gilt $\dim U = 2$ und $\dim W = 1$. Für $v \in U \cap W$ existieren $\lambda, \mu, \rho \in \mathbb{R}$ mit

$$v = \lambda(1, 1, 0) + \mu(0, 2, 1) = \rho(1, 1, 1).$$

Es folgt $(\lambda, \lambda + 2\mu, \mu) = (\rho, \rho, \rho)$. Ein Koeffizientenvergleich liefert $\lambda = \rho = \mu$ und $3\rho = \lambda + 2\mu = \rho$. Dies kann nur für $\rho = 0$ gelten. Also ist $v = 0(1, 1, 1) = 0$ und $U \cap W = \{0\}$. Man erhält $\dim(U + W) = \dim U + \dim W = 2 + 1 = 3$. Wegen $U + W \leq \mathbb{R}^3$ ist daher $\mathbb{R}^3 = U \oplus W$.

Satz 4.23. Für Vektorräume $U \leq V$ gilt $\dim V = \dim U + \dim(V/U)$.

Beweis. Sei W ein Komplement von U in V . Sei B eine Basis von W . Es genügt zu zeigen, dass $\bar{B} := \{b + U : b \in B\}$ eine Basis von V/U ist. Jedes Element in V hat die Form $v = w + u$ mit $u \in U$ und $w \in W$. Wegen $v + U = w + U$ ist \bar{B} ein Erzeugendensystem von V/U . Seien nun $\lambda_b \in K$ mit $\sum_{b \in B} \lambda_b (b + U) = 0_{V/U}$. Dann ist $\sum_{b \in B} \lambda_b b \in U$. Aus $U \cap W = \{0\}$ folgt $\lambda_b = 0$ für alle $b \in B$. Also ist \bar{B} linear unabhängig. □

5 Matrizen

5.1 Der Matrizen-Vektorraum

Bemerkung 5.1. Sofern nichts Gegenteiliges gesagt wird, setzen wir von nun an stillschweigend voraus, dass alle Vektorräume endlich-dimensional sind. Eine Matrix ist ein Schema zur expliziten Berechnung von Basen von Vektorräumen und Lösungen von linearen Gleichungssystemen. Matrizen treten auch als eigenständige Objekte in zahlreichen anderen Gebieten auf.

Definition 5.2. Sei K ein Körper und $n, m \in \mathbb{N}$. Eine $(n \times m)$ -Matrix über K ist ein rechteckig angeordnetes nm -Tupel

$$A = (a_{ij})_{i,j=1}^n = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}$$

mit $a_{ij} \in K$ für $i = 1, \dots, n$ und $j = 1, \dots, m$. Die Menge der $n \times m$ -Matrizen über K bezeichnet man mit $K^{n \times m}$. Im Fall $n = m$ nennt man A *quadratisch*.

Beispiel 5.3.

- (a) Die 1×1 -Matrizen entsprechen genau den Elementen aus K . Die Vektoren aus K^n kann man als $1 \times n$ -Matrizen auffassen. Man spricht dann von *Zeilenvektoren*. Die $n \times 1$ -Matrizen heißen demnach *Spaltenvektoren*. Wenn Missverständnisse ausgeschlossen sind, benutzen wir die Standardbasis e_1, \dots, e_n sowohl als Zeilen- als auch Spaltenvektoren.
- (b) Die $n \times m$ -Nullmatrix $0_{n \times m} := (0)_{i,j} \in K^{n \times m}$, $0_n := 0_{n \times n}$ (wie üblich lassen wir die Indizes weg, wenn Missverständnisse ausgeschlossen sind).
- (c) Die (quadratische) $n \times n$ -Einheitsmatrix

$$1_n = (\delta_{ij}) := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}.$$

Das Symbol δ_{ij} nennt man das *KRONECKER-Delta*. Es gilt

$$\delta_{ij} := \begin{cases} 1 & \text{falls } i = j, \\ 0 & \text{falls } i \neq j. \end{cases}$$

Die Zeilen von 1_n bilden die Standardbasis $\{e_1, \dots, e_n\}$ von K^n .

(d) Für $\lambda_1, \dots, \lambda_n \in K$ nennt man

$$\text{diag}(\lambda_1, \dots, \lambda_n) := (\delta_{ij}\lambda_i) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

eine *Diagonalmatrix*. Die Einträge $\lambda_1, \dots, \lambda_n$ bilden die *Hauptdiagonale*. Im Spezialfall $\lambda_1 = \dots = \lambda_n$ spricht man von *Skalarmatrizen*.

(e) Die $n \times m$ -Matrix E_{st} mit einer 1 an Position (s, t) und sonst nur Nullen. Man nennt sie *Standardmatrizen*. Mit dem Kronecker-Delta gilt $E_{st} = (\delta_{is}\delta_{jt})_{i,j}$.

(f) Für $A \in K^{n \times m}$ ist $A^t := (a_{ji})_{i,j} \in K^{m \times n}$ die zu A *transponierte Matrix*. Sie entsteht aus A durch Spiegelung an der Hauptdiagonale:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \longrightarrow A^t = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$$

Offensichtlich ist $(A^t)^t = A$. Durch Transponieren werden Zeilenvektoren zu Spaltenvektoren und umgekehrt.

(g) Quadratische Matrizen A heißen *symmetrisch*, falls $A^t = A$.

Lemma 5.4. *Mit komponentenweisen Verknüpfungen wird $K^{n \times m}$ zu einem nm -dimensionalen K -Vektorraum:*

$$A + B := (a_{ij} + b_{ij})_{i,j}, \\ \lambda \cdot A := (\lambda a_{ij})_{i,j}$$

für $A = (a_{ij}), B = (b_{ij}) \in K^{n \times m}$ und $\lambda \in K$. Die Standardmatrizen E_{st} bilden eine Basis von $K^{n \times m}$. Insbesondere ist $\dim(K^{n \times m}) = nm$.

Beweis. Die definierten Verknüpfungen auf $K^{n \times m}$ entsprechen genau den Verknüpfungen in K^{nm} , indem man die Vektoren aus K^{nm} als $n \times m$ -Matrix anordnet. Da K^{nm} ein Vektorraum ist, muss auch $K^{n \times m}$ ein Vektorraum sein. Die Standardbasisvektoren e_1, \dots, e_{nm} von K^{nm} entsprechen (bis auf Reihenfolge) genau den Standardmatrizen. \square

Beispiel 5.5.

(a) Achtung: Nur Matrizen vom gleichen Format können addiert werden. Zum Beispiel:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} + \begin{pmatrix} 0 & -1 & 1 \\ 1 & -2 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 4 \\ 5 & 3 & 8 \end{pmatrix} \quad 2 \cdot \begin{pmatrix} 1 & 2 \\ 0 & -3 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 0 & -6 \end{pmatrix}$$

Die Skalarmatrizen sind genau die skalaren Vielfachen der Einheitsmatrix.

(b) Die symmetrischen Matrizen bilden einen Unterraum S von $K^{n \times n}$. Eine Basis erhält man durch die Matrizen E_{11}, \dots, E_{nn} und $E_{ij} + E_{ji}$ für $i < j$. Insbesondere ist

$$\dim S = n + \frac{n(n-1)}{2} = \frac{n(n+1)}{2}.$$

5.2 Matrizenmultiplikation

Definition 5.6. Für $A = (a_{ij}) \in K^{n \times m}$ und $B = (b_{ij}) \in K^{m \times k}$ sei $A \cdot B := (c_{ij})_{i,j} \in K^{n \times k}$ mit

$$c_{ij} := \sum_{l=1}^m a_{il}b_{lj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{im}b_{mj}.$$

Bemerkung 5.7.

(a) Merkgel: c_{ij} entsteht, indem man die i -Zeile von A mit der j -Spalte von B „verrechnet“:

$$\begin{pmatrix} * & * & * \\ * & * & * \\ a & b & c \\ * & * & * \end{pmatrix} \cdot \begin{pmatrix} * & a' \\ * & b' \\ * & c' \end{pmatrix} = \begin{pmatrix} * & * \\ * & * \\ * & aa' + bb' + cc' \\ * & * \end{pmatrix}$$

(die Sterne bezeichnen beliebige Einträge). Oft ist es hilfreich sich folgendes Schema vorzustellen:

$$\boxed{4 \times 3} \cdot \boxed{3 \times 2} = \boxed{4 \times 2}$$

(b) Die Multiplikation von Diagonalmatrizen ist einfach:

$$\text{diag}(\lambda_1, \dots, \lambda_n) \cdot \text{diag}(\mu_1, \dots, \mu_n) = \text{diag}(\lambda_1\mu_1, \dots, \lambda_n\mu_n).$$

(c) Als Vektorraum ist $(K^{n \times n}, +)$ eine abelsche Gruppe. Das folgende Lemma zeigt, dass $(K^{n \times n}, +, \cdot)$ einige, aber nicht alle Körperaxiome erfüllt.¹

Lemma 5.8. Für alle Matrizen A, B, C mit „passendem“ Format und $\lambda \in K$ gilt

$$\begin{array}{lll} A \cdot 1_m = A = 1_n \cdot A, & (AB)^t = B^t A^t, & \lambda(AB) = (\lambda A)B = A(\lambda B), \\ A(BC) = (AB)C, & A(B + C) = AB + AC, & (A + B)C = AC + BC. \end{array}$$

Beweis. Wie üblich sei $A = (a_{ij})$, $B = (b_{ij})$ und $C = (c_{ij})$. Für eine beliebige Matrix M sei M_{ij} der Eintrag an Position (i, j) . Dann gilt

$$\begin{aligned} (A1_m)_{ij} &= \sum_{k=1}^m a_{ik}\delta_{kj} = a_{ij} = \sum_{k=1}^n \delta_{ik}a_{kj} = (1_n A)_{ij}, \\ ((AB)^t)_{ij} &= \sum_{k=1}^m a_{jk}b_{ki} = \sum_{k=1}^m (B^t)_{ik}(A^t)_{kj} = (B^t A^t)_{ij}, \\ (\lambda(AB))_{ij} &= \lambda \sum_{k=1}^m a_{ik}b_{kj} = \sum_{k=1}^m (\lambda a_{ik})b_{kj} = ((\lambda A)B)_{ij} = \sum_{k=1}^m a_{ik}(\lambda b_{kj}) = (A(\lambda B))_{ij}, \\ (A(BC))_{ij} &= \sum_{k=1}^m a_{ik}(BC)_{kj} = \sum_{k=1}^m a_{ik} \sum_{l=1}^n b_{kl}c_{lj} = \sum_{k=1}^m \sum_{l=1}^n a_{ik}b_{kl}c_{lj} \end{aligned}$$

¹Man nennt diese schwächere Struktur einen *Ring*.

$$\begin{aligned}
&= \sum_{l=1}^n \left(\sum_{k=1}^m a_{ik} b_{kl} \right) c_{lj} = \sum_{l=1}^n (AB)_{il} c_{lj} = ((AB)C)_{ij}, \\
(A(B+C))_{ij} &= \sum_{k=1}^m a_{ik} (B+C)_{kj} = \sum_{k=1}^m a_{ik} (b_{kj} + c_{kj}) = \sum_{k=1}^m (a_{ik} b_{kj} + a_{ik} c_{kj}) \\
&= \sum_{k=1}^m a_{ik} b_{kj} + \sum_{k=1}^m a_{ik} c_{kj} = (AB)_{ij} + (AC)_{ij}, \\
((A+B)C)_{ij} &= \sum_{k=1}^m (a_{ik} + b_{ik}) c_{kj} = \sum_{k=1}^m a_{ik} c_{kj} + \sum_{k=1}^m b_{ik} c_{kj} = (AC)_{ij} + (BC)_{ij}. \quad \square
\end{aligned}$$

Bemerkung 5.9.

- (a) Für $A \in K^{n \times m}$ gilt selbstverständlich $0_{k \times n} \cdot A = 0_{k \times m}$ und $A \cdot 0_{m \times k} = 0_{n \times k}$.
- (b) Wir nennen Matrizen $A, B \in K^{n \times n}$ *vertauschbar*, falls $AB = BA$. Das ist für $n \geq 2$ in der Regel nicht erfüllt:

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = 0_2 \neq \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Daher ist $(K^{n \times n}, +, \cdot)$ *kein* Körper. Wir sehen auch, dass nicht jede (von 0_n verschiedene) Matrix ein Inverses bzgl. \cdot besitzt (selbst die Kürzungsregel gilt nicht für Matrizen).

- (c) Eine quadratische Matrix $A \in K^{n \times n}$ heißt *invertierbar*, falls eine Matrix $B \in K^{n \times n}$ mit

$$AB = 1_n = BA$$

existiert.² Gilt auch $AC = 1_n = CA$, so ist $C = C1_n = C(AB) = (CA)B = 1_n B = B$. Daher ist B eindeutig bestimmt und man schreibt wie bei Gruppen $A^{-1} := B$. Man nennt A^{-1} die zu A *inverse* Matrix. Wir zeigen in Lemma 5.15, dass aus $AB = 1_n$ bereits die Invertierbarkeit folgt, d. h. $BA = 1_n$ muss nicht geprüft werden.

- (d) Ist A invertierbar, so auch A^t , denn

$$A^t(A^{-1})^t = (A^{-1}A)^t = 1_n^t = 1_n = (AA^{-1})^t = (A^{-1})^t A^t.$$

Man setzt daher $A^{-t} := (A^{-1})^t = (A^t)^{-1}$.

- (e) Manchmal ist es nützlich Matrizen in rechteckige *Blöcke* zu unterteilen:

$$A = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix} \quad \text{mit } A_1 \in K^{n_1 \times m_1}, A_2 \in K^{n_1 \times m_2}, A_3 \in K^{n_2 \times m_1}, A_4 \in K^{n_2 \times m_2}$$

Für eine weitere Matrix $B = \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix}$ mit Blöcken $B_1 \in K^{m_1 \times k_1}$, $B_2 \in K^{m_1 \times k_2}$, $B_3 \in K^{m_2 \times k_1}$, $B_4 \in K^{m_2 \times k_2}$ gilt dann

$$AB = \begin{pmatrix} A_1 B_1 + A_2 B_3 & A_1 B_2 + A_2 B_4 \\ A_3 B_1 + A_4 B_3 & A_3 B_2 + A_4 B_4 \end{pmatrix}.$$

Sind A_1 und A_4 quadratisch und $A_2 = A_3 = 0$, so nennt man $A = \begin{pmatrix} A_1 & 0 \\ 0 & A_4 \end{pmatrix} = \text{diag}(A_1, A_4)$ eine *Blockdiagonalmatrix*.

²Invertierbare (bzw. nicht invertierbare) Matrizen nennt man auch *regulär* (bzw. *singulär*).

Lemma 5.10. Die invertierbaren Matrizen in $K^{n \times n}$ bilden eine Gruppe $GL(n, K)$ bzgl. Multiplikation. Man nennt sie allgemeine lineare Gruppe vom Grad n über K .

Beweis. Das neutrale Element 1_n ist offenbar invertierbar. Mit A ist auch A^{-1} invertierbar. Für invertierbare Matrizen A und B gilt

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = A1_nA^{-1} = AA^{-1} = 1_n$$

und analog $(B^{-1}A^{-1})(AB) = 1_n$. Dies zeigt, dass auch AB invertierbar ist mit $(AB)^{-1} = B^{-1}A^{-1}$. Das Assoziativgesetz der Multiplikation folgt aus Lemma 5.8. \square

Beispiel 5.11. In $GL(2, \mathbb{F}_2)$ gilt

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} = 1_2 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

5.3 Der Rang einer Matrix

Satz 5.12. Seien z_1, \dots, z_n die Zeilen und s_1, \dots, s_m die Spalten einer Matrix $A \in K^{n \times m}$. Dann gilt $\dim\langle z_1, \dots, z_n \rangle = \dim\langle s_1, \dots, s_m \rangle$.

Beweis. Nach Bemerkung 4.20 existiert $I \subseteq \{1, \dots, n\}$, sodass $\{z_i : i \in I\}$ eine Basis von $\langle z_1, \dots, z_n \rangle$ ist. Sei analog $\{s_j : j \in J\}$ eine Basis von $\langle s_1, \dots, s_m \rangle$ ist. Sei $A = (a_{ij})$. Wir zeigen, dass die Zeilen der Matrix

$$B := (a_{ij})_{i \in I, j \in J} \in K^{|I| \times |J|}$$

linear unabhängig sind. Seien dazu $\lambda_i \in K$ mit $\sum_{i \in I} \lambda_i a_{ij} = 0$ für alle $j \in J$. Für $k \in \{1, \dots, m\} \setminus J$ existieren $\mu_j \in K$ mit $s_k = \sum_{j \in J} \mu_j s_j$, d. h. $a_{ik} = \sum_{j \in J} \mu_j a_{ij}$ für alle $i \in \{1, \dots, n\}$. Es folgt

$$\sum_{i \in I} \lambda_i a_{ik} = \sum_{i \in I} \lambda_i \sum_{j \in J} \mu_j a_{ij} = \sum_{j \in J} \mu_j \sum_{i \in I} \lambda_i a_{ij} = 0.$$

Also gilt $\sum_{i \in I} \lambda_i a_{ij} = 0$ für alle $j \in \{1, \dots, m\}$, d. h. $\sum_{i \in I} \lambda_i z_i = 0$. Aus der linearen Unabhängigkeit von $\{z_i : i \in I\}$ folgt $\lambda_i = 0$ für $i \in I$. Daher sind die Zeilen von B linear unabhängig. Da sie im $|J|$ -dimensionalen Vektorraum $K^{|J|}$ liegen, gilt $|I| \leq |J|$. Die Zeilen (bzw. Spalten) von A sind die Spalten (bzw. Zeilen) von A^t . Benutzt man das obige Argument mit A^t , so folgt $|J| \leq |I|$. Insgesamt ist $\dim\langle z_1, \dots, z_n \rangle = |I| = |J| = \dim\langle s_1, \dots, s_m \rangle$. \square

Definition 5.13. In der Situation von Satz 5.12 nennt man

$$\text{rk}(A) := \dim\langle z_1, \dots, z_n \rangle = \dim\langle s_1, \dots, s_m \rangle$$

den *Rang* von A . Im Fall $\text{rk}(A) = \min\{n, m\}$ sagt man: A hat *vollen Rang*.

Beispiel 5.14.

- (a) Die Einheitsmatrix 1_n hat (vollen) Rang n , denn ihre Zeilen bilden die Standardbasis. Es gilt $\text{rk}\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} = 1$, denn die zweite Zeile ist das Doppelte der ersten. Wir überlegen uns in Bemerkung 6.13 wie man den Rang einer beliebigen Matrix effizient berechnet.

(b) Für jede Matrix A gilt $\text{rk}(A) = \text{rk}(A^t)$ und $\text{rk}(A) = 0 \iff A = 0$.

Lemma 5.15.

(a) Für Matrizen A und B mit „passendem“ Format gilt $\text{rk}(AB) \leq \min\{\text{rk}(A), \text{rk}(B)\}$.

(b) Eine quadratische Matrix ist genau dann invertierbar, wenn sie vollen Rang hat.

Beweis.

(a) Seien s_1, \dots, s_m die Spalten von A und sei $(\lambda_1, \dots, \lambda_m)^t$ die i -te Spalte von B . Dann ist $\lambda_1 s_1 + \dots + \lambda_m s_m$ die i -te Spalte von AB . Also sind die Spalten von AB Linearkombinationen der Spalten von A . Dies zeigt $\text{rk}(AB) \leq \text{rk}(A)$. Aus Beispiel 5.14 folgt

$$\text{rk}(AB) = \text{rk}((AB)^t) = \text{rk}(B^t A^t) \leq \text{rk}(B^t) = \text{rk}(B).$$

(b) Ist $A \in K^{n \times n}$ invertierbar, so gilt

$$n = \text{rk}(1_n) = \text{rk}(AA^{-1}) \stackrel{(a)}{\leq} \text{rk}(A) \leq n,$$

d. h. A hat vollen Rang. Sei umgekehrt $\text{rk}(A) = n$. Dann lassen sich die Standardbasisvektoren e_1, \dots, e_n als Linearkombinationen der Spalten von A ausdrücken. Also existiert $B \in K^{n \times n}$ mit $AB = 1_n$. Wegen $\text{rk}(A^t) = \text{rk}(A)$ existiert auch ein $C \in K^{n \times n}$ mit $A^t C = 1_n$, d. h. $C^t A = (A^t C)^t = 1_n^t = 1_n$. Wegen $C^t = C^t(AB) = (C^t A)B = B$ ist A invertierbar. \square

6 Der Gauß-Algorithmus

6.1 Gleichungssysteme

Definition 6.1. Ein (*lineares*) *Gleichungssystem* ist eine Matrixgleichung der Form $Ax = b$, wobei die *Koeffizientenmatrix* $A \in K^{n \times m}$ und der Vektor $b \in K^{n \times 1}$ gegeben sind. Gesucht ist die *Lösungsmenge*

$$L := \{x \in K^{m \times 1} : Ax = b\} \subseteq K^{m \times 1}.$$

- Im Fall $L \neq \emptyset$ nennt man das Gleichungssystem *lösbar*.
- Im Fall $b = 0$ nennt man das Gleichungssystem *homogen* und anderenfalls *inhomogen*.
- Durch Anfügen der Spalte b zu A erhält man die *erweiterte* Koeffizientenmatrix $(A|b) \in K^{n \times (m+1)}$.

Beispiel 6.2. Das Gleichungssystem

$$\begin{array}{rcl} 2x_1 + 3x_2 & = & 5, \\ -x_1 & = & 2 \end{array}$$

entspricht der Matrixgleichung

$$\begin{pmatrix} 2 & 3 \\ -1 & 0 \end{pmatrix} x = \begin{pmatrix} 5 \\ 2 \end{pmatrix}$$

mit genau einer Lösung $x = (-2, 3)^t$.

Bemerkung 6.3. Jedes homogene Gleichungssystem ist lösbar, denn der Nullvektor ist eine Lösung.

Satz 6.4 (KRONECKER-CAPELLI¹). *Genau dann ist das Gleichungssystem $Ax = b$ lösbar, wenn $\text{rk}(A) = \text{rk}(A|b)$.*

Beweis. Seien s_1, \dots, s_m die Spalten von A . Dann gilt

$$\begin{aligned} \text{rk}(A) = \text{rk}(A|b) &\iff \dim\langle s_1, \dots, s_m \rangle = \dim\langle s_1, \dots, s_m, b \rangle \\ &\iff \langle s_1, \dots, s_m \rangle = \langle s_1, \dots, s_m, b \rangle \iff b \in \langle s_1, \dots, s_m \rangle \\ &\iff \exists x_1, \dots, x_m \in K : b = \sum_{i=1}^m x_i s_i \iff \exists x \in K^{m \times 1} : Ax = b. \quad \square \end{aligned}$$

¹mitunter auch nach Rouché und Frobenius benannt

Bemerkung 6.5. Sei $A \in K^{n \times m}$ mit vollem Rang $n \leq m$. Dann ist

$$\text{rk}(A) \leq \text{rk}(A|b) \leq \min\{n, m+1\} = n = \text{rk}(A)$$

und $Ax = b$ ist stets lösbar. Im Fall $n = m$ ist A invertierbar nach Lemma 5.15 und $x = A^{-1}b$ ist die einzige Lösung. Im Fall $n < m$ nennt man das Gleichungssystem $Ax = b$ *unterbestimmt*, d. h. es gibt weniger Gleichungen als Unbekannte. Wir zeigen, dass es dann mehrere Lösungen gibt.

Satz 6.6. Sei $A \in K^{n \times m}$ und $b \in K^{n \times 1}$.

(a) Die Lösungsmenge des homogenen Gleichungssystems $Ax = 0$ ist ein Unterraum L_0 von $K^{m \times 1}$ der Dimension $m - \text{rk}(A)$.

(b) Besitzt das Gleichungssystem $Ax = b$ eine Lösung \tilde{x} , so ist

$$\tilde{x} + L_0 := \{\tilde{x} + y : y \in L_0\}$$

die Lösungsmenge.

Beweis.

(a) Wegen $0 \in L_0$ ist $L_0 \neq \emptyset$. Für $x, y \in L_0$ und $\lambda \in K$ gilt

$$A(\lambda x + y) = \lambda Ax + Ay = 0,$$

d. h. $\lambda x + y \in L_0$. Dies zeigt $L_0 \leq K^{m \times 1}$ (Bemerkung 3.13). Sei b_1, \dots, b_k eine Basis von L_0 . Wir ergänzen zu einer Basis b_1, \dots, b_m von $K^{m \times 1}$. Die i -te Spalte von A ist Ae_i mit dem Standardbasisvektor e_i . Da e_i eine Linearkombination von b_1, \dots, b_m ist, liegt jede Spalte von A in $\langle Ab_1, \dots, Ab_m \rangle$. Insbesondere ist $\text{rk}(A) = \dim\langle Ab_1, \dots, Ab_m \rangle$. Aus $b_1, \dots, b_k \in L_0$ folgt $Ab_1 = \dots = Ab_k = 0$ und

$$\text{rk}(A) = \dim\langle Ab_{k+1}, \dots, Ab_m \rangle.$$

Es genügt zu zeigen, dass die $m - k$ Vektoren Ab_{k+1}, \dots, Ab_m linear unabhängig sind, denn dann ist $\text{rk}(A) = m - k$. Seien $\lambda_i \in K$ mit $\sum_{i=k+1}^m \lambda_i Ab_i = 0$. Dann ist auch $A \sum_{i=k+1}^m \lambda_i b_i = 0$, d. h. $\sum_{i=k+1}^m \lambda_i b_i \in L_0 = \langle b_1, \dots, b_k \rangle$. Da b_1, \dots, b_m eine Basis von $K^{m \times 1}$ ist, erhält man $\lambda_{k+1} = \dots = \lambda_m = 0$ wie gewünscht.

(b) Es gilt

$$Ax = b \iff Ax = A\tilde{x} \iff A(x - \tilde{x}) = 0 \iff x - \tilde{x} \in L_0 \iff x \in \tilde{x} + L_0. \quad \square$$

Bemerkung 6.7.

(a) Hat $A \in K^{n \times m}$ vollem Rang $m \leq n$, so besitzt das Gleichungssystem $Ax = b$ höchstens eine Lösung. Im Fall $m < n$ spricht man von *überbestimmten* Gleichungssystemen. Im Folgenden beschäftigen wir uns mit der expliziten Konstruktion der Lösungsmenge eines beliebigen Gleichungssystems.

(b) Da die Abbildung $L_0 \rightarrow \tilde{x} + L_0, v \mapsto \tilde{x} + v$ eine Bijektion ist, besitzt ein lösbares Gleichungssystem genauso viele Lösungen wie das entsprechende homogene Gleichungssystem. Ist K endlich (z. B. $K = \mathbb{F}_2$), so ist die Anzahl dieser Lösungen eine Potenz von $|K|$ (Satz 4.8). Für unendliche Körper wie \mathbb{Q} oder \mathbb{R} besitzt jedes Gleichungssystem keine, genau eine oder unendlich viele Lösungen.

6.2 Elementare Zeilenoperationen

Definition 6.8. Die folgenden Transformationen einer Matrix $A \in K^{n \times m}$ werden (*elementare*) *Zeilenoperationen* genannt:

- Multiplikation einer Zeile von A mit einem Skalar $\lambda \in K^\times$. Dies entspricht der Multiplikation mit einer *Elementarmatrix* der Form

$$\begin{pmatrix} 1_{s-1} & & 0 \\ & \lambda & \\ 0 & & 1_{n-s} \end{pmatrix} = 1_n + (\lambda - 1)E_{ss}$$

von links an A .

- Vertauschen zweier Zeilen von A . Dies entspricht der Multiplikation mit einer Elementarmatrix der Form

$$\begin{pmatrix} 1_{s-1} & & & & \\ & 0 & & 1 & \\ & & 1_{t-s-1} & & \\ & 1 & & & 0 \\ & & & & 1_{n-t} \end{pmatrix} = 1_n - E_{ss} - E_{tt} + E_{st} + E_{ts}$$

von links an A .

- Addieren eines Vielfaches einer Zeile von A zu einer anderen Zeile. Dies entspricht der Multiplikation mit einer Elementarmatrix der Form

$$\begin{pmatrix} 1_{s-1} & & & & \\ & 1 & & \lambda & \\ & & 1_{t-s-1} & & \\ & & & 1 & \\ & & & & 1_{n-t} \end{pmatrix} = 1_n + \lambda E_{st} \quad (\lambda \in K, s \neq t)$$

von links an A .

Matrizen A und B heißen *zeilen-äquivalent*, falls sich A durch endlich viele elementare Zeilenoperationen in B überführen lässt. Ggf. schreiben wir $A \sim B$.

Bemerkung 6.9.

- Alle elementaren Zeilenoperationen sind umkehrbar. Aus $A \sim B$ folgt somit $B \sim A$. Außerdem sind die Elementarmatrizen invertierbar. Nach Lemma 5.10 ist auch das Produkt von Elementarmatrizen invertierbar. Aus $A \sim B$ folgt daher die Existenz einer Matrix $S \in \text{GL}(n, K)$ mit $SA = B$.
- Nach (a) ist die Zeilen-Äquivalenz eine Äquivalenzrelation auf $K^{n \times m}$. Im nächsten Satz bestimmen wir ein besonders einfaches Repräsentantensystem für die Äquivalenzklassen.
- Analog definiert man (*elementare*) *Spaltenoperationen*. Diese entsprechen der Multiplikation von Elementarmatrizen von *rechts* an A (probieren Sie es aus). Spaltenoperationen lassen sich auch durch Zeilenoperation mit A^t realisieren. Wir benutzen die Schreibweise $A \sim B$ auch, wenn B aus A durch Spaltenoperationen hervorgeht.

Satz 6.10 (GAUSS-Algorithmus²). Jede Matrix $A \in K^{n \times m}$ ist zeilen-äquivalent zu genau einer Matrix \hat{A} in Zeilenstufenform³, d. h.

$$\hat{A} = \begin{pmatrix} 0 & \cdots & 0 & \boxed{1} & * & \cdots & * & 0 & * & \cdots & * & 0 & * & \cdots & * \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & \boxed{1} & * & \cdots & * & 0 & * & \cdots & * \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & \boxed{1} & * & \cdots & * \\ \vdots & & & & & & & & & & & & & & \vdots \end{pmatrix}.$$

Beweis. Der folgende Algorithmus überführt A in \hat{A} :

- (1) Setze $z := 1$ (Zeilenindex).
- (2) Für $s = 1, \dots, m$ (Spaltenindex):
 - Falls $\exists i \geq z : a_{is} \neq 0$, dann:
 - Tausche i -te mit z -ter Zeile. Anschließend gilt $a_{zs} \neq 0$.
 - Dividiere z -te Zeile durch a_{zs} . Anschließend gilt $a_{zs} = 1$.
 - Für $j = 1, \dots, z-1, z+1, \dots, n$ subtrahiere das a_{js} -Fache der z -ten Zeile von der j -ten Zeile. Anschließend gilt $a_{js} = 0$.
 - Erhöhe z um 1.

Für die Eindeutigkeit von \hat{A} seien B und C Matrizen in Zeilenstufenform mit $A \sim B$ und $A \sim C$. Dann ist auch $B \sim C$ und es existiert $S \in GL(n, K)$ mit $SB = C$. Sei b_i (bzw. s_i) die i -te Spalte von B (bzw. S). Dann ist $c_i = Sb_i$ die i -te Spalte von C . Sei e_1, \dots, e_n die Standardbasis von K^n . Sei $b_i \in \langle e_1, \dots, e_k \rangle$. Wir zeigen $b_i = c_i$ und $s_k = e_k^t$ durch Induktion nach k . Im Fall $k = 0$ ist $b_i \in \langle \emptyset \rangle = \{0\}$, also $b_i = 0$. Sicher ist dann auch $c_i = Sb_i = 0$. Sei nun die Behauptung bis $k-1$ bereits bewiesen. Die erste Spalte (von links) von B , die nicht in $\langle e_1, \dots, e_{k-1} \rangle$ liegt, ist $b_i = e_k$ wegen der Zeilenstufenform. Die Spalten von S sind linear unabhängig, da S invertierbar ist. Dies zeigt

$$c_i = Sb_i = Se_k^t = s_k \notin \langle s_1, \dots, s_{k-1} \rangle = \langle e_1, \dots, e_{k-1} \rangle.$$

Also ist c_i die erste Spalte von C , die nicht in $\langle e_1, \dots, e_{k-1} \rangle$ liegt, d. h. $c_i = e_k = s_k = b_i$ (Zeilenstufenform). Für jede weitere Spalte $b_j \in \langle e_1, \dots, e_k \rangle$ gilt nun $c_j = Sb_j = (e_1, \dots, e_k, s_{k+1}, \dots, s_n)b_j = b_j$. Somit ist $b_i = c_i$ für $i = 1, \dots, m$, d. h. $B = C$. \square

Beispiel 6.11.

$$\begin{pmatrix} 0 & 1 & 1 & 3 \\ 2 & 1 & 3 & 0 \\ 3 & -1 & 2 & 1 \end{pmatrix} \begin{matrix} \leftarrow \\ \leftarrow \\ \leftarrow \end{matrix} \sim \begin{pmatrix} 2 & 1 & 3 & 0 \\ 0 & 1 & 1 & 3 \\ 3 & -1 & 2 & 1 \end{pmatrix} \begin{matrix} | : 2 \\ \\ \end{matrix} \sim \begin{pmatrix} 1 & 1/2 & 3/2 & 0 \\ 0 & 1 & 1 & 3 \\ 3 & -1 & 2 & 1 \end{pmatrix} \begin{matrix} \leftarrow -3 \\ \\ \leftarrow + \end{matrix} \\ \sim \begin{pmatrix} 1 & 1/2 & 3/2 & 0 \\ 0 & 1 & 1 & 3 \\ 0 & -5/2 & -5/2 & 1 \end{pmatrix} \begin{matrix} \leftarrow + \\ \leftarrow -1/2 \\ \leftarrow + \end{matrix} \sim \begin{pmatrix} 1 & 0 & 1 & -3/2 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & 17/2 \end{pmatrix} \begin{matrix} \\ \\ | : 17/2 \end{matrix}$$

²auch *Gauß-Elimination* oder *Gauß-Jordan-Algorithmus* genannt
³Jede von 0 verschiedene Zeile enthält eine *führende Eins*. Alle Einträge links, ober- und unterhalb einer führenden Eins sind 0. Die führenden Einsen rutschen mit jeder Zeile weiter nach rechts. Nullzeilen (falls vorhanden) stehen unten.

$$\sim \begin{pmatrix} 1 & 0 & 1 & -3/2 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{array}{l} \leftarrow + \\ \leftarrow + \\ \leftarrow -3 \end{array} \begin{array}{l} + \\ + \\ -3/2 \end{array} \sim \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

6.3 Anwendungen

Satz 6.12. Sei $U := \langle u_1, \dots, u_n \rangle \leq K^m$. Sei $A \in K^{n \times m}$ die Matrix mit Zeilen u_1, \dots, u_n . Dann bilden die von 0 verschiedenen Zeilen von \hat{A} eine Basis von U . Insbesondere ist $\dim U = \text{rk}(\hat{A}) = \text{rk}(A)$.

Beweis. Durch elementare Zeilenoperationen werden Zeilen von A durch Linearkombinationen von Zeilen ersetzt. Die Zeilen von \hat{A} erzeugen daher einen Unterraum $W \leq U$. Da alle Zeilenoperationen umkehrbar sind, gilt sogar $W = U$ und $\dim U = \text{rk}(A) = \text{rk}(\hat{A})$. Sei k die Anzahl der von 0 verschiedenen Zeilen in \hat{A} . Dann gilt $\text{rk}(\hat{A}) \leq k$. Andererseits besitzt \hat{A} die linear unabhängigen Spalten e_1, \dots, e_k . Dies zeigt $\text{rk}(\hat{A}) = k$ und die Behauptung folgt. \square

Bemerkung 6.13. Zur Ermittlung einer Basis von U muss man beim Gauß-Algorithmus keine Nullen oberhalb der führenden Einsen erzeugen (die von 0 verschiedenen Zeilen sind trotzdem linear unabhängig, denn deren Anzahl bleibt gleich). Außerdem ist es ratsam Divisionen zu vermeiden, indem man mit Zeilen tauscht, die bereits eine führende Eins in der aktuellen Spalte haben. Ist man nur an $\dim U$ (oder allgemeiner an $\text{rk}(A)$) interessiert, so darf man wegen $\text{rk}(A) = \text{rk}(A^t)$ auch elementare Spaltenoperationen verwenden. Dies kann nützlich sein, wenn A weniger Spalten als Zeilen besitzt (viele Möglichkeiten führen zum Ziel).

Beispiel 6.14. Eine Art Schach-Rätsel: Rang in zwei Zügen!

$$\begin{pmatrix} 1 & -1 & 0 \\ 2 & 0 & 2 \\ 3 & -1 & 2 \\ -1 & 2 & 1 \end{pmatrix} \begin{array}{l} + \\ \leftarrow \downarrow \end{array} \sim \begin{pmatrix} 1 & 0 & 0 \\ 2 & 2 & 2 \\ 3 & 2 & 2 \\ -1 & 1 & 1 \end{pmatrix} \begin{array}{l} -1 + \\ \leftarrow \downarrow \end{array} \sim \begin{pmatrix} 1 & 0 & 0 \\ 2 & 2 & 0 \\ 3 & 2 & 0 \\ -1 & 1 & 0 \end{pmatrix} \implies \text{rk}(A) = 2$$

Satz 6.15. Sei $Ax = b$ ein Gleichungssystem mit $A \in K^{n \times m}$. Sei $(\hat{A}|c)$ die Zeilenstufenform von $(A|b)$. Dann gilt:

(a) Genau dann ist $Ax = b$ lösbar, wenn e_{m+1} keine Zeile von $(\hat{A}|c)$ ist.

Ggf. erhält man die Lösungsmenge wie folgt: Seien $(1, n_1), \dots, (k, n_k)$ die Positionen der führenden Einsen in $(\hat{A}|c)$. Die $n - k$ Nullzeilen werden gestrichen. Für alle $i \in \{1, \dots, m\} \setminus \{n_1, \dots, n_k\}$ fügen wir die Zeile $-e_i$ an Position i ein, sodass die resultierende Matrix $M \in K^{m \times (m+1)}$ auf der Hauptdiagonale nur Einträge ± 1 besitzt.

(b) Die letzte Spalte \tilde{x} von M erfüllt $A\tilde{x} = b$.

(c) Die $m - k$ Spalten von M , die nicht zu den Indizes n_1, \dots, n_k gehören, bilden eine Basis von $L_0 := \{x \in K^{m \times 1} : Ax = 0\}$.

(d) Die Lösungsmenge von $Ax = b$ ist $\tilde{x} + L_0$.

Beweis. Sei $S \in \text{GL}(n, K)$ mit $(SA|Sb) = S(A|b) = (\widehat{A}|c)$. Für $x \in K^{m \times 1}$ gilt

$$Ax = b \iff SAx = Sb \iff \widehat{A}x = c.$$

Ist e_{m+1} eine Zeile von $(\widehat{A}|c)$, so erhält man in der Gleichung $\widehat{A}x = c$ den Widerspruch $0 = 1$, d. h. es gibt keine Lösung.

Sei nun e_{m+1} keine Zeile von $(\widehat{A}|c)$. Wir verifizieren die Behauptungen an folgendem Beispiel

$$(\widehat{A}|c) = \left(\begin{array}{cc|cc|c} \cdot & 1 & a_1 & \cdot & a_2 & c_1 \\ \cdot & \cdot & \cdot & 1 & a_3 & c_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{array} \right) \quad M = \left(\begin{array}{cccc|cc} -1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 1 & a_1 & \cdot & a_2 & c_1 \\ \cdot & \cdot & -1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 & a_3 & c_2 \\ \cdot & \cdot & \cdot & \cdot & -1 & \cdot \end{array} \right)$$

(hier steht \cdot für 0 zur besseren Übersicht). Man sieht leicht, dass $\widehat{A}\tilde{x} = c$ gilt. Somit gilt auch $A\tilde{x} = b$. Damit sind (a) und (b) bewiesen. Genauso sieht man, dass die (rot markierten) Spalten s_1 , s_3 und s_5 von M in L_0 liegen. Die verschiedenen Positionen der Einträge -1 in s_i implizieren die lineare Unabhängigkeit von $\{s_1, s_3, s_5\}$. Nach Satz 6.6 und Bemerkung 6.13 ist andererseits

$$\dim L_0 = m - \text{rk}(A) = m - \text{rk}(\widehat{A}) = m - k = 3.$$

Dies zeigt (c). Aus Satz 6.6 folgt (d). □

Beispiel 6.16.

$$\begin{aligned} & \begin{pmatrix} -1 & 3 & 1 & 1 & 0 \\ -2 & 1 & -3 & 2 & -4 \\ 0 & 1 & 1 & 0 & 0 \end{pmatrix} x = \begin{pmatrix} 11 \\ -6 \\ 4 \end{pmatrix} \\ (A|b) &= \left(\begin{array}{ccccc|c} -1 & 3 & 1 & 1 & 0 & 11 \\ -2 & 1 & -3 & 2 & -4 & -6 \\ 0 & 1 & 1 & 0 & 0 & 4 \end{array} \right) \sim \dots \sim \begin{pmatrix} 1 & 0 & 2 & -1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{pmatrix} \\ M &= \begin{pmatrix} 1 & 0 & 2 & -1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 4 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{pmatrix} \\ L = \tilde{x} + L_0 &= \begin{pmatrix} 1 \\ 4 \\ 0 \\ 0 \\ 2 \end{pmatrix} + \left\langle \begin{pmatrix} 2 \\ 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 0 \\ -1 \\ 0 \end{pmatrix} \right\rangle \end{aligned}$$

Kontrolle:

$$A\tilde{x} = \begin{pmatrix} -1 & 3 & 1 & 1 & 0 \\ -2 & 1 & -3 & 2 & -4 \\ 0 & 1 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 4 \\ 0 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 11 \\ -6 \\ 4 \end{pmatrix} = b.$$

Satz 6.17 (Matrizeninversion). Für $A \in K^{n \times n}$ sei $(\hat{A}|B)$ die Zeilenstufenform von $(A|1_n) \in K^{n \times 2n}$. Genau dann ist A invertierbar, wenn $\hat{A} = 1_n$. Ggf. ist $A^{-1} = B$.

Beweis. Es gilt

$$A \text{ invertierbar} \xLeftrightarrow{5.15} \text{rk}(A) = n \xLeftrightarrow{6.12} \text{rk}(\hat{A}) = n \iff \hat{A} = 1_n.$$

Sei nun $S \in \text{GL}(n, K)$ mit

$$(SA|S) = S(A|1_n) = (\hat{A}|B) = (1_n|B).$$

Dann ist $B = S = S(AA^{-1}) = (SA)A^{-1} = 1_n A^{-1} = A^{-1}$. □

Folgerung 6.18. Jede invertierbare Matrix ist ein Produkt von Elementarmatrizen.

Beweis. Für $A \in \text{GL}(n, K)$ gilt $A \sim 1_n$ nach Satz 6.17. □

Beispiel 6.19.

$$\begin{aligned} (A|1_3) &= \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ -1 & -1 & 1 & 0 & 1 & 0 \\ -1 & 1 & -2 & 0 & 0 & 1 \end{array} \right) \begin{array}{l} \left[\begin{array}{l} \leftarrow \\ \leftarrow \end{array} \right]_+ \\ \leftarrow \end{array} \sim \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 & 1 & 0 \\ 0 & 1 & -3 & 1 & 0 & 1 \end{array} \right) \begin{array}{l} \leftarrow \\ \leftarrow \end{array} \\ &\sim \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & -3 & 1 & 0 & 1 \\ 0 & -1 & 0 & 1 & 1 & 0 \end{array} \right) \begin{array}{l} \left[\begin{array}{l} \leftarrow \\ \leftarrow \end{array} \right]_+ \\ \leftarrow \end{array} \sim \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & -3 & 1 & 0 & 1 \\ 0 & 0 & -3 & 2 & 1 & 1 \end{array} \right) \begin{array}{l} \leftarrow \\ \leftarrow \end{array} \mid : (-3) \\ &\sim \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & -3 & 1 & 0 & 1 \\ 0 & 0 & 1 & -2/3 & -1/3 & -1/3 \end{array} \right) \begin{array}{l} \left[\begin{array}{l} \leftarrow \\ \leftarrow \end{array} \right]_+ \\ \left[\begin{array}{l} \leftarrow \\ \leftarrow \end{array} \right]_3 \end{array} \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1/3 & -1/3 & -1/3 \\ 0 & 1 & 0 & -1 & -1 & 0 \\ 0 & 0 & 1 & -2/3 & -1/3 & -1/3 \end{array} \right) \\ A^{-1} &= \frac{1}{3} \begin{pmatrix} 1 & -1 & -1 \\ -3 & -3 & 0 \\ -2 & -1 & -1 \end{pmatrix} \end{aligned}$$

Bemerkung 6.20.

(a) Ergibt sich während des Gauß-Algorithmus ein „Versatz“ der Zeilen

$$\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & 1 & \\ & & & & & * \end{pmatrix},$$

so muss die Zeilenstufenform eine Nullzeile aufweisen. Die Matrix kann dann nicht invertierbar sein und man kann den Algorithmus vorzeitig abbrechen.

(b) Matrizen $A, B \in K^{n \times m}$ heißen *äquivalent*, falls man A durch Zeilen- und Spaltenoperationen in B überführen kann (und umgekehrt). Das bedeutet es existieren $S \in \text{GL}(n, K)$ und $T \in \text{GL}(m, K)$ mit $SAT = B$. Offenbar definiert dies eine Äquivalenzrelation. Man sieht leicht, dass jede Matrix A zu einer Matrix der Form

$$\begin{pmatrix} 1_r & 0 \\ 0 & 0 \end{pmatrix}$$

äquivalent ist. Dabei ist $r = \text{rk}(A)$ eindeutig bestimmt. Insbesondere sind A und B genau dann äquivalent, wenn $\text{rk}(A) = \text{rk}(B)$ gilt. Man braucht für die Äquivalenz also kein neues Symbol einführen. Die Anzahl der Äquivalenzklassen von $n \times n$ -Matrizen ist $n + 1$.

- (c) Der folgende Satz liefert einen effizienten Algorithmus um gleichzeitig Summe und Durchschnitt von Unterräumen zu bestimmen.

Satz 6.21 (ZASSENHAUS-Algorithmus). *Seien $U := \langle u_1, \dots, u_s \rangle \leq K^n$ und $W := \langle w_1, \dots, w_t \rangle \leq K^n$. Sei*

$$A := \begin{pmatrix} u_1 & u_1 \\ \vdots & \vdots \\ u_s & u_s \\ w_1 & 0 \\ \vdots & \vdots \\ w_t & 0 \end{pmatrix} \in K^{(s+t) \times 2n}, \quad \widehat{A} = \begin{pmatrix} b_1 & * \\ \vdots & \vdots \\ b_k & * \\ 0 & c_1 \\ \vdots & \vdots \\ \vdots & c_l \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{pmatrix},$$

wobei $b_1, \dots, b_k, c_1, \dots, c_l \in K^n$ mit $b_k \neq 0 \neq c_l$. Dann ist $\{b_1, \dots, b_k\}$ eine Basis von $U + W$ und $\{c_1, \dots, c_l\}$ ist eine Basis von $U \cap W$.

Beweis. Wegen $U + W = \langle u_1, \dots, u_s, w_1, \dots, w_t \rangle$ ist $\{b_1, \dots, b_k\}$ eine Basis von $U + W$ nach Satz 6.12. Außerdem ist jede Zeile der Form $(0, c_m)$ von \widehat{A} eine Linearkombination der Zeilen von A , sagen wir

$$(0, c_m) = \sum_{i=1}^s \lambda_i (u_i, u_i) + \sum_{j=1}^t \mu_j (w_j, 0)$$

mit $\lambda_1, \dots, \lambda_s, \mu_1, \dots, \mu_t \in K$. Dies zeigt

$$c_m = \sum_{i=1}^s \lambda_i u_i = - \sum_{j=1}^t \mu_j w_j \in U \cap W$$

für $m = 1, \dots, l$. Aufgrund der Zeilenstufenform ist $\{c_1, \dots, c_l\}$ linear unabhängig. Durch elementare Spaltenoperationen überführt man A zu

$$\begin{pmatrix} 0 & u_1 \\ \vdots & \vdots \\ 0 & u_s \\ w_1 & 0 \\ \vdots & \vdots \\ w_t & 0 \end{pmatrix}.$$

Aus Bemerkung 6.13 folgt nun leicht $\text{rk}(A) = \dim U + \dim W$. Die Dimensionsformel liefert daher

$$\dim(U \cap W) = \dim U + \dim W - \dim(U + W) = \text{rk}(A) - k = \text{rk}(\widehat{A}) - k = l.$$

Also ist $\{c_1, \dots, c_l\}$ eine Basis von $U \cap W$. □

Beispiel 6.22. Sei $U := \langle (1, 1, 1, 0), (0, -4, 1, 5) \rangle$ und $W := \langle (0, -2, 1, 2), (1, -1, 1, 3) \rangle$. Wie üblich muss man nicht alle Schritte des Gauß-Algorithmus durchführen:

$$\begin{aligned} & \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & -4 & 1 & 5 & 0 & -4 & 1 & 5 \\ 0 & -2 & 1 & 2 & 0 & 0 & 0 & 0 \\ 1 & -1 & 1 & 3 & 0 & 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & -2 & 1 & 2 & 0 & 0 & 0 & 0 \\ 0 & -4 & 1 & 5 & 0 & -4 & 1 & 5 \\ 0 & -2 & 0 & 3 & -1 & -1 & -1 & 0 \end{pmatrix} \\ \sim & \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & -2 & 1 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & -4 & 1 & 5 \\ 0 & 0 & -1 & 1 & -1 & -1 & -1 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & -2 & 1 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & -4 & 1 & 5 \\ 0 & 0 & 0 & 0 & -1 & 3 & -2 & -5 \end{pmatrix} \end{aligned}$$

Es folgt $U + W = \langle (1, 1, 1, 0), (0, -2, 1, 2), (0, 0, -1, 1) \rangle$ und $U \cap W = \langle (-1, 3, -2, -5) \rangle$.

7 Lineare Abbildungen

7.1 Definitionen und Beispiele

Bemerkung 7.1. Um verschiedene Vektorräume V und W in Beziehung zu setzen, studieren wir Abbildungen $V \rightarrow W$, die Addition und Skalarmultiplikation „respektieren“. Es wird sich zeigen, dass solche Abbildungen durch Matrizen beschrieben werden können.

Definition 7.2. Eine Abbildung $f: V \rightarrow W$ zwischen K -Vektorräumen V und W heißt *linear* oder *Homomorphismus*, falls für alle $u, v \in V$ und $\lambda \in K$ gilt:

$$f(\lambda u + v) = \lambda f(u) + f(v).$$

Die Menge der linearen Abbildungen $V \rightarrow W$ wird mit $\text{Hom}(V, W)$ bezeichnet. Ist f linear und bijektiv, so nennt man f einen *Isomorphismus*. Ggf. nennt man V und W *isomorph* und schreibt $V \cong W$.

Bemerkung 7.3.

(a) Eine Abbildung $f: V \rightarrow W$ ist genau dann linear, wenn

$$\begin{aligned} f(u + v) &= f(u) + f(v), \\ f(\lambda u) &= \lambda f(u) \end{aligned}$$

für alle $u, v \in V$ und $\lambda \in K$ gilt (setze $\lambda = 1$ bzw. $v = 0$ in Definition 7.2; vgl. Bemerkung 3.13). Isomorphe Vektorräume unterscheiden sich daher nur durch die Benennung ihrer Elemente.

(b) Für $f \in \text{Hom}(V, W)$ gilt

$$f(0_V) = f(0_K \cdot 0_V) = 0_K f(0_V) = 0_W.$$

(c) Sei $f \in \text{Hom}(U, V)$ und $g \in \text{Hom}(V, W)$. Für $u, u' \in U$ und $\lambda \in K$ gilt

$$(g \circ f)(\lambda u + u') = g(\lambda f(u) + f(u')) = \lambda g(f(u)) + g(f(u')) = \lambda(g \circ f)(u) + (g \circ f)(u').$$

Dies zeigt $g \circ f \in \text{Hom}(U, W)$.

(d) Ist $f: V \rightarrow W$ ein Isomorphismus, so ist auch $f^{-1}: W \rightarrow V$ ein Isomorphismus¹, denn für $w, w' \in W$ und $\lambda \in K$ gilt

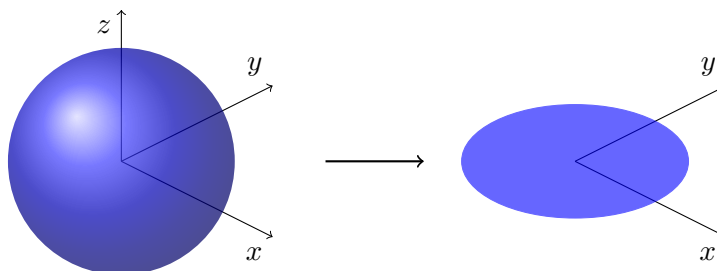
$$\begin{aligned} f^{-1}(\lambda w + w') &= f^{-1}(\lambda f(f^{-1}(w)) + f(f^{-1}(w'))) \\ &= f^{-1}(f(\lambda f^{-1}(w) + f^{-1}(w'))) = \lambda f^{-1}(w) + f^{-1}(w'). \end{aligned}$$

¹In der Analysis ist die Umkehrabbildung einer bijektiven stetigen Funktion im Allgemeinen nicht stetig. Daher gibt es den merkwürdigen Begriff *Homöomorphismus* (kein Tippfehler).

- (e) Die Isomorphie von Vektorräumen ist eine Äquivalenzrelation.² Die Reflexivität folgt aus dem Isomorphismus id_V (Beispiel 7.4), die Symmetrie folgt aus (c) und die Transitivität folgt aus (d).

Beispiel 7.4.

- (a) Die Nullabbildung $0: V \rightarrow W, v \mapsto 0_W$ ist stets linear. Die Identität $\text{id}_V: V \rightarrow V$ ist ein Isomorphismus.
- (b) Für $f \in \text{Hom}(V, W)$ und $U \leq V$ ist die Einschränkung $f|_U$ linear. Insbesondere ist die Inklusionsabbildung $U \rightarrow V$ als Einschränkung von id_V linear.
- (c) Für $n \leq m$ ist die Projektion $K^m \rightarrow K^n, (x_1, \dots, x_m) \mapsto (x_1, \dots, x_n)$ ein surjektiver Homomorphismus. Die Projektion $\mathbb{R}^3 \rightarrow \mathbb{R}^2$ reduziert ein 3-dimensionales Objekt auf seinen „Schatten“:



- (d) Sei $f: \mathbb{R} \rightarrow \mathbb{R}$ linear und $a := f(1)$. Für $x \in \mathbb{R}$ gilt $f(x) = f(x \cdot 1) = x f(1) = ax$. Der Graph von f beschreibt daher eine Gerade durch den Koordinatenursprung. Achtung: In der Schulmathematik werden mitunter auch Funktionen der Form $f(x) = ax + b$ als „linear“ bezeichnet (solche Abbildungen heißen *affin*³, siehe Aufgabe I.21).
- (e) Für $A \in K^{n \times m}$ ist die Abbildung $K^{m \times 1} \rightarrow K^{n \times 1}, x \mapsto Ax$ nach Lemma 5.8 linear. Wir zeigen in Satz 7.18, dass jede lineare Abbildung (nach Basiswahl) diese Form besitzt.
- (f) Die Transposition $K^{n \times m} \rightarrow K^{m \times n}, A \mapsto A^t$ ist ein Isomorphismus.
- (g) Für jede n -elementige Menge $M = \{x_1, \dots, x_n\}$ ist die Abbildung $\text{Abb}(M, V) \rightarrow V^n, f \mapsto (f(x_1), \dots, f(x_n))$ ein Isomorphismus.

Lemma 7.5. Für $f \in \text{Hom}(V, W)$, $V_1 \leq V$ und $W_1 \leq W$ ist $f(V_1) \leq W$ und $f^{-1}(W_1) \leq V$. Insbesondere ist $f(V) \leq W$ und $\text{Ker}(f) := f^{-1}(\{0\}) \leq V$.

Beweis. Wegen $0 = f(0) \in f(V_1)$ ist $f(V_1) \neq \emptyset$. Für $u, v \in V_1$ und $\lambda \in K$ gilt

$$\lambda f(u) + f(v) = f(\lambda u + v) \in f(V_1).$$

Dies zeigt $f(V_1) \leq W$ (Bemerkung 3.13). Wegen $0 \in f^{-1}(\{0\}) \subseteq f^{-1}(W_1)$ ist auch $f^{-1}(W_1) \neq \emptyset$. Für $u, w \in f^{-1}(W_1)$ und $\lambda \in K$ gilt $f(\lambda u + w) = \lambda f(u) + f(w) \in W_1$, d. h. $\lambda u + w \in f^{-1}(W_1)$. Dies zeigt $f^{-1}(W_1) \leq V$. \square

Definition 7.6. In der Situation von Lemma 7.5 nennt man $\text{Ker}(f)$ den *Kern* von f und

$$\text{rk}(f) := \dim f(V)$$

den *Rang* von f . Für $A \in K^{n \times m}$ sei $\text{Ker}(A) := \{x \in K^{m \times 1} : Ax = 0\}$ der *Kern* von A .

²Allerdings ist die Gesamtheit aller Vektorräume keine Menge, sondern eine *Klasse*.

³Ein Beispiel ist die Umrechnung von Grad Celsius nach Fahrenheit: $f(x) = \frac{9}{5}x + 32$.

Lemma 7.7. Genau dann ist $f \in \text{Hom}(V, W)$ injektiv, wenn $\text{Ker}(f) = \{0\}$. Ggf. ist $V \rightarrow f(V)$, $v \mapsto f(v)$ ein Isomorphismus.

Beweis. Sei f injektiv und $v \in \text{Ker}(f)$. Aus $f(v) = 0 = f(0)$ folgt $v = 0$, d. h. $\text{Ker}(f) = \{0\}$. Sei umgekehrt $\text{Ker}(f) = \{0\}$ und $u, v \in V$ mit $f(u) = f(v)$. Dann ist $f(u - v) = f(u) - f(v) = 0$, also $u - v \in \text{Ker}(f) = \{0\}$. Daher ist $u = v$ und f ist injektiv. Die zweite Aussage ist trivial. \square

Satz 7.8. Seien V und W Vektorräume. Sei b_1, \dots, b_n eine Basis von V und seien $c_1, \dots, c_n \in W$ beliebig. Dann existiert genau eine lineare Abbildung $f: V \rightarrow W$ mit $f(b_i) = c_i$ für $i = 1, \dots, n$. Dabei gilt:

- (a) f injektiv $\iff c_1, \dots, c_n$ linear unabhängig.
- (b) f surjektiv $\iff W = \langle c_1, \dots, c_n \rangle$.
- (c) f Isomorphismus $\iff c_1, \dots, c_n$ Basis von W .

Beweis. Jedes $u \in V$ lässt sich eindeutig in der Form $u = \sum_{i=1}^n \lambda_i b_i$ schreiben. Wir definieren

$$f(u) := \sum_{i=1}^n \lambda_i c_i \in W.$$

Für $v = \sum_{i=1}^n \mu_i b_i$ und $\rho \in K$ gilt

$$f(\rho u + v) = f\left(\sum_{i=1}^n (\rho \lambda_i + \mu_i) b_i\right) = \sum_{i=1}^n (\rho \lambda_i + \mu_i) c_i = \rho \sum_{i=1}^n \lambda_i c_i + \sum_{i=1}^n \mu_i c_i = \rho f(u) + f(v).$$

Also ist f linear mit $f(b_i) = c_i$ für $i = 1, \dots, n$. Ist auch $g \in \text{Hom}(V, W)$ mit $g(b_i) = c_i$ für $i = 1, \dots, n$, so gilt

$$g(u) = \sum_{i=1}^n \lambda_i g(b_i) = \sum_{i=1}^n \lambda_i c_i = \sum_{i=1}^n \lambda_i f(b_i) = f(u)$$

für alle $u \in V$. Also ist $g = f$ und f ist eindeutig bestimmt.

- (a) Sei f injektiv und $\sum_{i=1}^n \lambda_i c_i = 0$ für $\lambda_i \in K$. Dann ist

$$f\left(\sum_{i=1}^n \lambda_i b_i\right) = \sum_{i=1}^n \lambda_i c_i = 0.$$

Aus Lemma 7.7 folgt $\sum_{i=1}^n \lambda_i b_i \in \text{Ker}(f) = \{0\}$. Da b_1, \dots, b_n linear unabhängig sind, erhält man $\lambda_1 = \dots = \lambda_n = 0$. Also sind c_1, \dots, c_n linear unabhängig. Seien nun umgekehrt c_1, \dots, c_n linear unabhängig und $u := \sum_{i=1}^n \lambda_i b_i \in \text{Ker}(f)$. Dann ist $\sum_{i=1}^n \lambda_i c_i = f(u) = 0$ und man erhält $\lambda_1 = \dots = \lambda_n = 0$. Daher ist $u = 0$ und $\text{Ker}(f) = \{0\}$. Nach Lemma 7.7 ist f injektiv.

- (b) Sei f surjektiv und $w \in W$. Dann existiert ein $v = \sum_{i=1}^n \lambda_i b_i \in V$ mit $f(v) = w$. Es folgt

$$w = f(v) = \sum_{i=1}^n \lambda_i c_i \in \langle c_1, \dots, c_n \rangle.$$

Sei umgekehrt $W = \langle c_1, \dots, c_n \rangle$ und $w \in W$. Dann existieren $\lambda_i \in K$ mit $w = \sum_{i=1}^n \lambda_i c_i$. Für $v := \sum_{i=1}^n \lambda_i b_i \in V$ gilt dann $f(v) = w$, d. h. f ist surjektiv.

- (c) Folgt aus (a) und (b). \square

Bemerkung 7.9. Sei $f \in \text{Hom}(V, W)$. Im Fall $\dim V < \dim W$ ist f nicht surjektiv, denn das Bild einer Basis von V kann kein Erzeugendensystem von W sein. Im Fall $\dim V > \dim W$ ist f nicht injektiv, denn das Bild einer Basis kann nicht linear unabhängig sein. Für $\dim V = \dim W$ erhält man

$$f \text{ injektiv} \iff f \text{ surjektiv} \iff f \text{ bijektiv}$$

(vgl. Bemerkung 2.11(e)).

Satz 7.10. Zwei K -Vektorräume sind genau dann isomorph, wenn sie die gleiche Dimension haben. Insbesondere ist jeder n -dimensionale K -Vektorraum zu K^n isomorph.

Beweis. Sei $f: V \rightarrow W$ ein Isomorphismus von Vektorräumen und B eine Basis von V . Nach Satz 7.8 ist $f(B)$ eine Basis von W . Also gilt $\dim V = |B| = |f(B)| = \dim W$. Haben umgekehrt V und W die gleiche Dimension, so gibt es Basen $\{b_1, \dots, b_n\}$ und $\{c_1, \dots, c_n\}$ von V bzw. W . Nach Satz 7.8 existiert ein Isomorphismus $f: V \rightarrow W$ mit $f(b_i) = c_i$ für $i = 1, \dots, n$. Die zweite Behauptung folgt aus $\dim K^n = n$. Einen expliziten Isomorphismus erhält man durch die Koordinatendarstellung $V \rightarrow K^n$, $v \mapsto {}_B[v]$ (sie bildet B auf die Standardbasis von K^n ab). \square

Bemerkung 7.11.

- (a) Die Vektorräume $\{0\}, K, K^2, \dots$ bilden ein Repräsentantensystem für die Isomorphieklassen von endlich-dimensionalen K -Vektorräumen.
- (b) Für K -Vektorräume V_1, \dots, V_n mit $d_i := \dim V_i$ gilt $V_1 \times \dots \times V_n \cong K^{d_1} \times \dots \times K^{d_n} \cong K^{d_1 + \dots + d_n}$.
- (c) Obwohl \mathbb{Q} und \mathbb{Q}^2 gleichmächtig sind, gilt $\mathbb{Q} \not\cong \mathbb{Q}^2$ nach Satz 7.10.
- (d) Seien $U, V \leq W$ Vektorräume. Offenbar gilt $V \leq U + V$ und $U \cap V \leq U$. Aus Satz 4.23 und der Dimensionsformel folgt

$$\dim((U + V)/V) = \dim(U + V) - \dim(V) = \dim(U) - \dim(U \cap V) = \dim(U/(U \cap V)).$$

Mit Satz 7.10 erhält man den sogenannten *ersten Isomorphiesatz*⁴

$$\boxed{(U + V)/V \cong U/(U \cap V).}$$

- (e) Für Vektorräume $U \leq V \leq W$ gilt offenbar $V/U \leq W/U$. Aus

$$\begin{aligned} \dim((W/U)/(V/U)) &= \dim(W/U) - \dim(V/U) \\ &= \dim(W) - \dim(U) - \dim(V) + \dim(U) = \dim(W/V) \end{aligned}$$

erhält man den *zweiten Isomorphiesatz*⁵

$$\boxed{(W/U)/(V/U) \cong W/V}$$

(vgl. Aufgabe I.13).

⁴Merkregel: Auf einer Seite stehen zwei U , auf der anderen Seite zwei V .

⁵Merkregel: Kürzen eines Doppelbruchs.

Satz 7.12 (Homomorphiesatz). Für $f \in \text{Hom}(V, W)$ ist die Abbildung

$$\begin{aligned}\bar{f}: V/\text{Ker}(f) &\rightarrow f(V), \\ v + \text{Ker}(f) &\mapsto f(v)\end{aligned}$$

ein Isomorphismus. Also gilt $V/\text{Ker}(f) \cong f(V)$ und $\dim V = \text{rk}(f) + \dim \text{Ker}(f)$.

Beweis. Für $v, w \in V$ gilt

$$\begin{aligned}v + \text{Ker}(f) = w + \text{Ker}(f) &\iff v - w \in \text{Ker}(f) \iff f(v - w) = 0 \\ &\iff f(v) = f(w) \iff \bar{f}(v + \text{Ker}(f)) = \bar{f}(w + \text{Ker}(f)).\end{aligned}$$

Die Implikation \Rightarrow zeigt, dass \bar{f} wohldefiniert ist, während die Implikation \Leftarrow zeigt, dass \bar{f} injektiv ist. Offenbar ist \bar{f} auch linear und surjektiv. Die Gleichung für $\dim V$ folgt aus Satz 4.23. \square

7.2 Darstellungsmatrizen

Satz 7.13. Für K -Vektorräume V und W ist $\text{Hom}(V, W)$ ein Unterraum von $\text{Abb}(V, W)$.

Beweis. Sicher liegt das neutrale Element $f = 0$ von $\text{Abb}(V, W)$ in $\text{Hom}(V, W)$. Seien $f, g \in \text{Hom}(V, W)$, $u, v \in V$ und $\lambda, \mu \in K$. Wegen

$$\begin{aligned}(f + g)(\mu u + v) &= f(\mu u + v) + g(\mu u + v) = \mu f(u) + f(v) + \mu g(u) + g(v) = \mu(f + g)(u) + (f + g)(v), \\ (\lambda f)(\mu u + v) &= \lambda f(\mu u + v) = \lambda(\mu f(u) + f(v)) = \mu \lambda f(u) + \lambda f(v) = \mu(\lambda f)(u) + (\lambda f)(v)\end{aligned}$$

sind $f + g$ und λf linear. \square

Bemerkung 7.14. Ist B eine Basis von V , so ist die Einschränkungabbildung $\text{Hom}(V, W) \rightarrow \text{Abb}(B, W)$, $f \mapsto f|_B$ ein Isomorphismus nach Satz 7.8.

Definition 7.15. Seien V und W Vektorräume mit Basen $B = \{b_1, \dots, b_m\}$ bzw. $C = \{c_1, \dots, c_n\}$. Sei $f \in \text{Hom}(V, W)$ und $f(b_i) = \sum_{j=1}^n a_{ji} c_j$ mit $a_{ji} \in K$ für $i = 1, \dots, m$.

- (a) Man nennt ${}_C[f]_B := (a_{ij}) \in K^{n \times m}$ die *Darstellungsmatrix* von f bzgl. B und C .
- (b) Im Fall $V = W$ und $f = \text{id}_V$ nennt man ${}_C\Delta_B := {}_C[\text{id}_V]_B$ die *Basiswechselmatrix* bzgl. B und C .
- (c) Sind B und C die Standardbasen von $V = K^m$ und $W = K^n$, so setzt man $[f] := {}_C[f]_B$.
Merkregel: Die Spalten von $[f]$ sind die Bilder der Standardbasis.

Bemerkung 7.16. In der Situation von Definition 7.15 gilt ${}_B\Delta_B = 1_m$.

Beispiel 7.17. Die Abbildung

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}^3, \quad (x, y) \mapsto (2x - y, y, -3x)$$

ist linear mit Matrix

$$[f] = \begin{pmatrix} 2 & -1 \\ 0 & 1 \\ -3 & 0 \end{pmatrix}.$$

Offenbar sind $B := \{(1, -1), (0, 2)\}$ und $C := \{(1, 1, 1), (0, -1, 1), (1, 0, 1)\}$ Basen von \mathbb{R}^2 bzw. \mathbb{R}^3 . Wegen

$$\begin{aligned} f(1, -1) &= (3, -1, -3) = -7(1, 1, 1) - 6(0, -1, 1) + 10(1, 0, 1), \\ f(0, 2) &= (-2, 2, 0) = 4(1, 1, 1) + 2(0, -1, 1) - 6(1, 0, 1) \end{aligned}$$

ergibt sich

$$C[f]_B = \begin{pmatrix} -7 & 4 \\ -6 & 2 \\ 10 & -6 \end{pmatrix}$$

(im Zweifel müssen Sie die Einträge durch ein Gleichungssystem bestimmen).

Satz 7.18. Sei V ein m -dimensionaler Vektorraum mit Basis B und sei W ein n -dimensionaler Vektorraum mit Basis C . Dann ist die Abbildung

$$C[\cdot]_B: \text{Hom}(V, W) \rightarrow K^{n \times m}, \quad f \mapsto C[f]_B$$

ein Isomorphismus mit

$$\boxed{C[f(v)]^t = C[f]_B B[v]^t} \quad \begin{array}{ccc} V & \xrightarrow{f} & W \\ \downarrow B[\cdot] & & \downarrow C[\cdot] \\ K^m & \xrightarrow{C[f]_B} & K^n \end{array}$$

für alle $v \in V$. Insbesondere ist $\dim \text{Hom}(V, W) = nm$ und $\text{rk}(f) = \text{rk}(C[f]_B)$.

Beweis. Sei $B = \{b_1, \dots, b_m\}$ und $C = \{c_1, \dots, c_n\}$. Nach Satz 7.8 ist $C[\cdot]_B$ eine Bijektion. Seien $f, g \in \text{Hom}(V, W)$ mit $C[f]_B = (a_{ij})$ und $C[g]_B = (b_{ij})$. Für $i = 1, \dots, m$ und $\lambda \in K$ gilt

$$(\lambda f + g)(b_i) = \lambda f(b_i) + g(b_i) = \lambda \sum_{j=1}^n a_{ji} c_j + \sum_{j=1}^n b_{ji} c_j = \sum_{j=1}^n (\lambda a_{ji} + b_{ji}) c_j.$$

Dies zeigt $C[\lambda f + g]_B = \lambda C[f]_B + C[g]_B$. Also ist $C[\cdot]_B$ ein Isomorphismus. Sei $v = \sum_{i=1}^m v_i b_i$. Dann ist

$$f(v) = \sum_{i=1}^m v_i f(b_i) = \sum_{i=1}^m v_i \sum_{j=1}^n a_{ji} c_j = \sum_{j=1}^n \left(\sum_{i=1}^m a_{ji} v_i \right) c_j$$

und es folgt

$$C[f]_B B[v]^t = \left(\sum_{i=1}^m a_{ji} v_i \right)_j^t = C[f(v)]^t.$$

Offenbar ist $B[b_i] = e_i$ der i -te Standardbasisvektor. Also ist $C[f(b_i)] = C[f]_B B[b_i]^t$ die i -te Spalte von $C[f]_B$. Da $C[\cdot]$ ein Isomorphismus ist, gilt

$$\text{rk}(f) = \dim \langle f(b_1), \dots, f(b_m) \rangle = \dim \langle C[f(b_1)], \dots, C[f(b_m)] \rangle = \text{rk}(C[f]_B). \quad \square$$

Bemerkung 7.19.

- (a) Für $V = W$ und $f = \text{id}_V$ erhält man $C[v]^t = C \Delta_{BB} [v]^t$. Für $f: K^n \rightarrow K^m$ gilt $f(v)^t = [f]v^t$ bzgl. der Standardbasen.

(b) Sei $U \leq V$. Nach Folgerung 4.16 besitzt U ein Komplement W , d. h. $V = U \oplus W$. Die Projektion $f: V \rightarrow W$, $u + w \mapsto w$ für $u \in U$ und $w \in W$ ist ein Homomorphismus mit $\text{Ker}(f) = U$. Daher ist jeder Unterraum Kern eines Homomorphismus. Seien B und C Basen von V bzw. W . Für $A := {}_C[f]_B$ gilt $\text{Ker}(f) = \{v \in V : A_B[v]^t = 0\}$ nach Satz 7.18. Also lässt sich U als Lösungsmenge eines linearen Gleichungssystems beschreiben.

Beispiel 7.20. Seien f , B und C wie in Beispiel 7.17. Für $v := (2, 4) = 2(1, -1) + 3(0, 2) \in \mathbb{R}^2$ gilt

$$f(v)^t = [f]v^t = \begin{pmatrix} 2 & -1 \\ 0 & 1 \\ -3 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \\ -6 \end{pmatrix},$$

$${}_C[f(v)]^t = {}_C[f]_{BB}[v]^t = \begin{pmatrix} -7 & 4 \\ -6 & 2 \\ 10 & -6 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} -2 \\ -6 \\ 2 \end{pmatrix}.$$

Kontrolle: $-2(1, 1, 1) - 6(0, -1, 1) + 2(1, 0, 1) = (0, 4, -6) = f(v)$.

Folgerung 7.21. Sei $f \in \text{Hom}(V, W)$ und $A := {}_C[f]_B \in K^{n \times m}$ bzgl. beliebiger Basen. Dann gilt

(a) f injektiv $\iff \text{rk}(A) = m$.

(b) f surjektiv $\iff \text{rk}(A) = n$.

Beweis.

(a) f injektiv $\stackrel{7.7}{\iff} \text{Ker}(A) = \text{Ker}(f) = \{0\} \stackrel{6.6}{\iff} \text{rk}(A) = m$.

(b) f surjektiv $\iff f(V) = W \iff \text{rk}(A) = \text{rk}(f) = \dim W = n$. □

Satz 7.22. Seien U, V und W Vektorräume mit Basen B, C bzw. D . Seien $f \in \text{Hom}(U, V)$ und $g \in \text{Hom}(V, W)$. Dann ist

$$\boxed{{}_D[g \circ f]_B = {}_D[g]_{CC}{}_C[f]_B.}$$

Beweis. Nach Bemerkung 7.3 ist $g \circ f \in \text{Hom}(U, W)$. Sei $B = \{b_1, \dots, b_m\}$, $C = \{c_1, \dots, c_n\}$ und $D = \{d_1, \dots, d_k\}$. Sei ${}_C[f]_B = (a_{ij})$ und ${}_D[g]_C = (b_{lj})$. Dann gilt

$$(g \circ f)(b_i) = g\left(\sum_{j=1}^n a_{ji}c_j\right) = \sum_{j=1}^n a_{ji}g(c_j) = \sum_{j=1}^n a_{ji} \sum_{l=1}^k b_{lj}d_l = \sum_{l=1}^k \left(\sum_{j=1}^n b_{lj}a_{ji}\right)d_l.$$

Darin ist $\sum_{j=1}^n b_{lj}a_{ji}$ der Eintrag von ${}_D[g]_{CC}{}_C[f]_B$ an Position (l, i) wie gewünscht. □

Bemerkung 7.23. Merkregel: Die Komposition von linearen Abbildungen entspricht der Multiplikation von Matrizen. Für lineare Abbildungen f, g, h zwischen „passenden“ Räumen übertragen sich die Distributivgesetze von Matrizen:

$$(f + g) \circ h = (f \circ h) + (g \circ h) \qquad f \circ (g + h) = (f \circ g) + (f \circ h).$$

Das kann man natürlich auch direkt nachrechnen.

Beispiel 7.24. Seien f , B und C wie in Beispiel 7.17. Sei $g \in \text{Hom}(\mathbb{R}^3, \mathbb{R}^2)$ mit Matrix ${}_B[g]_C = -\begin{pmatrix} 3 & 0 & 2 \\ 1 & 1/2 & 1 \end{pmatrix}$. Dann gilt

$${}_B[g \circ f]_B = {}_B[g]_C {}_C[f]_B = -\begin{pmatrix} 3 & 0 & 2 \\ 1 & 1/2 & 1 \end{pmatrix} \begin{pmatrix} -7 & 4 \\ -6 & 2 \\ 10 & -6 \end{pmatrix} = 1_2 = {}_B[\text{id}_{\mathbb{R}^2}]_B,$$

d. h. $g \circ f = \text{id}_{\mathbb{R}^2}$. Umgekehrt ist $f \circ g \neq \text{id}_{\mathbb{R}^3}$, denn f kann nicht surjektiv sein.

Folgerung 7.25. Sei $f: V \rightarrow W$ ein Isomorphismus zwischen Vektorräumen V und W mit Basen B bzw. C . Dann ist $\boxed{{}_B[f^{-1}]_C = {}_C[f]_B^{-1}}$. Im Fall $V = W$ ist ${}_C\Delta_B^{-1} = {}_B\Delta_C$.

Beweis. Nach Bemerkung 7.3 ist $f^{-1} \in \text{Hom}(W, V)$. Aus Satz 7.22 folgt

$${}_C[f]_B {}_B[f^{-1}]_C = {}_C[\text{id}_W]_C = {}_C\Delta_C = 1_n$$

und ${}_B[f^{-1}]_C = {}_C[f]_B^{-1}$. Die zweite Aussage folgt mit $f = \text{id}_V$. \square

Bemerkung 7.26. Die Isomorphismen $f: V \rightarrow V$ bilden die *allgemeine lineare Gruppe* $\text{GL}(V)$. Sie entsprechen genau den invertierbaren Matrizen, d. h. ${}_B[\cdot]_B: \text{GL}(V) \rightarrow \text{GL}(n, K)$ ist ein Isomorphismus von Gruppen (anstatt von Vektorräumen).

Folgerung 7.27 (Basiswechsel). Seien B, B' Basen von V und C, C' Basen von W . Für $f \in \text{Hom}(V, W)$ gilt

$$\boxed{{}_{C'}[f]_{B'} = {}_{C'}\Delta_{CC} [f]_{BB} \Delta_{B'}.$$

Im Fall $V = W$ ist

$$\boxed{{}_{B'}[f]_{B'} = {}_{B'}\Delta_{BB} [f]_{BB} \Delta_{B'}^{-1}.$$

Beweis. Aus Satz 7.22 folgt

$${}_{C'}\Delta_{CC} [f]_{BB} \Delta_{B'} = {}_{C'}[\text{id}_W]_{CC} [f]_{BB} [\text{id}_V]_{B'} = {}_{C'}[\text{id}_W]_{CC} [f]_{B'} = {}_{C'}[f]_{B'}.$$

Für $V = W$, $C = B$ und $C' = B'$ erhält man ${}_{B'}[f]_{B'} = {}_{B'}\Delta_{BB} [f]_{BB} \Delta_{B'} = {}_{B'}\Delta_{BB} [f]_{BB} \Delta_{B'}^{-1}$ mit Folgerung 7.25. \square

Beispiel 7.28. Sei wieder $f: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ wie in Beispiel 7.17. Wir ersetzen die Basis $B = \{(1, -1), (0, 2)\}$ durch $B' := \{(0, 1), (1, 1)\}$. Wegen $(0, 1) = 0(1, -1) + \frac{1}{2}(0, 2)$ und $(1, 1) = (1, -1) + (0, 2)$ gilt

$${}_B\Delta_{B'} = \begin{pmatrix} 0 & 1 \\ 1/2 & 1 \end{pmatrix}, \quad {}_C[f]_{B'} = {}_C[f]_{BB} \Delta_{B'} = \begin{pmatrix} -7 & 4 \\ -6 & 2 \\ 10 & -6 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1/2 & 1 \end{pmatrix} = \begin{pmatrix} 2 & -3 \\ 1 & -4 \\ -3 & 4 \end{pmatrix}.$$

Definition 7.29. Matrizen $A, B \in K^{n \times n}$ heißen *ähnlich*, falls ein $S \in \text{GL}(n, K)$ mit $B = SAS^{-1}$ existiert. Wir schreiben ggf. $A \approx B$.

Bemerkung 7.30.

- (a) Offenbar ist $A \approx A$ (wähle $S = 1_n$). Aus $B = SAS^{-1}$ folgt $A = S^{-1}BS$. Aus $B = SAS^{-1}$ und $C = TBT^{-1}$ mit $T \in \text{GL}(n, K)$ folgt $C = TSAS^{-1}T^{-1} = (TS)A(TS)^{-1}$. Daher ist die Ähnlichkeit von Matrizen eine Äquivalenzrelation.
- (b) Nach Satz 7.18 und Folgerung 7.27 bestimmt jeder Endomorphismus von V durch Basiswahl eine Ähnlichkeitsklasse von Darstellungsmatrizen in $K^{n \times n}$. Dies erlaubt konkrete Berechnungen mit abstrakten Abbildungen. In der linearen Algebra II konstruieren wir spezielle Basen, sodass die Darstellungsmatrizen möglichst „einfache“ Gestalt haben (zum Beispiel Diagonalmatrizen). Dies beschleunigt Berechnungen.

Definition 7.31. Für $A = (a_{ij})_{i,j} \in K^{n \times n}$ nennen wir $\text{tr}(A) := \sum_{i=1}^n a_{ii}$ die *Spur* von A . Dies ist die Summe der Hauptdiagonaleinträge.

Lemma 7.32. Die Abbildung $\text{tr}: K^{n \times n} \rightarrow K$ ist linear mit $\text{tr}(A^t) = \text{tr}(A)$ und $\text{tr}(AB) = \text{tr}(BA)$ für $A, B \in K^{n \times n}$.

Beweis. Für $A = (a_{ij}), B = (b_{ij})$ und $\lambda \in K$ gilt

$$\text{tr}(\lambda A + B) = \text{tr}((\lambda a_{ij} + b_{ij})_{i,j}) = \sum_{i=1}^n (\lambda a_{ii} + b_{ii}) = \lambda \sum_{i=1}^n a_{ii} + \sum_{i=1}^n b_{ii} = \lambda \text{tr}(A) + \text{tr}(B).$$

Also ist tr linear. Da eine Spiegelung an der Hauptdiagonale diese selbst nicht ändert, gilt $\text{tr}(A^t) = \text{tr}(A)$. Außerdem gilt

$$\text{tr}(AB) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} b_{ji} = \sum_{j=1}^n \sum_{i=1}^n b_{ji} a_{ij} = \text{tr}(BA). \quad \square$$

Folgerung 7.33. Ähnliche Matrizen haben die gleiche Spur.

Beweis. Für $A \in K^{n \times n}$ und $S \in \text{GL}(n, K)$ gilt $\text{tr}(S(AS^{-1})) = \text{tr}((AS^{-1})S) = \text{tr}(A)$. □

Definition 7.34. Sei V ein Vektorraum mit Basis B und $f \in \text{Hom}(V, V)$. Wir nennen $\text{tr}(f) := \text{tr}_B[f]_B$ die *Spur* von f . Nach Folgerung 7.27 und Folgerung 7.33 hängt $\text{tr}(f)$ nicht von der Wahl der Basis ab.

Satz 7.35 (FILLMORE). Sei $A \in K^{n \times n} \setminus K1_n$ und $d_1, \dots, d_n \in K$ mit $\text{tr}(A) = d_1 + \dots + d_n$. Dann ist A zu einer Matrix mit Hauptdiagonale d_1, \dots, d_n ähnlich.

Beweis. Induktion nach n : Wegen $A \notin K1_n$ ist $n \geq 2$. Sei $A = (a_{ij})$. Gilt $a_{ij} \neq 0$ für gewisse $i \neq j$, so ist $Ae_j \notin \langle e_j \rangle$. Ist A eine Diagonalmatrix, so existieren $i \neq j$ mit $a_{ii} \neq a_{jj}$. In diesem Fall ist

$$A(e_i + e_j) = a_{ii}e_i + a_{jj}e_j \notin \langle e_i + e_j \rangle.$$

In jedem Fall existiert ein $b_1 \in K^n$, sodass b_1 und Ab_1 linear unabhängig sind. Dann sind auch b_1 und $b_2 := Ab_1 - d_1b_1$ linear unabhängig. Wir ergänzen b_1, b_2 zu einer Basis b_1, \dots, b_n von K^n . Bezüglich dieser Basis hat A die Form $\begin{pmatrix} d_1 & * \\ * & A_1 \end{pmatrix}$ mit $A_1 = (a'_{ij}) \in K^{(n-1) \times (n-1)}$. Im Fall $n = 2$ sind wir fertig, denn $\text{tr}(A_1) = \text{tr}(A) - d_1 = d_2$.

Sei nun $n \geq 3$. Ist $A_1 \in K1_{n-1}$, so gilt

$$A(b_1 + b_3) = d_1 b_1 + b_2 + A b_3 \in b_2 + \langle b_1, b_3 \rangle.$$

Indem wir b_3 durch $b_1 + b_3$ ersetzen, erreichen wir $a'_{23} = 1$ und $A_1 \notin K1_{n-1}$. Nach Induktion existiert $S \in \text{GL}(n-1, K)$, sodass SA_1S^{-1} Hauptdiagonale d_2, \dots, d_n hat. Nun hat

$$\begin{pmatrix} 1 & 0 \\ 0 & S \end{pmatrix} \begin{pmatrix} d_1 & * \\ * & A_1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & S^{-1} \end{pmatrix} = \begin{pmatrix} d_1 & * \\ * & SA_1S^{-1} \end{pmatrix}$$

Hauptdiagonale d_1, \dots, d_n . □

Beispiel 7.36. Sei $A := \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \in \mathbb{Q}^{2 \times 2}$ und $(d_1, d_2) = (0, 3)$. Wir setzen $b_1 := (1, 1)$ und $b_2 := Ab_1 = (1, 2)$. Sei E die Standardbasis von \mathbb{Q}^2 und $B := \{b_1, b_2\}$. Für $S := {}_E\Delta_B = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ gilt

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

nach Folgerung 7.27.

7.3 Dualräume

Definition 7.37. Für einen K -Vektorraum V nennt man $V^* := \text{Hom}(V, K)$ den *Dualraum* von V . Seine Elemente nennt man (*lineare*) *Funktionale*.

Lemma 7.38. Sei b_1, \dots, b_n eine Basis von V . Für $i = 1, \dots, n$ sei $b_i^* \in V^*$ mit $b_i^*(b_j) = \delta_{ij}$ für $j = 1, \dots, n$. Dann ist b_1^*, \dots, b_n^* eine Basis von V^* .

Beweis. Seien $\lambda_1, \dots, \lambda_n \in K$ mit $f := \lambda_1 b_1^* + \dots + \lambda_n b_n^* = 0$. Für $i = 1, \dots, n$ gilt $0 = f(b_i) = \lambda_i$. Daher sind b_1^*, \dots, b_n^* linear unabhängig. Die Behauptung folgt aus $\dim V^* = \dim V$ (Satz 7.18). □

Beispiel 7.39. Ist e_1, \dots, e_n die Standardbasis von $V = K^n$, so sind e_1^*, \dots, e_n^* die Projektionen, d. h. $e_i^*(v_1, \dots, v_n) = v_i$ für $i = 1, \dots, n$.

Bemerkung 7.40.

- (a) In der Situation von Lemma 7.38 nennt man b_1^*, \dots, b_n^* die zu b_1, \dots, b_n *duale Basis*.
- (b) Für unendlich-dimensionale Räume V ist $V^* \not\cong V$, denn V^* ist „größer“ als V (ohne Beweis).
- (c) Nach (dem Beweis von) Satz 7.13 existiert ein Isomorphismus $V \rightarrow V^*$, der eine Basis auf die entsprechende duale Basis abbildet. Allerdings gibt es keinen *kanonischen* Isomorphismus, der nicht von einer Basiswahl abhängt (vgl. Satz 12.7). Der nächste Satz zeigt, dass die Situation zwischen V und dem *Bidualraum* $V^{**} := (V^*)^*$ besser ist.

Satz 7.41. Für $v \in V$ sei $F_v: V^{**} \rightarrow K$, $f \mapsto f(v)$. Dann ist $F: V \rightarrow V^{**}$, $v \mapsto F_v$ ein *kanonischer Isomorphismus*.

Beweis. Für $f_1, f_2 \in V^*$ und $\lambda \in K$ ist

$$F_v(\lambda f_1 + f_2) = (\lambda f_1 + f_2)(v) = \lambda f_1(v) + f_2(v) = \lambda F_v(f_1) + F_v(f_2),$$

d. h. $F_v \in V^{**}$. Für $v, w \in V$ gilt

$$F_{\lambda v + w}(f) = f(\lambda v + w) = \lambda f(v) + f(w) = \lambda F_v(f) + F_w(f) = (\lambda F_v + F_w)(f),$$

d. h. F ist linear. Sei nun $F_v = 0$. Im Fall $v \neq 0$ kann man v zu einer Basis von V fortsetzen. Für die duale Basis wäre dann $0 = F_v(v^*) = v^*(v) = 1$. Dieser Widerspruch zeigt $\text{Ker}(F) = \{0\}$, d. h. F ist injektiv. Wegen $\dim V^{**} = \dim V^* = \dim V < \infty$ muss F auch surjektiv sein. \square

Definition 7.42. Für $U \leq V$ und $W \leq V^*$ sei

$$U^0 := \{f \in V^* : f(U) = \{0\}\} \subseteq V^*,$$

$$W_0 := \{v \in V : \forall f \in W : f(v) = 0\} \subseteq V.$$

Wir nennen U^0 (bzw. W_0) das *duale Komplement* von U (bzw. W).

Lemma 7.43.

- (a) Für $U \leq V$ ist $U^0 \leq V^*$ mit $\dim V = \dim U + \dim U^0$.
- (b) Für $U \leq V^*$ ist $U_0 \leq V$ mit $\dim V = \dim U + \dim U_0$.
- (c) Die Abbildungen $U \mapsto U^0$ und $U \mapsto U_0$ sind zueinander invers.
- (d) Für $U, W \leq V$ gilt $(U + W)^0 = U^0 \cap W^0$ und $(U \cap W)^0 = U^0 + W^0$.
- (e) Für $U, W \leq V^*$ gilt $(U + W)_0 = U_0 \cap W_0$ und $(U \cap W)_0 = U_0 + W_0$.
- (f) Es gilt $V = U \oplus W \iff V^* = U^0 \oplus W^0$.

Beweis.

- (a) Die Einschränkung $F: V^* \rightarrow U^*$, $f \mapsto f|_U$ ist ein Homomorphismus mit Kern $U^0 \leq V^*$. Da man jedes Funktional in U^* nach V^* fortsetzen kann (Basisergänzung), ist F surjektiv. Aus dem Homomorphiesatz folgt $\dim V = \dim V^* = \dim U^0 + \dim U^* = \dim U^0 + \dim U$.
- (b) Die Konstruktion aus (a) liefert zunächst $U^0 = \{f \in V^{**} : f(U) = \{0\}\} \leq V^{**}$ mit $\dim V = \dim U + \dim U^0$. Für den Isomorphismus $F: V \rightarrow V^{**}$, $v \mapsto F_v$ aus Satz 7.41 gilt

$$v \in U_0 \iff F_v(U) = \{0\} \iff F_v \in U^0.$$

Also ist $U_0 = F^{-1}(U^0) \leq V$ mit $\dim U_0 = \dim U^0$.

- (c) Nach Definition ist $U \leq (U^0)_0$ und $U \leq (U_0)^0$. Aus Dimensionsgründen gilt Gleichheit.
- (d) Es gilt $f \in (U + W)^0 \iff f(U) = \{0\} = f(W) \iff f \in U^0 \cap W^0$. Dies zeigt die erste Gleichung. Aus $U \cap W \leq U, W$ folgt $U^0 + W^0 = \langle U^0 \cup W^0 \rangle \subseteq (U \cap W)^0$. Nach (a) und der Dimensionsformel gilt

$$\begin{aligned} \dim(U^0 + W^0) &= \dim U^0 + \dim W^0 - \dim(U^0 \cap W^0) \\ &= 2 \dim V - \dim U - \dim W - \dim((U + W)^0) \end{aligned}$$

$$\begin{aligned}
&= \dim V - \dim U - \dim W + \dim(U + W) \\
&= \dim V - \dim(U \cap W) = \dim((U \cap W)^0).
\end{aligned}$$

Also ist $(U \cap W)^0 = U^0 + W^0$.

(e) Nach (c) und (d) gilt

$$\begin{aligned}
(U + W)_0 &= ((U_0)^0 + (W_0)^0)_0 = ((U_0 \cap W_0)^0)_0 = U_0 \cap W_0, \\
(U \cap W)_0 &= ((U_0)^0 \cap (W_0)^0)_0 = ((U_0 + W_0)^0)_0 = U_0 + W_0.
\end{aligned}$$

(f) Sei $V = U \oplus W$. Nach (d) ist $U^0 + W^0 = (U \cap W)^0 = \{0\}^0 = V^*$ und $U^0 \cap W^0 = (U + W)^0 = V^0 = \{0\}$. Also ist $V^* = U^* \oplus W^*$. Sei umgekehrt $V^* = U^0 \oplus W^0$. Aus (e) folgt

$$U + W = (U^0)_0 + (W^0)_0 = (U^0 \cap W^0)_0 = \{0\}_0 = V$$

und $U \cap W = (U^0)_0 \cap (W^0)_0 = (U^0 + W^0)_0 = (V^*)_0 = \{0\}$. Dies zeigt $V = U \oplus W$. \square

Folgerung 7.44. Sei V ein n -dimensionaler Vektorraum und $0 \leq k \leq n$. Dann gibt es eine Bijektion zwischen der Menge der k -dimensionalen Unterräume und der Menge der $(n - k)$ -dimensionalen Unterräume von V .

Beweis. Sei $F: V \rightarrow V^*$ ein beliebiger Isomorphismus. Sei $U \leq V$ mit Dimension k . Nach Lemma 7.43 haben $F(U)_0 \leq V$ und $F^{-1}(U^0) \leq V$ die Dimension $n - k$. Außerdem gilt

$$\begin{aligned}
F^{-1}((F(U)_0)^0) &= F^{-1}(F(U)) = U, \\
F(F^{-1}(U^0))_0 &= (U^0)_0 = U.
\end{aligned}$$

Daher sind die Abbildungen $U \mapsto F(U)_0$ und $U \mapsto F^{-1}(U^0)$ zueinander inverse Bijektionen zwischen der Menge der k -dimensionalen Unterräume und der Menge der $(n - k)$ -dimensionalen Unterräume von V . \square

Satz 7.45. Sei V ein n -dimensionaler Vektorraum über einem Körper K mit $q < \infty$ Elementen. Für $1 \leq k \leq n$ besitzt V genau

$$\frac{(q^n - 1)(q^{n-1} - 1) \dots (q^{n-k+1} - 1)}{(q^k - 1)(q^{k-1} - 1) \dots (q - 1)}$$

Unterräume der Dimension k .

Beweis. Jeder k -dimensionale Unterraum $U \leq V$ wird durch ein k -Tupel linear unabhängiger Vektoren (v_1, \dots, v_k) aufgespannt. Für $v_1 \in V \setminus \{0\}$ hat man $|V| - 1 = q^n - 1$ Möglichkeiten. Wegen der linearen Unabhängigkeit darf v_2 nicht in $\langle v_1 \rangle \cong K$ liegen. Daher gibt es $q^n - q$ Möglichkeiten für $v_2 \in V \setminus \langle v_1 \rangle$. Allgemein hat man $q^n - q^{i-1}$ Möglichkeiten für die Wahl von $v_i \in V \setminus \langle v_1, \dots, v_{i-1} \rangle$. Insgesamt gibt es $D(n, k) := (q^n - 1)(q^n - q) \dots (q^n - q^{k-1})$ linear unabhängige k -Tupel (v_1, \dots, v_k) in V . Allerdings spannen viele davon den gleichen Raum auf.

Das gleiche Argument mit U anstelle von V liefert genau $D(k, k) = (q^k - 1)(q^k - q) \dots (q^k - q^{k-1})$ linear unabhängige k -Tupel in U . Also spannen genau $D(k, k)$ der k -Tupel in V den gleichen Unterraum auf. Die Anzahl der k -dimensionalen Unterräume ist daher $\frac{D(n, k)}{D(k, k)}$. Die Behauptung folgt, indem man alle Faktoren q kürzt. \square

Bemerkung 7.46. Auf den ersten Blick ist nicht klar, warum die in Satz 7.45 angegebene Formel überhaupt eine ganze Zahl ist. Wir haben also eine arithmetische Aussage mit Hilfe der linearen Algebra bewiesen. Wegen

$$\frac{(q^n - 1)(q^{n-1} - 1) \dots (q^{n-k+1} - 1)}{(q^k - 1)(q^{k-1} - 1) \dots (q - 1)} = \frac{(q^n - 1)(q^{n-1} - 1) \dots (q^{k+1} - 1)}{(q^{n-k} - 1)(q^{n-k-1} - 1) \dots (q - 1)}$$

erhält man Folgerung 7.44 für endliche Körper.

Beispiel 7.47. Die Anzahl der 2-dimensionalen Unterräume von \mathbb{F}_2^5 ist

$$\frac{(2^5 - 1)(2^4 - 1)}{(2^2 - 1)(2 - 1)} = \frac{31 \cdot 15}{3} = 155.$$

Definition 7.48. Seien V, W Vektorräume und $f \in \text{Hom}(V, W)$. Dann heißt $f^* : W^* \rightarrow V^*$, $g \mapsto g \circ f$ die zu f *duale Abbildung*.

Satz 7.49. Für Vektorräume V, W ist $\text{Hom}(V, W) \rightarrow \text{Hom}(W^*, V^*)$, $f \mapsto f^*$ ein Isomorphismus. Seien B und C Basen von V bzw. W . Seien B^* und C^* die entsprechenden dualen Basen. Dann gilt

$$\boxed{B^*[f^*]_{C^*} = C[f]_B^t}$$

Insbesondere ist

- (a) f injektiv $\iff f^*$ surjektiv.
- (b) f surjektiv $\iff f^*$ injektiv.

Beweis. Für $g_1, g_2 \in W^*$ und $\lambda \in K$ gilt

$$f^*(\lambda g_1 + g_2) = (\lambda g_1 + g_2) \circ f = \lambda(g_1 \circ f) + (g_2 \circ f) = \lambda f^*(g_1) + f^*(g_2)$$

nach Bemerkung 7.23. Dies zeigt $f^* \in \text{Hom}(W^*, V^*)$. Für $f_1, f_2 \in \text{Hom}(V, W)$ gilt analog

$$(\lambda f_1 + f_2)^*(g) = g \circ (\lambda f_1 + f_2) = \lambda(g \circ f_1) + (g \circ f_2) = (\lambda f_1^* + f_2^*)(g).$$

Daher ist $f \mapsto f^*$ linear. Sei $B = \{b_1, \dots, b_m\}$ und $C = \{c_1, \dots, c_n\}$. Sei $f(b_i) = \sum_{j=1}^n a_{ji} c_j$ mit $a_{ji} \in K$. Dann gilt

$$f^*(c_i^*)(b_j) = (c_i^* \circ f)(b_j) = c_i^*\left(\sum_{k=1}^n a_{kj} c_k\right) = a_{ij} = \left(\sum_{k=1}^n a_{ik} b_k^*\right)(b_j)$$

für $j = 1, \dots, m$. Es folgt $f^*(c_i^*) = \sum_{k=1}^n a_{ik} b_k^*$ für $i = 1, \dots, n$. Dies zeigt $C[f]_B^t = (a_{ji})^t = (a_{ij}) = B^*[f^*]_{C^*}$. Nach Satz 7.18 ist $f \mapsto f^*$ ein Isomorphismus.

Mit Folgerung 7.21 ergibt sich

$$\begin{aligned} f \text{ injektiv} &\iff \text{rk}(C[f]_B) = m \iff \text{rk}(B^*[f^*]_{C^*}) = m \iff f^* \text{ surjektiv,} \\ f \text{ surjektiv} &\iff \text{rk}(C[f]_B) = n \iff \text{rk}(B^*[f^*]_{C^*}) = n \iff f^* \text{ injektiv.} \quad \square \end{aligned}$$

Bemerkung 7.50. Für $f \in \text{Hom}(V, W)$ und $g \in \text{Hom}(W, U)$ gilt $(g \circ f)^* = f^* \circ g^*$, denn

$$(g \circ f)^*(\varphi) = \varphi \circ (g \circ f) = (\varphi \circ g) \circ f = g^*(\varphi) \circ f = f^*(g^*(\varphi)) = (f^* \circ g^*)(\varphi)$$

für $\varphi \in U^*$. Alternativ kann man die Matrixidentität $(AB)^t = B^t A^t$ aus Lemma 5.8 benutzen.

8 Eigenwerte und Eigenvektoren

8.1 Definitionen und Beispiele

Bemerkung 8.1. In diesem Kapitel untersuchen wir Homomorphismen $f: V \rightarrow V$ zwischen den gleichen Räumen. Man spricht dann von *Endomorphismen* und schreibt $\text{End}(V) := \text{Hom}(V, V)$. Durch Wahl einer geeigneten Basis B von V werden wir erreichen, dass ${}_B[f]_B$ möglichst einfache Gestalt hat.¹

Definition 8.2. Sei V ein K -Vektorraum und $f \in \text{End}(V)$. Man nennt $\lambda \in K$ einen *Eigenwert*² von f , falls der *Eigenraum*

$$E_\lambda(f) := \{v \in V : f(v) = \lambda v\}$$

nicht der Nullraum ist. Ggf. nennt man $\dim E_\lambda(f)$ die *geometrische Vielfachheit* von λ . Die Vektoren $v \in E_\lambda(f) \setminus \{0\}$ heißen *Eigenvektoren* zum Eigenwert λ .

Bemerkung 8.3.

(a) Für $v \in V$ und $\lambda \in K$ gilt

$$f(v) = \lambda v \iff f(v) - \lambda v = 0 \iff (f - \lambda \text{id})(v) = 0 \iff v \in \text{Ker}(f - \lambda \text{id}).$$

Daher ist der Eigenraum $E_\lambda(f) = \text{Ker}(f - \lambda \text{id})$ tatsächlich ein Unterraum von V . Nach dem Homomorphiesatz ist $\dim(V) - \text{rk}(f - \lambda \text{id})$ die geometrische Vielfachheit von λ .

(b) Sei B eine Basis von V , $x := {}_B[v]^t \in K^{n \times 1}$ und $A := {}_B[f]_B$. Dann gilt

$$f(v) = \lambda v \stackrel{7.18}{\iff} Ax = \lambda x \iff (A - \lambda 1_n)x = 0.$$

Daher lässt sich $E_\lambda(f)$ durch Lösen des homogenen Gleichungssystems $(A - \lambda 1_n)x = 0$ berechnen. Wir sprechen dann auch von Eigenwerten, Eigenräumen und Eigenvektoren von A . Da ähnliche Matrizen denselben Endomorphismus (bzgl. verschiedener Basen) beschreiben, haben sie die gleichen Eigenwerte (aber nicht die gleichen Eigenvektoren).

(c) Aus Lemma 7.7 und Bemerkung 7.9 folgt: $f \in \text{End}(V)$ ist genau dann ein Isomorphismus, wenn 0 *kein* Eigenwert von f ist.

Beispiel 8.4. Sei $f \in \text{End}(\mathbb{R}^3)$ mit

$$A := [f] = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}.$$

¹Nach Bemerkung 6.20 findet man stets Basen B, C von V , sodass ${}_C[f]_B = \text{diag}(1_r, 0_{n-r})$ gilt. Das Produkt solcher Matrizen lässt sich allerdings nicht sinnvoll interpretieren.

²Obwohl auch im Englischen von *eigenvalues* die Rede ist, gab es keinen Namensgeber EIGEN.

Zieht man $\lambda = 1$ auf der Hauptdiagonale ab, so erhält man eine Matrix vom Rang 1 mit drei identischen Zeilen. Offenbar bilden $b_1 := (1, 0, -1)$ und $b_2 := (0, 1, -1)$ eine Basis von $E_1(f)$. Insbesondere hat $\lambda = 1$ geometrische Vielfachheit 2. Da die Zeilensummen von A konstant sind, ist $b_3 := (1, 1, 1)$ ein Eigenvektor zum Eigenwert $\lambda = 4$. Nun ist $B := \{b_1, b_2, b_3\}$ eine Basis von \mathbb{R}^3 mit ${}_B[f]_B = \text{diag}(1, 1, 4)$. Damit berechnet man leicht ${}_B[f \circ f \circ f]_B = {}_B[f]_B^3 = \text{diag}(1, 1, 4)^3 = \text{diag}(1, 1, 64)$.

8.2 Diagonalisierbarkeit

Definition 8.5. Man nennt $f \in \text{End}(V)$ *diagonalisierbar*, falls eine Basis B von V existiert, sodass ${}_B[f]_B$ eine Diagonalmatrix ist. Eine Matrix $A \in K^{n \times n}$ heißt *diagonalisierbar*, falls A zu einer Diagonalmatrix ähnlich ist.

Bemerkung 8.6. Offenbar ist $f \in \text{End}(V)$ genau dann diagonalisierbar, wenn V eine Basis aus Eigenvektoren von f besitzt. Eine Matrix A ist genau dann diagonalisierbar, wenn die entsprechende lineare Abbildung $f: K^{n \times 1} \rightarrow K^{n \times 1}$, $x \mapsto Ax$ diagonalisierbar ist (Folgerung 7.27).

Beispiel 8.7.

- (a) Die Abbildung aus Beispiel 8.4 ist diagonalisierbar.
- (b) Diagonalmatrizen sind offensichtlich diagonalisierbar.
- (c) Sei $A := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in K^{2 \times 2}$. Da $A - \lambda 1_2$ für $\lambda \neq 0$ vollen Rang hat, ist $\lambda = 0$ der einzige Eigenwert von A . Wegen $E_0(A) = \langle (1, 0) \rangle$ existiert keine Basis aus Eigenvektoren und A ist *nicht* diagonalisierbar.
- (d) Sei $A := \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix} \in \mathbb{Q}^{2 \times 2}$ und $\lambda \in \mathbb{Q}$. Dann ist

$$A - \lambda 1_2 = \begin{pmatrix} -\lambda & 2 \\ 1 & -\lambda \end{pmatrix} \begin{array}{c} \leftarrow \\ \rightarrow \end{array} \begin{array}{c} + \\ - \end{array} \begin{array}{c} \lambda \\ \lambda \end{array} \sim \begin{pmatrix} 0 & 2 - \lambda^2 \\ 1 & -\lambda \end{pmatrix}.$$

Wegen $\sqrt{2} \notin \mathbb{Q}$ (Beispiel 1.11) ist $2 - \lambda^2 \neq 0$. Also besitzt A keinen Eigenwert über \mathbb{Q} und kann nicht diagonalisierbar sein. Andererseits ist A als $\mathbb{R}^{2 \times 2}$ -Matrix diagonalisierbar (Beispiel 8.13).

Definition 8.8. Wir haben in Definition 4.2 die (direkte) Summe $U + W$ (bzw. $U \oplus W$) von zwei Unterräumen $U, W \leq V$ eingeführt. Für $U_1, \dots, U_n \leq V$ definiert man induktiv

$$U_1 + \dots + U_n := (U_1 + \dots + U_{n-1}) + U_n \leq V.$$

Offenbar besteht $U_1 + \dots + U_n$ aus den Elementen der Form $u_1 + \dots + u_n$ mit $u_i \in U_i$ für $i = 1, \dots, n$. Wir nennen die Summe *direkt* und schreiben $U_1 \oplus \dots \oplus U_n$, falls $U_1 + \dots + U_{n-1} = U_1 \oplus \dots \oplus U_{n-1}$ und $(U_1 + \dots + U_{n-1}) \cap U_n = \{0\}$. Handlicher ist die folgende Charakterisierung.

Lemma 8.9. Für Unterräume U_1, \dots, U_n eines Vektorraums V sind die folgenden Aussagen äquivalent:

- (1) $U_1 + \dots + U_n = U_1 \oplus \dots \oplus U_n$.
- (2) $\dim(U_1 + \dots + U_n) = \dim(U_1) + \dots + \dim(U_n)$.
- (3) Die Abbildung $U_1 \times \dots \times U_n \rightarrow U_1 + \dots + U_n$, $(u_1, \dots, u_n) \mapsto u_1 + \dots + u_n$ ist ein Isomorphismus.
- (4) Jedes Element $u \in U_1 + \dots + U_n$ lässt sich eindeutig in der Form $u = u_1 + \dots + u_n$ mit $u_i \in U_i$ für $i = 1, \dots, n$ schreiben.

(5) Ist $u_1 + \dots + u_n = 0$ mit $u_i \in U_i$ für $i = 1, \dots, n$, so folgt $u_1 = \dots = u_n = 0$.

Beweis.

(1) \Rightarrow (2): Für $n = 1$ ist (2) trivial. Induktiv dürfen wir annehmen, dass (2) bereits für $n - 1$ gilt. Aus dem Dimensionssatz folgt dann

$$\dim(U_1 + \dots + U_n) = \dim(U_1 + \dots + U_{n-1}) + \dim(U_n) = \dim(U_1) + \dots + \dim(U_n).$$

(2) \Rightarrow (3): Die gegebene Abbildung ist stets linear und surjektiv. Wegen

$$\dim(U_1 \times \dots \times U_n) \stackrel{7.11}{=} \dim(U_1) + \dots + \dim(U_n) = \dim(U_1 + \dots + U_n)$$

muss sie auch injektiv sein.

(3) \Rightarrow (4): Ergibt sich aus der Injektivität von $U_1 \times \dots \times U_n \rightarrow U_1 + \dots + U_n$.

(4) \Rightarrow (5): Die beiden Zerlegungen des Nullvektors $u_1 + \dots + u_n = 0 + \dots + 0$ müssen nach (4) identisch sein, d. h. $u_1 = \dots = u_n = 0$.

(5) \Rightarrow (1): Für $n = 1$ ist nichts zu zeigen. Induktiv können wir $U_1 + \dots + U_{n-1} = U_1 \oplus \dots \oplus U_{n-1}$ annehmen, denn die Voraussetzung (5) überträgt sich auf U_1, \dots, U_{n-1} . Sei nun $u = u_1 + \dots + u_{n-1} \in (U_1 + \dots + U_{n-1}) \cap U_n$. Dann ist

$$0 = u_1 + \dots + u_{n-1} - u \in U_1 + \dots + U_n$$

und (5) zeigt $u = 0$. Also gilt (1). □

Satz 8.10. Seien $\lambda_1, \dots, \lambda_k$ paarweise verschiedene Eigenwerte von $f \in \text{End}(V)$. Dann gilt

$$E_{\lambda_1}(f) + \dots + E_{\lambda_k}(f) = E_{\lambda_1}(f) \oplus \dots \oplus E_{\lambda_k}(f) \leq V. \quad (8.1)$$

Insbesondere ist $k \leq \dim V$.

Beweis. Induktion nach k : Für $k = 1$ ist nichts zu zeigen. Sei also $k \geq 2$ und (8.1) für $k - 1$ bereits bewiesen. Seien $v_i \in E_{\lambda_i}(f)$ mit $v_1 + \dots + v_k = 0$. Dann gilt

$$\begin{aligned} 0 &= f(v_1 + \dots + v_k) - \lambda_k(v_1 + \dots + v_k) = \lambda_1 v_1 + \dots + \lambda_k v_k - \lambda_k v_1 - \dots - \lambda_k v_k \\ &= (\lambda_1 - \lambda_k)v_1 + \dots + (\lambda_{k-1} - \lambda_k)v_{k-1} \in E_{\lambda_1}(f) \oplus \dots \oplus E_{\lambda_{k-1}}(f). \end{aligned}$$

Aus Lemma 8.9 folgt $(\lambda_i - \lambda_k)v_i = 0$ für $i = 1, \dots, k - 1$. Wegen $\lambda_i \neq \lambda_k$ gilt $v_1 = \dots = v_{k-1} = 0$. Schließlich ist auch $v_k = v_1 + \dots + v_k = 0$. Nun ergibt sich (8.1) aus Lemma 8.9. Die letzte Behauptung folgt aus $\dim E_{\lambda_i}(f) \geq 1$ für $i = 1, \dots, k$. □

Bemerkung 8.11. Merkgel: Eigenvektoren zu verschiedenen Eigenwerten sind linear unabhängig.

Folgerung 8.12. Besitzt $A \in K^{n \times n}$ genau n verschiedene Eigenwerte, so ist A diagonalisierbar.

Beweis. Für die verschiedenen Eigenwerte $\lambda_1, \dots, \lambda_n$ von A gilt

$$\dim(E_{\lambda_1}(A) \oplus \dots \oplus E_{\lambda_n}(A)) = \dim(E_{\lambda_1}(A)) + \dots + \dim(E_{\lambda_n}(A)) \geq n = \dim K^{n \times 1}$$

nach Satz 8.10. Dies zeigt $E_{\lambda_1}(A) \oplus \dots \oplus E_{\lambda_n}(A) = K^{n \times 1}$. Insbesondere besitzt $K^{n \times 1}$ eine Basis aus Eigenvektoren von A . □

Beispiel 8.13.

- (a) Die Einheitsmatrix zeigt, dass die Umkehrung von Folgerung 8.12 falsch ist. Wir leiten später eine genaue Charakterisierung der Diagonalisierbarkeit her (Satz 10.34).
- (b) Die Matrix $A := \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix} \in \mathbb{R}^{2 \times 2}$ aus Beispiel 8.7 besitzt die Eigenwerte $\pm\sqrt{2}$ und ist daher diagonalisierbar. Wir zeigen über den Umweg der Determinante, dass die Eigenwerte jeder Matrix Nullstellen von Polynomen mit Koeffizienten in K sind (Satz 10.32).

Definition 8.14. Man nennt $A = (a_{ij}) \in K^{n \times n}$ eine (*obere*³) *Dreiecksmatrix*, falls alle Einträge unterhalb der Hauptdiagonalen verschwinden, d. h. $a_{ij} = 0$ für alle $i > j$:

$$A = \begin{pmatrix} * & \cdots & * \\ & \ddots & \vdots \\ 0 & & * \end{pmatrix}.$$

Gilt zusätzlich $a_{ii} = 0$ für $i = 1, \dots, n$, so spricht man von einer *strikten* (oberen) Dreiecksmatrix.

Beispiel 8.15. Sei $A = (a_{ij}) \in K^{n \times n}$ eine obere Dreiecksmatrix. Für $\lambda \in \{a_{11}, \dots, a_{nn}\}$ hat $A - \lambda 1_n$ nicht vollen Rang, denn beim Gauß-Algorithmus tritt ein Versatz der Zeilen auf (Bemerkung 6.20). Ist andererseits $\lambda \notin \{a_{11}, \dots, a_{nn}\}$, so sind die Hauptdiagonaleinträge von $A - \lambda 1_n$ alle ungleich 0. Daher hat $A - \lambda 1_n$ vollen Rang. Dies zeigt, dass die Eigenwerte von A genau die Einträge auf der Hauptdiagonale sind. Insbesondere ist A diagonalisierbar, wenn a_{11}, \dots, a_{nn} paarweise verschieden sind.

³Analog definiert man *untere* Dreiecksmatrizen.

9 Determinanten

9.1 Rekursive Definition

Bemerkung 9.1.

- (a) Mathematiker versuchen oft komplizierte Objekte (wie $f \in \text{End}(V)$) durch einfachere (wie $\text{rk}(f)$ oder $\text{tr}(f)$) zu ersetzen, um wesentliche Informationen sichtbar zu machen. So haben wir in Lemma 5.15 gesehen, dass $\text{rk}(f)$ Rückschluss über die Bijektivität von f liefert, sofern man $\dim V$ kennt. Man nennt solche Größen *Invarianten*, wenn sie unter „natürlichen“ Umformungen (wie Basiswechsel) unverändert bleiben. Wir definieren in diesem Abschnitt als weitere Invariante die *Determinante* $\det(f)$. Wir zeigen, dass f genau dann bijektiv ist, wenn $\det(f) \neq 0$ gilt (im Gegensatz zu $\text{rk}(f)$ hängt dieses Kriterium nicht mehr von $\dim(V)$ ab.¹)
- (b) In der Maßtheorie versucht man möglichst vielen Mengen $S \subseteq \mathbb{R}^n$ ein „Volumen“ $\text{vol}(S) \in \mathbb{R}_{\geq 0}$ zuzuordnen. Sei $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ linear. Die in (a) beschriebene Zahl $\det(f)$ misst wie sehr sich das Volumen durch Anwenden von f verändert, d. h. es gilt $\text{vol}(f(S)) = |\det(f)| \text{vol}(S)$ sofern $\text{vol}(S)$ definiert ist. Dem n -dimensionalen *Hyperwürfel*

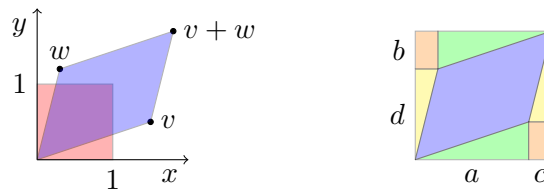
$$H := \{(x_1, \dots, x_n) \in \mathbb{R}^n : \forall i : 0 \leq x_i \leq 1\}$$

wird das Volumen $\text{vol}(H) = 1$ zugewiesen. Daraus folgt

$$|\det(f)| = \text{vol}(f(H)).$$

Das Vorzeichen von $\det(f)$ beschreibt, ob f orientierungserhaltend (Beispiel: Drehung) oder orientierungsumkehrend (Beispiel: Spiegelung) ist. Mehr dazu in Beispiel 11.22.

Beispiel 9.2. Sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ mit $v := f(e_1) = (a, b)$ und $w := f(e_2) = (c, d)$, d. h. $[f] = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$. Das Bild des (roten) Quadrats $H = \{(x, y) \in \mathbb{R}^2 : 0 \leq x, y \leq 1\}$ ist das von den Vektoren v und w aufgespannte (blaue) Parallelogramm $f(H)$:



Die Fläche von $f(H)$ ist

$$\text{vol}(f(H)) = (a + c)(b + d) - 2bc - ab - cd = ad - bc.$$

Die Fläche ist genau dann 0, wenn v und w auf einer Geraden liegen, also linear abhängig sind. Dies ist äquivalent zu $\text{rk}(f) \leq 1$.

¹Aus dem realen Leben: Wenn Sie Probleme haben sich das Alter einer Person zu merken, merken Sie sich stattdessen das Geburtsjahr, denn diese Invariante ändert sich nicht jedes Jahr.

Definition 9.3. Sei $A = (a_{ij}) \in K^{n \times n}$ und $1 \leq s, t \leq n$. Durch Streichen der s -ten Zeile und t -ten Spalte von A entsteht die Matrix $A_{st} \in K^{(n-1) \times (n-1)}$. Die *Determinante*² von A ist rekursiv definiert:

$$\det(A) := \begin{cases} a_{11} & \text{falls } n = 1, \\ \sum_{i=1}^n (-1)^{i+1} a_{i1} \det(A_{i1}) & \text{falls } n \geq 2. \end{cases}$$

Beispiel 9.4.

(a) Für $n = 2$ erhält man

$$\det \begin{pmatrix} a & c \\ b & d \end{pmatrix} = a \det(A_{11}) - b \det(A_{21}) = ad - bc$$

(vgl. Beispiel 9.2).

(b) Für jede obere Dreiecksmatrix $A = (a_{ij})$ gilt $\det(A) = a_{11} \dots a_{nn}$. Dies ist klar für $n = 1$. Sei induktiv die Behauptung für $n - 1$ bereits bewiesen. Da A_{11} auch eine obere Dreiecksmatrix ist, folgt

$$\det(A) = \sum_{i=1}^n (-1)^{i+1} a_{i1} \det(A_{i1}) = a_{11} \det(A_{11}) = a_{11} a_{22} \dots a_{nn}.$$

Da man mit dem Gauß-Algorithmus jede Matrix in eine obere Dreiecksmatrix überführen kann, untersuchen wir wie sich die Determinante bei elementaren Zeilenoperationen verändert.

Lemma 9.5. Die Abbildung $\det: K^{n \times n} \rightarrow K$ ist linear in jeder Zeile, d. h. für $a_1, \dots, a_n, b \in K^n$, $\lambda \in K$ und $1 \leq k \leq n$ gilt

$$\det \begin{pmatrix} a_1 \\ \vdots \\ \lambda a_k + b \\ \vdots \\ a_n \end{pmatrix} = \lambda \det \begin{pmatrix} a_1 \\ \vdots \\ a_k \\ \vdots \\ a_n \end{pmatrix} + \det \begin{pmatrix} a_1 \\ \vdots \\ b \\ \vdots \\ a_n \end{pmatrix}.$$

Beweis. Sei $a_i = (a_{i1}, \dots, a_{in})$ und $b = (b_1, \dots, b_n)$. Für $c \in K^n$ sei $M(c)$ die Matrix mit Zeilen $a_1, \dots, a_{k-1}, c, a_{k+1}, \dots, a_n$. Für $n = 1$ ist $k = 1$ und

$$\det(M(\lambda a_1 + b)) = \lambda a_{11} + b_1 = \lambda \det(M(a_1)) + \det(M(b))$$

wie behauptet. Sei nun $n \geq 2$ und die Behauptung für $n - 1$ bereits bewiesen. Streichen der k -ten Zeile ergibt

$$M(\lambda a_k + b)_{k1} = M(a_k)_{k1} = M(b)_{k1}.$$

Es folgt

$$\begin{aligned} \det(M(\lambda a_k + b)) &= (-1)^{k+1} (\lambda a_{k1} + b_1) \det(M(\lambda a_k + b)_{k1}) + \sum_{i \neq k} (-1)^{i+1} a_{i1} \det(M(\lambda a_k + b)_{i1}) \\ &= \lambda (-1)^{k+1} a_{k1} \det(M(a_k)_{k1}) + (-1)^{k+1} b_1 \det(M(b)_{k1}) \end{aligned}$$

²In manchen Büchern schreibt man $|A|$ anstelle von $\det(A)$. Das kann aber mit einer Matrixnorm verwechselt werden (Beispiel 17.61).

$$\begin{aligned}
& + \sum_{i \neq k} (-1)^{i+1} a_{i1} (\lambda \det(M(a_k)_{i1}) + \det(M(b)_{i1})) \\
& = \lambda \sum_{i=1}^n (-1)^{i+1} a_{i1} \det(M(a_k)_{i1}) + (-1)^{k+1} b_1 \det(M(b)_{k1}) + \sum_{i \neq k} (-1)^{i+1} a_{i1} \det(M(b)_{i1}) \\
& = \lambda \det(M(a_k)) + \det(M(b)). \quad \square
\end{aligned}$$

Satz 9.6. Für $A \in K^{n \times n}$ gilt:

- (a) Durch Multiplikation einer Zeile von A mit $\lambda \in K$ wird auch $\det(A)$ mit λ multipliziert.
- (b) Vertauschen von zwei Zeilen von A ändert das Vorzeichen von $\det(A)$.
- (c) Addieren eines Vielfachen einer Zeile zu einer anderen Zeile von A ändert $\det(A)$ nicht.

Beweis.

- (a) Setzt man zunächst $\lambda = 1$ und $b = 0$ in Lemma 9.5, so sieht man, dass die Determinante verschwindet, wenn A eine Nullzeile besitzt. Die Behauptung folgt nun, indem man λ beliebig und $b = 0$ in Lemma 9.5 wählt.
- (b) Hier ist $n \geq 2$. Seien a_1, \dots, a_n die Zeilen von A und $s < t$. Vertauschen von a_s und a_t liefert die Matrix A' . Für $n = 2$ ist $(s, t) = (1, 2)$ und

$$\det(A') = \det \begin{pmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{pmatrix} = a_{21}a_{12} - a_{22}a_{11} = -(a_{11}a_{22} - a_{12}a_{21}) = -\det(A)$$

nach Beispiel 9.4. Sei nun die Behauptung für $n - 1$ bereits bewiesen. Für $s \neq i \neq t$ entsteht A'_{i1} durch Zeilentausch aus A_{i1} . Also ist $\det(A'_{i1}) = -\det(A_{i1})$. Andererseits entsteht A'_{s1} aus A_{t1} durch die Vertauschungen $a_s \leftrightarrow a_{s+1} \leftrightarrow a_{s+2} \leftrightarrow \dots \leftrightarrow a_{t-1}$:

$$A_{t1} = \begin{pmatrix} \vdots \\ a_s \\ a_{s+1} \\ \vdots \\ a_{t-1} \\ a_{t+1} \\ \vdots \end{pmatrix} \begin{matrix} \leftarrow \\ \leftarrow \end{matrix} \sim \begin{pmatrix} \vdots \\ a_{s+1} \\ a_s \\ a_{s+2} \\ \vdots \\ a_{t-1} \\ a_{t+1} \\ \vdots \end{pmatrix} \begin{matrix} \leftarrow \\ \leftarrow \end{matrix} \sim \dots \sim \begin{pmatrix} \vdots \\ a_{s+1} \\ \vdots \\ a_{t-1} \\ a_s \\ a_{t+1} \\ \vdots \end{pmatrix} = A'_{s1}.$$

Dies zeigt $\det(A'_{s1}) = (-1)^{t-s-1} \det(A_{t1})$. Analog ist $\det(A'_{t1}) = (-1)^{t-s-1} \det(A_{s1})$. Wegen $(-1)^{t-s-1} = (-1)^{s-t-1}$ ergibt sich

$$\begin{aligned}
\det(A') & = (-1)^{s+1} a_{t1} \det(A'_{s1}) + (-1)^{t+1} a_{s1} \det(A'_{t1}) + \sum_{i \notin \{s,t\}} (-1)^{i+1} a_{i1} \det(A'_{i1}) \\
& = (-1)^t a_{t1} \det(A_{t1}) + (-1)^s a_{s1} \det(A_{s1}) + \sum_{i \notin \{s,t\}} (-1)^i a_{i1} \det(A_{i1}) \\
& = - \sum_{i=1}^n (-1)^{i+1} a_{i1} \det(A_{i1}) = -\det(A).
\end{aligned}$$

(c) Wir addieren λa_k zu Zeile a_l mit $k \neq l$ und erhalten

$$\det \begin{pmatrix} a_1 \\ \vdots \\ a_l + \lambda a_k \\ \vdots \\ a_n \end{pmatrix} \stackrel{9.5}{=} \det(A) + \lambda \det \begin{pmatrix} \vdots \\ a_k \\ \vdots \\ a_k \\ \vdots \end{pmatrix}.$$

Es genügt also $\det(A) = 0$ zu zeigen, falls A zwei identische Zeilen besitzt.³ Für $n = 2$ gilt

$$\det(A) = \det \begin{pmatrix} a & b \\ a & b \end{pmatrix} = ab - ba = 0.$$

Sei nun $n \geq 3$. Nach (b) können wir annehmen, dass die ersten beiden Zeilen von A identisch sind. Dann hat A_{i1} für $i \geq 3$ ebenfalls zwei identische Zeilen. Mit Induktion nach n folgt

$$\det(A) = \sum_{i=1}^n (-1)^{i+1} a_{i1} \det(A_{i1}) = a_{11} \det(A_{11}) - a_{21} \det(A_{21}) = 0. \quad \square$$

Beispiel 9.7.

$$\begin{aligned} \det \begin{pmatrix} -4 & -2 & -2 \\ 6 & 3 & 2 \\ 8 & 7 & 6 \end{pmatrix} & \quad | :(-2) \\ & = -2 \det \begin{pmatrix} 2 & 1 & 1 \\ 6 & 3 & 2 \\ 8 & 7 & 6 \end{pmatrix} \begin{array}{l} \left[\begin{array}{l} \leftarrow -3 \\ \leftarrow + \end{array} \right]^{-4} \\ \leftarrow + \end{array} \\ & = -2 \det \begin{pmatrix} 2 & 1 & 1 \\ 0 & 0 & -1 \\ 0 & 3 & 2 \end{pmatrix} \begin{array}{l} \leftarrow \\ \leftarrow \end{array} \\ & = 2 \det \begin{pmatrix} 2 & 1 & 1 \\ 0 & 3 & 2 \\ 0 & 0 & -1 \end{pmatrix} \stackrel{9.4}{=} -12 \end{aligned}$$

Bemerkung 9.8. Für $A \in K^{n \times n}$ und $\lambda \in K$ gilt $\boxed{\det(\lambda A) = \lambda^n \det(A)}$, denn jede der n Zeilen wird mit λ multipliziert.

Lemma 9.9. Sei $A = \begin{pmatrix} A_1 & A_2 \\ 0 & A_3 \end{pmatrix} \in K^{(n+m) \times (n+m)}$ mit $A_1 \in K^{n \times n}$, $A_2 \in K^{n \times m}$ und $A_3 \in K^{m \times m}$. Dann gilt $\det(A) = \det(A_1) \det(A_3)$.

Beweis. Führt man den Gauß-Algorithmus an A durch, so wird zuerst A_1 und dann A_2 in eine obere Dreiecksmatrix umgeformt. Am Ende ist auch A eine obere Dreiecksmatrix und die Behauptung folgt. \square

9.2 Eigenschaften

Satz 9.10. Für $A \in K^{n \times n}$ gilt

$$A \text{ invertierbar} \iff \text{rk}(A) = n \iff \det(A) \neq 0.$$

³Für $K = \mathbb{Q}$ folgt dies sofort aus (b), aber nicht für $K = \mathbb{F}_2$.

Beweis. Die erste Äquivalenz stammt aus Lemma 5.15. Da die Zeilenstufenform \widehat{A} eine obere Dreiecksmatrix ist, gilt

$$\det(A) \neq 0 \stackrel{9.6}{\iff} \det(\widehat{A}) \neq 0 \stackrel{9.4}{\iff} \widehat{A} = 1_n \iff \operatorname{rk}(A) = \operatorname{rk}(\widehat{A}) = n. \quad \square$$

Satz 9.11 (Determinantensatz). Für $A, B \in K^{n \times n}$ gilt $\boxed{\det(AB) = \det(A) \det(B)}$.

Beweis. Ist $\det(A) = 0$, so folgt $\operatorname{rk}(AB) \leq \operatorname{rk}(A) < n$ aus Lemma 5.15. Dann ist auch $\det(AB) = 0$ nach Satz 9.10. Wir können also $A \in \operatorname{GL}(n, K)$ annehmen. Nach Folgerung 6.18 ist A ein Produkt von Elementarmatrizen, sagen wir $A = A_1 \dots A_k$. Sei $M \in K^{n \times n}$ beliebig. Für die drei Arten von elementaren Zeilenoperationen gilt jeweils

$$\det(A_i M) = \begin{cases} \lambda \det(M) \\ -\det(M) \\ \det(M) \end{cases} = \det(A_i 1_n) \det(M) = \det(A_i) \det(M)$$

nach Satz 9.6. Insgesamt folgt

$$\begin{aligned} \det(AB) &= \det(A_1 \dots A_k B) = \det(A_1) \det(A_2 \dots A_k B) = \dots = \det(A_1) \dots \det(A_k) \det(B) \\ &= \dots = \det(A_1) \det(A_2 \dots A_k) \det(B) = \det(A) \det(B). \end{aligned} \quad \square$$

Folgerung 9.12.

(a) Für $A \in K^{n \times n}$ gilt $\boxed{\det(A^t) = \det(A)}$.

(b) Für $A \in \operatorname{GL}(n, K)$ gilt $\boxed{\det(A^{-1}) = \det(A)^{-1}}$.

(c) Ähnliche Matrizen haben die gleiche Determinante.

Beweis.

(a) Wegen $\operatorname{rk}(A) = \operatorname{rk}(A^t)$ können wir annehmen, dass A invertierbar ist (anderenfalls ist $\det(A) = 0 = \det(A^t)$). Wieder ist A ein Produkt von Elementarmatrizen $A = A_1 \dots A_k$. Für die ersten beiden Zeilenoperationen gilt $A_i^t = A_i$. Für die dritte Zeilenoperation ist $\det(A_i) = 1 = \det(A_i^t)$. Dies zeigt

$$\begin{aligned} \det(A^t) &\stackrel{5.8}{=} \det(A_k^t \dots A_1^t) = \det(A_k^t) \dots \det(A_1^t) = \det(A_k) \dots \det(A_1) \\ &= \det(A_1) \dots \det(A_k) = \det(A_1 \dots A_k) = \det(A). \end{aligned}$$

(b) Die Behauptung folgt aus $\det(A) \det(A^{-1}) = \det(AA^{-1}) = \det(1_n) = 1$.

(c) Für $A \in K^{n \times n}$ und $S \in \operatorname{GL}(n, K)$ gilt

$$\det(SAS^{-1}) = \det(S) \det(A) \det(S^{-1}) = \det(S) \det(S)^{-1} \det(A) = \det(A). \quad \square$$

Definition 9.13. Nach dem Determinantensatz und Folgerung 9.12 bilden die Matrizen mit Determinante 1 eine Untergruppe $\operatorname{SL}(n, K) \leq \operatorname{GL}(n, K)$. Man nennt $\operatorname{SL}(n, K)$ die *spezielle lineare Gruppe* vom Grad n über K . Es gilt $\operatorname{SL}(n, \mathbb{F}_2) = \operatorname{GL}(n, \mathbb{F}_2)$.

Bemerkung 9.14.

- (a) Wegen $\det(A^t) = \det(A)$ darf man bei der Berechnung von $\det(A)$ auch elementare Spaltenoperationen benutzen.
- (b) Die folgende Aussage verallgemeinert den Determinantensatz.

Satz 9.15 (CAUCHY-BINET-Formel). Für $A, B \in K^{n \times m}$ und $I \subseteq \{1, \dots, m\}$ sei $A_I := (a_{ij} : i = 1, \dots, n, j \in I)$. Dann gilt

$$\det(AB^t) = \sum_{\substack{I \subseteq \{1, \dots, m\} \\ |I|=n}} \det(A_I) \det(B_I).$$

Beweis. Im Fall $n > m$ ist die Summe leer und $\det(AB^t) = 0$ nach Lemma 5.15. Sei also $n \leq m$. Wir zerlegen die i -te Zeile von A als $a_i = a'_i + a''_i$. Ersetzt man a_i durch a'_i bzw. a''_i , so erhält man Matrizen A' bzw. A'' mit $\det(A_I) = \det(A'_I) + \det(A''_I)$ für alle $I \subseteq \{1, \dots, m\}$ mit $|I| = n$ nach Lemma 9.5. Die i -te Zeile von AB^t ist $a'_i B^t + a''_i B^t$. Daher gilt auch $\det(AB^t) = \det(A' B^t) + \det(A'' B^t)$. Es genügt also die Behauptung für A' bzw. A'' anstelle von A zu beweisen. Auf diese Weise erreicht man, dass A in jeder Zeile höchstens einen von 0 verschiedenen Eintrag besitzt. Also existiert höchstens ein $I \subseteq \{1, \dots, m\}$ mit $|I| = n$ und $\det(A_I) \neq 0$ (alle anderen A_I besitzen Nullspalten). Es gilt $AB^t = A_I(B_I)^t$ und $\det(AB^t) = \det(A_I) \det(B_I)$ nach Folgerung 9.12. □

9.3 Laplace-Entwicklung

Satz 9.16 (LAPLACE-Entwicklung). Sei $n \geq 2$ und $A \in K^{n \times n}$. Für $1 \leq k \leq n$ gilt

$$\det(A) = \sum_{i=1}^n (-1)^{i+k} a_{ik} \det(A_{ik}) = \sum_{i=1}^n (-1)^{i+k} a_{ki} \det(A_{ki}).$$

Beweis. Seien a_1, \dots, a_n die Spalten von A . Nach Bemerkung 9.14 gilt

$$\begin{aligned} \det \left(\begin{array}{cccc} \cdots & a_{k-1} & a_k & \cdots \end{array} \right) &= - \det \left(\begin{array}{cccc} \cdots & a_{k-2} & a_k & a_{k-1} & \cdots \end{array} \right) = \dots \\ &= (-1)^{k-1} \det \left(\begin{array}{cccc} a_k & a_1 & \cdots & a_{k-1} & a_{k+1} & \cdots \end{array} \right) = (-1)^{k-1} \sum_{i=1}^n (-1)^{i+1} a_{ik} \det(A_{ik}). \end{aligned}$$

Die zweite Gleichung folgt aus der ersten, indem man $\det(A^t) = \det(A)$ benutzt. □

Bemerkung 9.17. Die Gleichungen in Satz 9.16 nennt man *Entwicklung nach der k -ten Spalte/Zeile*. Die Vorzeichen $(-1)^{i+k}$ verteilen sich schachbrettartig:

$$\begin{pmatrix} + & - & + & \cdots \\ - & + & - & \cdots \\ + & - & + & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Beispiel 9.18. Wie die meisten rekursiven Verfahren ist auch die Laplace-Entwicklung in der Regel ineffizient. Sie eignet sich jedoch für sogenannte *dünnbesetzte* Matrizen, d. h. wenn viele Einträge 0 sind. Wir entwickeln zuerst nach der dritten Zeile und anschließend nach der zweiten Spalte:

$$\det \begin{pmatrix} 1 & 2 & -3 & 0 \\ -2 & 0 & 1 & 2 \\ 0 & 0 & 1 & 0 \\ -2 & 0 & 1 & 0 \end{pmatrix} = \det \begin{pmatrix} 1 & 2 & 0 \\ -2 & 0 & 2 \\ -2 & 0 & 0 \end{pmatrix} = -2 \det \begin{pmatrix} -2 & 2 \\ -2 & 0 \end{pmatrix} = -2 \cdot 4 = -8$$

Bemerkung 9.19. Für eine Folge von Zahlen $a_1, \dots, a_n \in K$ definiert das Produkt $\prod_{i=1}^n a_i = a_1 \cdot \dots \cdot a_n$ analog zum Summenzeichen \sum .

Satz 9.20 (VANDERMONDE). Für $x_1, \dots, x_n \in K$ nennt man

$$A := (x_i^{j-1}) = \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{pmatrix} \in K^{n \times n}$$

VANDERMONDE-Matrix.⁴ Es gilt

$$\det(A) = \prod_{1 \leq i < j \leq n} (x_j - x_i).$$

Insbesondere ist A genau dann invertierbar, wenn x_1, \dots, x_n paarweise verschieden sind.

Beweis. Induktion nach n : Im Fall $n = 1$ ist $A = 1_1$ und $\prod_{i < j} (x_j - x_i)$ ist das leere Produkt, welches man als 1 interpretiert (so wie die leere Summe als 0 interpretiert wird). Sei also $n \geq 2$. Wir subtrahieren das x_1 -Fache der vorletzten Spalte von der letzten Spalte. Anschließend subtrahieren wir das x_1 -Fache der $(n - 2)$ -ten Spalte von der vorletzten Spalte usw. Dadurch erhält man die Matrix

$$\begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 1 & x_2 - x_1 & (x_2 - x_1)x_2 & \cdots & (x_2 - x_1)x_2^{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n - x_1 & (x_n - x_1)x_n & \cdots & (x_n - x_1)x_n^{n-2} \end{pmatrix}$$

mit derselben Determinante (Satz 9.6). Durch Entwicklung nach der ersten Zeile kann man zur kleineren Matrix

$$\begin{pmatrix} x_2 - x_1 & (x_2 - x_1)x_2 & \cdots & (x_2 - x_1)x_2^{n-2} \\ \vdots & \vdots & & \vdots \\ x_n - x_1 & (x_n - x_1)x_n & \cdots & (x_n - x_1)x_n^{n-2} \end{pmatrix}$$

übergehen. Die Faktoren $(x_k - x_1)$ können für $k = 2, \dots, n$ aus der Determinante herausgezogen werden. Dies ergibt

$$\det(A) = (x_2 - x_1)(x_3 - x_1) \cdots (x_n - x_1) \det((x_{i+1}^{j-1})_{i,j=1}^{n-1}).$$

Nun folgt die Behauptung mit Induktion. □

⁴In manchen Büchern betrachtet man die transponierte Matrix.

Definition 9.21. Für $A \in K^{n \times n}$ nennt man

$$\tilde{A} := \begin{cases} 1_1 & \text{falls } n = 1 \\ ((-1)^{i+j} \det(A_{ji}))_{i,j} & \text{falls } n > 1 \end{cases} \in K^{n \times n}$$

die zu A komplementäre Matrix.⁵

Satz 9.22. Für alle $A \in K^{n \times n}$ gilt $A\tilde{A} = \det(A)1_n = \tilde{A}A$. Insbesondere ist $A^{-1} = \frac{1}{\det(A)}\tilde{A}$, falls $A \in \text{GL}(n, K)$.

Beweis. Für $n = 1$ ist die Behauptung klar. Sei also $n \geq 2$. Sei B_{kl} die Matrix, die aus A entsteht, indem man die l -te Zeile durch die k -te Zeile ersetzt. Für $k \neq l$ hat B_{kl} zwei identische Zeilen und es folgt $\det(B_{kl}) = 0$. Andererseits ist $B_{kk} = A$. Sei $A\tilde{A} = (c_{ij})$. Entwicklung nach der l -ten Zeile von B_{kl} ergibt

$$\delta_{kl} \det(A) = \det(B_{kl}) = \sum_{i=1}^n a_{ki} (-1)^{i+l} \det(A_{li}) = c_{kl}.$$

Dies zeigt $A\tilde{A} = \det(A)1_n$. Die Gleichung $\tilde{A}A = \det(A)1_n$ zeigt man analog durch Entwicklung nach einer Spalte. \square

Beispiel 9.23. Für jede invertierbare 2×2 -Matrix $A := \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ erhält man

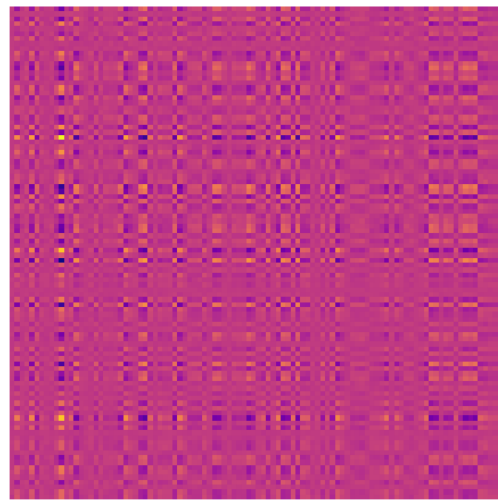
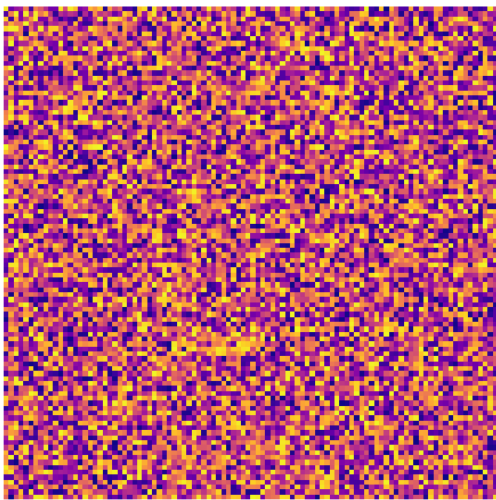
$$A^{-1} = \frac{1}{\det(A)}\tilde{A} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Bemerkung 9.24.

- (a) Die Formel $A^{-1} = \frac{1}{\det(A)}\tilde{A}$ zeigt, dass der Eintrag von A^{-1} an Position (i, j) von allen a_{st} , außer denen mit entweder $s = i$ oder $t = j$ abhängt. Daher weist die Inverse einer zufällig gewählten 100×100 -Matrix deutliche Strukturen auf.⁶

A

A^{-1}



⁵auch *Adjunkte* genannt; Verwechslungsgefahr mit der *adjungierten* Matrix A^* aus Satz 13.7

⁶Die Grafiken wurden mit sageMath generiert.

- (b) Für „große“ n hat diese Formel mehr theoretische als praktische Bedeutung (nutzen Sie Satz 6.17 zur Berechnung von A^{-1}). Ist beispielsweise $A \in \mathbb{Z}^{n \times n}$, so ist auch $\det(A)A^{-1} = \tilde{A} \in \mathbb{Z}^{n \times n}$. Insbesondere ist $A^{-1} \in \mathbb{Z}^{n \times n}$, falls $\det(A) = \pm 1$. Aus dem Gauß-Algorithmus ist diese Beobachtung nicht ersichtlich. Der nächste Satz liefert eine ähnliche Aussage für Gleichungssysteme.

Satz 9.25 (CRAMERSche Regel). Sei $A \in \text{GL}(n, K)$ und $b \in K^{n \times 1}$. Für $k = 1, \dots, n$ sei A_k die Matrix, die aus A entsteht, indem man die k -te Spalte durch b ersetzt. Für die eindeutige Lösung $x = (x_1, \dots, x_n)^t$ des Gleichungssystems $Ax = b$ gilt dann $x_k = \frac{\det(A_k)}{\det(A)}$ für $k = 1, \dots, n$.

Beweis. Wir benutzen Satz 9.22 und entwickeln A_k nach der k -ten Spalte:

$$(\det(A_k))_k = \left(\sum_{i=1}^n (-1)^{i+k} \det(A_{ik}) b_i \right)_k = \tilde{A}b = \tilde{A}Ax = \det(A)x = (\det(A)x_k)_k. \quad \square$$

9.4 Die Leibniz-Formel

Bemerkung 9.26. Führt man die Laplace-Entwicklung für $n \times n$ -Matrizen bis auf 1×1 -Matrizen zurück, so erhält man die Determinante als Summe von $n(n-1) \cdot \dots \cdot 2 \cdot 1 = n!$ Termen. Wir bestimmen diese Terme explizit.

Definition 9.27. Sei $n \in \mathbb{N}$ und $N := \{1, \dots, n\}$. Eine Bijektion der Form $N \rightarrow N$ heißt *Permutation* von N . Die Menge der Permutationen von N wird mit S_n bezeichnet.

Bemerkung 9.28.

- (a) Analog zu $\text{GL}(V)$ ist auch S_n eine Gruppe bzgl. Komposition von Abbildungen. Wir werden daher das Kompositionszeichen \circ oft einsparen. Man nennt S_n die *symmetrische Gruppe von Grad n* .
- (b) Sei $\sigma \in S_n$. Für die Wahl von $\sigma(1) \in \{1, \dots, n\}$ gibt es n Möglichkeiten. Da σ injektiv ist, gilt $\sigma(2) \neq \sigma(1)$. Für die Wahl von $\sigma(2)$ verbleiben also noch $n-1$ Möglichkeiten usw. Insgesamt hat man $n!$ Möglichkeiten eine Permutation zu definieren, d. h. $|S_n| = n!$.

Beispiel 9.29.

- (a) Für $k \geq 2$ nennt man $\sigma \in S_n$ einen (k) -Zyklus (oder Zyklus der Länge k), falls paarweise verschiedene $1 \leq a_1, \dots, a_k \leq n$ existieren, sodass

$$\sigma(x) = \begin{cases} a_{i+1} & \text{falls } x = a_i \text{ mit } i < k, \\ a_1 & \text{falls } x = a_k, \\ x & \text{sonst.} \end{cases} \quad \begin{array}{ccc} & a_1 & \\ \sigma \curvearrowright & & \curvearrowleft \sigma \\ a_2 & & a_4 \\ \sigma \curvearrowright & & \curvearrowleft \sigma \\ & a_3 & \end{array}$$

Man schreibt dann $\sigma = (a_1, \dots, a_k)$. Diese Schreibweise ist eindeutig bis auf „Rotation“, d. h.

$$\sigma = (a_2, \dots, a_k, a_1) = \dots = (a_k, a_1, \dots, a_{k-1}).$$

Die Komposition von Zyklen geschieht wie bei Abbildungen üblich von rechts nach links:

$$(1, 3, 4, 5) \circ (3, 4, 5) \circ (1, 3, 2) = (1, 5, 4)(2, 3).$$

Außerdem ist $(a_1, \dots, a_k)^{-1} = (a_k, a_{k-1}, \dots, a_1)$.

- (b) Zyklen der Länge 2 heißen *Transpositionen*. Eine Transposition vertauscht also zwei Elemente und lässt alle anderen Elemente fest. Jeder k -Zyklus ist eine Komposition von $k - 1$ Transpositionen:

$$(a_1, \dots, a_k) = (a_1, a_2)(a_2, a_3) \dots (a_{k-1}, a_k).$$

Allerdings gibt es viele Möglichkeiten für eine solche Komposition.

- (c) In S_3 ist jedes Element ein Zyklus: $S_3 = \{1, (1, 2), (1, 3), (2, 3), (1, 2, 3), (1, 3, 2)\}$.
 (d) Zyklen $\sigma = (a_1, \dots, a_k)$ und $\tau = (b_1, \dots, b_l)$ heißen *disjunkt*, falls

$$\{a_1, \dots, a_k\} \cap \{b_1, \dots, b_l\} = \emptyset.$$

Gegebenenfalls gilt $\sigma\tau = \tau\sigma$.

Satz 9.30. *Jede Permutation $\sigma \in S_n$ ist eine Komposition von paarweise disjunkten Zyklen $\sigma = \sigma_1 \dots \sigma_k$. Dabei sind $\sigma_1, \dots, \sigma_k$ bis auf die Reihenfolge eindeutig bestimmt.*

Beweis. Sei $\sigma \in S_n$. Im Fall $\sigma = \text{id}$ ist σ das leere Produkt von Zyklen. Sei also $\sigma \neq \text{id}$ und

$$a_1 := \min\{1 \leq i \leq n : \sigma(i) \neq i\}.$$

Sei $a_i := \sigma^{i-1}(a_1)$ für $i \geq 2$. Wegen $n < \infty$ existieren $i < j$ mit $a_i = a_j$, d. h. $\sigma^{j-i}(a_1) = a_1$. Daher existiert

$$k := \min\{1 \leq i \leq n : \sigma^i(a_1) = a_1\}.$$

Wegen $k \leq j - i$ sind die Elemente a_1, \dots, a_k paarweise verschieden. Also ist $\sigma_1 := (a_1, \dots, a_k)$ ein k -Zyklus. Für $\rho := \sigma\sigma_1^{-1}$ gilt

$$\rho(x) = \begin{cases} x & \text{falls } x \in \{a_1, \dots, a_k\}, \\ \sigma(x) & \text{sonst.} \end{cases}$$

Im Fall $\rho = \text{id}$ ist $\sigma = \sigma_1$ und wir sind fertig. Anderenfalls können wir das Verfahren mit ρ anstatt σ wiederholen. Da mit jeder Wiederholung die Anzahl der Fixpunkte von σ wächst, erreicht man nach endlich vielen Schritten $\sigma = \sigma_1 \dots \sigma_k$ mit paarweise disjunkten Zyklen $\sigma_1, \dots, \sigma_k$.

Sei auch $\sigma = \tau_1 \dots \tau_l$ eine Komposition von paarweise disjunkten Zyklen τ_1, \dots, τ_l . Dann existiert ein i mit $\tau_i(a_1) \neq a_1$. Da die Zyklen disjunkt sind, gilt $\tau_i(a_1) = \sigma(a_1) = \sigma_1(a_1) = a_2$, $\tau_i(a_2) = \sigma_1(a_2) = a_3$ usw. Also ist $\tau_i = \sigma_1$ und $\sigma_2 \dots \sigma_k = \tau_1 \dots \tau_{i-1} \tau_{i+1} \dots \tau_l$. Die Eindeutigkeit der σ_i folgt nun mit Induktion nach k . \square

Bemerkung 9.31.

- (a) Da nach Beispiel 9.29 jeder Zyklus eine Komposition von Transpositionen ist, ist sogar jede Permutation eine Komposition von (in Regel nicht disjunkten) Transpositionen.
 (b) Man kann die Schreibweise in disjunkte Zyklen

$$\sigma = (a_1, \dots, a_s)(b_1, \dots, b_t) \dots$$

vollständig eindeutig machen, indem man $a_1 = \min\{a_1, \dots, a_s\} < b_1 = \min\{b_1, \dots, b_t\} < \dots$ fordert.

Definition 9.32. Für $\sigma \in S_n$ nennt man

$$P_\sigma := (\delta_{i\sigma(j)})_{i,j} \in \mathbb{Q}^{n \times n}$$

die *Permutationsmatrix* von σ . Außerdem heißt $\text{sgn}(\sigma) := \det(P_\sigma)$ das *Signum* oder *Vorzeichen* von σ .

Bemerkung 9.33. Die Permutationsmatrix von σ entsteht, indem man die Zeilen der Einheitsmatrix (also die Standardbasis e_1, \dots, e_n) gemäß σ permutiert. Mit dem Gauß-Algorithmus lässt sich diese Permutation durch endlich viele Zeilenvertauschungen realisieren (das entspricht dem Sortieralgorithmus *Selectionsort*). Wegen $\det(1_n) = 1$ ist $\text{sgn}(\sigma) \in \{\pm 1\}$ also tatsächlich ein „Vorzeichen“.

Beispiel 9.34. Wir bestimmen die Permutationsmatrix und das Signum für die Permutationen in S_3 :

σ	id	(1, 2)	(1, 3)	(2, 3)	(1, 2, 3)	(1, 3, 2)
P_σ	1_3	$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$
$\text{sgn}(\sigma)$	1	-1	-1	-1	1	1

Satz 9.35. Für $\sigma, \tau \in S_n$ gilt $P_{\sigma \circ \tau} = P_\sigma P_\tau$ und $\boxed{\text{sgn}(\sigma \circ \tau) = \text{sgn}(\sigma) \text{sgn}(\tau)}$.

Beweis. Der Eintrag von $P_\sigma P_\tau$ an Position (i, j) ist

$$\sum_{k=1}^n \delta_{i\sigma(k)} \delta_{k\tau(j)} = \delta_{i,\sigma(\tau(j))} = \delta_{i,(\sigma \circ \tau)(j)}.$$

Dies zeigt die erste Gleichung. Die zweite folgt aus dem Determinantensatz. □

Bemerkung 9.36.

- (a) Die Permutationsmatrix einer Transposition ist genau die Elementarmatrix zur Vertauschung von Zeilen. Insbesondere hat jede Transposition Signum -1 . Nach Satz 9.35 ist das Produkt einer geraden Anzahl an Transpositionen niemals ein Produkt einer ungeraden Anzahl an Transpositionen.
- (b) Nach Beispiel 9.29 hat jeder k -Zyklus Signum $(-1)^{k-1}$. Ist σ ein Produkt von paarweise disjunkten Zyklen mit Längen l_1, \dots, l_k , so gilt

$$\boxed{\text{sgn}(\sigma) = (-1)^{l_1 + \dots + l_k - k}}.$$

Zum Beispiel ist

$$\text{sgn}((1, 2, 5, 6)(3, 7)(4, 9, 8)) = (-1)^{4+2+3-3} = 1.$$

Zählt man 1-Zyklen mit, so vereinfacht sich die Formel zu $\text{sgn}(\sigma) = (-1)^{n-k}$.

- (c) Aus Satz 9.35 folgt, dass die Permutationen mit Signum 1 eine Untergruppe $A_n \leq S_n$ bilden. Man nennt A_n die *alternierende* Gruppe vom Grad n . In gewisser Weise verhält sich A_n zu S_n so wie $\text{SL}(n, K)$ zu $\text{GL}(n, K)$.

Satz 9.37 (LEIBNIZ-Formel). Für $A = (a_{ij}) \in K^{n \times n}$ gilt

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}.$$

Beweis. Die Zeilen a_1, \dots, a_n von A lassen sich als Linearkombination der Standardbasis ausdrücken: $a_i = \sum_{j=1}^n a_{ij} e_j$. Da \det in jeder Zeile linear ist (Lemma 9.5), gilt

$$\begin{aligned} \det(A) &= \sum_{i_1=1}^n a_{1i_1} \det \begin{pmatrix} e_{i_1} \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \sum_{i_1=1}^n a_{1i_1} \sum_{i_2=1}^n a_{2i_2} \det \begin{pmatrix} e_{i_1} \\ e_{i_2} \\ a_3 \\ \vdots \\ a_n \end{pmatrix} = \dots \\ &= \sum_{1 \leq i_1, \dots, i_n \leq n} a_{1i_1} \cdots a_{ni_n} \det \begin{pmatrix} e_{i_1} \\ \vdots \\ e_{i_n} \end{pmatrix}. \end{aligned}$$

Existieren $s \neq t$ mit $i_s = i_t$, so verschwindet die entsprechende Determinante. Man muss also nur über die Tupel (i_1, \dots, i_n) mit paarweise verschiedenen Einträgen summieren. Jedes solche Tupel beschreibt eine Permutation $\sigma \in S_n$ mit $\sigma(j) = i_j$ mit $j = 1, \dots, n$. Es folgt

$$\det(A) = \sum_{\sigma \in S_n} a_{1\sigma(1)} \cdots a_{n\sigma(n)} \det(P_\sigma) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}. \quad \square$$

Folgerung 9.38 (SARRUS-Regel). Für 3×3 -Matrizen gilt

$$\det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = aei + bfg + cdh - gec - hfa - idb.$$

Beweis. Man benutze die Leibniz-Formel mit Beispiel 9.34. □

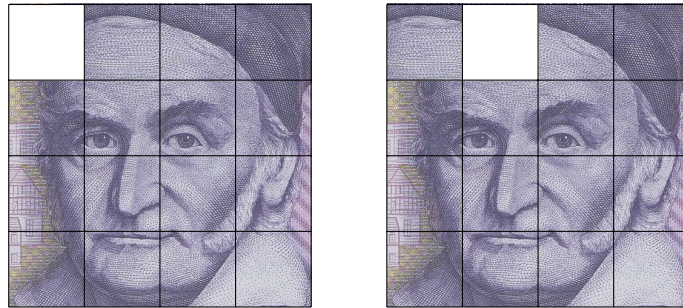
Bemerkung 9.39.

(a) Man kann sich die Sarrus-Regel mit folgendem Schema merken:

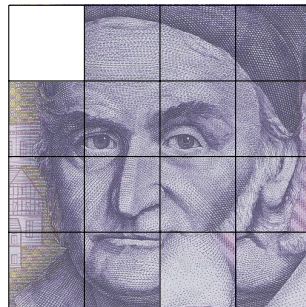
$$\begin{pmatrix} a & b & c & a & b \\ d & e & f & d & e \\ g & h & i & g & h \end{pmatrix}$$

- (b) Achtung: Die Sarrus-Regel gilt *nur* für 3×3 -Matrizen (für 4×4 -Matrizen braucht man $4! = 24$ Summanden).
- (c) Sei $A \in K^{n \times n}$ mit Eigenwert λ . Dann ist $E_\lambda(A) = \operatorname{Ker}(A - \lambda 1_n) \neq \{0\}$ und $\det(A - \lambda 1_n) = 0$. Nach der Leibniz-Formel ist $\det(A - \lambda 1_n)$ ein Polynom in λ . Auf diese Weise werden wir alle Eigenwerte von A berechnen.

Beispiel 9.40. Das folgende *Schiebepuzzle* besteht aus 15 beweglichen Quadraten und einem leeren Feld. Ein Quadrat, welches horizontal oder vertikal an das leere Feld grenzt, darf auf dieses geschoben werden (vgl. Cover):



Sam Loyd bot ein Preisgeld von 1000 \$, wem es gelingt die folgende Konfiguration in den Ausgangszustand zu überführen:⁷



Jeder Zug entspricht einer Transposition in S_{16} . Legt man ein Schachbrettmuster zugrunde, so wandert das leere Feld bei jedem Zug von schwarz nach weiß oder umgekehrt. Da das leere Feld in Loyds Konfiguration in der Ausgangsstellung liegt, benötigt man zur Lösung eine gerade Anzahl an Zügen. Andererseits unterscheidet sich Loyds Konfiguration nur um eine Transposition vom Ausgangszustand. Nach Bemerkung 9.36 ist diese Konfiguration also unlösbar und Loyd musste das Preisgeld nie auszahlen.

⁷siehe [D. Slocum und J. Sonneveld, *The 15 puzzle*, The Slocum Puzzle Foundation, Beverly Hills, 2006]

Aufgaben

Aufgabe I.1. Welche der folgenden Aussagen sind im Jahr 2025 wahr?

- (a) Es gibt einen Monat mit 28 Tagen.
- (b) Es gibt einen Monat mit genau 28 Tagen.
- (c) Es gibt genau einen Monat mit 28 Tagen.
- (d) Es gibt genau einen Monat mit genau 28 Tagen.

Aufgabe I.2. Seien $1 \leq a \leq b \leq 9$ natürliche Zahlen. Der Logiker (S)iegfried kennt nur die Summe $a + b$, während sein Kollege (P)etrus nur das Produkt ab kennt. Die beiden führen folgenden Dialog:

S: „Ich kenne a und b nicht.“ P: „Ich kenne a und b nicht.“
S: „Ich kenne a und b nicht.“ P: „Ich kenne a und b nicht.“
S: „Ich kenne a und b nicht.“ P: „Ich kenne a und b nicht.“
S: „Ich kenne a und b nicht.“ P: „Ich kenne a und b nicht.“
S: „Ich kenne a und b nicht.“ P: „Jetzt kenne ich a und b !“

Bestimmen Sie daraus a und b .

Aufgabe I.3. Für endliche Mengen A und B gilt $|A \cup B| = |A| + |B| - |A \cap B|$ nach Lemma 1.12. Finden und beweisen Sie eine analoge Gleichung für drei endliche Mengen.

Aufgabe I.4. Beweisen Sie mit vollständiger Induktion: Die Summe der ersten n ungeraden Zahlen ist n^2 .

Aufgabe I.5. Beweisen Sie, dass $\mathbb{N} \times \mathbb{N}$ abzählbar ist.

Aufgabe I.6. Konstruieren Sie Relationen mit folgenden Eigenschaften:

- (a) reflexiv, aber weder symmetrisch noch transitiv.
- (b) symmetrisch, aber weder reflexiv noch transitiv.
- (c) transitiv, aber weder reflexiv noch symmetrisch.

Aufgabe I.7. Sei $U := \{2z + 1 : z \in \mathbb{Z}\} \cup \{0\}$. Untersuchen Sie, ob $(U, +)$ eine Gruppe ist.

Aufgabe I.8. Seien $(G, *)$ und (H, \circ) Gruppen. Zeigen Sie, dass $G \times H$ mit der Verknüpfung

$$(g_1, h_1) \cdot (g_2, h_2) := (g_1 * g_2, h_1 \circ h_2)$$

zu einer Gruppe wird.

Aufgabe I.9. Konstruieren Sie einen Körper mit drei Elementen.

Hinweis: Was ist $1 + 1$?

Aufgabe I.10. Zeigen Sie:

- (a) Eine nichtleere Teilmenge H einer Gruppe G ist genau dann eine Untergruppe, wenn $x, y \in H \Rightarrow xy^{-1} \in H$ gilt.
- (b) Eine nichtleere Teilmenge U eines K -Vektorraums V genau dann ein Unterraum ist, wenn für alle $u, v \in U$ und $\lambda \in K$ gilt: $\lambda u + v \in U$.

Aufgabe I.11. (DEDEKIND-Identität) Seien X, Y, Z Unterräume eines Vektorraums V mit $X \subseteq Z$. Zeigen Sie: $(X + Y) \cap Z = X + (Y \cap Z)$.

Aufgabe I.12. Sei V ein Vektorraum, $S \subseteq V$ und $U, W \leq V$. Zeigen Sie:

- (a) $\langle S \rangle$ ist der Durchschnitt aller Unterräume von V , die S enthalten.
- (b) $U + W = \langle U \cup W \rangle$.
- (c) $U \cup W \leq V \iff U \cup W \in \{U, W\}$.

Aufgabe I.13 (Korrespondenzsatz). Seien $U \leq V$ Vektorräume.

- (a) Zeigen Sie, dass jeder Unterraum von V/U die Form W/U mit $U \leq W \leq V$ hat.
- (b) Zeigen Sie, dass die Abbildung $W \mapsto W/U$ eine Bijektion zwischen $\{W : U \leq W \leq V\}$ und der Menge der Unterräume von V/U ist.

Aufgabe I.14. Seien u, v, w Vektoren eines Vektorraums V . Beweisen oder widerlegen Sie: Genau dann ist $\{u, v, w\}$ linear unabhängig, wenn $\{u, v\}$, $\{u, w\}$ und $\{v, w\}$ linear unabhängig sind.

Aufgabe I.15. Offenbar ist \mathbb{R} ein \mathbb{Q} -Vektorraum, in dem die Skalarmultiplikation mit der üblichen Multiplikation in \mathbb{R} übereinstimmt (dies müssen Sie nicht prüfen). Zeigen Sie:

- (a) 1 und $\sqrt{2}$ sind linear unabhängig über \mathbb{Q} .
- (b) $\mathbb{Q}(\sqrt{2}) := \langle 1, \sqrt{2} \rangle = \mathbb{Q} + \mathbb{Q}\sqrt{2}$ ist ein Körper mit den gleichen Verknüpfungen wie in \mathbb{R} .

Aufgabe I.16. Zeigen Sie:

- (a) Das Vertauschen von zwei Zeilen einer Matrix lässt sich durch die beiden anderen elementaren Zeilenoperationen realisieren.
- (b) Jede $n \times n$ -Matrix ist ein Produkt von Matrizen der Form $1_n + \lambda E_{ij}$ mit $\lambda \in K$ und $1 \leq i, j \leq n$ (der Fall $i = j$ ist zugelassen).

Aufgabe I.17. Sei $n = n_1 + \dots + n_k$ und $\lambda_1, \dots, \lambda_k \in K$ paarweise verschieden. Sei

$$A := \text{diag}(\lambda_1 1_{n_1}, \dots, \lambda_k 1_{n_k}) \in K^{n \times n}$$

und $B \in K^{n \times n}$. Zeigen Sie:

- (a) Genau dann ist $AB = BA$, wenn $B = \text{diag}(B_1, \dots, B_k)$ mit $B_i \in K^{n_i \times n_i}$ für $i = 1, \dots, k$.

(b) Genau dann ist $AB = BA$ für alle $B \in K^{n \times n}$, wenn A eine Skalarmatrix ist (d. h. $k = 1$).

Aufgabe I.18. Seien $A \in \mathbb{Q}^{n \times m} \subseteq \mathbb{R}^{n \times m}$ und $b \in \mathbb{Q}^{n \times 1} \subseteq \mathbb{R}^{n \times 1}$. Begründen Sie:

- (a) Der Rang von A über \mathbb{Q} ist der Rang von A über \mathbb{R} .
- (b) Ist A über \mathbb{R} invertierbar, so auch über \mathbb{Q} .
- (c) Besitzt das Gleichungssystem $Ax = b$ eine Lösung in $\mathbb{R}^{m \times 1}$, so existiert auch eine Lösung in $\mathbb{Q}^{m \times 1}$.
- (d) Geben Sie ein Beispiel, in dem die Lösungsmengen von $Ax = b$ über \mathbb{Q} und \mathbb{R} unterschiedlich sind.

Aufgabe I.19. Seien U, V, W Vektorräume und $f: U \rightarrow V$, $g: V \rightarrow W$ und $h: W \rightarrow X$ lineare Abbildungen.

- (a) Zeigen Sie $\text{rk}(g \circ f) + \text{rk}(h \circ g) \leq \text{rk}(g) + \text{rk}(h \circ g \circ f)$ (FROBENIUS-Ungleichung).
Hinweis: Für $g(V) = g(f(U)) \oplus Y$ gilt $h(g(V)) = h(g(f(U))) + h(Y)$.
- (b) Folgern Sie Lemma 5.15(a) aus Teil (a).
- (c) Zeigen Sie $\text{rk}(A) + \text{rk}(B) \leq \text{rk}(AB) + n$ für $A \in K^{m \times n}$ und $B \in K^{n \times k}$ (SYLVESTER-Ungleichung).

Aufgabe I.20. Sei K ein Körper und $n \in \mathbb{N}$. Zeigen Sie:

- (a) Die Summe und das Produkt von (strikten) oberen (bzw. unteren) Dreiecksmatrizen in $K^{n \times n}$ sind wieder (strikte) obere (bzw. untere) Dreiecksmatrizen (Definition 8.14).
- (b) Die oberen (bzw. unteren) Dreiecksmatrizen bilden einen Unterraum U von $K^{n \times n}$. Berechnen Sie $\dim U$.
- (c) Die Menge der invertierbaren oberen (bzw. unteren) Dreiecksmatrizen in $K^{n \times n}$ ist eine Untergruppe von $\text{GL}(n, K)$.

Aufgabe I.21. Sei V ein K -Vektorraum. Für $\varphi \in \text{GL}(V)$ und $v \in V$ sei $f_{\varphi, v}: V \rightarrow V$, $w \mapsto \varphi(w) + v$. Die Abbildungen der Form $f_{\text{id}_V, v}$ nennt man *Translationen*. Zeigen Sie:

(a)

$$\text{Aff}(V) := \{f_{\varphi, v} : \varphi \in \text{GL}(V), v \in V\} \subseteq \text{Abb}(V, V)$$

ist eine Gruppe bzgl. Komposition von Abbildungen. Handelt es sich um eine Untergruppe von $\text{GL}(V)$?

(b) Die Translationen bilden eine Untergruppe von $\text{Aff}(V)$.

Bemerkung: Man nennt $\text{Aff}(V)$ die *affine* Gruppe von V .

Aufgabe I.22. Sei $A \in \text{GL}(n, K)$. Zeigen Sie, dass man die Matrix $\begin{pmatrix} A \\ I_n \end{pmatrix} \in K^{2n \times n}$ durch elementare Spaltenoperationen in die Form $\begin{pmatrix} 1_n \\ B \end{pmatrix}$ überführen kann. Dabei ist $B = A^{-1}$.

Aufgabe I.23. Seien V, W Vektorräume und $f \in \text{Hom}(V, W)$. Zeigen Sie:

- (a) Genau dann ist f injektiv, wenn ein $g \in \text{Hom}(W, V)$ mit $g \circ f = \text{id}_V$ existiert.
- (b) Genau dann ist f surjektiv, wenn ein $g \in \text{Hom}(W, V)$ mit $f \circ g = \text{id}_W$ existiert.

(c) Sind die Abbildungen g jeweils eindeutig bestimmt?

Aufgabe I.24. Sei $\lambda \in K$ ein Eigenwert von $A \in K^{n \times n}$ und $k \in \mathbb{N}$. Zeigen Sie, dass λ^k ein Eigenwert von A^k ist. Gilt auch die Umkehrung? Zeigen Sie, dass λ^{-1} ein Eigenwert von A^{-1} ist, falls A invertierbar ist.

Aufgabe I.25. Sei $A \in K^{n \times n}$ diagonalisierbar. Zeigen Sie, dass auch A^t diagonalisierbar ist mit den gleichen Eigenwerten. Stimmen auch die Eigenräume überein?

Aufgabe I.26. Seien $a, b \in K$ und

$$A = \begin{pmatrix} a & b & \cdots & b \\ b & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & b \\ b & \cdots & b & a \end{pmatrix} \in K^{n \times n}.$$

Zeigen Sie:

$$\det(A) = (a - b)^{n-1}(a + (n - 1)b).$$

Hinweis: Eigenwerte.

Aufgabe I.27. Ist das Schiebepuzzle auf dem Cover lösbar?

Aufgabe I.28. Transpositionen der Form $(a, a + 1) \in S_n$ nennt man *Basistransposition*. Zeigen Sie, dass jede Permutation ein Produkt von Basistranspositionen ist.

Bemerkung: Dies ist die Grundlage von *Bubblesort*.

Aufgabe I.29. Zeigen Sie $|A_n| = \frac{n!}{2}$ für $n \geq 2$.

Hinweis: Wenden Sie die Leibniz-Formel auf die Matrix $(1)_{i,j=1}^n \in \mathbb{Q}^{n \times n}$ an.

Aufgabe I.30. Seien $\sigma, \tau \in S_n$ mit $\text{sgn}(\sigma) \neq \text{sgn}(\tau)$. Zeigen Sie $\det(P_\sigma + P_\tau) = 0$.

Hinweis: $P_\sigma + P_\tau = P_\sigma(P_\sigma^{-1} + P_\tau^{-1})P_\tau$.

Aufgabe I.31. Sei $\sigma \in S_n$. Zeigen Sie

$$\text{sgn}(\sigma) = \prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i}.$$

Hinweis: Betrachten Sie zunächst Transpositionen σ .

Aufgabe I.32 (Allgemeiner Entwicklungssatz). Seien $A \in K^{r \times s}$, $B \in K^{s \times t}$, $I \subseteq \{1, \dots, r\}$ und $J \subseteq \{1, \dots, t\}$ mit $|I| = |J| = k$. Für eine Matrix $M = (m_{ij})$ sei $M_{IJ} := (m_{ij})_{i \in I, j \in J}$. Zeigen Sie:

$$\det((AB)_{IJ}) = \sum_{\substack{L \subseteq \{1, \dots, s\} \\ |L|=k}} \det(A_{IL}) \det(B_{LJ}).$$

Bemerkung: Dies verallgemeinert die Matrizenmultiplikation, den Determinantensatz und die Cauchy-Binet-Formel.

Aufgabe I.33. Beim Spiel *Lights Out* ist ein 5×5 -Grid von Lichtern gegeben, die an oder aus sein können. Berührt man ein Licht, so ändert das Licht zusammen mit seinen horizontal und vertikalen Nachbarn (oben/unten, rechts/links) den Zustand (an \leftrightarrow aus). Ziel des Spiels ist, alle Lichter eines gegebenen Zustands auszuschalten.

- (a) Überlegen Sie sich, dass eine Lösung eindeutig durch einen Vektor in \mathbb{F}_2^{25} gegeben ist, wobei jede Koordinate beschreibt, ob das entsprechende Licht berührt werden muss.
- (b) Modulieren Sie die Lösung des Spiels als Gleichungssystem mit Koeffizientenmatrix in $\mathbb{F}_2^{25 \times 25}$.
- (c) Prüfen Sie (mit Computer), wie viele der 2^{25} Zustände lösbar sind.
- (d) Wie viele Lösungen besitzt ein lösbarer Zustand und worin unterscheiden sich die Lösungen?
- (e) Wie viele „Züge“ (Lichtschalter-Berührungen) sind im worst case zur Lösung erforderlich?
- (f) Welche Zustände mit nur einem brennenden Licht sind lösbar?
- (g) Entwickeln Sie einen leicht zu lernenden Algorithmus zum Lösen des Spiels, der ohne Computerberechnungen auskommt.

Hinweis: Spielen Sie hier: <https://raw.org/research/solving-lightsout-using-linear-algebra>

Lineare Algebra II

10 Polynome

10.1 Der Vektorraum der Polynome

Bemerkung 10.1. In Bemerkung 9.39 haben wir angedeutet, dass die Eigenwerte einer Matrix A Lösungen gewisser (nicht-linearer) Polynomgleichungen sind. Wir definieren in dem Kapitel Polynome mit Koeffizienten in einem beliebigen Körper und untersuchen deren Nullstellen. Daraus leiten wir ein notwendiges und hinreichendes Kriterium für die Diagonalisierbarkeit eines Endomorphismus ab.

Definition 10.2. Ein (formales) *Polynom* über einem Körper K in der *Variablen* X ist eine Summe der Form

$$\alpha = \sum_{k=0}^d a_k X^k = a_0 + a_1 X + \dots + a_d X^d$$

mit *Koeffizienten* $a_0, \dots, a_d \in K$.¹

- Man nennt a_0 das *Absolutglied* von α .
- Sofern nicht alle Koeffizienten 0 sind, nennt man

$$\deg(\alpha) := \max\{d \in \mathbb{N}_0 : a_d \neq 0\}$$

den *Grad* von α und a_d den *führenden Koeffizienten*.² Im Fall $a_d = 1$ heißt α *normiert*.

- Für das *Nullpolynom* (alle Koeffizienten sind 0) setzt man $\deg(0) := -\infty$.
- Die Menge aller Polynome über K wird mit $K[X]$ bezeichnet.

Bemerkung 10.3.

- (a) Kennt man den Grad von $\alpha \in K[X]$ nicht, so schreibt man $\alpha = \sum_{k=0}^{\infty} a_k X^k = \sum a_k X^k$ unter der Annahme, dass nur endlich viele Koeffizienten ungleich 0 sind.
- (b) Polynome werden als gleich angesehen, wenn sie die gleichen Koeffizienten haben, d. h.

$$\sum_{k=0}^{\infty} a_k X^k = \sum_{k=0}^{\infty} b_k X^k \iff \forall k \in \mathbb{N}_0 : a_k = b_k.$$

- (c) Die Körperelemente $\lambda \in K$ werden mit den *konstanten* Polynomen $\lambda X^0 \in K[X]$ identifiziert. Dies sind genau die Polynome vom Grad ≤ 0 . Insbesondere gilt $0, 1 \in K \subseteq K[X]$.

Beispiel 10.4. Das Polynom $\alpha = X^2 - 3X + 1 \in \mathbb{Q}[X]$ ist normiert vom Grad 2 mit Absolutglied 1.

¹Formal: Ein Polynom ist eine Abbildung $\mathbb{N}_0 \rightarrow K$, $k \mapsto a_k$ mit $|\{k \in \mathbb{N}_0 : a_k \neq 0\}| < \infty$.

²auch *Leitkoeffizient* genannt

Satz 10.5. *Mit den Verknüpfungen*

$$\begin{aligned}\sum a_k X^k + \sum b_k X^k &:= \sum (a_k + b_k) X^k, \\ \lambda \sum a_k X^k &:= \sum (\lambda a_k) X^k\end{aligned}$$

wird $K[X]$ ein unendlich-dimensionaler K -Vektorraum mit Basis $1, X, X^2, \dots$

Beweis. Seien $\alpha = \sum a_k X^k$ und $\beta = \sum b_k X^k$ mit $d := \deg(\alpha) \geq \deg(\beta)$. Man kann (a_0, \dots, a_d) und (b_0, \dots, b_d) als Vektoren in K^{d+1} ansehen. Die Verknüpfungen in $K[X]$ entsprechen genau denen in K^{d+1} . Daher erfüllt $K[X]$ die Vektorraumaxiome. Nach Definition ist jedes Polynom eine endliche Linearkombination von $1, X, X^2, \dots$, d. h. $K[X] = \langle 1, X, X^2, \dots \rangle$. Aus der Eindeutigkeit der Koeffizienten (Bemerkung 10.3) folgt die lineare Unabhängigkeit von $\{1, X, X^2, \dots\}$. \square

Bemerkung 10.6.

- (a) Für $\alpha, \beta \in K[X]$ und $\lambda \in K$ gilt offenbar $\deg(\alpha + \beta) \leq \max\{\deg(\alpha), \deg(\beta)\}$ und $\deg(\lambda\alpha) \leq \deg(\alpha)$. Daher bilden die Polynome vom Grad kleiner d einen d -dimensionalen Unterraum mit Basis $1, X, \dots, X^{d-1}$.
- (b) Sie wissen vermutlich, dass man Polynome auch multiplizieren kann, z. B.

$$\begin{aligned}(2X^3 - X^2 + 5X - 1)(4X^2 + 3) &= 8X^5 - 4X^4 + (6 + 20)X^3 + (-3 - 4)X^2 + 15X - 3 \\ &= 8X^5 - 4X^4 + 26X^3 - 7X^2 + 15X - 3\end{aligned}$$

Dies lässt sich wie folgt formalisieren.

Satz 10.7. *Für Polynome $\alpha = \sum a_k X^k$, $\beta = \sum b_k X^k$ ist*

$$\alpha \cdot \beta := \sum_{k=0}^{\infty} \left(\sum_{l=0}^k a_l b_{k-l} \right) X^k$$

ein Polynom vom Grad $\deg(\alpha) + \deg(\beta)$. Es gelten folgende Rechenregeln:

$$\alpha\beta = \beta\alpha, \quad \alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma, \quad \alpha(\beta\gamma) = (\alpha\beta)\gamma.$$

Beweis. Im Fall $\alpha = 0$ oder $\beta = 0$ ist $\alpha\beta = 0$ und $\deg(\alpha\beta) = -\infty = \deg(\alpha) + \deg(\beta)$. Sei also $d := \deg(\alpha) \geq 0$ und $e := \deg(\beta) \geq 0$. Für $k > d + e$ ist $\sum_{l=0}^k a_l b_{k-l} = 0$ und $\deg(\alpha\beta) \leq d + e$. Für $k = d + e$ ist $\sum_{l=0}^k a_l b_{k-l} = a_d b_e \neq 0$. Dies zeigt $\deg(\alpha\beta) = d + e$. Insbesondere ist $\alpha\beta \in K[X]$. Für $\gamma = \sum c_k X^k$ gilt

$$\begin{aligned}\alpha\beta &= \sum_{k=0}^{\infty} \left(\sum_{l=0}^k a_l b_{k-l} \right) X^k = \sum_{k=0}^{\infty} \left(\sum_{l=0}^k b_l a_{k-l} \right) X^k = \beta\alpha \\ \alpha(\beta + \gamma) &= \sum_{k=0}^{\infty} \left(\sum_{l=0}^k a_l (b_{k-l} + c_{k-l}) \right) X^k = \sum_{k=0}^{\infty} \left(\sum_{l=0}^k a_l b_{k-l} \right) X^k + \sum_{k=0}^{\infty} \left(\sum_{l=0}^k a_l c_{k-l} \right) X^k = \alpha\beta + \alpha\gamma.\end{aligned}$$

Der Koeffizient von X^k in $\alpha(\beta\gamma)$ ist

$$\sum_{l=0}^k a_l \sum_{m=0}^{k-l} b_m c_{k-l-m} = \sum_{\substack{r,s,t \in \mathbb{N}_0 \\ r+s+t=k}} a_r b_s c_t = \sum_{l=0}^k \left(\sum_{m=0}^l a_m b_{l-m} \right) c_{k-l}.$$

Dies ist auch der Koeffizient von X^k in $(\alpha\beta)\gamma$. Also ist $\alpha(\beta\gamma) = (\alpha\beta)\gamma$. \square

Bemerkung 10.8. Im Gegensatz zur Matrizenmultiplikation ist die Multiplikation von Polynomen kommutativ. Das einzige Körperaxiom, welches $K[X]$ nicht erfüllt, ist die Existenz von Inversen. Zum Beispiel existiert kein $\alpha \in K[X]$ mit $X \cdot \alpha = 1$. Dennoch gilt die Kürzungsregel: $\alpha\beta = \alpha\gamma \Rightarrow \beta = \gamma$, falls $\alpha \neq 0$. Dies folgt aus

$$\deg(\beta - \gamma) \leq \deg(\alpha) + \deg(\beta - \gamma) = \deg(\alpha(\beta - \gamma)) = \deg(0) = -\infty.$$

Man kann in $K[X]$ also wie in \mathbb{Z} rechnen.

Satz 10.9 (Division mit Rest). Für $\alpha, \beta \in K[X]$ mit $\beta \neq 0$ existieren eindeutig bestimmte Polynome $\gamma, \delta \in K[X]$ mit $\alpha = \beta\gamma + \delta$ und $\deg \delta < \deg \beta$.

Beweis. Existenz: Wähle $\gamma \in K[X]$, sodass

$$\delta := \alpha - \beta\gamma = a_d X^d + a_{d-1} X^{d-1} + \dots + a_0$$

möglichst kleinen Grad $d \in \mathbb{N}_0 \cup \{-\infty\}$ hat. Sei $\beta = b_e X^e + \dots + b_0$ und $e := \deg \beta$. Gilt $d \geq e$, so ist

$$a_d b_e^{-1} X^{d-e} \beta = a_d b_e^{-1} (b_e X^d + b_{e-1} X^{d-1} + \dots + b_0 X^{d-e}) = a_d X^d + \dots$$

und es folgt

$$\deg(\alpha - \beta(\gamma + a_d b_e^{-1} X^{d-e})) = \deg(\delta - a_d b_e^{-1} X^{d-e} \beta) < d.$$

Dies ist ein Widerspruch zur Wahl von γ . Also ist $d < e$ und $\alpha = \beta\gamma + \delta$.

Eindeutigkeit: Sei nun $\alpha = \beta\tilde{\gamma} + \tilde{\delta}$ mit $\tilde{\gamma}, \tilde{\delta} \in K[X]$ und $\deg \tilde{\delta} < e$. Nach Satz 10.7 ist

$$e + \deg(\tilde{\gamma} - \gamma) = \deg(\beta) + \deg(\tilde{\gamma} - \gamma) = \deg(\beta(\tilde{\gamma} - \gamma)) = \deg(\delta - \tilde{\delta}) \leq \max\{\deg(\delta), \deg(\tilde{\delta})\} < e.$$

Es folgt $\deg(\tilde{\gamma} - \gamma) = -\infty = \deg(\delta - \tilde{\delta})$. Dies zeigt $\tilde{\gamma} = \gamma$ und $\tilde{\delta} = \delta$. □

Definition 10.10. In der Situation von Satz 10.9 nennt man δ den *Rest* bei der Division von α durch β . Im Fall $\delta = 0$ nennt man β einen *Teiler* von α und schreibt $\beta \mid \alpha$. Ggf. sagt man auch „ β teilt α “ oder „ α ist durch β teilbar“.

Beispiel 10.11.

$$\begin{array}{r} (2X^3 \quad -X^2 \quad +5X \quad +1) : (X^2 + 3) = 2X - 1 =: \gamma \\ -(2X^3 \quad \quad \quad +6X) \\ \hline \quad \quad -X^2 \quad -X \quad +1 \\ \quad -(-X^2 \quad \quad \quad -3) \\ \hline \quad \quad \quad -X \quad +4 =: \delta \end{array}$$

Also $\alpha = 2X^3 - X^2 + 5X + 1 = (X^2 + 3)(2X - 1) - X + 4 = \beta\gamma + \delta$ mit $\deg \delta = 1 < 2 = \deg \beta$.

Bemerkung 10.12. Die Division durch normierte Polynome vom Grad 1 lässt sich mit dem *HORNER-Schema*³ effizient gestalten. Sei dazu $\alpha = a_n X^n + a_{n-1} X^{n-1} + \dots + a_0$ und $\beta = X - b$. Wir berechnen $c_n := 0$, $c_k := a_{k+1} + bc_{k+1}$ für $k = n-1, \dots, 0$ und $d := a_0 + bc_0$:

$$\begin{array}{cccccc} & a_n & a_{n-1} & a_{n-2} & \cdots & a_0 \\ + & 0 & bc_{n-1} & bc_{n-2} & \cdots & bc_0 \\ \hline c_{n-1} & c_{n-2} & \cdots & c_0 & d & \end{array}$$

³auch RUFFINIS *Regel* genannt

Für $\gamma := c_{n-1}X^{n-1} + \dots + c_0$ gilt nun

$$\begin{aligned}\beta\gamma + d &= c_{n-1}X^n + (c_{n-2} - bc_{n-1})X^{n-1} + \dots + (c_0 - bc_1)X - bc_0 + d \\ &= a_nX^n + a_{n-1}X^{n-1} + \dots + a_0 = \alpha.\end{aligned}$$

Beispiel 10.13. Für $\alpha = 2X^3 - X^2 + 3X + 1$ und $\beta = X - 2$ erhält man:

$$\begin{array}{rcccc} & 2 & -1 & 3 & 1 \\ + & 0 & 4 & 6 & 18 \\ \hline & 2 & 3 & 9 & 19 \end{array}$$

Dies zeigt $\alpha = \beta(2X^2 + 3X + 9) + 19$.

10.2 Nullstellen

Definition 10.14. Sei $\alpha = \sum_{k=0}^d a_k X^k \in K[X]$. Man kann ein Element $x \in K$ für X in α einsetzen:

$$\alpha(x) := \sum_{k=0}^d a_k x^k \in K.$$

Man nennt x eine *Nullstelle* von α , falls $\alpha(x) = 0$.

Lemma 10.15. Für $\alpha, \beta \in K[X]$ und $x \in K$ gilt

$$\begin{aligned}(\alpha + \beta)(x) &= \alpha(x) + \beta(x), \\ (\alpha\beta)(x) &= \alpha(x)\beta(x).\end{aligned}$$

Beweis. Seien $\alpha = \sum a_k X^k$ und $\beta = \sum b_k X^k$. Dann gilt

$$\begin{aligned}(\alpha + \beta)(x) &= \sum (a_k + b_k)x^k = \sum a_k x^k + \sum b_k x^k = \alpha(x) + \beta(x), \\ (\alpha\beta)(x) &= \sum_{n=0}^{\infty} \sum_{k=0}^n a_k b_{n-k} x^n = \sum_{n=0}^{\infty} \sum_{k=0}^n (a_k x^k)(b_{n-k} x^{n-k}) = \sum_{n=0}^{\infty} a_n x^n \sum_{k=0}^n b_k x^k = \alpha(x)\beta(x). \quad \square\end{aligned}$$

Bemerkung 10.16. Merkgel: Es ist egal, ob Sie erst addieren/multiplizieren und danach einsetzen oder erst einsetzen und danach addieren/multiplizieren. Achtung: Im Allgemeinen ist $\alpha(x+y) \neq \alpha(x) + \alpha(y)$ und $\alpha(xy) \neq \alpha(x)\alpha(y)$ für $\alpha \in K[X]$ und $x, y \in K$.

Satz 10.17 (Interpolation). Seien $x_1, \dots, x_n \in K$ paarweise verschieden und $y_1, \dots, y_n \in K$ beliebig. Dann existiert genau ein Polynom α vom Grad $< n$ mit $\alpha(x_i) = y_i$ für $i = 1, \dots, n$.

Beweis. Sei $\alpha = a_0 + a_1 X + \dots + a_{n-1} X^{n-1}$. Die Bedingung $\alpha(x_i) = y_i$ für $i = 1, \dots, n$ bedeutet:

$$\begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

Die Koeffizientenmatrix dieses Gleichungssystems ist eine Vandermonde-Matrix. Da die x_i paarweise verschieden sind, ist die Matrix nach Satz 9.20 invertierbar. Also existiert genau eine Lösung (a_0, \dots, a_{n-1}) . \square

Bemerkung 10.18. Eine explizite Lösung der Interpolationsaufgabe ist durch das LAGRANGE-Polynom

$$\alpha := \sum_{i=1}^n y_i \prod_{j \neq i} \frac{X - x_j}{x_i - x_j} \in K[X]$$

gegeben (nachrechnen).

Beispiel 10.19.

- (a) Für $n = 2$ und $K = \mathbb{R}$ ist Satz 10.17 die geometrische Aussage, dass zwei verschiedene Punkte im \mathbb{R}^2 durch genau eine Gerade verbunden sind.
- (b) Wir suchen ein Polynom $\alpha \in \mathbb{R}[X]$ durch die Punkte $(-1, 2)$, $(0, 1)$ und $(1, 3)$. Der Beweis von Satz 10.17 führt auf das Gleichungssystem

$$\begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}$$

mit der eindeutigen Lösung $\alpha = 1 + \frac{1}{2}X + \frac{3}{2}X^2$.

Folgerung 10.20.

- (a) Jedes Polynom $\alpha \in K[X]$ vom Grad $d \geq 0$ besitzt höchstens d Nullstellen.
- (b) Sei $|K| = \infty$ und $\alpha, \beta \in K[X]$ mit $\alpha(x) = \beta(x)$ für alle $x \in K$. Dann gilt $\alpha = \beta$.

Beweis.

- (a) Angenommen α besitzt paarweise verschiedene Nullstellen $x_1, \dots, x_{d+1} \in K$. Nach Satz 10.17 mit $y_1 = \dots = y_{d+1} = 0$ ist α das einzige Polynom vom Grad $\leq d$ mit diesen Nullstellen. Andererseits hat das Nullpolynom auch diese Nullstellen. Also gilt $\alpha = 0$ und $d = -\infty$ im Widerspruch zur Annahme.
- (b) Wegen $|K| = \infty$ besitzt $\alpha - \beta$ unendlich viele Nullstellen. Aus (a) folgt $\alpha = \beta$. □

Bemerkung 10.21. Für $K = \mathbb{R}$ ist jedes Polynom $\alpha \in \mathbb{R}[X]$ eindeutig durch die (stetige) Funktion $\mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto \alpha(x)$ bestimmt, denn $|K| = \infty$. In der Analysis unterscheidet man daher nicht zwischen Polynom und Funktion. Über endlichen Körpern K würde man dabei Information verlieren, denn es gibt nur endlich viele Abbildungen $K \rightarrow K$, aber unendlich viele Polynome. Zum Beispiel entsprechen die Polynome $X, X^2, \dots \in \mathbb{F}_2[X]$ alle der Identität $\text{id}_{\mathbb{F}_2}$.

Lemma 10.22. Genau dann ist $x \in K$ eine Nullstelle von α , wenn $(X - x) \mid \alpha$.

Beweis. Division mit Rest liefert $\gamma, \delta \in K[X]$ mit $\alpha = (X - x)\gamma + \delta$ und $\deg \delta < \deg(X - x) = 1$, d. h. $\delta \in K$. Nun ist

$$\delta = \delta(x) = (\alpha - (X - x)\gamma)(x) \stackrel{10.15}{=} \alpha(x) - (x - x)\gamma(x) = \alpha(x). \quad \square$$

Definition 10.23. Sei $x \in K$ eine Nullstelle von α . Man nennt $X - x$ einen *Linearfaktor* von α . Die größte Zahl $e \in \mathbb{N}$ mit $(X - x)^e \mid \alpha$ nennt man die *algebraische Vielfachheit* der Nullstelle x . Im Fall $e = 1$ spricht man von einer *einfachen* Nullstelle und anderenfalls von einer *mehrfachen* Nullstelle.

Lemma 10.24. Seien $x_1, \dots, x_n \in K$. Dann hat jeder normierte Teiler von $(X - x_1) \dots (X - x_n) \in K[X]$ die Form $(X - x_{i_1}) \dots (X - x_{i_k})$ mit $1 \leq i_1 < \dots < i_k \leq n$.

Beweis. Im Fall $n = 1$ sind 1 (das leere Produkt mit $k = 0$) und $X - x_1$ die einzigen normierten Teiler. Sei also $n \geq 2$ und die Behauptung für $n - 1$ bereits bewiesen. Seien $\alpha, \beta \in K[X]$ mit $(X - x_1) \dots (X - x_n) = \alpha\beta$. Dann ist $\alpha(x_n)\beta(x_n) = (\alpha\beta)(x_n) = 0$, o. B. d. A. sei $\alpha(x_n) = 0$. Nach Lemma 10.22 gilt $\alpha = (X - x_n)\gamma$ für ein $\gamma \in K[X]$. Nach Bemerkung 10.8 darf man $X - x_n$ kürzen und erhält $(X - x_1) \dots (X - x_{n-1}) = \gamma\beta$. Die Behauptung folgt nun durch Induktion. \square

Beispiel 10.25.

(a) Sei $\alpha = X^3 + X^2 - 5X + 3 \in \mathbb{R}[X]$. Eine Nullstelle $x \in \mathbb{R}$ ist eine Lösung der Gleichung

$$x^3 + x^2 - 5x + 3 = 0.$$

Auch wenn es Lösungsformeln für solche Gleichungen (dritten und vierten Grades⁴) gibt, sind diese in der Praxis aufwendig. Wir werden unsere Beispiele (und Übungsaufgaben) daher so wählen, dass man „kleine“ ganzzahlige Nullstellen erraten kann. Angenommen es gibt eine Nullstelle $x \in \mathbb{Z}$. Wegen $x(x^2 + x - 5) = -3$ ist x ein Teiler von 3 , d. h. $x \in \{\pm 1, \pm 3\}$. Man prüft leicht, dass $x_1 = 1$ tatsächlich eine Nullstelle ist ($1^3 + 1^2 - 5 \cdot 1 + 3 = 0$). Polynomdivision (zum Beispiel mit dem Horner-Schema) ergibt

$$(X^3 + X^2 - 5X + 3) : (X - 1) = X^2 + 2X - 3 =: \gamma.$$

Für jede Nullstelle $y \in \mathbb{R}$ von γ gilt nun $\alpha(y) = (y - 1)\gamma(y) = 0$, d. h. y ist auch eine Nullstelle von α . Mit der p - q -Formel $\frac{1}{2}(-p \pm \sqrt{p^2 - 4q})$ für quadratische Gleichungen erhält man die Nullstellen von γ :

$$x_2 = \frac{1}{2}(-2 + \sqrt{4 + 12}) = 1, \quad x_3 = \frac{1}{2}(-2 - \sqrt{4 + 12}) = -3.$$

Daher ist $x_1 = x_2 = 1$ eine Nullstelle von α mit algebraischer Vielfachheit 2 (eine *doppelte* Nullstelle). Außerdem *zerfällt* α in Linearfaktoren $\alpha = (X - 1)^2(X + 3)$.

- (b) Offensichtlich ist $\alpha(0)$ das Absolutglied von $\alpha \in K[X]$. Also ist $x = 0$ genau dann eine Nullstelle von α , wenn das Absolutglied von α verschwindet.
- (c) Bekanntlich besitzt $X^2 + 1 \in \mathbb{R}[X]$ keine Nullstelle. Wir konstruieren später einen „größeren“ Körper, über dem auch dieses Polynom in Linearfaktoren zerfällt (Lemma 11.27).
- (d) Das Polynom $X^2 + X + 1 \in \mathbb{F}_2[X]$ hat keine Nullstelle in \mathbb{F}_2 , denn es kommen nur 0 und 1 in Frage.

10.3 Charakteristische Polynome

Bemerkung 10.26. Im Folgenden betrachten wir Matrizen mit Einträgen in $K[X]$. Aufgrund der Rechenregeln für Polynome (Satz 10.7) überlegt man sich leicht, dass die gewohnten Rechenregeln (Lemma 5.8) für Matrizen auch in $K[X]^{n \times n}$ gelten. Schließlich kann man sogar die Definition der Determinante auf Matrizen in $K[X]^{n \times n}$ anwenden (dabei werden Matrixeinträge nur addiert und multipliziert, aber niemals dividiert). Ebenso bleiben der Determinantensatz, die Laplace-Entwicklung, die Leibniz-Formel und der Satz 9.22 über die komplementäre Matrix in dieser größeren Allgemeinheit richtig. Andererseits funktioniert der Gauß-Algorithmus in $K[X]^{n \times n}$ nicht, denn hier muss dividiert werden.

⁴siehe Algebra-Skript

Definition 10.27. Für $A = (a_{ij}) \in K^{n \times n}$ betrachten wir die Matrix

$$X1_n - A = \begin{pmatrix} X - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & X - a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & -a_{n-1,n} \\ -a_{n1} & \cdots & -a_{n,n-1} & X - a_{nn} \end{pmatrix} \in K[X]^{n \times n}.$$

Man nennt $\chi_A := \det(X1_n - A) \in K[X]$ das *charakteristische Polynom* von A .⁵

Lemma 10.28. *Ähnliche Matrizen haben das gleiche charakteristische Polynom.*

Beweis. Für $A \in K^{n \times n}$ und $S \in \text{GL}(n, K)$ gilt

$$\chi_{SAS^{-1}} = \det(X1_n - SAS^{-1}) = \det(S(X1_n - A)S^{-1}) \stackrel{10.28}{=} \det(X1_n - A) = \chi_A. \quad \square$$

Definition 10.29. Sei V ein n -dimensionaler K -Vektorraum mit Basis B . Für $f \in \text{End}(V)$ definiert man

$$\det(f) := \det({}_B[f]_B), \quad \chi_f := \det(X1_n - {}_B[f]_B).$$

Nach Folgerung 7.27, Folgerung 9.12 und Lemma 10.28 hängen $\det(f)$ und χ_f nicht von der Wahl von B ab. In den nachfolgenden Sätzen kann man Matrizen also durch Endomorphismen ersetzen (und umgekehrt).

Beispiel 10.30. Das charakteristische Polynom von $A := \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \in \mathbb{Q}[X]$ ist

$$\chi_A = \det \begin{pmatrix} X - 1 & -2 \\ -3 & X - 4 \end{pmatrix} = (X - 1)(X - 4) - (-2)(-3) = X^2 - 5X - 2 = X^2 - \text{tr}(A)X + \det(A).$$

Lemma 10.31. *Für $A \in K^{n \times n}$ gilt $\chi_A = X^n - \text{tr}(A)X^{n-1} + \dots + (-1)^n \det(A)$. Insbesondere ist χ_A normiert vom Grad n .*

Beweis. Sei $A = (a_{ij})$. Nach der Leibniz-Formel gilt

$$\begin{aligned} \chi_A = \det(X1_n - A) &= (X - a_{11})(X - a_{22}) \cdots (X - a_{nn}) \\ &+ \sum_{\sigma \in S_n \setminus \{\text{id}\}} \text{sgn}(\sigma) (\delta_{1\sigma(1)}X - a_{1\sigma(1)}) \cdots (\delta_{n\sigma(n)}X - a_{n\sigma(n)}). \end{aligned}$$

Für $\sigma \in S_n \setminus \{\text{id}\}$ existiert ein $k \in \{1, \dots, n\}$ mit $l := \sigma(k) \neq k$. Da σ injektiv ist, gilt $\sigma(l) \neq \sigma(k) = l$. Daher ist $\delta_{k\sigma(k)} = 0 = \delta_{l\sigma(l)}$ und

$$(\delta_{1\sigma(1)}X - a_{1\sigma(1)}) \cdots (\delta_{n\sigma(n)}X - a_{n\sigma(n)})$$

ist ein Polynom vom Grad $\leq n - 2$. Insgesamt ist

$$\chi_A = (X - a_{11})(X - a_{22}) \cdots (X - a_{nn}) + \alpha$$

mit $\deg(\alpha) \leq n - 2$. Ausmultiplizieren zeigt $\chi_A = X^n - (a_{11} + \dots + a_{nn})X^{n-1} + \dots = X^n - \text{tr}(A)X^{n-1} + \dots$. Zur Berechnung des Absolutglieds setzt man $X = 0$ und erhält $\chi_A(0) \stackrel{10.15}{=} \det(-A) = (-1)^n \det(A)$ aus Bemerkung 9.8. \square

⁵In manchen Büchern definiert man χ_A durch $\det(A - X1_n) = (-1)^n \det(X1_n - A)$. Das macht keinen großen Unterschied, aber bringt den Nachteil, dass χ_A nicht normiert ist, wenn n ungerade ist.

Satz 10.32. Die Eigenwerte von $A \in K^{n \times n}$ sind die Nullstellen von χ_A .

Beweis. Es gilt

$$\text{Ker}(A - \lambda 1_n) \neq \{0\} \iff \det(A - \lambda 1_n) = 0 \stackrel{9.8}{\iff} \det(\lambda 1_n - A) = 0 \stackrel{10.15}{\iff} \chi_A(\lambda) = 0. \quad \square$$

Lemma 10.33. Sei $\lambda \in K$ ein Eigenwert von $f \in \text{End}(V)$. Dann ist die geometrische Vielfachheit von λ höchstens so groß wie die algebraische Vielfachheit von λ als Nullstelle von χ_f .

Beweis. Man ergänze eine Basis b_1, \dots, b_e von $E_\lambda(f)$ zu einer Basis $B := \{b_1, \dots, b_n\}$ von V . Dann gilt

$$\chi_f = \det(X 1_n - {}_B[f]_B) = \det \begin{pmatrix} (X - \lambda) 1_e & * \\ 0 & * \end{pmatrix}.$$

Lemma 9.9 zeigt $\chi_f = (X - \lambda)^e \beta$ für ein $\beta \in K[X]$. Also ist die algebraische Vielfachheit von λ als Nullstelle von χ_f mindestens e . \square

Satz 10.34. Genau dann ist $f \in \text{End}(V)$ diagonalisierbar, wenn χ_f in Linearfaktoren zerfällt und für jede Nullstelle von χ_f die algebraische Vielfachheit mit der geometrischen Vielfachheit übereinstimmt.

Beweis. Seien $\lambda_1, \dots, \lambda_k \in K$ die verschiedenen Nullstellen von χ_f mit algebraischen Vielfachheiten m_1, \dots, m_k . Dann existiert ein $\alpha \in K[X]$ mit $\chi_f = (X - \lambda_1)^{m_1} \dots (X - \lambda_k)^{m_k} \alpha$. Sei m'_i die geometrische Vielfachheit von λ_i als Eigenwert von f . Nach Lemma 10.33 gilt

$$m'_1 + \dots + m'_k \leq m_1 + \dots + m_k + \deg(\alpha) \stackrel{10.7}{=} \deg((X - \lambda_1)^{m_1} \dots (X - \lambda_k)^{m_k} \alpha) = \deg(\chi_f) = \dim V.$$

Nach Satz 8.10 besitzt V genau dann eine Basis aus Eigenvektoren, wenn $m'_1 + \dots + m'_k = \dim V$. Dies gilt genau dann, wenn χ_f in Linearfaktoren zerfällt (d. h. $\alpha = 1$) und die algebraischen Vielfachheiten mit den geometrischen Vielfachheiten übereinstimmen (d. h. $m_i = m'_i$ für $i = 1, \dots, k$). \square

Bemerkung 10.35. Zerfällt χ_A in Linearfaktoren, so gilt

$$\chi_A = (X - \lambda_1) \dots (X - \lambda_n) = X^n - (\lambda_1 + \dots + \lambda_n) X^{n-1} + \dots + (-1)^n \lambda_1 \dots \lambda_n.$$

Ein Vergleich mit Lemma 10.31 zeigt:

$$\begin{array}{l} \text{tr}(A) = \lambda_1 + \dots + \lambda_n, \\ \det(A) = \lambda_1 \dots \lambda_n, \end{array}$$

d. h. die Spur ist die Summe der Eigenwerte und die Determinante ist das Produkt der Eigenwerte (sofern diese existieren). Hat man bereits $\lambda_1, \dots, \lambda_{n-1}$ bestimmt, so erhält man $\lambda_n = \text{tr}(A) - \lambda_1 - \dots - \lambda_{n-1}$.

Satz 10.36 (MIRSKY). Sei $d_1, \dots, d_n, \lambda_1, \dots, \lambda_n \in K$ mit $d_1 + \dots + d_n = \lambda_1 + \dots + \lambda_n$. Dann existiert eine Matrix $A \in K^{n \times n}$ mit Hauptdiagonale d_1, \dots, d_n und Eigenwerten $\lambda_1, \dots, \lambda_n$.

Beweis. Im Fall $n = 1$ erfüllt $A = (d_1) = (\lambda_1)$ die Behauptung. Sei $n \geq 2$ und $A = (a_{ij}) \in K^{n \times n} \setminus K 1_n$ eine Dreiecksmatrix mit Hauptdiagonale $\lambda_1, \dots, \lambda_n$. Dann hat A Eigenwerte $\lambda_1, \dots, \lambda_n$. Nach dem Satz von Fillmore ist A zu einer Matrix mit Hauptdiagonale d_1, \dots, d_n ähnlich. Diese hat die gleichen Eigenwerte wie A . \square

Beispiel 10.37.

- (a) Wir suchen eine Matrix mit Eigenwerten 1, 1, 1 und Hauptdiagonale 0, 0, 3. Dafür wenden wir den Satz von Fillmore auf

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

an. Der Übergang zur Basis $\{(1, 0, 0), (1, 1, 0), (1, 2, 1)\}$ liefert

$$A \approx \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & -3 \\ 0 & 1 & 3 \end{pmatrix}.$$

- (b) Die FIBONACCI-Zahlen F_k sind rekursiv definiert:

$$F_k := k \quad (k = 0, 1) \quad F_{k+1} := F_k + F_{k-1} \quad (k \geq 1).$$

k	0	1	2	3	4	5	6	7	8	9	10
F_k	0	1	1	2	3	5	8	13	21	34	55

Wir suchen eine explizite Formel für F_k . Für $A := \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \in \mathbb{R}^{2 \times 2}$ gilt

$$\begin{pmatrix} F_{k+1} \\ F_k \end{pmatrix} = A \begin{pmatrix} F_k \\ F_{k-1} \end{pmatrix} = A^2 \begin{pmatrix} F_{k-1} \\ F_{k-2} \end{pmatrix} = \dots = A^k \begin{pmatrix} F_1 \\ F_0 \end{pmatrix} = A^k \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Um A^k zu berechnen, diagonalisieren wir A . Wegen $\chi_A = (X - 1)X - 1 = X^2 - X - 1$ hat A die Eigenwerte $\varphi := \frac{1+\sqrt{5}}{2}$ und $\psi := \frac{1-\sqrt{5}}{2}$ (man nennt $\varphi \approx 1.618$ den *goldenen Schnitt*). Man berechnet

$$E_\varphi(A) = \text{Ker}(A - \varphi 1_2) = \left\langle \begin{pmatrix} \varphi \\ 1 \end{pmatrix} \right\rangle,$$

$$E_\psi(A) = \left\langle \begin{pmatrix} \psi \\ 1 \end{pmatrix} \right\rangle.$$

Für $S := \begin{pmatrix} \varphi & \psi \\ 1 & 1 \end{pmatrix}$ gilt also $S^{-1}AS = \text{diag}(\varphi, \psi)$ und

$$A^k = (S \text{diag}(\varphi, \psi) S^{-1})^k = S \text{diag}(\varphi^k, \psi^k) S^{-1} = S \text{diag}(\varphi^k, \psi^k) S^{-1}.$$

Nach Beispiel 9.23 ist

$$S^{-1} = \frac{1}{\det(S)} \tilde{S} = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & -\psi \\ -1 & \varphi \end{pmatrix}.$$

Insgesamt erhält man

$$A^k = S \text{diag}(\varphi^k, \psi^k) S^{-1} = \frac{1}{\sqrt{5}} \begin{pmatrix} \varphi^{k+1} & \psi^{k+1} \\ \varphi^k & \psi^k \end{pmatrix} \begin{pmatrix} 1 & -\psi \\ -1 & \varphi \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} * & * \\ \varphi^k - \psi^k & * \end{pmatrix}$$

und

$$\boxed{F_k = \frac{1}{\sqrt{5}}(\varphi^k - \psi^k)} \quad (\text{BINET-Formel}^6)$$

Wegen $|\psi^k| \approx 0.618^k \rightarrow 0$ gilt $F_k \approx \frac{1}{\sqrt{5}}\varphi^k$, d. h. F_k wächst exponentiell.

⁶Man kann die Formel auch durch Induktion beweisen, sofern man sie zuvor erraten hat.

Lemma 10.38. Für $A \in K^{m \times n}$ und $B \in K^{n \times m}$ gilt $X^n \chi_{AB} = X^m \chi_{BA}$. Insbesondere haben AB und BA die gleichen von 0 verschiedenen Eigenwerte.

Beweis. Nach den Regeln für Blockmatrizen (Bemerkung 5.9, Lemma 9.9) und dem Determinantensatz gilt

$$\begin{aligned} X^n \chi_{AB} &= \det \left(\begin{pmatrix} 1_m & 0 \\ -B & X1_n \end{pmatrix} \begin{pmatrix} X1_m - AB & A \\ 0 & 1_n \end{pmatrix} \begin{pmatrix} 1_m & 0 \\ B & 1_n \end{pmatrix} \right) \\ &= \det \left(\begin{pmatrix} 1_m & 0 \\ -B & X1_n \end{pmatrix} \begin{pmatrix} X1_m & A \\ B & 1_n \end{pmatrix} \right) = \det \begin{pmatrix} X1_m & A \\ 0 & X1_n - BA \end{pmatrix} = X^m \chi_{BA}. \end{aligned}$$

Jeder Eigenwert $\lambda \in K$ von AB ist eine Nullstelle von $X^n \chi_{AB} = X^m \chi_{BA}$. Im Fall $\lambda \neq 0$ muss λ eine Nullstelle von χ_{BA} sein. Dann ist λ auch ein Eigenwert von BA (und umgekehrt). \square

Bemerkung 10.39. Im Fall $n = m$ gilt sogar $\chi_{AB} = \chi_{BA}$ in der Situation von Lemma 10.38.

10.4 Minimalpolynome

Bemerkung 10.40. Wir haben bereits Polynome in Matrizen eingesetzt. Wir setzen nun umgekehrt Matrizen in Polynome ein. Für $\alpha = \sum_{k=0}^d a_k X^k \in K[X]$ und $A \in K^{n \times n}$ definieren wir

$$\alpha(A) := \sum_{k=0}^d a_k A^k \in K^{n \times n}.$$

Die Regeln aus Lemma 10.15 gelten auch in dieser Allgemeinheit.

Satz 10.41. Für $A \in K^{n \times n}$ existiert genau ein normiertes Polynom $\mu_A \in K[X] \setminus \{0\}$ mit $\mu_A(A) = 0_n$ und $\deg(\mu_A)$ minimal.

Beweis. Wegen $\dim K^{n \times n} = n^2$ (Lemma 5.4) sind die Potenzen $1_n = A^0, A, A^2, \dots, A^{n^2}$ linear abhängig in $K^{n \times n}$. Also existieren $a_0, \dots, a_{n^2} \in K$ (nicht alle 0) mit $\sum_{k=0}^{n^2} a_k A^k = 0$. Für $\alpha = \sum a_k X^k \in K[X]$ gilt somit $\alpha(A) = 0$. Indem man durch den führenden Koeffizienten von α teilt, kann man annehmen, dass α normiert ist. Dies zeigt, dass μ_A existiert. Sei auch $\tilde{\mu} \in K[X]$ normiert mit $\tilde{\mu}(A) = 0$ und $\deg(\tilde{\mu}) = \deg(\mu_A)$ minimal. Dann ist $(\mu_A - \tilde{\mu})(A) = \mu_A(A) - \tilde{\mu}(A) = 0$ und $\deg(\mu_A - \tilde{\mu}) < \deg(\mu_A)$. Die Minimalität von $\deg(\mu_A)$ zeigt $\mu_A - \tilde{\mu} = 0$, d. h. μ_A ist eindeutig bestimmt. \square

Definition 10.42. Man nennt μ_A das *Minimalpolynom* von A .

Beispiel 10.43. Sei $A := \begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix} \in \mathbb{Q}^{2 \times 2}$. Da A keine Skalarmatrix ist, gilt $\deg \mu_A \geq 2$. Wir machen den Ansatz

$$A^2 + xA + y1_2 = \begin{pmatrix} -1 & -1 \\ 2 & -2 \end{pmatrix} + x \begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix} + y \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

mit $x, y \in \mathbb{Q}$. Ein Vergleich der Matrixeinträge an Position $(1, 2)$ zeigt $x = -1$. Tatsächlich gilt die Gleichung nun für $y = 2$. Daher ist $\mu_A = X^2 - X + 2$.

Lemma 10.44. Sei $A \in K^{n \times n}$ und $\alpha \in K[X]$ mit $\alpha(A) = 0$. Dann gilt $\mu_A \mid \alpha$.

Beweis. Wir dividieren mit Rest: $\alpha = \mu_A \gamma + \delta$ mit $\gamma, \delta \in K[X]$ und $\deg(\delta) < \deg(\mu_A)$. Dann ist

$$\delta(A) = (\alpha - \mu_A \gamma)(A) = \alpha(A) - \mu_A(A) \gamma(A) = 0.$$

Aus der Minimalität von $\deg(\mu_A)$ folgt $\delta = 0$ und $\mu_A \mid \alpha$. □

Lemma 10.45. *Ähnliche Matrizen haben das gleiche Minimalpolynom.*

Beweis. Sei $A \in K^{n \times n}$ mit $\mu_A = \sum a_k X^k$. Für $S \in \text{GL}(n, K)$ gilt

$$\mu_A(SAS^{-1}) = \sum a_k (SAS^{-1})^k = \sum a_k SA^k S^{-1} = S \left(\sum a_k A^k \right) S^{-1} = S \mu_A(A) S^{-1} = 0_n.$$

Aus Lemma 10.44 folgt $\mu_{SAS^{-1}} \mid \mu_A$. Da Ähnlichkeit eine symmetrische Relation ist, gilt auch $\mu_A \mid \mu_{SAS^{-1}}$. Da beide Minimalpolynome normiert sind, müssen sie gleich sein. □

Definition 10.46. Sei V ein K -Vektorraum mit Basis B . Für $f \in \text{End}(V)$ sei wie üblich $\mu_f := \mu_{B[f]_B}$. Nach Lemma 10.45 hängt μ_f nicht von der Wahl von B ab. Die folgenden Sätze über Matrizen gelten sinngemäß auch für Endomorphismen.

Bemerkung 10.47. Aus dem Beweis von Satz 10.41 erhält man $\deg(\mu_A) \leq n^2$. Der nächste Satz impliziert $\deg(\mu_A) \leq n$.

Satz 10.48 (CAYLEY-HAMILTON). *Für $A \in K^{n \times n}$ gilt $\chi_A(A) = 0$ und $\mu_A \mid \chi_A$.*

Beweis. Sei $B := X1_n - A \in K[X]^{n \times n}$ und $\tilde{B} \in K[X]^{n \times n}$ die zu B komplementäre Matrix. Aus jedem Eintrag von \tilde{B} extrahieren wir den Koeffizienten von X^k und bilden daraus die Matrix $B_k \in K^{n \times n}$. Es gilt nun

$$\tilde{B} = \sum_{k=0}^{\infty} B_k X^k,$$

wobei nur endlich viele der B_k ungleich 0 sind. Sei $\chi_A = \sum a_k X^k$. Nach Satz 9.22 gilt

$$\sum_{k=0}^{\infty} a_k 1_n X^k = \chi_A 1_n = \det(B) 1_n = \tilde{B} B = \sum_{k=0}^{\infty} B_k X^k (X 1_n - A) = \sum_{k=0}^{\infty} (B_{k-1} - B_k A) X^k,$$

wobei $B_{-1} := 0_n$. Ein Koeffizientenvergleich ergibt $a_k 1_n = B_{k-1} - B_k A$ für $k = 0, 1, \dots$. Daher ist

$$\chi_A(A) = \sum_{k=0}^{\infty} a_k A^k = \sum_{k=0}^{\infty} (B_{k-1} A^k - B_k A^{k+1}) = \sum_{k=0}^{\infty} B_{k-1} A^k - \sum_{k=0}^{\infty} B_k A^{k+1} = 0.$$

Die zweite Behauptung folgt aus Lemma 10.44. □

Beispiel 10.49.

(a) Für $A \in K^{2 \times 2}$ gilt $A^2 - \text{tr}(A)A + \det(A)1_2 = 0$ nach Lemma 10.31.

(b) Sei $A \in \text{GL}(n, K)$ mit $\chi_A = \mu_A \gamma$ für ein $\gamma \in K[X]$. Nach Lemma 10.31 gilt

$$\mu_A(0)\gamma(0) = \chi_A(0) = \det(A) \neq 0.$$

Also hat μ_A die Form $\mu_A = X^d + a_{d-1}X^{d-1} + \dots + a_0$ mit $a_0 \neq 0$. Man kann nun die Gleichung $A^d + a_{d-1}A^{d-1} + \dots + a_01_n = 0$ auf beiden Seiten mit A^{-1} multiplizieren und erhält $A^{d-1} + a_{d-1}A^{d-2} + \dots + a_11_n + a_0A^{-1} = 0$. Dies liefert eine Formel für die Inverse

$$A^{-1} = -\frac{1}{a_0}(A^{d-1} + a_{d-1}A^{d-2} + \dots + a_11_n).$$

Speziell für $n = 2$:

$$A^{-1} \stackrel{(a)}{=} \frac{1}{\det(A)}(\text{tr}(A)1_2 - A) = \frac{1}{\det(A)}\tilde{A}$$

(vgl. Beispiel 9.23).

Satz 10.50. Die Eigenwerte von $A \in K^{n \times n}$ sind die Nullstellen von μ_A , d. h. χ_A und μ_A haben die gleichen Nullstellen (nicht unbedingt mit den gleichen Vielfachheiten).

Beweis. Nach Cayley-Hamilton ist jede Nullstelle von μ_A auch eine Nullstelle von χ_A und damit ein Eigenwert von A (Satz 10.32). Sei umgekehrt $\lambda \in K$ ein Eigenwert von A mit Eigenvektor $v \in K^{n \times 1}$. Für $k \in \mathbb{N}_0$ ist $A^k v = A^{k-1} \lambda v = \dots = \lambda^k v$. Sei $\mu_A = \sum a_k X^k$. Dann gilt

$$0 = \mu_A(A)v = \sum a_k A^k v = \sum a_k \lambda^k v = \mu_A(\lambda)v.$$

Wegen $v \neq 0$ ist $\mu_A(\lambda) = 0$, d. h. λ ist eine Nullstelle von μ_A . □

Bemerkung 10.51. Wegen $\deg \mu_A \leq \deg \chi_A$ vereinfacht Satz 10.50 die Bestimmung der Eigenwerte. Andererseits ist nicht klar, wie man μ_A effizient berechnet. Der nächste Satz verbessert Folgerung 8.12.

Satz 10.52. Genau dann ist $A \in K^{n \times n}$ diagonalisierbar, wenn μ_A in paarweise verschiedene Linearfaktoren zerfällt.

Beweis. Sei $A \in K^{n \times n}$ diagonalisierbar. Dann existiert $S \in \text{GL}(n, K)$ mit $S^{-1}AS = \text{diag}(\lambda_1, \dots, \lambda_n)$. Die λ_i lassen sich sortieren, indem man die Spalten von S (d. h. die Eigenvektoren von A) entsprechend anordnet. Nach Lemma 10.45 können wir

$$A = \begin{pmatrix} \lambda_1 1_{n_1} & & 0 \\ & \ddots & \\ 0 & & \lambda_k 1_{n_k} \end{pmatrix}$$

annehmen, wobei $n = n_1 + \dots + n_k$ und $\lambda_i \neq \lambda_j$ für $i \neq j$. Dann gilt

$$(A - \lambda_1 1_n) \dots (A - \lambda_k 1_n) = \text{diag}(0_{n_1}, *, \dots, *) \text{diag}(*, \dots, *, 0_{n_2}, *, \dots, *) \dots \text{diag}(*, \dots, *, 0_{n_k}) = 0_n.$$

Nach Lemma 10.44 ist μ_A ein Teiler von $(X - \lambda_1) \dots (X - \lambda_k)$. Andererseits sind $\lambda_1, \dots, \lambda_k$ Eigenwerte und damit Nullstellen von μ_A . Dies zeigt $\mu_A = (X - \lambda_1) \dots (X - \lambda_k)$.

Nehmen wir umgekehrt $\mu_A = (X - \lambda_1) \dots (X - \lambda_k)$ mit paarweise verschiedenen $\lambda_1, \dots, \lambda_k$ an. Sei P_k der k -dimensionale Vektorraum aller Polynome vom Grad kleiner k (Bemerkung 10.6). Für $i = 1, \dots, k$ sei

$$\gamma_i := (X - \lambda_1) \dots (X - \lambda_{i-1})(X - \lambda_{i+1}) \dots (X - \lambda_k) \in P_k.$$

Seien $a_1, \dots, a_k \in K$ mit $a_1\gamma_1 + \dots + a_k\gamma_k = 0$. Dann gilt

$$a_i(\lambda_i - \lambda_1) \dots (\lambda_i - \lambda_{i-1})(\lambda_i - \lambda_{i+1}) \dots (\lambda_i - \lambda_n) = a_i\gamma_i(\lambda_i) = (a_1\gamma_1 + \dots + a_k\gamma_k)(\lambda_i) = 0$$

und es folgt $a_i = 0$ für $i = 1, \dots, k$. Also sind $\gamma_1, \dots, \gamma_k$ linear unabhängig in P_k . Wegen $\dim P_k = k$ bilden sie sogar eine Basis. Insbesondere existieren $b_1, \dots, b_k \in K$ mit $\gamma := b_1\gamma_1 + \dots + b_k\gamma_k = X^0 = 1$. Für $v \in K^{n \times 1}$ gilt

$$(A - \lambda_i 1_n)\gamma_i(A)v = \mu_A(A)v = 0,$$

d. h. $\gamma_i(A)v$ liegt in $E_{\lambda_i}(A)$. Andererseits ist

$$v = 1_n v = A^0 v = \gamma(A)(v) = b_1\gamma_1(A)(v) + \dots + b_k\gamma_k(A)(v).$$

Dies zeigt $K^{n \times 1} = E_{\lambda_1}(A) + \dots + E_{\lambda_k}(A)$. Nach Bemerkung 8.6 ist A diagonalisierbar. \square

Beispiel 10.53. Sei $A \in K^{n \times n}$ mit genau zwei verschiedenen Eigenwerten. Angenommen wir finden einen Vektor $v \in K^{n \times 1}$, sodass v, Av, A^2v linear unabhängig sind. Dann sind auch $1_n, A, A^2$ linear unabhängig. Dies zeigt $\deg \mu_A \geq 3$. Nach Satz 10.52 ist A nicht diagonalisierbar.

Satz 10.54. Für $A \in K^{n \times n}$ gilt $\chi_A \mid \mu_A^n$.

Beweis. Sei $\mu_A = \sum a_i X^i$. Für $i \geq 1$ ist

$$(X^i 1_n - A^i) = (X 1_n - A)(X^{i-1} 1_n + X^{i-2} A + \dots + X A^{i-2} + A^{i-1}).$$

Es folgt

$$\mu_A 1_n = \mu_A(X 1_n) - \mu_A(A) = \sum_{i \geq 1} a_i (X^i 1_n - A^i) = (X 1_n - A)B$$

für ein $B \in K[X]^{n \times n}$. Bildet man auf beiden Seiten die Determinante, so ergibt sich $\mu_A^n = \chi_A \det(B)$. \square

Bemerkung 10.55. Die algebraische Vielfachheit eines Eigenwerts von $A \in K^{n \times n}$ beträgt höchstens n . Zerfällt $\mu_A = (X - \lambda_1) \dots (X - \lambda_k)$ in Linearfaktoren, so gilt daher $\chi_A \mid (X - \lambda_1)^n \dots (X - \lambda_k)^n = \mu_A^n$ (der Beweis von Satz 10.54 kommt ohne diese Annahme aus).

11 Euklidische Geometrie

11.1 Skalarprodukte

Bemerkung 11.1. Wir betrachten in diesem Kapitel $K = \mathbb{R}$. Im Gegensatz zu beliebigen Körpern kann man \mathbb{R} in positive und negative Zahlen unterteilen. Für $x \in \mathbb{R}$ sei wie üblich

$$|x| := \begin{cases} x & \text{falls } x \geq 0, \\ -x & \text{falls } x < 0 \end{cases}$$

der *Betrag* von x . Wir benutzen außerdem, dass jede positive reelle Zahl genau eine positive Quadratwurzel besitzt. Es gilt also $|x| = \sqrt{x^2}$ für alle $x \in \mathbb{R}$.

Definition 11.2. Sei V ein \mathbb{R} -Vektorraum. Eine Abbildung $V \times V \rightarrow \mathbb{R}$, $(v, w) \mapsto [v, w]$ heißt *Skalarprodukt*¹, falls folgende Bedingungen für alle $u, v, w \in V$ und $\lambda \in \mathbb{R}$ gelten:

- $[v, v] \geq 0$ mit Gleichheit genau dann, wenn $v = 0$ (*positiv definit*),
- $[v, w] = [w, v]$ (*symmetrisch*),
- $[\lambda u + v, w] = \lambda[u, w] + [v, w]$ (*bilinear*).

Zusammen mit einem Skalarprodukt wird V ein *euklidischer Raum*. Vektoren $v, w \in V$ heißen *orthogonal*, falls $[v, w] = 0$. Man nennt $|v| := \sqrt{[v, v]} \geq 0$ die *Norm* von v . Im Fall $|v| = 1$ nennt man v *normiert*.

Bemerkung 11.3.

(a) Die Symmetrie des Skalarprodukts zeigt

$$[u, \lambda v + w] = [\lambda v + w, u] = \lambda[v, u] + [w, u] = \lambda[u, v] + [u, w]$$

für alle $u, v, w \in V$ und $\lambda \in \mathbb{R}$. Für ein festes $x \in V$ sind also die Abbildungen $V \rightarrow \mathbb{R}$, $v \mapsto [v, x]$ und $V \rightarrow \mathbb{R}$, $v \mapsto [x, v]$ linear (dies erklärt den Begriff *bilinear*). Insbesondere ist $[v, 0] = 0 = [0, v]$ für alle $v \in V$. Dennoch ist die Abbildung $V \times V \rightarrow \mathbb{R}$, $(v, w) \mapsto [v, w]$ *nicht* linear, also kein Funktional, denn $[v, v] > 0 = [0, v] + [v, 0]$ für $v \neq 0$.

(b) Jeder Unterraum eines euklidischen Raums ist selbst ein euklidischer Raum mit dem eingeschränkten Skalarprodukt.

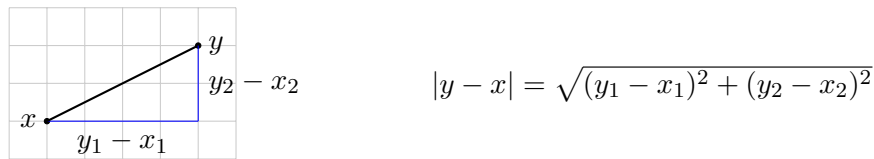
¹Die Schreibweise $[v, w]$ ist in der Literatur nicht einheitlich. Man findet auch $\langle v, w \rangle$ (Verwechslung mit Spann), $(v | w)$ u. ä. In allgemeinerem Kontext (Definition 17.48) benutzen $\|v\|$ anstatt $|v|$.

Beispiel 11.4.

(a) Das wichtigste Beispiel eines euklidischen Raums ist $V = \mathbb{R}^n$ mit dem *Standardskalarprodukt*

$$[x, y] := xy^t = \sum_{i=1}^n x_i y_i \quad (x, y \in \mathbb{R}^n).$$

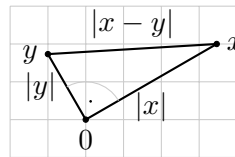
Man überprüft leicht die drei Eigenschaften (positiv definit, symmetrisch und bilinear). Im Fall $n = 1$ ist $|x| = \sqrt{x_1^2}$ der gewöhnliche Betrag (dies rechtfertigt die Verwendung der Betragsstriche). Nach dem Satz des Pythagoras² entspricht die Norm $|y - x|$ im \mathbb{R}^2 dem geometrischen Abstand zwischen x und y :



Sind x und y orthogonal, so erhält man

$$|x - y|^2 = [x - y, x - y] = [x, x] - 2[x, y] + [y, y] = |x|^2 + |y|^2.$$

Nach der Umkehrung vom Satz des Pythagoras bilden x und y einen rechten Winkel, d. h. sie stehen senkrecht aufeinander (man schreibt $x \perp y$):



Im Allgemeinen gilt die *Parallelogrammgleichung*:

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

(b) Vektoren in \mathbb{R}^n sind „diskrete“ Funktionen $\{1, \dots, n\} \rightarrow \mathbb{R}$. In der Analysis betrachtet man eine „kontinuierliche“ Variante: Die stetigen Abbildungen $[0, 1] \rightarrow \mathbb{R}$ auf dem abgeschlossenen Intervall $[0, 1]$ bilden einen (unendlich-dimensionalen) Unterraum $V \leq \text{Abb}([0, 1], \mathbb{R})$ mit Skalarprodukt

$$[f, g] := \int_0^1 f(x)g(x) dx \quad (f, g \in V).$$

Lemma 11.5. Sei V ein euklidischer Raum, $v, w \in V$ und $\lambda \in \mathbb{R}$. Dann gilt:

- (a) $|\lambda v| = |\lambda||v|$ (Homogenität).
- (b) $|[v, w]| \leq |v||w|$ mit Gleichheit genau dann, wenn v und w linear abhängig sind (CAUCHY-SCHWARZ-Ungleichung).
- (c) $||v| - |w|| \leq |v + w| \leq |v| + |w|$ (Dreiecksungleichung).

Beweis.

(a) $|\lambda v| = \sqrt{[\lambda v, \lambda v]} = \sqrt{\lambda^2[v, v]} = \sqrt{\lambda^2} \sqrt{[v, v]} = |\lambda||v|.$

²Man könnte auch den Abstand zwischen x und y durch $|y - x|$ definieren und damit den Satz des Pythagoras beweisen.

(b) O. B. d. A. sei $w \neq 0$. Sei $\lambda := \frac{[v,w]}{[w,w]}$. Nach den Eigenschaften des Skalarprodukts gilt

$$0 \leq |v - \lambda w|^2 = [v - \lambda w, v - \lambda w] = [v, v] - 2\lambda[v, w] + \lambda^2[w, w] = |v|^2 - \frac{[v, w]^2}{|w|^2}.$$

Es folgt $[v, w]^2 \leq |v|^2|w|^2$ und $|[v, w]| \leq |v||w|$. Gleichheit impliziert $v = \lambda w$, d. h. v und w sind linear abhängig. Sind umgekehrt v und w linear abhängig gegeben, dann existiert ein $\mu \in \mathbb{R}$ mit $v = \mu w$ und $|[v, w]| = |\mu||w|^2 \stackrel{(a)}{=} |\mu w||w| = |v||w|$.

(c) Zunächst ist

$$|v + w|^2 = [v + w, v + w] = [v, v] + 2[v, w] + [w, w] \stackrel{(b)}{\leq} |v|^2 + 2|v||w| + |w|^2 = (|v| + |w|)^2$$

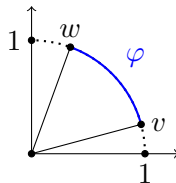
und $|v + w| \leq |v| + |w|$. Daraus folgt $|v| = |v + w - w| \leq |v + w| + |w|$ und $|v| - |w| \leq |v + w|$. Vertauschen von v und w liefert $-(|v| - |w|) = |w| - |v| \leq |v + w|$, also $||v| - |w|| \leq |v + w|$. \square

Bemerkung 11.6.

- (a) Sind $v, w \in \mathbb{R}^n$ linear unabhängig, so bilden $0, v$ und $v + w$ ein Dreieck mit Seiten $|v|, |w|$ und $|v + w|$. Die Dreiecksungleichung besagt, dass die Summe von je zwei Seiten größer ist als die dritte Seite.
- (b) Die Cauchy-Schwarz-Ungleichung impliziert $-1 \leq \frac{[v,w]}{|v||w|} \leq 1$ für $v, w \in V \setminus \{0\}$. Dieser Bruch verändert sich durch positive Skalierung von v und w nicht:

$$\frac{[\lambda v, \mu w]}{|\lambda v||\mu w|} = \frac{\lambda\mu[v, w]}{|\lambda||\mu||v||w|} = \frac{[v, w]}{|v||w|} \quad (\lambda, \mu > 0)$$

Seien also v und w normiert. Dann definiert man den *Winkel* φ (im Bogenmaß) zwischen v und w als die Länge des Bogens auf dem Einheitskreis³ zwischen v und w :



Die Länge des Halbkreisbogens nennt man π und berechnet $\pi \approx 3.14$ (Aufgabe II.10). Der *Kosinus* von φ wird durch $\cos \varphi := [v, w]$ definiert.⁴ Es gilt

$$\cos 0 = [e_1, e_1] = 1, \quad \cos(\pi/2) = [e_1, e_2] = 0, \quad \cos \pi = [e_1, -e_1] = -1.$$

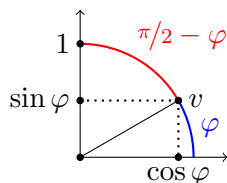
Durch $\cos(\varphi + 2k\pi) = \cos \varphi$ für $k \in \mathbb{Z}$ setzt sich \cos periodisch auf ganz \mathbb{R} fort. Dabei gilt $\cos(-\varphi) = \cos \varphi$ und $\cos(\pi - \varphi) = -\cos \varphi$. Die „verschobene“ Funktion

$$\sin \varphi := \cos(\varphi - \pi/2) = \cos(\pi/2 - \varphi)$$

³in der Ebene $\langle v, w \rangle$

⁴Man kann zeigen, dass diese Definition mit der analytischen Definition als Potenzreihe übereinstimmt.

für $\varphi \in \mathbb{R}$ nennt man den *Sinus* von φ . Für einen beliebigen normierten Vektor $v = (x, y)$ gilt $x = [v, e_1] = \cos \varphi$ und $y = [v, e_2] = \cos(\pi/2 - \varphi) = \sin \varphi$:



11.2 Orthonormalbasen

Definition 11.7. Sei V ein n -dimensionaler euklidischer Raum. Vektoren b_1, \dots, b_n bilden eine *Orthonormalbasis* von V , falls sie normiert und paarweise orthogonal sind, d. h. $[b_i, b_j] = \delta_{ij}$ für $1 \leq i, j \leq n$.

Bemerkung 11.8. Eine Orthonormalbasis b_1, \dots, b_n von V ist tatsächlich eine Basis. Dafür genügt es die lineare Unabhängigkeit zu prüfen. Seien $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ mit $\lambda_1 b_1 + \dots + \lambda_n b_n = 0$. Dann gilt

$$\lambda_i = \sum_{j=1}^n \lambda_j [b_j, b_i] = \left[\sum_{j=1}^n \lambda_j b_j, b_i \right] = [0, b_i] = 0$$

für $i = 1, \dots, n$.

Beispiel 11.9. Die Standardbasis $e_1, \dots, e_n \in \mathbb{R}^n$ ist eine Orthonormalbasis bzgl. des Standardskalarprodukts. Jede Permutation einer Orthonormalbasis ist wieder eine Orthonormalbasis.

Satz 11.10 (GRAM-SCHMIDT-Verfahren). *Seien $v_1, \dots, v_k \in V$ linear unabhängig. Wir definieren rekursiv:*

$$b_s := v_s - \sum_{i=1}^{s-1} \frac{[v_s, b_i]}{[b_i, b_i]} b_i \quad (s = 1, \dots, k).$$

Dann sind b_1, \dots, b_k paarweise orthogonal mit $\langle v_1, \dots, v_k \rangle = \langle b_1, \dots, b_k \rangle$. Folglich ist $\frac{1}{|b_1|} b_1, \dots, \frac{1}{|b_k|} b_k$ eine Orthonormalbasis von $\langle v_1, \dots, v_k \rangle$.

Beweis. Induktion nach k : Für $k = 1$ ist $b_1 = v_1 \neq 0$. Sei nun $k \geq 2$ und die Behauptung für $k - 1$ bereits bewiesen, d. h. $\langle v_1, \dots, v_{k-1} \rangle = \langle b_1, \dots, b_{k-1} \rangle$ und $[b_i, b_j] = 0$ für $1 \leq i < j \leq k - 1$. Wegen $\sum_{i=1}^{k-1} \frac{[v_k, b_i]}{[b_i, b_i]} b_i \in \langle b_1, \dots, b_{k-1} \rangle$ gilt

$$\langle v_1, \dots, v_k \rangle = \langle b_1, \dots, b_{k-1}, v_k \rangle = \langle b_1, \dots, b_k \rangle.$$

Für $i = 1, \dots, k - 1$ ist außerdem

$$[b_k, b_i] = [v_k, b_i] - \sum_{j=1}^{k-1} \frac{[v_k, b_j]}{[b_j, b_j]} [b_j, b_i] = [v_k, b_i] - [v_k, b_i] = 0.$$

Damit ist die erste Behauptung bewiesen. Wegen $|\frac{1}{|b_i|} b_i| = \frac{1}{|b_i|} |b_i| = 1$ folgt die zweite Behauptung. \square

Folgerung 11.11. *Jeder euklidische Raum besitzt (mindestens) eine Orthonormalbasis.*

Beweis. Man wende das Gram-Schmidt-Verfahren auf eine beliebige Basis an. \square

Beispiel 11.12. Seien $v_1 := (1, 0, 1)$, $v_2 := (0, 1, 1)$ und $v_3 := (-1, 2, 0)$ linear unabhängig in \mathbb{R}^3 . Bezüglich des Standardskalarprodukts erhält man

$$\begin{aligned} b_1 &:= v_1 = (1, 0, 1), \\ b_2 &:= v_2 - \frac{[v_2, b_1]}{[b_1, b_1]} b_1 = (0, 1, 1) - \frac{1}{2}(1, 0, 1) = \frac{1}{2}(-1, 2, 1), \\ b_3 &:= v_3 - \frac{[v_3, b_1]}{[b_1, b_1]} b_1 - \frac{[v_3, b_2]}{[b_2, b_2]} b_2 = (-1, 2, 0) + \frac{1}{2}(1, 0, 1) - \frac{5}{6}(-1, 2, 1) = \frac{1}{3}(1, 1, -1). \end{aligned}$$

Man beachte, dass Skalierungsfaktoren in dieser Rechnung keine Rolle spielen. Man kann b_3 also etwas bequemer mit $b_2 = (-1, 2, 1)$ ausrechnen. Nach Normierung ist $\frac{1}{\sqrt{2}}(1, 0, 1)$, $\frac{1}{\sqrt{6}}(-1, 2, 1)$, $\frac{1}{\sqrt{3}}(1, 1, -1)$ eine Orthonormalbasis von \mathbb{R}^3 . Um die Einträge der Vektoren zu minimieren, kann es nützlich sein zuerst den Gauß-Algorithmus anzuwenden, bevor man das Gram-Schmidt-Verfahren startet. In diesem Beispiel würde man am Ende die Standardbasis von \mathbb{R}^3 erhalten.

Definition 11.13. Sei V ein euklidischer Raum und $S \subseteq V$. Dann nennt man

$$S^\perp := \{v \in V : \forall s \in S : [v, s] = 0\}$$

das *orthogonale Komplement* von S in V . Im Fall $S = \{s\}$ schreibt man $s^\perp := S^\perp$.

Bemerkung 11.14. Für $v, w \in S^\perp$ und $\lambda \in \mathbb{R}$ gilt $[\lambda v + w, s] = \lambda[v, s] + [w, s] = 0$ für alle $s \in S$, d.h. $\lambda v + w \in S^\perp$. Daher ist S^\perp ein Unterraum, selbst wenn S nur eine Teilmenge ist. Außerdem gilt $S^\perp = \langle S \rangle^\perp$.

Lemma 11.15. Für Unterräume U, W eines euklidischen Raums V gilt

- (a) $V = U \oplus U^\perp$ und $\dim V = \dim U + \dim U^\perp$.
- (b) $(U^\perp)^\perp = U$.
- (c) $U \subseteq W \iff W^\perp \subseteq U^\perp$.

Beweis.

- (a) Für $v \in U \cap U^\perp$ gilt $|v|^2 = [v, v] = 0$ und $v = 0$. Also ist $U \cap U^\perp = \{0\}$ und $U + U^\perp = U \oplus U^\perp$. Wir können eine Basis v_1, \dots, v_k von U zu einer Basis v_1, \dots, v_n von V ergänzen. Das Gram-Schmidt-Verfahren liefert eine Orthonormalbasis b_1, \dots, b_n von V mit $\langle v_1, \dots, v_k \rangle = \langle b_1, \dots, b_k \rangle$. Daher ist $b_{k+1}, \dots, b_n \in U^\perp$ und $V = U \oplus U^\perp$.
- (b) Nach Definition ist $U \subseteq (U^\perp)^\perp$. Nach (a) ist $\dim(U^\perp)^\perp = \dim V - \dim U^\perp = \dim U$. Also ist $U = (U^\perp)^\perp$.
- (c) Nach Definition gilt

$$U \subseteq W \implies W^\perp \subseteq U^\perp \stackrel{(b)}{\implies} U = (U^\perp)^\perp \subseteq (W^\perp)^\perp = W. \quad \square$$

Beispiel 11.16.

- (a) In $V = \mathbb{R}^n$ lässt sich ein orthogonales Komplement von $U \leq V$ bzgl. des Standardskalarprodukts mit dem Gauß-Algorithmus bestimmen: Man schreibt die Vektoren eines Erzeugendensystems von U als Zeilen in eine Matrix $A \in \mathbb{R}^{k \times n}$. Die Lösungsmenge L_0 des homogenen Gleichungssystems $Ax = 0$ ist U^\perp , denn nach Satz 6.6 gilt $\dim L_0 = n - \text{rk}(A) = n - \dim U = \dim U^\perp$.
- (b) Für $v = (x, y) \in \mathbb{R}^2 \setminus \{0\}$ gilt $v^\perp = \langle (y, -x) \rangle$.
- (c) Seien $v, w \in \mathbb{R}^3$ linear unabhängig. Wir ergänzen zu einer Basis u, v, w von \mathbb{R}^3 , sodass die Matrix

$$A := \begin{pmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{pmatrix}$$

Determinante 1 hat (das geht immer, indem man u geeignet skaliert). Es gibt genau einen Vektor $x \in \langle v, w \rangle^\perp$ mit $Ax^t = e_1$ und zwar

$$x^t = A^{-1} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \stackrel{9.22}{=} \tilde{A} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \det(A_{11}) \\ -\det(A_{12}) \\ \det(A_{13}) \end{pmatrix} = \begin{pmatrix} v_2w_3 - v_3w_2 \\ v_3w_1 - v_1w_3 \\ v_1w_2 - v_2w_1 \end{pmatrix} =: v \times w.$$

Man nennt $v \times w$ das *Kreuzprodukt* von v und w . Nach Konstruktion gilt $\langle v, w \rangle^\perp = \langle v \times w \rangle$. Die Richtung von $v \times w$ lässt sich mit der *Rechte-Hand-Regel* bestimmen: Zeigt v in Richtung Daumen und w in Richtung Zeigefinger, so zeigt $v \times w$ in Richtung des Mittelfingers der rechten Hand.

11.3 Symmetrische und orthogonale Abbildungen

Definition 11.17. Sei V ein euklidischer Raum und $f \in \text{End}(V)$. Man nennt f

- *symmetrisch*,⁵ falls $[f(v), w] = [v, f(w)]$ für alle $v, w \in V$.
- *orthogonal*,⁶ falls $[f(v), f(w)] = [v, w]$ für alle $v, w \in V$.

Bemerkung 11.18.

- (a) Die Nullabbildung ist symmetrisch. Sind $f, g \in \text{End}(V)$ symmetrisch und $\lambda \in \mathbb{R}$, dann ist offenbar auch $\lambda f + g$ symmetrisch. Die symmetrischen Abbildungen bilden also einen Unterraum von $\text{End}(V)$.
- (b) Orthogonale Abbildungen $f \in \text{End}(V)$ sind Isomorphismen, denn aus $v \in \text{Ker}(f)$ folgt $|v|^2 = [v, v] = [f(v), f(v)] = 0$, also $v = 0$. Wegen $|f(v - w)| = |v - w|$ und

$$\frac{[f(v), f(w)]}{|f(v)||f(w)|} = \frac{[v, w]}{|v||w|}$$

erhält f Abstände und Winkel (Bemerkung 11.6). Insbesondere bildet f Orthonormalbasen auf Orthonormalbasen ab. Mit f, g sind auch $f \circ g$ und f^{-1} orthogonal. Die Menge der orthogonalen Abbildungen bildet daher eine Untergruppe $O(V)$ von $\text{GL}(V)$. Man nennt $O(V)$ die *orthogonale Gruppe* von V .

⁵oder *selbstadjungiert*, siehe Abschnitt 13.2

⁶oder *isometrisch*

- (c) Der folgende Satz zeigt, dass langen-erhaltende Abbildung automatisch orthogonal (insbesondere linear) sind.

Satz 11.19 (MAZUR-ULAM). *Sei V ein euklidischer Raum und $f \in \text{Abb}(V, V)$ mit $|f(v)| = |v|$ fur alle $v \in V$. Dann ist $f \in \text{O}(V)$.*

Beweis. Fur $v, w \in V$ gilt

$$[f(v), f(w)] = \frac{1}{2}(|f(v)|^2 + |f(w)|^2 - |f(v) - f(w)|^2) = \frac{1}{2}(|v|^2 + |w|^2 - |v - w|^2) = [v, w].$$

Daraus folgt

$$\begin{aligned} |f(\lambda v + w) - \lambda f(v) - f(w)|^2 &= [f(\lambda v + w) - \lambda f(v) - f(w), f(\lambda v + w) - \lambda f(v) - f(w)] \\ &= [\lambda v + w - \lambda v - w, \lambda v + w - \lambda v - w] = 0 \end{aligned}$$

fur $\lambda \in \mathbb{R}$. Also ist f linear und orthogonal. □

Lemma 11.20. *Sei V euklidisch mit Orthonormalbasis B , $f \in \text{End}(V)$ und $A := {}_B[f]_B$. Dann gilt:*

(a) f symmetrisch $\iff A^t = A$.

(b) f orthogonal $\iff A^t = A^{-1}$.

Beweis. Sei $B = \{b_1, \dots, b_n\}$ und $A = (a_{ij})$. Dann ist $f(b_i) = \sum_{j=1}^n a_{ji} b_j$ fur $i = 1, \dots, n$.

- (a) Ist f symmetrisch, so gilt $a_{ji} = [f(b_i), b_j] = [b_i, f(b_j)] = a_{ij}$ fur $1 \leq i, j \leq n$. Also ist $A = A^t$. Sei umgekehrt $A = A^t$. Dann folgt $[f(b_i), b_j] = [b_i, f(b_j)]$ fur $1 \leq i, j \leq n$. Fur beliebige $v = \sum \lambda_i b_i$ und $w = \sum \mu_j b_j$ in V gilt

$$[f(v), w] = \sum_{i,j=1}^n \lambda_i \mu_j [f(b_i), b_j] = \sum_{i,j=1}^n \lambda_i \mu_j [b_i, f(b_j)] = [v, f(w)].$$

Also ist f symmetrisch.

- (b) Ist f orthogonal, so gilt

$$\sum_{k=1}^n a_{ki} a_{kj} = \left[\sum_{k=1}^n a_{ki} b_k, \sum_{k=1}^n a_{kj} b_k \right] = [f(b_i), f(b_j)] = [b_i, b_j] = \delta_{ij}.$$

Dies zeigt $A^t A = 1_n$ und $A^t = A^{-1}$. Ist umgekehrt $A^t A = 1_n$, so gilt $[f(b_i), f(b_j)] = \delta_{ij} = [b_i, b_j]$. Wie in (a) folgt $[f(v), f(w)] = [v, w]$ fur alle $v, w \in V$, d. h. f ist orthogonal. □

Bemerkung 11.21.

- (a) Fur einen beliebigen Korper K nennt man Matrizen $A \in \text{GL}(n, K)$ *orthogonal*, falls $A^t = A^{-1}$. Man zeigt leicht, dass die orthogonalen Matrizen eine Untergruppe $\text{O}(n, K)$ von $\text{GL}(n, K)$ bilden. Wie ublich entspricht $\text{O}(n, \mathbb{R})$ der Gruppe $\text{O}(V)$ durch Basiswahl (so wie sich $\text{GL}(V)$ und $\text{GL}(n, K)$ entsprechen). Fur jede orthogonale Matrix A gilt

$$\det(A)^2 \stackrel{9.12}{=} \det(A) \det(A^t) = \det(AA^t) = \det(1_n) = 1$$

und $\det(A) = \pm 1$. Man nennt

$$\text{SO}(n, K) := \text{O}(n, K) \cap \text{SL}(n, K) \leq \text{O}(n, K) \leq \text{GL}(n, K).$$

die *spezielle orthogonale Gruppe* vom Grad n uber K .

- (b) Die Gleichung $A^t A = 1_n = A A^t$ bedeutet für orthogonale reelle Matrizen, dass die Spalten (bzw. Zeilen) von A eine Orthonormalbasis von \mathbb{R}^n bzgl. des Standardskalarprodukts bilden. Sei v ein Eigenvektor von A zum Eigenwert $\lambda \in \mathbb{R}$. Dann gilt

$$|v|^2 = v^t v = v^t A^t A v = (A v)^t (A v) = \lambda^2 v^t v = \lambda^2 |v|^2$$

und $\lambda = \pm 1$.

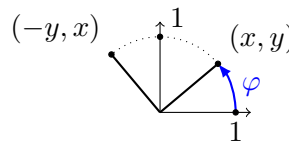
Beispiel 11.22.

- (a) Jede Permutationsmatrix ist orthogonal, denn die Zeilen bilden eine Orthonormalbasis (nämlich eine Permutation der Standardbasis).
- (b) Für $A \in O(2, \mathbb{R})$ gilt

$$A = \begin{pmatrix} x & \mp y \\ y & \pm x \end{pmatrix}$$

mit $x^2 + y^2 = 1 = \pm \det(A)$ (vgl. Beispiel 11.16).

- Im Fall $\det(A) = 1$ beschreibt A eine *Drehung* um den Winkel φ zwischen e_1 und (x, y) :



Es gilt dann

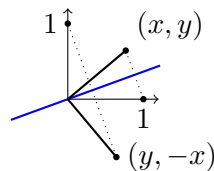
$$A = D(\varphi) := \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}.$$

Für zwei Winkel φ und ψ erhält man $D(\varphi + \psi) = D(\varphi)D(\psi)$, woraus die bekannten *Additionstheoreme* folgen:

$$\begin{aligned} \cos(\varphi + \psi) &= \cos(\varphi) \cos(\psi) - \sin(\varphi) \sin(\psi), \\ \sin(\varphi + \psi) &= \sin(\varphi) \cos(\psi) + \sin(\psi) \cos(\varphi). \end{aligned}$$

Offenbar besitzt $D(\varphi)$ nur dann (reelle) Eigenwerte, wenn $\varphi \in \{0, \pi\}$, d. h. $D(0) = 1_2$ und $D(\pi) = -1_2$.

- Im Fall $\det(A) = -1$ beschreibt A eine *Spiegelung* an der Winkelhalbierenden zwischen e_1 und (x, y) :



Es gilt dann

$$A = S(\varphi) := \begin{pmatrix} \cos \varphi & \sin \varphi \\ \sin \varphi & -\cos \varphi \end{pmatrix}.$$

Die Spiegelachse wird von einem Eigenvektor zum Eigenwert 1 aufgespannt. Orthogonal dazu steht ein Eigenvektor zum Eigenwert -1 (beachte: $\det(A)$ ist das Produkt der Eigenwerte). Nach Folgerung 8.12 ist $S(\varphi)$ diagonalisierbar. Tatsächlich ist

$$D(\varphi/2)^{-1}S(\varphi)D(\varphi/2) = D(-\varphi/2)S(\varphi)D(\varphi/2) = S(0) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Für spezielle Winkel erhält man (vgl. Aufgabe II.9):

φ	$\pi/6$	$\pi/4$	$\pi/3$	$\pi/2$	π
$D(\varphi)$	$\frac{1}{2} \begin{pmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{pmatrix}$	$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$	$\frac{1}{2} \begin{pmatrix} 1 & -\sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix}$	$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$	-1_2
$S(\varphi)$	$\frac{1}{2} \begin{pmatrix} \sqrt{3} & 1 \\ 1 & -\sqrt{3} \end{pmatrix}$	$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$	$\frac{1}{2} \begin{pmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$

Damit lässt sich auch

$$D(\pi/12) = D(\pi/3 - \pi/4) = D(\pi/3)D(\pi/4)^{-1} = D(\pi/3)D(\pi/4)^t$$

berechnen.

Bemerkung 11.23. Um zu zeigen, dass symmetrische Endomorphismen diagonalisierbar sind, müssen wir die reellen Zahlen vorübergehend verlassen.

11.4 Komplexe Zahlen

Lemma 11.24. Der \mathbb{R} -Vektorraum $\mathbb{C} := \mathbb{R}^2$ wird durch die Multiplikation

$$(a, b) \cdot (c, d) := (ac - bd, ad + bc) \quad (a, b, c, d \in \mathbb{R})$$

zu einem Körper.

Beweis. Als Vektorraum ist $(\mathbb{C}, +)$ bereits eine abelsche Gruppe. Die Multiplikation in \mathbb{C} führen wir auf die Matrizenmultiplikation zurück.⁷ Dafür betrachten wir die injektive Abbildung

$$\Gamma: \mathbb{C} \rightarrow \mathbb{R}^{2 \times 2}, \quad z = (a, b) \mapsto \Gamma(z) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

Für $x, y, z \in \mathbb{C}$ gilt $\Gamma(x) + \Gamma(y) = \Gamma(x + y)$ und $\Gamma(x)\Gamma(y) = \Gamma(x \cdot y)$ (nachrechnen). Aus Lemma 5.8 folgt

$$\Gamma(x(yz)) = \Gamma(x)\Gamma(yz) = \Gamma(x)(\Gamma(y)\Gamma(z)) = (\Gamma(x)\Gamma(y))\Gamma(z) = \Gamma(xy)\Gamma(z) = \Gamma((xy)z)$$

und $x(yz) = (xy)z$, da Γ injektiv ist. Analog beweist man das Kommutativ- und Distributivgesetz. Wegen $\Gamma(1, 0) = 1_2$ ist $(1, 0)$ das Einselement in \mathbb{C} . Es bleibt zu zeigen, dass $z \neq 0$ invertierbar ist. Es gilt $\det(\Gamma(z)) = a^2 + b^2 > 0$ und

$$\Gamma(z)^{-1} \stackrel{9.23}{=} \frac{1}{\det(\Gamma(z))} \widetilde{\Gamma(z)} = \frac{1}{a^2 + b^2} \begin{pmatrix} a & b \\ -b & a \end{pmatrix} = \Gamma\left(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2}\right). \quad \square$$

⁷Man kann die Axiome auch direkt mit der Definition nachrechnen.

Definition 11.25. Man nennt \mathbb{C} den Körper der *komplexen Zahlen*.

- Mittels der Abbildung $\mathbb{R} \rightarrow \mathbb{C}$, $a \mapsto (a, 0)$ werden wir \mathbb{R} als Teilmenge von \mathbb{C} auffassen. Die Verknüpfungen in \mathbb{R} entsprechen genau denen in \mathbb{C} mit den gleichen neutralen Elementen.
- Man nennt $i := (0, 1) \in \mathbb{C} \setminus \mathbb{R}$ die *imaginäre Einheit*. Es gilt $i^2 = (-1, 0) = -1$. Da $1, i$ eine Basis von \mathbb{C} bilden, lässt sich jede komplexe Zahl eindeutig in der Form $z = a + bi$ schreiben, wobei $\operatorname{Re}(z) := a \in \mathbb{R}$ der *Realteil* und $\operatorname{Im}(z) := b \in \mathbb{R}$ der *Imaginärteil* von z ist.
- Man nennt $|z| := \sqrt{\operatorname{Re}(z)^2 + \operatorname{Im}(z)^2} \geq 0$ den *Betrag* von z (dies entspricht der Norm in \mathbb{R}^2).

Bemerkung 11.26. Im Gegensatz zu \mathbb{R} ist \mathbb{C} kein *angeordneter* Körper, d. h. es gibt keine Ordnungsrelation \leq auf \mathbb{C} mit

$$\begin{aligned} a \leq b &\implies a + c \leq b + c, \\ a, b \geq 0 &\implies ab \geq 0 \end{aligned}$$

für alle $a, b, c \in \mathbb{C}$. Denn angenommen es gilt $i > 0$. Dann wäre $-1 = i^2 > 0$ und $1 = (-1)^2 > 0$. Nun erhält man den Widerspruch $0 = 1 - 1 > 0 + 0 = 0$. Ist hingegen $i < 0$, so wäre $0 = i - i < -i$ und $-1 = (-i)^2 > 0$. Dies führt zum selben Widerspruch.

Lemma 11.27. Für $z \in \mathbb{C} \setminus \{0\}$ und $n \in \mathbb{N}$ existieren paarweise verschiedene n -te Wurzeln $\zeta_1, \dots, \zeta_n \in \mathbb{C}$ mit $\zeta_1^n = \dots = \zeta_n^n = z$.

Beweis. Sei $z_0 := \frac{1}{|z|}z \in \mathbb{C}$ und $\Gamma(z_0) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \in \mathbb{R}^{2 \times 2}$ wie im Beweis von Lemma 11.24. Wegen $a^2 + b^2 = |z_0|^2 = 1$ ist $\Gamma(z_0) = D(\varphi) \in O(2, \mathbb{R})$ für einen Winkel φ mit $a = \cos \varphi$ und $b = \sin \varphi$. Wir definieren $\varphi_k := \frac{\varphi + 2k\pi}{n}$ für $k = 1, \dots, n$. Dann gilt

$$D(\varphi_k)^n = D(n\varphi_k) = D(\varphi + 2k\pi) = D(\varphi) = \Gamma(z_0).$$

Die Zahlen $z_k := \cos \varphi_k + i \sin \varphi_k \in \mathbb{C}$ erfüllen also $z_k^n = z_0$. Wegen $|z| > 0$ existiert $\sqrt[n]{|z|} \in \mathbb{R}_{>0}$ (Analysis). Für $\zeta_k := \sqrt[n]{|z|}z_k$ erhält man $\zeta_k^n = |z|z_0 = z$ für $k = 1, \dots, n$. Sei nun $\zeta_k = \zeta_l$ für $1 \leq k < l \leq n$. Dann unterscheiden sich φ_k und φ_l nur um ein Vielfaches von 2π . Dies zeigt $2\pi(l - k) = 2\pi cn$ für ein $c \in \mathbb{N}_0$. Aus $0 \leq l - k < n$ folgt $k = l$. Daher sind die n -ten Wurzeln ζ_1, \dots, ζ_n paarweise verschieden. \square

Beispiel 11.28. Die (n -ten) Wurzeln aus 1 nennt man *Einheitswurzeln*. Sie entsprechen Drehungen um $2\pi k/n$ mit $k \in \mathbb{Z}$. Die vierten Einheitswurzeln sind $1, i, -1, -i$.

Folgerung 11.29. Sei $A \in \mathbb{C}^{n \times n}$ und $k \in \mathbb{N}$ mit $A^k = A$. Dann ist A diagonalisierbar.

Beweis. Das Minimalpolynom μ_A teilt $X^k - X = X(X^{k-1} - 1)$ nach Lemma 10.44. Die Nullstellen von $X^k - X$ sind die $(k-1)$ -ten Einheitswurzeln und 0. Nach Lemma 10.24 zerfällt μ_A in paarweise verschiedene Linearfaktoren. Die Behauptung folgt nun aus Satz 10.52. \square

Beispiel 11.30. Sei $k \in \mathbb{N}$ und $A := D(2\pi/k)$. Wegen $A^{k+1} = AA^k = AD(2\pi) = A$ ist A über \mathbb{C} diagonalisierbar, aber nicht unbedingt über \mathbb{R} .

Definition 11.31. Die Abbildung $\mathbb{C} \rightarrow \mathbb{C}$, $a + bi \mapsto a - bi =: \overline{a + bi}$ heißt *komplexe Konjugation*.

Lemma 11.32. Für $z, w \in \mathbb{C}$ gilt $|z|^2 = z\bar{z}$, $\overline{z+w} = \bar{z} + \bar{w}$ und $\overline{z\bar{w}} = \bar{z} \cdot \bar{w}$.

Beweis. Für $z = a + bi$ und $w = c + di$ gilt

$$\begin{aligned} |z|^2 &= a^2 + b^2 = (a + bi)(a - bi) = z\bar{z}, \\ \overline{z+w} &= a + c - (b + d)i = a - bi + c - di = \bar{z} + \bar{w}, \\ \overline{z\bar{w}} &= ac - bd - (ad + bc)i = (a - bi)(c - di) = \bar{z} \cdot \bar{w}. \end{aligned} \quad \square$$

Bemerkung 11.33.

(a) Für Matrizen $A = (a_{ij}) \in \mathbb{C}^{n \times m}$ definieren wir $\bar{A} := (\bar{a}_{ij}) \in \mathbb{C}^{n \times m}$. Für $B = (b_{ij}) \in \mathbb{C}^{m \times k}$ gilt

$$\overline{AB} = \left(\sum_{l=1}^m a_{il} b_{lj} \right)_{ij} \stackrel{11.32}{=} \left(\sum_{l=1}^m \overline{a_{il} b_{lj}} \right)_{ij} \stackrel{11.32}{=} \left(\sum_{l=1}^m \bar{a}_{il} \bar{b}_{lj} \right)_{ij} = \bar{A} \cdot \bar{B}.$$

(b) Nach Lemma 11.27 hat das Polynom $X^n - z$ für jedes $z \in \mathbb{C}$ eine Nullstelle (insbesondere ist i eine Nullstelle von $X^2 + 1$). Erstaunlicherweise besitzt sogar jedes nicht-konstante Polynom in $\mathbb{C}[X]$ eine Nullstelle.⁸

Satz 11.34 (Fundamentalsatz der Algebra). Jedes Polynom $\alpha \in \mathbb{C}[X] \setminus \mathbb{C}$ besitzt eine Nullstelle in \mathbb{C} .

Beweis. Der Beweis benutzt Analysis (genauer die Vollständigkeit von \mathbb{R}) und ist für diese Vorlesung zu schwierig. Siehe Algebra-Skript. \square

Folgerung 11.35. Jedes normierte Polynom $\alpha \in \mathbb{C}[X] \setminus \mathbb{C}$ zerfällt in Linearfaktoren.

Beweis. Induktion nach $d := \deg(\alpha) \geq 1$. Im Fall $d = 1$ ist α selbst ein Linearfaktor, da α normiert ist. Sei nun $d \geq 2$. Nach dem Fundamentalsatz besitzt α eine Nullstelle $x \in \mathbb{C}$. Nach Lemma 10.22 existiert ein $\beta \in \mathbb{C}[X]$ mit $\alpha = (X - x)\beta$ und $\deg(\beta) = d - 1$. Da α normiert ist, muss auch β normiert sein. Nach Induktion zerfällt β in Linearfaktoren und damit auch α . \square

Beispiel 11.36. Sei $\alpha \in \mathbb{R}[X]$ mit ungeradem Grad. Um zu zeigen, dass α eine Nullstelle hat, können wir annehmen, dass α normiert ist. Dann gilt $\lim_{x \rightarrow \pm\infty} \alpha(x) = \pm\infty$. Nach dem Zwischenwertsatz der Analysis besitzt die stetige Funktion $\mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto \alpha(x)$ eine reelle Nullstelle. In der Praxis lässt sich eine solche Nullstelle allerdings nur näherungsweise berechnen. Sei konkret $\alpha := X^5 - 4X + 2$. Anhand des Graphen von α vermuten wir eine Nullstelle in der Nähe von $x_0 := 0.5$. Sei $\alpha' = 5X^4 - 4$ die Ableitung von α (Analysis). Beim *Newton-Verfahren* berechnet man die rekursive Folge

$$x_{n+1} := x_n - \frac{\alpha(x_n)}{\alpha'(x_n)} \quad (n \geq 0),$$

also $x_1 = 0.50847\dots$, $x_2 = 0.50849948\dots$ usw. Ist x_0 (so wie hier) „gut“ gewählt, so konvergiert die Folge $(x_n)_n$ quadratisch gegen eine Nullstelle, d. h. mit jeder Iteration verdoppelt sich die Anzahl der korrekten Dezimalstellen. Tatsächlich sind bereits alle angegebenen Dezimalstellen von x_2 korrekt.

⁸Man sagt: \mathbb{C} ist *algebraisch abgeschlossen*.

Bemerkung 11.37. Sei $x \in \mathbb{C}$ eine Nullstelle von $\alpha = \sum a_k X^k \in \mathbb{C}[X]$. Wir definieren $\bar{\alpha} := \sum \bar{a}_k X^k \in \mathbb{C}[X]$. Es gilt dann

$$\bar{\alpha}(\bar{x}) = \sum \overline{a_k x^k} = \overline{\alpha(x)} = \bar{0} = 0,$$

d. h. \bar{x} ist eine Nullstelle von $\bar{\alpha}$. Im Fall $\alpha \in \mathbb{R}[X]$ gilt also: $\alpha(x) = 0 \iff \alpha(\bar{x}) = 0$. Ggf. ist

$$(X - x)(X - \bar{x}) = X^2 - (x + \bar{x})X + x\bar{x} = X^2 - 2\operatorname{Re}(x)X + |x|^2 \in \mathbb{R}[X]$$

ein Teiler von α .

11.5 Der Hauptsatz

Satz 11.38 (Hauptsatz). *Sei V ein euklidischer Raum und $f \in \operatorname{End}(V)$. Genau dann ist f symmetrisch, wenn V eine Orthonormalbasis aus Eigenvektoren von f besitzt. Insbesondere sind symmetrische Endomorphismen diagonalisierbar.*

Beweis. Sei B eine Orthonormalbasis aus Eigenvektoren von f . Dann ist ${}_B[f]_B$ eine Diagonalmatrix und daher symmetrisch. Nach Lemma 11.20 ist f symmetrisch. Sei nun umgekehrt f symmetrisch. Wir argumentieren durch Induktion nach $n := \dim V$. Im Fall $n = 1$ kann jeder normierte Vektor für eine Orthonormalbasis aus Eigenvektoren verwendet werden. Sei also $n \geq 2$ und die Behauptung für $n - 1$ bereits bewiesen. Sei zunächst B eine beliebige Orthonormalbasis von V . Dann können wir die symmetrische Matrix $A := {}_B[f]_B$ auch als komplexe Matrix auffassen. Nach dem Fundamentalsatz der Algebra besitzt $\chi_A \in \mathbb{C}[X] \setminus \mathbb{C}$ eine Nullstelle $\lambda \in \mathbb{C}$. Also ist λ ein Eigenwert von A . Sei $v = (v_1, \dots, v_n)^t \in \mathbb{C}^{n \times 1}$ ein entsprechender Eigenvektor. Wegen

$$\lambda \sum_{i=1}^n |v_i|^2 = \lambda \bar{v}^t v = \bar{v}^t A v = (\bar{v}^t A v)^t = v^t A^t \bar{v} = v^t A \bar{v} = v^t \overline{A v} = \bar{\lambda} v^t \bar{v} = \bar{\lambda} \sum_{i=1}^n |v_i|^2$$

ist $\lambda = \bar{\lambda} \in \mathbb{R}$. Nun ist λ auch ein Eigenwert von f und wir können einen entsprechenden Eigenvektor $b_1 \in V$ wählen. Nach Normierung ist $|b_1| = 1$. Für $U := b_1^\perp$ gilt $V = \langle b_1 \rangle \oplus U$ nach Lemma 11.15. Für $u \in U$ gilt

$$[f(u), b_1] = [u, f(b_1)] = \lambda [u, b_1] = 0,$$

d. h. $f(u) \in U$. Daher liegt die Einschränkung $g := f|_U$ in $\operatorname{End}(U)$. Offenbar ist g auch symmetrisch. Wegen $\dim U = n - 1$ besitzt U nach Induktion eine Orthonormalbasis $b_2, \dots, b_n \in U$ aus Eigenvektoren von g . Wegen $f(b_i) = g(b_i)$ sind b_2, \dots, b_n auch Eigenvektoren von f . Insgesamt ist b_1, \dots, b_n eine Orthonormalbasis von V aus Eigenvektoren von f . \square

Folgerung 11.39. *Genau dann ist $A \in \mathbb{R}^{n \times n}$ symmetrisch, wenn eine Matrix $S \in O(n, \mathbb{R})$ mit $S^t A S = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ existiert. Ggf. sind $\lambda_1, \dots, \lambda_n$ die Eigenwerte von A .*

Beweis. Ist $D := S^t A S = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, so gilt

$$A^t = (S D S^t)^t = S D^t S^t = S D S^t = A,$$

d. h. A ist symmetrisch. Für die Umkehrung sei $V := \mathbb{R}^n$ und $f \in \operatorname{End}(V)$ mit $[f] = A$. Sei A symmetrisch. Nach Lemma 11.20 ist f symmetrisch. Nach dem Hauptsatz existiert eine Orthonormalbasis B von V bestehend aus Eigenvektoren von f . Sei E die Standardbasis von V und $S := {}_E \Delta_B \in \operatorname{GL}(n, \mathbb{R})$. Da die Spalten von S aus den Vektoren von B bestehen, gilt $S \in O(n, \mathbb{R})$. Aus Folgerung 7.27 folgt

$$S^t A S = S^{-1} A S = {}_B[f]_B = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$$

mit den Eigenwerten $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ von f . \square

Bemerkung 11.40. Sei $f \in \text{End}(V)$ symmetrisch. Seien $v, w \in V$ Eigenvektoren von f zu verschiedenen Eigenwerten λ bzw. μ . Dann gilt $\lambda[v, w] = [f(v), w] = [v, f(w)] = \mu[v, w]$ und es folgt $[v, w] = 0$. Merkregel: Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal. Man kann die gesuchte Orthonormalbasis von V also berechnen, indem man das Gram-Schmidt-Verfahren auf jeden Eigenraum anwendet.

Beispiel 11.41. In Beispiel 8.4 hatten wir die Abbildung $f \in \text{End}(\mathbb{R}^3)$ mit symmetrischer Matrix

$$A := [f] = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

untersucht. Es gilt

$$E_1(A) = \left\langle \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right\rangle, \quad E_4(A) = \left\langle \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\rangle$$

(beachte: $\text{tr}(A) = 6 = 1 + 1 + 4$). Das Gram-Schmidt-Verfahren für $E_1(A)$ liefert

$$b_1 := (1, 0, -1), \quad b_2 := (0, 1, -1) - \frac{1}{2}(1, 0, -1) = \frac{1}{2}(-1, 2, -1).$$

Der Eigenvektor zum Eigenwert 4 muss nur normiert werden. Insgesamt erhält man die Orthonormalbasis $\frac{1}{\sqrt{2}}(1, 0, -1), \frac{1}{\sqrt{6}}(-1, 2, -1), \frac{1}{\sqrt{3}}(1, 1, 1)$ aus Eigenvektoren von f . Für die orthogonale Matrix

$$S := \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \\ -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \end{pmatrix}$$

gilt $S^t A S = S^{-1} A S = \text{diag}(1, 1, 4)$. Man kann das Ergebnis nutzen, um Wurzeln von Matrizen zu ziehen. Für $\sqrt{A} := S \text{diag}(1, 1, 2) S^t$ gilt nämlich

$$\sqrt{A}^2 = S \text{diag}(1, 1, 2) S^t S \text{diag}(1, 1, 2) S^t = S \text{diag}(1, 1, 2)^2 S^t = S \text{diag}(1, 1, 4) S^t = A.$$

Mehr dazu in Satz 12.46 und Satz 14.42.

Satz 11.42 (EULER). Sei V ein 3-dimensionaler euklidischer Raum und $f \in O(V)$. Dann existiert eine Orthonormalbasis B von V und ein Winkel φ mit

$${}_B[f]_B = \begin{pmatrix} \pm 1 & 0 \\ 0 & D(\varphi) \end{pmatrix}.$$

Beweis. Da 3 ungerade ist, besitzt χ_f eine Nullstelle $\lambda \in \mathbb{R}$ nach Beispiel 11.36 (oder Bemerkung 11.37). Nach Bemerkung 11.21 ist $\lambda \in \{\pm 1\}$. Sei b_1 ein entsprechender normierter Eigenvektor und $U := b_1^\perp$. Für $u \in U$ gilt

$$[f(u), b_1] = [f(u), \lambda^2 b_1] = \lambda [f(u), f(b_1)] = \lambda [u, b_1] = 0$$

und $f(u) \in U$. Daher liegt die Einschränkung $g := f|_U$ in $O(U)$. Für eine Orthonormalbasis $C := \{b_2, b_3\}$ von U ist ${}_C[g]_C \in O(2, \mathbb{R})$ nach Lemma 11.20. Im Fall $\det(g) = 1$ ist ${}_C[g]_C = D(\varphi)$ nach Beispiel 11.22 und die Behauptung folgt mit $B := \{b_1, b_2, b_3\}$. Sei nun $\det(g) = -1$. Nach Beispiel 11.22 ist g eine Spiegelung mit Eigenwerten 1 und -1 . Da dies auch Eigenwerte von f sind, können wir λ durch $-\lambda$ ersetzen und b_1 entsprechend wählen. Anschließend ist $\det(g) = 1$ und die Behauptung folgt wie zuvor. \square

Bemerkung 11.43. Die Matrizen in Satz 11.42 mit Determinante 1 (also von der Form $\text{diag}(1, D(\varphi))$) entsprechen Drehungen um den Winkel φ , wobei die Drehachse vom ersten Basisvektor aufgespannt wird. Nach dem Determinantensatz ist die Komposition von Drehungen wieder eine Drehung (im 2-dimensionalen Raum ist dies klar wegen $D(\varphi)D(\psi) = D(\varphi + \psi)$). Achtung: Die orthogonalen Abbildungen mit Determinante -1 sind nicht unbedingt Spiegelungen. Es gibt auch sogenannte *Drehspiegelungen*, d. h. Kompositionen einer Drehung mit einer Spiegelung.

Beispiel 11.44. Während eines Fußballspiels gibt es einen Punkt auf der Oberfläche des Fußballs, der sich zu zwei verschiedenen Zeitpunkten exakt am gleichen Ort befindet. Begründung: Der Mittelpunkt des Fußballs liegt zu Beginn der ersten und zweiten Halbzeit auf dem Anstoßpunkt. Dazwischen führt der Ball mit jedem Schuss eine Drehung (und Translation) aus. Da die Komposition von Drehungen wieder eine Drehung ist, wird auch die Transformation auf dem Anstoßpunkt durch eine Drehung f beschrieben. Die Drehachse von f schneidet die Oberfläche des Balls an zwei Punkten. Diese bleiben also fest.

Definition 11.45. Sei V ein euklidischer Raum und $v \in V \setminus \{0\}$. Man nennt

$$S_v: V \rightarrow V, \quad w \mapsto w - 2 \frac{[w, v]}{[v, v]} v$$

die *Spiegelung* an v^\perp .

Bemerkung 11.46. Für $\lambda \in \mathbb{R}^\times$ ist $S_{\lambda v} = S_v$. Wir können daher annehmen, dass v normiert ist. Dann gilt

$$S_v(w) = w - 2[w, v]v$$

für alle $w \in V$. Offenbar ist S_v linear und $S_v(w) = w$ gilt genau dann, wenn $w \in v^\perp$. Außerdem ist $S_v(v) = -v$. Geometrisch entspricht S_v also einer Spiegelung an der Hyperebene v^\perp . Bzgl. einer geeigneten Basis hat S_v die Darstellungsmatrix $\text{diag}(-1, 1, \dots, 1)$ (vgl. Aufgabe II.11). Insbesondere ist S_v orthogonal und $\det S_v = -1$. Außerdem gilt $S_v \circ S_v = \text{id}_V$.

Satz 11.47 (CARTAN-DIEUDONNÉ). Sei V ein n -dimensionaler euklidischer Raum und $f \in \text{O}(V)$. Dann ist f eine Komposition von höchstens n Spiegelungen.

Beweis. Im Fall $f = \text{id}_V$ ist f die Komposition von 0 Spiegelungen. Sei daher $f \neq \text{id}_V$ und $w \in V$ mit $f(w) \neq w$. Im Fall $n = 1$ ist $f = -\text{id}_V = S_1$. Sei nun $n \geq 2$ und die Behauptung für $n - 1$ bereits beweisen. Für $v := f(w) - w \neq 0$ gilt

$$[f(w) + w, v] = [f(w) + w, f(w) - w] = [f(w), f(w)] - [w, w] = 0,$$

d. h. $f(w) + w \in v^\perp$. Also ist

$$(S_v \circ f)(w) = \frac{1}{2} \left(S_v(f(w) + w) + S_v(v) \right) = \frac{1}{2} (f(w) + w - v) = w.$$

Sei $U := w^\perp$. Wegen $g := S_v \circ f \in \text{O}(V)$ ist $g(U) = U$. Nach Induktion ist $g|_U$ eine Komposition von höchstens $n - 1$ Spiegelungen S_{w_1}, \dots, S_{w_k} mit $w_1, \dots, w_k \in U$. Man kann S_{w_i} mit der gleichen Formel auch als Spiegelung von V auffassen. Wegen $w \in U^\perp$ gilt $S_{w_i}(w) = w$ für $i = 1, \dots, k$. Damit ist $f = S_v \circ g$ ein Produkt von höchstens n Spiegelungen. \square

Bemerkung 11.48. Nach Bemerkung 11.46 ist $f \in \text{SO}(V)$ ein Produkt einer geraden Anzahl an Spiegelungen. Insbesondere benötigt man im Fall $\dim V = 2n + 1$ nur $2n$ Spiegelungen.

12 Bilinearformen

12.1 Gram-Matrizen

Bemerkung 12.1. Wir verallgemeinern in diesem Kapitel Teile der euklidischen Geometrie auf beliebige Körper. Stets sei V ein endlich-dimensionaler K -Vektorraum.

Definition 12.2.

- Eine *Bilinearform* auf V ist eine Abbildung $\beta: V \times V \rightarrow K$, die in der ersten und zweiten Komponente linear ist, d. h.

$$\begin{aligned}\beta(\lambda u + v, w) &= \lambda\beta(u, w) + \beta(v, w), \\ \beta(u, \lambda v + w) &= \lambda\beta(u, v) + \beta(u, w)\end{aligned}$$

für alle $u, v, w \in V$ und $\lambda \in K$.

- Man nennt β
 - *symmetrisch*, falls $\beta(v, w) = \beta(w, v)$ für alle $v, w \in V$ gilt.
 - *antisymmetrisch*, falls $\beta(v, w) = -\beta(w, v)$ für alle $v, w \in V$ gilt.
 - *alternierend*, falls $\beta(v, v) = 0$ für alle $v \in V$ gilt.
 - *ausgeartet*, falls $\exists v \in V \setminus \{0\} : \forall w \in V : \beta(v, w) = 0$.

Beispiel 12.3.

- Die triviale Bilinearform $\beta(v, w) = 0$ für alle $v, w \in V$ ist symmetrisch, antisymmetrisch und alternierend. Für $V \neq \{0\}$ ist sie ausgeartet. In der Regel interessieren wir uns für nicht-ausgeartete Bilinearformen.
- Ein Skalarprodukt $\beta(v, w) = [v, w]$ auf einem euklidischen Raum V ist eine symmetrische Bilinearform. Für $V \neq \{0\}$ ist β nicht-ausgeartet, denn $[v, v] = |v|^2 > 0$ für $v \neq 0$.
- Für $V = K^n$ und $A \in K^{n \times n}$ definiert $\beta_A(v, w) := vAw^t$ eine Bilinearform. Dies folgt aus den Rechenregeln für Matrizen. Wir werden in Satz 12.9 sehen, dass jede Bilinearform diese Form hat.
- Die Wahl $A = 1_n$ in (c) führt zur symmetrischen Bilinearform

$$\beta(v, w) = vw^t = \sum_{i=1}^n v_i w_i.$$

Im Fall $K = \mathbb{R}$ erhält man das Standardskalarprodukt.

- Sei $V := K^2$ und $\beta(v, w) := \det\left(\begin{smallmatrix} v \\ w \end{smallmatrix}\right)$ für $v, w \in V$. Nach Lemma 9.5 ist β eine alternierende Bilinearform. Im Allgemeinen ist \det eine *Multilinearform*.

Bemerkung 12.4.

(a) Jede alternierende Bilinearform β ist antisymmetrisch, denn aus

$$0 = \beta(v + w, v + w) = \beta(v, v) + \beta(v, w) + \beta(w, v) + \beta(w, w) = \beta(v, w) + \beta(w, v)$$

folgt $\beta(v, w) = -\beta(w, v)$. Für $K \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$ gilt auch die Umkehrung, denn aus $\beta(v, v) = -\beta(v, v)$ folgt $\beta(v, v) = 0$. Für $K = \mathbb{F}_2$ ist diese Schlussweise ungültig, denn $1 + 1 = 0$. Hier ist symmetrisch und antisymmetrisch sogar gleichbedeutend. Insbesondere ist die Bilinearform $\beta(v, w) = vw^t$ auf $V = K^n$ (anti)symmetrisch, aber nicht alternierend. Um diesen Fall zu vermeiden, werden wir im Folgenden oft $1 + 1 \neq 0$ annehmen.¹

(b) Jede symmetrische Bilinearform β definiert eine *quadratische Form* $q: V \rightarrow K$ durch $q(v) := \beta(v, v)$. Ist $1 + 1 \neq 0$, so kann β aus q durch *Polarisierung* zurückgewinnen:

$$\beta(v, w) = \frac{1}{2} \left(q(v + w) - q(v) - q(w) \right) \quad (v, w \in V).$$

Satz 12.5. Die Menge $\text{Bil}(V)$ aller Bilinearformen auf V ist Unterraum von $\text{Abb}(V \times V, K)$.

Beweis. Die Nullabbildung ist die triviale Bilinearform. Sei $\beta, \gamma \in \text{Bil}(V)$ und $\lambda \in K$. Wie in Satz 7.13 zeigt man, dass $\beta + \gamma$ und $\lambda\beta$ in der ersten und zweiten Komponenten linear sind (allerdings ist $\text{Bil}(V) \not\subseteq \text{Hom}(V \times V, K)$). Dies zeigt $\beta + \gamma, \lambda\beta \in \text{Bil}(V)$. \square

Bemerkung 12.6. Offenbar bilden die symmetrischen (antisymmetrischen, alternierenden) Bilinearformen jeweils einen Unterraum von $\text{Bil}(V)$. Satz 12.9 gibt Aufschluss über die Dimension dieser Unterräume.

Satz 12.7. Sei $\beta \in \text{Bil}(V)$. Für $v \in V$ sei $F_v: V \rightarrow K, w \mapsto \beta(v, w)$. Dann ist $F: V \rightarrow V^*, v \mapsto F_v$ ein Homomorphismus. Genau dann ist β nicht-ausgeartet, wenn F ein Isomorphismus ist.

Beweis. Wegen der Bilinearität von β ist $F_v \in V^*$ und F ein Homomorphismus. Offenbar ist β genau dann nicht-ausgeartet, wenn F injektiv ist. Wegen $\dim V^* = \dim V$ ist injektiv äquivalent zu bijektiv. \square

Definition 12.8. Sei β eine Bilinearform auf V . Sei $B := \{b_1, \dots, b_n\}$ eine Basis von V . Man nennt

$${}_B[\beta]_B := \left(\beta(b_i, b_j) \right)_{i,j=1}^n \in K^{n \times n}$$

die *Gram-Matrix* von β bzgl. B . Ist $V = K^n$ und B die Standardbasis, so sei $[\beta] := {}_B[\beta]_B$.

Satz 12.9. Sei B eine Basis von V mit $|B| = n$. Dann ist die Abbildung

$${}_B[\cdot]_B: \text{Bil}(V) \rightarrow K^{n \times n}, \quad \beta \mapsto {}_B[\beta]_B$$

ein Isomorphismus. Für $A := {}_B[\beta]_B$ gilt:

(a) β symmetrisch $\iff A$ symmetrisch.

(b) β antisymmetrisch $\iff A^t = -A$.²

¹Neben \mathbb{F}_2 gibt es viele (auch unendliche) Körper mit $1 + 1 = 0$.

²Man nennt A *antisymmetrisch* oder *schiefsymmetrisch*.

(c) β nicht-ausgeartet $\iff A$ invertierbar.

Beweis. Für $\beta, \gamma \in \text{Bil}(V)$ und $\lambda \in K$ gilt

$${}_B[\lambda\beta + \gamma]_B = ((\lambda\beta + \gamma)(b_i, b_j)) = \lambda(\beta(b_i, b_j)) + (\gamma(b_i, b_j)) = \lambda{}_B[\beta]_B + {}_B[\gamma]_B,$$

d. h. ${}_B[\cdot]_B$ ist linear. Ist ${}_B[\beta]_B = 0$, so gilt

$$\beta\left(\sum \lambda_i b_i, \sum \mu_j b_j\right) = \sum_{i,j} \lambda_i \mu_j \beta(b_i, b_j) = 0$$

für alle $\lambda_i, \mu_j \in K$. Also ist $\beta = 0$ und ${}_B[\cdot]_B$ ist injektiv. Für die Surjektivität sei $A \in K^{n \times n}$ gegeben. Aus der Linearität der Koordinatendarstellung $v \mapsto {}_B[v]$ (siehe Beweis von Satz 7.10) folgt, dass

$$\beta(v, w) := {}_B[v] A {}_B[w]^t$$

eine Bilinearform definiert. Wegen $\beta(b_i, b_j) = e_i A e_j^t = a_{ij}$ ist ${}_B[\beta]_B = A$. Also ist ${}_B[\cdot]_B$ ein Isomorphismus.

- (a) Ist β symmetrisch, so gilt $a_{ij} = \beta(b_i, b_j) = \beta(b_j, b_i) = a_{ji}$ für alle $1 \leq i, j \leq n$. Daher ist A symmetrisch. Sei umgekehrt A symmetrisch. Dann ist $\beta(b_i, b_j) = a_{ij} = a_{ji} = \beta(b_j, b_i)$. Für ${}_B[v] = (v_1, \dots, v_n)$ und ${}_B[w] = (w_1, \dots, w_n)$ folgt

$$\beta(v, w) = \sum_{1 \leq i, j \leq n} v_i w_j \beta(b_i, b_j) = \sum_{1 \leq i, j \leq n} v_i w_j \beta(b_j, b_i) = \beta(w, v),$$

d. h. β ist symmetrisch.

- (b) Ist β antisymmetrisch, so gilt $a_{ij} = \beta(b_i, b_j) = -\beta(b_j, b_i) = -a_{ji}$ und $A^t = -A$. Die Umkehrung zeigt man wie in (a).

- (c) Sei β ausgeartet und $v \neq 0$ mit $\beta(v, w) = 0$ für alle $w \in W$. Für $x := {}_B[v]^t$ und $i = 1, \dots, n$ gilt dann

$$e_i A x = \sum_{j=1}^n \beta(b_i, b_j) x_j = \beta(b_i, v) = 0.$$

Dies zeigt $Ax = 0$. Nach Satz 6.6 ist A nicht invertierbar. Ist umgekehrt A nicht invertierbar, so existiert $x \in K^{n \times 1} \setminus \{0\}$ mit $Ax = 0$. Für $v = \sum_{i=1}^n x_i b_i$ gilt dann $\beta(v, w) = 0$ für alle $w \in V$. Daher ist β ausgeartet. \square

Beispiel 12.10. Sei $1 + 1 \neq 0$ in K und $n := \dim V$ ungerade. Sei $\beta \in \text{Bil}(V)$ antisymmetrisch mit Gram-Matrix A bzgl. einer beliebigen Basis. Nach Satz 12.9 gilt

$$\det(A) = \det(A^t) = \det(-A) \stackrel{9.8}{=} (-1)^n \det(A) = -\det(A)$$

und $\det(A) = 0$. Daher muss β ausgeartet sein.

Satz 12.11. Seien B und C Basen von V . Für alle $\beta \in \text{Bil}(V)$ gilt

$$\boxed{{}_C[\beta]_C = {}_B \Delta_C^t {}_B[\beta]_B \Delta_C.}$$

Beweis. Sei $B = \{b_1, \dots, b_n\}$, $C = \{c_1, \dots, c_n\}$ und $S = (s_{ij}) = {}_B\Delta_C$. Für $1 \leq i, j \leq n$ gilt

$$\beta(c_i, c_j) = \sum_{k,l=1}^n s_{ki}\beta(b_k, b_l)s_{lj} = \sum_{k=1}^n s_{ki}({}_B[\beta]_BS)_{kj} = (S^t{}_B[\beta]_BS)_{ij}. \quad \square$$

Definition 12.12. Matrizen $A, B \in K^{n \times n}$ heißen *kongruent*, falls ein $S \in \text{GL}(n, K)$ mit $B = SAS^t$ existiert.³

Bemerkung 12.13.

- (a) Wie in Bemerkung 7.30 zeigt man, dass die Kongruenz von Matrizen eine Äquivalenzrelation ist. Nach Satz 12.9 und Satz 12.11 bestimmt jede Bilinearform auf V eine Kongruenzklasse von Gram-Matrizen. Wir zeigen im nächsten Abschnitt, dass man Bilinearformen ähnlich wie Endomorphismen diagonalisieren kann.
- (b) Nach Lemma 5.15 ist

$$\text{rk}(B) = \text{rk}(SAS^t) \leq \min\{\text{rk}(S), \text{rk}(AS^t)\} = \text{rk}(AS^t) \leq \min\{\text{rk}(A), \text{rk}(S^t)\} = \text{rk}(A).$$

Aus Symmetriegründen folgt $\text{rk}(B) = \text{rk}(A)$, d. h. kongruente Matrizen haben den gleichen Rang. Andererseits haben A und B nicht unbedingt die gleiche Determinante, denn $\det(B) = \det(SAS^t) = \det(S)^2 \det(A)$. Wegen Bemerkung 10.35 haben A und B in der Regel auch nicht die gleichen Eigenwerte. Bereits im Fall $n = 1$ sieht man, dass A und B ebenso nicht die gleiche Spur haben.

- (c) Nach dem Hauptachsensatz ist jede symmetrische reelle Matrix kongruent und ähnlich zu einer Diagonalmatrix.

12.2 Sylvesters Trägheitssatz

Definition 12.14. Sei $\beta \in \text{Bil}(V)$ und $S \subseteq V$. Wie in euklidischen Räumen definieren wir das *orthogonale Komplement* von S durch

$$S^\perp := \{v \in V : \forall s \in S : \beta(v, s) = 0\}.$$

Bemerkung 12.15.

- (a) In der Definition von S^\perp müsste man streng genommen zwischen „links-orthogonal“ und „rechts-orthogonal“ ($\forall s \in S : \beta(s, v) = 0$) unterscheiden. In der Regel werden wir annehmen, dass β (anti)symmetrisch ist, sodass beide Versionen äquivalent sind.
- (b) Für $U \leq V$ ist $U^\perp \leq V$. Genau dann ist β ausgeartet, falls $V^\perp \neq \{0\}$ gilt.

Satz 12.16. Sei $\beta \in \text{Bil}(V)$ nicht-ausgeartet. Für $U \leq V$ gilt $\dim V = \dim U + \dim U^\perp$.

Beweis. Sei $F: V \rightarrow V^*$, $v \mapsto F_v$ der Isomorphismus aus Satz 12.7. Dann gilt

$$F_v \in F(U^\perp) \iff \forall u \in U : F_v(u) = \beta(v, u) = 0 \iff F_v \in U^0,$$

d. h. $F(U^\perp) = U^0$. Die Behauptung folgt aus Lemma 7.43. □

³Im Gegensatz zu Ähnlichkeit führen wir kein Symbol für die Kongruenz von Matrizen ein.

Bemerkung 12.17. Achtung: Im Gegensatz zu euklidischen Räumen gilt im Allgemeinen weder $U \cap U^\perp = \{0\}$ noch $V = U \oplus U^\perp$. Sei zum Beispiel $V = \mathbb{F}_2^2$ und $[\beta] = 1_2$. Für $U = \langle (1, 1) \rangle$ gilt $U = U^\perp$.

Definition 12.18. Sei $\beta \in \text{Bil}(V)$ symmetrisch. Eine Basis $B = \{b_1, \dots, b_n\}$ von V heißt *Orthogonalbasis* bzgl. β , falls $\beta(b_i, b_j) = 0$ für $i \neq j$ gilt. Ist zusätzlich $\beta(b_i, b_i) = 1$, so nennt man B eine *Orthonormalbasis* bzgl. β .

Bemerkung 12.19.

- (a) Offenbar ist B genau dann *Orthogonalbasis* (bzw. *Orthonormalbasis*) bzgl. β , falls ${}_B[\beta]_B$ eine Diagonalmatrix (bzw. ${}_B[\beta]_B = 1_n$) ist.
- (b) Nach Folgerung 11.11 besitzt jedes Skalarprodukt auf einem euklidischen Raum eine Orthonormalbasis.

Satz 12.20. Sei $1+1 \neq 0$ in K . Dann besitzt jede symmetrische Bilinearform auf V eine Orthogonalbasis.

Beweis. Sei $\beta \in \text{Bil}(V)$ symmetrisch. Wir argumentieren durch Induktion nach $n := \dim V$. Im Fall $n \leq 1$ oder $\beta = 0$ ist jede Basis eine Orthogonalbasis. Sei also $n \geq 2$ und $\beta(v, w) \neq 0$ für gewisse $v, w \in V$. Im Fall $\beta(v, v) = \beta(w, w) = 0$ ist

$$\beta(v+w, v+w) = \beta(v, v) + \beta(v, w) + \beta(w, v) + \beta(w, w) = 2\beta(v, w) \neq 0$$

(beachte $1+1 \neq 0$). In jedem Fall existiert also ein $b_1 \in V \setminus \{0\}$ mit $\beta(b_1, b_1) \neq 0$. Sei $U := \langle b_1 \rangle$. Wegen $U \cap U^\perp = \{0\}$ ist $V = U \oplus U^\perp$. Die Einschränkung von β auf $U \times U$ ist eine symmetrische Bilinearform. Nach Induktion besitzt U^\perp eine Orthogonalbasis b_2, \dots, b_n . Nun ist b_1, \dots, b_n eine Orthogonalbasis von V . □

Bemerkung 12.21.

- (a) Die Matrix-Version von Satz 12.20 lautet: Jede symmetrische Matrix ist zu einer Diagonalmatrix kongruent (falls $1+1 \neq 0$). Wir werden sehen, dass man die Diagonaleinträge speziell wählen kann.
- (b) Ist $1+1 = 0$ in K , so wird Satz 12.20 falsch: Die Bilinearform β mit $[\beta] = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ besitzt keine Orthogonalbasis, denn $\beta(v, v) = 0$ für alle $v \in K^2$.
- (c) Das Gram-Schmidt-Verfahren zur Berechnung von Orthogonalbasen funktioniert in dieser Allgemeinheit nicht immer, denn man muss durch $\beta(b_i, b_i)$ teilen, aber dieser Wert kann 0 sein. Wir modifizieren stattdessen den Gauß-Algorithmus wie folgt:
 - (1) Sei $A := [\beta]$. Nehmen wir induktiv an, dass die Zeilen und Spalten $1, \dots, k-1$ von A bereits die gewünschte Diagonalgestalt haben (zu Beginn sei $k=1$).
 - (2) Nehmen wir zunächst $a_{kk} \neq 0$ an. Dann kann man $a_{ik} = 0$ für $i = k+1, \dots, n$ erreichen, indem man wie gewohnt ein Vielfaches der k -ten Zeile auf die darunterliegenden Zeilen addiert. Dies entspricht der Multiplikation mit Elementarmatrizen S_1, \dots, S_{n-k} von links. Anschließend setze man $a_{ki} = 0$ für $i = k+1, \dots, n$. Dies entspricht der Multiplikation von S_1^t, \dots, S_{n-k}^t von rechts (in beliebiger Reihenfolge). Insgesamt wird A zu SAS^t , wobei $S = S_1 \dots S_n$. Wegen $(SAS^t)^t = (S^t)^t A^t S^t = SAS^t$ bleibt A durch dieses Vorgehen symmetrisch.

- (3) Sei nun $a_{kk} = 0$. Existiert ein $l > k$ mit $a_{ll} \neq 0$, so tausche man die Zeilen k und l und anschließend die Spalten k und l . Dadurch werden a_{kk} und a_{ll} vertauscht und A bleibt symmetrisch. Man ist nun in Situation (2).
- (4) Sei schließlich $a_{ii} = 0$ für $i = k, \dots, n$. Ist $a_{ij} = 0$ für alle $k \leq i, j \leq n$, so sind wir fertig. Sei daher $a_{ij} \neq 0$ für gewisse $k \leq i < j \leq n$. Wir addieren die i -te Zeile zur j -ten Zeile und anschließend die i -te Spalte zur j -ten Spalte. Der Eintrag an Position (j, j) wird dadurch zu $2a_{ij} \neq 0$ (beachte $1 + 1 \neq 0$). Also ist man in Situation (3).
- (5) Zur Bestimmung der Orthogonalbasis B führe man alle Zeilenoperationen (aber nicht die Spaltenoperationen) an der Einheitsmatrix aus. Dadurch entsteht die Matrix $S \in GL(n, K)$, sodass SAS^t eine Diagonalmatrix ist (vgl. Satz 6.17). Nach Satz 12.11 sind die Zeilen von S die Vektoren von B .

Beispiel 12.22.

$$\begin{aligned}
 (A|1_3) &= \left(\begin{array}{ccc|ccc} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{array} \right) \begin{array}{l} \leftarrow + \\ \leftarrow + \\ \leftarrow + \end{array} \sim \left(\begin{array}{ccc|ccc} 2 & 1 & 2 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 2 & 1 & 1 & 0 & 0 & 1 \end{array} \right) \begin{array}{l} \leftarrow -1/2 \\ \leftarrow + \\ \leftarrow + \end{array} \\
 &\sim \left(\begin{array}{ccc|ccc} 2 & 1 & 2 & 1 & 1 & 0 \\ 0 & -1/2 & 0 & -1/2 & 1/2 & 0 \\ 0 & 0 & -1 & -1 & -1 & 1 \end{array} \right) \sim \left(\begin{array}{ccc|ccc} 2 & 0 & 0 & 1 & 1 & 0 \\ 0 & -1/2 & 0 & -1/2 & 1/2 & 0 \\ 0 & 0 & -1 & -1 & -1 & 1 \end{array} \right) \\
 S &= \begin{pmatrix} 1 & 1 & 0 \\ -1/2 & 1/2 & 0 \\ -1 & -1 & 1 \end{pmatrix}
 \end{aligned}$$

Satz 12.23 (SYLVESTERS Trägheitssatz). Sei β eine symmetrische Bilinearform auf $V = \mathbb{R}^n$. Dann existiert eine Basis B von V mit

$${}_B[\beta]_B = \text{diag}(1_r, -1_s, 0_t).$$

Die Zahlen r, s, t hängen nicht von B ab.

Beweis. Nach Satz 12.20 existiert eine Orthogonalbasis $B = \{b_1, \dots, b_n\}$ von V . Nach Umsortierung können wir

$$\beta(b_i, b_i) \begin{cases} > 0 & \text{für } i = 1, \dots, r, \\ < 0 & \text{für } i = r + 1, \dots, r + s, \\ = 0 & \text{für } i = r + s + 1, \dots, r + s + t = n \end{cases}$$

annehmen. Für $i = 1, \dots, r + s$ ersetzen wir b_i durch $\frac{1}{\sqrt{|\beta(b_i, b_i)|}} b_i$. Dann gilt $\beta(b_i, b_i) = 1$ für $i = 1, \dots, r$ und $\beta(b_i, b_i) = -1$ für $i = r + 1, \dots, r + s$. Damit ist die Existenz von B gezeigt.

Wegen $V^\perp = \langle b_{r+s+1}, \dots, b_n \rangle$ hängt $t = \dim V^\perp$ nicht von B ab. Sei $V_+ := \langle b_1, \dots, b_r \rangle$. Für $v = \sum_{i=1}^r \lambda_i b_i \in V_+$ mit $\lambda_i \in \mathbb{R}$ gilt $\beta(v, v) = \sum_{i=1}^r \lambda_i^2 \geq 0$ mit Gleichheit genau dann, wenn $v = 0$. Sei $B' = \{b'_1, \dots, b'_n\}$ eine weitere Basis von V mit

$${}_{B'}[\beta]_{B'} = \text{diag}(1_{r'}, -1_{s'}, 0_t).$$

Sei $V'_{\leq 0} := \langle b'_{r'+1}, \dots, b'_n \rangle$. Für $v \in V_+ \cap V'_{\leq 0}$ gilt einerseits $\beta(v, v) \geq 0$ und andererseits $\beta(v, v) \leq 0$. Dies zeigt $\beta(v, v) = 0$ und $v = 0$. Also ist $V_+ \cap V'_{\leq 0} = \{0\}$ und

$$r + s' + t = \dim(V_+ \oplus V'_{\leq 0}) \leq \dim V = n = r + s + t.$$

Es folgt $s' \leq s$. Aus Symmetriegründen ist $s' = s$ und $r' = r$. □

Definition 12.24. In der Situation von Satz 12.23 nennt man $\text{ind}(\beta) := (r, s, t)$ den *Index* von β .⁴ Für eine symmetrische Matrix $A \in \mathbb{R}^{n \times n}$ definiert man $\text{ind}(A) := \text{ind}(\beta)$ durch der Bilinearform β mit $[\beta] = A$.

Beispiel 12.25.

- (a) Jedes Skalarprodukt auf \mathbb{R}^n hat Index $(n, 0, 0)$.
- (b) Ist $\text{ind}(\beta) = (r, s, t)$, so ist β genau dann ausgeartet, wenn $t > 0$ gilt.
- (c) In der Relativitätstheorie betrachtet man den *Minkowski-Raum* \mathbb{R}^4 bzgl. einer Bilinearform β mit Index $(3, 1, 0)$. Die vierte Dimension beschreibt dabei die Zeit.

Bemerkung 12.26. Zur Berechnung des Index einer Matrix A kann man zunächst eine Orthogonalbasis \tilde{B} mit dem modifizierten Gauß-Algorithmus aus Bemerkung 12.21 bestimmen (beachte: $1 + 1 \neq 0$ in \mathbb{R}). Anschließend zählt man die positiven und negativen Einträge auf der Hauptdiagonale. Dividiert man die Vektoren in \tilde{B} jeweils durch die Wurzel des entsprechenden Diagonaleintrags, so erhält man die Basis B wie im Beweis von Satz 12.23. Die Matrix aus Beispiel 12.22 hat zum Beispiel Index $(1, -2, 0)$ bzgl.

$$B = \left\{ \frac{1}{\sqrt{2}}(1, 1, 0), \frac{1}{\sqrt{2}}(-1, 1, 0), (-1 - 1, 1) \right\}.$$

Satz 12.27. Sei $A \in \mathbb{R}^{n \times n}$ symmetrisch mit Index (r, s, t) . Dann ist r die Anzahl der positiven Eigenwerte und s die Anzahl der negativen Eigenwerte von A jeweils gezählt mit (algebraischen) Vielfachheiten.

Beweis. Nach der Matrix-Version des Hauptsatzes ist A zu $\text{diag}(\lambda_1, \dots, \lambda_n)$ kongruent, wobei $\lambda_1, \dots, \lambda_n$ die Eigenwerte von A sind (die algebraische Vielfachheit jedes Eigenwerts stimmt mit der geometrischen Vielfachheit überein). Man kann das Argument aus dem Beweis von Satz 12.23 auf diese Matrix anwenden. Die positiven λ_i werden dabei zu 1 transformiert und die negativen zu -1 . □

Folgerung 12.28. Für symmetrische Matrizen $A, B \in \mathbb{R}^{n \times n}$ sind die folgenden Aussagen äquivalent:

- (1) A und B sind kongruent.
- (2) $\text{ind}(A) = \text{ind}(B)$.
- (3) A und B haben die gleiche Anzahl an positiven Eigenwerten und die gleiche Anzahl an negativen Eigenwerten.⁵

⁴Dieser Begriff ist in der Literatur nicht einheitlich! Manche Autoren nennen $r - s$ die *Signatur* von β . Aus Dimension, Rang und Signatur lässt sich der Index bestimmen.

⁵Das beschreibt die *Trägheit* in Sylvesters Satz.

Beweis. Da kongruente Matrizen die gleiche Bilinearform beschreiben, gilt (1) \Rightarrow (2). Aus Satz 12.27 folgt (2) \Rightarrow (3). Gilt (3), so haben A und B den gleichen Index nach Satz 12.27. Also sind A und B kongruent, d. h. (1) gilt. \square

Bemerkung 12.29.

- (a) Für jedes $m \leq n$ gibt es genau $m + 1$ Kongruenzklassen von symmetrischen $n \times n$ -Matrizen mit Rang m (nämlich mit Index $(i, m - i, n - m)$ für $i = 0, \dots, m$). Daher gibt es

$$\sum_{m=0}^n (m + 1) = \sum_{m=1}^{n+1} m = \frac{(n + 1)(n + 2)}{2}$$

symmetrische Bilinearformen auf \mathbb{R}^n bis auf Basiswahl (vgl. Beispiel 1.16).

- (b) Arbeitet man über \mathbb{C} anstatt \mathbb{R} , so kann man stets ${}_B[\beta]_B = \text{diag}(1_r, 0_t)$ erreichen. Man kann nämlich jeden Basisvektor b mit $\beta(b, b) < 0$ durch $\frac{i}{\sqrt{|\beta(b, b)|}}b$ ersetzen. Daher gibt es nur $n + 1$ symmetrische Bilinearformen auf \mathbb{C}^n bis auf Basiswahl. Wir beweisen in Satz 13.25 Sylvesters Trägheitssatz über \mathbb{C} mit einem anderen Kongruenzbegriff.

Satz 12.30. Sei V ein beliebiger K -Vektorraum. Sei $\beta \in \text{Bil}(V)$ alternierend und nicht-ausgeartet. Dann existiert eine Basis B von V mit ${}_B[\beta]_B = \begin{pmatrix} 0_n & 1_n \\ -1_n & 0_n \end{pmatrix}$.

Beweis. Induktion nach $\dim V$.⁶ Sei $b_1 \in V \setminus \{0\}$. Da β nicht-ausgeartet ist, existiert $c_1 \in V$ mit $\beta(b_1, c_1) \neq 0$. Indem man c_1 durch $\beta(b_1, c_1)^{-1}c_1$ ersetzt, erreicht man $\beta(b_1, c_1) = 1$. Sei $U := \langle b_1, c_1 \rangle$. Für $v = \lambda b_1 + \mu c_1 \in U$ gilt $\beta(v, b_1) = \mu\beta(c_1, b_1) = -\mu$ und $\beta(v, c_1) = \lambda$. Dies zeigt $U \cap U^\perp = 0$. Nach Satz 12.16 gilt $V = U \oplus U^\perp$. Für $u \in U^\perp$ existiert ein $v \in V$ mit $\beta(u, v) \neq 0$. Schreibt man $v = v_1 + v_2$ mit $v_1 \in U$ und $v_2 \in U^\perp$, so folgt $\beta(u, v_2) = \beta(u, v) \neq 0$. Daher ist auch die Einschränkung β' von β auf $U^\perp \times U^\perp$ nicht-ausgeartet und alternierend. Nach Induktion besitzt U^\perp eine Basis $B' = \{b_2, \dots, b_n, c_2, \dots, c_n\}$ mit

$${}_{B'}[\beta']_{B'} = \begin{pmatrix} 0_{n-1} & 1_{n-1} \\ -1_{n-1} & 0_{n-1} \end{pmatrix}.$$

Nun ist $B = \{b_1, \dots, b_n, c_1, \dots, c_n\}$ eine Basis mit der gewünschten Eigenschaft. \square

Bemerkung 12.31. Ein Vektorraum V mit einer alternierenden, nicht-ausgearteten Bilinearform nennt man einem *symplektischen Raum*. Darauf gehen wir nicht weiter ein.

12.3 Positiv definite Matrizen

Bemerkung 12.32. In diesem Abschnitt verallgemeinern wir die positive Definitheit von Skalarprodukten in euklidischen Räumen. Stets sei V ein \mathbb{R} -Vektorraum.

Definition 12.33. Eine symmetrische Bilinearform $\beta \in \text{Bil}(V)$ nennt man

- *positiv (semi)definit*, falls $\beta(v, v) > 0$ (bzw. $\beta(v, v) \geq 0$) für alle $v \in V \setminus \{0\}$,
- *negativ (semi)definit*, falls $\beta(v, v) < 0$ (bzw. $\beta(v, v) \leq 0$) für alle $v \in V \setminus \{0\}$,

⁶Nach Beispiel 12.10 ist $\dim V$ gerade, aber das wird hier nicht benutzt.

- *indefinit*, falls $\exists v, w \in V : \beta(v, v) > 0 > \beta(w, w)$.

Diese Begriffe übertragen sich auf symmetrische Matrizen A , indem man ein β mit $[\beta] = A$ wählt.

Bemerkung 12.34.

- (a) Offensichtlich ist jede positiv definite Bilinearform auch positiv semidefinit. Ist β positiv (semi)definit, so ist $-\beta$ negativ (semi)definit. Daher kann man sich in Regel auf die positive Eigenschaft beschränken.
- (b) Ist β positiv (semi)definit, so ist die dazu gehörige quadratische Form aus Bemerkung 12.4 positiv (bzw. nicht-negativ).
- (c) Die Summe von positiv (semi)definiten Bilinearformen und Matrizen ist wieder positiv (semi)definit. Man kann positiv semidefinite Matrizen als mehrdimensionale Verallgemeinerung von nicht-negativen reellen Zahlen ansehen (siehe Satz 12.46, Beispiel 18.20 und Aufgabe III.11). Allerdings ist das Produkt von positiv (semi)definiten Matrizen in der Regel nicht symmetrisch und kann nach unserer Definition nicht positiv (semi)definit sein.

Beispiel 12.35.

- (a) Sei

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

Für $x \in \mathbb{R}^n \setminus \{0\}$ gilt

$$xAx^t = 2 \sum_{i=1}^n x_i^2 - 2 \sum_{i=1}^{n-1} x_i x_{i+1} = x_1^2 + \sum_{i=1}^{n-1} (x_i - x_{i+1})^2 + x_n^2.$$

Im Fall $x_1 \neq 0$ oder $x_n \neq 0$ folgt $xAx^t \geq x_1^2 + x_n^2 > 0$. Anderenfalls existiert ein $i \geq 0$ mit $x_i = 0 \neq x_{i+1}$. Dann ist ebenfalls $xAx^t \geq (x_i - x_{i+1})^2 > 0$. Also ist A positiv definit.

- (b) In der Analysis bildet man aus den zweiten partiellen Ableitungen einer Funktion $f: \mathbb{R}^n \rightarrow \mathbb{R}$ die *Hesse-Matrix* $(\frac{\partial^2 f(x)}{\partial x_i \partial x_j}) \in \mathbb{R}^{n \times n}$, welche an einem Punkt $x \in \mathbb{R}^n$ nur dann positiv definit sein kann, wenn x ein lokales Minimum der Funktion ist.

Satz 12.36. Für jede symmetrische Bilinearform $\beta \in \text{Bil}(V)$ mit Index (r, s, t) gilt

- (a) β positiv semidefinit $\iff s = 0$.
- (b) β positiv definit $\iff s = t = 0$.
- (c) β indefinit $\iff s, t > 0$.

Beweis. Sei B die Basis aus Sylvesters Trägheitssatz. Ist $s > 0$, so existiert ein $b \in B$ mit $\beta(b, b) = -1$. Dann kann β nicht positiv semidefinit sein. Ist $t > 0$, so existiert $b \in B$ mit $\beta(b, b) = 0$. Dann kann β nicht positiv definit sein. Ist $s = 0$ (bzw. $s = t = 0$) so gilt $\beta(b, b) \geq 0$ (bzw. $\beta(b, b) > 0$) für alle $b \in B$. Da B eine Orthogonalbasis ist, folgt leicht, dass β positiv (semi)definit ist. Genau dann ist β indefinit, falls $b, c \in B$ mit $\beta(b, b) = 1 = -\beta(c, c)$ existieren. Das bedeutet $s, t > 0$. \square

Folgerung 12.37. Eine symmetrische Matrix $A \in \mathbb{R}^{n \times n}$ ist genau dann positiv (semi)definit, wenn alle Eigenwerte von A positiv (bzw. nicht-negativ) sind.

Beweis. Bekanntlich sind alle Eigenwerte von A reell. Die Behauptung folgt aus Satz 12.27. \square

Beispiel 12.38. Für

$$A = (1 + \delta_{ij}) = \begin{pmatrix} 2 & 1 & \cdots & 1 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & 2 \end{pmatrix} \in \mathbb{R}^{n \times n}$$

funktioniert der Trick aus Beispiel 12.35 nicht. Wegen $\text{rk}(A - 1_n) = 1$ ist 1 ein Eigenwert von A mit (geometrischer) Vielfachheit $n - 1$ (vgl. Aufgabe I.26). Der fehlende Eigenwert muss $\text{tr}(A) - (n - 1) = 2n - n + 1 = n + 1$ sein. Daher ist A positiv definit.

Bemerkung 12.39. Da sich die Eigenwerte in der Praxis nur schwer berechnen lassen, ist es nützlich andere Kriterien für die positive Definitheit zu kennen. Eine notwendige (aber nicht hinreichende) Bedingung ist, dass alle Diagonaleinträge positiv sind, denn $a_{ii} = e_i A e_i^t$.

Lemma 12.40. Genau dann ist $A \in \mathbb{R}^{n \times n}$ positiv semidefinit, wenn eine Matrix $S \in \mathbb{R}^{n \times n}$ mit $A = SS^t$ existiert. Genau dann ist A positiv definit, wenn S invertierbar ist.

Beweis. Sei $A = SS^t$ für ein $S \in \mathbb{R}^{n \times n}$. Für $v \in \mathbb{R}^n$ gilt $vAv^t = vS(vS)^t = |vS|^2 \geq 0$. Daher ist A positiv semidefinit. Ist S invertierbar, so gilt $|vS| > 0$ für alle $v \neq 0$. Dann ist A positiv definit. Sei umgekehrt A positiv semidefinit. Nach Sylvesters Trägheitssatz und Satz 12.36 existiert ein $U \in \text{GL}(n, \mathbb{R})$ mit $A = UDU^t$, wobei $D = \text{diag}(1_r, 0_t)$. Für $S = UD$ gilt $SS^t = A$ wegen $D^2 = D$. Ist A positiv definit, so ist $S = U$ invertierbar. \square

Satz 12.41 (SYLVESTER-Kriterium). Sei $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ symmetrisch und $A_k := (a_{ij})_{i,j=1}^k$ für $1 \leq k \leq n$. Genau dann ist A positiv definit, wenn $\det(A_k) > 0$ für $k = 1, \dots, n$ gilt.

Beweis. Induktion nach n : Im Fall $n = 1$ ist A genau dann positiv definit, falls $\det(A) = \det(A_1) = a_{11} > 0$. Sei nun $n \geq 2$ und $1 \leq k \leq n$. Sei A positiv definit und $v \in \mathbb{R}^k \setminus \{0\}$. Für $w := (v_1, \dots, v_k, 0, \dots, 0) \in \mathbb{R}^n$ gilt $vA_k v^t = wAw^t > 0$. Also ist A_k positiv definit. Nach Lemma 12.40 existiert ein $S \in \text{GL}(k, \mathbb{R})$ mit $A_k = SS^t$. Es folgt $\det(A_k) = \det(S)^2 > 0$.

Nehmen wir umgekehrt $\det(A_k) > 0$ für $k = 1, \dots, n$ an. Nach Induktion ist A_{n-1} positiv definit. Sei $\beta \in \text{Bil}(\mathbb{R}^n)$ mit $[\beta] = A$. Dann ist die Einschränkung β_1 von β auf $\mathbb{R}^{n-1} \times \mathbb{R}^{n-1}$ positiv definit, denn $[\beta_1] = A_{n-1}$. Sei $b'_1, \dots, b'_{n-1} \in \mathbb{R}^{n-1}$ eine Orthonormalbasis von β_1 . Durch Anfügen einer 0 erhält man die Vektoren $b_i := (b'_i, 0) \in \mathbb{R}^n$ mit

$$\beta(b_i, b_j) = b_i A b_j^t = b'_i A_{n-1} b'_j = \delta_{ij}$$

für $i = 1, \dots, n - 1$. Sei $V_1 := \langle b_1, \dots, b_{n-1} \rangle$. Da β auf V_1 positiv definit ist, gilt $V_1 \cap V_1^\perp = \{0\}$. Mit $b_n \in V_1^\perp \setminus \{0\}$ ist $B = \{b_1, \dots, b_n\}$ eine Orthogonalbasis von \mathbb{R}^n und

$$D := B[\beta]_B = \text{diag}(1, \dots, 1, \lambda)$$

für ein $\lambda \in \mathbb{R}$. Da A und D kongruent sind, existiert ein $S \in \text{GL}(n, \mathbb{R})$ mit $D = SAS^t$. Es folgt $\lambda = \det(D) = \det(S)^2 \det(A) > 0$. Mit D ist nun auch A positiv definit. \square

Beispiel 12.42.

(a) Die Matrix $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ ist genau dann positiv definit, wenn $a > 0$ und $ac > b^2$.

(b) Sei

$$A := \begin{pmatrix} 2 & -1 & \cdot & \cdot & -1 \\ -1 & 2 & -1 & \cdot & \cdot \\ \cdot & -1 & 2 & -1 & \cdot \\ \cdot & \cdot & -1 & 2 & -1 \\ -1 & \cdot & \cdot & -1 & 2 \end{pmatrix}$$

Nach Beispiel 12.35 gilt $\det(A_k) > 0$ für $k = 1, \dots, 4$. Da die Zeilensummen von A verschwinden, ist $(1, \dots, 1)^t$ ein Eigenvektor zum Eigenwert 0. Dies zeigt $\det(A) = 0$. Der Beweis von Satz 12.41 zeigt, dass A positiv semidefinit ist, aber nicht positiv definit.

Bemerkung 12.43.

(a) Man nennt $\det(A_k)$ die *Hauptminoren* von A .

(b) Achtung: Aus $\det(A_k) \geq 0$ für $k = 1, \dots, n$ folgt nicht, dass A positiv semidefinit ist (betrachte zum Beispiel $A = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$).

(c) Ebenso ist die negative Definitheit nicht zu $\det(A_k) < 0$ äquivalent. Bekanntlich ist A genau dann negativ definit, wenn $-A$ positiv definit ist. Dies ist äquivalent zu $(-1)^k \det(A_k) > 0$ für $k = 1, \dots, n$.

Satz 12.44. Eine symmetrische Matrix $A \in \mathbb{R}^{n \times n}$ mit charakteristischem Polynom

$$\chi_A = X^n + a_1 X^{n-1} + \dots + a_n \in \mathbb{R}[X]$$

ist genau dann positiv definit, falls die Koeffizienten von χ_A alternierende Vorzeichen haben, d. h. $a_i(-1)^i > 0$ für $i = 1, \dots, n$.

Beweis. Sei $a_i(-1)^i > 0$ für $i = 1, \dots, n$. Für $\lambda < 0$ gilt dann

$$\chi_A(\lambda) = (-1)^n (|\lambda|^n + |a_1||\lambda|^{n-1} + \dots + |a_n|) \neq 0.$$

Also besitzt A nur positive Eigenwerte. Nach Folgerung 12.37 ist A positiv definit.

Sei umgekehrt A positiv definit mit Eigenwerten $\lambda_1, \dots, \lambda_n > 0$. Dann gilt

$$\chi_A = \prod_{i=1}^n (X - \lambda_i) = X^n - \left(\sum_{i=1}^n \lambda_i \right) X^{n-1} + \left(\sum_{i < j} \lambda_i \lambda_j \right) X^{n-2} + \dots + (-1)^n \lambda_1 \dots \lambda_n$$

und

$$a_k(-1)^k = \sum_{i_1 < \dots < i_k} \lambda_{i_1} \dots \lambda_{i_k} > 0$$

für $k = 1, \dots, n$. □

Beispiel 12.45. Für

$$A = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 1 \end{pmatrix}$$

gilt

$$\chi_A = (X - 2)^2(X - 1) + 2 - 2(X - 2) - (X - 1) = X^3 - 5X^2 + 5X + 3.$$

Da die letzten beiden Koeffizienten das gleiche Vorzeichen haben, kann A nicht positiv definit sein.

Satz 12.46. Sei $A \in \mathbb{R}^{n \times n}$ positiv (semi)definit und $k \in \mathbb{N}$. Dann besitzt A genau eine positiv (semi)definite k -te Wurzel $W \in \mathbb{R}^{n \times n}$, d. h. $W^k = A$.

Beweis. Nach dem Hauptsatz existiert ein $S \in O(n, \mathbb{R})$ mit $S^t A S = \text{diag}(\lambda_1, \dots, \lambda_n)$. Da A positiv semidefinit ist, gilt $\lambda_1, \dots, \lambda_n \geq 0$ nach Folgerung 12.37. Nun ist $W := S \text{diag}(\sqrt[k]{\lambda_1}, \dots, \sqrt[k]{\lambda_n}) S^t$ positiv (semi)definit mit $W^k = A$. Sei $\alpha \in \mathbb{R}[X]$ mit $\alpha(\lambda_i) = \sqrt[k]{\lambda_i}$ für $i = 1, \dots, n$ (Interpolation). Wegen $(S^t A S)^i = S^t A^i S$ für $i \in \mathbb{N}$ gilt

$$\alpha(A) = S \alpha(S^t A S) S^t = W.$$

Sei auch $B \in \mathbb{R}^{n \times n}$ positiv (semi)definit mit $B^k = A$. Der Hauptsatz liefert ein $T \in O(n, \mathbb{R})$ mit $T^t B T = \text{diag}(\mu_1, \dots, \mu_n)$ und $\mu_1, \dots, \mu_n \geq 0$. Dabei sind μ_1^k, \dots, μ_n^k die Eigenwerte von $B^k = A$. Also existiert eine Permutationsmatrix P mit $P^t T^t B T P = S^t W S$ und $P^t T^t A T P = S^t A S$. Daher ist $T P S^t$ mit A und mit $\alpha(A) = W$ vertauschbar. Dies zeigt $B = T P S^t W S P^t T^t = W$. \square

13 Unitäre Räume

13.1 Sesquilinearformen

Bemerkung 13.1. Wir entwickeln in diesem Kapitel eine Geometrie für die komplexen Zahlen. Anstelle des euklidischen Skalarprodukts tritt eine „schiefsymmetrische“ Abbildung, die nur noch in der ersten Koordinate linear ist. Das Gegenstück des Hauptsatzesatzes ist der Spektralsatz.

Definition 13.2. Ein *Skalarprodukt* auf einem \mathbb{C} -Vektorraum V ist eine Abbildung $V \times V \rightarrow \mathbb{C}$ mit folgenden Eigenschaften ($u, v, w \in V, \lambda \in \mathbb{C}$):

- $[v, v] \in \mathbb{R}$ und $[v, v] \geq 0$ mit Gleichheit genau dann, wenn $v = 0$ (*positiv definit*),
- $[v, w] = \overline{[w, v]}$ (*schiefsymmetrisch*),
- $[\lambda u + v, w] = \lambda[u, w] + [v, w]$.

Zusammen mit einem Skalarprodukt wird V ein *unitärer Raum*. Vektoren $v, w \in V$ heißen *orthogonal*, falls $[v, w] = 0$. Man nennt $|v| := \sqrt{[v, v]} \geq 0$ die *Norm* von v . Im Fall $|v| = 1$ nennt man v *normiert*.

Beispiel 13.3. Das *Standardskalarprodukt* auf $V = \mathbb{C}^n$ ist definiert durch

$$[v, w] := v\bar{w}^t = \sum_{i=1}^n v_i \bar{w}_i$$

für $v, w \in V$. Man prüft leicht, dass die Eigenschaften des Skalarprodukts erfüllt sind.

Bemerkung 13.4.

- (a) Für $v, w \in V$ und $\lambda \in \mathbb{C}$ gilt

$$[u, \lambda v + w] = \overline{[\lambda v + w, u]} = \overline{\lambda[v, u] + [w, u]} = \bar{\lambda}[u, v] + [u, w],$$

d. h. das Skalarprodukt ist in der zweiten Komponente nicht linear. Man spricht daher von einer *Sesquilinearform*.¹

- (b) Die meisten der in Abschnitt 11.1 bewiesenen Sätze hängen nicht wesentlich von der Bilinearität des Skalarprodukts ab und können daher problemlos übertragen werden.

Lemma 13.5. Sei V ein unitärer Raum, $v, w \in V$ und $\lambda \in \mathbb{C}$. Dann gilt

- (a) $|\lambda v| = |\lambda||v|$ (Homogenität).
(b) $|[v, w]| \leq |v||w|$ mit Gleichheit genau dann, wenn v und w linear abhängig sind (CAUCHY-SCHWARZ-Ungleichung).

¹sesqui ist lateinisch für eineinhalb.

(c) $\left| |v| - |w| \right| \leq |v + w| \leq |v| + |w|$ (Dreiecksungleichung).

Beweis.

(a) $|\lambda v| = \sqrt{[\lambda v, \lambda v]} = \sqrt{\lambda \bar{\lambda} [v, v]} = \sqrt{|\lambda|^2} \sqrt{[v, v]} = |\lambda| |v|.$

(b) O. B. d. A. sei $w \neq 0$ und $\lambda := \frac{[v, w]}{[w, w]}$. Wegen $\bar{\lambda} = \frac{[w, v]}{[w, w]}$ gilt

$$0 \leq |v - \lambda w|^2 = [v - \lambda w, v - \lambda w] = [v, v] - \lambda [w, v] - \bar{\lambda} [v, w] + |\lambda|^2 [w, w] = |v|^2 - \frac{|[v, w]|^2}{|w|^2}.$$

Es folgt $|[v, w]|^2 \leq |v|^2 |w|^2$ und $|[v, w]| \leq |v| |w|$. Gleichheit impliziert $v = \lambda w$, d. h. v und w sind linear abhängig. Sind umgekehrt v und w linear abhängig gegeben, dann existiert ein $\mu \in \mathbb{C}$ mit $v = \mu w$ und $|[v, w]| = |\mu| |w|^2 \stackrel{(a)}{=} |\mu w| |w| = |v| |w|.$

(c) Sei $[v, w] = a + bi$. Dann gilt

$$[v, w] + [w, v] = [v, w] + \overline{[v, w]} = 2a \leq 2\sqrt{a^2 + b^2} = 2|[v, w]|.$$

Es folgt

$$|v + w|^2 = [v + w, v + w] \leq [v, v] + 2|[v, w]| + [w, w] \stackrel{(b)}{\leq} |v|^2 + 2|v| |w| + |w|^2 = (|v| + |w|)^2$$

und $|v + w| \leq |v| + |w|$. Also ist $|v| = |v + w - w| \leq |v + w| + |w|$ und $|v| - |w| \leq |v + w|$. Vertauschen von v und w liefert $-(|v| - |w|) = |w| - |v| \leq |v + w|$, d. h. $||v| - |w|| \leq |v + w|$. \square

Bemerkung 13.6. Sei V ein unitärer Raum.

(a) Eine Basis b_1, \dots, b_n von V heißt *Orthonormalbasis*, falls $[b_i, b_j] = \delta_{ij}$ für $1 \leq i, j \leq n$ gilt. Das Gram-Schmidt-Verfahren zur Berechnung einer Orthonormalbasis funktioniert unverändert für unitäre Räume. Insbesondere besitzt jeder unitäre Raum eine Orthonormalbasis.

(b) Für $U \leq V$ definiert man wie üblich $U^\perp := \{v \in V : \forall u \in U : [v, u] = 0\} \leq V$. Die Regeln aus Lemma 11.15

- $V = U \oplus U^\perp,$
- $(U^\perp)^\perp = U,$
- $U \subseteq W \iff W^\perp \subseteq U^\perp,$

gelten auch für unitäre Räume.

13.2 Adjungierte Abbildungen

Satz 13.7. Sei V ein unitärer Raum und $f \in \text{End}(V)$. Dann existiert genau ein $f^* \in \text{End}(V)$ mit $[f(v), w] = [v, f^*(w)]$ für alle $v, w \in V$.

Beweis. Sei b_1, \dots, b_n eine Orthonormalbasis von V . Wir definieren $f^* \in \text{End}(V)$ durch

$$f^*(b_j) := \sum_{i=1}^n [f(b_i), b_j] b_i$$

für $j = 1, \dots, n$. Für $v = \sum \lambda_i b_i$ und $w = \sum \mu_j b_j$ gilt

$$[f(v), w] = \sum_{i,j=1}^n \lambda_i \overline{\mu_j} [f(b_i), b_j] = \sum_{i,j=1}^n \lambda_i \overline{\mu_j} [b_i, f^*(b_j)] = [v, f^*(w)].$$

Sei auch $f_1 \in \text{End}(V)$ mit $[f(v), w] = [v, f_1(w)]$ für alle $v, w \in V$. Dann ist $[v, f^*(w) - f_1(w)] = 0$. Für $v := f^*(w) - f_1(w)$ ergibt sich $f^*(w) = f_1(w)$ für alle $w \in V$. Dies zeigt $f_1 = f^*$. \square

Definition 13.8. In der Situation von Satz 13.7 nennt man f^* die zu f adjungierte Abbildung.

Bemerkung 13.9.

- (a) Für $v, w \in V$ gilt $[v, f(w)] = \overline{[f(w), v]} = \overline{[w, f^*(v)]} = [f^*(v), w]$. Insbesondere ist $(f^*)^* = f$.
- (b) Für $f, g \in \text{End}(V)$ und $\lambda \in \mathbb{C}$ gilt $(\lambda f + g)^* = \overline{\lambda} f^* + g^*$.
- (c) Achtung: Wir benutzen für die adjungierte Abbildung von f die gleiche Bezeichnung wie für die zu f duale Abbildung. Die duale Abbildung liegt jedoch in $\text{End}(V^*)$.

Definition 13.10. Sei V ein unitärer Raum. Man nennt $f \in \text{End}(V)$

- *hermitesch*, falls $f = f^*$, d. h. $[f(v), w] = [v, f(w)]$ für alle $v, w \in V$.
- *unitär*, falls $[f(v), f(w)] = [v, w]$ für alle $v, w \in V$.
- *normal*, falls $f \circ f^* = f^* \circ f$.

Bemerkung 13.11.

- (a) Sei $f \in \text{End}(V)$ hermitesch. Sei $\lambda \in \mathbb{C}$ ein Eigenwert von f mit Eigenvektor $v \in V$. Dann gilt

$$\lambda |v|^2 = [\lambda v, v] = [f(v), v] = [v, f(v)] = [v, \lambda v] = \overline{\lambda} |v|^2$$

und $\lambda = \overline{\lambda} \in \mathbb{R}$.

- (b) Sei $f \in \text{End}(V)$ unitär. Für $v \in \text{Ker}(f)$ ist $|v|^2 = [v, v] = [f(v), f(v)] = 0$, d. h. $v = 0$. Daher ist f ein Isomorphismus. Wegen $[f(v), w] = [v, f^{-1}(w)]$ für alle $v, w \in V$ folgt außerdem $f^* = f^{-1}$. Sind $f, g \in \text{End}(V)$ unitär, so auch $f \circ g$ und f^{-1} . Daher bilden die unitären Abbildungen eine Untergruppe $U(V)$ von $GL(V)$. Man nennt $U(V)$ die *unitäre Gruppe* vom Grad n .
- (c) Sei λ ein Eigenwert einer unitären Abbildung f mit Eigenvektor $v \in V$. Dann gilt $|\lambda|^2 |v|^2 = [\lambda v, \lambda v] = [f(v), f(v)] = [v, v] = |v|^2$ und $|\lambda| = 1$.
- (d) Hermitesche und unitäre Abbildungen sind offenbar normal.

Lemma 13.12. Sei V unitär mit Orthonormalbasis B und $f \in \text{End}(V)$. Dann gilt $B[f^*]_B = \overline{B[f]_B}^t$.

Beweis. Sei $B = \{b_1, \dots, b_n\}$. Sei $f(b_i) = \sum_{j=1}^n a_{ji} b_j$ und $f^*(b_i) = \sum_{j=1}^n a_{ji}^* b_j$ für $i = 1, \dots, n$. Die Behauptung folgt aus

$$\overline{a_{ji}^*} = [b_j, f^*(b_i)] = [f(b_j), b_i] = a_{ij}. \quad \square$$

Definition 13.13. Für $A \in \mathbb{C}^{n \times m}$ sei $A^* := \overline{A}^t = \overline{A}^t$ die zu A adjungierte Matrix. Für $A \in \text{GL}(n, \mathbb{C})$ benutzen wir die Abkürzung $A^{-*} := (A^{-1})^* = (A^*)^{-1}$.

Folgerung 13.14. Sei V unitär mit Orthonormalbasis B , $f \in \text{End}(V)$ und $A := {}_B[f]_B$. Dann gilt

- (a) f hermitesch $\iff A^* = A \iff A^t = \overline{A}$.
- (b) f unitär $\iff A^* = A^{-1} \iff A^t = \overline{A}^{-1}$.
- (c) f normal $\iff A^*A = AA^* \iff A^t\overline{A} = \overline{A}A^t$.

Beweis. Es gilt

$$f \text{ hermitesch} \iff f^* = f \stackrel{13.12}{\iff} A^* = A \iff \overline{A^*} = \overline{A} \iff A^t = \overline{A}.$$

Die anderen Aussagen beweist man analog. □

Definition 13.15. Eine Matrix $A \in \mathbb{C}^{n \times n}$ heißt

- hermitesch, falls $A^* = A$.
- unitär, falls $A^* = A^{-1}$.
- normal, falls $AA^* = A^*A$.

Beispiel 13.16.

- (a) Hermitesche und unitäre Matrizen sind normal. Diagonalmatrizen sind normal, aber nicht unbedingt hermitesch oder unitär.
- (b) Reelle Matrizen sind genau dann hermitesch (bzw. unitär), wenn sie symmetrisch (bzw. orthogonal) sind.
- (c) Invertieren, Transponieren und Komplex-Konjugieren sind miteinander vertauschbare Operatoren auf $\mathbb{C}^{n \times n}$ (vgl. Bemerkung 5.9). Ist A hermitesch (bzw. unitär, normal), so auch \overline{A} , A^t und A^* (nachrechnen).

Bemerkung 13.17. Die unitären Matrizen bilden eine Untergruppe $U(n, \mathbb{C})$ von $\text{GL}(n, \mathbb{C})$, die wie üblich $U(V)$ entspricht. Für $A \in U(n, \mathbb{C})$ gilt $|\det(A)|^2 = \det(A)\det(\overline{A}) = \det(A^t\overline{A}) = 1$. Nach Bemerkung 13.11 haben auch alle Eigenwerte von A den Betrag 1. Man nennt

$$\text{SU}(n, \mathbb{C}) := U(n, \mathbb{C}) \cap \text{SL}(n, \mathbb{C}) \leq U(n, \mathbb{C})$$

die *spezielle unitäre Gruppe* vom Grad n .

13.3 Der Spektralsatz

Satz 13.18 (Spektralsatz²). *Sei V unitär und $f \in \text{End}(V)$. Genau dann ist f normal, wenn V eine Orthonormalbasis aus Eigenvektoren von f besitzt. Insbesondere sind normale Endomorphismen diagonalisierbar.*

Beweis. Sei B eine Orthonormalbasis aus Eigenvektoren von f . Dann ist $A := {}_B[f]_B$ eine Diagonalmatrix. Sicher ist A^* auch eine Diagonalmatrix. Insbesondere sind A und A^* vertauschbar. Nach Folgerung 13.14 ist f normal. Sei umgekehrt f normal. Wir argumentieren durch Induktion nach $n := \dim V$. Im Fall $n = 1$ liefert jeder normierte Vektor eine Orthonormalbasis aus Eigenvektoren. Sei $n \geq 2$. Nach dem Fundamentalsatz der Algebra besitzt f einen Eigenwert $\lambda \in \mathbb{C}$ mit Eigenvektor $b_1 \in V$. O. B. d. A. sei $|b_1| = 1$. Wegen

$$\begin{aligned} |f^*(b_1) - \bar{\lambda}b_1|^2 &= [f^*(b_1), f^*(b_1)] - \lambda[f^*(b_1), b_1] - \bar{\lambda}[b_1, f^*(b_1)] + |\lambda|^2[b_1, b_1] \\ &= [f(f^*(b_1)), b_1] - \lambda[b_1, f(b_1)] - \bar{\lambda}[f(b_1), b_1] + |\lambda|^2 = [f^*(f(b_1)), b_1] - |\lambda|^2 \\ &= \lambda[f^*(b_1), b_1] - |\lambda|^2 = \lambda[b_1, f(b_1)] - |\lambda|^2 = 0 \end{aligned}$$

ist $\bar{\lambda}$ ein Eigenwert von f^* zum Eigenvektor b_1 . Für $U := \langle b_1 \rangle$ gilt $V = U \oplus U^\perp$. Für $u \in U^\perp$ ist $[f(u), b_1] = [u, f^*(b_1)] = \lambda[u, b_1] = 0$ und $f(u) \in U^\perp$. Daher ist die Einschränkung von f auf U^\perp ein normaler Endomorphismus. Nach Induktion existiert eine Orthonormalbasis b_2, \dots, b_n von U^\perp aus Eigenvektoren von f . Nun ist b_1, \dots, b_n eine Orthonormalbasis von V aus Eigenvektoren von f . \square

Bemerkung 13.19. Ist $A \in \mathbb{R}^{n \times n}$ symmetrisch, so liefert der Spektralsatz ein $S \in U(n, \mathbb{C})$ mit $S^*AS = \text{diag}(\lambda_1, \dots, \lambda_n)$. Im Gegensatz zum Hauptachsensatz erhält man nicht, dass S orthogonal (also reell) ist.

Folgerung 13.20. *Sei V unitär und $f \in \text{End}(V)$ normal mit Eigenwerten $\lambda_1, \dots, \lambda_n \in \mathbb{C}$. Dann gilt*

- (a) f hermitesch $\iff \lambda_1, \dots, \lambda_n \in \mathbb{R}$.
- (b) f unitär $\iff |\lambda_1| = \dots = |\lambda_n| = 1$.

Beweis. Nach dem Spektralsatz existiert eine Orthonormalbasis B von V mit $D := {}_B[f]_B = \text{diag}(\lambda_1, \dots, \lambda_n)$. Nach Folgerung 13.14 ist f genau dann hermitesch (bzw. unitär), wenn $\bar{D} = D^t = D$ (bzw. $1_n = D\bar{D} = \text{diag}(|\lambda_1|^2, \dots, |\lambda_n|^2)$) gilt. Dies zeigt die Behauptung. \square

Bemerkung 13.21. Sei $f \in \text{End} V$ normal. Seien $v, w \in V$ Eigenvektoren zu verschiedenen Eigenwerten λ bzw. μ . Wie im Beweis des Spektralsatz ist w ein Eigenvektor von f^* zum Eigenwert $\bar{\mu}$. Aus

$$\lambda[v, w] = [f(v), w] = [v, f^*(w)] = [v, \bar{\mu}w] = \mu[v, w]$$

folgt $[v, w] = 0$. Daher sind Eigenvektoren zu verschiedenen Eigenwerten orthogonal (wie bei symmetrischen Matrizen, siehe Bemerkung 11.40). Die Orthonormalbasis im Spektralsatz lässt sich also wie beim Hauptachsensatz mit dem Gram-Schmidt-Verfahren bestimmen.

²Die Menge der Eigenwerte von $f \in \text{End}(V)$ nennt man das *Spektrum* von f .

Beispiel 13.22. Nehmen wir an, dass

$$A := \begin{pmatrix} 5i & -4 & 2 \\ 4 & 5i & 2i \\ -2 & 2i & 8i \end{pmatrix}$$

normal ist (anderenfalls wird sich ein Widerspruch ergeben). Man berechnet

$$\chi_A = (X - 5i)^2(X - 8i) - 32i + 8(X - 5i) + 16(X - 8i) = X^3 - 18iX^2 - 81X = X(X - 9i)^2.$$

Wegen

$$A \sim \begin{pmatrix} 5i & -4 & 2 \\ 0 & 9i/5 & 18i/5 \\ 0 & 18i/5 & 36i/5 \end{pmatrix} \sim \begin{pmatrix} 5i & -4 & 2 \\ 0 & 9i/5 & 18i/5 \\ 0 & 0 & 0 \end{pmatrix}$$

ist $(2i, -2, 1)$ ein Eigenvektor zum Eigenwert 0. Nach Normierung sei $b_1 := \frac{1}{3}(2i, -2, 1)$. Der Eigenraum zum Eigenwert $9i$ wird durch $c_2 := (0, 1, 2)$ und $c_3 := (i, 1, 0)$ aufgespannt. Wir stellen fest, dass diese Vektoren in der Tat orthogonal zu b_1 sind. Daher muss A normal sein. Das Gram-Schmidt-Verfahren liefert:

$$\begin{aligned} b_2 &:= \frac{1}{\sqrt{5}}(0, 1, 2), \\ c'_3 &:= c_3 - [c_3, b_2]b_2 = (i, 4/5, -2/5), \\ b_3 &:= \frac{1}{3\sqrt{5}}(5i, 4, -2). \end{aligned}$$

Nun ist $B = \{b_1, b_2, b_3\}$ eine Orthonormalbasis aus Eigenvektoren von A .

Satz 13.23 (SCHUR-Zerlegung). Für $A \in \mathbb{C}^{n \times n}$ existiert ein $S \in U(n, \mathbb{C})$, sodass S^*AS eine Dreiecksmatrix ist. Genau dann ist A normal, wenn S^*AS eine Diagonalmatrix ist.

Beweis. O. B. d. A. sei $n \geq 2$. Sei $\lambda \in \mathbb{C}$ ein Eigenwert von A und $v \in \mathbb{C}^n$ ein normierter Eigenvektor zu λ . Wir ergänzen v mit Gram-Schmidt zu einer Orthonormalbasis von \mathbb{C}^n . Nach Wechsel zu dieser Basis gilt $A = \begin{pmatrix} \lambda & * \\ 0 & A_1 \end{pmatrix}$ mit $A_1 \in \mathbb{C}^{(n-1) \times (n-1)}$. Induktiv existiert ein $S_1 \in U(n-1, \mathbb{C})$, sodass $S_1^*A_1S_1$ eine obere Dreiecksmatrix ist. Jetzt ist $S := \text{diag}(1, S_1) \in U(n, \mathbb{C})$ und S^*AS ist eine obere Dreiecksmatrix. Eine untere Dreiecksmatrix erhält man, indem man A^t in eine obere Dreiecksmatrix transformiert und anschließend transponiert.

Ist A normal, so ist $D := (d_{ij}) = S^*AS$ eine normale obere Dreiecksmatrix. Wegen

$$|d_{ii}|^2 = (DD^*)_{ii} = (D^*D)_{ii} = |d_{1i}|^2 + \dots + |d_{ii}|^2$$

ist $a_{ji} = 0$ für $j = 1, \dots, i-1$ und $i = 1, \dots, n$. Dies zeigt, dass D eine Diagonalmatrix ist. Ist umgekehrt D eine Diagonalmatrix, so ist A nach dem Spektralsatz normal. \square

Beispiel 13.24. Die Matrix

$$A := \begin{pmatrix} -1 & -1 & 0 \\ 2 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix}.$$

hat offenbar den Eigenvektor e_3 zum Eigenwert 1. Übergang zur Orthonormalbasis $\{e_3, e_2, e_1\}$ ergibt

$$A \approx \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 0 & -1 & -1 \end{pmatrix}.$$

Die Matrix $A_1 = \begin{pmatrix} 1 & 2 \\ -1 & -1 \end{pmatrix}$ im Beweis von Satz 13.23 hat den Eigenvektor $(1 + i, -1)$ zum Eigenwert i . Orthogonal dazu steht $(1, 1 - i)$. Nach Normierung erhält man

$$A \approx \frac{1}{3} \begin{pmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 - i & -1 \\ 0 & 1 & 1 + i \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 0 & -1 & -1 \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 + i & 1 \\ 0 & -1 & 1 - i \end{pmatrix} = \begin{pmatrix} 1 & (2+2i)/\sqrt{3} & 2/\sqrt{3} \\ 0 & i & 1 - 2i \\ 0 & 0 & -i \end{pmatrix}.$$

Satz 13.25 (Trägheitssatz für hermitesche Matrizen). *Für jede hermitesche Matrix $A \in \mathbb{C}^{n \times n}$ existiert ein $S \in \text{GL}(n, \mathbb{C})$ mit*

$$S^*AS = \text{diag}(1_r, -1_s, 0_t).$$

Die Zahlen r, s, t sind eindeutig bestimmt.

Beweis. Nach dem Spektralsatz können wir $A = \text{diag}(\lambda_1, \dots, \lambda_n)$ annehmen. Nach Folgerung 13.20 sind die Eigenwerte $\lambda_1, \dots, \lambda_n$ von A reell. Sei o. B. d. A. $\lambda_1, \dots, \lambda_r > 0$, $\lambda_{r+1}, \dots, \lambda_{r+s} < 0$ und $\lambda_{r+s+1} = \dots = \lambda_n = 0$. Die Gleichung gilt nun mit

$$S := \text{diag}(\sqrt{|\lambda_1|}^{-1}, \dots, \sqrt{|\lambda_{r+s}|}^{-1}, 1_t) \in \text{GL}(n, \mathbb{C}).$$

Mit $\text{rk}(A) = r + s$ ist t eindeutig bestimmt. Für alle $v \in \langle e_1, \dots, e_r \rangle \setminus \{0\}$ gilt $\bar{v}Av^t > 0$. Sei umgekehrt $U \leq \mathbb{C}^n$ mit $\bar{u}Au^t > 0$ für alle $u \in U \setminus \{0\}$. Für $u \in U \cap \langle e_{r+1}, \dots, e_n \rangle$ gilt einerseits $\bar{u}Au^t \geq 0$ und andererseits $\bar{u}Au^t \leq 0$. Dies zeigt $u = 0$ und $U \cap \langle e_{r+1}, \dots, e_n \rangle = \{0\}$. Es folgt $\dim U \leq r$. Somit ist r die maximale Dimension eines Unterraums $U \leq \mathbb{C}^n$ mit $\bar{u}Au^t > 0$ für alle $u \in U \setminus \{0\}$. Auf diese Weise ist r eindeutig durch A bestimmt. Nun ist auch $s = n - r - t$ eindeutig bestimmt. \square

Beispiel 13.26. Die Zahlen r, s, t in Satz 13.25 kann man mit dem modifizierten Gauß-Verfahren aus Bemerkung 12.21 bestimmen (beachte: $1 + 1 \neq 0$ in \mathbb{C}). Zum Beispiel:

$$\begin{aligned} (A|1_3) &= \left(\begin{array}{ccc|ccc} 1 & i & 2 & 1 & 0 & 0 \\ -i & 1 & -1 & 0 & 1 & 0 \\ 2 & -1 & 3 & 0 & 0 & 1 \end{array} \right) \sim \left(\begin{array}{ccc|ccc} 1 & i & 2 & 1 & 0 & 0 \\ 0 & 0 & 2i-1 & i & 1 & 0 \\ 0 & -2i-1 & -1 & -2 & 0 & 1 \end{array} \right) \\ &\sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & -2i-1 & -2 & 0 & 1 \\ 0 & 2i-1 & 0 & i & 1 & 0 \end{array} \right) \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & -2 & 0 & 1 \\ 0 & 0 & 5 & 2-3i & 1 & 2i-1 \end{array} \right) \\ &\sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & (2-3i)/\sqrt{5} & 1/\sqrt{5} & (2i-1)/\sqrt{5} \\ 0 & 0 & -1 & -2 & 0 & 1 \end{array} \right) \\ S^* &= \frac{1}{\sqrt{5}} \begin{pmatrix} \sqrt{5} & 0 & 0 \\ 2-3i & 1 & 2i-1 \\ -2\sqrt{5} & 0 & \sqrt{5} \end{pmatrix} \end{aligned}$$

Bemerkung 13.27. Der Satz von Mirsky gibt eine notwendige und hinreichende Bedingung für die Existenz von Matrizen mit vorgegebenen Eigenwerten und Hauptdiagonalelementen (deren Summe muss gleich sein). Für hermitesche Matrizen gelten stärkere Einschränkungen an diese Zahlen.

Satz 13.28 (SCHUR-HORN).

(a) Ist $A \in \mathbb{C}^{n \times n}$ hermitesch mit Hauptdiagonale $d_1 \geq \dots \geq d_n$ und Eigenwerten $\lambda_1 \geq \dots \geq \lambda_n$, so gilt

$$\sum_{i=1}^k d_i \leq \sum_{i=1}^k \lambda_i$$

für $k = 1, \dots, n$ mit Gleichheit im Fall $k = n$.

(b) Sind reelle Zahlen $d_1 \geq \dots \geq d_n$ und $\lambda_1 \geq \dots \geq \lambda_n$ mit $\sum_{i=1}^k d_i \leq \sum_{i=1}^k \lambda_i$ für $k = 1, \dots, n$ und $\sum_{i=1}^n d_i = \sum_{i=1}^n \lambda_i$ gegeben, so existiert eine reelle symmetrische Matrix mit Hauptdiagonale d_1, \dots, d_n und Eigenwerten $\lambda_1, \dots, \lambda_n$.

Beweis (CHAN-LI).

(a) Nach dem Spektralsatz existiert eine unitäre Matrix $S := (s_{ij})$ mit

$$A = (a_{ij}) = S^* \operatorname{diag}(\lambda_1, \dots, \lambda_n) S.$$

Für $1 \leq k \leq n$ gilt

$$\sum_{i=1}^k d_i = \sum_{i=1}^k a_{ii} = \sum_{i=1}^k \sum_{j=1}^n \overline{s_{ji}} s_{ji} \lambda_j = \sum_{j=1}^n \lambda_j \sum_{i=1}^k |s_{ji}|^2. \quad (13.1)$$

Da S unitär ist, gilt $t_j := \sum_{i=1}^k |s_{ji}|^2 \leq 1$ mit Gleichheit, falls $k = n$. Es folgt

$$\begin{aligned} \sum_{j=1}^n t_j &= \sum_{i=1}^k \sum_{j=1}^n |s_{ji}|^2 = k, \\ \sum_{i=1}^k (d_i - \lambda_i) &= \sum_{j=1}^n \lambda_j t_j - \sum_{j=1}^k \lambda_j + \lambda_k \left(k - \sum_{i=1}^n t_i \right) \\ &= \sum_{j=1}^k \underbrace{(\lambda_j - \lambda_k)}_{\geq 0} \underbrace{(t_j - 1)}_{\leq 0} + \sum_{i=k+1}^n t_i \underbrace{(\lambda_i - \lambda_k)}_{\leq 0} \leq 0. \end{aligned}$$

Also ist $\sum_{i=1}^k d_i \leq \sum_{i=1}^k \lambda_i$ und $\sum_{i=1}^n d_i = \operatorname{tr}(A) = \sum_{i=1}^n \lambda_i$.

(b) Induktion nach n : Im Fall $n = 1$ erfüllt $A := (d_1) = (\lambda_1)$ die Behauptung. Sei $n = 2$. Dann gilt $\lambda_1 \geq d_1 \geq d_2 = \lambda_1 + \lambda_2 - d_1 \geq \lambda_2$. Im Fall $\lambda_1 = \lambda_2$ können wir $A := d_1 1_2$ wählen. Im Fall $\lambda_1 > \lambda_2$ ist

$$S := \frac{1}{\sqrt{\lambda_1 - \lambda_2}} \begin{pmatrix} \sqrt{d_1 - \lambda_2} & -\sqrt{\lambda_1 - d_1} \\ \sqrt{\lambda_1 - d_1} & \sqrt{d_1 - \lambda_2} \end{pmatrix} \in O(2, \mathbb{R}).$$

Für $A = (a_{ij}) = S^t \operatorname{diag}(\lambda_1, \lambda_2) S$ gilt

$$a_{11} = \frac{1}{\lambda_1 - \lambda_2} ((d_1 - \lambda_2)\lambda_1 + (\lambda_1 - d_1)\lambda_2) = d_1$$

nach (13.1). Es folgt $a_{22} = \lambda_1 + \lambda_2 - d_1 = d_2$. Sei nun $n \geq 3$ und

$$\lambda'_2 := \lambda_1 + \lambda_2 - d_1 \geq \lambda_2.$$

Nach Induktion existiert $S \in O(2, \mathbb{R})$, sodass $S^t \operatorname{diag}(\lambda_1, \lambda_2) S$ Hauptdiagonale (d_1, λ'_2) hat. Da die Folgen $(\lambda'_2, \lambda_3, \dots, \lambda_n)$ und (d_2, \dots, d_n) die gleichen Voraussetzungen erfüllen, existiert nach

Induktion ein $T \in O(n-1, \mathbb{R})$, sodass $T^t \text{diag}(\lambda'_2, \dots, \lambda_n)T$ Hauptdiagonale d_2, \dots, d_n hat. Für $U := \text{diag}(S, 1_{n-2}) \text{diag}(1_1, T) \in O(n, \mathbb{R})$ hat

$$A := U^t \text{diag}(\lambda_1, \dots, \lambda_n)U = \begin{pmatrix} 1 & 0 \\ 0 & T^t \end{pmatrix} \begin{pmatrix} d_1 & * & & & 0 \\ * & \lambda'_2 & & & \\ & & \lambda_3 & & \\ & & & \ddots & \\ 0 & & & & \lambda_n \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & T \end{pmatrix}$$

Hauptdiagonale (d_1, \dots, d_n) und Eigenwerte $\lambda_1, \dots, \lambda_n$. □

Beispiel 13.29. Wir konstruieren eine symmetrische Matrix mit Eigenwerten 3, 2, 1 und Hauptdiagonale 2, 2, 2. Mit der Bezeichnung aus dem obigen Beweis ist $\lambda'_2 = 3$. Man erhält

$$\begin{aligned} A &= \begin{pmatrix} 1 & \cdot & \cdot \\ \cdot & 1/\sqrt{2} & 1/\sqrt{2} \\ \cdot & -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} \cdot & 1 & \cdot \\ -1 & \cdot & \cdot \\ \cdot & \cdot & 1 \end{pmatrix} \begin{pmatrix} 3 & \cdot & \cdot \\ \cdot & 2 & \cdot \\ \cdot & \cdot & 1 \end{pmatrix} \begin{pmatrix} \cdot & -1 & \cdot \\ 1 & \cdot & \cdot \\ \cdot & \cdot & 1 \end{pmatrix} \begin{pmatrix} 1 & \cdot & \cdot \\ \cdot & 1/\sqrt{2} & -1/\sqrt{2} \\ \cdot & 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} \sqrt{2} & \cdot & \cdot \\ \cdot & 1 & 1 \\ \cdot & -1 & 1 \end{pmatrix} \begin{pmatrix} 2 & \cdot & \cdot \\ \cdot & 3 & \cdot \\ \cdot & \cdot & 1 \end{pmatrix} \begin{pmatrix} \sqrt{2} & \cdot & \cdot \\ \cdot & 1 & -1 \\ \cdot & 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & \cdot & \cdot \\ \cdot & 2 & -1 \\ \cdot & -1 & 2 \end{pmatrix}. \end{aligned}$$

14 Die Jordan-Normalform

14.1 Haupträume

Bemerkung 14.1. Nach Satz 10.34 gibt es zwei Gründe, warum ein Endomorphismus nicht diagonalisierbar ist:

- Das charakteristische Polynom zerfällt nicht in Linearfaktoren.
- Die Eigenräume sind zu „klein“, d. h. die geometrische Vielfachheit eines Eigenwerts ist kleiner als die entsprechende algebraische Vielfachheit.

Über \mathbb{C} tritt das erste Problem nach dem Fundamentalsatz der Algebra nicht auf. Um das zweite Problem zu umgehen, ersetzen wir Eigenräume durch größere Unterräume. Im Folgenden sei V ein endlich-dimensionaler K -Vektorraum.

Definition 14.2. Sei $f \in \text{End}(V)$. Ein Unterraum $U \leq V$ heißt *f-invariant*, falls $f(U) \subseteq U$ gilt.

Beispiel 14.3. Für $f \in \text{End}(V)$ sind $\text{Ker}(f)$ und $f(V)$ stets f -invariante Unterräume, denn $f(\text{Ker}(f)) = \{0\} \subseteq \text{Ker}(f)$ und $f(f(V)) \subseteq f(V)$.

Bemerkung 14.4. Sei $U \leq V$ ein f -invarianter Unterraum.

(a) Für $v, w \in V$ gilt

$$v + U = w + U \implies v - w \in U \implies f(v) - f(w) = f(v - w) \in U \implies f(v) + U = f(w) + U.$$

Daher ist $\bar{f}: V/U \rightarrow V/U, v + U \mapsto f(v) + U$ ein wohldefinierter Endomorphismus.

(b) Sei B_1 eine Basis von U und $B = B_1 \dot{\cup} B_2$ eine Basis von V . Aus Dimensionsgründen ist dann $\{b + U : b \in B_2\}$ eine Basis von V/U . Außerdem ist ${}_B[f]_B = \begin{pmatrix} A & C \\ 0 & D \end{pmatrix}$, wobei $A = {}_{B_1}[f|_U]_{B_1}$ und $D = {}_{B_2}[\bar{f}]_{B_2}$. Besitzt U ein f -invariantes Komplement W (d. h. $V = U \oplus W$), so kann man $C = 0$ durch geeignete Basiswahl erreichen. Wir versuchen daher V in möglichst kleine f -invariante Unterräume zu zerlegen.

Definition 14.5. Man nennt $f \in \text{End}(V)$ *trigonalisierbar*, falls eine Basis B von V existiert, sodass ${}_B[f]_B$ eine Dreiecksmatrix ist. Analog heißt $A \in K^{n \times n}$ *trigonalisierbar*, falls A zu einer Dreiecksmatrix ähnlich ist.

Bemerkung 14.6. Sei $B = \{b_1, \dots, b_n\}$ eine Basis von V , sodass ${}_B[f]_B$ eine obere Dreiecksmatrix ist. Dann ist ${}_{B'}[f]_{B'}$ mit $B' = \{b_n, b_{n-1}, \dots, b_1\}$ eine untere Dreiecksmatrix. Man muss in Definition 14.5 also nicht spezifizieren, ob man obere oder untere Dreiecksmatrizen betrachtet (vgl. Bemerkung 15.26).

Beispiel 14.7. Nach der Schur-Zerlegung ist jede komplexe quadratische Matrix trigonalisierbar. Der nächste Satz erweitert Satz 10.52.

Satz 14.8. Eine Abbildung $f \in \text{End}(V)$ ist genau dann trigonalisierbar, wenn μ_f (oder χ_f) in Linearfaktoren zerfällt.

Beweis. Ist ${}_B[f]_B$ (für eine Basis B von V) eine Dreiecksmatrix mit Diagonale $\lambda_1, \dots, \lambda_n$, so zerfällt $\chi_f = (X - \lambda_1) \dots (X - \lambda_n)$ in Linearfaktoren. Nach Cayley-Hamilton und Lemma 10.24 zerfällt auch μ_f in Linearfaktoren.

Nehmen wir umgekehrt an, dass μ_f in Linearfaktoren zerfällt. Nach Satz 10.54 zerfällt auch χ_f in Linearfaktoren. O. B. d. A. sei $V \neq \{0\}$. Dann existiert ein Eigenvektor $v \in V$ von f . Offenbar ist $U := \langle v \rangle$ f -invariant. Sei $\bar{V} := V/U$ und \bar{f} wie in Bemerkung 14.4. Dann ist $\mu_{\bar{f}} \mid \mu_f$. Nach Lemma 10.24 zerfällt auch $\mu_{\bar{f}}$ in Linearfaktoren. Durch Induktion nach $\dim V$ erhalten wir eine Basis $\bar{B} = \{\bar{b}_1, \dots, \bar{b}_{n-1}\}$ von \bar{V} , sodass ${}_{\bar{B}}[\bar{f}]_{\bar{B}}$ eine untere Dreiecksmatrix ist. Wir wählen $b_i \in \bar{b}_i$ für $i = 1, \dots, n-1$. Dann gilt $f(b_i) \in \langle b_i, \dots, b_{n-1}, v \rangle$ für $i = 1, \dots, n-1$. Daher ist $B := \{b_1, \dots, b_{n-1}, v\}$ eine Basis von V , sodass ${}_B[f]_B$ eine untere Dreiecksmatrix ist. \square

Lemma 14.9. Sei $f \in \text{End}(V)$.

- (a) Ist $U \leq V$ f -invariant, so ist $U^0 \leq V^*$ f^* -invariant.
- (b) Ist $U \leq V^*$ f^* -invariant, so ist $U_0 \leq V$ f -invariant.

Beweis. Die dualen Komplemente U^0, U_0 und die duale Abbildung f^* wurden in Abschnitt 7.3 definiert.

- (a) Sei $\gamma \in U^0$. Für alle $u \in U$ gilt $f^*(\gamma)(u) = \gamma(f(u)) = 0$, d. h. $f^*(\gamma) \in U^0$.
- (b) Sei $v \in U_0$. Für alle $\gamma \in U$ gilt $\gamma(f(v)) = f^*(\gamma)(v) = 0$, d. h. $f(v) \in U_0$. \square

Lemma 14.10. Sei $f \in \text{End}(V)$ diagonalisierbar und $U \leq V$ ein f -invarianter Unterraum. Dann ist auch die Einschränkung $f|_U$ diagonalisierbar.

Beweis. Für das Minimalpolynom μ_1 von $f|_U$ gilt $\mu_1 \mid \mu_f$ nach Lemma 10.44. Nach Satz 10.52 zerfällt μ_f in paarweise verschiedene Linearfaktoren. Nach Lemma 10.24 gilt dies auch für μ_1 . \square

Lemma 14.11 (Simultane Diagonalisierung). Seien $f, g \in \text{End}(V)$ diagonalisierbar. Genau dann gilt $f \circ g = g \circ f$, wenn f und g simultan diagonalisierbar sind, d. h. es existiert eine Basis B von V , sodass ${}_B[f]_B$ und ${}_B[g]_B$ Diagonalmatrizen sind.

Beweis. Sei B eine Basis von V , sodass $D_f := {}_B[f]_B$ und $D_g := {}_B[g]_B$ Diagonalmatrizen sind. Aus $D_f D_g = D_g D_f$ folgt dann $f \circ g = g \circ f$. Sei umgekehrt $f \circ g = g \circ f$. Für die Eigenwerte $\lambda_1, \dots, \lambda_k \in K$ von f gilt $V = E_{\lambda_1}(f) \oplus \dots \oplus E_{\lambda_k}(f)$, da f diagonalisierbar ist. Für $v \in U := E_{\lambda_i}(f)$ gilt

$$f(g(v)) = g(f(v)) = \lambda_i g(v)$$

und $g(v) \in U$. Daher ist U ein g -invarianter Unterraum. Nach Lemma 14.10 ist $g|_U$ diagonalisierbar. Man kann also eine Basis von V wählen, die aus gemeinsamen Eigenvektoren von f und g besteht. \square

Lemma 14.12 (Simultane Trigonalisierung). Seien $f, g \in \text{End}(V)$ vertauschbar und trigonalisierbar. Dann sind f und g simultan trigonalisierbar, d. h. es existiert eine Basis B von V , sodass ${}_B[f]_B$ und ${}_B[g]_B$ untere Dreiecksmatrizen sind.

Beweis. O. B. d. A. sei $V \neq \{0\}$. Nach Satz 14.8 besitzt f einen Eigenwert $\lambda \in K$. Wie im Beweis von Lemma 14.11 ist $U := E_\lambda(f)$ ein g -invarianter Unterraum. Das Minimalpolynom von $g|_U$ teilt μ_g und zerfällt daher in Linearfaktoren. Daher existiert ein Eigenvektor $u \in U$ von g . Nun ist $W := \langle u \rangle$ sowohl f -invariant und g -invariant. Sei $\bar{V} := V/W$, $\bar{f} \in \text{End}(\bar{V})$ und $\bar{g} \in \text{End}(\bar{V})$ wie im Beweis von Satz 14.8. Wegen $\mu_{\bar{f}} \mid \mu_f$ und $\mu_{\bar{g}} \mid \mu_g$ kann man induktiv annehmen, dass \bar{f} und \bar{g} simultan trigonalisierbar sind. Die Behauptung folgt nun wie im Beweis von Satz 14.8. \square

Beispiel 14.13. Im Gegensatz zur simultanen Diagonalisierung ist die Umkehrung von Lemma 14.12 falsch: $\text{diag}(1, 0)$ und $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ sind trivialerweise trigonalisierbar, aber nicht vertauschbar.

Definition 14.14. Sei $\dim V = n$ und $f \in \text{End}(V)$ mit Eigenwert $\lambda \in K$. Man nennt

$$H_\lambda(f) := \text{Ker}((f - \lambda \text{id}_V)^n) \leq V$$

den *Hauptraum* von f zu λ . Ist λ Eigenwert einer Matrix $A \in K^{n \times n}$, so definiert man analog $H_\lambda(A) := \text{Ker}((A - \lambda 1_n)^n)$.

Beispiel 14.15. Sei $f \in \text{End}(K^2)$ mit $A := [f] = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Dann ist $E_1(f) = \langle e_1 \rangle$. Wegen $(A - 1_2)^2 = 0$ ist $H_1(f) = K^2$, d. h. der Hauptraum ist größer als der Eigenraum.

Bemerkung 14.16. Für $v \in E_\lambda(f)$ gilt

$$(f - \lambda \text{id}_V)^n(v) = (f - \lambda \text{id}_V)^{n-1}((f - \lambda \text{id}_V)(v)) = (f - \lambda \text{id}_V)^{n-1}(0) = 0.$$

Dies zeigt $E_\lambda(f) \subseteq H_\lambda(f)$. Offenbar ist f mit $f - \lambda \text{id}$ und $(f - \lambda \text{id})^n$ vertauschbar. Für $v \in H_\lambda(f)$ gilt daher

$$(f - \lambda \text{id}_V)^n(f(v)) = ((f - \lambda \text{id}_V)^n \circ f)(v) = f((f - \lambda \text{id}_V)^n(v)) = f(0) = 0.$$

Also ist $f(v) \in H_\lambda(f)$ und $H_\lambda(f)$ ist f -invariant. Wir zeigen als Nächstes, dass $H_\lambda(f)$ ein f -invariantes Komplement besitzt.

Lemma 14.17 (FITTING). Sei $\dim V = n$ und $f \in \text{End}(V)$. Dann ist

$$\boxed{V = f^n(V) \oplus \text{Ker}(f^n)}$$

eine Zerlegung in f -invariante Unterräume.

Beweis. Für $v \in \text{Ker}(f^k)$ gilt $f^{k+1}(v) = f(f^k(v)) = f(0) = 0$. Dies zeigt $\text{Ker}(f) \leq \text{Ker}(f^2) \leq \dots$. Da die Dimension dieser Unterräume durch n beschränkt ist, existiert ein $k \leq n$ mit $\text{Ker}(f^k) = \text{Ker}(f^{k+1})$. Für $v \in \text{Ker}(f^{k+2})$ gilt $f^{k+1}(f(v)) = f^{k+2}(v) = 0$, also $f(v) \in \text{Ker}(f^{k+1}) = \text{Ker}(f^k)$. Dies zeigt $f^{k+1}(v) = 0$ und $v \in \text{Ker}(f^{k+1})$. Induktiv erhält man

$$\text{Ker}(f^k) = \text{Ker}(f^{k+1}) = \dots = \text{Ker}(f^n) = \text{Ker}(f^{n+1}) = \dots$$

Für $v \in \text{Ker}(f^n) \cap f^n(V)$ existiert ein $w \in V$ mit $v = f^n(w)$. Wegen $f^{2n}(w) = f^n(v) = 0$ ist $w \in \text{Ker}(f^{2n}) = \text{Ker}(f^n)$ und es folgt $v = f^n(w) = 0$. Also ist $\text{Ker}(f^n) \cap f^n(V) = \{0\}$. Der Homomorphiesatz zeigt $V = \text{Ker}(f^n) \oplus f^n(V)$. Wegen $f(\text{Ker}(f^n)) \subseteq \text{Ker}(f^{n-1}) \subseteq \text{Ker}(f^n)$ und $f(f^n(V)) = f^n(f(V)) \subseteq f^n(V)$ sind beide Unterräume f -invariant. \square

Beispiel 14.18. Seien $\lambda, \lambda' \in K$ verschiedene Eigenwerte von $f \in \text{End}(V)$. Für $U := H_\lambda(f) \cap H_{\lambda'}(f)$ gilt

$$(f - \lambda \text{id})^n(U) = \{0\} = (f - \lambda' \text{id})^n(U).$$

Das Minimalpolynom der Einschränkung $g := f|_U \in \text{End}(U)$ teilt daher $(X - \lambda)^n$ und $(X - \lambda')^n$ nach Lemma 10.44. Aus Lemma 10.24 folgt $\mu_g = 1$ und $U = \{0\}$.

Satz 14.19 (Hauptraumzerlegung). *Sei $f \in \text{End}(V)$, sodass μ_f in Linearfaktoren zerfällt. Dann gilt*

$$V = H_{\lambda_1}(f) \oplus \dots \oplus H_{\lambda_k}(f)$$

für die verschiedenen Eigenwerte $\lambda_1, \dots, \lambda_k \in K$ von f .

Beweis. Wir argumentieren durch Induktion nach $n := \dim V$. Der Fall $n = 1$ ist trivial. Sei also $n \geq 2$ und die Behauptung für $n - 1$ bereits bewiesen. Da μ_f in Linearfaktoren zerfällt, besitzt f einen Eigenwert $\lambda = \lambda_1 \in K$ nach Satz 10.50. Für die Abbildung $g := f - \lambda \text{id}_V$ gilt

$$V = g^n(V) \oplus \text{Ker}(g^n) = g^n(V) \oplus H_\lambda(f)$$

nach Fitting. Für $U := g^n(V)$ gilt $f(U) = (g + \lambda \text{id})(U) \subseteq g(U) + U \subseteq U$, d. h. U ist f -invariant. Wir betrachten nun die Einschränkung $h := f|_U \in \text{End}(U)$. Nach Lemma 10.44 ist $\mu_h \mid \mu_f$ und μ_h zerfällt in Linearfaktoren (Lemma 10.24). Wegen $\dim H_\lambda(f) \geq \dim E_\lambda(f) \geq 1$ ist $\dim U < n$. Nach Induktion gilt daher

$$U = H_{\lambda_2}(h) \oplus \dots \oplus H_{\lambda_k}(h)$$

für die verschiedenen Eigenwerte $\lambda_2, \dots, \lambda_k$ von h . Wir zeigen $H_{\lambda_i}(h) = H_{\lambda_i}(f)$ für $i = 2, \dots, k$. Wegen $E_{\lambda_i}(h) \subseteq E_{\lambda_i}(f)$ sind $\lambda_2, \dots, \lambda_k$ auch Eigenwerte von f . Wegen $E_{\lambda_i}(h) \cap H_\lambda(f) \subseteq U \cap H_\lambda(f) = \{0\}$ ist $\lambda \neq \lambda_i$ für $i = 2, \dots, k$. Sicher ist $H_{\lambda_i}(h) \subseteq H_{\lambda_i}(f)$. Sei umgekehrt $v \in H_{\lambda_i}(f)$ für ein $i \geq 2$. Dann existieren $u \in U$ und $w \in H_\lambda(f)$ mit $v = u + w$. Es folgt

$$0 = (f - \lambda_i \text{id})^n(v) = (f - \lambda_i \text{id})^n(u) + (f - \lambda_i \text{id})^n(w) \in U \oplus H_\lambda(f).$$

Aus Lemma 8.9 erhält man $(f - \lambda_i \text{id})^n(u) = 0 = (f - \lambda_i \text{id})^n(w)$. Nach Beispiel 14.18 ist $w \in H_\lambda(f) \cap H_{\lambda_i}(f) = \{0\}$ und $v = u \in H_{\lambda_i}(f) \cap U = H_{\lambda_i}(h)$. Dies zeigt

$$V = H_\lambda(f) \oplus U = H_\lambda(f) \oplus H_{\lambda_2}(h) \oplus \dots \oplus H_{\lambda_k}(h) = H_{\lambda_1}(f) \oplus \dots \oplus H_{\lambda_k}(f). \quad \square$$

Bemerkung 14.20. Zerfällt μ_f in paarweise verschiedene Linearfaktoren, so ist f diagonalisierbar (Satz 10.52). Die Hauptraumzerlegung ist dann genau die Zerlegung in Eigenräume, denn $E_\lambda(f) \subseteq H_\lambda(f)$.

Beispiel 14.21. Sei

$$A := \begin{pmatrix} \cdot & 1 & 1 & 1 \\ 1 & 1 & 1 & \cdot \\ 1 & 1 & \cdot & 1 \\ \cdot & \cdot & 1 & 1 \end{pmatrix} \in \mathbb{F}_2^{4 \times 4}.$$

Da 0 und 1 die einzig möglichen Eigenwerte sind, berechnen wir probeweise

$$A^2 = \begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & \cdot \\ 1 & \cdot & 1 & \cdot \\ 1 & 1 & 1 & \cdot \end{pmatrix}, \quad (A - 1_4)^2 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & \cdot & 1 & \cdot \\ 1 & 1 & 1 & 1 \\ \cdot & \cdot & 1 & \cdot \end{pmatrix}^2 = \begin{pmatrix} 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot & \cdot \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

Man sieht $(0, 0, 0, 1), (1, 0, 1, 0) \in H_0(A)$ und $(0, 1, 1, 0), (0, 1, 0, 1) \in H_1(A)$. Aus Dimensionsgründen folgt

$$\mathbb{F}_2^4 = H_0(A) \oplus H_1(A) = \langle (0, 0, 0, 1), (1, 0, 1, 0) \rangle \oplus \langle (0, 1, 1, 0), (0, 1, 0, 1) \rangle.$$

14.2 Jordanblöcke

Bemerkung 14.22. Nach der Hauptraumzerlegung interessieren wir uns für Basen von Haupträumen.

Definition 14.23. Für $\lambda \in K$ und $n \geq 1$ nennt man

$$J_n(\lambda) := \begin{pmatrix} \lambda & & & 0 \\ & \ddots & & \\ 1 & & \ddots & \\ & \ddots & & \ddots \\ 0 & & & 1 & \lambda \end{pmatrix} \in K^{n \times n}$$

einen *Jordanblock* zum Eigenwert λ .¹

Bemerkung 14.24. Offenbar hat $A := J_n(\lambda)$ das charakteristische Polynom $\chi_A = (X - \lambda)^n$, d. h. λ hat algebraische Vielfachheit n . Andererseits hat λ geometrische Vielfachheit $\dim E_\lambda(A) = 1$. Somit ist A besonders weit davon entfernt diagonalisierbar zu sein (Satz 10.34). Für den Standardbasisvektor $e_1 \in K^n$ gilt $Ae_1 = (*, 1, 0, \dots, 0)^t$ und induktiv

$$A^k e_1 = A(A^{k-1} e_1) = A(\underbrace{(*, \dots, *)}_{k-1}, 1, 0, \dots, 0)^t = (\underbrace{(*, \dots, *)}_k, 1, 0, \dots, 0)^t.$$

Also sind $e_1, Ae_1, \dots, A^{n-1}e_1$ linear unabhängig. Daher müssen auch die Matrizen $1_n, A, \dots, A^{n-1}$ linear unabhängig sein. Dies zeigt $\deg \mu_A \geq n$. Aus Cayley-Hamilton folgt $\mu_A = \chi_A = (X - \lambda)^n$.

Definition 14.25. Man nennt $f \in \text{End}(V)$ (bzw. $A \in K^{n \times n}$) *nilpotent*, falls $f^m = f \circ \dots \circ f = 0$ (bzw. $A^m = 0_n$) für ein $m \in \mathbb{N}$ gilt.

Beispiel 14.26. Sei A eine strikte (obere oder untere) Dreiecksmatrix. Dann ist $\chi_A = X^n$. Nach Cayley-Hamilton ist $\mu_A \mid X^n$ und daher $A^n = 0_n$. Also ist A nilpotent. Insbesondere sind Jordanblöcke zum Eigenwert 0 nilpotent.

Bemerkung 14.27. Ist $A \in K^{n \times n}$ nilpotent, so ist $\mu_A \mid X^m$ für ein $m \in \mathbb{N}$. Wegen $\deg \mu_A \leq n$ ist $\mu_A \mid X^n$ und $A^n = 0$. Der folgende Satz liefert ein kanonisches Repräsentantensystem für die Ähnlichkeitsklassen von nilpotenten Matrizen. Die Anzahl der Ähnlichkeitsklassen ist die Anzahl $p(n)$ der *Partitionen* von n , d. h. Zerlegungen der Form $n = n_1 + \dots + n_k$ mit $n_1 \geq \dots \geq n_k$. Zum Beispiel gilt $p(5) = 7$, denn

$$5 = 4 + 1 = 3 + 2 = 3 + 1 + 1 = 2 + 2 + 1 = 2 + 1 + 1 + 1 = 1 + 1 + 1 + 1 + 1.$$

Man kennt keine einfache Formel für $p(n)$.

Satz 14.28. Sei $f \in \text{End}(V)$ nilpotent. Dann existiert eine Basis B von V , sodass

$${}_B[f]_B = \text{diag}(J_{n_1}(0), \dots, J_{n_s}(0)).$$

Die Zahlen $n_1 \geq \dots \geq n_s \geq 1$ sind durch die Gleichungen

$$\boxed{|\{1 \leq i \leq s : n_i \geq k\}|} = \text{rk}(f^{k-1}) - \text{rk}(f^k) \quad (k = 1, \dots, m) \quad (14.1)$$

eindeutig bestimmt. Insbesondere ist $s = \dim \text{Ker}(f)$.

¹In manchen Büchern benutzt man $J_n(\lambda)^t$. Das macht keinen wesentlichen Unterschied (Satz 14.44).

Beweis. Sei $m \in \mathbb{N}$ mit $f^m = 0 \neq f^{m-1}$. Für $k \in \mathbb{N}$ gilt $f(\text{Ker}(f^k)) \subseteq \text{Ker}(f^{k-1})$. Nach Folgerung 4.16 existieren Unterräume U_1, \dots, U_m mit

$$\begin{aligned} V &= \text{Ker}(f^m) = \text{Ker}(f^{m-1}) \oplus U_1, \\ \text{Ker}(f^{m-1}) &= (\text{Ker}(f^{m-2}) + f(U_1)) \oplus U_2, \\ &\vdots \\ \text{Ker}(f) &= (f^{m-1}(U_1) + \dots + f(U_{m-1})) \oplus U_m. \end{aligned}$$

Wir zeigen, dass alle Summen direkt sind. Dies ist für die erste Summe gegeben. Sei induktiv bereits gezeigt:

$$\text{Ker}(f^{m-k+1}) = \text{Ker}(f^{m-k}) \oplus f^{k-1}(U_1) \oplus f^{k-2}(U_2) \oplus \dots \oplus f(U_{k-1}) \oplus U_k. \quad (14.2)$$

Sei $w + f^k(u_1) + \dots + f(u_k) = 0$ mit $w \in \text{Ker}(f^{m-k-1})$ und $u_i \in U_i$ für $i = 1, \dots, k$. Dann folgt

$$\begin{aligned} f^{m-k}(f^{k-1}(u_1) + \dots + u_k) &= f^{m-k-1}(f^k(u_1) + \dots + f(u_k)) \\ &= f^{m-k-1}(w + f^k(u_1) + \dots + f(u_k)) = f^{m-k-1}(0) = 0 \end{aligned}$$

und

$$f^{k-1}(u_1) + \dots + u_k \in \text{Ker}(f^{m-k}) \cap (f^{k-1}(U_1) \oplus \dots \oplus f(U_{k-1}) \oplus U_k) \stackrel{(14.2)}{=} \{0\}.$$

Dies zeigt $f^{k-1}(u_1) = \dots = u_k = 0$ (Lemma 8.9) und es folgt $w = 0$. Also ist

$$\text{Ker}(f^{m-k}) = \text{Ker}(f^{m-k-1}) \oplus f^k(U_1) \oplus f^{k-1}(U_2) \oplus \dots \oplus f(U_k) \oplus U_{k+1}$$

wie gewünscht.

Insgesamt ist

$$V = U_1 \oplus f(U_1) \oplus \dots \oplus f^{m-1}(U_1) \oplus U_2 \oplus f(U_2) \oplus \dots \oplus f^{m-2}(U_2) \oplus \dots \oplus U_m.$$

Sei b_{i1}, \dots, b_{ik_i} eine Basis von U_i für $i = 1, \dots, m$ (der Fall $U_i = \{0\}$ mit $k_i = 0$ ist zugelassen). Wegen $\text{Ker}(f^j) \cap U_i = \{0\}$ ist die Einschränkung $f_{|U_i}^j$ für $j = 1, \dots, m-i$ injektiv (Lemma 7.7). Insbesondere ist $f^j(b_{i1}), \dots, f^j(b_{ik_i})$ eine Basis von $f^j(U_i)$. Also ist

$$B := \bigcup_{i=1}^m \bigcup_{j=1}^{k_i} \{b_{ij}, f(b_{ij}), \dots, f^{m-i}(b_{ij})\}$$

eine Basis von V . Wegen $U_i \subseteq \text{Ker}(f^{m-i+1})$ ist $f^{m-i+1}(b_{ij}) = 0$. Somit entsprechen die Elemente $b_{ij}, f(b_{ij}), \dots, f^{m-i}(b_{ij})$ dem Jordanblock $J_{m-i+1}(0)$ in ${}_B[f]_B$. Insgesamt hat ${}_B[f]_B$ nun die gewünschte Form. Außerdem ist

$$\begin{aligned} k_1 + \dots + k_l &= \dim(f^{l-1}(U_1) \oplus f^{l-2}(U_2) \oplus \dots \oplus U_l) = \dim \text{Ker}(f^{m-l+1}) - \dim \text{Ker}(f^{m-l}) \\ &\stackrel{7.12}{=} \text{rk}(f^{m-l}) - \text{rk}(f^{m-l+1}) \end{aligned}$$

die Anzahl der Jordanblöcke $J_{n_i}(0)$ mit $n_i \geq m-l+1$. Indem man $m-l+1$ durch k ersetzt, folgt (14.1). Die letzte Behauptung erhält man mit $k=1$ in (14.1). \square

Bemerkung 14.29. In der Situation von Satz 14.28 gilt

$$\begin{aligned} |\{1 \leq i \leq s : n_i = k\}| &= |\{1 \leq i \leq s : n_i \geq k\}| - |\{1 \leq i \leq s : n_i \geq k+1\}| \\ &= \text{rk}(f^{k-1}) + \text{rk}(f^{k+1}) - 2 \text{rk}(f^k) \end{aligned}$$

für $k = 1, \dots, n$. Insbesondere ist $2 \text{rk}(f^k) \leq \text{rk}(f^{k+1}) + \text{rk}(f^{k-1})$, d. h. die Folge $\text{rk}(f), \text{rk}(f^2), \dots$ kann nicht zu „schnell“ fallen.

Beispiel 14.30. Sei $f \in \text{End}(\mathbb{R}^5)$ mit

$$A := {}_B[f]_B = \begin{pmatrix} 2 & -3 & -1 & -1 & 2 \\ 1 & -2 & 0 & 0 & 1 \\ 0 & 2 & -1 & -1 & 0 \\ 1 & -4 & 1 & 1 & 1 \\ 0 & -1 & 1 & 1 & 0 \end{pmatrix}.$$

Man berechnet

$$A^2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & -2 & 0 & 0 & 1 \\ -1 & 2 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad A^3 = 0.$$

Also ist $\text{rk}(f^2) = 1$. Wie im Beweis von Satz 14.28 können wir $b_1 := e_1 \notin \text{Ker}(f^2)$ und $U_1 := \langle b_1 \rangle$ mit $\mathbb{R}^5 = \text{Ker}(f^2) \oplus \langle b_1 \rangle$ wählen. Weiter ist

$$A \sim \begin{pmatrix} 1 & -2 & 0 & 0 & 1 \\ 0 & 1 & -1 & -1 & 0 \\ 0 & 2 & -1 & -1 & 0 \\ 0 & -2 & 1 & 1 & 0 \\ 0 & -1 & 1 & 1 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -2 & -2 & 1 \\ 0 & 1 & -1 & -1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\text{Ker}(f) \oplus f(U_1) = \left\langle \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle.$$

Mit $b_2 := e_3 \in \text{Ker}(f^2) \setminus (\text{Ker}(f) \oplus f(U_1))$ und $U_2 := \langle b_2 \rangle$ gilt $\text{Ker}(f^2) = \text{Ker}(f) \oplus f(U_1) \oplus U_2$. Schließlich ist $\text{Ker}(f) = f^2(U_1) \oplus f(U_2)$ (also $U_3 := \{0\}$). Bezüglich der Basis $B := \{b_1, f(b_1), f^2(b_1), b_2, f(b_2)\}$ hat f die Darstellungsmatrix $\text{diag}(J_3(0), J_2(0))$.

Satz 14.31. Sei $f \in \text{End}(V)$, sodass μ_f in Linearfaktoren zerfällt. Dann existiert eine Basis B von V mit

$${}_B[f]_B = \text{diag}(J_{n_1}(\lambda_1), \dots, J_{n_s}(\lambda_s)), \quad (\text{JORDAN-Normalform})$$

wobei $\lambda_1, \dots, \lambda_s \in K$ die Eigenwerte von f sind (möglicherweise mit Vielfachheiten). Die Jordanblöcke $J_{n_i}(\lambda_i)$ sind bis auf die Reihenfolge eindeutig bestimmt.

Beweis. Wir fügen alle Puzzleteile zusammen: Seien $\lambda_1, \dots, \lambda_k \in K$ die verschiedenen Eigenwerte von f und $H_i := H_{\lambda_i}(f)$ für $i = 1, \dots, k$. Nach der Hauptraumzerlegung gilt $V = H_1 \oplus \dots \oplus H_k$. Für $g_i := (f - \lambda_i \text{id}_V)|_{H_i}$ gilt $g_i^n = 0$. Nach Satz 14.28 existiert eine Basis B_i von H_i mit

$${}_{B_i}[g_i]_{B_i} = \text{diag}(J_{m_1}(0), \dots, J_{m_t}(0)).$$

Nach Satz 7.18 ist

$$\begin{aligned} {}_{B_i}[f|_{H_i}]_{B_i} &= {}_{B_i}[g_i + \lambda_i \text{id}_{H_i}]_{B_i} = {}_{B_i}[g_i]_{B_i} + \lambda_i {}_{B_i}[\text{id}_{H_i}]_{B_i} = \text{diag}(J_{m_1}(0), \dots, J_{m_t}(0)) + \lambda_i 1 \\ &= \text{diag}(J_{m_1}(\lambda_i), \dots, J_{m_t}(\lambda_i)). \end{aligned}$$

Also ist $B := B_1 \cup \dots \cup B_k$ eine geeignete Basis. Für ein fest gewähltes Paar (n_i, λ_i) ergibt sich die Anzahl der Blöcke $J_{n_i}(\lambda_i)$ aus Bemerkung 14.29 angewendet auf g_i . \square

Folgerung 14.32. Jede Matrix $A \in \mathbb{C}^{n \times n}$ besitzt eine Jordan-Normalform (genauer: A ist zu einer Jordan-Normalform ähnlich).

Beweis. Nach Folgerung 11.35 zerfällt μ_A in Linearfaktoren. □

Bemerkung 14.33.

- (a) Da man die Eigenwerte $\lambda_1, \dots, \lambda_n$ in der Regel nicht sinnvoll ordnen kann (vgl. Bemerkung 11.26), gibt es im Gegensatz zu Satz 14.28 keine kanonische Reihenfolge der Jordanblöcke. Man kann jedoch die *lexikografische* Ordnung

$$x <_l y \iff \operatorname{Re}(x) < \operatorname{Re}(y) \vee (\operatorname{Re}(x) = \operatorname{Re}(y) \wedge \operatorname{Im}(x) < \operatorname{Im}(y))$$

auf \mathbb{C} definieren und die Jordanblöcke zunächst nach Größe und dann nach Eigenwert bzgl. $<_l$ sortieren. Im Folgenden sprechen wir etwas ungenau von *der* Jordan-Normalform von f .

- (b) Sei $A \in K^{n \times n}$. In Satz 15.39 konstruieren wir einen Körper $L \supseteq K$, sodass μ_A in $L[X]$ in Linearfaktoren zerfällt. Man kann dann A als Matrix in $L^{n \times n}$ auffassen und dort die Jordan-Normalform bestimmen.

Beispiel 14.34. Sei $f \in \operatorname{End}(\mathbb{C}^3)$ mit

$$A := [f] = \begin{pmatrix} 5 & 0 & 1 \\ -5 - i & -i & -1 \\ -9 & 0 & -1 \end{pmatrix}.$$

Durch Laplace-Entwicklung nach der zweiten Spalte erhält man

$$\chi_f = \chi_A = (X + i)((X - 5)(X + 1) + 9) = (X + i)(X^2 - 4X + 4) = (X + i)(X - 2)^2.$$

Also sind $\lambda_1 = 2$ und $\lambda_2 = -i$ die Eigenwerte von f . Offensichtlich ist $b_3 := e_2$ ein Eigenvektor zu λ_2 . Wegen

$$A - 2 \cdot 1_3 = \begin{pmatrix} 3 & 0 & 1 \\ -5 - i & -2 - i & -1 \\ -9 & 0 & -3 \end{pmatrix} \sim \begin{pmatrix} 3 & 0 & 1 \\ -5 - i & -2 - i & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

ist $\operatorname{rk}(f - 2 \operatorname{id}) = 2$, d. h. $\lambda_1 = 2$ hat geometrische Vielfachheit 1. Also ist f nicht diagonalisierbar und die Jordan-Normalform von f muss $\operatorname{diag}(J_2(2), J_1(-i))$ sein. Wegen

$$(A - 2 \cdot 1_3)^2 = \begin{pmatrix} 0 & 0 & 0 \\ 3 + 4i & 3 + 4i & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

können wir $b_1 := e_3 \in \operatorname{Ker}((f - 2 \operatorname{id})^2) \setminus \operatorname{Ker}(f - 2 \operatorname{id})$ wie in Beispiel 14.30 wählen. Für

$$b_2 := (f - 2 \operatorname{id})(b_1) = (1, -1, -3)$$

gilt $f(b_1) = 2b_1 + b_2$ und $f(b_2) = (f - 2 \operatorname{id})(b_2) + 2b_2 = 2b_2$. Für die Basis $B := \{b_1, b_2, b_3\}$ erhält man

$${}_B[f]_B = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & -i \end{pmatrix} = \operatorname{diag}(J_2(2), J_1(-i)).$$

14.3 Anwendungen

Satz 14.35. Sei $J := \text{diag}(J_{n_1}(\lambda_1), \dots, J_{n_s}(\lambda_s))$ die Jordan-Normalform von $A \in K^{n \times n}$. Seien ρ_1, \dots, ρ_k die verschiedenen Eigenwerte von A . Für $i = 1, \dots, k$ sei a_i , m_i und s_i die Anzahl, das Maximum und die Summe der n_j mit $\lambda_j = \rho_i$. Dann ist a_i die geometrische Vielfachheit von ρ_i und

$$\chi_A = (X - \rho_1)^{s_1} \dots (X - \rho_k)^{s_k} \quad \mu_A = (X - \rho_1)^{m_1} \dots (X - \rho_k)^{m_k}.$$

Insbesondere ist A genau dann diagonalisierbar, wenn J eine Diagonalmatrix ist.

Beweis. Man sieht leicht $\text{rk}(J - \mu_i 1_n) = n - a_i$, d. h. a_i ist die geometrische Vielfachheit von ρ_i (vgl. Satz 14.28). Da J eine untere Dreiecksmatrix ist, ist s_i die algebraische Vielfachheit von ρ_i . Da μ_A nicht von der Basiswahl abhängt, gilt

$$0 = \mu_A(J) = \text{diag}(\mu_A(J_{n_1}(\lambda_1)), \dots, \mu_A(J_{n_s}(\lambda_s))).$$

Aus Bemerkung 14.24 folgt $\mu_A = (X - \rho_1)^{m_1} \dots (X - \rho_k)^{m_k}$. Die zweite Behauptung gilt nach Satz 10.52. \square

Beispiel 14.36. Für $A = \text{diag}(J_3(1), J_2(1), J_4(i))$ gilt $\chi_A = (X - 1)^5(X - i)^4$ und $\mu_A = (X - 1)^3(X - i)^4$.

Folgerung 14.37. Sei $A, B \in \mathbb{C}^{n \times n}$. Im Fall $n \leq 3$ gilt

$$A \approx B \iff \chi_A = \chi_B, \mu_A = \mu_B.$$

Im Fall $n \leq 6$ sind folgende Aussagen äquivalent:

- (1) $A \approx B$.
- (2) $\chi_A = \chi_B$, $\mu_A = \mu_B$ und die geometrischen Vielfachheiten der Eigenwerte von A stimmen mit denen von B überein.

Beweis. Bekanntlich impliziert $A \approx B$ die jeweils andere Aussage. Sei nun $n \leq 3$, $\chi_A = \chi_B$ und $\mu_A = \mu_B$. Dann haben A und B die gleichen Eigenwerte mit den gleichen algebraischen Vielfachheiten. Sei λ ein Eigenwert mit algebraischer Vielfachheit $r \leq 3$. Für die entsprechenden Größen der Jordanblöcke $n_1 \geq \dots \geq n_k$ von A gilt $r = n_1 + \dots + n_k$. Da n_1 durch μ_A festgelegt ist, sind auch k und n_2, \dots, n_k durch r eindeutig bestimmt. Also haben A und B die gleiche Jordan-Normalform. Es folgt $A \approx B$.

Sei jetzt $n \leq 6$ und (2) erfüllt. Neben r und n_1 ist nun auch die geometrische Vielfachheit k von λ eindeutig bestimmt. Es gibt folgende Möglichkeiten:

- $k = 1$: Hier ist $n_1 = r$.
- $k = 2$: Hier ist $n_2 = r - n_1$.
- $k = 3$: Wegen $r \leq 6$ ist $(n_1, n_2, n_3) \in \{(1, 1, 1), (2, 1, 1), (2, 2, 1), (2, 2, 2), (3, 1, 1), (3, 2, 1), (4, 1, 1)\}$. Diese Tripel sind durch r und n_1 eindeutig bestimmt.
- $k = 4$: Hier ist $(n_1, n_2, n_3, n_4) \in \{(1, 1, 1, 1), (2, 1, 1, 1), (2, 2, 1, 1), (3, 1, 1, 1)\}$.
- $k = 5$: Hier ist $n_1 = r - 4$ und $n_2 = \dots = n_5 = 1$.
- $k = 6$: Hier ist $n_1 = \dots = n_6 = 1$.

In allen Fällen folgt $A \approx B$. \square

Beispiel 14.38. Die Matrizen

$$\begin{aligned} \text{diag}(J_2(0), J_2(0)) &\not\approx \text{diag}(J_2(0), J_1(0), J_1(0)), \\ \text{diag}(J_3(0), J_2(0), J_2(0)) &\not\approx \text{diag}(J_3(0), J_3(0), J_1(0)) \end{aligned}$$

zeigen, dass man Folgerung 14.37 nicht auf $n = 4$ bzw. $n = 7$ erweitern kann.

Satz 14.39. Eine Matrix $A \in \mathbb{C}^{n \times n}$ ist genau dann nilpotent, wenn $\text{tr}(A^k) = 0$ für $k = 1, \dots, n$ gilt.

Beweis. O. B. d. A. sei A in Jordan-Normalform mit paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_s \in \mathbb{C}$. Sei m_i die algebraische Vielfachheit von λ_i . Die Eigenwerte von A^k sind dann $\lambda_1^k, \dots, \lambda_s^k$ mit den entsprechenden Vielfachheiten. Nach Bemerkung 10.35 gilt

$$\text{tr}(A^k) = m_1 \lambda_1^k + \dots + m_s \lambda_s^k = 0.$$

Ist A nilpotent, so gilt $s = 1$, $\lambda_1 = 0$ und $\text{tr}(A^k) = 0$ für $k = 1, \dots, n$.

Sei umgekehrt $\text{tr}(A^k) = m_1 \lambda_1^k + \dots + m_s \lambda_s^k = 0$ für $k = 1, \dots, n$. Nehmen wir indirekt an, dass A nicht nilpotent ist. O. B. d. A. sei $\lambda_i \neq 0$ für $i = 1, \dots, s$. Dann ist $(m_1, \dots, m_s)^t$ eine Lösung des homogenen linearen Gleichungssystems mit Koeffizientenmatrix $V = (\lambda_j^i) \in K^{s \times s}$. Nach Vandermonde ist allerdings $\det(V) = \lambda_1 \dots \lambda_s \det(\lambda_j^{i-1}) \neq 0$. Widerspruch. \square

Bemerkung 14.40. Satz 14.39 gilt nicht über beliebigen Körpern. Zum Beispiel ist $\text{tr}(1_2^k) = \text{tr}(1_2) = 0$ in \mathbb{F}_2 für alle $k \in \mathbb{N}$.

Lemma 14.41. Für $\lambda \in \mathbb{C}$, $n \in \mathbb{N}$ und $k \in \mathbb{N}_0$ gilt

$$J_n(\lambda)^k = \begin{pmatrix} \lambda^k & & & & 0 \\ k\lambda^{k-1} & \ddots & & & \\ \binom{k}{2}\lambda^{k-2} & \ddots & \ddots & & \\ \vdots & \ddots & \ddots & \ddots & \\ \binom{k}{n-1}\lambda^{k-n+1} & \dots & \binom{k}{2}\lambda^{k-2} & k\lambda^{k-1} & \lambda^k \end{pmatrix},$$

wobei $\binom{k}{l} = \frac{k(k-1)\dots(k-l+1)}{l!}$ für $l = 1, \dots, n-1$.

Beweis. Im Fall $\lambda = 0$ sieht man leicht:

$$J_n(0)^2 = \begin{pmatrix} 0 & & & 0 \\ 0 & \ddots & & \\ 1 & \ddots & \ddots & \\ & \ddots & \ddots & \ddots \\ 0 & & 1 & 0 & 0 \end{pmatrix}, \dots, J_n(0)^{n-1} = \begin{pmatrix} 0 & & & 0 \\ \vdots & \ddots & & \\ 0 & & \ddots & \\ 1 & 0 & \dots & 0 \end{pmatrix}, J_n(0)^n = 0.$$

Der allgemeine Fall folgt aus der binomischen Formel²

$$J_n(\lambda)^k = (\lambda 1_n + J_n(0))^k = \sum_{l=0}^k \binom{k}{l} \lambda^{k-l} J_n(0)^l. \quad \square$$

²Die Formel ist anwendbar, da 1_n und $J_n(0)$ vertauschbar sind.

Satz 14.42. Für $k \in \mathbb{N}$ und $A \in \text{GL}(n, \mathbb{C})$ existiert ein $W \in \mathbb{C}^{n \times n}$ mit $W^k = A$.

Beweis. Wegen $(SWS^{-1})^k = SW^kS^{-1}$ für $S \in \text{GL}(n, \mathbb{C})$ können wir annehmen, dass A eine Jordan-Normalform ist. Es genügt eine k -te Wurzel für jeden Jordanblock zu konstruieren. Sei also $A = J_n(\lambda)$ mit $\lambda \in \mathbb{C}$. Da A invertierbar ist, gilt $\lambda \neq 0$. Nach Lemma 11.27 existiert ein $\mu \in \mathbb{C}$ mit $\mu^k = \lambda$. Sei $J := J_n(\mu)$. Nach Lemma 14.41 besitzt J^k direkt unterhalb der Hauptdiagonale Einträge $k\mu^{k-1} \neq 0$. Also ist $\mu^k = \lambda$ der einzige Eigenwert von J^k und die geometrische Vielfachheit beträgt 1. Nach Satz 14.35 ist A die Jordan-Normalform von J^k . Insbesondere ist $A \approx J^k$. \square

Beispiel 14.43.

(a) Der Beweis zeigt $J_n(1)^k \approx J_n(1)$ für alle $k, n \in \mathbb{N}$. Wir suchen eine dritte Wurzel von $J_3(1)$. Für

$$S := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 3 & 9 \end{pmatrix}$$

gilt

$$S^{-1}J_3(1)^3S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & -1/9 & 1/9 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 3 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 3 & 9 \end{pmatrix} = J_3(1).$$

Für

$$W := S^{-1}J_3(1)S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & -1/9 & 1/9 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 3 & 0 \\ 0 & 6 & 9 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1/3 & 1 & 0 \\ -1/9 & 1/3 & 1 \end{pmatrix}$$

gilt $W^3 = J_3(1)$. In Satz 15.35 bestimmen wir die Anzahl der k -ten Wurzeln von $J_n(\lambda)$ für alle $\lambda \in \mathbb{C}$ und $n, k \in \mathbb{N}$.

(b) Die nilpotente Matrix $J_2(0)$ besitzt keine Quadratwurzel, denn für $A \in K^{2 \times 2}$ mit $A^2 = J_2(0)$ wäre $A^4 = 0$ und damit $A^2 = 0$ (Bemerkung 14.27).

(c) Man rechnet leicht nach, dass auch $J_2(1) \in \text{GL}(2, \mathbb{F}_2)$ keine Quadratwurzel besitzt.

Satz 14.44. Für alle $A \in \mathbb{C}^{n \times n}$ gilt $A \approx A^t$.

Beweis. Wie üblich können wir $A = J_n(\lambda)$ annehmen. Da λ der einzige Eigenwert von A^t ist und die geometrische Vielfachheit 1 beträgt, ist A die Jordan-Normalform von A^t . \square

Bemerkung 14.45. Man kann Satz 14.44 auch direkt nachrechnen:

$$\begin{pmatrix} 0 & & 1 \\ & \ddots & \\ 1 & & 0 \end{pmatrix}^{(-1)} J_n(\lambda) \begin{pmatrix} 0 & & 1 \\ & \ddots & \\ 1 & & 0 \end{pmatrix} = \begin{pmatrix} 0 & & 1 \\ & \ddots & \\ 1 & & 0 \end{pmatrix} \begin{pmatrix} 0 & & \lambda \\ & \ddots & 1 \\ \lambda & 1 & 0 \end{pmatrix} = J_n(\lambda)^t.$$

In Satz 15.25 beweisen wir Satz 14.44 über beliebigen Körpern.

15 Die Frobenius-Normalform

15.1 Irreduzible Polynome

Bemerkung 15.1. Wir konstruieren in diesem Kapitel eine Normalform für Endomorphismen, die im Gegensatz zur Jordan-Normalform über jedem Körper existiert. Da im Allgemeinen nicht jedes Polynom in Linearfaktoren zerfällt, kann man nicht erwarten, dass die Normalform eine Dreiecksmatrix ist (Satz 14.8). Dennoch erzielen wir die gleiche Anzahl an Nulleinträgen wie in der Jordan-Normalform. Die Grundidee ist, das Minimalpolynom in möglichst „kleine“ Polynome zu faktorisieren. Wir immer sei V ein endlich-dimensionaler K -Vektorraum.

Definition 15.2.

- Ein normiertes Polynom $\alpha \in K[X] \setminus K$ heißt *irreduzibel*, falls es keine Faktorisierung $\alpha = \beta\gamma$ mit $\beta, \gamma \in K[X] \setminus K$ gibt (vgl. Primzahl).
- Wir nennen $\alpha, \beta \in K[X]$ *teilerfremd*, falls jeder gemeinsame Teiler von α und β konstant ist, also in K liegt.

Beispiel 15.3.

- (a) Normierte Polynome vom Grad 1 sind irreduzibel, denn $\deg(\alpha\beta) = \deg(\alpha) + \deg(\beta)$ für $\alpha, \beta \in K[X]$. In $\mathbb{C}[X]$ hat umgekehrt jedes irreduzible Polynom Grad 1 nach dem Fundamentalsatz der Algebra und Lemma 10.22. Im Allgemeinen sind normierte Polynome vom Grad ≤ 3 genau dann irreduzibel, wenn sie keine Nullstelle besitzen. Zum Beispiel ist $X^2 + X + 1 \in \mathbb{F}_2[X]$ irreduzibel.
- (b) Das Polynom $X^2 - 2$ ist irreduzibel in $\mathbb{Q}[X]$, aber nicht in $\mathbb{R}[X]$, denn $X^2 - 2 = (X + \sqrt{2})(X - \sqrt{2})$. Analog ist $X^2 + 1$ irreduzibel in $\mathbb{R}[X]$, aber nicht in $\mathbb{C}[X]$ wegen $X^2 + 1 = (X + i)(X - i)$.
- (c) Zwei verschiedene irreduzible Polynome sind teilerfremd. Ist $\gamma \in K[X] \setminus K$ ein gemeinsamer Teiler von α und β , so ist auch jeder irreduzible Teiler von γ ein gemeinsamer Teiler von α und β . Also sind α und β genau dann teilerfremd, wenn sie keine irreduziblen gemeinsamen Teiler haben.

Satz 15.4 (Euklidischer Algorithmus). *Seien $\alpha, \beta \in K[X]$ mit $\deg \alpha \geq \deg \beta$. Der folgende Algorithmus bestimmt, ob α und β teilerfremd sind:*

(1) Setze $\alpha_0 := \alpha$, $\alpha_1 := \beta$ und $i := 1$.

(2) Wiederhole:

- *Division mit Rest:* $\alpha_{i-1} = \alpha_i \gamma_i + \alpha_{i+1}$ mit $\deg(\alpha_{i+1}) < \deg(\alpha_i)$.
- *Gilt $\alpha_{i+1} = 0$, so breche ab.*
- *Erhöhe i um 1.*

(3) *Genau dann sind α und β teilerfremd, wenn $\alpha_i \in K$ gilt.*

Beweis. Jeder gemeinsame Teiler von α_{i-1} und α_i teilt auch $\alpha_{i-1} - \alpha_i \gamma_i = \alpha_{i+1}$ für $i = 1, 2, \dots$. Da $\deg(\alpha_i)$ mit jeder Division mit Rest kleiner wird, muss der Algorithmus nach endlich vielen Schritten abbrechen. Am Ende gilt $\alpha_{i+1} = 0 \neq \alpha_i$, d. h. $\alpha_{i-1} = \alpha_i \gamma_i$. Daher sind α und β genau dann teilerfremd, wenn α_i konstant ist. \square

Beispiel 15.5. Sei $\alpha = X^5 + X^4 + X^3 + 1 \in \mathbb{F}_2[X]$ und $\beta = X^4 + X^3 + X + 1 \in \mathbb{F}_2[X]$. Es gilt

$$\begin{aligned} X^5 + X^4 + X^3 + 1 &= (X^4 + X^3 + X + 1)X + X^3 + X^2 + X + 1 && \implies \alpha_2 = X^3 + X^2 + X + 1, \\ X^4 + X^3 + X + 1 &= (X^3 + X^2 + X + 1)X + X^2 + 1 && \implies \alpha_3 = X^2 + 1 \notin \mathbb{F}_2, \\ X^3 + X^2 + X + 1 &= (X^2 + 1)(X + 1) && \implies \alpha_4 = 0. \end{aligned}$$

Also ist $X^2 + 1$ ein gemeinsamer Teiler von α und β .

Lemma 15.6 (BÉZOUT). Sind $\alpha, \beta \in K[X]$ teilerfremd, so existieren $\tilde{\alpha}, \tilde{\beta} \in K[X]$ mit $\alpha\tilde{\alpha} + \beta\tilde{\beta} = 1$.

Beweis. Nach Voraussetzung sind α und β nicht beide 0. Also existieren $\tilde{\alpha}, \tilde{\beta} \in K[X]$, sodass $\rho := \alpha\tilde{\alpha} + \beta\tilde{\beta} \neq 0$ minimalen Grad hat. Division mit Rest liefert $\gamma, \delta \in K[X]$ mit $\alpha = \gamma\rho + \delta$ und $\deg \delta < \deg \rho$. Also ist

$$\delta = \alpha - \gamma\rho = \alpha - \gamma\alpha\tilde{\alpha} - \gamma\beta\tilde{\beta} = \alpha(1 - \gamma\tilde{\alpha}) - \beta(\gamma\tilde{\beta}).$$

Die Wahl von ρ zeigt $\delta = 0$. Es folgt $\rho \mid \alpha$. Analog ergibt sich $\rho \mid \beta$. Da α und β teilerfremd sind, ist $\rho \in K^\times$. Nach Normierung kann man $\rho = 1$ annehmen. \square

Bemerkung 15.7. Die Polynome $\tilde{\alpha}$ und $\tilde{\beta}$ in Lemma 15.6 lassen sich berechnen, indem man den euklidischen Algorithmus rückwärts liest

$$\alpha_i = \alpha_{i-2} - \alpha_{i-1}\gamma_{i-1} = \alpha_{i-2} - (\alpha_{i-3} - \alpha_{i-2}\gamma_{i-2})\gamma_{i-1} = \alpha_{i-3}\gamma_{i-2} + \alpha_{i-2}(1 - \gamma_{i-2}\gamma_{i-1}) = \dots$$

und durch α_i teilt.

Beispiel 15.8. Sei $\alpha := X^3 + X^2 + 1$ und $\beta := X^2 - X$. Der euklidische Algorithmus liefert

$$\begin{aligned} X^3 + X^2 + 1 &= (X^2 - X)(X + 2) + 2X + 1 \\ X^2 - X &= (2X + 1)\frac{1}{4}(2X - 3) + \frac{3}{4}. \end{aligned}$$

Also ist

$$\begin{aligned} \frac{3}{4} &= \beta - \frac{1}{4}(2X + 1)(2X - 3) = \beta - \frac{1}{4}(\alpha - \beta(X + 2))(2X - 3) \\ &= -\frac{1}{4}(2X - 3)\alpha + \frac{1}{4}(4 + (X + 2)(2X - 3))\beta \end{aligned}$$

und

$$-\frac{1}{3}(2X - 3)\alpha + \frac{1}{3}(2X^2 + X - 2)\beta = 1.$$

Satz 15.9 (Primfaktorzerlegung in $K[X]$). Jedes normierte Polynom in $K[X] \setminus K$ ist ein Produkt von irreduziblen Faktoren, die bis auf die Reihenfolge eindeutig bestimmt sind.

Beweis. Sei $\alpha \in K[X]$ normiert. Induktion nach $d := \deg(\alpha)$. Ist α irreduzibel (zum Beispiel $d = 1$), so sind wir fertig. Anderenfalls ist $\alpha = \beta\gamma$ mit $\deg(\beta), \deg(\gamma) < d$. Wir können annehmen, dass β und γ normiert sind. Nach Induktion sind β und γ Produkte von irreduziblen Faktoren und daher auch α .

Sei nun $\alpha = \sigma_1 \dots \sigma_n = \tau_1 \dots \tau_m$ mit irreduziblen Polynomen $\sigma_1, \dots, \sigma_n, \tau_1, \dots, \tau_m \in K[X]$. Induktion nach m . Im Fall $m = 1$ ist $n = 1$ und $\sigma_1 = \tau_1$. Sei jetzt $m \geq 2$. Im Fall $\sigma_1 \neq \tau_1$ sind σ_1 und τ_1 teilerfremd. Nach Bézout existieren $\tilde{\sigma}, \tilde{\tau} \in K[X]$ mit $\sigma_1 \tilde{\sigma} + \tau_1 \tilde{\tau} = 1$. Es folgt

$$\sigma_1(\tilde{\sigma}\tau_2 \dots \tau_m + \sigma_2 \dots \sigma_n \tilde{\tau}) = \tau_2 \dots \tau_m.$$

Induktiv erhält man $\sigma_1 = \tau_i$ für ein $i \in \{1, \dots, m\}$. Dann ist $\sigma_2 \dots \sigma_n = \tau_1 \dots \tau_{i-1} \tau_{i+1} \dots \tau_m$ und Induktion liefert $\{\sigma_1, \dots, \sigma_n\} = \{\tau_1, \dots, \tau_m\}$. \square

Beispiel 15.10. Wie bei natürlichen Zahlen werden wir in der Primfaktorzerlegung von Polynomen gleiche irreduzible Faktoren in Potenzen zusammenfassen. Zum Beispiel ist

$$X^6 + X^5 - X^4 - X^3 = X^3(X+1)^2(X-1).$$

Aus algorithmischer Sicht ist die Bestimmung der Primfaktorzerlegung (von Zahlen wie Polynomen) ein sehr schwieriges Problem.

Folgerung 15.11. Seien $\alpha, \beta \in K[X]$ und γ ein irreduzibler Teiler von $\alpha\beta$. Dann gilt $\gamma \mid \alpha$ oder $\gamma \mid \beta$.

Beweis. Wegen $\gamma \mid \alpha\beta$ muss γ in der Primfaktorzerlegung von $\alpha\beta$ auftauchen. Diese Primfaktorzerlegung erhält man, indem man die Primfaktorzerlegungen von α und β zusammenfasst. Also muss γ in der Zerlegung von α oder β auftreten. \square

Satz 15.12 (Primärzerlegung). Sei $f \in \text{End}(V)$ und $\mu_f = \gamma_1^{a_1} \dots \gamma_k^{a_k}$ die Primfaktorzerlegung von μ_f in $K[X]$. Dann ist

$$V = \text{Ker}(\gamma_1^{a_1}(f)) \oplus \dots \oplus \text{Ker}(\gamma_k^{a_k}(f))$$

eine Zerlegung in f -invariante Unterräume. Außerdem ist $\gamma_i^{a_i}$ das Minimalpolynom der Einschränkung von f auf $\text{Ker}(\gamma_i^{a_i}(f))$ für $i = 1, \dots, k$.

Beweis. Für $1 \leq i \leq k$ sei $V_i := \text{Ker}(\gamma_i^{a_i}(f))$. Für $v \in V_i$ gilt

$$\gamma_i^{a_i}(f)(f(v)) = f(\gamma_i^{a_i}(f)(v)) = f(0) = 0,$$

d. h. $f(v) \in V_i$. Also ist V_i f -invariant. Zum Nachweis der direkten Zerlegung argumentieren wir durch Induktion nach k . Für $k = 1$ ist $V_1 = \text{Ker}(\mu_f(f)) = \text{Ker}(0) = V$. Sei also $k \geq 2$. Die Polynome $\alpha := \gamma_1^{a_1}$ und $\beta := \gamma_2^{a_2} \dots \gamma_k^{a_k}$ sind teilerfremd, da sie keinen irreduziblen gemeinsamen Teiler haben. Nach Bézout existieren $\tilde{\alpha}, \tilde{\beta} \in K[X]$ mit $\alpha\tilde{\alpha} + \beta\tilde{\beta} = 1$. Sei $V_\beta := \text{Ker}(\beta(f))$. Wegen $(\alpha\beta)(f) = 0 = (\beta\alpha)(f)$ ist $\beta(f)(V) \subseteq V_1$ und $\alpha(f)(V) \subseteq V_\beta$. Es folgt

$$V = \text{id}(V) = (\alpha\tilde{\alpha} + \beta\tilde{\beta})(f)(V) \subseteq \alpha(f)(V) + \beta(f)(V) \subseteq V_\beta + V_1.$$

Für $v \in V_1 \cap V_\beta$ ist andererseits

$$v = (\alpha\tilde{\alpha} + \beta\tilde{\beta})(f)(v) = \tilde{\alpha}(f)(\alpha(f)(v)) + \tilde{\beta}(f)(\beta(f)(v)) = 0.$$

Dies zeigt $V = V_1 \oplus V_\beta$.

Das Minimalpolynom α_1 der Einschränkung $f|_{V_1}$ teilt α , denn $\alpha(f)(V_1) = 0$. Insbesondere gilt $\deg(\alpha_1) \leq \deg(\alpha)$. Genauso ist β durch das Minimalpolynom β_1 von $f_\beta := f|_{V_\beta}$ teilbar. Wegen $V = V_1 \oplus V_\beta$ gilt andererseits $(\alpha_1\beta_1)(f) = 0$ und $\mu_f \mid \alpha_1\beta_1$. Aus

$$\deg(\alpha) + \deg(\beta) = \deg(\mu_f) \leq \deg(\alpha_1\beta_1) = \deg(\alpha_1) + \deg(\beta_1)$$

folgt $\alpha_1 = \alpha$ sowie $\beta_1 = \beta$. Wir können nun die Induktionsvoraussetzung auf $f_\beta \in \text{End}(V_\beta)$ anwenden. Für $i = 2, \dots, k$ gilt $\text{Ker}(\gamma_i^{a_i}(f_\beta)) = V_i \cap V_\beta = V_i$. Dies zeigt

$$V_\beta = V_2 \oplus \dots \oplus V_k.$$

Die Behauptung folgt aus Lemma 8.9. □

Bemerkung 15.13. Nehmen wir an, dass μ_f in Linearfaktoren zerfällt, d. h. es gilt $\gamma_i = X - \lambda_i$ für $i = 1, \dots, k$, wobei $\lambda_1, \dots, \lambda_k$ die verschiedenen Eigenwerte von f sind. Wegen

$$\text{Ker}(\gamma_i^{a_i}(f)) = \text{Ker}((f - \lambda_i \text{id})^{a_i}) \subseteq \text{Ker}((f - \lambda_i \text{id})^n) = H_{\lambda_i}(f)$$

stimmt die Primärzerlegung mit der Hauptraumzerlegung aus Satz 14.19 überein.

Definition 15.14. Für $v \in V$ und $f \in \text{End}(V)$ nennt man $U := \langle f^i(v) : i \in \mathbb{N}_0 \rangle \leq V$ einen *zyklischen* Unterraum. Offenbar ist U f -invariant. Wir bezeichnen das Minimalpolynom von $f|_U$ mit μ_v .

Bemerkung 15.15. Sei $f \in \text{End}(V)$ und $v \in V \setminus \{0\}$. Sei $d \in \mathbb{N}$ minimal, sodass $v, f(v), \dots, f^d(v)$ linear abhängig sind. Dann existieren $a_0, \dots, a_{d-1} \in K$ mit $f^d(v) + a_{d-1}f^{d-1}(v) + \dots + a_1f(v) + a_0v = 0$. Sei $\alpha := X^d + a_{d-1}X^{d-1} + \dots + a_0 \in K[X]$. Dann gilt

$$\alpha(f)(f^i(v)) = f^i(\alpha(f)(v)) = f^i(0) = 0$$

für alle $i \in \mathbb{N}_0$. Also ist $\mu_v \mid \alpha$. Da $v, f(v), \dots, f^{d-1}(v)$ linear unabhängig, gilt andererseits $\deg(\mu_v) \geq d$. Dies zeigt $\mu_v = \alpha$.

Lemma 15.16. Für $f \in \text{End}(V)$ existiert ein $v \in V$ mit $\mu_v = \mu_f$.

Beweis. Sei $V_i := \text{Ker}(\gamma_i^{a_i}(f))$ für $i = 1, \dots, k$ wie in Satz 15.12. Für $v \in V_i$ ist $\langle f^j(v) : j \in \mathbb{N}_0 \rangle \subseteq V_i$. Daher ist μ_v ein Teiler von $\gamma_i^{a_i}$. Nach der eindeutigen Primfaktorzerlegung gilt $\mu_v = \gamma_i^{b_i}$ für ein $b_i \leq a_i$. Da $\gamma_i^{a_i}$ das Minimalpolynom von $f|_{V_i}$ ist, muss ein $v_i \in V_i$ mit $\mu_{v_i} = \gamma_i^{a_i}$ existieren. Wir setzen $v := v_1 + \dots + v_k$. Da V_i f -invariant ist, gilt $\mu_v(f)(v_i) = 0$ für $i = 1, \dots, k$. Dies zeigt $\gamma_i^{a_i} = \mu_{v_i} \mid \mu_v$ und es folgt $\mu_v = \mu_f$. □

15.2 Begleitmatrizen

Bemerkung 15.17. Wir definieren das Gegenstück zu den Jordanblöcke über beliebigen Körpern.

Definition 15.18. Für $\alpha := X^n + a_{n-1}X^{n-1} + \dots + a_0 \in K[X] \setminus K$ nennt man

$$B(\alpha) := \begin{pmatrix} 0 & & & -a_0 \\ 1 & \ddots & & \vdots \\ & \ddots & 0 & -a_{n-2} \\ 0 & & 1 & -a_{n-1} \end{pmatrix} \in K^{n \times n}$$

die *Begleitmatrix* von α .

Beispiel 15.19. Offenbar ist $B_{X^n} = J_n(0)$ ein Jordanblock und $B_{X^{n-1}} = P_\sigma$ die Permutationsmatrix des n -Zyklus $\sigma = (1, \dots, n)$.

Lemma 15.20. Sei $\alpha \in K[X] \setminus K$ normiert und $k \in \mathbb{N}$. Dann gilt:

- (a) $\chi_{B(\alpha)} = \mu_{B(\alpha)} = \alpha$.
- (b) $\alpha(B(\alpha^k)) \approx \text{diag}(J_k(0), \dots, J_k(0))$.

Beweis. Sei $\alpha = X^n + a_{n-1}X^{n-1} + \dots + a_0$.

- (a) Sei $B := B(\alpha)$. Für $i = 1, \dots, n-1$ gilt $Be_i = e_{i+1}$. Daher ist $e_1, Be_1, \dots, B^{n-1}e_1$ die Standardbasis von K^n ist und $\deg(\mu_B) \geq \deg(\mu_{e_1}) \geq n$. Andererseits gilt

$$\alpha(B)e_i = B^{i-1}\alpha(B)e_1 = B^{i-1}(Be_n + a_{n-1}e_n + \dots + a_1e_2 + a_0e_1) = 0$$

für $i = 1, \dots, n$. Dies zeigt $\mu_B \mid \alpha$. Insgesamt ist $\mu_B = \alpha$. Nach Cayley-Hamilton ist $\chi_B = \mu_B$ wegen $\deg(\chi_B) = n$.

- (b) Sei $d := \deg(\alpha^k) = k \deg(\alpha) = kn$. Für $A := B(\alpha^k)$ und $N := \alpha(A)$ gilt $N^k = \alpha^k(A) = 0$ nach (a). Nach Satz 14.28 besitzt N eine Jordan-Normalform J bestehend aus Blöcken der Form $J_l(0)$ mit $l \leq k$. Wie in (a) gilt $e_{i+1} = Ae_i$ für $i = 1, \dots, d-1$. Daher sind die Vektoren

$$Ne_i = A^n e_i + a_{n-1}A^{n-1}e_i + \dots + e_i \in e_{n+i} + \langle e_1, \dots, e_{n+i-1} \rangle$$

für $i = 1, \dots, d-n$ linear unabhängig. Dies zeigt $\text{rk}(N) \geq d-n$ und $\dim \text{Ker}(N) \leq n$ nach dem Homomorphiesatz. Somit besitzt J höchstens n Jordanblöcke. Diese müssen folglich alle die Größe $k \times k$ haben. \square

Satz 15.21. Sei $f \in \text{End}(V)$ mit $\chi_f = \mu_f$. Dann existiert eine Basis B von V mit $B[f]_B = B(\mu_f)$.

Beweis. Sei $n := \dim V$. Nach Lemma 15.16 existiert ein $v \in V$ mit $\mu_f = \mu_v$. Für $U := \langle f^i(v) : i \in \mathbb{N}_0 \rangle \leq V$ gilt

$$n = \deg(\chi_f) = \deg(\mu_f) = \deg(\mu_v) \leq \dim U.$$

Also ist $U = V$. Nach Bemerkung 15.15 ist $B := \{v, f(v), \dots, f^{n-1}(v)\}$ eine Basis von V und $B[f]_B = B(\mu_f)$. \square

Satz 15.22. Für jede Matrix $A \in K^{n \times n}$ gilt:

- (a) Es existieren eindeutig bestimmte normierte Polynome $\alpha_1, \dots, \alpha_k \in K[X] \setminus K$ mit $\alpha_k \mid \alpha_{k-1} \mid \dots \mid \alpha_1$ und

$$A \approx \text{diag}(B(\alpha_1), \dots, B(\alpha_k)). \quad (\text{FROBENIUS-Normalform}^1)$$

Dabei ist $\mu_A = \alpha_1$ und $\chi_A = \alpha_1 \dots \alpha_k$.

- (b) Es existieren irreduzible Polynome $\gamma_1, \dots, \gamma_s \in K[X]$ und $a_1, \dots, a_s \in \mathbb{N}$ mit

$$A \approx \text{diag}(B(\gamma_1^{a_1}), \dots, B(\gamma_s^{a_s})). \quad (\text{WEIERSTRASS-Normalform})$$

Dabei sind die Potenzen $\gamma_1^{a_1}, \dots, \gamma_s^{a_s}$ bis auf die Reihenfolge eindeutig bestimmt und $\chi_A = \gamma_1^{a_1} \dots \gamma_s^{a_s}$.

¹Im Englischen: *rational canonical form*

Beweis (JACOB). Sei $V := K^n$ und $f \in \text{End}(V)$ mit $[f] = A$. Nach Lemma 15.16 existiert $v \in V$ mit $\alpha_1 := \mu_v = \mu_f$. Sei $d := \deg(\alpha_1)$ und

$$U := \langle f^i(v) : i \in \mathbb{N}_0 \rangle \leq V.$$

Nach Bemerkung 15.15 ist $B(\alpha_1)$ die Darstellungsmatrix von $f|_U$ bzgl. $\{b_i := f^{i-1}(v) : i = 1, \dots, d\}$. Im Fall $U = V$ sind wir fertig. Sei also $U < V$. Wir ergänzen die b_i zu einer Basis b_1, \dots, b_n von V . Sei b_1^*, \dots, b_n^* die duale Basis von V^* und $f^* \in \text{End}(V^*)$ die zu f duale Abbildung. Nach Lemma 14.9 ist $U^0 \leq V^*$ f^* -invariant. Der zyklische Unterraum

$$L := \langle (f^*)^i(b_d^*) : i \in \mathbb{N}_0 \rangle \leq V^*$$

ist ebenfalls f^* -invariant. Wegen $[f^*] = A^t$ (Satz 7.49) ist $\mu_{b_d^*} | \mu_{f^*} = \mu_f$. Insbesondere ist $\dim L \leq d$. Angenommen es existieren $\lambda_1, \dots, \lambda_t \in K$ mit $t \leq d$, $\lambda_t \neq 0$ und

$$v^* := \sum_{i=1}^t \lambda_i (f^*)^{i-1}(b_d^*) \in L \cap U^0.$$

Nach Bemerkung 7.50 gilt $(f^*)^i = (f^i)^*$ für $i \in \mathbb{N}_0$. Es folgt der Widerspruch

$$0 = v^*(b_{d-t+1}) = b_d^* \left(\sum_{i=1}^t \lambda_i f^{i-1}(b_{d-t+1}) \right) = b_d^*(\lambda_t b_d) = \lambda_t.$$

Daher ist $\{(f^*)^i(b_d^*) : i = 0, \dots, d-1\}$ eine Basis von L und $L \cap U^0 = \{0\}$. Aus $\dim U^0 = n - \dim U = n - d$ folgt $V^* = L \oplus U^0$. Nach Lemma 7.43 ist

$$V = L_0 \oplus U,$$

wobei L_0 nach Lemma 14.9 f -invariant ist. Das Minimalpolynom von $f_1 := f|_{L_0}$ teilt α_1 . Durch Induktion nach n besitzt f_1 eine Frobenius-Normalform $\text{diag}(B(\alpha_2), \dots, B(\alpha_k))$ mit $\alpha_k | \alpha_{k-1} | \dots | \alpha_2 = \mu_{f_1} | \alpha_1$. Insgesamt existiert eine Frobenius-Normalform für f . Aus der Blockdiagonalform und Lemma 15.20 ergibt sich $\chi_A = \alpha_1 \dots \alpha_k$.

Für die Eindeutigkeit konstruieren wir zunächst die Weierstraß-Normalform. Dafür sei $\mu_f = \gamma_1^{c_1} \dots \gamma_l^{c_l}$ die Primfaktorzerlegung, $V_i := \text{Ker}(\gamma_i^{c_i}(f))$ und $f_i := f|_{V_i}$ für $i = 1, \dots, l$. Nach der Primärzerlegung ist

$$V = V_1 \oplus \dots \oplus V_l$$

und $\mu_{f_i} = \gamma_i^{c_i}$. Eine Frobenius-Normalform von f_i hat somit die Form $\text{diag}(B(\gamma_i^{a_1}), \dots, B(\gamma_i^{a_s}))$ mit $1 \leq a_1, \dots, a_s \leq c_i$. Daraus erhält man eine Weierstraß-Normalform für f . Nach Lemma 15.20 ist

$$\text{diag}(\underbrace{J_{a_1}(0), \dots, J_{a_1}(0)}_{\deg(\gamma_i)}, \dots, \underbrace{J_{a_s}(0), \dots, J_{a_s}(0)}_{\deg(\gamma_i)})$$

die Jordan-Normalform von $\gamma_i(f_i)$. Daher sind die Zahlen a_1, \dots, a_s eindeutig bestimmt. Sei umgekehrt $B(\rho)$ mit $\rho \in K[X]$ ein Block einer beliebigen Weierstraß-Normalform W von f . Dann existiert ein f -invarianter Unterraum $U \leq V$, sodass ρ das Minimalpolynom von $f|_U$ ist. Es folgt $\rho | \mu_f$. Nach Folgerung 15.11 ist $\rho | \gamma_i^{c_i}$ für ein i und $U \leq V_i$. Die entsprechenden Blöcke von W liefern also auch eine Weierstraß-Normalform von f_i . Da die Zahlen a_1, \dots, a_s eindeutig bestimmt sind, ist Weierstraß-Normalform von f ebenfalls eindeutig bestimmt (bis auf Reihenfolge der Blöcke).

Wir betrachten schließlich die f -invariante Zerlegung $V = U_1 \oplus \dots \oplus U_k$ einer Frobenius-Normalform von f wie oben. Sei $\alpha_i = \gamma_{i1}^{a_{i1}} \dots \gamma_{is}^{a_{is}}$ die Primfaktorzerlegung des Minimalpolynoms von $f|_{U_i}$. In der Weierstraß-Normalform von $f|_{U_i}$ müssen dann die Blöcke $B(\gamma_{ij}^{a_{ij}})$ mit $j = 1, \dots, s$ auftreten. Da α_i auch das charakteristische Polynom von $f|_{U_i}$ ist (Lemma 15.20), können keine weiteren Blöcke auftreten. Zusammen ergeben alle $B(\gamma_{ij}^{a_{ij}})$ die eindeutige Weierstraß-Normalform von f . Auf diese Weise sind auch $\alpha_1, \dots, \alpha_k$ eindeutig bestimmt (siehe Beispiel 15.24). \square

Bemerkung 15.23.

- (a) Im Gegensatz zur Jordan-Normalform sind die Blöcke $B(\alpha_i)$ der Frobenius-Normalform in einer festen Reihenfolge. Außerdem hängt die Frobenius-Normalform nicht von der Faktorisierung des Minimalpolynoms ab. Insbesondere verändert sich die Frobenius-Normalform nicht, wenn man K durch einen größeren Körper ersetzt (zum Beispiel $\mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$).
- (b) Die Polynome $\alpha_1, \dots, \alpha_k$ in der Frobenius-Normalform lassen sich alternativ durch die Folge $\beta_1 := \alpha_1/\alpha_2, \beta_2 := \alpha_2/\alpha_3, \dots, \beta_k := \alpha_k$ beschreiben. Sind umgekehrt beliebige normierte Polynome β_1, \dots, β_k gegeben, so erhält man die Frobenius-Normalform

$$\text{diag}(\beta_1 \dots \beta_k, \beta_2 \dots \beta_k, \dots, \beta_k) \in K^{n \times n}$$

mit

$$n = \deg(\beta_1) + 2 \deg(\beta_2) + \dots + k \deg(\beta_k).$$

Auf diese Weise lassen sich die Ähnlichkeitsklassen von Matrizen systematisch aufzählen. Möchte man nur invertierbare Matrizen zählen, so müssen alle β_i ein Absolutglied $\neq 0$ haben (anderenfalls wäre X ein Teiler des charakteristischen Polynoms)

- (c) Ist $q := |K| < \infty$, so gibt es genau q^d normierte Polynome vom Grad $d \geq 0$. Unter diesen haben $q^d - q^{d-1}$ ein Absolutglied $\neq 0$. Sei $n = 4$. Mit den Bezeichnungen aus (b) gibt es folgende Möglichkeiten:

$\deg(\beta_1)$	$\deg(\beta_2)$	$\deg(\beta_3)$	$\deg(\beta_4)$	Anzahl	invertierbar
0	0	0	1	q	$q - 1$
1	0	1		q^2	$(q - 1)^2$
2	1			q^3	$(q^2 - q)(q - 1)$
0	2			q^2	$q^2 - q$
4				q^4	$q^4 - q^3$

Insgesamt gibt es $q^4 + q^3 + 2q^2 + q$ Ähnlichkeitsklassen von Matrizen in $K^{4 \times 4}$. Davon bestehen $q^4 - q$ aus invertierbaren Matrizen.

- (d) Ist A nilpotent, so stimmen Jordan-, Frobenius- und Weierstraß-Normalform überein.
- (e) Ein Nachteil der Frobenius-Normalform ist, dass man Begleitmatrizen weniger leicht multiplizieren kann als Jordanblöcke. Ein weiterer Nachteil ist, dass die Frobenius-Normalform F nur dann eine Diagonalmatrix ist, wenn $A = F$ eine Skalarmatrix ist (beachte $\deg(\mu_A) = \deg(\alpha_1) = 1$). Man kann am ersten Block $B(\alpha_1)$ jedoch erkennen, ob A diagonalisierbar ist (Satz 10.52).
- (f) Man beachte, dass die irreduziblen Polynome γ_i in der Weierstraß-Normalform nicht unbedingt verschieden sind. Da diese Polynome aus der Primfaktorzerlegung der α_i entstehen, hängt die Weierstraß-Normalform, im Gegensatz zu Frobenius-Normalform, von K ab (in einem größeren Körper zerfallen die γ_i möglicherweise). Genau dann ist A diagonalisierbar, wenn die Weierstraß-Normalform eine Diagonalmatrix ist. In diesem Fall stimmt die Weierstraß-Normalform also mit der Jordan-Normalform überein.

Beispiel 15.24.

- (a) Seien $\alpha, \beta, \gamma \in K[X]$ irreduzibel. Die Umrechnung zwischen Frobenius-Normalform und Weierstraß-Normalform geht wie folgt:

$$\text{diag}(B(\alpha^2 \beta^3 \gamma), B(\alpha^2 \beta^2), B(\alpha)) \approx \text{diag}(B(\alpha), B(\alpha^2), B(\alpha^2), B(\beta^2), B(\beta^3), B(\gamma)).$$

- (b) Die Berechnung der Frobenius/Weierstraß-Normalform (von Hand) ist naturgemäß sehr aufwendig. Unser Beweis von Lemma 15.16 ist zum Beispiel nicht konstruktiv, aber es gibt entsprechende Algorithmen.² Oft kann man Ad-hoc-Argumente benutzen. Die Matrix

$$A := \begin{pmatrix} 7 & -1 & -3 \\ 30 & -4 & -15 \\ 0 & 0 & 1 \end{pmatrix} \in \mathbb{Q}^{3 \times 3}$$

hat das charakteristische Polynom

$$\chi_A = ((X - 7)(X + 4) + 30)(X - 1) = (X^2 - 3X + 2)(X - 1) = (X - 1)^2(X - 2).$$

Wegen

$$(A - 1_2)(A - 2 \cdot 1_2) = \begin{pmatrix} 6 & -1 & -3 \\ 30 & -5 & -15 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 5 & -1 & -3 \\ 30 & -6 & -15 \\ 0 & 0 & -1 \end{pmatrix} = 0$$

ist $\mu_A = (X - 1)(X - 2)$. Also ist

$$\begin{pmatrix} B(X - 1) & 0 \\ 0 & B(\mu) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -2 \\ 0 & 1 & 3 \end{pmatrix}$$

die Frobenius-Normalform von A . Die Weierstraß-Normalform ist hingegen $\text{diag}(1, 1, 2)$.

- (c) Wir beschreiben die Ähnlichkeitsklassen von Matrizen in $\mathbb{F}_2^{3 \times 3}$ mit der Folge $(\beta_1, \dots, \beta_k)$ aus Bemerkung 15.23:

$$\begin{array}{lll} (1, 1, X) = 0_3, & (1, 1, X + 1) = 1_3, & (X, X), \\ (X, X + 1) \approx \text{diag}(1, 1, 0), & (X + 1, X) \approx \text{diag}(1, 0, 0), & (X + 1, X + 1) = P_{(1,2)}, \\ (X^3) = J_3(0), & (X^3 + 1) = P_{(1,2,3)}, & (X^3 + X), \\ (X^3 + X + 1), & (X^3 + X^2), & (X^3 + X^2 + 1), \\ (X^3 + X^2 + X), & (X^3 + X^2 + X + 1). & \end{array}$$

Also gibt es genau 14 Ähnlichkeitsklassen, von denen sechs zu invertierbaren Matrizen gehören (welche?).

Satz 15.25. Für alle $A \in K^{n \times n}$ gilt $A \approx A^t$.

Beweis. Nach der Frobenius-Normalform kann man $A = B(\alpha)$ für ein normiertes Polynom $\alpha \in K[X] \setminus K$ annehmen. Nach Lemma 15.20 und Aufgabe II.4 gilt $\chi_{A^t} = \chi_A = \mu_A = \mu_{A^t}$. Nach Satz 15.21 ist $A^t \approx A$. \square

Bemerkung 15.26. Einen direkten Beweis von Satz 15.25 findet man in Aufgabe II.27.

²siehe [M. Geck, *On Jacob's construction of the rational canonical form of a matrix*, Electron. J. Linear Algebra 36 (2020), 177–182]

15.3 Zentralisatoren

Definition 15.27. Für $f \in \text{End}(V)$ und $A \in K^{n \times n}$ sei

$$\begin{aligned} C(f) &:= \{g \in \text{End}(V) : f \circ g = g \circ f\} \subseteq \text{End}(V), \\ C(A) &:= \{B \in K^{n \times n} : AB = BA\} \subseteq K^{n \times n} \end{aligned}$$

der *Zentralisator* von f bzw. A .

Bemerkung 15.28. Offenbar sind Zentralisatoren Unterräume. Für $C, D \in C(A)$ ist außerdem $CD \in C(A)$. Für $S \in \text{GL}(n, K)$ gilt $C(SAS^{-1}) = SC(A)S^{-1}$. Daher ist $\dim C(A)$ eine Invariante der Ähnlichkeitsklasse von A . Es gilt

$$\{\alpha(A) : \alpha \in K[X]\} = \langle A^i : i \in \mathbb{N}_0 \rangle \subseteq C(A).$$

Wir untersuchen, wann Gleichheit gilt.

Beispiel 15.29.

- (a) Sei $A \in K^{n \times n}$ eine Diagonalmatrix. Nach Permutation der Basisvektoren können wir $A = \text{diag}(\lambda_1 1_{d_1}, \dots, \lambda_k 1_{d_k})$ mit paarweise verschiedenen $\lambda_1, \dots, \lambda_k \in K$ annehmen. Nach Aufgabe I.17 gilt $C(A) = \{\text{diag}(A_1, \dots, A_k) : \forall i : A_i \in K^{d_i \times d_i}\}$. Insbesondere ist $\dim C(A) = \sum_{i=1}^k d_i^2$. Dies wird in Aufgabe II.30 und Bemerkung 15.33 verallgemeinert.
- (b) Sei $A \in K^{d \times d}$ und $D := \text{diag}(A, \dots, A) \in K^{dk \times dk}$. Sei $B = (B_{ij}) \in K^{dk \times dk}$ mit Blöcken $B_{ij} \in K^{d \times d}$. Dann gilt

$$\begin{pmatrix} AB_{11} & \cdots & AB_{1k} \\ \vdots & & \vdots \\ AB_{k1} & \cdots & AB_{kk} \end{pmatrix} = DB = BD = \begin{pmatrix} B_{11}A & \cdots & B_{1k}A \\ \vdots & & \vdots \\ B_{k1}A & \cdots & B_{kk}A \end{pmatrix} \iff \forall i, j : B_{ij} \in C(A).$$

Insbesondere ist $\dim C(D) = k^2 \dim C(A)$.

- (c) Nach Aufgabe II.22 gilt $\dim C(J_n(\lambda)) = n$ für alle $\lambda \in K$. Dies wird in Folgerung 15.34 verallgemeinert.

Lemma 15.30. Seien V, W Vektorräume und $f \in \text{End}(V)$, $g \in \text{End}(W)$ mit $\chi_f = \mu_f \mid \chi_g = \mu_g$. Dann ist

$$H := \{h \in \text{Hom}(V, W) : h \circ f = g \circ h\}$$

ein Vektorraum mit $\dim H = \dim V$.

Beweis. Sei $d := \dim V$ und $e := \dim W$. Wegen $\chi_f = \mu_f$ und $\chi_g = \mu_g$ existieren $v \in V$ und $w \in W$ mit

$$V = \langle f^i(v) : i = 0, \dots, d-1 \rangle, \quad W = \langle g^i(w) : i = 0, \dots, e-1 \rangle.$$

Sei $h \in H$. Dann existiert genau ein Polynom $\alpha \in K[X]$ mit $h(v) = \alpha(g)(w)$ und $\deg(\alpha) < e$. Für $i = 1, \dots, d-1$ folgt $h(f^i(v)) = g^i(h(v))$. Also ist h durch α eindeutig bestimmt. Nach Lemma 10.15 ist die Abbildung $H \rightarrow K[X]$, $h \mapsto \alpha$ linear und injektiv. Wegen

$$0 = h(\mu_f(f)(v)) = \mu_f(g)(h(v)) = \mu_f(g)(\alpha(g)(w)) = (\alpha\mu_f)(g)(w)$$

gilt außerdem $\mu_g \mid \alpha\mu_f$, d. h. $\tau := \frac{\mu_g}{\mu_f} \mid \alpha$. Sei $P_d \subseteq K[X]$ der Vektorraum aller Polynome vom Grad $< d$. Dann ist

$$\Gamma: H \rightarrow P_d, \quad h \mapsto \alpha/\tau$$

eine injektive lineare Abbildung.

Sei umgekehrt $\beta \in P_d$ gegeben und $\alpha := \beta\tau$. Dann gilt $\mu_g \mid \alpha\mu_f$. Wir definieren $h \in \text{Hom}(V, W)$ durch $h(v) = \alpha(g)(w)$ und $h(f^i(v)) = g^i(h(v))$ für $i = 1, \dots, d-1$. Für $i = 0, \dots, d-2$ gilt dann

$$(h \circ f)(f^i(v)) = g^{i+1}(h(v)) = g(g^i(h(v))) = (g \circ h)(f^i(v)).$$

Sei $\mu_f = X^d + a_{d-1}X^{d-1} + \dots + a_0$. Wegen $\mu_f(f) = 0$ gilt

$$\begin{aligned} (h \circ f)(f^{d-1}(v)) &= h(f^d(v)) = h(-a_{d-1}f^{d-1}(v) - \dots - a_0v) = -a_{d-1}g^{d-1}(h(v)) - \dots - a_0h(v) \\ &= g^d(h(v)) - \mu_f(g)(h(v)) = g^d(h(v)) - (\mu_f\alpha)(g)(w) = g^d(h(v)) = (g \circ h)(f^{d-1}(v)). \end{aligned}$$

Dies zeigt $h \in H$ mit $\Gamma(h) = \beta$. Also ist Γ ein Isomorphismus und $\dim H = \dim P_d = d = \dim V$. \square

Beispiel 15.31. Sei $V = K$ und $f = \lambda \text{id}_V$ für ein $\lambda \in K$. Für $h \in H$ gilt $\lambda h(1) = h(f(1)) = g(h(1))$, d. h. $h(1)$ ist ein Eigenvektor von g zum Eigenwert λ . Außerdem ist $H \rightarrow E_\lambda(g)$, $h \mapsto h(1)$ ein Isomorphismus. Wegen $\dim H = 1$ hat λ geometrische Vielfachheit 1. Das kann man natürlich auch direkt aus der Bedingung $\chi_f = X - \lambda \mid \chi_g = \mu_g$ ableiten.

Satz 15.32 (FROBENIUS). Sei $A \in K^{n \times n}$ mit Frobenius-Normalform $\text{diag}(B(\alpha_1), \dots, B(\alpha_k))$. Dann gilt

$$\dim C(A) = \sum_{i=1}^k (2i-1) \deg(\alpha_i).$$

Insbesondere ist $\dim C(A) \geq n$ mit Gleichheit genau dann, wenn $\chi_A = \mu_A$.

Beweis. Sei $d_i := \deg(\alpha_i)$ und $A_i := B(\alpha_i) \in K^{d_i \times d_i}$ für $i = 1, \dots, k$. Nach Bemerkung 15.28 können wir $A = \text{diag}(A_1, \dots, A_k)$ annehmen. Sei $C = (C_{ij}) \in K^{n \times n}$ eine Blockmatrix mit $C_{ij} \in K^{d_i \times d_j}$ für $1 \leq i, j \leq k$. Dann gilt

$$C \in C(A) \iff CA = AC \iff \forall i, j : C_{ij}A_j = A_iC_{ij}.$$

Für $i \leq j$ gilt $\chi_{A_j} = \mu_{A_j} = \alpha_j \mid \alpha_i = \chi_{A_i} = \mu_{A_i}$ nach Lemma 15.20. Nach Lemma 15.30 gibt es d_j linear unabhängige Möglichkeiten für die Wahl von C_{ij} . Im Fall $i > j$ können wir

$$C_{ij}^t A_i^t = (A_i C_{ij})^t = (C_{ij} A_j)^t = A_j^t C_{ij}^t$$

betrachten. Wegen $\chi_{A_i^t} = \alpha_i = \mu_{A_i^t}$ (Aufgabe II.4) gibt es d_i linear unabhängige Möglichkeiten für C_{ij}^t (und damit auch für C_{ij}). Da die Blöcke C_{ij} in C unabhängig voneinander gewählt werden können, erhält man

$$\dim C(A) = \sum_{i,j=1}^k \min\{d_i, d_j\} = d_1 + 3d_2 + 5d_3 + \dots + (2k-1)d_k \geq d_1 + \dots + d_k = n$$

mit Gleichheit genau dann, wenn $k = 1$ und $\mu_A = \alpha_1 = \chi_A$ (Satz 15.22). \square

Bemerkung 15.33. In Bemerkung 15.23 haben wir die Frobenius-Normalform durch die Polynome $\beta_1 := \alpha_1/\alpha_2$, $\beta_2 := \alpha_2/\alpha_3, \dots, \beta_k := \alpha_k$ beschrieben. Es gilt $\deg(\alpha_i) = \deg(\beta_i\beta_{i+1}\dots\beta_k) = \sum_{j=i}^k \deg(\beta_j)$ und $\sum_{i=1}^m (2i-1) = m^2$ nach Aufgabe I.4. Dies zeigt

$$\dim C(A) = \sum_{l=1}^k l^2 \deg(\beta_l).$$

Folgerung 15.34. Sei $A \in K^{n \times n}$ mit $\chi_A = \mu_A$. Dann gilt $C(A) = \langle A^i : i = 0, \dots, n-1 \rangle$.

Beweis. Aus $\deg(\mu_A) = \deg(\chi_A) = n$ folgt $\dim C(A) \geq \dim \langle A^i : i \in \mathbb{N}_0 \rangle = n$. Nach Frobenius ist andererseits $\dim C(A) = n$. \square

15.4 Zerfällungskörper

Satz 15.35. Für $n, k \in \mathbb{N}$ und $\lambda \in \mathbb{C}$ existieren genau k Matrizen $W \in \mathbb{C}^{n \times n}$ mit $W^k = J_n(\lambda)$.

Beweis. Sei $J := J_n(\lambda)$. Nach Satz 14.42 existiert zumindest eine k -te Wurzel W mit $W^k = J$. Für jede k -te Einheitswurzel $\zeta \in \mathbb{C}$ gilt auch $(\zeta W)^k = J$. Nach Lemma 11.27 existieren also mindestens k Wurzeln von J . Nach Folgerung 15.34 gilt $W \in C(J) = \langle J^i : i \in \mathbb{N}_0 \rangle$. Insbesondere ist W eine untere Dreiecksmatrix mit Diagonale (μ, \dots, μ) für eine k -te Einheitswurzel μ (vgl. Aufgabe II.22). Indem man W durch $\mu^{-1}W$ ersetzt, kann man $\mu = 1$ annehmen.

Sei umgekehrt $A \in \mathbb{C}^{n \times n}$ mit $A^k = J$. Dann ist auch A eine untere Dreiecksmatrix mit Diagonale (ζ, \dots, ζ) für eine k -te Einheitswurzel ζ . Da $B := \zeta W \in C(J)$ ein Polynom in J ist, sind A und B vertauschbar. Es folgt

$$0 = A^k - B^k = (A - B) \underbrace{(A^{k-1} + A^{k-2}B + \dots + AB^{k-2} + B^{k-1})}_{=: C}.$$

Offenbar ist C eine untere Dreiecksmatrix mit Hauptdiagonale $k\zeta^{k-1}(1, \dots, 1) \neq 0$. Insbesondere ist C invertierbar und $A - B = 0 \cdot C^{-1} = 0$, d. h. $A = B$. \square

Satz 15.36 (SCHURS Lemma). Sei $f \in \text{End}(V)$. Genau dann ist χ_f irreduzibel, wenn $\{0\}$ und V die einzigen f -invarianten Unterräume sind.

Beweis. Seien $\{0\}$ und V die einzigen f -invarianten Unterräume von V . Nach der Frobenius-Normalform ist $\chi_f = \mu_f$. Sei γ ein irreduzibler Teiler von μ_f . Dann ist $U := \text{Ker}(\gamma(f)) \leq V$ ein f -invarianter Unterraum. Für $v \in V$ mit $\mu_v = \mu_f$ (Lemma 15.16) gilt $0 \neq \frac{\mu_f}{\gamma}(f)(v) \in U$. Dies zeigt $U = V$ und $\chi_f = \mu_f = \gamma$ ist irreduzibel.

Sei umgekehrt χ_f irreduzibel und $U < V$ f -invariant. Wir ergänzen eine Basis B_1 von U zu einer Basis B von V . Dann gilt

$${}_B[f]_B = \begin{pmatrix} {}_{B_1}[f]_{B_1} & * \\ 0 & * \end{pmatrix}.$$

Dies zeigt $\chi_{f|_U} \mid \chi_f$ und $\chi_{f|_U} \in K$. Es folgt $U = \{0\}$. \square

Satz 15.37. Für jedes irreduzible Polynom $\gamma \in K[X]$ ist $L := C(B(\gamma))$ ein Körper mit $\dim_K L = \deg(\gamma)$.

Beweis. Sei $n := \deg(\gamma)$, $V = K^n$ und $f \in \text{End}(V)$ mit $A := B(\gamma) = [f]$. Nach Lemma 15.20 ist $\chi_f = \mu_f = \gamma$. Nach Folgerung 15.34 ist $L = \langle A^i : i \in \mathbb{N}_0 \rangle \subseteq K^{n \times n}$. Insbesondere ist die (Matrizen-) Multiplikation in L kommutativ. Für $g \in C(f) \setminus \{0\}$ ist $\text{Ker}(g) \leq V$ wie üblich ein f -invarianter Unterraum. Nach Schurs Lemma gilt $\text{Ker}(g) = \{0\}$ und $g \in \text{GL}(V)$. Mit g ist auch $g^{-1} \in C(f)$. Also besitzt jedes nicht-triviale Element in L ein Inverses bzgl. Multiplikation. Die verbleibenden Körperaxiome für $C(A)$ folgen aus den Rechenregeln in $K^{n \times n}$. Nach Folgerung 15.34 ist $\dim L = n$. \square

Beispiel 15.38.

- (a) Bekanntlich ist $\gamma := X^2 + 1 \in \mathbb{R}[X]$ irreduzibel. Sei $A := B(\gamma) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathbb{R}^{2 \times 2}$. Nach Satz 15.37 ist $C(A) = \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} : a, b \in \mathbb{R} \right\}$ ein Körper. Tatsächlich stimmt $C(A)$ mit der Konstruktion von \mathbb{C} im Beweis von Lemma 11.24 überein.
- (b) Für $\gamma = X^2 - 2 \in \mathbb{Q}[X]$ ist $C(B(\gamma)) = \mathbb{Q}(\sqrt{2})$ der Körper aus Aufgabe I.15.
- (c) Für $\gamma := X^2 + X + 1 \in \mathbb{F}_2[X]$ ist

$$C(B(\gamma)) = \left\{ 0_2, 1_2, \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \right\}$$

ein Körper mit vier Elementen.

Satz 15.39. *Für jedes normierte Polynom $\alpha \in K[X]$ existiert ein Körper $L \supseteq K$, sodass α in $L[X]$ in Linearfaktoren zerfällt.*

Beweis. Im Fall $\deg(\alpha) = 1$ ist α selbst ein Linearfaktor. Sei also $\deg(\alpha) \geq 2$. Sei γ ein irreduzibler Teiler von α mit $d := \deg(\gamma)$. Nach Satz 15.37 ist $L_1 := C(B(\gamma_1))$ ein Körper. Wir können K mit den Skalarmatrizen $\{\lambda 1_d \in L_1 : \lambda \in K\}$ identifizieren (die Rechenregeln für Skalarmatrizen entsprechen denen in K). Für $x := B(\gamma) \in L_1$ gilt $\gamma(x) = 0$, d. h. x ist eine Nullstelle von γ . Nach Lemma 10.22 gilt $\alpha = (X - x)\beta$ für ein $\beta \in L_1[X]$ mit $\deg(\beta) = \deg(\alpha) - 1$. Durch Induktion nach $\deg(\alpha)$ existiert ein Körper $L \supseteq L_1$, in dem β in Linearfaktoren zerfällt. Offenbar zerfällt auch α in $L[X]$ in Linearfaktoren. \square

Definition 15.40. In der Situation von Satz 15.39 nennt man L einen *Zerfällungskörper* von α (nicht eindeutig bestimmt).

16 Die Jordan-Chevalley-Zerlegung

16.1 Der chinesische Restsatz

Bemerkung 16.1. Wir erweitern unsere Kenntnisse über Polynome. Mit diesem Werkzeug zerlegen wir eine Matrix auf eindeutige Weise in einen diagonalisierbaren Teil (über einem Zerfällungskörper) und einen nilpotenten Teil.

Definition 16.2. Seien $\alpha, \beta, \delta \in K[X]$ mit $\delta \neq 0$. Wir sagen α ist kongruent zu β modulo δ , falls $\delta \mid \alpha - \beta$. Ggf. schreiben wir $\alpha \equiv \beta \pmod{\delta}$.

Lemma 16.3. Die Kongruenz modulo $\delta \in K[X] \setminus \{0\}$ ist eine Äquivalenzrelation auf $K[X]$. Für $\alpha_i, \beta_i \in K[X]$ mit $\alpha_i \equiv \beta_i \pmod{\delta}$ ($i = 1, 2$) gilt $\alpha_1 + \alpha_2 \equiv \beta_1 + \beta_2 \pmod{\delta}$ und $\alpha_1 \alpha_2 \equiv \beta_1 \beta_2 \pmod{\delta}$.

Beweis. Wegen $\delta \mid 0 = \alpha - \alpha$ ist $\alpha \equiv \alpha \pmod{\delta}$. Aus $\alpha \equiv \beta \pmod{\delta}$ folgt $\beta \equiv \alpha \pmod{\delta}$. Sei nun $\alpha \equiv \beta \pmod{\delta}$ und $\beta \equiv \gamma \pmod{\delta}$. Dann gilt $\delta \mid (\alpha - \beta) + (\beta - \gamma) = \alpha - \gamma$ und $\alpha \equiv \gamma \pmod{\delta}$. Daher ist \equiv eine Äquivalenzrelation. Aus $\alpha_i \equiv \beta_i \pmod{\delta}$ folgt

$$\begin{aligned}\delta \mid (\alpha_1 - \beta_1) + (\alpha_2 - \beta_2) &= (\alpha_1 + \alpha_2) - (\beta_1 + \beta_2), \\ \delta \mid (\alpha_1 - \beta_1)\alpha_2 + (\alpha_2 - \beta_2)\beta_1 &= \alpha_1\alpha_2 - \beta_1\beta_2.\end{aligned}$$

Dies zeigt die zweite Behauptung. □

Beispiel 16.4.

- (a) Für $\delta \in K^\times$ ist $\alpha \equiv \beta \pmod{\delta}$ für alle $\alpha, \beta \in K[X]$, d. h. \equiv ist die triviale Relation.
- (b) Es gilt $\alpha \equiv \beta \pmod{X}$ genau dann, wenn α und β das gleiche Absolutglied haben.
- (c) Für alle $\alpha \in K[X]$ existiert ein $\beta \in K[X]$ mit $\alpha \equiv \beta \pmod{\delta}$ und $\deg(\beta) < \deg(\delta)$. Dies ergibt sich aus der Division mit Rest.
- (d) Lemma 16.3 vereinfacht viele Teilbarkeitsbetrachtungen: Um zu prüfen, ob

$$\alpha = (X^2 + X + 2)^4 + (X^5 - 1)(X^3 + X) \in \mathbb{Q}[X]$$

durch $\delta = X + 1 \in \mathbb{Q}[X]$ teilbar ist, kann man die Kongruenz mit jedem einzelnen Term durchführen. Wegen

$$\begin{aligned}X^2 + X + 2 &= X(X + 1) + 2 \equiv 2 \pmod{\delta}, \\ X^5 - 1 &= (X + 1)(X^4 - X^3 + X^2 - X + 1) - 2 \equiv -2 \pmod{\delta}, \\ X^3 + X + 10 &= (X + 1)(X^2 - X + 2) + 8 \equiv 8 \pmod{\delta}\end{aligned}$$

gilt $\alpha \equiv 2^4 + (-2)8 \equiv 0 \pmod{\delta}$, d. h. $\delta \mid \alpha$.

Satz 16.5 (Chinesischer Restsatz). Seien $\alpha_1, \dots, \alpha_n \in K[X] \setminus \{0\}$ paarweise teilerfremd und $\beta_1, \dots, \beta_n \in K[X]$ beliebig. Dann existiert ein $\gamma \in K[X]$ mit $\gamma \equiv \beta_i \pmod{\alpha_i}$ für $i = 1, \dots, n$.

Beweis. Für $i = 1, \dots, n$ sei $\alpha'_i := \prod_{j \neq i} \alpha_j$. Nach Folgerung 15.11 muss jeder irreduzible Teiler von α'_i ein α_j mit $j \neq i$ teilen. Da α_i und α_j teilerfremd sind, müssen auch α_i und α'_i teilerfremd sein. Nach Bézout existieren $\gamma_1, \dots, \gamma_n \in K[X]$ mit $\alpha'_i \gamma_i \equiv 1 \pmod{\alpha_i}$. Sei

$$\gamma := \sum_{i=1}^n \alpha'_i \beta_i \gamma_i.$$

Dann gilt $\gamma \equiv \alpha'_i \beta_i \gamma_i \equiv \beta_i \pmod{\alpha_i}$ für $i = 1, \dots, n$. □

Bemerkung 16.6. Mit γ erfüllt auch $\gamma + \rho \prod_{i=1}^n \alpha_i$ für alle $\rho \in K[X]$ die Kongruenzen aus dem chinesischen Restsatz. Daher gibt es unendlich viele Lösungen.

Beispiel 16.7. Da die Polynome $X^2 - 2, X^2 + 1 \in \mathbb{Q}[X]$ keine gemeinsamen Nullstellen haben, sind sie teilerfremd. Gesucht sei $\gamma \in \mathbb{Q}[X]$ mit

$$\gamma \equiv X \pmod{X^2 - 2}, \quad \gamma \equiv 3 \pmod{X^2 + 1}.$$

Bézout liefert $\frac{1}{3}(X^2 + 1) \equiv 1 \pmod{X^2 - 2}$ und $-\frac{1}{3}(X^2 - 2) \equiv 1 \pmod{X^2 + 1}$ (im Zweifel muss man den euklidischen Algorithmus bemühen). Wie im Beweis von Satz 16.5 ist

$$\gamma = \frac{1}{3}(X^2 + 1)X - \frac{1}{3}(X^2 - 2)3 = \frac{1}{3}X^3 - X^2 + \frac{1}{3}X + 2$$

eine Lösung der Kongruenzen.

Definition 16.8.

- Für $\alpha = \sum_{k=0}^{\infty} a_k X^k \in K[X]$ definieren wir die (formale) *Ableitung*

$$\alpha' = \sum_{k=1}^{\infty} k a_k X^{k-1}$$

von α wie in der Analysis.

- Ein irreduzibles Polynom α heißt *separabel*, wenn $\alpha' \neq 0$. Ein beliebiges Polynom α heißt *separabel*, wenn alle irreduziblen Teiler von α separabel sind (das Gegenteil ist *inseparabel*).

Beispiel 16.9.

- Die konstanten Polynome sind separabel, da sie keine irreduziblen Teiler haben.
- Ist die Abbildung $\mathbb{N} \rightarrow K, n \mapsto n \cdot 1_K$ injektiv, so gilt $\alpha' \neq 0$ für alle $\alpha \in K[X] \setminus K$. Insbesondere ist jedes Polynom in $\mathbb{C}[X]$ separabel.

- (c) Über endlichen Körpern ist die Bedingung $\alpha' \neq 0$ weniger offensichtlich. Zum Beispiel gilt $(X^2)' = 0$ in $\mathbb{F}_2[X]$. Sei allgemein $\alpha = \sum a_k X^k \in \mathbb{F}_2[X]$ mit $\alpha' = \sum k a_k X^{k-1} = 0$. Für ungerade k folgt $a_k X^{k-1} = k a_k X^{k-1} = 0$, also $a_k = 0$. Dies zeigt

$$\alpha = a_0 + a_2 X^2 + \dots + a_{2n} X^{2n}$$

für ein $n \in \mathbb{N}$. Wegen $(x+y)^2 = x^2 + 2xy + y^2 = x^2 + y^2$ in \mathbb{F}_2 erhält man induktiv

$$\alpha = (a_0 + a_2 X + \dots + a_{2n} X^n)^2.$$

Insbesondere ist α reduzibel (d. h. nicht irreduzibel). Damit ist jedes Polynom in $\mathbb{F}_2[X]$ separabel.

- (d) Man konstruiert \mathbb{Q} aus \mathbb{Z} , indem man Brüche einführt. Auf die gleiche Weise kann man den Körper der *rationalen Funktionen*

$$K(X) := \left\{ \frac{\alpha}{\beta} : \alpha, \beta \in K[X], \beta \neq 0 \right\}$$

aus $K[X]$ gewinnen, indem man Brüche von Polynomen einführt. Sei speziell $K := \mathbb{F}_2(X)$. Man kann zeigen, dass $\alpha := X^2 + Y \in K[Y]$ irreduzibel und inseparabel ist.

Lemma 16.10. Für $\alpha, \beta \in K[X]$ gilt

$$(\alpha + \beta)' = \alpha' + \beta', \quad (\text{Summenregel})$$

$$(\alpha\beta)' = \alpha'\beta + \alpha\beta'. \quad (\text{Produktregel})$$

Beweis. Für $\alpha = \sum a_k X^k$ und $\beta = \sum b_k X^k$ ist

$$\begin{aligned} (\alpha + \beta)' &= \left(\sum (a_k + b_k) X^k \right)' = \sum k(a_k + b_k) X^{k-1} \\ &= \sum k a_k X^{k-1} + \sum k b_k X^{k-1} = \alpha' + \beta'. \end{aligned}$$

Daraus folgt

$$\begin{aligned} (\alpha\beta)' &= \left(\sum_{k,l=0}^{\infty} a_k b_l X^{k+l} \right)' = \sum_{k,l} a_k b_l (X^{k+l})' = \sum_{k,l} (k+l) a_k b_l X^{k+l-1} \\ &= \sum_{k,l} k a_k b_l X^{k-1+l} + \sum_{k,l} l a_k b_l X^{l-1+k} = \alpha'\beta + \alpha\beta'. \quad \square \end{aligned}$$

Lemma 16.11. Sei $\gamma \in K[X]$ irreduzibel. Genau dann ist γ separabel, wenn γ keine mehrfachen Nullstellen in einem Zerfällungskörper besitzt.

Beweis. Sei $\lambda \in L$ eine Nullstelle von γ in einem Zerfällungskörper L . Dann existiert ein $\rho \in L[X]$ mit $\gamma = (X - \lambda)\rho$. Aus der Produktregel folgt

$$\gamma' = \rho + (X - \lambda)\rho'$$

und $\gamma'(\lambda) = \rho(\lambda)$. Ist γ inseparabel, so ist $\rho(\lambda) = \gamma'(\lambda) = 0$, d. h. λ ist eine mehrfache Nullstelle. Sei umgekehrt γ separabel, d. h. $\gamma' \neq 0$. Dann gilt $0 \leq \deg(\gamma') < \deg(\gamma)$. Da γ irreduzibel ist, sind γ' und γ teilerfremd. Nach Bézout existieren $\alpha, \beta \in K[X]$ mit $\alpha\gamma + \beta\gamma' = 1$. Es folgt $\beta(\lambda)\gamma'(\lambda) = 1$ und $\rho(\lambda) = \gamma'(\lambda) \neq 0$. Also ist λ eine einfache Nullstelle. Da λ beliebig war, sind alle Nullstellen einfach. \square

16.2 Separable und halbeinfache Abbildungen

Definition 16.12. Wir nennen $f \in \text{End}(V)$ (bzw. $A \in K^{n \times n}$)

- *separabel*, falls μ_f (bzw. μ_A) separabel ist.
- *halbeinfach*, falls μ_f (bzw. μ_A) in paarweise verschiedene irreduzible Polynome zerfällt.

Beispiel 16.13.

- (a) Über \mathbb{Q} , \mathbb{R} , \mathbb{C} und \mathbb{F}_2 sind nach Beispiel 16.9 alle Endomorphismen separabel.
- (b) Sei $K = \mathbb{C}$. Nach dem Fundamentalsatz der Algebra und Satz 10.52 ist $f \in \text{End}(V)$ genau dann halbeinfach, wenn f diagonalisierbar. Dies wird in Lemma 16.15 verallgemeinert. Andererseits sind $B(X^2 - 2) \in \mathbb{Q}^{2 \times 2}$ und $B(X^2 + X + 1) \in \mathbb{F}_2^{2 \times 2}$ halbeinfach, aber nicht diagonalisierbar.

Satz 16.14. *Genau dann ist $f \in \text{End}(V)$ halbeinfach, wenn jeder f -invariante Unterraum $U \leq V$ ein f -invariantes Komplement besitzt.*

Beweis. Sei f halbeinfach und $\mu_f = \gamma_1 \dots \gamma_k$ mit paarweise verschiedenen irreduziblen Polynomen $\gamma_1, \dots, \gamma_k$. Die Weierstraß-Normalform von f liefert eine Zerlegung $V = V_1 \oplus \dots \oplus V_s$ in f -invariante Unterräume, sodass die Darstellungsmatrix von f auf V_i durch $B(\gamma_j)$ für ein $1 \leq j \leq k$ gegeben ist (beachte $k \leq s$). Nach Lemma 15.20 und Schurs Lemma besitzt V_i keine echten nicht-trivialen f -invarianten Unterräume für $i = 1, \dots, s$. Sei $U \leq V$ f -invariant. Sei $W \leq V$ f -invariant mit $U \cap W = \{0\}$, sodass $\dim W$ möglichst groß ist (notfalls $W = \{0\}$). Angenommen es gilt $U \oplus W < V$. Dann existiert ein i mit $V_i \not\subseteq U + W$. Offenbar ist

$$L := V_i \cap (U + W) < V_i$$

f -invariant. Aus Schurs Lemma folgt $L = \{0\}$. Sei $u = w + v \in U \cap (W + V_i)$ mit $w \in W$ und $v \in V_i$. Dann gilt $v = u - w \in V_i \cap (U + W) = \{0\}$ und $u = w \in U \cap W = \{0\}$. Also ist $U \cap (W + V_i) = \{0\}$ im Widerspruch zur Wahl von W . Dies zeigt $V = U \oplus W$.

Nehmen wir umgekehrt an, dass jeder f -invariante Unterraum von V ein f -invariantes Komplement besitzt. Wir nehmen indirekt $\mu_f = \gamma^2 \delta$ für ein irreduzibles Polynom γ an. Nach Voraussetzung besitzt $U := \text{Ker}(\gamma(f)) \leq V$ ein f -invariantes Komplement W . Für $w \in W$ gilt

$$(\gamma \delta)(f)(w) \in U \cap W = \{0\}.$$

Wegen $(\gamma \delta)(f)(U) = \{0\}$ gilt sogar $(\gamma \delta)(f)(v) = 0$ für alle $v \in V$. Dann wäre aber $\mu_f \mid \gamma \delta$. Widerspruch. \square

Lemma 16.15. *Sei $A \in K^{n \times n}$ separabel. Genau dann ist A halbeinfach, wenn ein Körper $L \supseteq K$ existiert, sodass A in $L^{n \times n}$ diagonalisierbar ist.*

Beweis. Sei A halbeinfach und $\mu_A = \gamma_1 \dots \gamma_k$ mit paarweise verschiedenen (separablen) irreduziblen Polynomen $\gamma_1, \dots, \gamma_k$. Sei L ein Zerfällungskörper von μ_A . Nach Lemma 16.11 hat jedes γ_i keine mehrfachen Nullstellen in L . Für $i \neq j$ existieren $\alpha, \beta \in K[X]$ mit $\alpha \gamma_i + \beta \gamma_j = 1$ nach Bézout. Daher haben γ_i und γ_j keine gemeinsamen Nullstellen in L . Insgesamt zerfällt μ_A in paarweise verschiedene Linearfaktoren in $L[X]$. Nach Satz 10.52 ist A in $L^{n \times n}$ diagonalisierbar. Ist A nicht halbeinfach, so hat μ_A offensichtlich mehrfache Nullstellen in $L[X]$. Damit kann A nicht diagonalisierbar sein. \square

Lemma 16.16. Sind $A, B \in K^{n \times n}$ vertauschbar, separabel und halbeinfach, so ist auch $A + B$ halbeinfach.

Beweis. Sei L ein Zerfällungskörper von $\mu_A \mu_B$. Nach Lemma 16.15 und Lemma 14.11 sind A und B in $L^{n \times n}$ simultan diagonalisierbar. Also ist auch $A + B$ in $L^{n \times n}$ diagonalisierbar. Nach Satz 10.52 hat μ_{A+B} keine mehrfachen Nullstellen in L . Nach Lemma 16.11 ist $A + B$ separabel und nach Lemma 16.15 halbeinfach. \square

16.3 Verallgemeinerte Jordanblöcke

Bemerkung 16.17. Ist $\gamma \in K[X]$ irreduzibel, so kann man $\lambda := B(\gamma)$ als Element des Körpers $L := C(B(\gamma))$ auffassen. Im folgenden Lemma studieren wir den Jordanblock $J_k(\lambda) \in L^{k \times k}$ als Matrix in $K^{n \times n}$, wobei $n := k \deg(\gamma)$. Für $\gamma = X - \mu$ erhält man $J_k(\gamma) = J_k(\mu)$.

Satz 16.18 (Verallgemeinerter Jordanblock). Sei $\gamma \in K[X]$ irreduzibel und separabel vom Grad d und $k \in \mathbb{N}$. Dann gilt

$$J_k(\gamma) := \begin{pmatrix} B(\gamma) & & & 0 \\ 1_d & \ddots & & \\ & \ddots & \ddots & \\ 0 & & 1_d & B(\gamma) \end{pmatrix} \approx B(\gamma^k).$$

Beweis. Sei $J := J_k(\gamma)$. Eine einfache Induktion zeigt

$$J^m = \begin{pmatrix} B(\gamma)^m & & & 0 \\ mB(\gamma)^{m-1} & \ddots & & \\ & \ddots & \ddots & \\ * & & mB(\gamma)^{m-1} & B(\gamma)^m \end{pmatrix}$$

für $m \in \mathbb{N}_0$ (vgl. Satz 14.42). Es folgt

$$\gamma(J) = \begin{pmatrix} 0 & & & 0 \\ \gamma'(B(\gamma)) & \ddots & & \\ & \ddots & \ddots & \\ * & & \gamma'(B(\gamma)) & 0 \end{pmatrix}.$$

Insbesondere ist $\gamma(J)$ nilpotent und $\mu_J \mid \gamma^k$. Eine weitere Induktion zeigt

$$\gamma(J)^2 = \begin{pmatrix} 0 & & & 0 \\ 0 & \ddots & & \\ \gamma'(B(\gamma))^2 & \ddots & \ddots & \\ & \ddots & \ddots & \\ * & & \gamma'(B(\gamma))^2 & 0 & 0 \end{pmatrix}, \dots, \gamma(J)^{k-1} = \begin{pmatrix} 0 & & & 0 \\ \vdots & \ddots & & \\ 0 & & \ddots & \\ \gamma'(B(\gamma))^{k-1} & 0 & \dots & 0 \end{pmatrix}.$$

Da γ separabel ist, gilt $0 \leq \deg(\gamma') < \deg(\gamma)$. Da γ irreduzibel ist, müssen γ und γ' teilerfremd sind. Nach Bézout existieren $\alpha, \beta \in K[X]$ mit $\alpha\gamma + \beta\gamma' = 1$. Es folgt

$$1 = \alpha(B(\gamma))\gamma(B(\gamma)) + \beta(B(\gamma))\gamma'(B(\gamma)) = \beta(B(\gamma))\gamma'(B(\gamma)).$$

Also ist $\gamma'(B(\gamma))$ invertierbar und $\gamma'(B(\gamma))^{k-1} \neq 0$. Dies zeigt $\mu_J = \gamma^k$. Die Behauptung folgt aus Satz 15.21. \square

Satz 16.19 (JORDAN-CHEVALLEY-Zerlegung). *Für jede separable Matrix $A \in K^{n \times n}$ existieren eindeutig bestimmte Matrizen $D, N \in K^{n \times n}$ mit folgenden Eigenschaften:*

- (a) $A = D + N$ und $DN = ND$.
- (b) D ist halbeinfach und N nilpotent.

Ggf. existiert ein $\alpha \in K[X]$ mit $\alpha(A) = D$.

Beweis. Wir transformieren A zunächst in die Weierstraß-Normalform. Da μ_A separabel ist, können wir anschließend jeden Block $B(\gamma^a)$ in der Weierstraß-Normalform zu $J_a(\gamma)$ umformen (Satz 16.18). Sei also $S \in \text{GL}(n, K)$ mit

$$W := SAS^{-1} = \text{diag}(J_{a_1}(\gamma_1), \dots, J_{a_s}(\gamma_s))$$

Für $i = 1, \dots, s$ sei $d_i := \deg(\gamma_i)$ und

$$\begin{aligned} D_i &:= \text{diag}(B(\gamma_i), \dots, B(\gamma_i)) \in K^{a_i d_i \times a_i d_i}, & N_i &:= J_{a_i}(\gamma_i) - D_i, \\ \check{D} &:= \text{diag}(D_1, \dots, D_s), & \check{N} &:= \text{diag}(N_1, \dots, N_s). \end{aligned}$$

Dann gilt $W = \check{D} + \check{N}$. O. B. d. A. seien $\gamma_1, \dots, \gamma_k$ die *verschiedenen* Primteiler von μ_A . Für $1 \leq i \leq k$ und $1 \leq j \leq s$ mit $\gamma_i = \gamma_j$ sei außerdem $a_i \geq a_j$. Nach Lemma 15.20 ist $\mu_{D_i} = \gamma_i$ und $\mu_{\check{D}} = \gamma_1 \dots \gamma_k$, d. h. \check{D} ist halbeinfach. Andererseits ist \check{N} eine strikte untere Dreiecksmatrix und daher nilpotent (Beispiel 14.26). Eine Rechnung wie in Satz 16.18 zeigt $D_i N_i = N_i D_i$ und $\check{D} \check{N} = \check{D} \check{N}$. Da das Minimalpolynom nicht von der Basiswahl abhängt, ist $D := S^{-1} \check{D} S$ halbeinfach und $N := S^{-1} \check{N} S$ nilpotent. Außerdem gilt $DN = ND$ und $A = S^{-1} W S = D + N$.

Wegen $J_i := J_{a_i}(\gamma_i) = D_i + N_i$ ist $D_i \in C(J_i)$. Nach Satz 16.18 und Folgerung 15.34 existiert ein $\alpha_i \in K[X]$ mit $\alpha_i(J_i) = D_i$ für $i = 1, \dots, k$. Sei $1 \leq j \leq s$ mit $\gamma_j = \gamma_i$. Dann ist $a_j \leq a_i$ und $J_i = \begin{pmatrix} J_j & 0 \\ * & * \end{pmatrix}$. Es folgt

$$\begin{pmatrix} D_j & 0 \\ * & * \end{pmatrix} = D_i = \alpha_i(J_i) = \begin{pmatrix} \alpha_i(J_j) & 0 \\ * & * \end{pmatrix}$$

und $\alpha_i(J_j) = D_j$. Nach den chinesischen Restsatz existiert ein $\alpha \in K[X]$ mit $\alpha \equiv \alpha_i \pmod{\gamma_i^{a_i}}$ für $i = 1, \dots, k$. Aus $\gamma_i^{a_i}(J_i) = 0$ folgt

$$\alpha(W) = \text{diag}(\alpha_1(J_1), \dots, \alpha_s(J_s)) = \text{diag}(D_1, \dots, D_s) = \check{D}$$

und $\alpha(A) = S^{-1} \alpha(W) S = D$. Damit ist die Existenzaussage bewiesen. Da μ_D die gleichen Primteiler wie μ_A hat, ist D separabel.

Sei nun $A = \check{D} + \check{N}$ mit den gleichen Eigenschaften. Wir zeigen zunächst, dass \check{D} separabel ist. Über einem Zerfällungskörper sind \check{D} und \check{N} nach Satz 14.8 und Lemma 14.12 simultan trigonalisierbar. Bzgl. einer geeigneten Basis ist dann A eine Dreiecksmatrix. Als nilpotente Matrix muss \check{N} bzgl. dieser Basis eine strikte Dreiecksmatrix sein. Daher haben A und \check{D} die gleiche Hauptdiagonale und damit das gleiche charakteristische Polynom (beachte: χ_A hängt nicht vom Zerfällungskörper ab). Da A separabel ist, muss auch \check{D} separabel sein. Nach Voraussetzung ist \check{D} mit A und mit $\alpha(A) = D$ vertauschbar.

Nach Lemma 16.16 ist $D - \dot{D}$ halbeinfach. Analog sind auch \dot{N} und N vertauschbar. Multipliziert man $(\dot{N} - N)^{2n}$ aus, so erhält man Summanden der Form $\dot{N}^i N^{2n-i}$ mit $0 \leq i \leq 2n$. Im Fall $i \geq n$ ist $\dot{N}^i = 0$. Anderenfalls ist $2n - i \geq n$ und $N^{2n-i} = 0$. In jedem Fall ist $(\dot{N} - N)^{2n} = 0$. Insgesamt ist $D - \dot{D} = \dot{N} - N$ halbeinfach und nilpotent. Das geht nur mit dem Minimalpolynom X , d. h. $\dot{D} = D$ und $\dot{N} = N$. \square

Folgerung 16.20. Für jede Matrix $A \in \mathbb{C}^{n \times n}$ existieren eindeutig bestimmte Matrizen $D, N \in \mathbb{C}^{n \times n}$ mit folgenden Eigenschaften:

- (a) $A = D + N$ und $DN = ND$.
- (b) D ist diagonalisierbar und N nilpotent.

Beweis. Nach Beispiel 16.13 ist A separabel und D ist genau dann halbeinfach, wenn D diagonalisierbar ist. \square

Bemerkung 16.21. Man kann die Jordan-Chevalley-Zerlegung von $A \in K^{n \times n}$ berechnen, indem man A über einem Zerfällungskörper von μ_A betrachtet und dort die Jordan-Normalform benutzt. Da die Jordan-Chevalley-Zerlegung eindeutig bestimmt ist, liegen D und N (dennoch) in $K^{n \times n}$.

Beispiel 16.22. Sei $\gamma := X^2 + 1 \in \mathbb{Q}[X]$ und

$$A := B(\gamma^2) = \begin{pmatrix} \cdot & \cdot & \cdot & -1 \\ 1 & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & -2 \\ \cdot & \cdot & 1 & \cdot \end{pmatrix} \in \mathbb{Q}[X].$$

Wir bestimmen die Jordan-Chevalley-Zerlegung von A ohne den in Bemerkung 16.21 beschriebenen Umweg über \mathbb{C} (oder $\mathbb{Q}(i)$, vgl. Aufgabe I.15). Nach Satz 16.18 gilt $A \approx J_2(\gamma)$. Um diesen Basiswechsel zu realisieren, benötigen wir ein Basisvektor $b_3 \in \mathbb{Q}^4$ mit $\mu_{b_3} = \gamma$. Dafür eignet sich

$$b_3 := \gamma(A)e_1 = e_1 + e_3 = (1, 0, 1, 0)^t.$$

Dann ist $b_4 := Ab_3 = (0, 1, 0, 1)^t$. Für b_1 und b_2 soll gelten: $Ab_1 = b_2 + b_3$ und $Ab_2 = -b_1 + b_4$. Damit erhält man das Gleichungssystem

$$\gamma(A)b_1 = (A^2 + 1_4)b_1 = 2b_4 = (0, 2, 0, 2)^t$$

mit der Lösung $b_1 = 2e_2$. Schließlich ist $b_2 = Ab_1 - b_3 = (-1, 0, 1, 0)^t$. Für

$$S := \begin{pmatrix} \cdot & -1 & 1 & \cdot \\ 2 & \cdot & \cdot & 1 \\ \cdot & 1 & 1 & \cdot \\ \cdot & \cdot & \cdot & 1 \end{pmatrix} \in \text{GL}(4, \mathbb{Q})$$

gilt $S^{-1}AS = J_2(\gamma)$. Dies ergibt die Jordan-Chevalley-Zerlegung $A = D + N$ mit

$$D := S \text{diag}(B(\gamma), B(\gamma))S^{-1} = \frac{1}{2} \begin{pmatrix} \cdot & -1 & \cdot & -1 \\ 3 & \cdot & -1 & \cdot \\ \cdot & 1 & \cdot & -3 \\ 1 & \cdot & 1 & \cdot \end{pmatrix},$$

$$N := S \begin{pmatrix} 0_2 & 0_2 \\ 1_2 & 0_2 \end{pmatrix} S^{-1} = \frac{1}{2} \begin{pmatrix} \cdot & 1 & \cdot & -1 \\ -1 & \cdot & 1 & \cdot \\ \cdot & 1 & \cdot & -1 \\ -1 & \cdot & 1 & \cdot \end{pmatrix}.$$

Satz 16.23. Sei $A \in K^{n \times n}$ separabel mit Jordan-Chevalley-Zerlegung $A = D + N$. Dann gilt $C(A) = C(D) \cap C(N)$.

Beweis. Für $B \in C(D) \cap C(N)$ gilt $BA = BD + BN = DB + NB = AB$ und $B \in C(A)$. Sei $\alpha \in K[X]$ mit $\alpha(A) = D$. Für $B \in C(A)$ gilt dann $BD = B\alpha(A) = \alpha(A)B = DB$ und $BN = B(A - \alpha(A)) = (A - \alpha(A))B = NB$. Dies zeigt $B \in C(D) \cap C(N)$. \square

Aufgaben

Aufgabe II.1. Seien $\alpha, \beta \in K[X]$ mit $\alpha \mid \beta \mid \alpha$. Zeigen Sie, dass ein $c \in K^\times$ mit $\alpha = c\beta$ existiert.

Aufgabe II.2. Sei $n \in \mathbb{N}$. Zeigen Sie:

- (a) $A \in \mathbb{F}_2^{n \times n}$ ist genau dann diagonalisierbar, wenn $A^2 = A$.
- (b) $A \in \text{GL}(n, \mathbb{F}_2)$ ist genau dann diagonalisierbar, wenn $A = 1_n$.
Hinweis: Satz 10.52.
- (c) $|\text{GL}(n, \mathbb{F}_2)| = (2^n - 1)(2^n - 2) \dots (2^n - 2^{n-1})$.
Hinweis: Satz 7.45.

Aufgabe II.3. Sei $A \in K^{n \times m}$ und $B \in K^{m \times n}$. Beweisen Sie die *Determinanten-Formel* von SYLVESTER

$$\det(1_n + AB) = \det(1_m + BA).$$

Hinweis: Lemma 10.38.

Aufgabe II.4. Seien $A, B \in K^{n \times n}$.

- (a) Beweisen oder widerlegen Sie:

$$\chi_A = \chi_{A^t}, \quad \mu_A = \mu_{A^t}, \quad \mu_{AB} = \mu_{BA}.$$

- (b) Wie lassen sich $\chi_{A^{-1}}$ und $\mu_{A^{-1}}$ aus χ_A und μ_A berechnen, falls A invertierbar ist?

Aufgabe II.5. Zeigen Sie $\chi_A = X^3 - \text{tr}(A)X^2 + \frac{1}{2}(\text{tr}(A)^2 - \text{tr}(A^2))X - \det(A)$ für alle $A \in K^{3 \times 3}$.

Aufgabe II.6. Sei V ein euklidischer Raum und $U, W \leq V$. Zeigen Sie:

- (a) $(U + W)^\perp = U^\perp \cap W^\perp$.
- (b) $(U \cap W)^\perp = U^\perp + W^\perp$.

Hinweis: Man kann Lemma 16.3 benutzen.

Aufgabe II.7. Sei K ein Körper, $n \in \mathbb{N}$ und $S \in K^{n \times n}$. Zeigen Sie, dass

$$\{A \in \text{GL}(n, K) : ASA^t = S\}$$

eine Untergruppe von $\text{GL}(n, K)$ ist. Für $S = 1_n$ erhält man $O(n, K)$.

Aufgabe II.8. Seien $v, w \in \mathbb{R}^2$ linear unabhängig. Dann bilden $0, v, w$ die Eckpunkte eines Dreiecks mit Seitenlängen $A := |v|$, $B := |w|$ und $C := |v - w|$. Seien α, β, γ die Winkel gegenüber von A, B, C . Zeigen Sie:

- (a) (Sinussatz) $\frac{\sin \alpha}{A} = \frac{\sin \beta}{B} = \frac{\sin \gamma}{C}$.
 (b) (Kosinussatz) $C^2 = A^2 + B^2 - 2AB \cos \gamma$.
 (c) (Trigonometrischer Pythagoras) $\sin(\alpha)^2 + \cos(\alpha)^2 = 1$.

Aufgabe II.9. Für $\zeta := \cos(\pi/5) + i \sin(\pi/5) \in \mathbb{C}$ gilt $\zeta^5 = -1$ (siehe Beweis von Lemma 11.27). Sei $\omega := \zeta + \zeta^{-1} = 2\operatorname{Re}(\zeta) \in \mathbb{R}$. Zeigen Sie:

- (a) $\omega^2 - \omega - 1 = 0$.
Hinweis: $X^5 + 1 = (X + 1)(X^4 - X^3 + X^2 - X + 1)$.

(b) $\omega = \frac{1}{2}(\sqrt{5} + 1)$.

(c)

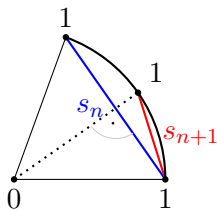
$$D(\pi/5) = \frac{1}{4} \begin{pmatrix} \frac{\sqrt{5} + 1}{\sqrt{10 - 2\sqrt{5}}} & -\frac{\sqrt{10 - 2\sqrt{5}}}{\sqrt{5} + 1} \\ \sqrt{10 - 2\sqrt{5}} & \sqrt{5} + 1 \end{pmatrix}.$$

Hinweis: $\cos(\varphi)^2 + \sin(\varphi)^2 = 1$.

Aufgabe II.10. Wir approximieren die Halbkreisbogenlänge π durch den halben Umfang eines regelmäßigen 2^n -Ecks mit „Radius“ 1. Dafür sei s_n die Seitenlänge des regelmäßigen 2^n -Ecks.

(a) Zeigen Sie $s_2 = \sqrt{2}$.

(b) Zeigen Sie $s_{n+1} = \sqrt{2 - \sqrt{4 - s_n^2}}$ durch zweimalige Anwendung von Pythagoras:



(c) Zeigen Sie

$$2^n \underbrace{\sqrt{2 - \sqrt{2 + \sqrt{2 + \dots \sqrt{2}}}}}_{n \text{ Wurzeln}} = \lim_{n \rightarrow \infty} 2^n s_{n+1} = \pi.$$

Aufgabe II.11. Sei $V := \mathbb{R}^n$ der euklidische Raum bzgl. des Standardskalarprodukts. Sei $v \in V$ normiert und $S_v \in O(V)$ die Spiegelung an v^\perp . Zeigen Sie $[S_v] = 1_n - 2v^t v$.

Bemerkung: Solche Matrizen nennt man *Householder-Transformationen*.

Aufgabe II.12. Sei $\alpha \in \mathbb{R}[X]$. Zeigen Sie, dass Polynome $\gamma_1, \dots, \gamma_k \in \mathbb{R}[X]$ mit $\alpha = \gamma_1 \dots \gamma_k$ und $\deg(\gamma_i) \leq 2$ für $i = 1, \dots, k$ existieren.

Hinweis: Bemerkung 11.37.

Aufgabe II.13. Zeigen Sie, dass man im Hauptsatz und im Spektralsatz die Transformationsmatrix S mit $\det(S) = 1$ wählen kann.

Aufgabe II.14. Sei V ein \mathbb{R} -Vektorraum und $\beta \in \text{Bil}(V)$ symmetrisch mit $\text{ind}(\beta) = (r, s, t)$. Zeigen Sie, dass r (bzw. s) die maximale Dimension eines Unterraums $U \leq V$ ist, sodass die Einschränkung von β auf $U \times U$ positiv (bzw. negativ) definit ist.

Aufgabe II.15. Sei $A \in \mathbb{R}^{n \times n}$ symmetrisch mit $\text{ind}(A) = (r, s, t)$ und $\lambda \in \mathbb{R}$. Zeigen Sie

$$\text{ind}(\lambda A) = \begin{cases} (r, s, t) & \text{falls } \lambda > 0, \\ (s, r, t) & \text{falls } \lambda < 0, \\ (0, 0, n) & \text{falls } \lambda = 0. \end{cases}$$

Aufgabe II.16. Sei $A \in \mathbb{R}^{n \times n}$ symmetrisch. Wie kann man an χ_A ablesen, ob A positiv semidefinit ist? Beweisen Sie ein Kriterium in Analogie zu Satz 12.44.

Aufgabe II.17. Zeigen Sie für alle $A \in \mathbb{C}^{n \times m}$:

- (a) $\text{Ker}(A) = \text{Ker}(A^*A)$.
- (b) $\text{rk}(A) = \text{rk}(A^*A)$.

Aufgabe II.18. Eine Matrix $A \in \mathbb{C}^{n \times n}$ heißt *positiv (semi)definit*, falls $vAv^* > 0$ (bzw. $vAv^* \geq 0$) für alle $v \in \mathbb{C}^n \setminus \{0\}$ gilt. Zeigen Sie:

- (a) Jede positiv (semi)definite Matrix ist hermitesch.
Bemerkung: Die analoge Aussage für reelle Matrizen ist falsch.
- (b) Eine hermitesche Matrix A ist genau dann positiv (semi)definit, wenn alle Eigenwerte von A positiv (bzw. nicht-negativ) sind.
- (c) Für $k \in \mathbb{N}$ besitzt jede positiv (semi)definite Matrix genau eine positiv (semi)definite k -te Wurzel.

Aufgabe II.19. Seien $A_1, \dots, A_k \in K^{n \times n}$ diagonalisierbar und paarweise vertauschbar. Zeigen Sie, dass A_1, \dots, A_k simultan diagonalisierbar sind.

Hinweis: Die Induktion nach k ist nicht-trivial! Man kann Aufgabe I.17 benutzen.

Aufgabe II.20. Seien $A_1, \dots, A_k \in K^{n \times n}$ paarweise vertauschbare trigonalisierbare Matrizen. Zeigen Sie, dass A_1, \dots, A_k simultan trigonalisierbar sind.

Aufgabe II.21 (Reelle Schur-Zerlegung). Zeigen Sie, dass für jede Matrix $A \in \mathbb{R}^{n \times n}$ eine orthogonale Matrix $Q \in O(n, \mathbb{R})$ existiert, sodass

$$Q^t A Q = \begin{pmatrix} R_{11} & & * \\ & \ddots & \\ 0 & & R_{kk} \end{pmatrix}$$

mit $R_{ii} \in \mathbb{R} \cup \mathbb{R}^{2 \times 2}$ für $i = 1, \dots, k$. Im Fall $R_{ii} \in \mathbb{R}^{2 \times 2}$ hat R_{ii} zwei komplex konjugierte (nicht-reelle) Eigenwerte. Insbesondere ist $Q^t A Q$ eine obere Dreiecksmatrix, falls A lauter reelle Eigenwerte besitzt.

Aufgabe II.22. Sei $\lambda \in K$ und $A := J_n(\lambda)$ ein Jordanblock. Zeigen Sie, dass $C \in K^{n \times n}$ genau dann mit A vertauschbar ist, wenn $c_1, \dots, c_n \in K$ mit

$$C = \begin{pmatrix} c_1 & & & 0 \\ c_2 & \ddots & & \\ \vdots & \ddots & \ddots & \\ c_n & \cdots & c_2 & c_1 \end{pmatrix}$$

existieren.

Aufgabe II.23. Sei V ein K -Vektorraum und $f \in \text{End}(V)$. Sei b_1, \dots, b_n eine Basis von V . Zeigen Sie, dass μ_f das kleinste gemeinsame Vielfache von $\mu_{b_1}, \dots, \mu_{b_n}$ ist, d. h. es gibt kein normiertes Polynom kleineren Grades, das durch $\mu_{b_1}, \dots, \mu_{b_n}$ teilbar ist.

Aufgabe II.24. Sei V ein K -Vektorraum und $f \in \text{End}(V)$. Sei $\lambda \in K$ ein Eigenwert von f , der mit Vielfachheit k als Nullstelle von μ_f auftritt. Zeigen Sie

$$E_\lambda(f) \subsetneq \text{Ker}((f - \lambda \text{id})^2) \subsetneq \dots \subsetneq \text{Ker}((f - \lambda \text{id})^k) = H_\lambda(f).$$

Hinweis: Man kann die Weierstraß-Normalform benutzen.

Aufgabe II.25. Sei $\alpha \in K[X] \setminus K$ normiert. Zeigen Sie $\chi_{B(\alpha)} = \alpha$ mit der Definition des charakteristischen Polynoms (und nicht über das Minimalpolynom wie in Lemma 15.20).

Aufgabe II.26. Sei $\alpha = (X - \lambda_1) \dots (X - \lambda_n) \in K[X]$ mit paarweise verschiedenen $\lambda_1, \dots, \lambda_n$. Sei $V := (\lambda_i^{j-1}) \in K^{n \times n}$ die Vandermonde-Matrix. Zeigen Sie $VB(\alpha)V^{-1} = \text{diag}(\lambda_1, \dots, \lambda_n)$.

Aufgabe II.27. Sei $\alpha = X^n + a_{n-1}X^{n-1} + \dots + a_0$ und

$$S = \begin{pmatrix} a_1 & a_2 & \cdots & a_{n-1} & 1 \\ a_2 & a_3 & \ddots & & \\ \vdots & \ddots & \ddots & & \\ a_{n-1} & 1 & & & \\ 1 & & & & 0 \end{pmatrix} \in \text{GL}(n, K)$$

(a) Zeigen Sie $S^{-1}B(\alpha)S = B(\alpha)^t$.

(b) (TAUSSKY) Folgern Sie, dass jede quadratische Matrix das Produkt von zwei symmetrischen Matrizen ist.

Aufgabe II.28. Sei $A \in \mathbb{C}^{n \times n}$ mit lauter reellen Eigenwerten. Zeigen Sie, dass A zu einer reellen Matrix ähnlich ist.

Hinweis: Frobenius-Normalform.

Aufgabe II.29. Beweisen Sie Folgerung 14.37 über einen beliebigen Körper K anstelle von \mathbb{C} .

Aufgabe II.30.

(a) Seien $A \in K^{n \times n}$ und $B \in K^{m \times m}$, sodass μ_A und μ_B teilerfremd sind. Zeigen Sie

$$C(\text{diag}(A, B)) = \{\text{diag}(C, D) : C \in C(A), D \in C(B)\} \cong C(A) \times C(B).$$

Hinweis: Lemma von Bézout.

(b) Sei A halbeinfach mit $\chi_A = \gamma_1^{a_1} \dots \gamma_k^{a_k}$ (Primfaktorzerlegung). Zeigen Sie

$$\dim C(A) = \sum_{i=1}^k a_i^2 \deg(\gamma_i)$$

mit (a). Vergleichen Sie mit Bemerkung 15.33.

Aufgabe II.31. Sei V ein K -Vektorraum und $F \subseteq \text{End}(V)$. Ein Unterraum $U \leq V$ heißt F -invariant, falls $f(U) \subseteq U$ für alle $f \in F$ gilt. Sei

$$C(F) := \bigcap_{f \in F} C(f)$$

der Zentralisator von F (analog für Matrizen).

(a) Angenommen $\{0\}$ und V sind die einzigen F -invarianten Unterräume. Zeigen Sie, dass alle $g \in C(F) \setminus \{0\}$ invertierbar sind.

(b) Sei $F = \{f_1, f_2\}$ mit

$$f_1 := \text{diag}(B(X^2 + 1), B(X^2 + 1)) \in \mathbb{R}^{4 \times 4}, \quad f_2 := \begin{pmatrix} 0_2 & B(X^2 - 1) \\ -B(X^2 - 1) & 0_2 \end{pmatrix} \in \mathbb{R}^{4 \times 4}.$$

Zeigen Sie, dass $\mathbb{H} := C(F) \subseteq \mathbb{R}^{4 \times 4}$ ein 4-dimensionaler \mathbb{R} -Vektorraum ist, in dem jedes 0 verschiedene Element invertierbar ist.

Bemerkung: Im Gegensatz zu Satz 15.39 ist die Multiplikation in \mathbb{H} nicht kommutativ. Man nennt \mathbb{H} den HAMILTONSchen Schiefkörper.

Lineare Algebra III

17 Numerische Verfahren

17.1 Effiziente Arithmetik

Bemerkung 17.1. In der linearen Algebra I und II haben wir Sachverhalte hauptsächlich aus theoretischer Sicht untersucht. Zum Beispiel wurden Eigenwerte als Nullstellen des charakteristischen Polynoms berechnet, obwohl dies für größere Matrizen unpraktikabel ist (vgl. Beispiel 10.25). Wir beschreiben in diesem Kapitel Algorithmen, mit denen man in der Praxis Probleme der linearen Algebra effizient und robust (gegenüber Rundungsfehler) löst. Eigenwerte werden mehr oder weniger über die Definition berechnet (Satz 17.70).

Beispiel 17.2. Auf Computern ist die Multiplikation zweier Zahlen in der Regel aufwendiger als die Addition (es gibt jedoch Ausnahmen: Die Multiplikation mit 2 entspricht einem Shift der Binärfolge um eine Ziffer nach links.). Den folgenden Laufzeitanalysen werden wir daher Additionen vernachlässigen. Die Multiplikation zweier n -stelliger Dezimalzahlen mit der „Schulmethode“ erfordert n^2 Ziffer-Multiplikationen (das kleine Einmaleins kann persistent auf dem Chip gespeichert werden):

$$\begin{array}{r} 123 \cdot 567 = 69741 \\ \hline 861 \\ + 738 \\ + 615 \\ \hline 69741 \end{array}$$

Das geht schneller.

Satz 17.3 (KARATSUBA-Algorithmus). *Seien $x, y \in \mathbb{N}$ Dezimalzahlen mit n Ziffern.*

(1) *Teile x und y in zwei Hälften der Länge $m \approx n/2$:*

$$x = x_1 10^m + x_0, \quad y = y_1 10^m + y_0 \quad (x_0, y_0 < 10^m)$$

(2) *Multipliziere rekursiv die m -stelligen Zahlen:*

$$z_0 := x_0 y_0, \quad z_2 := x_1 y_1, \quad z_1 := (x_1 - x_0)(y_0 - y_1) + z_0 + z_2.$$

(3) *Dann gilt $xy = 10^{2m} z_2 + 10^m z_1 + z_0$.*

Dieser Algorithmus benötigt ca. $n^{\log_2(3)} \approx n^{1.58} < n^2$ Ziffer-Multiplikationen.

Beweis. Es gilt $z_1 = x_1 y_0 + x_0 y_1 - x_1 y_1 - x_0 y_0 + z_0 + z_2 = x_1 y_0 + x_0 y_1$ und

$$xy = (x_1 10^m + x_0)(y_1 10^m + y_0) = 10^{2m} x_1 y_1 + 10^m (x_1 y_0 + x_0 y_1) + x_0 y_0 = 10^{2m} z_2 + 10^m z_1 + z_0.$$

Für die zweite Behauptung argumentieren wir durch Induktion nach n : Für $n = 1$ braucht man $1 = 1^{\log_2(3)}$ Ziffer-Multiplikation. Die Berechnung von z_0, z_1, z_2 für $n \geq 2$ erfordert 3 Multiplikationen m -stelliger Zahlen. Induktiv erhält man

$$3m^{\log_2(3)} \approx 3(n/2)^{\log_2(3)} = n^{\log_2(3)}$$

Ziffer-Multiplikationen. □

Beispiel 17.4. Für $x = 87 = 80 + 7$ und $y = 91 = 90 + 1$ erhält man

$$\begin{aligned} z_0 &= 7 \cdot 1 = 7, & z_2 &= 8 \cdot 9 = 72, & z_1 &= (8 - 7)(1 - 9) + z_0 + z_2 = -8 + 7 + 72 = 71, \\ xy &= 7200 + 710 + 7 = 7917. \end{aligned}$$

Bemerkung 17.5. Vor der Einführung elektronischer Taschenrechner hat man die Multiplikation von Zahlen $x, y \in \mathbb{R}$ mittels *Logarithmentafeln* auf die einfachere Addition $\log(x) + \log(y)$ zurückgeführt. Grundlage dafür ist die Funktionalgleichung $\log(xy) = \log(x) + \log(y)$. Eine ähnliche Reduktion mit Hilfe der FOURIER-Transformation ist heute noch relevant.

Definition 17.6. Für $n \in \mathbb{N}$ sei

$$\zeta_n := \zeta := \cos(2\pi/n) + i \sin(2\pi/n)$$

eine n -te Einheitswurzel (Beispiel 11.28). Die symmetrische Vandermonde-Matrix

$$W_n := W := (\zeta^{ij})_{i,j=0}^{n-1} \in \mathbb{C}^{n \times n}$$

nennt man die n -te *Fourier-Matrix*. Die Abbildung $\mathcal{F}_n: \mathbb{C}^n \rightarrow \mathbb{C}^n$, $x \mapsto xW =: \hat{x}$ heißt *diskrete Fourier-Transformation*.

Beispiel 17.7. Nach Beispiel 11.28 gilt

$$W_4 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{pmatrix}.$$

Lemma 17.8. Für $x \in \mathbb{C}^n$ gilt $\mathcal{F}_n^{-1}(x) = \frac{1}{n} \overline{\mathcal{F}_n(\bar{x})}$.

Beweis. Wegen $\zeta \bar{\zeta} = |\zeta|^2 = 1$ ist $\zeta^{-1} = \bar{\zeta}$. Für $1 \leq i, j \leq n$ ist auch $\sigma := \zeta^{i-j}$ eine n -te Einheitswurzel. Im Fall $i = j$ ist $\sigma = 1$ und anderenfalls $\sum_{k=0}^{n-1} \sigma^k = \frac{\sigma^n - 1}{\sigma - 1} = 0$ nach der Formel für die geometrische Reihe. Daher gilt

$$(W_n \overline{W_n})_{ij} = \sum_{k=0}^{n-1} \zeta^{ik} \zeta^{-kj} = \sum_{k=0}^{n-1} \sigma^k = \begin{cases} n & \text{falls } i = j, \\ 0 & \text{falls } i \neq j. \end{cases}$$

Dies zeigt $W_n^{-1} = \frac{1}{n} \overline{W_n}$ und $n\mathcal{F}_n^{-1}(x) = x \overline{W_n} = \overline{\mathcal{F}_n(\bar{x})}$. □

Satz 17.9 (Faltungssatz). Für $\alpha = \sum a_k X^{k-1} \in \mathbb{C}[X]$ sei $[\alpha] := (a_1, \dots, a_n)$. Für alle $\alpha, \beta \in \mathbb{C}[X]$ mit $\deg(\alpha) + \deg(\beta) < n$ gilt

$$\mathcal{F}_n([\alpha\beta]) = (\hat{a}_1 \hat{b}_1, \dots, \hat{a}_n \hat{b}_n).$$

Beweis. Wegen

$$\alpha\beta = \sum_{k=2}^{n+1} \left(\sum_{i+j=k} a_i b_j \right) X^{k-2} = \sum_{k=1}^n \left(\sum_{i+j=k+1} a_i b_j \right) X^{k-1}$$

gilt $[\alpha\beta]_k = \sum_{i+j=k+1} a_i b_j$ für $1 \leq k \leq n$. Es folgt

$$\hat{a}_k \hat{b}_k = \sum_{i,j=1}^n a_i b_j \zeta^{(k-1)(i+j-2)} = \sum_{l=2}^{2n} \left(\zeta^{(k-1)(l-2)} \sum_{i+j=l} a_i b_j \right) = \sum_{l=1}^n [\alpha\beta]_l \zeta^{(k-1)(l-1)} = \mathcal{F}_n([\alpha\beta])_k. \quad \square$$

Bemerkung 17.10.

- (a) Um zwei Zahlen $a, b \in \mathbb{N}$ (mit $ab < 10^n$) zu multiplizieren, fasst man sie als Polynome in $X = 10$ auf $a = \sum a_k 10^{k-1}$, $b = \sum b_k 10^{k-1}$ und wendet den Faltungssatz an:

$$ab = \mathcal{F}_n^{-1}(\hat{a}_1 \hat{b}_1, \dots, \hat{a}_n \hat{b}_n) \cdot (1, 10, \dots, 10^{n-1})^t.$$

Man beachte, dass Überträge wie bei der Schulmultiplikation hier nicht aufgelöst werden. Überträge lassen sich generell vermeiden, indem man a und b *binär*, d. h. als Polynom in $X = 2$ darstellt. Diese Art der Berechnung liefert allerdings noch keine Beschleunigung.

- (b) Sei $n = 2m$ gerade. Mit den Bezeichnungen $y_k := x_{2k-1}$ und $z_k := x_{2k}$ gilt

$$\begin{aligned} \mathcal{F}_n(x)_i &= \sum_{k=1}^n x_k \zeta_n^{(i-1)(k-1)} = \sum_{k=1}^m y_k \zeta_m^{(i-1)(k-1)} + \zeta_n^{i-1} \sum_{k=1}^m z_k \zeta_m^{(i-1)(k-1)} \\ &= \mathcal{F}_m(y)_i + \zeta_n^{i-1} \mathcal{F}_m(z)_i \end{aligned} \tag{17.1}$$

für $i = 1, \dots, n$, wobei $\mathcal{F}_m(y)_i = \mathcal{F}_m(y)_{i-m}$ für $i > m$. Mit Lemma 17.8 ist analog

$$\mathcal{F}_n^{-1}(x)_i = \frac{1}{2} \left(\mathcal{F}_m^{-1}(y)_i + \zeta_n^{1-i} \mathcal{F}_m^{-1}(z)_i \right).$$

Bei der *schnellen* Fourier-Transformation (FFT) geht man davon aus, dass n eine 2-Potenz ist (was durch Anfügen von Nullen immer möglich ist). Wie beim Karatsuba-Algorithmus kann man \mathcal{F}_n bis auf $\mathcal{F}_1(x) = x$ reduzieren. Für ein festes n können die n -ten Einheitswurzeln einmalig berechnet und gespeichert werden. Die Addition solcher Zahlen kann günstig mit beschränkter Genauigkeit durchgeführt werden.

- (c) Der weitverbreitete SCHÖNHAGE-STRASSEN-Algorithmus führt die Multiplikation n -stelliger Zahlen mittels FFT in einer asymptotischen Laufzeit von $n \log(n) \log \log(n)$ durch. 2019 konnte HARVEY-VAN DER HOEVEN die Laufzeit auf $n \log(n)$ verbessert, wobei es sich jedoch um einen *galaktischen* Algorithmus, d. h. die Zeitersparnis wird erst für extrem große Zahlen jenseits jeglicher praktischer Bedeutung relevant. Man vermutet, dass $n \log(n)$ asymptotisch die bestmögliche Schranke ist.
- (d) Die *kontinuierliche* Fourier-Transformation weist einer integrierbaren Funktion $f: \mathbb{R}^n \rightarrow \mathbb{R}$ das Integral

$$\mathcal{F}(y) = \frac{1}{\sqrt{2\pi}^n} \int_{\mathbb{R}^n} f(x) e^{2\pi i [x,y]} dx$$

zu.

Beispiel 17.11. Wir multiplizieren $342 \cdot 87$ mit der schnellen Fourier-Transformation. Nach (17.1) gilt

$$\mathcal{F}_4(2, 4, 3, 0)^t = \begin{pmatrix} \mathcal{F}_2(2, 3)_1 + \mathcal{F}_2(4, 0)_1 \\ \mathcal{F}_2(2, 3)_2 + i\mathcal{F}_2(4, 0)_2 \\ \mathcal{F}_2(2, 3)_1 - \mathcal{F}_2(4, 0)_1 \\ \mathcal{F}_2(2, 3)_2 - i\mathcal{F}_2(4, 0)_2 \end{pmatrix} = \begin{pmatrix} (2+3) + 4 \\ (2-3) + 4i \\ (2+3) - 4 \\ (2-4) - 4i \end{pmatrix} = \begin{pmatrix} 9 \\ -1 + 4i \\ 1 \\ -1 - 4i \end{pmatrix},$$

$$\mathcal{F}_4(7, 8, 0, 0) = (15, 7 + 8i, -1, 7 - 8i),$$

$$\mathcal{F}_4^{-1}(135, -39 + 20i, -1, -39 - 20i)^t = \frac{1}{4} \begin{pmatrix} 135 - 1 - 2 \cdot 39 \\ 135 + 1 - i40i \\ 135 - 1 + 2 \cdot 39 \\ 135 + 1 + i40i \end{pmatrix} = (14, 44, 53, 24),$$

$$342 \cdot 87 = \mathcal{F}_4^{-1}(\dots)(1, 10, 100, 1000)^t = 14 + 440 + 5.300 + 24.000 = 29.754.$$

Bemerkung 17.12. Die direkte Multiplikation zweier $n \times n$ -Matrizen A, B erfordert n^3 Zahl-Multiplikationen $a_{ij}b_{jk}$ mit $1 \leq i, j, k \leq n$, die man wiederum mit einem der obigen Algorithmen durchführen kann. Auch das geht besser.

Satz 17.13 (STRASSEN-Algorithmus). *Seien $A, B \in K^{n \times n}$.*

(1) *Durch Anfügen einer Nullzeile und -spalte kann man $n = 2m$ annehmen.*

(2) *Teile A und B in $m \times m$ -Blöcke:*

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \quad B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$

(3) *Berechne rekursiv:*

$$\begin{aligned} C_1 &:= (A_{11} + A_{22})(B_{11} + B_{22}), & C_2 &:= (A_{21} + A_{22})B_{11}, \\ C_3 &:= A_{11}(B_{12} - B_{22}), & C_4 &:= A_{22}(B_{21} - B_{11}), \\ C_5 &:= (A_{11} + A_{12})B_{22}, & C_6 &:= (A_{21} - A_{11})(B_{11} + B_{12}), \\ C_7 &:= (A_{12} - A_{22})(B_{21} + B_{22}). \end{aligned}$$

(4) *Dann gilt*

$$AB = \begin{pmatrix} C_1 + C_4 - C_5 + C_7 & C_3 + C_5 \\ C_2 + C_4 & C_1 - C_2 + C_3 + C_6 \end{pmatrix}.$$

Dieser Algorithmus benötigt ca. $n^{\log_2(7)} \approx n^{2.81}$ Zahl-Multiplikationen.

Beweis. Die Formel für AB folgt aus

$$\begin{aligned} C_1 + C_4 - C_5 + C_7 &= (A_{11} + A_{22})(B_{11} + B_{22}) + A_{22}(B_{21} - B_{11}) - (A_{11} + A_{12})B_{22} \\ &\quad + (A_{12} - A_{22})(B_{21} + B_{22}) = A_{11}B_{11} + A_{12}B_{21}, \\ C_3 + C_5 &= A_{11}(B_{12} - B_{22}) + (A_{11} + A_{12})B_{22} = A_{11}B_{12} + A_{12}B_{22}, \\ C_2 + C_4 &= (A_{21} + A_{22})B_{11} + A_{22}(B_{21} - B_{11}) = A_{21}B_{11} + A_{22}B_{21}, \\ C_1 - C_2 + C_3 + C_6 &= (A_{11} + A_{22})(B_{11} + B_{22}) - (A_{21} + A_{22})B_{11} + A_{11}(B_{12} - B_{22}) \\ &\quad + (A_{21} - A_{11})(B_{11} + B_{12}) = A_{21}B_{12} + A_{22}B_{22}. \end{aligned}$$

Die Berechnung jedes C_i erfordert induktiv ca. $m^{\log_2(7)} = \frac{1}{7}n^{\log_2(7)}$ Zahl-Multiplikationen. Daher benötigt man für AB ca. $n^{\log_2(7)}$ Zahl-Multiplikationen. \square

Bemerkung 17.14.

- (a) Auch der Strassen-Algorithmus wurde weiter verbessert. Zuletzt wurde 2025 ein Algorithmus mit $n^{2.371339}$ Zahl-Multiplikationen entdeckt.¹ Aktuell werden *Large-Language-Models* wie AlphaEvolve eingesetzt. Die asymptotisch bestmögliche Komplexität ist ein offenes Problem der Informatik.
- (b) Matrixpotenzen lassen sich durch iteriertes Quadrieren effizient berechnen: $A^{11} = A(A^5)^2 = A(A(A^2)^2)^2$ benötigt nur fünf Multiplikationen anstatt zehn (Aufgabe III.1).
- (c) Die Multiplikation beliebiger (nicht unbedingt quadratischer) Matrizen $A \in K^{n \times m}$ und $B \in K^{m \times k}$ erfordert nmk Zahl-Multiplikationen. Werden mehr als zwei solche Matrizen multipliziert, so hat die Klammersetzung einen erheblichen Einfluss auf die Anzahl der Zahl-Multiplikationen. Das zeigt sich besonders deutlich im Extremfall $n = k > 1 = m = l$:

$$(AB)C = \begin{pmatrix} a_{11}b_{11} & \cdots & a_{11}b_{1n} \\ \vdots & & \vdots \\ a_{n1}b_{11} & \cdots & a_{n1}b_{1n} \end{pmatrix} \begin{pmatrix} c_{11} \\ \vdots \\ c_{n1} \end{pmatrix} \quad (2n^2 \text{ Multiplikationen}),$$
$$A(BC) = \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix} (b_{11}c_{11} + \dots + b_{1n}c_{11}) \quad (2n \text{ Multiplikationen}).$$

Satz 17.15. Seien $A \in K^{n \times m}$, $B \in K^{m \times k}$ und $C \in K^{k \times l}$. Die Berechnung von $A(BC)$ benötigt genau dann weniger Zahl-Multiplikationen als $(AB)C$, wenn $\frac{1}{n} + \frac{1}{k} < \frac{1}{m} + \frac{1}{l}$.

Beweis. Die Berechnung von $A(BC)$ bzw. $(AB)C$ erfordert $mk l + nml = ml(n+k)$ bzw. $nmk + nkl = nk(m+l)$ Zahl-Multiplikationen. Es gilt

$$ml(n+k) < nk(m+l) \iff \frac{n+k}{nk} < \frac{m+l}{ml} \iff \frac{1}{n} + \frac{1}{k} < \frac{1}{m} + \frac{1}{l}. \quad \square$$

Bemerkung 17.16. Bei der konkreten Implementierung ist es empfehlenswert, bewährte Bibliotheken wie BLAS, LAPACK, Armadillo, NumPy und SciPy, Programmiersprachen wie Julia oder Programme wie MATLAB, Scilab und Octave zu benutzen.

17.2 Die Konditionszahl

Beispiel 17.17. Auch wenn lineare Abbildungen stetig im Sinn der Analysis sind, so können kleine Änderungen der Argumente dennoch große Auswirkungen auf die Werte nehmen (in der ϵ - δ -Definition der Stetigkeit muss $\delta \ll \epsilon$ gewählt werden):

$$Ax = \begin{pmatrix} 9 & 8 \\ 8 & 7 \end{pmatrix} x = \begin{pmatrix} 17 \\ 15 \end{pmatrix} \implies x = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$
$$\begin{pmatrix} 9 & 8 \\ 8 & 7 \end{pmatrix} x = \begin{pmatrix} 17 \\ 15.1 \end{pmatrix} \implies x = \begin{pmatrix} 1.8 \\ 0.1 \end{pmatrix}.$$

¹Siehe More Asymmetry Yields Faster Matrix Multiplication

Definition 17.18. Für $A \in \mathbb{C}^{n \times m}$ nennt man

$$\kappa(A) := \frac{\max\{|Ax| : x \in \mathbb{C}^{m \times 1}, |x| = 1\}}{\min\{|Ax| : x \in \mathbb{C}^{m \times 1}, |x| = 1\}} \geq 1$$

die *Konditionszahl* von A , wobei $|\cdot|$ die Norm des Standardskalarprodukts (Beispiel 13.3) bezeichnet. Ist $\kappa(A)$ „klein“ (bzw. „groß“), so ist A *gut* (bzw. *schlecht*) *konditioniert*.

Bemerkung 17.19.

- (a) Wie wollen messen wie sich kleine Änderungen an einem Vektor b auf die Lösung des Systems $Ax = b$ auswirken. Sei dafür $\tilde{b} \approx b$ und $A\tilde{x} = \tilde{b}$. Der relative Fehler von x ist

$$\frac{|x - \tilde{x}|}{|x|} = \frac{|x - \tilde{x}|}{|A(x - \tilde{x})|} \frac{|Ax|}{|x|} \frac{|b - \tilde{b}|}{|b|} \leq \frac{\max\{|Ax|/|x| : 0 \neq x \in \mathbb{C}^{m \times 1}\}}{\min\{|Ay|/|y| : 0 \neq y \in \mathbb{C}^{m \times 1}\}} \frac{|b - \tilde{b}|}{|b|} = \kappa(A) \frac{|b - \tilde{b}|}{|b|}.$$

Die Konditionszahl ist also der maximale Faktor, mit dem sich der relative Fehler von b nach x verstärken kann.

- (b) In der Analysis zeigt man, dass die Abbildung $\mathbb{C}^m \rightarrow \mathbb{C}$, $x \mapsto |Ax|$ und stetig ist. Da die Menge $\{x \in \mathbb{C}^m : |x| = 1\}$ kompakt ist, wird das Maximum/Minimum über $\{|Ax| : |x| = 1\}$ tatsächlich angekommen (d. h. man braucht kein Supremum/Infimum). Für \mathbb{R} anstelle von \mathbb{C} beschreibt die Konditionszahl wie sehr eine Matrix die n -dimensionale Einheitskugel deformiert.

Beispiel 17.20.

- (a) Hat A nicht vollen Rang, so existiert ein normiertes $x \in \mathbb{C}^{n \times 1}$ mit $Ax = 0$. In diesem Fall interpretieren wir $\kappa(A) = \infty$, d. h. A ist besonders schlecht konditioniert.
- (b) Für unitäre Matrizen $S \in U(n, \mathbb{C})$ gilt bekanntlich $|Sx| = |x|$ für alle $x \in \mathbb{C}^{n \times 1}$. Für beliebige $A \in \mathbb{C}^{n \times m}$ folgt $\kappa(SA) = \kappa(A)$ und analog $\kappa(AS) = \kappa(A)$, falls $S \in U(m, \mathbb{C})$. Insbesondere ist $\kappa(S) = \kappa(1_n) = 1$, d. h. unitäre Matrizen sind gut konditioniert.
- (c) Für beliebige Matrizen $A, S \in \mathbb{C}^{n \times n}$ gilt im Allgemeinen $\kappa(SA) \neq \kappa(A)$. Auf diese Weise kann man die Konditionszahl der Koeffizientenmatrix eines Gleichungssystems $Ax = b$ verbessern. Die notwendige Ersetzung von b durch Sb kann jedoch aus dem gleichen Grund fehleranfällig sein.
- (d) Ein klassisches Beispiel einer schlecht konditionierte Matrix ist die symmetrische *Hilbert-Matrix*

$$H_n := \left(\frac{1}{i+j-1} \right)_{i,j} = \begin{pmatrix} 1 & 1/2 & \cdots & 1/n \\ 1/2 & 1/3 & \cdots & 1/(n+1) \\ \vdots & \vdots & \ddots & \vdots \\ 1/n & 1/(n+1) & \cdots & 1/(2n-1) \end{pmatrix}.$$

Man kann zeigen, dass $\kappa(H_n)$ exponentiell in n wächst. Zum Beispiel ist $\kappa(H_4) \approx 15.514$.

- (e) Für normale Matrizen A kann man mit dem Spektralsatz und (b) die Berechnung von $\kappa(A)$ auf eine Diagonalmatrix zurückführen. Der folgende Satz erlaubt dies für beliebige (insbes. nicht-quadratische) Matrizen.

Satz 17.21 (Singulärwertzerlegung). Für $A \in \mathbb{C}^{n \times m}$ existieren $U \in U(n, \mathbb{C})$ und $V \in U(m, \mathbb{C})$, sodass UAV eine reelle Diagonalmatrix² mit nicht-negativen Einträgen ist. Die positiven Einträge (auf der Hauptdiagonale) nennt man Singulärwerte von A . Sie sind bis auf die Reihenfolge eindeutig bestimmt.

²nicht unbedingt quadratisch

Beweis. Existenz: Die positiv semidefinite Matrix $B := A^*A \in \mathbb{C}^{m \times m}$ besitzt nach Aufgabe II.18 lauter nicht-negative Eigenwerte

$$\lambda_1 \geq \dots \geq \lambda_k > \lambda_{k+1} = \dots = \lambda_m = 0,$$

wobei $k = \text{rk}(B) = \text{rk}(A) \leq \min\{n, m\}$ nach Aufgabe II.17. Nach dem Spektralsatz existiert ein $V \in U(m, \mathbb{C})$ mit $V^*BV = \text{diag}(\lambda_1, \dots, \lambda_m)$. Wir schreiben $V = (V_1, V_2)$ mit $V_1 \in \mathbb{C}^{m \times k}$. Aus $(AV_2)^*(AV_2) = V_2^*BV_2 = 0$ folgt $AV_2 = 0$. Wir setzen

$$U_1 := AV_1 \text{diag}(\lambda_1^{-1/2}, \dots, \lambda_k^{-1/2}) \in \mathbb{C}^{n \times k}.$$

Wegen $U_1^*U_1 = 1_k$ kann man U_1 zu $U = (U_1, U_2) \in U(n, \mathbb{C})$ mit Gram-Schmidt ergänzen. Nun ist

$$\begin{aligned} U_2^*AV_1 &= U_2^*U_1 \text{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_k}) = 0, \\ U^*AV &= \begin{pmatrix} U_1^* \\ U_2^* \end{pmatrix} A \begin{pmatrix} V_1 & V_2 \end{pmatrix} = \begin{pmatrix} U_1^*AV_1 & U_1^*AV_2 \\ U_2^*AV_1 & U_2^*AV_2 \end{pmatrix} = (\delta_{ij}\sqrt{\lambda_i}), \end{aligned}$$

wobei $\sqrt{\lambda_1}, \dots, \sqrt{\lambda_k}$ die Singulärwerte von A sind.

Eindeutigkeit: Sei SAT mit $S \in U(n, \mathbb{C})$ und $T \in U(m, \mathbb{C})$ ebenfalls eine reelle Diagonalmatrix mit nicht-negativen Einträgen. Dabei seien μ_1, \dots, μ_l die positiven Einträge. Dann sind μ_1^2, \dots, μ_l^2 die positiven Eigenwerte von

$$(SAT)^2 = (SAT)^*(SAT) = T^*BT.$$

Bei geeigneter Nummerierung gilt also $k = l$ und $\mu_1^2 = \lambda_1, \dots, \mu_k^2 = \lambda_k$. Aus $\mu_i > 0$ folgt $\mu_i = \sqrt{\lambda_i}$ für $i = 1, \dots, k$. Also sind die Singulärwerte bis auf die Reihenfolge eindeutig bestimmt. \square

Bemerkung 17.22. Der Beweis zeigt: Sind $\lambda_1, \dots, \lambda_k$ die von 0 verschiedenen Eigenwerte der positiv semidefiniten Matrix A^*A , so sind $\sqrt{\lambda_1}, \dots, \sqrt{\lambda_k}$ die Singulärwerte von A .

Folgerung 17.23. Sei $A \in \mathbb{C}^{n \times m}$ mit vollem Rang und Singulärwerten $\sigma_1 \geq \dots \geq \sigma_k$. Dann gilt $\kappa(A) = \sigma_1/\sigma_k$.

Beweis. Sei $D := (\delta_{ij}\sigma_i) \in \mathbb{R}^{n \times m}$ die Matrix aus der Singulärwertzerlegung von A . Nach Beispiel 17.20 gilt $\kappa(A) = \kappa(D) = \kappa(\text{diag}(\sigma_1, \dots, \sigma_k))$. Für $x \in \mathbb{C}^k$ mit $|x| = 1$ gilt

$$|(\sigma_1 x_1, \dots, \sigma_k x_k)|^2 = \sigma_1^2 |x_1|^2 + \dots + \sigma_k^2 |x_k|^2 \leq \sigma_1^2 |x|^2 = \sigma_1^2$$

mit Gleichheit für $x = e_1$. Dies zeigt $\kappa(\text{diag}(\sigma_1, \dots, \sigma_k)) = \sigma_1/\sigma_k$. \square

Folgerung 17.24. Für alle $A \in \mathbb{C}^{n \times m}$ gilt $\kappa(A) = \kappa(A^*)$.

Beweis. Nach Lemma 10.38 haben A^*A und AA^* die gleichen von 0 verschiedenen Eigenwerte. Daher folgt die Behauptung aus Bemerkung 17.22. \square

Satz 17.25. Ist $A \in \mathbb{C}^{n \times n}$ normal, so sind die Singulärwerte von A die Beträge der von 0 verschiedenen Eigenwerte.

Beweis. Nach dem Spektralsatz kann man $A = \text{diag}(\lambda_1, \dots, \lambda_n)$ mit den Eigenwerten $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ annehmen. Sei

$$\tilde{\lambda}_i := \begin{cases} \overline{\lambda_i}/|\lambda_i| & \text{falls } \lambda_i \neq 0, \\ 1 & \text{falls } \lambda_i = 0. \end{cases}$$

Dann ist $U := \text{diag}(\tilde{\lambda}_1, \dots, \tilde{\lambda}_n) \in U(n, \mathbb{C})$ und $UA = \text{diag}(|\lambda_1|, \dots, |\lambda_n|)$. Aus der Eindeutigkeit der Singulärwerte folgt die Behauptung. \square

Beispiel 17.26. Für die Matrix aus Beispiel 17.17 gilt $\kappa(A) = \frac{8+\sqrt{65}}{8-\sqrt{65}} \approx 258$.

Satz 17.27 (MOORE-PENROSE). Für $A \in \mathbb{C}^{n \times m}$ existiert genau ein $A^+ \in \mathbb{C}^{m \times n}$ mit den folgenden Eigenschaften:

- (a) $AA^+A = A$ und $A^+AA^+ = A^+$,
- (b) $(AA^+)^* = AA^+$ und $(A^+A)^* = A^+A$.

Man nennt A^+ Pseudoinverse³ von A .

Beweis. Sei $P := UAV = (\delta_{ij}\sigma_i)$ eine Singulärwertzerlegung von A mit Singulärwerten $\sigma_1, \dots, \sigma_k > 0$. Für $i = 1, \dots, m$ definieren wir

$$\tilde{\sigma}_i := \begin{cases} \sigma_i^{-1} & \text{falls } i \leq k, \\ 0 & \text{falls } i > k \end{cases}$$

und $Q := (\delta_{ij}\tilde{\sigma}_i) \in \mathbb{C}^{m \times n}$. Setze $A^+ := VQU \in \mathbb{C}^{m \times n}$. Die vier angegebenen Eigenschaften reduzieren sich auf $PQP = P$, $QPQ = Q$, $(PQ)^* = PQ$ und $(QP)^* = QP$. Dies gilt offenbar. Seien nun B und C Pseudoinverse von A . Dann gilt

$$B = BAB = B(ACA)B = (BA)^*(CA)^*B = (ABA)^*C^*B = (CA)^*B = CAB = \dots = CAC = C. \quad \square$$

Beispiel 17.28.

- (a) Trivialerweise ist $0_{n \times m}^+ = 0_{m \times n}$.
- (b) Ist A invertierbar, so erfüllt A^{-1} die Bedingungen der Pseudoinversen und es folgt $A^+ = A^{-1}$.
- (c) Für $A \in \mathbb{C}^{n \times m}$ und $\lambda \in \mathbb{C}^\times$ ist $(\lambda A)^+ = \lambda^{-1}A^+$.
- (d) Der Beweis von Satz 17.27 führt die Berechnung von A^+ auf die Singulärwertzerlegung und damit auf den Spektralsatz zurück. Für $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ liefert der Hauptsatzensatz $S := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ mit $S^{-1} = S^t = S$ und $SAS = \text{diag}(2, 0)$. Also ist $A^+ = S \text{diag}(1/2, 0)S = \frac{1}{4}A$.
- (e) Offenbar ist $(A^t)^+ = (A^+)^t$ und $\overline{A^+} = \overline{A^+}$. Insbesondere ist A^+ symmetrisch (bzw. reell, hermitesch), falls A symmetrisch (bzw. reell, hermitesch) ist.
- (f) Hat $A \in \mathbb{C}^{n \times m}$ vollen Rang, so ist A^*A (falls $n \geq m$) oder AA^* (falls $n \leq m$) invertierbar. Ggf. gilt $A^+ = (A^*A)^{-1}A^*$ bzw. $A^+ = A^*(AA^*)^{-1}$. Für $n \geq m$ kann man auf diese Weise eine (Nährungs)lösung des Gleichungssystems $Ax = b$ berechnen: $x \approx A^+b$. Beispiel:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix} x = \begin{pmatrix} 3 \\ 5 \\ 7 \end{pmatrix}, \quad A^+ = \frac{1}{3} \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} A^* = \frac{1}{3} \begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \end{pmatrix}, \quad x \approx \frac{1}{3} \begin{pmatrix} 8 \\ 14 \end{pmatrix}.$$

Mehr dazu in Satz 17.42.

³oder Moore-Penrose-Inverse

17.3 Stabile Varianten des Gauß-Algorithmus

Bemerkung 17.29. Auf Computern lassen sich reelle Zahlen nur näherungsweise durch *Gleitkommazahlen* darstellen. Selbst endliche Dezimalbrüche kann man nicht exakt speichern, wenn die interne Binärdarstellung unendlich ist:

$$0.1 = 2^{-4} + 2^{-5} + 2^{-8} + 2^{-9} + \dots \approx 0.0996.$$

Der IEEE-Standard 754 definiert den 32-bit Datentyp `float` (simple precision) durch folgende Aufteilung: 1 Bit für das Vorzeichen, 23 Bits für die *Mantisse* (Binärentwicklung) und 8 Bits für den Exponenten. Damit lassen sich Zahlen der Form

$$\pm \left(2^e + \sum_{i=1}^{23} a_i 2^{e-i} \right) \quad (a_i \in \{0, 1\}, -126 \leq e \leq 127)$$

repräsentieren ($e = -127$ und $e = 128$ sind reserviert für ∞ und `NaN`). Im Folgenden betrachten wir ein einfacheres Modell mit nur drei Dezimalziffern, also Zahlen der Form

$$\pm \sum_{i=0}^2 a_i 10^{e-i} \quad (0 \leq a_i \leq 9)$$

mit $a_0 \neq 0$ (da Über- und Unterläufe selten vorkommen, beschränken wir den Exponenten e nicht). Beliebige reellen Zahlen werden wie folgt konvertiert:

$$1001 \rightsquigarrow 1.00\text{e}4, \quad 0.00123 \rightsquigarrow 1.23\text{e}-3.$$

Subtraktion fast gleichgroßer Zahlen führt zum Verlust von Genauigkeit (*Auslöschung*):

$$1.23\text{e}0 - 1.22\text{e}0 = 1.??\text{e}-3.$$

Die mit einem Fragezeichen gekennzeichneten Ziffern enthalten keine sinnvolle Information. Die Addition von Gleitkommazahlen ist außerdem nicht assoziativ:

$$(1.00\text{e}-3 + 1.00\text{e}1) - 1.00\text{e}1 = 0.00\text{e}0 \neq 1.00\text{e}-3 = 1.00\text{e}-3 + (1.00\text{e}1 - 1.00\text{e}1).$$

Ebenfalls problematisch ist die Multiplikation mit großen Zahlen, weil dadurch vorhandene Rundungsfehler verstärkt werden.

Beispiel 17.30. Das Gleichungssystem

$$\begin{pmatrix} 10^{-4} & 1 \\ 1 & 1 \end{pmatrix} x = \begin{pmatrix} 10^4 \\ 1 \end{pmatrix}$$

besitzt die eindeutige Lösung $x = (-10^4, 10^4 + 1)$. Der Gauß-Algorithmus mit Gleitkommazahlen liefert jedoch

$$\left(\begin{array}{cc|c} 1\text{e}-4 & 1 & 1\text{e}4 \\ 1 & 1 & 1 \end{array} \right) \sim \left(\begin{array}{cc|c} 1 & 1\text{e}4 & 1\text{e}8 \\ 1 & 1 & 1 \end{array} \right) \sim \left(\begin{array}{cc|c} 1 & 1\text{e}4 & 1\text{e}8 \\ 0 & -1\text{e}4 & -1\text{e}8 \end{array} \right) \implies x = (0, 1\text{e}4).$$

Der Gauß-Algorithmus in Reinform ist also numerisch *instabil*.

Bemerkung 17.31. Ein Ansatz die Stabilität eines Gleichungssystems zu verbessern, ist die Singulärwertzerlegung der Koeffizientenmatrix A durchzuführen. Dabei wird A mit unitären Matrizen U und V von links/rechts multipliziert. Als erste Näherung kann man Permutationsmatrizen für U und V wählen. Dies realisiert Permutationen der Zeilen und Spalten von A .

Satz 17.32 (Pivotisierung). Gegeben sei das lineare Gleichungssystem $Ax = b$ mit $A \in \mathbb{C}^{n \times m}$ und $b \in \mathbb{C}^{n \times 1}$. Die folgende Modifikation des Gauß-Algorithmus auf $(A|b)$ vermindert Rundungsfehler:

(1) Setze $z := 1$ (Zeilenindex).

(2) Für $s = 1, \dots, m$ (Spaltenindex):

- Bestimme ein betragsmäßig größtes Element a_{ij} (Pivot) mit $i \geq z$ und $j \geq s$.
- Ist $a_{ij} = 0$, so sind wir fertig. Anderenfalls tausche i -te mit z -ter Zeile und j -te mit s -ter Spalte. Tausche x_j mit x_s im Lösungsvektor.
- Dividiere z -te Zeile durch a_{zs} .
- Eliminiere wie gewohnt $a_{ws} = 0$ für $w \neq z$.
- Erhöhe z um 1.

Beweis. Wendet man das angegebene Verfahren wie in Satz 6.15 an, so erhält man die besonders einfache erweiterte Matrix

$$M = \begin{pmatrix} 1 & & s_{11} & \cdots & s_{1,m-k} & c_1 \\ & \ddots & \vdots & & \vdots & \vdots \\ & & 1 & s_{k1} & \cdots & s_{k,m-k} & c_k \\ & & & -1 & & & 0 \\ & & & & \ddots & & \vdots \\ & & & & & -1 & 0 \end{pmatrix}$$

mit der Lösungsmenge

$$L = \left\langle \begin{pmatrix} c_1 \\ \vdots \\ c_k \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \left\langle \begin{pmatrix} s_{11} \\ \vdots \\ s_{k1} \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} s_{1,m-k} \\ \vdots \\ s_{k,m-k} \\ 0 \\ \vdots \\ -1 \end{pmatrix} \right\rangle \right\rangle.$$

Die Lösung des ursprünglichen Systems erhält man durch Permutation der Koordinaten gemäß der gewählten Vertauschungen $j \leftrightarrow s$. Die Wahl des Pivots garantiert, dass während des Algorithmus nicht mit großen Zahlen multipliziert wird, die eventuelle Rundungsfehler verstärken würden. \square

Beispiel 17.33. In Beispiel 17.30 vertauschen wir die Spalten:

$$\left(\begin{array}{cc|c} 1 & 1e-4 & 1e4 \\ 1 & 1 & 1 \end{array} \right) \sim \left(\begin{array}{cc|c} 1 & 1e-4 & 1e4 \\ 0 & 1 & -1e4 \end{array} \right) \implies \begin{cases} y = (1e4, -1e4), \\ x = (-1e4, 1e4) \approx (-10^4, 10^4 + 1). \end{cases}$$

Bemerkung 17.34. Für beliebige Körper besagt der folgende Satz, dass man die Pivotisierung theoretisch nur einmal am Anfang vornehmen muss.

Satz 17.35 (LR-Zerlegung⁴). Für jede Matrix $A \in K^{n \times n}$ existieren eine Permutationsmatrix P , eine untere Dreiecksmatrix L mit Einsen auf der Hauptdiagonale und eine obere Dreiecksmatrix R mit $A = PLR$.

Beweis. Induktion nach n : Für $n = 1$ kann man $P = L = 1_1$ und $R = A$ wählen. Sei $n \geq 2$. Ist die erste Spalte von A der Nullvektor, so setzen wir $P_1 = L_1 = 1_n$. Anderenfalls existiert eine Permutationsmatrix P_1 , die die erste Zeile von A mit einer anderen Zeile tauscht, sodass anschließend $a_{11} \neq 0$ gilt. Die Elimination von a_{i1} für $i > 1$ erreichbar man durch Multiplikation mit Elementarmatrizen von links. Das Produkt dieser Elementarmatrizen hat die Form $L_1 = \begin{pmatrix} 1 & 0 \\ * & 1_{n-1} \end{pmatrix}$. Nun gilt

$$L_1 P_1 A = \begin{pmatrix} a_{11} & * \\ 0 & A_2 \end{pmatrix}$$

mit $A_2 \in K^{(n-1) \times (n-1)}$. Nach Induktion existieren P_2, L_2 und R_2 mit $A_2 = P_2 L_2 R_2$. Sei $\hat{P}_2 := \text{diag}(1, P_2)$ und $\hat{L}_2 := \text{diag}(1, L_2)$. Wir können R_2 ebenfalls zu einer oberen Dreiecksmatrix \hat{R}_2 erweitern, sodass $L_1 P_1 A = \hat{P}_2 \hat{L}_2 \hat{R}_2$ gilt. Die Inverse L_1^{-1} hat die gleiche Gestalt wie L_1 (lediglich die Vorzeichen der Einträge unterhalb der Hauptdiagonale sind invertiert). Da \hat{P}_2 die erste Zeile nicht vertauscht, gilt $L_1^{-1} \hat{P}_2 = \hat{P}_2 L'_1$ für eine untere Dreiecksmatrix L'_1 mit Einsen auf der Hauptdiagonale. Insgesamt gilt

$$A = P_1^{-1} L_1^{-1} \hat{P}_2 \hat{L}_2 \hat{R}_2 = P_1 \hat{P}_2 L'_1 \hat{L}_2 \hat{R}_2 = PLR. \quad \square$$

Beispiel 17.36.

$$\begin{aligned} \begin{pmatrix} 0 & 0 & 2 \\ 2 & 0 & 1 \\ 2 & 1 & -1 \end{pmatrix} &= P_{(1,2)} \begin{pmatrix} 2 & 0 & 1 \\ 0 & 0 & 2 \\ 2 & 1 & -1 \end{pmatrix} = P_{(1,2)} \begin{pmatrix} 1 & . & . \\ . & 1 & . \\ 1 & . & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 1 \\ 0 & 0 & 2 \\ 0 & 1 & -2 \end{pmatrix} \\ &= P_{(1,2)} L_1 P_{(2,3)} \begin{pmatrix} 2 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 2 \end{pmatrix} = P_{(1,2)} P_{(2,3)} \begin{pmatrix} 1 & . & . \\ 1 & 1 & . \\ . & . & 1 \end{pmatrix} R = P_{(1,2,3)} LR \end{aligned}$$

Bemerkung 17.37.

- Da es in der Regel viele Möglichkeiten gibt im Gauß-Algorithmus zu pivotieren, ist die LR-Zerlegung nicht eindeutig. Kann man jedoch auf P verzichten, so wird die Zerlegung eindeutig (Aufgabe III.9).
- Sei $A \in \text{GL}(n, K)$. Der folgende Algorithmus liefert eine Variante der LR-Zerlegung, bei der P zwischen L und R steht: Wähle das kleinste j_1 mit $a_{1,j_1} \neq 0$ (existiert, da A invertierbar ist). Durch Linksmultiplikation mit einer unteren Dreiecksmatrix L_1 mit Einsen auf der Hauptdiagonale erreicht man $a_{i,j_1} = 0$ für $i \geq 2$. Analog erhält man $a_{1,j_1} = 1$ und $a_{1j} = 0$ für $j > j_1$ durch Rechtsmultiplikation mit einer oberen Dreiecksmatrix. Nun wählt man das minimal j_2 mit $a_{2,j_2} \neq 0$ und verfährt wie zuvor. Am Ende erhält man eine untere Dreiecksmatrix $L' := L_n \dots L_1$ mit Einsen auf der Hauptdiagonale und eine obere Dreiecksmatrix $R' := R_1 \dots R_n$, sodass $L' A R' = P_\sigma$ die Permutationsmatrix mit $\sigma^{-1}(k) = j_k$ für $k = 1, \dots, n$ ist. Also hat man eine Zerlegung der Form $A = L P R$ (vgl. Aufgabe III.10).
- Der nächste Satz liefert eine weitere Annäherung an die Singulärwertzerlegung.

⁴LR steht für links-rechts. Im Englischen benutzt man LU für lower-upper.

Satz 17.38 (QR-Zerlegung). Für jedes $A \in \mathbb{C}^{n \times m}$ existieren $Q \in U(n, \mathbb{C})$ und eine obere Dreiecksmatrix $R \in \mathbb{C}^{n \times m}$ mit reellen, nicht-negativen Diagonaleinträgen und $A = QR$.

(a) Hat A vollen Rang, so ist R eindeutig bestimmt.

(b) Ist A reell, so kann man $Q \in O(n, \mathbb{R})$ und $R \in \mathbb{R}^{n \times m}$ wählen.

Beweis. Seien a_1, \dots, a_m die Spalten von A und $s := \text{rk}(A)$. Wir wählen von links nach rechts die ersten s linear unabhängigen Spalten a_{i_1}, \dots, a_{i_s} von A . Jede weitere Spalte a_j lässt sich dann als Linearkombination der a_{i_k} mit $i_k < j$ darstellen. Wir ergänzen a_{i_1}, \dots, a_{i_s} mit reellen Vektoren zu einer Basis von \mathbb{C}^n . Das Gram-Schmidt-Verfahren (Bemerkung 13.6) überführt diese Basis in eine Orthonormalbasis q_1, \dots, q_n mit $q_1 = \frac{1}{|a_{i_1}|} a_{i_1}$, $q_2 = \lambda a_{i_1} + \mu a_{i_2}$ mit $\mu \in \mathbb{R}_{>0}$ usw. Durch Umstellen erhält man

$$a_{i_k} = \sum_{j=1}^k \lambda_j q_j \quad (\lambda_k \in \mathbb{R}_{>0}).$$

Indem man die übrigen Spalten von A ebenfalls bzgl. q_1, \dots, q_n darstellt, erhält man eine obere Dreiecksmatrix $R = (r_{ij}) \in \mathbb{C}^{n \times m}$ mit $a_k = \sum_{i=1}^k r_{ik} q_i$ für $k = 1, \dots, m$ und $r_{ii} \in \mathbb{R}_{\geq 0}$ für $i = 1, \dots, n$. Für $Q = (q_1, \dots, q_n) \in U(n, \mathbb{C})$ gilt nun $A = QR$.

(a) O. B. d. A. sei A selbst eine obere Dreiecksmatrix mit positiven Diagonaleinträgen. Für die erste Spalte gilt $a_{11} e_1 = r_{11} q_1$. Aus $a_{11}, r_{11} > 0$ und $|q_1| = 1$ folgt $q_1 = e_1$ und $a_{11} = r_{11}$. Induktiv sei bereits $r_{ij} = a_{ij}$ und $q_j = e_j$ für $i = 1, \dots, n$ und $j = 1, \dots, k-1$ bewiesen. Für die k -te Spalte von A gilt dann

$$\sum_{i=1}^k a_{ik} e_i = r_{kk} q_k + \sum_{i=1}^{k-1} r_{ik} e_i.$$

Dies zeigt $q_{ik} = 0$ für $i > k$. Wegen $[e_i, q_k] = [q_i, q_k] = 0$ für $i < k$ folgt $q_k = e_k$ und $r_{ik} = a_{ik}$ für $i = 1, \dots, n$. Induktiv erhält man $A = R$ (aber nicht unbedingt $Q = 1_n$ falls $n > m$).

(b) Ist A reell, so lässt sich das Gram-Schmidt-Verfahren in \mathbb{R}^n durchführen (man beachte, dass wir a_{i_1}, \dots, a_{i_s} mit reellen Vektoren zu einer Basis ergänzt haben). \square

Bemerkung 17.39.

(i) Ist $\text{rk}(A) = n \leq m$, so zeigt der Beweis, dass Q und R in der QR-Zerlegung eindeutig bestimmt sind.

(ii) Im Fall $n = m = 1$ erhält man die *Polardarstellung* einer komplexen Zahl $z = e^{i\varphi}|z|$. Im Allgemeinen hat die Polardarstellung einer Matrix eine andere Struktur (Aufgabe III.11).

(iii) Wir sehen in Abschnitt 17.7, dass die direkte Anwendung des Gram-Schmidt-Verfahrens instabil ist.

Beispiel 17.40.

$$\begin{pmatrix} 1 & i \\ 1 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \sqrt{2} & i-1 \\ \sqrt{2} & 1-i \end{pmatrix} \begin{pmatrix} \sqrt{2} & (1+i)/\sqrt{2} \\ 0 & 1 \end{pmatrix}$$

Bemerkung 17.41. In der Praxis entstehen durch ungenaue Messungen oft überbestimmte Gleichungssysteme $Ax = b$ ohne exakte Lösung. Die Güte einer Näherungslösung \tilde{x} lässt sich durch $|\tilde{x} - x|$ bzw. $|\tilde{x} - x|^2 = \sum (\tilde{x}_i - x_i)^2$ quantifizieren.

Satz 17.42 (Methode der kleinsten Quadrate). Sei $A \in \mathbb{C}^{n \times m}$ mit vollem Rang $m \leq n$ und $b \in \mathbb{C}^{n \times 1}$. Dann besitzt das Gleichungssystem $A^*Ax = A^*b$ eine eindeutige Lösung \tilde{x} und für alle $x \neq \tilde{x}$ gilt $|A\tilde{x} - b| < |Ax - b|$.

Beweis. Da A vollen Rang $m \leq n$ hat, ist A^*A positiv definit, also insbesondere invertierbar. Dies zeigt die Existenz und Eindeutigkeit von \tilde{x} . Für $y \neq 0$ gilt

$$[Ay, A\tilde{x} - b] = y^* A^* (A\tilde{x} - b) = 0 = [A\tilde{x} - b, Ay],$$

$$|A(\tilde{x} + y) - b|^2 = [(A\tilde{x} - b) + Ay, (A\tilde{x} - b) + Ay] = |A\tilde{x} - b|^2 + |Ay|^2 > |A\tilde{x} - b|^2. \quad \square$$

Bemerkung 17.43.

- (a) In der Neujahrsnacht 1801 entdeckte der Astronom PIAZZI den Zwergplaneten *Ceres*. Die Position wurde 40 Tage aufgezeichnet bis *Ceres* aus dem Blickfeld verschwand. Gauß hat mit der Methode der kleinsten Quadrate die Bahn von *Ceres* mit den vorhandenen Messdaten extrapoliert und konnte so die Position erfolgreich vorhersagen. Er erlangte dadurch internationale Bekanntheit.
- (b) Nach Beispiel 17.28 gilt $\tilde{x} = A^+b$ in der Situation von Satz 17.42.
- (c) Die Methode der kleinsten Quadrate führt auf das sogenannte *Normalgleichungssystem* $A^*Ax = A^*b$ mit positiv definiten Koeffizientenmatrix. Dafür gibt es spezielle Verfahren. Die folgende Zerlegung präzisiert Lemma 12.40.

Satz 17.44 (CHOLESKY-Zerlegung). Genau dann ist $A \in \mathbb{C}^{n \times n}$ positiv definit,⁵ falls eine obere Dreiecksmatrix $R \in \mathbb{C}^{n \times n}$ mit reellen, positiven Hauptdiagonaleinträgen und $A = R^*R$ existiert. Gegebenenfalls ist R eindeutig bestimmt.

Beweis. Für $R \in GL(n, \mathbb{C})$ ist R^*R bekanntlich positiv definit. Sei nun umgekehrt A positiv definit und $\sqrt{A} = QR$ die (eindeutige) QR-Zerlegung der Wurzel aus A (Aufgabe II.18). Dann gilt

$$A = \sqrt{A}^2 = \sqrt{A}^* \sqrt{A} = R^* Q^* Q R = R^* R.$$

Sei $P \in \mathbb{C}^{n \times n}$ eine weitere obere Dreiecksmatrix mit positiver Hauptdiagonale und $A = P^*P$. Dann ist $P^{-*}R^* = P^{-*}AR^{-1} = PR^{-1}$ eine obere und untere Dreiecksmatrix, also eine Diagonalmatrix mit positiven Hauptdiagonaleinträgen. Wegen

$$(PR^{-1})^* PR^{-1} = R^{-*} AR^{-1} = 1_n$$

ist PR^{-1} unitär. Es folgt $P = R$. □

Bemerkung 17.45.

- (a) Die Matrix $R = (r_{ij})$ lässt sich ohne den Umweg über die QR-Zerlegung iterativ berechnen: Wegen $a_{11} = e_1 A e_1^t > 0$ ist $r_{11} = \sqrt{a_{11}}$. Sei bereits $R_1 \in \mathbb{C}^{(n-1) \times (n-1)}$ mit $A_{nn} = R_1^* R_1$ bestimmt. Mit dem Ansatz $A = \begin{pmatrix} A_{nn} & a \\ a^* & a_{nn} \end{pmatrix}$ und $R = \begin{pmatrix} R_1 & v \\ 0 & r_{nn} \end{pmatrix}$ erhält man $R_1^* v = a$ und $r_{nn} = \sqrt{a_{nn} - v^* v}$.

⁵im Sinn von Aufgabe II.18

- (b) Das System $Ax = b$ mit $A = R^*R$ kann man bequem in zwei Schritten lösen: Zunächst löst man $R^*y = b$ durch *Vorwärtssubstitution*:

$$\begin{pmatrix} \overline{r_{11}} & & 0 \\ \vdots & \ddots & \\ \overline{r_{1n}} & \cdots & \overline{r_{nn}} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \implies \begin{cases} y_1 = \frac{b_1}{\overline{r_{11}}} \\ y_2 = \frac{b_2 - \overline{r_{12}}y_1}{\overline{r_{22}}} \\ \vdots \end{cases}$$

und anschließend $Rx = y$ durch *Rückwärtssubstitution*:

$$\begin{pmatrix} r_{11} & \cdots & r_{1n} \\ & \ddots & \vdots \\ 0 & & r_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \implies \begin{cases} x_n = \frac{y_n}{r_{nn}} \\ x_{n-1} = \frac{y_{n-1} - r_{n-1,n}x_n}{r_{n-1,n-1}} \\ \vdots \end{cases}$$

Beispiel 17.46. Wir betrachten das überbestimmte System

$$Ax = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 0 & -1 \end{pmatrix} x = \begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix}.$$

Das Normalgleichungssystem lautet $\begin{pmatrix} 2 & 3 \\ 3 & 6 \end{pmatrix} x = -\begin{pmatrix} 1 \\ 4 \end{pmatrix}$. Die Matrix R erhält man durch $r_{11} = \sqrt{2}$, $r_{12} = \frac{3}{\sqrt{2}}$ und $r_{22} = \sqrt{6 - 9/2} = \sqrt{3/2}$. Nun ist $y = -\frac{1}{\sqrt{2}}(1, 5/\sqrt{3})^t$ die Lösung von

$$\begin{pmatrix} \sqrt{2} & 0 \\ 3/\sqrt{2} & \sqrt{3/2} \end{pmatrix} y = R^t y = -\begin{pmatrix} 1 \\ 4 \end{pmatrix}$$

und $x = (2, -5/3)^t$ die Lösung von

$$\begin{pmatrix} \sqrt{2} & 3/\sqrt{2} \\ 0 & \sqrt{3/2} \end{pmatrix} x = Rx = y = -\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 5/\sqrt{3} \end{pmatrix}.$$

Für das ursprüngliche System erhält man die Näherung $Ax = \frac{1}{3}(1, -4, 5)^t$. In der Praxis werden die Wurzelausdrücke durch Gleitkommazahlen approximiert.

17.4 Iterative Verfahren

Bemerkung 17.47.

- (a) Für großen Matrizen sind Faktorisierungsmethoden wie die QR-Zerlegung langsam und speicherintensiv. Zum Lösen von linearen Gleichungssystemen kann man in diesen Fällen auf schnellere iterative Verfahren ausweichen. Für Konvergenzbeweise benötigen wir einige Hilfsmittel aus der Analysis.
- (b) Bisher gingen Normen stets aus Skalarprodukten hervor (Definition 11.2 und Definition 13.2). Sie können aber auch direkt über die Eigenschaften aus Lemma 13.5 definiert werden.

Definition 17.48. Sei V ein \mathbb{C} -Vektorraum. Eine Abbildung $V \rightarrow \mathbb{R}$, $v \mapsto \|v\|$ heißt *Norm*, falls für alle $v, w \in V$ und $\lambda \in \mathbb{C}$ gilt:

- (a) $\|v\| \geq 0$ mit Gleichheit genau dann, wenn $v = 0$ (*positiv definit*).

- (b) $\|\lambda v\| = |\lambda| \|v\|$ (*Homogenität*).
 (c) $\|v + w\| \leq \|v\| + \|w\|$ (*Dreiecksungleichung*).

Beispiel 17.49.

- (a) Ist $[\cdot, \cdot]$ ein Skalarprodukt auf V , so definiert $\|v\| := \sqrt{[v, v]}$ eine Norm nach Lemma 13.5.
 (b) Auf $V = \mathbb{C}^n$ definiert $\|x\|_\infty := \max\{|x_1|, \dots, |x_n|\}$ eine Norm. Diese Norm kann nicht aus einem Skalarprodukt gewonnen werden, denn die Parallelogrammgleichung

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

aus Beispiel 11.4 ist für $x = (1, 0)$ und $y = (0, 1)$ nicht erfüllt.

- (c) Für $V = \mathbb{C}^n$ und $p \geq 1$ definiert

$$\|x\|_p := \sqrt[p]{\sum_{i=1}^n |x_i|^p}$$

eine Norm (Aufgabe III.13). Für $p = 1$ erhält man $\|x\|_1 = |x_1| + \dots + |x_n|$. Für $p = 2$ erhält man die „euklidische“ Norm $\|x\| = \sqrt{[x, x]}$. Mit $p \rightarrow \infty$ nähert sich $\|\cdot\|_p$ der Norm $\|\cdot\|_\infty$ an.

- (d) Ist $\|\cdot\|$ eine Norm auf \mathbb{C}^n und $A \in \text{GL}(n, \mathbb{C})$, so definiert auch $\|x\|_A := \|Ax^t\|$ eine Norm auf \mathbb{C}^n .

Definition 17.50. Normen $\|\cdot\|$ und $\|\cdot\|'$ auf einem \mathbb{C} -Vektorraum V heißen *äquivalent*, falls $\lambda, \mu > 0$ mit

$$\lambda \|v\| \leq \|v\|' \leq \mu \|v\|$$

für alle $v \in V$ existieren.

Lemma 17.51. *Je zwei Normen auf einem endlich-dimensionalen \mathbb{C} -Vektorraum sind äquivalent.*

Beweis. O. B. d. A. sei $V = \mathbb{C}^n$. Es genügt zu zeigen, dass jede Norm $\|\cdot\|$ zu $\|\cdot\|_1$ äquivalent ist. Für $\lambda := \max\{\|e_1\|, \dots, \|e_n\|\} > 0$ gilt

$$\|x\| = \left\| \sum_{i=1}^n x_i e_i \right\| \leq \sum_{i=1}^n |x_i| \|e_i\| \leq \lambda \|x\|_1$$

für alle $x \in \mathbb{R}^n$. Sei $y_i := \bar{x}_i / |x_i|$ für $i = 1, \dots, n$. Nach der Cauchy-Schwarz-Ungleichung gilt $\|x\|_1 = [x, y] \leq \|x\| \|y\| = \sqrt{n} \|x\|$. Aus der (umgekehrten) Dreiecksungleichung folgt

$$\| \|x\| - \|y\| \| \leq \|x - y\| \leq \lambda \|x - y\|_1 \leq \lambda \sqrt{n} \|x - y\|.$$

Dies zeigt, dass die Abbildung $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $x \mapsto \|x\|$ bzgl. der euklidischen Norm stetig ist. Daher nimmt f auf der kompakten Menge $\{x \in \mathbb{R}^n : \|x\|_1 = 1\} \neq \emptyset$ sein Minimum $\mu > 0$ an. Für $x \neq 0$ gilt nun

$$\|x\| = \|x\|_1 f(\|x\|_1^{-1} x) \geq \|x\|_1 \mu. \quad \square$$

Bemerkung 17.52. Bekanntlich ist die euklidische Norm auf \mathbb{C}^n abgeschlossen (d. h. jede Cauchyfolge konvergiert). Nach Lemma 17.51 ist daher jede Norm auf einem endlich-dimensionalen \mathbb{C} -Vektorraum V abgeschlossen, d. h. V ist ein *Banachraum*.

Satz 17.53 (BANACHs Fixpunktsatz). Sei $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine Kontraktion bzgl. einer Norm $\|\cdot\|$, d. h. es gibt eine Konstante $c < 1$ mit $\|f(x) - f(y)\| \leq c\|x - y\|$ für alle $x, y \in \mathbb{R}^n$. Dann konvergiert die Folge $x_{k+1} := f(x_k)$ für alle $x_0 \in \mathbb{R}^n$ gegen den einzigen Fixpunkt von f .

Beweis. Für $k, l \in \mathbb{N}$ gilt

$$\begin{aligned} \|x_{k+l} - x_k\| &= \left\| \sum_{i=0}^{l-1} (x_{k+i+1} - x_{k+i}) \right\| \leq \sum_{i=0}^{l-1} \|x_{k+i+1} - x_{k+i}\| = \sum_{i=0}^{l-1} \|f^{k+i}(x_1) - f^{k+i}(x_0)\| \\ &\leq \|x_1 - x_0\| \sum_{i=0}^{l-1} c^{k+i} = \|x_1 - x_0\| c^k \frac{1 - c^l}{1 - c} \xrightarrow{k \rightarrow \infty} 0, \end{aligned}$$

d. h. ist $(x_k)_k$ eine Cauchyfolge. Nach Bemerkung 17.52 existiert der Grenzwert $\tilde{x} := \lim_{k \rightarrow \infty} x_k$. Da f als Kontraktion (gleichmäßig) stetig ist (setze $\delta := \frac{\epsilon}{c}$), gilt $f(\tilde{x}) = \lim_{k \rightarrow \infty} f(x_k) = \tilde{x}$. Wäre auch $y \neq \tilde{x}$ ein Fixpunkt von f , so hätte man den Widerspruch

$$\|\tilde{x} - y\| = \|f(\tilde{x}) - f(y)\| \leq c\|\tilde{x} - y\| < \|\tilde{x} - y\|. \quad \square$$

Bemerkung 17.54.

- (a) Eine lineare Abbildung $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ ist genau dann eine Kontraktion, wenn ein $c < 1$ mit $\|f(x)\| \leq c\|x\|$ für alle $x \in \mathbb{R}^n$ existiert.
- (b) Das nächste Lemma liefert eine explizite Äquivalenz zwischen der euklidischen Norm und der in Beispiel 17.49 eingeführten Norm $|\cdot|_A$.

Lemma 17.55. Sei $A \in \mathbb{C}^{n \times n}$ normal mit Eigenwerten $\lambda_1, \dots, \lambda_n \in \mathbb{C}$. Sei

$$\lambda_{\min} := \min\{|\lambda_i| : i = 1, \dots, n\}, \quad \lambda_{\max} := \max\{|\lambda_i| : i = 1, \dots, n\}.$$

Für $x \in \mathbb{C}^n$ gilt $\lambda_{\min}|x| \leq |Ax| \leq \lambda_{\max}|x|$ und $|x^*Ax| \leq \lambda_{\max}|x|^2$. Ist A positiv definit, so gilt $|x^*Ax| \geq \lambda_{\min}|x|^2$.

Beweis. Sei $S \in U(n, \mathbb{C})$ mit $S^*AS = \text{diag}(\lambda_1, \dots, \lambda_n)$ (Spektralsatz). Wegen $|x| = |Sx|$ und $|Ax| = |S^*Ax|$ können wir $A = \text{diag}(\lambda_1, \dots, \lambda_n)$ annehmen. Dann gilt

$$|Ax|^2 = (Ax)^*Ax = \sum_{i=1}^n |\lambda_i|^2 |x_i|^2 \leq \lambda_{\max}^2 \sum_{i=1}^n |x_i|^2 = \lambda_{\max}^2 |x|^2$$

und analog $|Ax| \geq \lambda_{\min}|x|$. Ebenso ist

$$|x^*Ax| = \left| \sum_{i=1}^n \lambda_i |x_i|^2 \right| \leq \sum_{i=1}^n |\lambda_i| |x_i|^2 \leq \lambda_{\max} |x|^2.$$

Ist A positiv definit, so sind $\lambda_1, \dots, \lambda_n$ reell und positiv. Dann folgt wie zuvor $|x^tAx| \geq \lambda_{\min}|x|^2$. \square

Satz 17.56 (GAUSS-SEIDEL-Verfahren). Sei $A = L + R \in \mathbb{C}^{n \times n}$ positiv definit, wobei L eine untere Dreiecksmatrix und R eine strikte obere Dreiecksmatrix ist. Sei $b \in \mathbb{C}^{n \times 1}$. Dann konvergiert die Folge

$$x_{k+1} := L^{-1}b - L^{-1}Rx_k$$

für alle Startwerte $x_0 \in \mathbb{C}^{n \times 1}$ gegen die Lösung von $Ax = b$.

Beweis. Da A reelle, positive Diagonaleinträge besitzt (Bemerkung 12.39) und R eine strikte obere Dreiecksmatrix ist, muss L invertierbar sein. Für $A\tilde{x} = b$ gilt

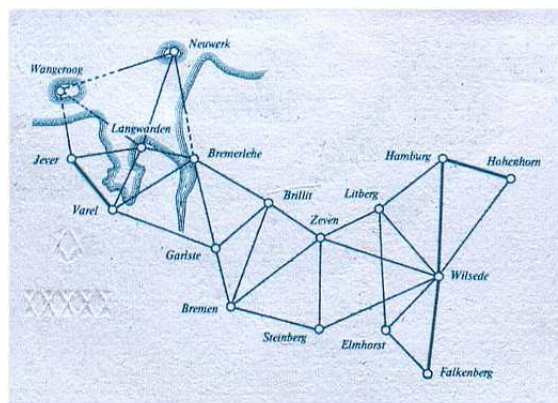
$$\tilde{x} = L^{-1}L\tilde{x} = L^{-1}(b - R\tilde{x}) = L^{-1}b - L^{-1}R\tilde{x},$$

d. h. \tilde{x} ist ein Fixpunkt der Abbildung $f(x) := L^{-1}b - L^{-1}Rx$. Da A positiv definit ist, definiert $[x, y] := y^*Ax$ ein Skalarprodukt mit Norm $\|x\| := \sqrt{[x, x]}$. Nach Banachs Fixpunktsatz genügt es zu zeigen, dass f bzgl. $\|\cdot\|$ eine Kontraktion ist. Da $L^{-1}b$ konstant ist, können wir die lineare Abbildung $g(x) := L^{-1}Rx$ anstelle von f betrachten. Nach Bemerkung 17.54 und Lemma 17.55 müssen wir zeigen, dass jeder Eigenwert $\lambda \in \mathbb{C}$ von $L^{-1}R$ betragsmäßig kleiner als 1 ist. Sei $x \in \mathbb{C}^{n \times 1}$ ein entsprechender Eigenvektor und $y := x + g(x) = L^{-1}Ax \neq 0$. Nach Voraussetzung ist $D := L^* - R = \text{diag}(a_{11}, \dots, a_{nn})$ positiv definit. Also gilt

$$\begin{aligned} |\lambda|\|x\| &= \|g(x)\| = \|x - y\| = \|x\|^2 - x^*A^*y - y^*Ax + y^*Ay \\ &= \|x\|^2 - y^*L^*y - y^*Ly + y^*(L + R)y = \|x\|^2 - y^*Dy < \|x\|. \end{aligned}$$

Dies zeigt $|\lambda| < 1$. □

Bemerkung 17.57. Gauß hat mit diesem Verfahren das Königreich Hannover vermessen, wovon eine Zeichnung auf dem 10-DM-Schein zeugt:



Er schrieb in einem Brief:

„Ich empfehle Ihnen diesen Modus zur Nachahmung. Schwerlich werden Sie je wieder direct [sic] eliminieren, wenigstens nicht, wenn Sie mehr als 2 Unbekannte haben. Das indirecte Verfahren lässt sich halb im Schlafe ausführen, oder man kann während desselben an andere Dinge denken.“

Beispiel 17.58. Das System $H_4x = (1, 1, 1, 1)^t$ mit der schlechte konditionierten Hilbert-Matrix H_4 hat die Lösung $x = (-4, 60, -180, 140)^t$. Das Gauß-Seidel-Verfahren konvergiert nur langsam:

$$\begin{aligned} x_0 &:= (0, 0, 0, 0) \\ x_{10} &\approx (-1.90, -2.97, 4.45, 8.06) \\ x_{100} &\approx (2.84, -14.20, -5.17, 27.94) \\ x_{1000} &\approx (-1.10, 28.86, -106.96, 93.31) \\ x_{10000} &\approx (-4.000, 59.995, -179.988, 139.992) \end{aligned}$$

17.5 Matrixnormen

Bemerkung 17.59. Die Menge der $n \times m$ -Matrizen bildet bekanntlich einen Vektorraum der Dimension nm . Es liegt daher nahe, die in Definition 17.48 definierten Normen auch für Matrizen einzuführen.

Definition 17.60. Eine Abbildung $\mathbb{C}^{n \times m} \rightarrow \mathbb{R}$, $A \mapsto \|A\|$ heißt *Matrixnorm*, falls für $A, B \in \mathbb{C}^{n \times m}$ und $\lambda \in \mathbb{C}$ gilt

- $\|A\| \geq 0$ mit Gleichheit genau dann, wenn $A = 0_{n \times m}$.
- $\|\lambda A\| = |\lambda| \|A\|$.
- $\|A + B\| \leq \|A\| + \|B\|$.

Im Fall $n = m$ nennen wir die Norm *submultiplikativ*, falls $\|AB\| \leq \|A\| \|B\|$ gilt.

Beispiel 17.61.

(a) Die „euklidische“ Norm auf $\mathbb{C}^{n \times m}$ nennt man *Frobenius-Norm*

$$|A| := \sqrt{\sum_{i=1}^n \sum_{j=1}^m |a_{ij}|^2} = \sqrt{\operatorname{tr}(A^* A)}.$$

Für $A, B \in \mathbb{C}^{n \times n}$ folgt aus der Cauchy-Schwarz-Ungleichung

$$|AB|^2 = \sum_{i,j} \left| \sum_{k=1}^n a_{ik} b_{kj} \right|^2 \leq \sum_{i,j} \left(\sum_{k=1}^n |a_{ik}| |b_{kj}| \right)^2 \leq \sum_{i,j} \left(\sum_{k=1}^n |a_{ik}|^2 \right) \left(\sum_{s=1}^n |b_{sj}|^2 \right) = |A|^2 |B|^2,$$

d. h. $|\cdot|$ ist submultiplikativ. Sind $\sigma_1, \dots, \sigma_k$ die Singulärwerte von A , so ist

$$|A| = \sqrt{\operatorname{tr}(A^* A)} = \sqrt{\sigma_1^2 + \dots + \sigma_k^2}$$

nach Bemerkung 17.22.

(b) Jede „natürliche“⁶ Vektornorm $\|\cdot\|$ induziert durch⁷

$$\|A\| = \max_{0 \neq x \in \mathbb{C}^{m \times 1}} \frac{\|Ax\|}{\|x\|} = \max_{\|x\|=1} \|Ax\|$$

eine Matrixnorm (Aufgabe III.14). Im Fall $m = 1$ stimmen die Normen überein. Für alle $x \in \mathbb{C}^{m \times 1}$ gilt $\|Ax\| \leq \|A\| \|x\|$. Für $A, B \in \mathbb{C}^{n \times n}$ folgt

$$\|AB\| = \max_{\|x\|=1} \|ABx\| \leq \max_{\|x\|=1} \|A\| \|Bx\| = \|A\| \|B\|,$$

d. h. $\|\cdot\|$ ist submultiplikativ.

(c) Nach (dem Beweis von) Folgerung 17.23 ist $\|A\|_2$ der größte Singulärwert von A (oder 0). Insbesondere ist $\|A\|_2 \leq |A|$ mit Gleichheit genau dann, wenn $\operatorname{rk}(A) \leq 1$.

(d) Nicht jede Matrixnorm ist submultiplikativ: Für $\|A\|_{\max} := \max_{i,j} |a_{ij}|$ und $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ gilt $\|A^2\|_{\max} = 2 > 1 = \|A\|_{\max}^2$.

⁶definiert für beliebige Dimension

⁷Wie in Bemerkung 17.19 wird das Maximum tatsächlich angenommen.

Lemma 17.62. Für $A \in \mathbb{C}^{n \times m}$ gilt

- (a) $\|A\|_1 = \max_{1 \leq j \leq m} \sum_{i=1}^n |a_{ij}|$ (Spaltensummen-Norm).
- (b) $\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^m |a_{ij}| = \|A^t\|_1$ (Zeilensummen-Norm).
- (c) $\|A\|_2 \leq |A| \leq \sqrt{\min\{m, n\}} \|A\|_2$.
- (d) $\frac{1}{\sqrt{m}} \|A\|_\infty \leq \|A\|_2 \leq \sqrt{n} \|A\|_\infty$.
- (e) $\frac{1}{\sqrt{n}} \|A\|_1 \leq \|A\|_2 \leq \frac{1}{\sqrt{m}} \|A\|_1$.

Beweis.

(a) Nach der Dreiecksungleichung gilt

$$\|A\|_1 = \max_{\|x\|_1=1} \sum_{i=1}^n \left| \sum_{j=1}^m a_{ij} x_j \right| \leq \max_{\sum |x_i|=1} \sum_{j=1}^m |x_j| \sum_{i=1}^n |a_{ij}| \leq \max_{1 \leq j \leq m} \sum_{i=1}^n |a_{ij}|.$$

Wird das Maximum rechts für j angenommen, so erhält man mit $x = e_j$ Gleichheit.

(b) Es gilt

$$\|A\|_\infty = \max_{\|x\|_\infty=1} \max_{1 \leq i \leq n} \left| \sum_{j=1}^m a_{ij} x_j \right| \leq \max_i \max_{|x_i|=1} \sum_{j=1}^m |a_{ij}| |x_j| \leq \max_{1 \leq i \leq n} \sum_{j=1}^m |a_{ij}|.$$

Wird das Maximum rechts für i angenommen, so erhält man mit $x_j = \bar{a}_{ij}/|a_{ij}|$ für $a_{ij} \neq 0$ und $x_j = 0$ sonst, Gleichheit.

(c) Sind $\sigma_1 \geq \dots \geq \sigma_k$ die Singulärwerte von A , so gilt

$$\|A\|_2 = \sigma_1 \leq \sqrt{\sigma_1^2 + \dots + \sigma_k^2} = |A| \leq \sqrt{k} \sigma_1 \leq \sqrt{\min\{n, m\}} \|A\|_2.$$

(d) Für $x \in \mathbb{C}^n$ gilt

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i| \leq \sqrt{|x_1|^2 + \dots + |x_n|^2} = |x| \leq \sqrt{n} \|x\|_\infty.$$

Dies zeigt

$$\begin{aligned} \|A\|_\infty &= \max_{0 \neq x \in \mathbb{C}^{m \times 1}} \frac{\|Ax\|_\infty}{\|x\|_\infty} \leq \max_{0 \neq x \in \mathbb{C}^{m \times 1}} \frac{|Ax|}{|x|/\sqrt{m}} = \sqrt{m} \|A\|_2, \\ \|A\|_2 &= \max_{0 \neq x \in \mathbb{C}^{m \times 1}} \frac{|Ax|}{|x|} \leq \max_{0 \neq x \in \mathbb{C}^{m \times 1}} \frac{\sqrt{n} \|Ax\|_\infty}{\|x\|_\infty} = \sqrt{n} \|A\|_\infty. \end{aligned}$$

(e) Aus der Singulärwertzerlegung ergibt sich $\|A^t\|_2 = \|A\|_2$. Daher folgt die Behauptung aus (b) und (d). \square

Lemma 17.63. Die Konditionszahl von $A \in \text{GL}(n, \mathbb{C})$ ist $\kappa(A) = \|A\|_2 \|A^{-1}\|_2$.

Beweis. Nach Definition ist

$$\begin{aligned}\|A\|_2 &= \max_{x \in \mathbb{C}^{n \times 1} \setminus \{0\}} |Ax|/|x|, \\ \|A^{-1}\|_2 &= \max_x |A^{-1}x|/|x| = \max_y |y|/|Ay| = (\min_y |Ay|/|y|)^{-1}.\end{aligned}$$

Die Behauptung folgt aus der Definition der Konditionszahl. \square

Definition 17.64. Eine Folge von Matrizen $(A_k)_k \in \mathbb{C}^{n \times m}$ konvergiert gegen eine Matrix $A \in \mathbb{C}^{n \times m}$, falls $\lim_{k \rightarrow \infty} \|A - A_k\| = 0$ für eine Matrixnorm gilt. Da alle Normen nach Lemma 17.51 äquivalent sind, hängt diese Definition nicht von der Wahl der Norm ab. Ggf. schreiben wir $A = \lim_{k \rightarrow \infty} A_k$

Bemerkung 17.65.

- (a) Wie in der Analysis zeigt man, dass die Folge $A_k = (a_{ij}^{(k)})$ genau dann gegen $A = (a_{ij})$ konvergiert, wenn $\lim_{k \rightarrow \infty} a_{ij}^{(k)} = a_{ij}$ für alle i, j gilt.
- (b) Offenbar sind die Abbildungen $A \rightarrow A^t$, $A \rightarrow \bar{A}$ und tr stetig (bzgl. jeder Matrixnorm). Bekanntlich sind Addition und Multiplikation komplexer Zahlen stetige Abbildungen. Daher ist auch die Matrizenmultiplikation stetig. Für $A = \lim_{k \rightarrow \infty} A_k$ und $B = \lim_{k \rightarrow \infty} B_k$ gilt also $\lim_{k \rightarrow \infty} A_k B_k = AB$. Die Leibniz-Formel zeigt, dass auch \det stetig ist. Die Darstellung $A^{-1} = \det(A)^{-1} \tilde{A}$ (Satz 9.22) liefert, dass die Inversenbildung $A \mapsto A^{-1}$ stetig ist.
- (c) Konvergiert die Folge der Partialsummen $B_k := \sum_{i=1}^k A_i$, so schreiben wir $\sum_{k=1}^{\infty} A_k := \lim_{k \rightarrow \infty} B_k$ wie üblich.

Satz 17.66. Für $A \in \mathbb{C}^{n \times n}$ konvergiert die Folge $(A^k)_k$ genau dann, wenn die folgenden beiden Aussagen gelten:

- (a) Für jeden Eigenwert λ von A gilt $|\lambda| < 1$ oder $\lambda = 1$.
- (b) Ist 1 ein Eigenwert, so stimmt die algebraische Vielfachheit mit der geometrischen Vielfachheit überein.

Beweis. O. B. d. A. sei $A = \text{diag}(J_{n_1}(\lambda_1), \dots, J_{n_s}(\lambda_s))$ in Jordan-Normalform. Gilt $|\lambda_i| > 1$, so besitzt A^k den Eintrag $\lambda_i^k \xrightarrow{k \rightarrow \infty} \infty$. Sei nun $|\lambda_i| < 1$. Die Beträge der Einträge von $J_{n_i}(\lambda_i)^k$ haben nach Lemma 14.41 die Form

$$\binom{k}{l} |\lambda_i|^{k-l} \leq k^n |\lambda_i|^{k-n} \xrightarrow{k \rightarrow \infty} 0 \quad (0 \leq l < n).$$

Dies zeigt $J_{n_i}(\lambda_i)^k \rightarrow 0$. Sei nun $|\lambda_i| = 1$. Damit A^k konvergieren kann, muss ein k mit $\lambda_i^k = \lambda_i^{k+1} = \dots$ existieren, d. h. $\lambda_i = 1$. Ist $n_i > 1$, so gilt $J_{n_i}(\lambda_i)_{21} = k \rightarrow \infty$. Also muss $n_i = 1$ für alle i mit $\lambda_i = 1$ gelten. O. B. d. A. sei $\lambda_1 = \dots = \lambda_t = 1 > |\lambda_j|$ für $j > t$. Offenbar konvergiert (A^k) dann gegen $\text{diag}(1_t, 0_{n-t})$. \square

Folgerung 17.67. Die Bedingungen aus Satz 17.66 seien für $A \in \mathbb{C}^{n \times n}$ mit $\dim E_1(A) = 1$ erfüllt. Seien $v, w \in \mathbb{C}^{n \times 1}$ Eigenvektoren von A bzw. A^* zum Eigenwert 1 mit $[v, w] = 1$. Dann gilt $\lim_{k \rightarrow \infty} A^k = vw^*$.

Beweis. Wegen $\chi_{A^*} = \overline{\chi_A}$ ist 1 tatsächlich ein Eigenwert von A^* . Bekanntlich (zum Beispiel nach der Jordan-Normalform) existiert ein $S \in GL(n, \mathbb{C})$ mit erster Spalte v , sodass $S^{-1}AS = \text{diag}(1, B)$ mit $\lim_{k \rightarrow \infty} B^k = 0$. Sei u die erste Zeile von S^{-1} . Dann gilt

$$C := \lim_{k \rightarrow \infty} A^k = \lim_{k \rightarrow \infty} S \text{diag}(1, B^k) S^{-1} = S \text{diag}(1, 0_{n-1}) S^{-1} = vu.$$

Andererseits ist $w = C^*w = u^*v^*w = u^*[v, w] = u^*$. □

17.6 Eigenwertberechnung

Bemerkung 17.68.

- (a) Die Determinante ist im Allgemeinen schlecht konditioniert (d. h. sensibel gegenüber kleinen Änderungen):

$$\det \begin{pmatrix} 1 & 33 \\ 3 & 100 \end{pmatrix} = 1 \qquad \det \begin{pmatrix} 1.1 & 33 \\ 3 & 100 \end{pmatrix} = 11$$

Da $\pm \det(A)$ das Absolutglied von χ_A ist, ist auch χ_A schlecht konditioniert. Noch schlechter ist μ_A konditioniert, denn hier können kleinere Änderungen sogar $\deg \mu_A$ ändern.

- (b) Die Eigenwerte als Nullstellen von χ_A (oder μ_A) sind ebenfalls schlecht konditioniert:

$$\begin{array}{ll} X^2 - 21X + 110 & \text{Nullstellen: } 10, 11 \\ X^2 - 21.1X + 110 & \text{Nullstellen: } \approx 9.41, 11.69 \\ X^2 - 20.9X + 110 & \text{Nullstellen: } \approx 10.45 \pm 0.89i \end{array}$$

Abweichungen an A wirken sich also „doppelt“ auf die Eigenwertberechnung aus. Außerdem gibt es keine explizite Formel für die Nullstellen eines Polynoms vom Grad ≥ 5 . Dennoch ist die Bestimmung der Eigenwerte von *normalen* Matrizen – ohne den Umweg über χ_A – gut konditioniert nach folgendem Satz.

Satz 17.69 (BAUER-FIKE). Sei $A \in \mathbb{C}^{n \times n}$ diagonalisierbar mit $S^{-1}AS = \text{diag}(\lambda_1, \dots, \lambda_n)$. Sei $X \in \mathbb{C}^{n \times n}$ und $\mu \in \mathbb{C}$ ein Eigenwert von $A + X$. Dann gilt

$$\min_{i=1, \dots, n} |\mu - \lambda_i| \leq \kappa(S) \|X\|_2.$$

Insbesondere ist $\min_{i=1, \dots, n} |\mu - \lambda_i| \leq \|X\|_2$, wenn A normal ist.

Beweis. O. B. d. A. sei $\mu \notin \{\lambda_1, \dots, \lambda_n\}$. Sei $D := \text{diag}(\lambda_1, \dots, \lambda_n)$. Dann ist $D - \mu 1_n$ invertierbar. Nach Voraussetzung ist $A + X - \mu 1_n$ nicht invertierbar. Daher ist auch

$$(D - \mu 1_n)^{-1} S^{-1} (A - \mu 1_n + X) S = 1_n + (D - \mu 1_n)^{-1} S^{-1} X S$$

nicht invertierbar. Aus Satz 18.8 und Lemma 17.63 folgt

$$1 \leq \rho((D - \mu 1_n)^{-1} S^{-1} X S) \leq \|(D - \mu 1_n)^{-1} S^{-1} X S\|_2 \leq \|(D - \mu 1_n)^{-1}\|_2 \kappa(S) \|X\|_2.$$

Nach Beispiel 17.61 ist $\|(D - \mu 1_n)^{-1}\|_2 = \max_{i=1, \dots, n} (|\lambda_i - \mu|)^{-1}$. Dies zeigt

$$\min_{i=1, \dots, n} |\mu - \lambda_i| = \|(D - \mu 1_n)^{-1}\|_2^{-1} \leq \kappa(S) \|X\|_2.$$

Ist A normal, so kann man $S \in U(n, \mathbb{C})$ nach dem Spektralsatz wählen. Dann ist $\kappa(S) = 1$. □

Satz 17.70 (Potenzmethode). Sei $A \in \mathbb{C}^{n \times n}$ mit Eigenwerten $\lambda = \lambda_1, \dots, \lambda_n \in \mathbb{C}$, sodass $|\lambda| > |\lambda_i|$ für $i \geq 2$ gilt. Dann konvergiert die Folge

$$x_{k+1} := \frac{Ax_k}{|Ax_k|}$$

für „fast alle“ Startvektoren $x_0 \in \mathbb{C}^{n \times 1}$ gegen einen Eigenvektor zu λ . Ggf. ist $\lim_{k \rightarrow \infty} x_k^* Ax_k = \lambda$.

Beweis. Nach Voraussetzung hat die Jordan-Normalform von A die Form

$$J = \text{diag}(\lambda, J_{n_2}(\lambda_2), \dots, J_{n_s}(\lambda_s)),$$

sofern man $\lambda_2, \dots, \lambda_n$ geeignet anordnet. Sei $b_1, \dots, b_n \in \mathbb{C}^{n \times 1}$ eine entsprechende Basis mit $Ab_1 = \lambda b_1$. Für $S := (b_1, \dots, b_n) \in \text{GL}(n, \mathbb{C})$ gilt $S^{-1}AS = J$. Sei $x_0 = \alpha_1 b_1 + \dots + \alpha_n b_n$ zufällig gewählt mit $\alpha_1, \dots, \alpha_n \in \mathbb{C}$. In fast allen Fällen ist $\alpha_1 \neq 0$. Dies nehmen wir ab jetzt an. Dann gilt

$$A^k x_0 = A^k S(\alpha_1, \dots, \alpha_n)^t = S J^k (\alpha_1, \dots, \alpha_n)^t = \lambda^k (\alpha_1 b_1 + \beta_2 b_2 + \dots + \beta_n b_n).$$

Für $i \geq 2$ existieren nach Lemma 14.41 $2 \leq j \leq s$ und $0 \leq t \leq i - 2$ mit

$$\beta_i = \sum_{l=0}^t \binom{k}{l} \lambda_j^{-l} \alpha_{i-l} \left(\frac{\lambda_j}{\lambda}\right)^k \xrightarrow{k \rightarrow \infty} 0$$

wegen $|\lambda_j| < |\lambda|$. Mit $A^k x_0$ strebt auch x_k gegen ein Vielfaches von b_1 . Die zweite Behauptung folgt aus $|x_k| = 1$ für alle $k \geq 1$. \square

Bemerkung 17.71. Hat man einen Eigenvektor b_1 zum betragsmäßig größten Eigenwert λ von A (näherungsweise) bestimmt, so kann man b_1 zu einer Basis b_1, \dots, b_n ergänzen und aus diesen Spalten die Matrix S bilden. Es gilt $S^{-1}AS = \begin{pmatrix} \lambda & * \\ 0 & A_1 \end{pmatrix}$, wobei jeder Eigenwert von $A_1 \in \mathbb{R}^{(n-1) \times (n-1)}$ auch ein Eigenwert von A ist. Haben alle Eigenwerte von A paarweise verschiedene Beträge, so kann man Satz 17.70 auf A_1 anwenden und iterieren. Wir werden sehen, dass man dies mit nur einer Iteration bewerkstelligen kann.

Lemma 17.72. Sei (A_k) eine Folge invertierbarer Matrizen mit $\lim_{k \rightarrow \infty} A_k = 1_n$. Sei $A_k = Q_k R_k$ die QR-Zerlegung. Dann gilt $\lim_{k \rightarrow \infty} Q_k = 1_n = \lim_{k \rightarrow \infty} R_k$.

Beweis. Wegen $|Q_k| = \sqrt{\text{tr}(Q_k^* Q_k)} = \sqrt{\text{tr}(1_n)} = \sqrt{n}$ ist die Folge $(Q_k)_k$ beschränkt. Nach Bolzano-Weierstraß existiert eine konvergente Teilfolge $(Q_{k_i})_i$ mit $Q := \lim_{i \rightarrow \infty} Q_{k_i}$. Wegen

$$Q^* Q = \lim_{i \rightarrow \infty} Q_{k_i}^* Q_{k_i} = 1_n$$

ist Q unitär. Aus der Stetigkeit der Matrizenmultiplikation erhält man

$$\lim_{i \rightarrow \infty} R_{k_i} = \lim_{i \rightarrow \infty} Q_{k_i}^* (Q_{k_i} R_{k_i}) = Q^* \lim_{i \rightarrow \infty} A_{k_i} = Q^*.$$

Da die Menge der oberen Dreiecksmatrizen mit nicht-negativen Hauptdiagonaleinträgen abgeschlossen ist, muss auch Q^* zu dieser Menge gehören. Dies geht offenbar nur, falls $Q = 1_n$. Insbesondere konvergiert jede konvergente Teilfolge von $(Q_k)_k$ gegen den gleichen Grenzwert. Daher muss auch (Q_k) gegen 1_n konvergieren und folglich auch (R_k) . \square

Definition 17.73. Eine Folge (A_k) von Matrizen konvergiert *quasi* gegen eine Matrix A , falls eine Folge von unitären Diagonalmatrizen (D_k) mit $\lim_{k \rightarrow \infty} D_k^* A_k D_k = A$ existiert.

Bemerkung 17.74. In der Situation von Definition 17.73 gilt $D_k = \text{diag}(d_1, \dots, d_n)$ mit $|d_i| = 1$ für $i = 1, \dots, n$. Die Hauptdiagonalen von A_k und $D_k^* A_k A_k$ sind für alle k identisch. Daher konvergiert die Hauptdiagonale von A_k gegen einen Vektor in \mathbb{C}^n . Die Einträge von A_k außerhalb der Hauptdiagonalen können hingegen in jeder Iteration um einen Faktoren vom Betrag 1 „oszillieren“.

Satz 17.75 (FRANCIS-Algorithmus⁸). Sei $A_1 := A \in \mathbb{C}^{n \times n}$ invertierbar, sodass die Eigenwerte von A paarweise verschiedene Beträge haben. Für $k = 1, 2, \dots$ sei $A_k = Q_k R_k$ die QR-Zerlegung und $A_{k+1} := R_k Q_k$. Dann konvergiert $(A_k)_k$ quasi gegen eine obere Dreiecksmatrix mit den Eigenwerten auf A auf der Hauptdiagonale.

Beweis (WILKINSON). Sei $\mathcal{Q}_k := Q_1 \dots Q_k$ und $\mathcal{R}_k := R_k \dots R_1$. Aus $Q_k^* A_k Q_k = R_k Q_k = A_{k+1}$ folgt $Q_k^* A Q_k = A_{k+1}$. Wir zeigen $\mathcal{Q}_k \mathcal{R}_k = A^k$ für alle $k \geq 1$. Dies ist klar für $k = 1$. Für $k \geq 2$ gilt induktiv

$$\mathcal{Q}_k \mathcal{R}_k = \mathcal{Q}_{k-1} Q_k R_k \mathcal{R}_{k-1} = \mathcal{Q}_{k-1} A_k \mathcal{R}_{k-1} = \mathcal{Q}_{k-1} Q_{k-1}^* A Q_{k-1} \mathcal{R}_{k-1} = A A^{k-1} = A^k.$$

Für die Eigenwerte $\lambda_1, \dots, \lambda_n$ von A gilt $|\lambda_1| > \dots > |\lambda_n| > 0$ nach Voraussetzung. Nach Folgerung 8.12 existiert ein $S \in \text{GL}(n, \mathbb{C})$ mit $S^{-1} A S = \text{diag}(\lambda_1, \dots, \lambda_n) =: D$. Sei $S^{-1} = L_{S'} P R_{S'}$ die modifizierte LR-Zerlegung von S^{-1} aus Bemerkung 17.37. Sei $SP = Q_S R_S$ die QR-Zerlegung von SP . Für $L_{S'} = (x_{ij})$ gilt

$$D^k L_{S'} D^{-k} = ((\lambda_i / \lambda_j)^k x_{ij})_{ij} \xrightarrow{k \rightarrow \infty} 1_n$$

wegen $x_{ij} = \delta_{ij}$ für $i \leq j$ und $|\lambda_i| < |\lambda_j|$ für $i > j$. Damit strebt auch $M_k := R_S P^t D^k L_{S'} D^{-k} P R_S^{-1}$ gegen 1_n . Für die QR-Zerlegung $M_k = Q_{M_k} R_{M_k}$ gilt $\lim_{k \rightarrow \infty} Q_{M_k} = 1_n$ nach Lemma 17.72. Insgesamt ist

$$\begin{aligned} \mathcal{Q}_k \mathcal{R}_k &= A^k = S D^k S^{-1} = S P P^t D^k L_{S'} P R_{S'} = Q_S R_S P^t D^k L_{S'} D^{-k} D^k P R_{S'} \\ &= Q_S M_k R_S P^t D^k P R_{S'} = Q_S Q_{M_k} R_{M_k} P^t D^k P R_{S'}. \end{aligned}$$

Mit D^k ist auch $P^t D^k P$ eine Diagonalmatrix. Folglich ist $T_k := (t_{ij}) = R_{M_k} P^t D^k P R_{S'}$ eine obere Dreiecksmatrix. Für

$$U_k := \text{diag}(\overline{t_{11}}/|t_{11}|, \dots, \overline{t_{nn}}/|t_{nn}|) \in \text{U}(n, \mathbb{C})$$

ist $U_k T_k$ eine obere Dreiecksmatrix mit positiver Hauptdiagonale und $Q_S Q_{M_k} U_k^* \in \text{U}(n, \mathbb{C})$. Aus der Eindeutigkeit der QR-Zerlegung ergibt sich $\mathcal{Q}_k = Q_S Q_{M_k} U_k^*$. Damit gilt

$$\lim_{k \rightarrow \infty} U_k^* A_{k+1} U_k = \lim_{k \rightarrow \infty} U_k^* \mathcal{Q}_k^* A \mathcal{Q}_k U_k = \lim_{k \rightarrow \infty} Q_{M_k}^* Q_S^* S D S^{-1} Q_S Q_{M_k} = R_S P^t D P R_S^{-1},$$

d. h. (A_k) konvergiert quasi gegen die obere Dreiecksmatrix $R_S P^t D P R_S^{-1} \approx D$. Auf der Hauptdiagonale stehen die Eigenwerte von D bzw. A . \square

Bemerkung 17.76.

- Offenbar konvergiert (A_k) quasi gegen eine Schur-Zerlegung (Satz 13.23) von A . Sei A reell. Dann folgt aus der Voraussetzung, dass auch die Eigenwerte reell sind, denn $|\lambda| = |\overline{\lambda}|$ für $\lambda \in \mathbb{C}$. Ist diese Voraussetzung nicht erfüllt, so kann man erreichen, dass (A_k) gegen eine reelle Blockdiagonalmatrix mit 2×2 -Blöcken konvergiert (Aufgabe II.21).
- In seiner Reinform konvergiert Francis' Algorithmus nur langsam. In der Praxis beschleunigt man das Verfahren, indem man A zunächst in eine *Hessenberg-Matrix* (d. h. $a_{ij} = 0$ für $i > j + 1$) transformiert und in jeder Iteration A_k um ein geeignetes Vielfaches der Einheitsmatrix verschiebt. Dieser Algorithmus wird (neben der FFT und dem Simplex-Algorithmus) zu den wichtigsten Algorithmen des 20. Jahrhunderts gezählt.⁹

⁸auch QR-Verfahren genannt

⁹Siehe [Dongarra-Sullivan, *Top Ten Algorithms of the Century*, Comput. Sci. Eng. 2 (2000), 22–23]

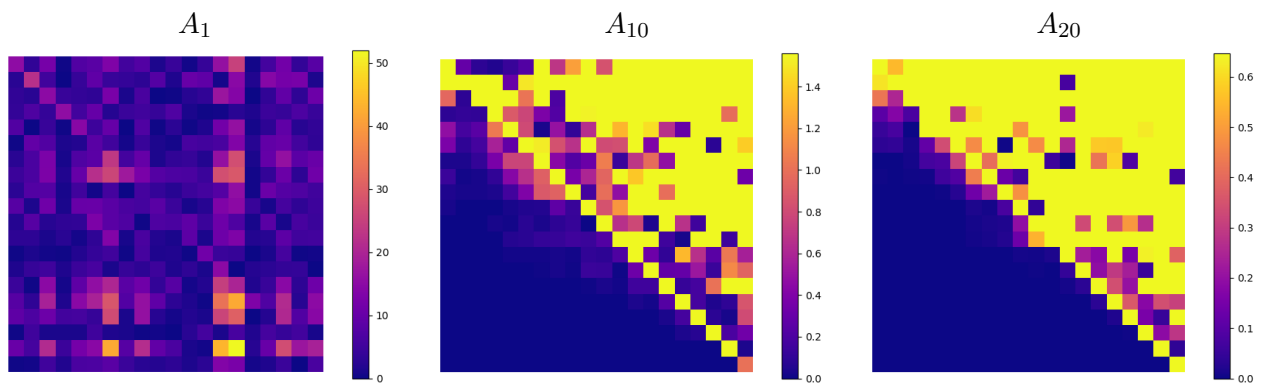
Beispiel 17.77.

(a) Für

$$A = \begin{pmatrix} 2 & -1 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} = Q_1 R_1$$

gilt $A_2 = R_1 Q_1 = \begin{pmatrix} 2 & -1 \\ 0 & 1 \end{pmatrix}$ und induktiv $A_k = \begin{pmatrix} 2 & (-1)^k \\ 0 & -1 \end{pmatrix}$. Dies zeigt das Phänomen der Quasi-Konvergenz.

(b) Sei $A = S^{-1} \text{diag}(1, \dots, 20) S \in \mathbb{R}^{20 \times 20}$ für eine zufällig gewählte Matrix $S \in \text{GL}(20, \mathbb{R})$. Die folgende Grafik illustriert Francis-Algorithmus für A . Die Farbskala repräsentiert die Beträge von Einträgen unterhalb der Hauptdiagonale (die Struktur von A wurde in Bemerkung 9.24 begründet).



Bemerkung 17.78. Für positiv definite Matrizen kann man die QR-Zerlegung in Francis' Algorithmus durch die Cholesky-Zerlegung ersetzen.

Satz 17.79 (CHOLESKY-Verfahren). Sei $A = A_1 \in \mathbb{C}^{n \times n}$ positiv definit. Für $k = 1, 2, \dots$ sei $A_k = R_k^* R_k$ die Cholesky-Zerlegung und $A_{k+1} := R_k R_k^*$. Dann konvergiert $(A_k)_k$ gegen eine Diagonalmatrix mit den Eigenwerten von A auf der Hauptdiagonale.

Beweis (SCHATZMAN). Wegen $A_{k+1} = R_k A_k R_k^{-1}$ haben alle A_k die gleichen Eigenwerte. Sei $A_k = (a_{ij}^{(k)})$ und $R_k = (r_{ij}^{(k)})$. Wegen

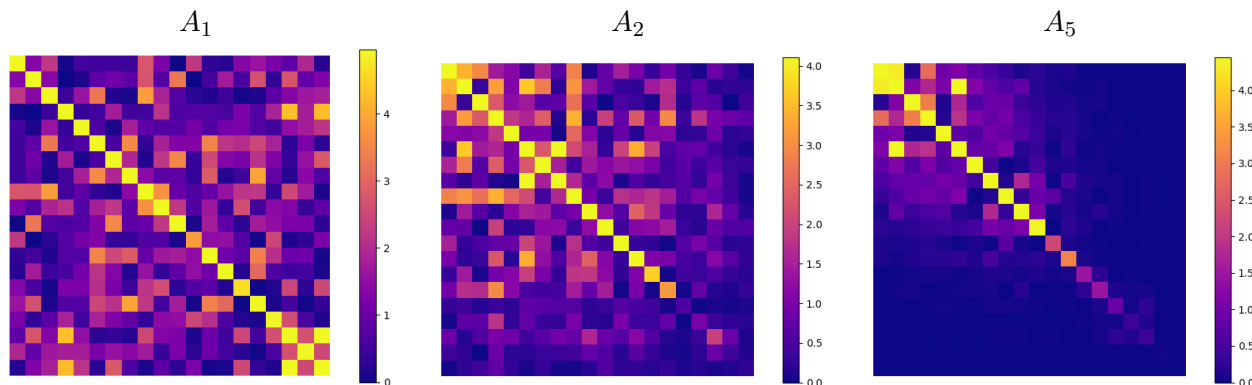
$$\delta_k(s) := \sum_{i=1}^s a_{ii}^{(k)} = \sum_{i=1}^s \sum_{j=1}^i |r_{ji}^{(k)}|^2 = \sum_{i=1}^s \sum_{j=i}^s |r_{ij}^{(k)}|^2 \leq \sum_{i=1}^s \sum_{j=i}^n |r_{ij}^{(k)}|^2 = \delta_{k+1}(s)$$

sind $(\delta_k(s))_k$ für $s = 0, \dots, n$ monoton steigende Folgen. Wegen $\delta_k(s) \leq \delta_k(n) = \text{tr}(A_k) = \text{tr}(A)$ sind die Folgen beschränkt. Daher existiert $\delta(s) := \lim_{k \rightarrow \infty} \delta_k(s)$ für $s = 0, \dots, n$. Es folgt $\lim_{k \rightarrow \infty} a_{ss}^{(k)} = \delta(s) - \delta(s-1)$ für $s = 1, \dots, n$. Da die Differenzen

$$\delta_{k+1}(s) - \delta_k(s) = \sum_{i=1}^s \sum_{j=s+1}^n |r_{ij}^{(k)}|^2$$

mit k gegen 0 streben, muss $\lim_{k \rightarrow \infty} r_{ij}^{(k)} = 0$ für alle $i \neq j$ gelten. Mit R_k muss auch A_k gegen eine Diagonalmatrix konvergieren. \square

Beispiel 17.80. Sei $A \in \mathbb{R}^{20 \times 20}$ eine zufällig gewählte positiv definite Matrix (konstruiert als $A = S^t S$). Die folgende Grafik illustriert den Fortschritt des Cholesky-Verfahrens mit A . Die Farbskala repräsentiert die Beträge aller Einträge außerhalb der Hauptdiagonale.



Offenbar arbeitet das Verfahren von rechts unten nach links oben.

Bemerkung 17.81. Die ungefähre Lage von Eigenwerten einer Matrix lässt sich ohne Rechenaufwand mit folgendem Satz einkreisen.

Satz 17.82 (GERSHGORIN). Für jeden Eigenwert $\lambda \in \mathbb{C}$ von $(a_{ij}) \in \mathbb{C}^{n \times n}$ existiert ein $i \in \{1, \dots, n\}$ mit $|\lambda - a_{ii}| \leq \sum_{j \neq i} |a_{ij}|$.

Beweis. Sei $x = (x_1, \dots, x_n) \in \mathbb{C}^n$ ein Eigenvektor zum Eigenwert λ . Sei

$$|x_i| = \max\{|x_j| : j = 1, \dots, n\} > 0.$$

Nach Normierung können wir $x_i = 1$ und $|x_j| \leq 1$ für $j \neq i$ annehmen. Nach der Dreiecksungleichung gilt dann

$$|\lambda - a_{ii}| = |(Ax)_i - a_{ii}x_i| = \left| \sum_{j=1}^n a_{ij}x_j - a_{ii}x_i \right| = \left| \sum_{j \neq i} a_{ij}x_j \right| \leq \sum_{j \neq i} |a_{ij}|. \quad \square$$

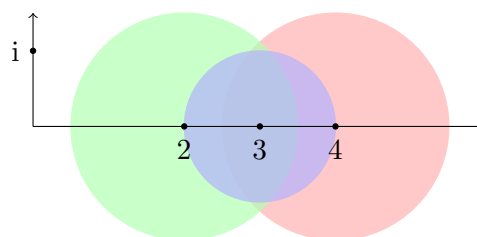
Folgerung 17.83. Jede Matrix $A \in \mathbb{C}^{n \times n}$ mit $|a_{ii}| > \sum_{j \neq i} |a_{ij}|$ für $i = 1, \dots, n$ ist invertierbar.¹⁰

Beweis. Nach Gershgorin ist 0 kein Eigenwert von A . □

Beispiel 17.84.

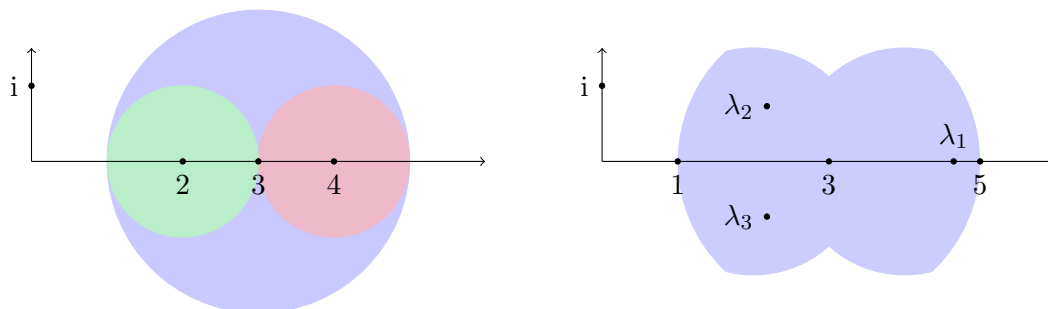
(a) Jeder Eigenwert liegt in einem der abgebildeten Kreise:

$$A = \begin{pmatrix} 3 & 1 & 0 \\ 0.5 & 4 & 1 \\ 1.5 & 0 & 2 \end{pmatrix}$$



¹⁰Diese Eigenschaft heißt *Diagonal-Dominanz*.

Da A und A^t die gleichen Eigenwerte besitzen, kann man auch die Spalten benutzen und anschließend beide Mengen schneiden:



(b) Wendet man Satz 17.82 auf die Begleitmatrix von $\alpha = X^n + a_1X^{n-1} + \dots + a_n \in \mathbb{C}[X]$ an, so erhält man

$$|x| \leq \max\{|a_n|, |a_{n-1}| + 1, \dots, |a_1| + 1\}$$

für jede Nullstelle $x \in \mathbb{C}$ von α (beachte $|x + a_1| \leq 1 \implies |x| \leq |a_1| + 1$).

17.7 Orthonormalisierung

Bemerkung 17.85. In unserer Formulierung des Gram-Schmidt-Verfahrens (11.10) haben wir zunächst eine Orthogonalbasis konstruiert und anschließend durch Normierung eine Orthonormalbasis erhalten. Dadurch können jedoch sehr große oder sehr kleine Faktoren auftreten, die zu numerischen Instabilitäten führen. Günstiger ist es also jeden gefundenen Vektor sofort zu normieren:

$$b'_s := v_s - \sum_{i=1}^{s-1} [v_s, b_i] b_i, \quad b_s := \frac{b'_s}{|b'_s|}$$

In der Summenbildung können sich zusätzlich Rundungsfehler akkumulieren. Stabiler ist daher ein iteriertes Vorgehen.

Satz 17.86 (Modifiziertes GRAM-SCHMIDT-Verfahren). Sei V ein unitärer Raum und $v_1, \dots, v_k \in V$ linear unabhängig. Für $s = 1, \dots, k$ definieren wir

$$b_{s,0} := v_s, \quad b_{s,i} := b_{s,i-1} - [b_{s,i-1}, b_i] b_i \quad (i = 1, \dots, s-1), \quad b_s := \frac{b_{s,s-1}}{|b_{s,s-1}|}.$$

Dann ist b_1, \dots, b_k eine Orthonormalbasis von $\langle v_1, \dots, v_k \rangle$.

Beweis. Es genügt zu zeigen, dass die Vektoren b_1, \dots, b_k mit den in Satz 11.10 gewonnenen Vektoren übereinstimmen. Dies ist klar für $b_1 = \frac{v_1}{|v_1|}$. Sei $s \geq 2$ und die Behauptung für b_1, \dots, b_{s-1} bereits gezeigt. Dann ist $b_{s,1} = v_s - [v_s, b_1] b_1$. Induktiv sei bereits

$$b_{s,i} = v_s - \sum_{j=1}^i [v_s, b_j] b_j$$

für ein $i < s-1$ gezeigt. Wegen $[b_j, b_{i+1}] = 0$ für $j \leq i$ ist

$$b_{s,i+1} = v_s - \sum_{j=1}^i [v_s, b_j] b_j - [v_s, b_{i+1}] b_{i+1} = v_s - \sum_{j=1}^{i+1} [v_s, b_j] b_j.$$

Dies zeigt, dass $b_{s,s-1}$, der in Satz 11.10 berechnete Vektor ist (vor Normierung). \square

Bemerkung 17.87. Bei der Berechnung der QR-Zerlegung einer invertierbaren Matrix $A \in \text{GL}(n, \mathbb{C})$ benötigt man neben den orthonormalisierten Spalten von A (als Spalten von Q) auch die Koeffizientenmatrix R . Dafür gibt es zwei gängige Methoden, die wir für reelle A vorstellen (siehe Aufgabe III.15 für den allgemeinen Fall):

- (a) Wir führen die Orthonormalisierung als Komposition von Spiegelungen durch. Sei dafür a_1 die erste Spalte von A und $b := a_1 - |a_1|e_1$. Die Spiegelung $Q_1 \in \text{O}(n, \mathbb{R})$ an der Hyperebene $\langle b \rangle^\perp$ lässt sich nach Aufgabe II.11 als HOUSEHOLDER-Transformation berechnen:

$$Q_1 = 1_n - \frac{2}{|b|^2} bb^t$$

(für $b = 0$ sei $Q_1 = 1_n$). Wegen

$$[a_1 + |a_1|e_1, b] = |a_1|^2 - |a_1|[a_1, e_1] + |a_1|[e_1, a_1] - |a_1|^2 = 0$$

ist $a_1 + |a_1|e_1 \in \langle b \rangle^\perp$ und

$$Q_1 a_1 = \frac{1}{2}(Q_1 b + Q_1(a_1 + |a_1|e_1)) = \frac{1}{2}(-b + a_1 + |a_1|e_1) = |a_1|e_1.$$

Es folgt $Q_1 A = \begin{pmatrix} |a_1| & * \\ 0 & * \end{pmatrix}$. Mit den Spalten $2, \dots, n$ geht man analog vor und erhält entsprechende Matrizen $Q_2, \dots, Q_n \in \text{O}(n, \mathbb{R})$ (tatsächlich reicht es die verkürzten Spalten $(a_{ii}, \dots, a_{ni})^t$ zu betrachten). Jetzt ist $R := Q_n \dots Q_1 A$ eine obere Dreiecksmatrix und $Q := Q_1 \dots Q_n \in \text{O}(n, \mathbb{R})$ mit $A = QR$, denn $Q_i^{-1} = Q_i$.

- (b) Anstelle von Spiegelungen kann man GIVENS-Rotationen benutzen (vgl. Bemerkung 11.43):

$$D_{st}(\varphi) := \begin{pmatrix} 1_{s-1} & & & & \\ & \cos \varphi & & -\sin \varphi & \\ & & 1_{t-s-1} & & \\ & \sin \varphi & & \cos \varphi & \\ & & & & 1_{n-t} \end{pmatrix} \in \text{O}(n, \mathbb{R}).$$

Wir gehen wie beim Gauß-Algorithmus vor. Angenommen die ersten $k-1$ Spalten von A sind bereits in oberer Dreiecksgestalt. Nach Zeilenvertauschung kann man $a_{kk} \neq 0$ annehmen (A invertierbar). Um den Eintrag a_{ik} für $i > k$ zu eliminieren, setzen wir $c := |(a_{kk}, a_{ik})| = \sqrt{|a_{kk}|^2 + |a_{ik}|^2}$ und

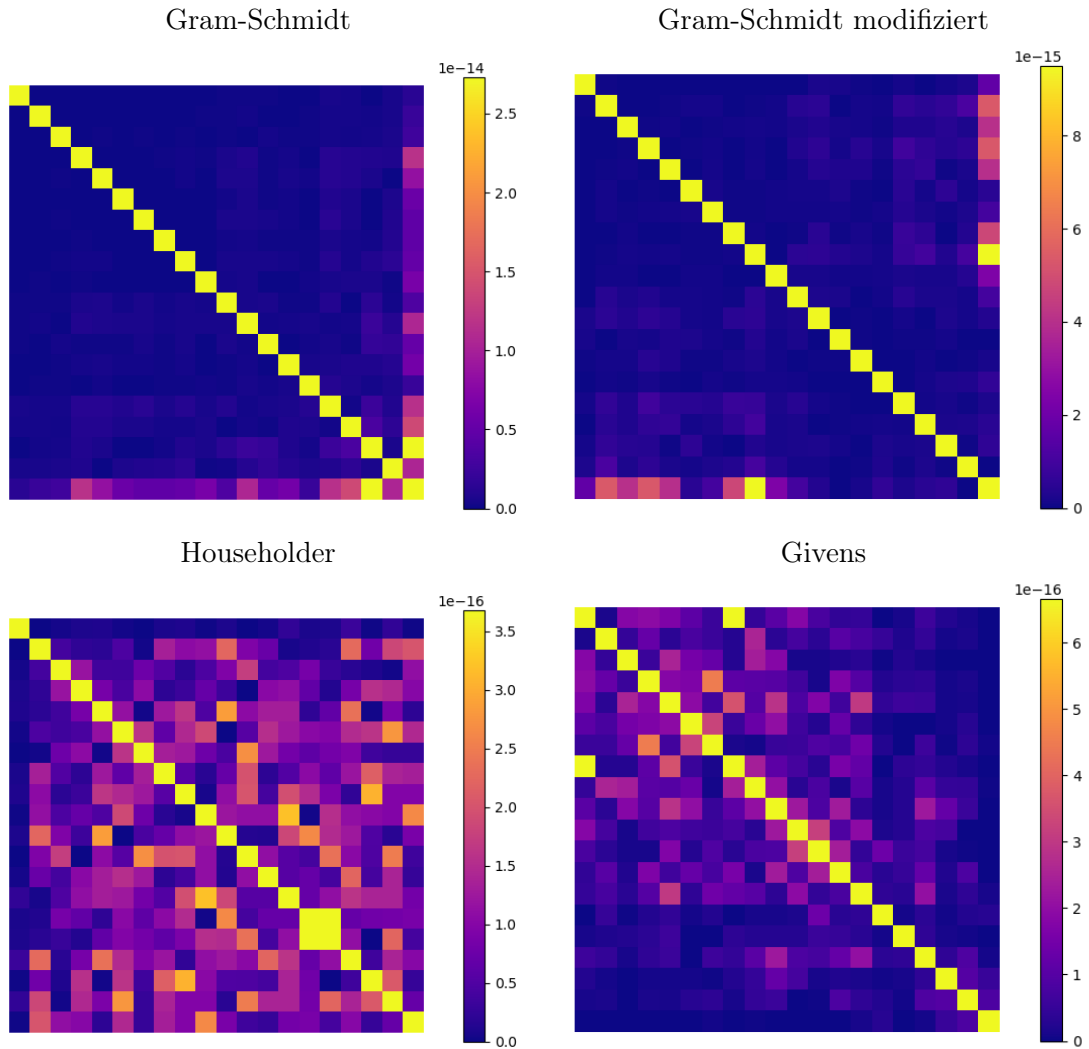
$$\cos \varphi := \frac{[(a_{kk}, a_{ik}), (1, 0)]}{c} = \frac{a_{kk}}{c}$$

Wegen $\sin^2 \varphi = 1 - \cos^2 \varphi = a_{ik}^2/c^2$ können wir $\sin \varphi = -a_{ik}/c$ annehmen (ersetze notfalls φ durch $-\varphi$). Nun ist

$$D_{ki}(\varphi) = \begin{pmatrix} 1_{k-1} & & & & \\ & a_{kk}/c & & a_{ik}/c & \\ & & 1_{i-k-1} & & \\ & -a_{ik}/c & & a_{kk}/c & \\ & & & & 1_{n-i} \end{pmatrix}.$$

Indem man, A durch $D_{ki}(\varphi)A$ ersetzt, wird $a_{ik} = 0$. Auf diese Weise erhält man eine Folge von Givens-Rotationen $Q_1, \dots, Q_l \in \text{O}(n, \mathbb{R})$, sodass $R := Q_l \dots Q_1 A$ eine obere Dreiecksmatrix ist. Für $Q := Q_l^\dagger \dots Q_1^\dagger \in \text{O}(n, \mathbb{R})$ gilt $A = QR$. Dies benötigt zwar deutlich mehr Matrizen als (a), aber die Q_i sind dünn-besetzt, sodass die Multiplikationen wenig kosten.

Beispiel 17.88. Wir berechnen die QR-Zerlegung $A = QR$ einer zufälligen (invertierbaren) Matrix $A \in \mathbb{R}^{20 \times 20}$ mit den vier vorgestellten Verfahren. Die Qualität der Ergebnisse zeigt sich durch einen Vergleich von $Q^t Q$ mit 1_{20} . Die Farbskala bezieht sich auf die Beträge der Einträge außerhalb der Hauptdiagonale (je kleiner, desto genauer):



18 Analytische Aspekte

18.1 Eigenwertabschätzungen

Bemerkung 18.1. Eine hermitesche Matrix $A \in \mathbb{C}^{n \times n}$ hat nach Folgerung 13.20 reelle Eigenwerte $\lambda_1 \geq \dots \geq \lambda_n$. Wir schreiben $\lambda_k(A) := \lambda_k$ für den k . größten Eigenwert. Der folgende *Min-Max-Satz* setzt diese Zahlen in Zusammenhang mit den sogenannten *RAYLEIGH-Quotienten* $\frac{vAv^*}{vv^*}$ und verallgemeinert Lemma 17.55.

Satz 18.2 (COURANT-FISCHER). Für jede hermitesche Matrix $A \in \mathbb{C}^{n \times n}$ gilt

$$\lambda_k(A) = \max_{\substack{V \leq \mathbb{C}^n \\ \dim V = k}} \min_{0 \neq v \in V} \frac{vAv^*}{vv^*} = \min_{\substack{V \leq \mathbb{C}^n \\ \dim V = n-k+1}} \max_{0 \neq v \in V} \frac{vAv^*}{vv^*}.$$

Beweis. Sei $\mu_k(A)$ der mittlere Teil der Formel. Nach dem Spektralsatz existiert ein $S \in U(n, \mathbb{C})$ mit $SAS^* = \text{diag}(\lambda_1, \dots, \lambda_n)$ und $\lambda_1 \geq \dots \geq \lambda_n$. Indem man $V \leq \mathbb{C}^n$ durch $\{vS : v \in V\}$ ersetzt, kann man $A = \text{diag}(\lambda_1, \dots, \lambda_n)$ annehmen. Außerdem kann man sich auf normierte $v \in V$ beschränken, indem man v durch $v/|v|$ ersetzt. Für $V := \langle e_1, \dots, e_k \rangle$ gilt

$$\mu_k(A) \geq \min_{\substack{v \in V \\ |v|=1}} vAv^* = \min_{\substack{v \in V \\ |v|=1}} \sum_{i=1}^k \lambda_k |v_i|^2 \geq \min_{|v|=1} \lambda_k |v|^2 = \lambda_k.$$

Sei umgekehrt $V \leq \mathbb{C}^n$ mit $\dim V = k$ beliebig. Nach der Dimensionsformel existiert ein normierter Vektor $w \in V \cap \langle e_k, \dots, e_n \rangle$. Es folgt

$$\min_{\substack{v \in V \\ |v|=1}} vAv^* \leq wAw^* = \sum_{i=k}^n \lambda_i |w_i|^2 \leq \lambda_k.$$

Da V beliebig ist, gilt $\mu_k(A) \leq \lambda_k = \lambda_k(A)$. Für die zweite Gleichheit betrachtet man analog $V := \langle e_k, \dots, e_n \rangle$ und $V \cap \langle e_1, \dots, e_k \rangle$. \square

Satz 18.3 (CAUCHYs Reißverschluss-Satz). Sei $A = (a_{ij}) \in \mathbb{C}^{n \times n}$ hermitesch und $P \in \mathbb{C}^{n \times k}$ mit orthonormalen Spalten ($k \leq n$). Dann gilt $\lambda_i(A) \geq \lambda_i(P^*AP) \geq \lambda_{n-k+i}(A)$ für $i = 1, \dots, k$.

Beweis. Mit A ist auch $B := P^*AP \in \mathbb{C}^{k \times k}$ hermitesch. Für $U \leq \mathbb{C}^k$ sei $P(U) := \{uP^* : u \in U\} \leq \mathbb{C}^n$. Da P vollen Rang hat, gilt $\dim U = \dim P(U)$. Außerdem ist $|uP^*| = |u|$ für alle $u \in U$. Nach Courant-Fischer folgt

$$\lambda_i(B) = \max_{\substack{U \leq \mathbb{C}^k \\ \dim U = i}} \min_{\substack{v \in P(U) \\ |v|=1}} vAv^* \leq \max_{\substack{V \leq \mathbb{C}^n \\ \dim V = i}} \min_{\substack{v \in V \\ |v|=1}} vAv^* = \lambda_i(A).$$

Sei b_j ein Eigenvektor von B zum Eigenwert $\lambda_j(B)$ und $U := \langle b_i, \dots, b_k \rangle$. Dann gilt

$$\lambda_{n-k+i}(A) = \min_{\substack{V \subseteq \mathbb{C}^n \\ \dim V = k-i+1}} \max_{\substack{v \in V \\ |v|=1}} vAv^* \leq \max_{\substack{v \in P(U) \\ |v|=1}} vAv^* = \max_{\substack{u \in U \\ |u|=1}} uBu^* = \lambda_i(B)$$

wie im Beweis von Satz 18.2. □

Beispiel 18.4. Für $I \subseteq \{1, \dots, n\}$ hat $P = (e_i : i \in I)$ orthonormale Spalten. Cauchys Reißverschluss-Satz gilt dann mit der Untermatrix $B := P^*AP = (a_{ij} : i, j \in I)$. Im Fall $|I| = n - 1$ erhält man

$$\lambda_1(A) \geq \lambda_1(B) \geq \lambda_2(A) \geq \lambda_2(B) \geq \dots \geq \lambda_{n-1}(B) \geq \lambda_n(A).$$

18.2 Der Spektralradius

Definition 18.5. Für $A \in \mathbb{C}^{n \times n}$ heißt

$$\rho(A) := \max\{|\lambda| : \lambda \in \mathbb{C} \text{ Eigenwert von } A\}$$

der *Spektralradius* von A .¹

Beispiel 18.6.

- (a) Für unitäre Matrizen A gilt $\rho(A) = 1$.
- (b) Genau dann ist $A \in \mathbb{C}^{n \times n}$ nilpotent, wenn $\rho(A) = 0$.
- (c) Aus der Jordan-Normalform folgt $\rho(A^t) = \rho(A)$ und $\rho(A^k) = \rho(A)^k$ für $k \in \mathbb{N}_0$.

Satz 18.7. Für alle $A \in \mathbb{C}^{n \times n}$ gilt:

- (a) $\|A\| \geq \rho(A)$ für jede submultiplikative Matrixnorm.
- (b) Für alle $\epsilon > 0$ existiert eine submultiplikative Matrixnorm mit $\|A\| \leq \rho(A) + \epsilon$.

Beweis.

- (a) Sei $\lambda \in \mathbb{C}$ ein Eigenwert von A mit Betrag $\rho(A)$ und v ein entsprechender Eigenvektor. Sei $B \in \mathbb{C}^{n \times n}$ mit erster Spalte v und sonst nur Nullen. Dann gilt

$$\rho(A)\|B\| = \|\lambda B\| = \|AB\| \leq \|A\|\|B\|$$

und $\rho(A) \leq \|A\|$.

- (b) Sei $J = S^{-1}AS$ die Jordan-Normalform von A mit $S \in \text{GL}(n, \mathbb{C})$. Sei $D := \text{diag}(\epsilon, \epsilon^2, \dots, \epsilon^n)$. Die Hauptdiagonalen von $N := DJD^{-1}$ und J sind identisch, während die Einträge unterhalb der Hauptdiagonale ϵ oder 0 sind. Für die Matrixnorm $\|B\| := \|DS^{-1}BSD^{-1}\|_1$ gilt nun $\|A\| := \|N\|_1 \leq \rho(A) + \epsilon$ nach Lemma 17.62. □

Satz 18.8 (Geometrische Reihe²). Für alle $A \in \mathbb{C}^{n \times n}$ sind die folgenden Aussagen äquivalent:

- (1) $\rho(A) < 1$.

¹Alle Eigenwerte liegen in der komplexen Ebene innerhalb des Kreises mit Mittelpunkt 0 und Radius $\rho(A)$.

²In diesem Kontext auch NEUMANN-*Reihe* genannt.

$$(2) \lim_{k \rightarrow \infty} A^k = 0_n.$$

$$(3) \sum_{k=0}^{\infty} A^k = (1_n - A)^{-1}.$$

Beweis.

(1) \Rightarrow (2): Sei $\epsilon > 0$ mit $q := \rho(A) + \epsilon < 1$. Nach Satz 18.7 existiert eine submultiplikative Matrixnorm mit $\|A\| \leq q$. Es folgt $\|A^k\| \leq \|A\|^k \leq q^k \rightarrow 0$ mit $k \rightarrow \infty$.

(2) \Rightarrow (1): Sei λ ein Eigenwert von A mit Betrag $\rho(A)$ mit Eigenvektor v . Wegen $\lambda^k v = A^k v \rightarrow 0$ ist $\rho(A) = |\lambda| < 1$.

(1) \Rightarrow (3): Wie zuvor existiert eine submultiplikative Matrixnorm mit $q := \|A\| < 1$. Sei $B_m := \sum_{k=0}^m A^k$. Nach der Dreiecksungleichung gilt

$$\|B_m\| \leq \sum_{k=0}^m q^k \leq \sum_{k=0}^{\infty} q^k = \frac{1}{1-q}.$$

Nach Bolzano-Weierstraß besitzt die beschränkte Folge $(B_m)_m$ eine konvergente Teilfolge $(B_{m_k})_k$. Für $B := \lim_{k \rightarrow \infty} B_{m_k}$ gilt

$$(1_n - A)B = \lim_{k \rightarrow \infty} (1_n - A) \sum_{i=0}^{m_k} A^i = \lim_{k \rightarrow \infty} (1_n - A^{m_k+1}) \stackrel{(2)}{=} 1_n,$$

d. h. $B = (1_n - A)^{-1}$. Da alle konvergente Teilfolgen von (B_m) den gleichen Grenzwert besitzt, muss auch die gesamte Folge gegen B konvergieren.

(3) \Rightarrow (2): Da die Partialsummen B_m konvergieren, müssen deren Differenzen $B_m - B_{m-1} = A^m$ eine Nullfolge bilden. \square

Folgerung 18.9. Sei $A \in \mathbb{C}^{n \times n}$ mit $\|A\| < 1$ für eine submultiplikative Matrixnorm. Dann gilt $\sum_{k=0}^{\infty} A^k = (1_n - A)^{-1}$.

Beweis. Aus Satz 18.7 folgt $\rho(A) < 1$. \square

Beispiel 18.10. Die Submultiplikativität in Folgerung 18.9 ist unentbehrlich: Für $A = (3/4)_{i,j=1}^2$ gilt $\|A\|_{\max} = 3/4 < 1$ mit der Matrixnorm aus Beispiel 17.61, aber $A^k = \left(\frac{3}{2}\right)^{k-1} A$.

Satz 18.11. Für jede Matrixnorm $\|\cdot\|$ und alle $A \in \mathbb{C}^{n \times n}$ gilt

$$\rho(A) = \lim_{k \rightarrow \infty} \sqrt[k]{\|A^k\|}.$$

Beweis. Für eine beliebige (nicht unbedingt submultiplikative) Matrixnorm $\|\cdot\|'$ existieren nach Lemma 17.51 Konstanten $\lambda, \mu > 0$ mit $\lambda\|A^k\| \leq \|A^k\|' \leq \mu\|A^k\|$ für alle k . Es folgt

$$\sqrt[k]{\lambda} \sqrt[k]{\|A^k\|} \leq \sqrt[k]{\|A^k\|'} \leq \sqrt[k]{\mu} \sqrt[k]{\|A^k\|}.$$

Wegen $\lim_{k \rightarrow \infty} \sqrt[k]{\lambda} = 1 = \lim_{k \rightarrow \infty} \sqrt[k]{\mu}$ genügt es die Aussage für eine bestimmte Matrixnorm zu beweisen. Sei

$$S^{-1}AS = \text{diag}(J_{n_1}(\lambda_1), \dots, J_{n_s}(\lambda_s))$$

die Jordan-Normalform von A mit $S \in \text{GL}(n, \mathbb{C})$. Wir definieren $\|B\| := \|S^{-1}BS\|_{\max}$ für $B \in \mathbb{C}^{n \times n}$ mit der Norm aus Beispiel 17.61. Jeder von 0 verschiedene Eintrag von A^k hat nach Lemma 14.41 die Form $\binom{k}{l} \lambda_i^{k-l}$ mit $1 \leq i \leq s$ und $0 \leq l \leq k$. Für jedes $\epsilon > 0$ gilt $k < (1 + \epsilon)^k$, wenn k groß genug ist. Es folgt

$$\lim_{k \rightarrow \infty} \sqrt[k]{\binom{k}{l} |\lambda_i|^{k-l}} = |\lambda_i| \lim_{k \rightarrow \infty} \sqrt[k]{k(k-1) \dots (k-l+1)} \lim_{k \rightarrow \infty} |\lambda_i|^{-l/k} = |\lambda_i|.$$

Dies zeigt $\lim_{k \rightarrow \infty} \sqrt[k]{\|A^k\|} = \rho(A)$. □

18.3 Die Exponentialfunktion einer Matrix

Definition 18.12. Für $A \in \mathbb{C}^{n \times n}$ sei

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

die *Exponentialfunktion* von A .

Bemerkung 18.13.

- (a) Wegen $\left| \frac{A^k}{k!} \right| \leq \frac{|A|^k}{k!} \rightarrow 0$ ist $\exp(A)$ nach Satz 18.8 wohldefiniert.
- (b) Durch $\exp(\lambda 1_n) = e^\lambda 1_n$ für $\lambda \in \mathbb{C}$ setzt \exp die gewöhnliche Exponentialfunktion auf \mathbb{C} fort.
- (c) Sei

$$J := S^{-1}AS = \text{diag}(J_{n_1}(\lambda_1), \dots, J_{n_s}(\lambda_s))$$

die Jordan-Normalform von A mit $S \in \text{GL}(n, \mathbb{C})$. Wegen der Stetigkeit der Matrizenmultiplikation gilt

$$S^{-1} \exp(A) S = \exp(J) = \text{diag}(\exp(J_{n_1}(\lambda_1)), \dots, \exp(J_{n_s}(\lambda_s))).$$

Nach Lemma 14.41 haben die von 0 verschiedenen Einträge von $\exp(J_{n_i}(\lambda_i))$ die Form

$$\sum_{k=l}^{\infty} \frac{\binom{k}{l} \lambda_i^{k-l}}{k!} = \sum_{k=l}^{\infty} \frac{\lambda_i^{k-l}}{l!(k-l)!} = \frac{e^{\lambda_i}}{l!}.$$

Also gilt

$$L_m(\lambda) := \exp(J_m(\lambda)) = e^\lambda \begin{pmatrix} 1 & & & \\ \frac{1}{1!} & \ddots & & \\ \vdots & \ddots & \ddots & \\ \frac{1}{(m-1)!} & \cdots & \frac{1}{1!} & 1 \end{pmatrix} \in \mathbb{C}^{m \times m}$$

und

$$\exp(A) = S \text{diag}(L_{n_1}(\lambda_1), \dots, L_{n_s}(\lambda_s)) S^{-1}.$$

Beispiel 18.14.

- (a) Für die Permutationsmatrix $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ gilt

$$\exp(P) = \sum_{k=0}^{\infty} \frac{1}{(2k)!} 1_n + \sum_{k=0}^{\infty} \frac{1}{(2k+1)!} P = \begin{pmatrix} \cosh(1) & \sinh(1) \\ \sinh(1) & \cosh(1) \end{pmatrix} = \begin{pmatrix} 1.54\dots & 1.17\dots \\ 1.17\dots & 1.54\dots \end{pmatrix}$$

mit den hyperbolischen Winkelfunktionen.

(b) Für die Matrix

$$A = \begin{pmatrix} 5 & 0 & 1 \\ -5 - i & -i & -1 \\ -9 & 0 & -1 \end{pmatrix}$$

aus Beispiel 14.34 gilt

$$\exp(A) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & -3 & 0 \end{pmatrix} \begin{pmatrix} e^2 & 0 & 0 \\ e^2 & e^2 & 0 \\ 0 & 0 & e^{-i} \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & -3 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 4e^2 & 0 & e^2 \\ -4e^2 + e^{-i} & e^{-i} & -e^2 \\ -9e^2 & 0 & -2e^2 \end{pmatrix}.$$

(c) Nach Bemerkung 18.13 gilt $\exp(A) = 1_n$ genau dann, wenn A diagonalisierbar ist und die Eigenwerte ganzzahlige Vielfache von $2\pi i$ sind.

Satz 18.15 (JACOBI). Für $A \in \mathbb{C}^{n \times n}$ gilt $\det(\exp(A)) = e^{\operatorname{tr}(A)}$.

Beweis. O. B. d. A. sei A in Jordan-Normalform mit Eigenwerten $\lambda_1, \dots, \lambda_n \in \mathbb{C}$. Dann ist $\exp(A)$ eine untere Dreiecksmatrix mit Diagonale $(e^{\lambda_1}, \dots, e^{\lambda_n})$. Dies zeigt $\det(\exp(A)) = e^{\lambda_1 + \dots + \lambda_n} = e^{\operatorname{tr}(A)}$. \square

Beispiel 18.16. Ist A reell, so ist $\det(\exp(A)) > 0$.

Satz 18.17 (Funktionalgleichung). Für vertauschbare Matrizen $A, B \in \mathbb{C}^{n \times n}$ gilt

$$\exp(A + B) = \exp(A) \exp(B).$$

Insbesondere ist $\exp(A)$ invertierbar mit $\exp(A)^{-1} = \exp(-A)$.

Beweis. Wir benutzen die Jordan-Chevalley-Zerlegung $A = D_A + N_A$ und $B = D_B + N_B$ aus Folgerung 16.20.³ Da alle beteiligten Matrizen Polynome in A bzw. B sind, sind sie paarweise vertauschbar. Nach Lemma 14.11 lassen sich D_A und D_B simultan diagonalisieren. O. B. d. A. sei also $D_A = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ und $D_B = \operatorname{diag}(\mu_1, \dots, \mu_n)$. Sei $D := D_A + D_B$ und $N := N_A + N_B$. Aus der Funktionalgleichung der gewöhnlichen Exponentialfunktion erhält man

$$\exp(D) = \operatorname{diag}(e^{\lambda_1 + \mu_1}, \dots, e^{\lambda_n + \mu_n}) = \operatorname{diag}(e^{\lambda_1}, \dots, e^{\lambda_n}) \operatorname{diag}(e^{\mu_1}, \dots, e^{\mu_n}) = \exp(D_A) \exp(D_B).$$

Da N_A und N_B nilpotent sind, gilt

$$\exp(N) = \sum_{k=0}^{\infty} \frac{1}{k!} (N_A + N_B)^k = \sum_{k=0}^{2n} \sum_{l=0}^k \frac{1}{k!} \binom{k}{l} N_A^l N_B^{k-l} = \sum_{k=0}^{2n} \sum_{l=0}^k \frac{N_A^l}{l!} \frac{N_B^{k-l}}{(k-l)!} = \exp(N_A) \exp(N_B).$$

Analog ist

$$\exp(D + N) = \sum_{k=0}^{\infty} \sum_{l=0}^k \frac{D^{k-l}}{(k-l)!} \frac{N^l}{l!} = \sum_{l=0}^n \frac{N^l}{l!} \sum_{k=l}^{\infty} \frac{D^{k-l}}{(k-l)!} = \exp(N) \exp(D).$$

Die entsprechenden Gleichungen gelten auch für $D_A + N_A$ und $D_B + N_B$. Insgesamt folgt

$$\begin{aligned} \exp(A + B) &= \exp(D + N) = \exp(D) \exp(N) = \exp(D_A) \exp(D_B) \exp(N_A) \exp(N_B) \\ &= \exp(D_A) \exp(N_A) \exp(D_B) \exp(N_B) = \exp(A) \exp(B). \end{aligned}$$

Die zweite Behauptung folgt aus $\exp(A) \exp(-A) = \exp(A - A) = \exp(0_n) = 1_n$. \square

³Das lässt sich vermeiden, indem man stattdessen das Cauchy-Produkt für absolut konvergente Matrix-Reihen benutzt.

Bemerkung 18.18.

- (a) Auf die Vertauschbarkeit von A und B in der Funktionalgleichung kann nicht verzichtet werden, denn sonst wäre $\exp(A)\exp(B) = \exp(A+B) = \exp(B+A) = \exp(B)\exp(A)$ für alle A, B im Widerspruch zu Satz 18.19. Tatsächlich gilt die GOLDEN-THOMPSON-Ungleichung

$$|\exp(A+B)| \leq |\exp(A)\exp(B)|$$

für alle hermiteschen Matrizen A, B mit Gleichheit genau dann, wenn $AB = BA$ (ohne Beweis).

- (b) Bekanntlich lässt sich jedes $z \in \mathbb{C}^\times$ eindeutig in der Form $|z|e^{i\varphi}$ mit $-\pi < \varphi \leq \pi$ schreiben. Die Abbildung $\log: \mathbb{C}^\times \rightarrow \mathbb{C}$, $z \mapsto \ln(|z|) + i\varphi$ nennt man den *Hauptzweigs des natürlichen Logarithmus*.

Satz 18.19. Für jedes $A \in \text{GL}(n, \mathbb{C})$ existiert genau ein $B \in \mathbb{C}^{n \times n}$ mit $\exp(B) = A$ und $-\pi < \text{Im}(\lambda) \leq \pi$ für jeden Eigenwert λ von B .

Beweis. Für die Existenz können wir nach Bemerkung 18.13 $A = J_n(\mu)$ für ein $\mu \neq 0$ annehmen. Sei $\lambda = \log(\mu)$ und $B := J_n(\lambda)$. Nach Bemerkung 18.13 ist μ ein Eigenwert von $\exp(B) = L_n(\lambda)$ mit algebraischer Vielfachheit n und geometrischer Vielfachheit 1. Aus Satz 14.28 folgt $\exp(B) \approx A$. Daher existiert ein $S \in \text{GL}(n, \mathbb{C})$ mit $\exp(S^{-1}BS) = S^{-1}\exp(B)S = A$.

Sei nun $A \in \text{GL}(n, \mathbb{C})$ beliebig und B, C mit den angegebenen Eigenschaften. Nach Basiswechsel können wir $B = \text{diag}(J_{n_1}(\lambda_1), \dots, J_{n_s}(\lambda_s))$ annehmen. Nach Bemerkung 18.13 ist

$$A = \exp(B) = \text{diag}(L_{n_1}(\lambda_1), \dots, L_{n_s}(\lambda_s)).$$

Durch die spezielle Wahl der Eigenwerte mit dem Hauptzweig des Logarithmus muss C die gleiche Jordan-Normalform wie B haben. Sei also $S \in \text{GL}(n, \mathbb{C})$ mit $S^{-1}BS = C$. Aus

$$A = \exp(C) = S^{-1}\exp(B)S = S^{-1}AS$$

folgt $S \in C(A)$. Wir müssen $S \in C(B)$ zeigen. Offenbar ist $\chi_A = (X - e^{\lambda_1})^{n_1} \dots (X - e^{\lambda_s})^{n_s}$. O. B. d. A. seien $\lambda_1, \dots, \lambda_s$ so sortiert, dass gleiche λ_i hintereinanderstehen. Nach Aufgabe II.30 ist $C = \text{diag}(C_1, \dots, C_t)$, wobei jedes C_k zu einem Block $\text{diag}(L_{n_i}(\lambda_i), \dots, L_{n_j}(\lambda_j))$ mit $\lambda_i = \dots = \lambda_j$ gehört. Wir können daher $\lambda_1 = \dots = \lambda_s$ mit $n_1 \geq \dots \geq n_s$ annehmen. Nun ist A eine Linearkombination der Potenzen von $J := \text{diag}(J_{n_1}(0), \dots, J_{n_s}(0))$. Insbesondere ist $C(J) \subseteq C(A)$. Nach Frobenius (Satz 15.32) ist

$$\dim C(A) = \sum_{i=1}^s (2i - 1)n_i = \dim C(J).$$

Es folgt $C(J) = C(A)$. Andererseits ist auch B eine Linearkombination von J . Dies zeigt $S \in C(A) = C(J) \subseteq C(B)$, wie gewünscht. \square

Beispiel 18.20. Die Eigenwerte einer unitären Matrix U lassen sich nach Folgerung 13.20 in der Form $e^{i\varphi_1}, \dots, e^{i\varphi_n}$ mit $\varphi_1, \dots, \varphi_n \in \mathbb{R}$ schreiben. Nach dem Spektralsatz existiert ein $S \in \text{U}(n, \mathbb{C})$ mit $U = S \text{diag}(e^{i\varphi_1}, \dots, e^{i\varphi_n})S^*$. Für die hermitesche Matrix $H := S \text{diag}(\varphi_1, \dots, \varphi_n)S^*$ gilt

$$U = \exp(iH).$$

Dieser Zusammenhang spielt in der Quantenmechanik eine Rolle. Verzichtet man auf die spezielle Wahl der Eigenwerte in Satz 18.19, so kann man $\varphi_1, \dots, \varphi_n > 0$ wählen. Dann ist H sogar positiv definit.

Bemerkung 18.21.

(a) In der Analysis kann man $\log(1 - x)$ für $x \in \mathbb{R}$ mit $|x| < 1$ durch die *Mercator-Reihe*

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

berechnen. Für $\rho(A) < 1$ konvergiert nach Satz 18.11 auch die entsprechende Reihe mit A anstelle von x . Ist A nilpotent (d. h. $\rho(A) = 0$), so bricht die Reihe nach $n - 1$ Summanden ab. Ggf. ist $B := \log(1 - A)$ die eindeutig bestimmte Matrix aus Satz 18.19 (Aufgabe III.18).

(b) Ein System von gewöhnlichen homogenen *Differentialgleichungen* erster Ordnung hat die Form $f'(t) = Af(t)$, wobei $f: \mathbb{R} \rightarrow \mathbb{R}^n$ eine differenzierbare Funktion mit Ableitung f' ist und $A \in \mathbb{R}^{n \times n}$. Man kann zeigen, dass bei gegebenem A alle Lösungen f die Form $f(t) = \exp(At)c$ mit $c \in \mathbb{R}^n$ haben (beachte: $t \in \mathbb{R}$). Das ist eine Verallgemeinerung der bekannten Ableitungsregel $(e^{at})' = ae^{at}$.

18.4 Nicht-negative Matrizen

Bemerkung 18.22. In der Wahrscheinlichkeitsrechnung und anderen praxisnahen Gebieten treten reelle Matrizen mit lauter nicht-negativen Einträgen auf. Wir zeigen, dass die Eigenwerte und Eigenvektoren solcher Matrizen eine besondere Struktur aufweisen. Dies ist die theoretische Grundlage wichtiger Anwendungen wie des Google-Suchalgorithmus.

Definition 18.23. Man nennt $A = (a_{ij}) \in \mathbb{R}^{n \times m}$

- *positiv* (bzw. *nicht-negativ*), falls $a_{ij} > 0$ (bzw. $a_{ij} \geq 0$) für alle i, j gilt.
- *zerlegbar*, falls $n = m$ und $\emptyset \neq I \subsetneq \{1, \dots, n\}$ mit $a_{ij} = 0$ für alle $i \in I$ und $j \notin I$ existiert.
- *unzerlegbar*, falls $n = m$ und A nicht zerlegbar ist.

Wir schreiben $A < B$ (bzw. $A \leq B$), falls $B - A$ positiv (bzw. nicht-negativ) ist. Außerdem sei $A_+ := (|a_{ij}|)_{ij}$.

Bemerkung 18.24. Man sieht leicht, dass \leq eine Ordnungsrelation auf $\mathbb{R}^{n \times m}$ definiert. Wir benutzen die Notation $v \geq w$ und v_+ auch für Vektoren v, w (aufgefasst als $n \times 1$ - oder $1 \times n$ -Matrizen). Im Gegensatz zu \mathbb{R} ist \leq auf $\mathbb{R}^{n \times m}$ nicht total. Zum Beispiel ist weder $(1, -1) \leq 0$ noch $(1, -1) \geq 0$.

Beispiel 18.25.

(a) In der Stochastik untersucht man die langfristige Entwicklung zufälliger Ereignisse mittels *Markov-Ketten*. Im einfachsten Fall besteht eine Markov-Kette aus Zuständen Z_1, \dots, Z_n . Die Wahrscheinlichkeit, dass ein Übergang von Z_i zu Z_j stattfindet sei w_{ij} . Man nennt $W = (w_{ij})$ die *Übergangsmatrix* der Markov-Kette. Startet der Prozess in Z_i , so beschreibt der Vektor $e_i W^k$ die Zustandswahrscheinlichkeiten nach k Zeiteinheiten. Man interessiert sich daher für $\lim_{k \rightarrow \infty} W^k$. Offenbar ist W nicht-negativ und jede Zeilensumme ist 1. Matrizen mit dieser Eigenschaft nennt man (Zeilen-) *stochastisch*. Offenbar ist $v = (1, \dots, 1)^t$ ein Eigenvektor von W zum Eigenwert 1. Nach Aufgabe III.20 ist auch W^k für $k \geq 0$ stochastisch. Aus Lemma 17.62 und Satz 18.11 folgt

$$\rho(W) = \lim_{k \rightarrow \infty} \sqrt[k]{\|W^k\|_{\infty}} = 1.$$

- (b) Offenbar ist jede positive Matrix unzerlegbar.
- (c) Eine Permutationsmatrix $P_\sigma = (\delta_{i\sigma(j)})$ ist genau dann unzerlegbar, wenn $\sigma \in S_n$ ein n -Zyklus ist. Denn ist $I \subseteq \{1, \dots, n\}$ die Menge der Ziffern eines Zyklus von σ , so gilt $\delta_{i\sigma(j)} = 0$ für alle $i \in I$ und $j \notin I$.
- (d) Für jede reduzible Matrix $A \in \mathbb{R}^{n \times m}$ existiert eine Permutationsmatrix P und $1 \leq k < n$ mit $PAP^t = \begin{pmatrix} A_1 & A_2 \\ 0_{k \times (n-k)} & A_3 \end{pmatrix}$.
- (e) Mit A ist auch A^t unzerlegbar.

Satz 18.26 (PERRON). Für jede nicht-negative Matrix $A \in \mathbb{R}^{n \times n}$ ist $\rho(A)$ ein Eigenwert mit einem nicht-negativen Eigenvektor.

Beweis.

Schritt 1: $\rho(A)$ ist ein Eigenwert von A .

Im Fall $\rho(A) = 0$ ist $\rho(A)$ ein Eigenwert. Sei also $\rho(A) > 0$. Indem man A durch $\rho(A)^{-1}A \geq 0$ ersetzt, kann man $\rho(A) = 1$ annehmen. Für $0 < t < 1$ und $m \in \mathbb{N}$ ist $\rho(tA) = t < 1$ und

$$(1 - tA)^{-1} \stackrel{18.8}{=} \sum_{k=0}^{\infty} (tA)^k \geq 1_n + tA + \dots + (tA)^m.$$

Ist 1 kein Eigenwert von A , so ist $1 - A$ invertierbar und es gilt

$$(1 - A)^{-1} = \left(\lim_{t \rightarrow 1^-} (1 - tA) \right)^{-1} = \lim_{t \rightarrow 1^-} (1 - tA)^{-1} \geq \lim_{t \rightarrow 1^-} (1_n + tA + \dots + (tA)^m) = 1_n + A + \dots + A^m$$

für alle $m \in \mathbb{N}$. Insbesondere ist $\lim_{m \rightarrow \infty} A^m = 0$ im Widerspruch zu Satz 18.8.

Schritt 2: Jede positive Matrix A besitzt einen positiven Eigenvektor zum Eigenwert $\rho(A)$.⁴

Wie oben können wir $\rho(A) = 1$ annehmen. Sei $v = (v_1, \dots, v_n)$ ein Eigenvektor zum Eigenwert 1 von A . Dann ist

$$v_+ = (Av)_+ \leq A_+ v_+ = Av_+,$$

also $w := (A - 1_n)v_+ \geq 0$. Im Fall $w = 0$ ist $v_+ = Av_+ > 0$ ein positiver Eigenvektor von A . Sei nun $w \neq 0$. Dann ist $Aw > 0$ wegen $A > 0$. Daher existiert ein $\epsilon > 0$ mit $Aw \geq \epsilon v_+$. Für $z := Av_+ > 0$ gilt

$$(A - 1_n)z = A(A - 1_n)v_+ = Aw \geq \epsilon z,$$

also $Az \geq (1 + \epsilon)z$. Für $B := (1 + \epsilon)^{-1}A$ folgt $Bz \geq z$ und $B^k z \geq z$ für alle $k \in \mathbb{N}$. Andererseits ist $\rho(B) = (1 + \epsilon)^{-1}\rho(A) < 1$ und $\lim_{k \rightarrow \infty} B^k = 0$ nach Satz 18.8. Widerspruch.

Schritt 3: Jede nicht-negative Matrix A besitzt einen nicht-negativen Eigenvektor zum Eigenwert $\rho(A)$.

Für $k \in \mathbb{N}$ sei $A_k := A + \frac{1}{k}J > 0$, wobei $J = (1)_{i,j=1}^n$. Sicher ist $A_1 > A_2 > \dots > A$ und $A_1^m > A_2^m > \dots > A^m$ für alle $m \in \mathbb{N}$. Aus Satz 18.11 (angewendet beispielsweise mit der euklidischen Norm) folgt $\rho(A_1) \geq \rho(A_2) \geq \dots \geq \rho(A)$. Insbesondere existiert

$$\mu := \lim_{k \rightarrow \infty} \rho(A_k) \geq \rho(A).$$

⁴Das hat Perron ursprünglich bewiesen.

Nach Schritt 2 existieren positive Eigenvektoren v_k von A_k zum Eigenwert $\rho(A_k)$. Nach Normierung gilt $|v_k| = 1$. Nach Bolzano-Weierstraß besitzt $(v_k)_k$ eine konvergente Teilfolge. O. B. d. A. sei also $v := \lim_{k \rightarrow \infty} v_k \geq 0$. Wegen $|v| = 1$ ist $v \neq 0$. Außerdem gilt

$$Av = \lim_{k \rightarrow \infty} A_k \lim_{k \rightarrow \infty} v_k = \lim_{k \rightarrow \infty} A_k v_k = \lim_{k \rightarrow \infty} \rho(A_k) v_k = \mu v.$$

Wegen $\mu \geq \rho(A)$ folgt $\mu = \rho(A)$. □

Satz 18.27 (PERRON-FROBENIUS). *Für jede nicht-negative unzerlegbare Matrix $A \in \mathbb{R}^{n \times n}$ gilt:*

- (a) *Die algebraische Vielfachheit von $\rho(A)$ als Eigenwert ist 1.*
- (b) *$E_{\rho(A)}(A) = \langle v \rangle$ mit $v > 0$.*
- (c) *Bis auf Skalarmultiplikation ist v der einzige nicht-negative Eigenvektor von A .*

Beweis.

(a,b) Nach Perron existiert ein Eigenvektor $v \geq 0$ zum Eigenwert $\rho(A)$. Sei $I := \{1 \leq i \leq n : v_i = 0\}$. Für $i \in I$ gilt

$$0 = \rho(A)v_i = (Av)_i = \sum_{j=1}^n a_{ij}v_j = \sum_{j \notin I} a_{ij}v_j.$$

Dies zeigt nur, falls $a_{ij} = 0$ für alle $i \in I$ und $j \notin I$. Da A unzerlegbar ist, muss $I = \emptyset$ gelten, d. h. $v > 0$.

Sei nun auch $Aw = \rho(A)w$ mit $w \in \mathbb{R}^{n \times 1}$. Dann existiert ein $\lambda \in \mathbb{R}$, sodass $w - \lambda v \geq 0$ an mindestens einer Koordinate verschwindet. Da wir bereits gesehen haben, dass jeder nicht-negative Eigenvektor zum Eigenwert $\rho(A)$ positiv ist, folgt $w = \lambda v$. Daher ist zumindest die geometrische Vielfachheit von $\rho(A)$ gleich 1. Die Jordan-Normalform von A besitzt also nur einen Block zu $\rho(A)$. Es genügt daher

$$\text{Ker}((A - \rho(A)1_n)^2) = \langle v \rangle$$

zu zeigen. Sei $w \in \text{Ker}((A - \rho(A)1_n)^2)$. Wegen $(A - \rho(A)1_n)w \in \langle v \rangle$ existiert ein $\lambda \in \mathbb{C}$ mit $Aw - \rho(A)w = \lambda v$. Da auch A^t unzerlegbar ist mit $\rho(A^t) = \rho(A)$, existiert $u > 0$ mit $A^t u = \rho(A)u$. Es folgt

$$\lambda[v, u] = [Aw, u] - \rho(A)[w, u] = w^t A^t u - \rho(A)[w, u] = \rho(A)([w, u] - [w, u]) = 0.$$

Wegen $[v, u] > 0$ ist $\lambda = 0$ und $w \in \langle v \rangle$ wie gewünscht.

(c) Sei $w \geq 0$ ein beliebiger Eigenvektor von A zum Eigenwert $\lambda \in \mathbb{C}$. Wie oben folgt $v > 0$ aus der Unzerlegbarkeit von A . Wie zuvor sei $u > 0$ mit $A^t u = \rho(A)u$. Wegen $[w, u] > 0$ und

$$\lambda[w, u] = [Aw, u] = w^t A^t u = \rho(A)[w, u]$$

gilt $\lambda = \rho(A)$. Also ist v bis auf Skalierung der einzige nicht-negative Eigenvektor von A . □

Bemerkung 18.28.

- (a) Nach Schur-Horn existiert für gegebene $\lambda_1, \dots, \lambda_n \geq 0$ stets eine positiv semidefinite reelle Matrix mit Eigenwerten $\lambda_1, \dots, \lambda_n$. Die Bestimmung aller möglichen Eigenwertmengen von (symmetrischen) nicht-negativen Matrizen ist hingegen ein offenes Problem.

(b) Der nächste Satz verallgemeinert die Abschätzung $\rho(A) \leq \|A\|_\infty$ aus Satz 18.7 (setze $x = (1, \dots, 1)$ und benutze Lemma 17.62).

Satz 18.29 (COLLATZ-WIELANDT). Sei $A = (a_{ij}) \in \mathbb{R}_{\geq 0}^{n \times n}$ unzerlegbar und $x > 0$. Dann gilt

$$\min \left\{ \sum_{j=1}^n a_{ij} x_j / x_i : i = 1, \dots, n \right\} \leq \rho(A) \leq \max \left\{ \sum_{j=1}^n a_{ij} x_j / x_i : i = 1, \dots, n \right\}.$$

Beweis. Sei $u > 0$ mit $A^t u = \rho(A)u$. Sei $y_i := \sum_{j=1}^n a_{ij} x_j$ und $z_i := y_i / x_i$. Dann gilt

$$\sum_{j=1}^n (z_j - \rho(A)) x_j u_j = \sum_{j=1}^n y_j u_j - \sum_{j=1}^n x_j \sum_{k=1}^n a_{kj} u_k = \sum_{j=1}^n y_j u_j - \sum_{k=1}^n u_k \sum_{j=1}^n a_{kj} x_j = 0.$$

Wegen $x, u > 0$ existieren $1 \leq s, t \leq n$ mit $z_s \leq \rho(A)$ und $z_t \geq \rho(A)$. Also ist $\min_i z_i \leq \rho(A) \leq \max_i z_i$. \square

Satz 18.30 (VON MISES). Sei $A \in \mathbb{R}^{n \times n}$ nicht-negativ und unzerlegbar. Dann konvergiert die Folge

$$x_{k+1} := \frac{Ax_k}{|Ax_k|} \quad (k = 1, 2, \dots)$$

für jeden positiven Startvektor $x_1 \in \mathbb{R}^{n \times 1}$ gegen einen positiven Eigenvektor zu $\rho(A)$. Insbesondere ist $\rho(A) = \lim_{k \rightarrow \infty} |Ax_k|$.

Beweis. Nach Perron-Frobenius sind die Voraussetzungen der Potenzmethode (Satz 17.70) erfüllt. Wir müssen jedoch prüfen, dass die Iteration für jeden positiven Startvektor x_1 konvergiert. Seien $v, u > 0$ mit $Av = \rho(A)v$ und $A^t u = \rho(A)u$. Wegen $[u, v] > 0$ ist $\mathbb{R}^n = \langle v \rangle \oplus u^\perp$. Für $x \in u^\perp$ gilt

$$[Ax, u] = x^t A^t u = \rho(A)[x, u] = 0.$$

Dies zeigt, dass u^\perp A -invariant ist. Ergänzt man v mit Vektoren aus u^\perp zu einer Basis von \mathbb{R}^n , so wird A in die Form $\begin{pmatrix} \rho(A) & 0 \\ 0 & * \end{pmatrix}$ überführt. Wegen $[x_1, u] > 0$ gilt $x_1 \notin u^\perp$. Der Beweis von Satz 17.70 zeigt nun, dass die Potenzmethode für x_1 konvergiert. \square

Lemma 18.31. Sei $A \geq 0$ unzerlegbar und λ ein Eigenwert von A mit Betrag $\rho(A)$. Dann existiert eine Diagonalmatrix $U \in U(n, \mathbb{C})$ mit $\rho(A)A = \lambda U A U^*$.

Beweis. O. B. d. A. sei $\rho(A) > 0$. Sei $w \in \mathbb{C}^{n \times 1}$ ein Eigenvektor von A zum Eigenwert λ . Für $z := w_+ \geq 0$ gilt

$$\rho(A)z = (\lambda w)_+ = (Aw)_+ \leq Az.$$

Sei $u > 0$ mit $A^t u = \rho(A)u$. Aus

$$0 < \rho(A)[z, u] \leq [Az, u] = z^t A^t u = \rho(A)[z, u]$$

folgt $Az = \rho(A)z$ und $z > 0$. Sei $U := \text{diag}(w_1/z_1, \dots, w_n/z_n) \in U(n, \mathbb{C})$. Für $B := (b_{ij}) = \rho(A)\lambda^{-1}U^*AU \in \mathbb{C}^{n \times n}$ gilt

$$Bz = \frac{\rho(A)}{\lambda} U^* A w = \rho(A) U^* w = \rho(A)z = Az.$$

Da sich die Einträge von B und A nur um Faktoren vom Betrag 1 unterscheiden, gilt $B_+ = A$. Für $i = 1, \dots, n$ ist

$$(Az)_i = (Bz)_i = \sum_{j=1}^n b_{ij} z_j \stackrel{\in \mathbb{R}}{=} \sum_{j=1}^n \operatorname{Re}(b_{ij}) z_j \leq \sum_{j=1}^n |\operatorname{Re}(b_{ij})| |z_j| \leq \sum_{j=1}^n |b_{ij}| |z_j| = (Az)_i.$$

Dies geht nur falls $B = B_+ = A$. □

Bemerkung 18.32. Man beachte, dass A und UAU^* in Lemma 18.31 die gleiche Hauptdiagonale haben. Ist also $a_{ii} \neq 0$ für ein $1 \leq i \leq n$, so ist $\rho(A)$ der einzige Eigenwert vom Betrag $\rho(A)$.

Folgerung 18.33. Für jede positive Matrix A ist $\rho(A)$ der einzige Eigenwert vom Betrag $\rho(A)$.

Beweis. Folgt aus Bemerkung 18.32. □

Beispiel 18.34. Wir betrachten eine Markov-Kette mit Übergangsmatrix

$$W = \frac{1}{10} \begin{pmatrix} 5 & 1 & 4 \\ 2 & 4 & 4 \\ 0 & 9 & 1 \end{pmatrix}.$$

Offenbar ist W irreduzibel. Nach Bemerkung 18.32 ist $\rho(A) = 1$ der einzige Eigenwert vom Betrag 1. Nach Satz 17.66 existiert $W_\infty := \lim_{k \rightarrow \infty} W^k$. Zur Berechnung von W_∞ müssen wir nach Folgerung 17.67 Eigenvektoren von W und W^t bestimmen. Da W stochastisch ist, ist $v := (1, 1, 1)^t$ ein Eigenvektor von W zum Eigenwert 1. Man rechnet nach, dass $w := \frac{1}{91}(18, 45, 28)^t$ ein Eigenvektor von W^t ist mit $[v, w] = 1$. Nun gilt

$$W_\infty = vw^t = \frac{1}{91} \begin{pmatrix} 18 & 45 & 28 \\ 18 & 45 & 28 \\ 18 & 45 & 28 \end{pmatrix}.$$

Auf lange Sicht befindet man sich also mit Wahrscheinlichkeit $\frac{18}{91} \approx 20\%$ bzw. $\frac{45}{91} \approx 49\%$, $\frac{28}{91} \approx 31\%$ in Zustand Z_1 bzw. Z_2, Z_3 .

Satz 18.35 (WIELANDT). Sei $A \in \mathbb{R}^{n \times n}$ nicht-negativ und unzerlegbar mit genau k Eigenwerten vom Betrag $\rho(A)$. Dann gilt:

- (a) Die Eigenwerte vom Betrag $\rho(A)$ sind $e^{2\pi i j/k} \rho(A)$ für $j = 1, \dots, k$, d. h. sie verteilen sich gleichmäßig auf einem Kreis in der komplexen Ebene.
- (b) Ist μ ein beliebiger Eigenwert von A , so ist auch $e^{2\pi i/k} \mu$ ein Eigenwert von A .

Beweis.

- (a) Seien $\lambda_1 \rho(A), \dots, \lambda_k \rho(A) \in \mathbb{C}$ die Eigenwerte von A vom Betrag $\rho(A)$, d. h. $|\lambda_1| = \dots = |\lambda_k| = 1$. Nach Lemma 18.31 existiert für jedes i ein $U_i \in \operatorname{GL}(n, \mathbb{C})$ mit $\lambda_i A = U_i^{-1} A U_i \approx A$ (unitär wird nicht gebraucht). Ein Vergleich der Eigenwerte zeigt $\{\lambda_i \lambda_j : j = 1, \dots, k\} = \{\lambda_1, \dots, \lambda_k\}$ für $i = 1, \dots, k$. Es folgt

$$\lambda_i^k \prod_{j=1}^k \lambda_j = \prod_{j=1}^k (\lambda_i \lambda_j) = \prod_{j=1}^k \lambda_j$$

und $\lambda_i^k = 1$ für $i = 1, \dots, k$. Daher sind $\lambda_1, \dots, \lambda_k$ genau die in Beispiel 11.28 und Definition 17.6 definierten k -ten Einheitswurzeln.

(b) O. B. d. A. sei $\lambda_1 = e^{2\pi i/k}$. Die Ähnlichkeit $\lambda_1 A \approx A$ zeigt, dass $\lambda_1 \mu$ ein Eigenwert von A ist. \square

Satz 18.36. Sei $A \in \mathbb{R}^{n \times n}$ nicht-negativ. Dann existiert für jeden Eigenwert $\lambda \in \mathbb{C}$ von A vom Betrag $\rho(A)$ ein $m \leq n$ mit $\lambda^m = \rho(A)^m$.

Beweis. Nach Beispiel 18.25 existiert eine Permutationsmatrix P , sodass

$$P^t A P = \begin{pmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_k \end{pmatrix}$$

mit quadratischen nicht-negativen unzerlegbaren Matrizen A_1, \dots, A_k . Jeder Eigenwert von A_i ist auch ein Eigenwert von A (man ergänze einen entsprechenden Eigenvektor mit Nullen). Da A genau n komplexe Eigenwerte hat, existiert für jeden Eigenwert λ von A ein i , sodass λ ein Eigenwert von A_i ist. Die Behauptung folgt mit Satz 18.35. \square

Beispiel 18.37. Sei P_σ eine Permutationsmatrix und $\sigma = \sigma_1 \dots \sigma_k$ die Zerlegung in paarweise disjunkte Zyklen (einschließlich 1-Zyklen). Sei l_i die Länge von σ_i . Nach Beispiel 15.19 besitzt P die Eigenwerte

$$\zeta_{l_1}, \zeta_{l_1}^2, \dots, \zeta_{l_1}^{l_1}, \zeta_{l_2}, \dots, \zeta_{l_2}^{l_2}, \dots, \zeta_{l_k}^{l_k}$$

mit $\zeta_m = e^{2\pi i/m}$ für $m \in \mathbb{N}$.

18.5 Der Page-Rang

Bemerkung 18.38. Die Suchmaschine Google weist zur Sortierung von Ergebnisse jeder Webseite w ein Gewicht zu, das misst, wie viele andere Webseiten auf w verweisen. Wir schreiben $w_i \rightarrow w_j$, falls die Webseite w_i einen Hyperlink nach w_j enthält. Sei l_i die Anzahl der (indizierten) Hyperlinks auf w_i . Man „definiert“ den *PAGE-Rang*⁵ von w_i durch

$$p_i := \frac{1-d}{n} + d \sum_{\substack{1 \leq j \leq n \\ w_j \rightarrow w_i}} \frac{p_j}{l_j}, \quad (18.1)$$

wobei n die Anzahl aller indizierten Seiten und $0 < d < 1$ ein konstanter Dämpfungsfaktor ist.⁶ Der Page-Rang von w_i ist also „groß“, wenn viele Seiten w_j (mit hohem p_j und niedrigen l_j) auf w_i verweisen. Durch $d < 1$ wird verhindert, dass isolierte Webseiten (auf die nicht verwiesen wird) Page-Rang 0 erhalten. Die p_i lassen sich durch (18.1) nicht direkt berechnen. Es ist nicht einmal klar, ob sie durch (18.1) eindeutig bestimmt sind.

Definition 18.39. Mit den obigen Bezeichnungen nennt man $G = (g_{ij})$ mit

$$g_{ij} = \begin{cases} \frac{d}{l_i} + \frac{1-d}{n} & \text{falls } w_i \rightarrow w_j, \\ \frac{1-d}{n} & \text{sonst} \end{cases}$$

die *Google-Matrix*.

⁵Benannt nach einem der Google-Gründer Larry Page (nicht nach „web page“).

⁶Google nutzt $d = 0.85$. Der genaue Wert von n ist unbekannt, aber dürfte größer als 10^{10} sein.

Satz 18.40. Die Google-Matrix ist stochastisch. Außerdem ist (p_1, \dots, p_n) der einzige Eigenvektor von G^t zum Eigenwert 1 mit $p_1 + \dots + p_n = 1$. Insbesondere sind p_1, \dots, p_n eindeutig bestimmt.

Beweis. Wegen $0 < d < 1$ ist $G > 0$. Die i -te Zeilensumme von G ist

$$l_i \left(\frac{d}{l_i} + \frac{1-d}{n} \right) + (n - l_i) \frac{1-d}{n} = 1,$$

d. h. G ist stochastisch. Nach Definition von p_i gilt

$$\sum_{i=1}^n p_i = 1 - d + d \sum_{i=1}^n \sum_{\substack{1 \leq j \leq n \\ w_j \rightarrow w_i}} \frac{p_j}{l_j} = 1 - d + d \sum_{j=1}^n p_j \sum_{\substack{1 \leq i \leq n \\ w_j \rightarrow w_i}} \frac{1}{l_j} = 1 - d + d \sum_{j=1}^n p_j.$$

Es folgt $\sum_{i=1}^n p_i = 1$. Für $v := (p_1, \dots, p_n)$ ist

$$(G^t v)_i = \sum_{j=1}^n g_{ji} p_j = \frac{1-d}{n} \sum_{i=j}^n p_j + \sum_{\substack{1 \leq j \leq n \\ w_j \rightarrow w_i}} \frac{d p_j}{l_j} = p_i,$$

d. h. v ist ein Eigenvektor von G^t zum Eigenwert $\rho(G^t) = \rho(G) = 1$. Nach Perron-Frobenius ist v durch $\sum p_i = 1$ eindeutig bestimmt. \square

Bemerkung 18.41.

- (a) In der Praxis wird p_i mit dem von-Mises-Satz (näherungsweise) bestimmt. Man kann zeigen, dass $|\lambda| \leq d$ für alle Eigenwerte $\lambda \neq 1$ von G gilt. Damit konvergiert die Potenzmethode mindestens mit dem Faktor $\frac{1}{d}$ (siehe Beweis von Satz 17.70).
- (b) Man kann G auch als Übergangsmatrix einer Markov-Kette betrachten, bei der ein Surfer auf w_i mit Wahrscheinlichkeit d auf einen der Hyperlinks auf w_i klickt und mit Wahrscheinlichkeit $1 - d$ eine zufällige andere Webseite aufruft. Hat man die p_i bestimmt, so kann man mit Folgerung 17.67 die langfristigen Aufenthaltswahrscheinlichkeiten des Surfers ermitteln (siehe Beispiel 18.34).

19 Lineare Optimierung

19.1 Lineare Programme

Bemerkung 19.1.

- (a) In der linearen Optimierung maximiert (oder minimiert) man eine *Zielfunktion* (zum Beispiel Gewinn oder Kosten eines Unternehmens) unter Nebenbedingungen (zum Beispiel Nachfrage oder Kapazität).
- (b) Wir werden der Einheitlichkeit halber nur Spaltenvektoren benutzen. Abweichend von früheren Kapiteln sei \mathbb{R}^n der Raum der reellen *Spaltenvektoren*.

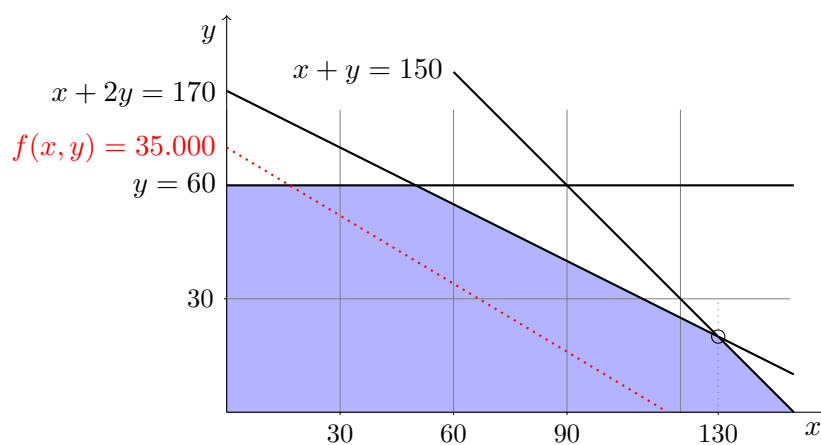
Beispiel 19.2. Ein Unternehmen stellt mit drei Maschinen A, B, C zwei Produkte P und Q her. Die Maschinen haben im Produktionszeitraum eine maximale Laufzeit von 170 Stunden (A), 150 Stunden (B) bzw. 180 Stunden (C). Die Herstellung von P und Q benötigt folgende Ressourcen:

Produkt	Laufzeit A	Laufzeit B	Laufzeit C	Erlös
P	1h	1h	0	300€
Q	2h	1h	3h	500€

In welchem Verhältnis sollte man Einheiten von P und Q erzeugen, um den Gesamterlös zu maximieren? Seien x und y die Anzahl der zu produzierenden Produkte P und Q . Die Zielfunktion $f: \mathbb{N}_0^2 \rightarrow \mathbb{R}$, $(x, y) \mapsto 300x + 500y$ soll unter den Nebenbedingungen

$$x, y \geq 0, \quad x + 2y \leq 170, \quad x + y \leq 150, \quad 3y \leq 180$$

maximiert werden. Diese Ungleichungen beschreiben die blau gekennzeichnete Fläche in \mathbb{R}^2 :



Die Parameter für einen Erlös von $f(x, y) = 35.000$ sind mit der rot gepunkteten Linie gekennzeichnet. Da die Steigung $-3/5$ dieser Geraden nicht von Erlös abhängt, sieht man leicht, dass f für $(x, y) = (130, 20)$ maximal ist. Der Erlös beträgt dann 49.000 € . Wir haben Glück, dass die Lösung (x, y) hier eindeutig und ganzzahlig ist. Für Probleme mit deutlich mehr Parametern ist dieser grafische Ansatz ungeeignet.

Bemerkung 19.3.

- (i) Eine „lineare“ Zielfunktion hat die Form $f: \mathbb{R}^m \rightarrow \mathbb{R}, x \mapsto c^t x + a$ für ein $c \in \mathbb{R}^m$ und $a \in \mathbb{R}$. Da Konstanten keinen Einfluss auf die Position von Extrema haben, kann man $a = 0$ annehmen. Wegen $\max f = -\min(-f)$ kann man sich auf die Suche nach Maxima beschränken.
- (ii) Lineare Nebenbedingungen an $x \in \mathbb{R}^m$ können in der Form $c^t x \leq b, c^t x = b$ oder $c^t x \geq b$ mit $c \in \mathbb{R}^m$ und $b \in \mathbb{R}$ auftreten. Wegen $c^t x \geq b \iff (-c)^t x \leq -b$ kommt man ohne die dritte Variante aus. Durch Hinzunahme einer sogenannten *Schlupfvariable* y kann man $c^t x \leq b$ durch $c^t x + y = b$ und $y \geq 0$ ersetzen. In der Zielfunktion kommen Schlupfvariablen nicht vor. Generell kann man $x_i \geq 0$ für $i = 1, \dots, m$ annehmen, indem man x_i durch zwei Variablen $x_i^+, x_i^- \geq 0$ mit $x = x_+ - x_-$ ersetzt. Alle Nebenbedingungen lassen sich nun in Matrixform $Ax = b$ mit $b \in \mathbb{R}^n$ und $x \geq 0$ zusammenfassen. Mit dem Gauß-Algorithmus kann man A in Zeilenstufenform überführen und Nullzeilen streichen. Insbesondere kann man annehmen, dass A vollen Rang hat. Im Fall $n \geq m$ kann $Ax = b$ nach Bemerkung 6.7 höchstens eine Lösung haben. Die Suche nach einem Maximum der Zielfunktion ist in diesem Fall uninteressant. Wir werden daher $n < m$ annehmen. Indem man Zeilen mit -1 multipliziert, erreicht man $b \geq 0$.
- (iii) Alternativ lassen sich Nebenbedingungen der Form $c^t x = b$ aufspalten in $c^t x \leq b$ und $-c^t x \leq -b$. Man erreicht damit $Ax \leq b$ ohne Hinzunahme von Schlupfvariablen und ohne Einschränkungen an x . Darauf kommen wir in Definition 19.14 zurück.

Definition 19.4. Ein *lineares Programm* (in Standardform) $L = (A, b, c)$ besteht aus $A \in \mathbb{R}^{n \times m}$ mit Rang $n < m, b \in \mathbb{R}^n$ mit $b \geq 0$ und $c \in \mathbb{R}^m$. Ein nicht-negativer Vektor $x \in \mathbb{R}^m$ heißt *zulässig*, wenn x die Nebenbedingung $Ax = b$ erfüllt. Gesucht sind zulässige x_{\max} , die die Zielfunktion $f: \mathbb{R}^m \rightarrow \mathbb{R}, x \mapsto c^t x$ maximieren, d. h. $f(x_{\max}) \geq f(x)$ für alle zulässigen $x \in \mathbb{R}^m$. Außerdem nennt man L

- *lösbar*, falls mindestens ein x_{\max} existiert.
- *unzulässig*, falls es keine zulässigen x gibt.
- *unbeschränkt*, falls f auf der Menge der zulässigen x beliebig groß wird.

Bemerkung 19.5. Sei L ein lineares Programm mit Zielfunktion f . Angenommen f ist auf der Menge M der zulässigen x beschränkt. Nach Bolzano-Weierstraß existiert eine Folge zulässiger Punkte $(x_i)_i$ mit $s := \lim_{i \rightarrow \infty} f(x_i) = \sup_{x \in M} f(x)$. Da M abgeschlossen ist, existiert $x_{\max} := \lim_{i \rightarrow \infty} x_i$. Da f als lineare Funktion stetig ist, gilt $f(x_{\max}) = s = \max_{x \in M} f(x)$, d. h. L ist lösbar. Dies zeigt, dass jedes lineare Programm entweder lösbar oder unzulässig oder unbeschränkt ist.

Beispiel 19.6. Das Problem aus Beispiel 19.2 lässt sich mit drei Schlupfvariablen als lineares Programm schreiben:

$$A = \begin{pmatrix} 1 & 2 & 1 & . & . \\ 1 & 1 & . & 1 & . \\ . & 3 & . & . & 1 \end{pmatrix}, \quad b = (170, 150, 180)^t, \quad c = (300, 500, 0, 0, 0)^t.$$

19.2 Konvexe Mengen

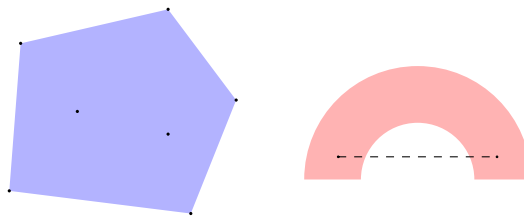
Definition 19.7.

- Eine Teilmenge $\Delta \subseteq \mathbb{R}^n$ heißt *konvex*, falls $\lambda x + (1 - \lambda)y \in \Delta$ für alle $x, y \in \Delta$ und $0 \leq \lambda \leq 1$ gilt. Dies bedeutet anschaulich, dass zu je zwei Punkten in Δ auch deren Verbindungsstrecke in Δ liegt.
- Für $v_1, \dots, v_k \in \mathbb{R}^n$ und $\lambda_1, \dots, \lambda_k \in \mathbb{R}_{>0}$ mit $\lambda_1 + \dots + \lambda_k = 1$ nennt man $\lambda_1 v_1 + \dots + \lambda_k v_k$ eine *Konvexkombination* von v_1, \dots, v_k .
- Die Menge aller Konvexkombinationen von Elementen aus Δ nennt man die *konvexe Hülle* von Δ und schreibt dafür $\text{con}(\Delta)$ (bzw. $\text{con}(x_1, \dots, x_k)$ falls $\Delta = \{x_1, \dots, x_k\}$). Offenbar ist $\text{con}(\Delta)$ die „kleinste“ konvexe Teilmenge, die Δ enthält, d. h.

$$\text{con}(\Delta) = \bigcap_{\substack{\Delta \subseteq \Gamma \subseteq \mathbb{R}^n \\ \Gamma \text{ konvex}}} \Gamma.$$

Beispiel 19.8.

- (a) Konvexe Mengen in \mathbb{R} sind (offene oder (halb)abgeschlossene) Intervalle. In \mathbb{R}^2 haben konvexe Mengen weder Löcher noch nach innen gerichtete „Knicks“. Hier die konvexe Hülle von sieben Punkten (blau) und eine nicht konvexe Menge (rot):



- (b) Für jede Norm ist $\{x \in \mathbb{R}^n : \|x\| \leq 1\}$ nach der Dreiecksungleichung konvex. Für $\|\cdot\|_2$ erhält man die Einheitskugel, für $\|\cdot\|_\infty$ den Würfel Kantenlänge 2 und für $\|\cdot\|_1$ den sogenannten *Standard-Simplex* $\text{con}(\pm e_1, \dots, \pm e_n)$ (für $n = 3$ ist dies ein Oktaeder).

Satz 19.9 (CARATHÉODORY). Sei $\Delta \subseteq \mathbb{R}^n$ und $x \in \text{con}(\Delta)$. Dann ist x eine Konvexkombination von höchstens $n + 1$ Elementen in Δ .

Beweis. Sei $x = \lambda_1 v_1 + \dots + \lambda_k v_k$ eine Konvexkombination mit $v_1, \dots, v_k \in \Delta$ und k minimal. Angenommen es gilt $k > n + 1$. Wir bilden die Matrix $A \in \mathbb{R}^{(n+1) \times k}$ mit den Spalten v_1, \dots, v_k und der zusätzlichen Zeile $(1, \dots, 1)$. Wegen $\text{rk}(A) \leq n + 1 < k$ existiert ein $y \in \mathbb{R}^k \setminus \{0\}$ mit $Ay = 0$. Damit gilt $y_1 v_1 + \dots + y_k v_k = 0$ und $y_1 + \dots + y_k = 0$. Sei $1 \leq s \leq k$ mit $\frac{y_s}{\lambda_s} \geq \frac{y_i}{\lambda_i}$ für $i = 1, \dots, k$. Für $\lambda'_i := \lambda_i \left(1 - \frac{y_i \lambda_s}{\lambda_i y_s}\right) \geq 0$ gilt

$$\sum_{i \neq s} \lambda'_i v_i = \sum_{i=1}^k \lambda_i v_i - \frac{\lambda_s}{y_s} \sum_{i=1}^k y_i v_i = x,$$

$$\sum_{i \neq s} \lambda'_i = \sum_{i=1}^k \lambda_i - \frac{\lambda_s}{y_s} \sum_{i=1}^k y_i = 1$$

im Widerspruch zur Wahl von k . □

Beispiel 19.10. Im Gegensatz zu Unterräumen lässt sich $\text{con}(\Delta) \subseteq \mathbb{R}^n$ nicht immer durch endlich (oder abzählbar) viele Elemente in Δ „aufspannen“ (als Konvexkombination). Betrachten wir dazu den konvexen Einheitskreis $\Delta := \{x \in \mathbb{R}^2 : |x| \leq 1\}$. Angenommen es existieren abzählbar viele $v_1, v_2, \dots \in \Delta$, sodass jedes Element in Δ eine Konvexkombination der v_i ist. Bekanntlich gibt es überabzählbar viele $x \in \Delta$ mit $|x| = 1$ (parametrisiert durch $(\cos(\varphi), \sin(\varphi))$ mit $\varphi \in \mathbb{R}$). Wir können $x \neq \pm v_i$ für $i \in \mathbb{N}$ annehmen. Dann gilt o. B. d. A. $x = \lambda_1 v_1 + \dots + \lambda_k v_k$ mit $k \geq 2$, $\lambda_i > 0$ und $\lambda_1 + \dots + \lambda_k = 1$. Nach der Dreiecksungleichung und der Cauchy-Schwarz-Ungleichung gilt

$$1 = [x, x] \leq \sum_{i=1}^k \lambda_i |[x, v_i]| \leq \sum_{i=1}^k \lambda_i |x| |v_i| \leq \sum_{i=1}^k \lambda_i = 1$$

nur, wenn x zu jedem v_i linear abhängig ist und $|v_1| = \dots = |v_k| = 1$. Dann wäre aber $x = \pm v_1$. Also benötigt man überabzählbar viele Elemente aus Δ , um jedes Element als Konvexkombination darzustellen.

Lemma 19.11 (FARKAS). *Für alle $U \leq \mathbb{R}^n$ gilt genau einer der folgenden Aussagen:*

- (1) *Es existiert ein $u \in U$ mit $u \geq 0$ und $u_1 > 0$.*
- (2) *Es existiert ein $v \in U^\perp$ mit $v \geq 0$ und $v_1 > 0$.*

Beweis. Gelten (1) und (2) gleichzeitig mit u bzw. v , so wäre $0 = [u, v] = u_1 v_1 + \dots + u_n v_n \geq u_1 v_1 > 0$. Wir argumentieren durch Induktion nach n . Für $n = 1$ ist $U = \mathbb{R}$ oder $U^\perp = \mathbb{R}$. Sei also $n \geq 2$ und

$$\begin{aligned} \tilde{U} &:= \{x \in \mathbb{R}^{n-1} : (x, 0) \in U\} \leq \mathbb{R}^{n-1}, \\ \hat{U} &:= \{x \in \mathbb{R}^{n-1} : \exists y \in \mathbb{R} : (x, y) \in U\} \leq \mathbb{R}^{n-1}. \end{aligned}$$

Aus Dimensionsgründen existiert ein $u_0 \in U$ mit $U = (\tilde{U} \times \{0\}) + \langle u_0 \rangle$. Aus Aufgabe II.6 folgt

$$U^\perp = (\tilde{U}^\perp \times \mathbb{R}) \cap u_0^\perp. \quad (19.1)$$

Nach Induktion gibt es drei Alternativen:

- (a) Es existiert ein $\tilde{u} \in \tilde{U}$ mit $\tilde{u} \geq 0$ und $\tilde{u}_1 > 0$. Dann gilt (1) mit $u := (\tilde{u}, 0) \in U$.
- (b) Es existiert ein $\hat{v} \in \hat{U}^\perp$ mit $\hat{v} \geq 0$ und $\hat{v}_1 > 0$. Für $v := (\hat{v}, 0)$ und alle $u = (\hat{u}, u_n) \in U$ gilt $[u, v] = [\hat{u}, \hat{v}] = 0$, d. h. (2) gilt für v .
- (c) Es existieren $\tilde{v} \in \tilde{U}^\perp$ und $\hat{u} \in \hat{U}$ mit $\tilde{v}, \hat{u} \geq 0$ und $\tilde{v}_1, \hat{u}_1 > 0$. Sei $y \in \mathbb{R}$ mit $u = (\hat{u}, y) \in U$. Im Fall $y \geq 0$ gilt (1) für u . Sei daher $y < 0$. Nach (19.1) existiert ein $z \in \mathbb{R}$ mit $v = (\tilde{v}, z) \in u_0^\perp \subseteq U^\perp$. Wegen $\hat{u}, \tilde{v} \geq 0$ ist

$$0 = [u, v] = [\hat{u}, \tilde{v}] + yz \geq yz$$

und $z \geq 0$. Also gilt (2) für v . □

Beispiel 19.12. In \mathbb{R}^2 besagt Farkas' Lemma, dass zu jeder Geraden g durch den Ursprung entweder g oder die dazu senkrechte Gerade g^\perp durch den ersten Quadranten ($x, y \geq 0$) verläuft.

Folgerung 19.13. *Für $A \in \mathbb{R}^{n \times m}$ und $b \in \mathbb{R}^n$ gilt genau eine der folgenden Aussagen:*

- (1) *Es existiert ein $x \in \mathbb{R}^m$ mit $x \geq 0$ und $Ax = b$.*
- (2) *Es existiert ein $y \in \mathbb{R}^n$ mit $b^t y < 0$ und $A^t y \geq 0$.*

Beweis. Sind x und y Lösungen von (1) bzw. (2), so wäre $0 \leq (A^t y)^t x = y^t (Ax) = y^t b < 0$. Es kann also höchstens eine der Aussagen gelten. Wir fügen $-b$ links als neue Spalte an A und erhalten $S := (-b|A) \in \mathbb{R}^{n \times (m+1)}$. Sei $U := \text{Ker}(S) \leq \mathbb{R}^{m+1}$. Angenommen es existiert ein $\tilde{x} = (x_1, x)^t \in U$ mit $x \geq 0$ und $x_1 > 0$. Nach Skalierung können wir $x_1 = 1$ annehmen. Dann ist x^t eine Lösung von (1).

Nach Farkas' Lemma können wir nun voraussetzen, dass ein $\tilde{z} = (z_1, z)^t \in U^\perp$ mit $z \geq 0$ und $z_1 > 0$ existiert. Nach Beispiel 11.16 ist $U^\perp = \{S^t y : y \in \mathbb{R}^n\}$. Sei also $y \in \mathbb{R}^n$ mit $S^t y = \tilde{z}$. Dann gilt $b^t y = -z_1 < 0$ und $A^t y = z \geq 0$. Somit gilt (2) für y . \square

Definition 19.14. Sei $L = (A, b, c)$ ein lineares Programm. Das zu L duale lineare Programm L^* minimiert die Zielfunktion $f^* : \mathbb{R}^n \rightarrow \mathbb{R}$, $y \mapsto b^t y$ unter der Nebenbedingung $A^t y \geq c$. Analog zu Definition 19.4 definiert man Lösbarkeit, Zulässigkeit und Beschränktheit von L^* .

Beispiel 19.15. Das duale Programm zu $L = (A, b, c)$ aus Beispiel 19.6 ist

$$\min_{y \in \mathbb{R}^3} (170, 150, 180)y \quad \begin{pmatrix} 1 & 1 & . \\ 2 & 1 & 3 \\ 1 & . & . \\ . & 1 & . \\ . & . & 1 \end{pmatrix} y \geq \begin{pmatrix} 300 \\ 500 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Wegen $y, c \geq 0$ ist L^* beschränkt. Offenbar ist y nur dann eine minimale Lösung, wenn in den ersten beiden Zeilen Gleichheit gilt, d. h. $y_1 + y_2 = 300$ und $2y_1 + y_2 + 3y_3 = 500$. Dann ist $y = (y_1, 300 - y_1, (200 - y_1)/3)$ und

$$f^*(y) = c^t y = 170y_1 + 150(300 - y_1) + 60(200 - y_1) = 57.000 - 40y_1.$$

Das Minimum $f^*(y) = 49.000$ erhält man für $y_1 = 200$ wegen $y_3 \geq 0$. Dies ist genau der Maximalwert von L . Der nächste Satz zeigt, dass dies kein Zufall ist.

Satz 19.16 (Dualitätssatz). *Für jedes lineare Programm $L = (c, A, b)$ gilt:*

- (a) L ist genau dann lösbar, wenn L^* lösbar ist. Für Lösungen x_{\max} und y_{\min} gilt $f(x_{\max}) = f^*(y_{\min})$.
- (b) Ist L (bzw. L^*) unbeschränkt, so ist L^* (bzw. L) unzulässig ist.

Beweis.

- (b) Ist $x \in \mathbb{R}^m$ zulässig für L und $y \in \mathbb{R}^n$ zulässig für L^* , so gilt

$$f(x) = c^t x \leq y^t Ax = y^t b = f^*(y). \quad (19.2)$$

Ist L (bzw. L^*) unbeschränkt, so muss also L^* (bzw. L) unzulässig sein.

- (a) Sei $x_{\max} \in \mathbb{R}^m$ eine Lösung von L . Sei

$$A_1 := \begin{pmatrix} 0 & A \\ -1 & c^t \end{pmatrix} \in \mathbb{R}^{(n+1) \times (m+1)}$$

und $U := \text{Ker}(A_1)$. Existiert ein $\begin{pmatrix} t \\ w \end{pmatrix} \in \text{Ker}(A_1)$ mit $t > 0$ und $w \in \mathbb{R}_{\geq 0}^m$, so wäre $x_{\max} + w \geq 0$, $A(x_{\max} + w) = b$ und

$$f(x_{\max} + w) = f(x_{\max}) + t > f(x_{\max}).$$

Nach Farkas' Lemma und Beispiel 11.16 existiert ein $y = A_1^t \begin{pmatrix} u \\ s \end{pmatrix} \geq 0$ mit $u \in \mathbb{R}^n$ und $y_1 > 0$. Nun ist $s < 0$ und $A^t u \geq -sc$. Für $v := -\frac{1}{s}u$ gilt schließlich $A^t v \geq c$, d. h. v ist zulässig für L^* .

Jetzt betrachten wir

$$A_2 = \begin{pmatrix} c & -A^t & A^t & 1_m & 0 & 0 \\ -b & 0 & 0 & 0 & A & 0 \\ 0 & b^t & -b^t & 0 & -c^t & 1 \end{pmatrix} \in \mathbb{R}^{(n+m+1) \times 2(n+m+1)}.$$

Angenommen es existieren $x \in \mathbb{R}^m$, $u \in \mathbb{R}^n$ und $s \in \mathbb{R}$ mit $(x^t, u^t, s)A_2 \geq 0$ und positiver ersten Koordinate. Dann ist

$$c^t x > u^t b, \quad Ax \leq sb, \quad Ax \geq sb, \quad x, s \geq 0, \quad A^t u \geq sc.$$

Sei dabei $s > 0$. Durch Übergang zu $\frac{1}{s}(x^t, u^t, s)$ kann man $s = 1$ annehmen. Dann ist x für L zulässig und u für L^* , aber (19.2) ist verletzt. Also gilt $s = 0$ und $Ax = 0$. Für alle $\lambda, \mu > 0$ ist $x_{\max} + \lambda x$ zulässig für L und $v + \mu u$ zulässig für L^* . Indem man λ oder μ (falls $c^t x = 0$) genügend groß wählt, wird (19.2) wieder ungültig. Daher kann (x^t, u^t, s) nicht existieren.

Mit Farkas' Lemma findet man einen nicht-negativen Vektor $(1, u_+, u_-, z, x, s) \in \text{Ker}(A_2)$. Das bedeutet

$$Ax = b, \quad A^t(u_+ - u_-) = c + z \geq c, \quad b^t(u_+ - u_-) = c^t x - s \leq c^t x.$$

Also ist x zulässig für L und $u := u_+ - u_-$ zulässig für L^* . Wegen $f(x) = c^t x \geq b^t u = f^*(u)$ folgt $s = 0$ aus (19.2). Außerdem muss u eine Lösung für L^* sein und $f(x_{\max}) = c^t x = b^t u = f^*(u)$.

Sei schließlich y_{\min} eine Lösung für L^* . Gilt (2) in Folgerung 19.13 für y , so wäre $A^t(y_{\min} + y) \geq c$ und $f^*(y_{\min} + y) = b^t(y_{\min} + y) < b^t y_{\min} = f^*(y_{\min})$. Also muss (1) gelten, d. h. L ist lösbar. \square

Beispiel 19.17. Es kann vorkommen, dass sowohl L als auch L^* unzulässig sind. Zum Beispiel für $A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$, $b = c = (1, 1)^t$.

19.3 Der Simplex-Algorithmus

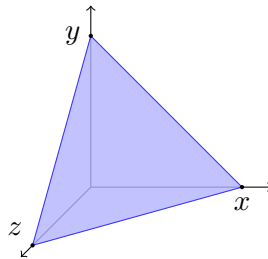
Definition 19.18.

- Die zulässige Menge $P := \{x \in \mathbb{R}_{\geq 0}^m : Ax = b\}$ eines linearen Programms $L = (A, b, c)$ beschreibt einen *Polyeder*. Offenbar ist P konvex. Man nennt $x \in P$ eine *Ecke* von P oder L , wenn auch $P \setminus \{x\}$ konvex ist.
- Für $x \in P$ nennt man $\text{supp}(x) := \{i : x_i > 0\} \subseteq \{1, \dots, m\}$ den *Träger*¹ von x .
- Für $I \subseteq \{1, \dots, m\}$ und $x \in \mathbb{R}^m$ sei $A_I := (a_{ij} : i = 1, \dots, n, j \in I)$, $x_I := (x_i : i \in I)$ und $I' := \{1, \dots, m\} \setminus I$. Man nennt I eine *Basismenge* von L , falls A_I invertierbar ist. Ist zusätzlich $A_I^{-1}b \geq 0$, so nennt man I *zulässig*.

¹engl. *support*

Bemerkung 19.19.

- (a) In der Geometrie definiert man Polyeder als Mengen der Form $P = \{x \in \mathbb{R}^n : Ax \leq b\}$. Beschränkte Polyeder heißen *Polytope*. In \mathbb{R}^2 sind Polytope konvexe Vielecke wie der zulässige Bereich in Beispiel 19.2. Die Eckpunkte sind genau diejenigen Punkte, die man umgangssprachlich als Ecken bezeichnen würde. Die Bedingung, dass $P \setminus \{x\}$ konvex bleibt, bedeutet anschaulich, dass man Ecken „abrunden“ kann.
- (b) Wir hatten in Bemerkung 19.3 gesehen, wie man durch Einführen von Schlupfvariablen Polyeder der Form $\{x : Ax \leq b\}$ in Polyeder der Form $\{x \geq 0 : Ax = b\}$ transformieren kann. Die geometrische Vorstellung der in Definition 19.18 definierten Polyeder bezieht sich also auf einen echten Unterraum. Zum Beispiel beschreibt $P = \{(x, y, z) \geq 0 : x + y + z = 1\}$ ein gleichseitiges Dreieck auf einer Ebene im \mathbb{R}^3 mit den Eckpunkten e_1, e_2 und e_3 .



Lemma 19.20. Für jeden zulässigen Punkt x eines linearen Programms $L = (A, b, c)$ sind äquivalent:

- (1) x ist eine Ecke.
- (2) Für $I := \text{supp}(x)$ gilt $\text{rk}(A_I) = |I|$.

Beweis.

- (1) \Rightarrow (2): Angenommen die Spalten von A_I sind linear abhängig. Dann existiert ein $y \in \mathbb{R}^m$ mit $\text{supp}(y) \subseteq I$ und $Ay = 0$. Außerdem existiert ein $\epsilon > 0$ mit $x \pm \epsilon y \geq 0$. Nun liegen $x \pm \epsilon y$ im Polyeder P von L und $x = \frac{1}{2}(x - \epsilon y) + \frac{1}{2}(x + \epsilon y)$. Dann wäre $P \setminus \{x\}$, aber nicht konvex. Widerspruch.
- (2) \Rightarrow (1): Angenommen $P \setminus \{x\}$ ist nicht konvex. Dann existieren $y, z \in P \setminus \{x\}$ und $0 < \lambda < 1$ mit $x = \lambda y + (1 - \lambda)z$. Wegen $y, z \geq 0$ gilt $\text{supp}(y), \text{supp}(z) \subseteq I$. Dies zeigt $A_I y = b = A_I z$ und $y - z \in \text{Ker}(A_I)$. Da A_I vollen Rang hat, wäre nun $y = z = x$. \square

Beispiel 19.21.

- (a) Im Fall $b = 0$ ist $x = 0$ eine Ecke mit $I = \emptyset$.
- (b) Ist I eine zulässige Basismenge für L , so existiert genau eine zulässige Ecke x mit $\text{supp}(x) \subseteq I$, denn $x_I = A_I^{-1}b \geq 0$. Im Fall $\text{supp}(x) \subsetneq I$ kann x zu mehreren Basismengen gehören.

Folgerung 19.22. Jedes lineare Programm in m Variablen besitzt höchstens $\sum_{k=0}^m \binom{m}{k} \leq 2^m$ Ecken.

Beweis. Jede Ecke x ist durch $I := \text{supp}(x)$ eindeutig bestimmt, denn $A_I x_I = b$ besitzt nach Lemma 19.20 (höchstens) eine Lösung. Die Anzahl der Teilmengen $I \subseteq \{1, \dots, m\}$ ist $|I| \leq m$ ist bekanntlich

$$\sum_{k=0}^m \binom{m}{k} \leq \sum_{k=0}^m \binom{m}{k} = (1 + 1)^m = 2^m. \quad \square$$

Bemerkung 19.23. Die Abschätzung in Folgerung 19.22 ist für $n \geq 2$ nicht optimal. Ist nämlich x eine Ecke, so führt jede Indexmenge I mit $\text{supp}(x) \subseteq I$ und $\text{rk}(A_I) = |I|$ zur gleichen Ecke. Man braucht also nur sogenannte *Antiketten* $\{I_1, \dots, I_k\}$ von $\{1, \dots, m\}$ zu betrachten, d. h. $I_s \not\subseteq I_t$ für alle $s \neq t$. Ein Satz von SPERNER liefert die schärfere Abschätzung $k \leq \binom{m}{\lfloor m/2 \rfloor}$.²

Trotzdem kann die Zahl der Ecken exponentiell mit m wachsen. Zum Beispiel besitzt das lineare Programm mit $A = (1_n, 1_n) \in \mathbb{R}^{n \times 2n}$ und $b = (1, \dots, 1)^t$ genau $2^n = \sqrt{2}^m$ Ecken (jedes I enthält genau ein Element aus $\{i, i+n\}$ für $i = 1, \dots, n$).

Satz 19.24 (Hauptsatz der linearen Optimierung). *Ist das lineare Programm $L = (A, b, c)$ lösbar, so wird das Maximum an einer Ecke angenommen.*

Beweis. Sei x_{\max} eine Lösung von L und $I := \text{supp}(x_{\max})$. Wir können annehmen, dass x_{\max} keine Ecke ist. Wie im Beweis von Lemma 19.20 existieren $y \in \mathbb{R}^m \setminus \{0\}$ und $\epsilon > 0$ mit $\text{supp}(y) \subseteq I$, $Ay = 0$ und $x_{\pm} := x_{\max} \pm \epsilon y \geq 0$. Wegen

$$f(x_{\max}) \pm \epsilon f(y) = f(x_{\pm}) \leq f(x_{\max})$$

ist $f(y) = 0$. Wir können ϵ so wählen, dass $\text{supp}(x_+) \subsetneq I$ oder $\text{supp}(x_-) \subsetneq I$ gilt. Ggf. ist x_+ bzw. x_- eine Lösung von L mit kleinerem Träger. Wenn wir auf diese Weise fortfahren, gelangen wir nach endlich vielen Schritten zu einer Ecke x mit $f(x) = f(x_{\max})$. \square

Bemerkung 19.25. Nach Satz 19.24 und Folgerung 19.22 muss man zur Lösung eines linearen Programms nur endlich viele (aber möglicherweise exponentiell viele) zulässige Ecken betrachten. Trotzdem kann es sein, dass das Maximum an unendlich vielen Punkten angenommen wird, zum Beispiel im trivialen Fall $c = 0$. Mit dem folgenden Satz lässt sich prüfen, ob eine gefundene Ecke eine Lösung ist.

Satz 19.26 (Simplex-Kriterium). *Sei $L = (A, b, c)$ ein lineares Programm mit zulässiger Basismenge I . Ist*

$$\gamma(I) := c_{I'}^t - c_I^t A_I^{-1} A_{I'} \leq 0,$$

so ist die zu I gehörende Ecke eine Lösung von L .

Beweis. O. B. d. A. sei $I = \{1, \dots, n\}$ und $A = (A_I, A_{I'})$. Für jeden zulässigen Vektor x gilt $A_I x_I + A_{I'} x_{I'} = b$. Es folgt $x_I + A_I^{-1} A_{I'} x_{I'} = A_I^{-1} b$ und

$$f(x) = c^t \begin{pmatrix} x_I \\ x_{I'} \end{pmatrix} = c_I^t (A_I^{-1} b - A_I^{-1} A_{I'} x_{I'}) + c_{I'}^t x_{I'} = c_I^t A_I^{-1} b + \gamma(I) x_{I'} \stackrel{x_{I'} \geq 0}{\leq} c_I^t A_I^{-1} b$$

mit Gleichheit, wenn $x_{I'} = 0$, d. h. wenn x die Ecke mit $\text{supp}(x) \subseteq I$ ist. \square

Bemerkung 19.27. In der Praxis ist es aufwendig, alle Ecken zu enumerieren. Die Idee des Simplex-Verfahrens ist es, von einer Ecke x zu einer „benachbarten“ Ecke y mit $f(y) > f(x)$ überzugehen. Benachbart bedeutet anschaulich, dass die Verbindungsstrecke $S := \{\lambda x + (1 - \lambda)y : 0 \leq \lambda \leq 1\}$ eine Kante des Polyeders P bildet, d. h. $P \setminus S$ ist konvex.

²siehe Skript zu Logik und Mengenlehre

Lemma 19.28. Sei $L = (A, b, c)$ ein lineares Programm mit zulässiger Basismenge I . Sei x die dazugehörige Ecke (Beispiel 19.21), $y := x_I$ und $M := (m_{ij}) = A_I^{-1}A_{I'}$. Sei $j \in I'$ mit $\gamma(I)_j > 0$.³ Dann gilt:

(a) Ist die j -te Spalte von M nicht-positiv (d. h. ≤ 0), so ist L unbeschränkt.

(b) Anderenfalls sei $i \in I$ mit

$$\frac{y_i}{m_{ij}} = \min \left\{ \frac{y_k}{m_{kj}} : k \in I, m_{kj} > 0 \right\}.$$

Dann ist auch $J := I \setminus \{i\} \cup \{j\}$ eine zulässige Basismenge und für die entsprechende Ecke x' gilt $f(x') \geq f(x)$.

(c) Ist $y > 0$, so gilt $f(x') > f(x)$ in (b).

Beweis.

(a) Sei $\mu > 0$ und $z \in \mathbb{R}^n$ mit $z_{I'} = \mu e_j \geq 0$ und $z_I = y - \mu M_j \geq y = x_I \geq 0$, wobei $M_j \leq 0$ die j -te Spalte von M bezeichnet. Wegen

$$Az = A_I z_I + A_{I'} z_{I'} = Ax - \mu A_I M_j + \mu A_{I'} e_j = b$$

ist z zulässig und

$$f(z) = c_I^t z_I + \mu c_{I'}^t e_j = c^t x - \mu c_I^t M_j + \mu c_{I'}^t e_j = c^t x + \mu \gamma(I)_j \rightarrow \infty \quad (\mu \rightarrow \infty).$$

Also ist L unbeschränkt.

(b,c) Offenbar ist $S := A_I^{-1}A_J = (e_1, \dots, e_{i-1}, M_j, e_{i+1}, \dots, e_n)$. Durch Entwicklung nach der i -ten Zeile erhält man $\det(S) = m_{ij} > 0$. Insbesondere ist S invertierbar. Damit muss auch A_J invertierbar sein, d. h. J ist eine Basismenge. Wir definieren $x' \in \mathbb{R}^m$ durch

$$x'_k = \begin{cases} y_k - \frac{y_i}{m_{ij}} m_{kj} & \text{falls } k \in I \setminus \{i\}, \\ \frac{1}{m_{ij}} y_i & \text{falls } k = j, \\ 0 & \text{sonst.} \end{cases}$$

Nach Wahl von i gilt $y_k \geq \frac{y_i}{m_{ij}} m_{kj}$ für $k \in I \setminus \{i\}$. Dies zeigt $x' \geq 0$. Außerdem ist

$$\begin{aligned} A_I^{-1} A x' &= A_I^{-1} A_J x'_J = \sum_{k \neq i} \left(y_k - \frac{y_i}{m_{ij}} m_{kj} \right) e_k + \frac{y_i}{m_{ij}} M_j \\ &= \sum_{k \in I} \left(y_k - \frac{y_i}{m_{ij}} m_{kj} \right) e_k + \frac{y_i}{m_{ij}} M_j = y = x_I \end{aligned}$$

und $A x' = A_I x_I = Ax = b$. Also ist x' eine zulässige Ecke für J . Schließlich gilt

$$f(x') = \sum_{k \in I} c_k \left(y_k - \frac{y_i}{m_{ij}} m_{kj} \right) + \frac{y_i}{m_{ij}} c_j = f(x) - \frac{y_i}{m_{ij}} c_I^t M_j + \frac{y_i}{m_{ij}} c_j = f(x) + \frac{y_i}{m_{ij}} \gamma(I)_j \geq f(x)$$

mit Gleichheit genau dann, wenn $y_i = 0$. □

³Wir indizieren die Komponenten von $\gamma(I)$ und die Spalten von M mit I' anstelle von $1, \dots, n - m$.

Bemerkung 19.29 (Simplex-Algorithmus). Gegeben sei ein lineares Programm $L = (A, b, c)$.

- (1) Man bestimme eine zulässige Basismenge I_1 und die dazugehörige Ecke x_1 . Dies lässt sich notfalls erreichen, indem man b als Spalte von A anfügt und $\{m + 1\}$ zu einer Basismenge ergänzt.
- (2) Für $k = 1, 2, \dots$ wiederhole man:
 - (a) Ist das Simplex-Kriterium $\gamma(I_k) \leq 0$ erfüllt, so ist x_k eine Lösung für L .
 - (b) Anderenfalls kann man Lemma 19.28 anwenden. Ist Lemma 19.28(a) erfüllt, so ist L unbeschränkt und man kann abbrechen.
 - (c) Anderenfalls findet man mit Lemma 19.28(a) eine Ecke x_{k+1} mit $f(x_{k+1}) \geq f(x_k)$.
- (3) In der Praxis kann es vorkommen, dass die Schleife in einen unendlichen Zyklus von Ecken $x_k, x_{k+1}, \dots, x_l, x_k, \dots$ gerät. In diesen Fall kann man mit einer neuen Startecke x_1 beginnen oder künstliche Störungen einbauen.

Man kann Beispiele konstruieren, in denen der Simplex-Algorithmus alle (exponentiell viele) Ecken durchläuft, aber in der Praxis tritt dies nur sehr selten auf.

Beispiel 19.30. Wir betrachten wieder das lineare Programm aus Beispiel 19.6:

$$A = \begin{pmatrix} 1 & 2 & 1 & . & . \\ 1 & 1 & . & 1 & . \\ . & 3 & . & . & 1 \end{pmatrix}, \quad b = (170, 150, 180), \quad c = (300, 500, 0, 0, 0).$$

Offenbar ist $I_1 = \{3, 4, 5\}$ eine zulässige Basismenge mit Ecke $x_1 = (0, 0, 170, 150, 180)^t$, $f(x_1) = 0$ und $A_{I_1} = 1_3$. Es gilt $\gamma(I) = (300, 500)$ und

$$M_1 = A_{I_1} = \begin{pmatrix} 1 & 2 \\ 1 & 1 \\ 0 & 3 \end{pmatrix}.$$

Wir können $j = 1$ in Lemma 19.28 wählen. Dann ist $i = 4$ wegen $170 > 150$. Nun ist $I_2 = \{1, 3, 5\}$ und $x_2 = (150, 0, 20, 0, 180)^t$ mit $f(x_2) = 45.000$. Man berechnet

$$M_2 = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 2 & 0 \\ 1 & 1 \\ 3 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 1 & 1 \\ 3 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 3 & 0 \end{pmatrix}$$

und $\gamma(I_2) = (500, 0) - (300, 0, 0)M = (200, -300)$. Hier muss man $j = 2$ und $i = 3$ wählen. Also ist $I_3 = \{1, 2, 5\}$ und $x_3 = (130, 20, 0, 0, 120)^t$ mit $f(x_3) = 49.000$. Weiter ist

$$M_3 = \begin{pmatrix} 1 & 2 & 0 \\ 1 & 1 & 0 \\ 0 & 3 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 2 & 0 \\ 1 & -1 & 0 \\ -3 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 2 \\ 1 & -1 \\ -3 & 3 \end{pmatrix}$$

und $\gamma(I_3) = (0, 0) - (300, 500, 0)M = (-200, -100)$. Nach dem Simplex-Kriterium ist x_3 eine Lösung. Bis auf die Schlupfvariable x_5 ist dies die in Beispiel 19.2 gefundene Lösung.

20 Gitter und quadratische Formen

20.1 Gitter

Bemerkung 20.1. In den Vektorraumaxiomen wird nicht benutzt, dass der Körper K inverse Elemente bzgl. Multiplikation besitzt. Man kann daher K durch einen beliebigen Ring R ersetzen und spricht dann von R -Moduln (genauer: *Linksmoduln*). Im Allgemeinen ist die Theorie der Moduln beliebig kompliziert und hat nicht mehr viel mit linearer Algebra zu tun. Selbst \mathbb{Z} -Moduln besitzen in der Regel keine Basis (zum Beispiel ist \mathbb{F}_2 auf natürliche Weise ein \mathbb{Z} -Modul, in dem jedes Element linear abhängig ist wegen $2 \cdot x = 0$ für $x \in \mathbb{F}_2$). Wir untersuchen in diesem Abschnitt eine Familie von \mathbb{Z} -Moduln innerhalb \mathbb{R}^n , die per definitionem eine Basis besitzt.

Definition 20.2. Ein *Gitter* ist eine Untergruppe $\Delta \leq (\mathbb{R}^n, +)$ der Form

$$\Delta = \left\{ \sum_{i=1}^m \lambda_i b_i : \lambda_1, \dots, \lambda_m \in \mathbb{Z} \right\} = \mathbb{Z}b_1 + \dots + \mathbb{Z}b_m,$$

wobei $B := \{b_1, \dots, b_m\} \subseteq \mathbb{R}^n$ linear unabhängig ist. Ggf. nennt man B eine *Basis* von Δ und $\text{rk}(\Delta) := |B|$ den *Rang* von Δ . Im Fall $m = n$ sagt man: Δ hat *vollen* Rang. Die Matrix $A \in \mathbb{R}^{m \times n}$ mit Zeilen b_1, \dots, b_m nennt man eine *Erzeugermatrix* von Δ . Außerdem nennt man

$${}_B[\Delta]_B := AA^t = ([b_i, b_j])_{ij} \in \mathbb{R}^{m \times m}$$

die *Gram-Matrix* von Δ bzgl. B .

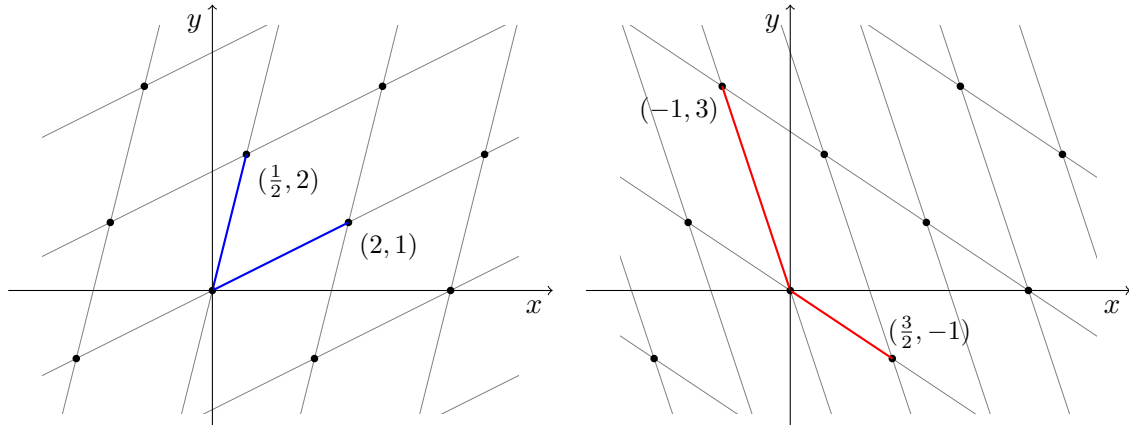
Bemerkung 20.3.

- (a) Für jede Basis B eines Gitters Δ gilt $|B| = \dim\langle B \rangle = \dim\langle \Delta \rangle$. Daher hängt $\text{rk}(\Delta)$ nicht von der Wahl der Basis ab. Indem man \mathbb{R}^n durch $\langle \Delta \rangle$ ersetzt, kann man häufig annehmen, dass Δ vollen Rang hat.
- (b) Da die Erzeugermatrix A vollen Rang hat, ist die Gram-Matrix AA^t positiv definit (Lemma 12.40).

Beispiel 20.4.

- (a) Das *triviale* Gitter mit der Standardbasis $B := \{e_1, \dots, e_n\}$ besteht aus allen Punkten in \mathbb{R}^n mit ganzzahligen Koordinaten. Die Erzeugermatrix und die Gram-Matrix bzgl. B sind 1_n .

(b) Ein Ausschnitt eines Gitters $\Delta \subseteq \mathbb{R}^2$ mit zwei verschiedenen Basen:



(c) Sei $A \in \mathbb{R}^{n \times n}$ positiv definit mit Cholesky-Zerlegung $A = R^t R$. Dann existiert ein Gitter mit Erzeugermatrix R^t und Gram-Matrix A . Auf diese Weise kann man A ein „kanonisches“ Gitter zuordnen.

(d) Für ein Gitter $\Delta \subseteq \mathbb{R}^n$ mit Erzeugermatrix $A \in \mathbb{R}^{m \times n}$ sei

$$\Delta^* := \{x \in \langle \Delta \rangle : \forall d \in \Delta : [x, d] \in \mathbb{Z}\} \subseteq \mathbb{R}^n.$$

Für jedes $x \in \langle \Delta \rangle$ existiert ein $y \in \mathbb{R}^m$ mit $x = yA$. Es gilt

$$x \in \Delta^* \iff \forall b \in B : [x, b] \in \mathbb{Z} \iff xA^t \in \mathbb{Z}^m \iff z := yAA^t \in \mathbb{Z}^m.$$

Für $S := (AA^t)^{-1}A = {}_B[\Delta]_B^{-1}A \in \mathbb{R}^{m \times n}$ gilt daher $\Delta^* = \{zS : z \in \mathbb{Z}^m\}$.¹ Dies zeigt, dass Δ^* ein Gitter mit Erzeugermatrix S ist. Man nennt Δ^* das zu Δ *duale* Gitter. Die Gram-Matrix von Δ^* bzgl. S ist

$$SS^t = {}_B[\Delta]_B^{-1}AA^t{}_B[\Delta]_B^{-1} = {}_B[\Delta]_B^{-1}.$$

Das triviale Gitter Δ ist offenbar *selbstdual*, d. h. $\Delta^* = \Delta$. Im Allgemeinen ist $(\Delta^*)^* = \Delta$, denn

$$({}_B[\Delta]_B^{-1})^{-1}S = {}_B[\Delta]_B S = A.$$

Definition 20.5. Für $n \geq 1$ sei $\text{GL}(n, \mathbb{Z}) := \{A \in \text{GL}(n, \mathbb{Q}) \cap \mathbb{Z}^{n \times n} : A^{-1} \in \mathbb{Z}^{n \times n}\}$.

Lemma 20.6. Für $A \in \mathbb{Z}^{n \times n}$ gilt $A \in \text{GL}(n, \mathbb{Z})$ genau dann, wenn $\det A = \pm 1$.

Beweis. Für $A \in \text{GL}(n, \mathbb{Z})$ gilt $\det A \in \mathbb{Z}$ und $\det(A)^{-1} = \det(A^{-1}) \in \mathbb{Z}$, also $\det A = \pm 1$. Die umgekehrte Inklusion folgt aus Bemerkung 9.24. \square

Lemma 20.7. Seien B und C Basen eines Gitters $\Delta \subseteq \mathbb{R}^n$. Dann existiert ein $S \in \text{GL}(m, \mathbb{Z})$ mit $S^t{}_B[\Delta]_B S = {}_C[\Delta]_C$.

Beweis. Seien $G_B, G_C \in \mathbb{R}^{m \times n}$ die Erzeugermatrizen von Δ bzgl. B bzw. C . Dann existiert $S \in \mathbb{Z}^{m \times m}$ mit $SG_B = G_C$. Es folgt

$${}_C[\Delta]_C = G_C G_C^t = SG_B G_B^t S^t = S_B[\Delta]_B S^t$$

und $\det({}_C[\Delta]_C) = \det(S)^2 \det({}_B[\Delta]_B) \geq \det({}_B[\Delta]_B)$. Aus Symmetriegründen gilt auch $\det({}_B[\Delta]_B) \geq \det({}_C[\Delta]_C)$. Dies zeigt $\det(S) = \pm 1$ und $S \in \text{GL}(m, \mathbb{Z})$ nach Lemma 20.6. \square

¹Übrigens ist $S^t = A^+$ die Pseudoinverse von A .

Definition 20.8.

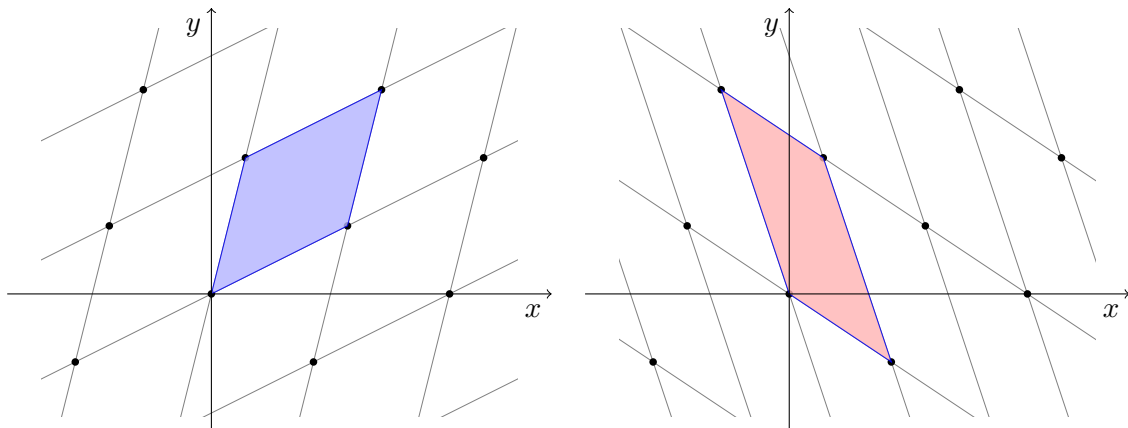
- Ein Gitter Δ heißt *ganz*, wenn ${}_B[\Delta]_B$ für eine Basis B ganzzahlig ist. Nach Lemma 20.7 hängt diese Eigenschaft nicht von der Wahl von B ab.
- Man nennt $\text{disc}(\Delta) := \sqrt{|\det({}_B[\Delta]_B)|}$ die *Diskriminante* von Δ . Auch dies ist unabhängig von B .

Beispiel 20.9.

- (a) Hat Δ vollen Rang mit Erzeugermatrix A , so gilt $\text{disc}(\Delta) = |\det(A)|$. Geometrisch ist $\text{disc}(\Delta)$ in diesem Fall das Volumen der sogenannten *Fundamentalmasche*

$$\{\lambda_1 b_1 + \dots + \lambda_n b_n : 0 \leq \lambda_1, \dots, \lambda_n \leq 1\}$$

für eine Basis b_1, \dots, b_n von Δ (vgl. Beispiel 9.2). Insbesondere gilt $\text{disc}(\Delta) = 7/2$ für das Gitter aus Beispiel 20.4(b).



- (b) Ein Gitter Δ ist genau dann ganz, wenn $\Delta \subseteq \Delta^*$. Im Allgemeinen gilt $\text{disc}(\Delta^*) = \text{disc}(\Delta)^{-1}$.

20.2 Die Minimal-Norm

Definition 20.10. Für ein Gitter $\Delta \subseteq \mathbb{R}^n$ sei

$$\min \Delta := \min\{|x| : x \in \Delta \setminus \{0\}\}$$

die *Minimal-Norm* von Δ .

Bemerkung 20.11.

- (a) Da ein Gitter Δ eine Gruppe ist, gilt $\min \Delta = \min\{|x - y| : x, y \in \Delta, x \neq y\}$.
- (b) Sei A eine Erzeugermatrix von Δ und $G := AA^t$ die Gram-Matrix. Für $x \in \Delta$ existiert ein $s \in \mathbb{Z}^m$ mit $x = sA$. Es folgt $|x| = |sA| = \sqrt{sGs^t}$. Sei $\lambda (> 0)$ der kleinste Eigenwert von G . Nach Lemma 17.55 gilt $|x| \geq \sqrt{\lambda}|s|$. Insbesondere gilt $\min \Delta \geq \sqrt{\lambda}$. Gleichheit wird in der Regel nicht eintreten, da es nicht unbedingt einen normierten ganzzahligen Eigenvektor von G gibt. Allerdings existieren nur endlich viele Vektoren $x \in \Delta$ mit $|x| = \min \Delta$. Man nennt sie *kürzeste* Vektoren. Sowohl die Bestimmung der kürzesten Vektoren als auch die Bestimmung von $\min \Delta$ sind für $n > 10$ schwierige algorithmische Probleme. Darauf beruhen moderne Verschlüsselungsverfahren wie LWE²,

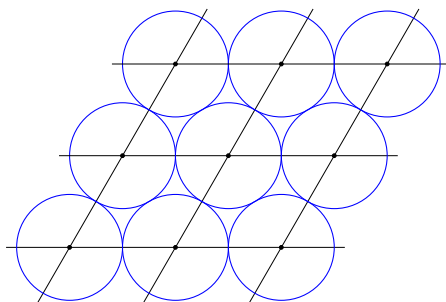
²Learning with errors

die – im Gegensatz zu klassischen Verfahren wie RSA – bislang nicht von Quantencomputern gebrochen werden können.

- (c) HARRIOT hat 1587 untersucht wie man (Kanonen-)Kugeln möglichst dicht im \mathbb{R}^n anordnen kann. Ordnet man die Einheitskugeln regelmäßig an, so bilden die Mittelpunkte ein Gitter Δ mit $\min \Delta \geq 2$ (anderenfalls würden sich Kugeln überschneiden). Die Dichte der Anordnung lässt sich messen, indem man das Volumen der Einheitskugel τ_n durch das Volumen der Fundamentalmasche $\text{disc}(\Delta)$ teilt: $\rho(\Delta) = \frac{\tau_n}{\text{disc}(\Delta)}$.

(1) Für $n = 1$ ist $\Delta = \mathbb{Z}2$ optimal mit $\rho(\Delta) = 1$.

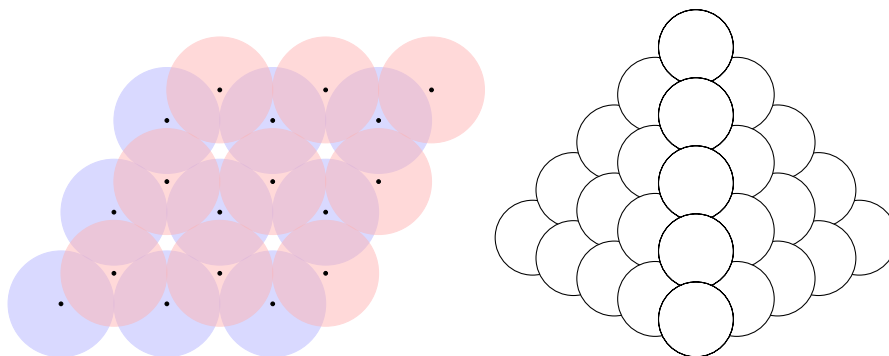
(2) Für $n = 2$ erhält man eine optimale Anordnung durch das hexagonale Gitter $\Delta = \mathbb{Z}(2, 0) + \mathbb{Z}(1, \sqrt{3})$



mit $\min \Delta = 2$ und $\rho(\Delta) = \frac{\pi}{2\sqrt{3}} \approx 0.91$. Das quadratische Gitter $\mathbb{Z}(2, 0) + \mathbb{Z}(0, 2)$ hat nur Dichte $\pi/4 \approx 0.79$.

(3) Für $n = 3$ nimmt man die optimale Anordnung in der Ebene und stapelt diese jeweils versetzt übereinander:

$$\Delta = \mathbb{Z}(2, 0, 0) + \mathbb{Z}(1, \sqrt{3}, 0) + \mathbb{Z}(1, 1/\sqrt{3}, \sqrt{8/3})$$



Es gilt $\min \Delta = 2$ und $\rho(\Delta) = \frac{\pi}{3\sqrt{2}} \approx 0.74$. Gauß bewies, dass es kein Gitter mit höherer Dichte gibt (Satz A.77, vgl. Aufgabe III.27). KEPLER vermutete 1611, dass es generell unmöglich ist, Kugeln noch dichter anzuordnen (auch unregelmäßig).³ Dies wurde erst 1998 durch HALES mit Computer bewiesen.⁴

³Es gibt unendlich viele Möglichkeiten die Ebenen mit der gleichen Dichte versetzt zu stapeln, aber dies sind formal keine Gitter.

⁴Seit 2017 gibt es einen mit Computer verifizierten „formalen“ Beweis.

- (4) Für $4 \leq n \leq 7$ kennt man die dichteste Kugelpackung nicht. Für $n = 8$ bewies VIAZOVSKA 2016, dass das sogenannte E_8 -Gitter mit Erzeuger- und Gram-Matrix

$$\begin{pmatrix} 2 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ -1 & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & -1 & 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & -1 & 1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & -1 & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & -1 & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & -1 & 1 & \cdot \\ 1/2 & 1/2 & 1/2 & 1/2 & 1/2 & 1/2 & 1/2 & 1/2 \end{pmatrix} \begin{pmatrix} 4 & -2 & \cdot & \cdot & \cdot & \cdot & \cdot & 1 \\ -2 & 2 & -1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & -1 & 2 & -1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & -1 & 2 & -1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & -1 & 2 & -1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & -1 & 2 & -1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & -1 & 2 & \cdot \\ 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 2 \end{pmatrix}$$

die dichteste Anordnung ist. Es gilt $\rho(E_8) = \frac{\pi^4}{384} \approx 0.25$. Man sieht, dass E_8 aus den Vektoren x und $x + \frac{1}{2}(1, \dots, 1)$ für $x \in \mathbb{Z}^8$ mit $2 \mid x_1 + \dots + x_8$ besteht. Insbesondere gilt $\min E_8 = 2$. Ein Jahr später wurde auch der Fall $n = 24$ gelöst. Viazovska bekam für diese Leistungen 2022 die Fields-Medaille.

Beispiel 20.12. Für das Gitter $\Delta = \mathbb{Z}(2, 1) + \mathbb{Z}(1/2, 2)$ aus Beispiel 20.4(b) sind $\pm(3/2, -1)$ die kürzesten Vektoren mit $\min \Delta = \sqrt{13}/2$.

Satz 20.13. Eine Untergruppe $U \leq (\mathbb{R}^n, +)$ ist genau dann ein Gitter, wenn für alle $d \in \mathbb{N}$ nur endlich viele $x \in U$ mit $|x| \leq d$ existieren.

Beweis. Für jedes Gitter gilt die angegebene Beschränkung nach Bemerkung 20.11. Sei umgekehrt $U \leq (\mathbb{R}^n, +)$, sodass die Bedingung gilt. Sei $u_1, \dots, u_k \in U$ eine größtmögliche Menge linear unabhängiger Vektoren. Dann ist $U \subseteq \langle u_1, \dots, u_k \rangle =: V$. Im Fall $k = 0$ ist $U = \{0\} = \langle \emptyset \rangle$ ein Gitter vom Rang 0. Sei nun die Behauptung für $k - 1$ bereits bewiesen. Da auch $W := U \cap \langle u_1, \dots, u_{k-1} \rangle$ die Voraussetzung des Satzes erfüllt, ist W ein Gitter mit Basis b_1, \dots, b_{k-1} . Wir betrachten

$$S := \{\lambda_1 b_1 + \dots + \lambda_{k-1} b_{k-1} + \lambda u_k : 0 \leq \lambda_1, \dots, \lambda_{k-1} < 1, 0 \leq \lambda \leq 1\} \cap U.$$

Nach der Dreiecksungleichung ist S beschränkt und daher endlich nach Voraussetzung. Wegen $u_k \in S$ existiert ein $b_n = \lambda_1 b_1 + \dots + \lambda_{k-1} b_{k-1} + \lambda u_k \in S$ mit $\lambda > 0$ minimal. Offenbar sind $b_1, \dots, b_n \in U$ linear unabhängig. Jedes Element $u \in U$ hat die Form $u = \mu_1 b_1 + \dots + \mu_k b_k$ mit $\mu_1, \dots, \mu_k \in \mathbb{R}$. Sei $x_k \in \mathbb{Z}$ mit $0 \leq \mu_k - x_k \lambda < \lambda$. Dann existieren $x_1, \dots, x_{k-1} \in \mathbb{Z}$, sodass

$$u - x_1 b_1 - \dots - x_k b_k = \lambda'_1 b_1 + \dots + \lambda'_{k-1} b_{k-1} + \lambda'_k u_k \in S$$

mit $0 \leq \lambda'_1, \dots, \lambda'_{k-1} < 1$ und $0 \leq \lambda'_k < \lambda$ gilt. Nach Wahl von b_n folgt $\lambda'_k = 0$ und $\lambda'_1 b_1 + \dots + \lambda'_{k-1} b_{k-1} \in W$. Da b_1, \dots, b_{k-1} eine Basis von W ist, muss $\lambda'_1 = \dots = \lambda'_{k-1} = 0$ gelten. Dies zeigt $u = x_1 b_1 + \dots + x_k b_k$. Also ist U ein Gitter mit Basis b_1, \dots, b_k . \square

Folgerung 20.14. Sind $\Delta, \Lambda \subseteq \mathbb{R}^n$ Gitter, so auch $\Delta \cap \Lambda$.

Beweis. Bekanntlich ist $\Delta \cap \Lambda$ eine Untergruppe von $(\mathbb{R}^n, +)$. Mit Δ kann auch $\Delta \cap \Lambda$ nur endlich viele Elemente mit beschränkter Norm enthalten. \square

20.3 Ganzzahlige Matrizen

Bemerkung 20.15. Die Untersuchung von ganzen Gittern führt zur Matrizen mit ganzzahligen Einträgen. Wir hatten in Abschnitt 10.3 bereits Matrizen mit Polynom-Einträgen konstruiert und festgehalten, dass der Gauß-Algorithmus nicht anwendbar, weil dafür dividiert werden müsste. Mit etwas Zahlentheorie⁵ entwickeln wir einen Ersatz für den Gauß-Algorithmus in $\mathbb{Z}^{n \times m}$, womit sich Gleichungssysteme über \mathbb{Z} lösen lassen.

Definition 20.16. Für $a, b \in \mathbb{Z}$ schreiben wir $a \mid b$ (a teilt b), falls ein $c \in \mathbb{Z}$ mit $b = ac$ existiert. Ein $d \in \mathbb{Z}$ heißt *gemeinsamer Teiler* von $a_1, \dots, a_n \in \mathbb{Z}$, falls $d \mid a_1, \dots, d \mid a_n$. Sei $\text{gT}(a_1, \dots, a_n)$ die Menge der gemeinsamen Teiler von a_1, \dots, a_n . Man nennt $d \in \text{gT}(a_1, \dots, a_n)$ *größten gemeinsamen Teiler* und schreibt $\text{ggT}(a_1, \dots, a_n) := d$, falls $d \geq 0$ und $e \mid d$ für alle $e \in \text{gT}(a_1, \dots, a_n)$ gilt. Im Fall $d = 1$ nennen wir a_1, \dots, a_n *teilerfremd* (wie bei Polynomen).

Bemerkung 20.17.

- (a) Sind g und g' größte gemeinsame Teiler von a_1, \dots, a_n , so gilt $g \mid g' \mid g$ und $g = \pm g'$. Wegen $g, g' \geq 0$ ist also $g = g'$, d. h. es existiert höchstens ein gemeinsamer Teiler von a_1, \dots, a_n (dies rechtfertigt die Schreibweise ggT).
- (b) Entgegen seines Namens ist der ggT nicht unbedingt der *größte* gemeinsame Teiler. So ist $\text{gT}(0, 0) = \mathbb{Z}$, aber $\text{ggT}(0, 0) = 0$ (beachte $0 \mid 0$). Allgemeiner ist $\text{ggT}(a_1, \dots, a_n) = 0$ genau dann, wenn $a_1 = \dots = a_n = 0$.
- (c) Man vergleiche folgendes Lemma mit Lemma 15.6.

Lemma 20.18 (BÉZOUT). Für $a_1, \dots, a_n \in \mathbb{Z}$ existieren $b_1, \dots, b_n \in \mathbb{Z}$ mit $a_1 b_1 + \dots + a_n b_n = \text{ggT}(a_1, \dots, a_n)$ und $\text{ggT}(b_1, \dots, b_n) = 1$.

Beweis. Induktion nach n : Im Fall $n = 1$ setzt man $b_1 = 1$. Sei $n \geq 2$ und die Behauptung für $n - 1$ bereits bewiesen. Dann existieren $b_1, \dots, b_{n-1} \in \mathbb{Z}$ mit $\text{ggT}(a_1, \dots, a_{n-1}) = a_1 b_1 + \dots + a_{n-1} b_{n-1}$. Wegen $\text{gT}(a_1, \dots, a_n) = \text{gT}(\text{ggT}(a_1, \dots, a_{n-1}), a_n)$ ist $\text{ggT}(a_1, \dots, a_n) = \text{ggT}(a_1 b_1 + \dots + a_{n-1} b_{n-1}, a_n)$. Daher können wir $n = 2$ annehmen. Sei

$$d := \min\{e \in \mathbb{N} : \exists b_1, b_2 \in \mathbb{Z} : e = a_1 b_1 + a_2 b_2\}.$$

Für $e \in \text{gT}(a_1, a_2)$ gilt dann $e \mid d$, also auch $\text{ggT}(a_1, a_2) \mid d$. Division mit Rest liefert $a_1 = qd + r$ mit $q \in \mathbb{Z}$ und $0 \leq r < d$. Wegen $r = a_1 - qd \in \mathbb{Z}a_1 + \mathbb{Z}a_2$ folgt $r = 0$ aus der Minimalität von d . Also ist $d \mid a_1$ und analog $d \mid a_2$. Dies zeigt $d \mid \text{ggT}(a_1, a_2)$ und $\text{ggT}(a_1, a_2) = d = a_1 b_1 + a_2 b_2$. Für $g := \text{ggT}(b_1, b_2)$ gilt

$$d \mid a_1 \frac{b_1}{g} + a_2 \frac{b_2}{g} = \frac{d}{g},$$

also $g = 1$. □

Lemma 20.19. Für $a_1, \dots, a_n \in \mathbb{Z}$ existiert eine Matrix in $\mathbb{Z}^{n \times n}$ mit erster Zeile (oder Spalte) (a_1, \dots, a_n) und Determinante $\text{ggT}(a_1, \dots, a_n)$.

⁵Mehr Details findet man in meinem Zahlentheorie-Skript

Beweis. Im Fall $a_1 = \dots = a_n = 0$ erfüllt 0_n die Behauptung. Sei also o. B. d. A. $a_1 \neq 0$ (tausche notfalls Spalten/Zeilen). Da die Determinante linear in jeder Zeile (und Spalte) ist, können $\text{ggT}(a_1, \dots, a_n) = 1$ annehmen. Wir argumentieren durch Induktion nach n . Für $n = 1$ wähle man $A = (a_1)$. Sei nun $n \geq 2$. Nach Induktionsvoraussetzung existiert $A_1 \in \mathbb{Z}^{(n-1) \times (n-1)}$ mit erster Zeile (a_1, \dots, a_{n-1}) und

$$g := \det(A_1) = \text{ggT}(a_1, \dots, a_{n-1}) > 0.$$

Nach Bézout existieren $s, t \in \mathbb{Z}$ mit $gs + a_n t = 1$. Sei $b_i := a_i t / g \in \mathbb{Z}$ für $i = 1, \dots, n-1$. Wir konstruieren $A_2 \in \mathbb{Z}^{(n-1) \times (n-1)}$, indem wir die erste Zeile von A_1 streichen und stattdessen $(b_1, \dots, b_{n-1}) = \frac{t}{g}(a_1, \dots, a_{n-1})$ als letzte Spalte ergänzen. Dann gilt $\det(A_2) = (-1)^n t$. Sei

$$A := \begin{pmatrix} & A_1 & & a_n \\ & & & 0 \\ b_1 & \dots & b_{n-1} & s \end{pmatrix}.$$

Durch Entwicklung nach der letzten Spalte, sieht man

$$\det(A) = (-1)^n a_n \det(A_2) + s \det(A_1) = a_n t + gs = 1.$$

Die analoge Aussage für Spalten erhält man durch Transposition. □

Bemerkung 20.20. Der nächste Satz liefert eine ganzzahlige Version der Zeilenstufenform (vgl. Satz 6.10).

Satz 20.21 (HERMITE-Normalform). Für jede Matrix $A \in \mathbb{Z}^{n \times m}$ existiert genau eine nicht-negative Matrix $H \in \mathbb{Z}^{n \times m}$ mit folgenden Eigenschaften:

- (i) $H = SA$ für ein $S \in \text{GL}(n, \mathbb{Z})$.
- (ii) Nullzeilen befinden sich am unteren Ende von H .
- (iii) Für $\tau_i := \min\{1 \leq j \leq m : a_{ij} \neq 0\}$ gilt $\tau_1 < \tau_2 < \dots$ (falls definiert) und $a_{j, \tau_i} < a_{i, \tau_i}$ für $j = 1, \dots, i-1$.

Insbesondere ist H eine obere Dreiecksmatrix.

Beweis. Existenz: Der folgende Algorithmus überführt A in einer Matrix H mit den gewünschten Eigenschaften. Wir durchlaufen die Spalten von A von links nach rechts. Sei (a_1, \dots, a_n) die erste Spalte von A . Nach Bézout existieren $b_1, \dots, b_n \in \mathbb{Z}$ mit $d := a_1 b_1 + \dots + a_n b_n$ und $\text{ggT}(b_1, \dots, b_n) = 1$. Nach Lemma 20.19 existiert ein $S \in \text{GL}(n, \mathbb{Z})$ mit erster Zeile (b_1, \dots, b_n) . Indem wir A durch SA ersetzen, erreichen wir $a_1 = d \geq 0$. Dabei verändern sich jedoch auch a_2, \dots, a_n . Wenn wir diesen Schritt wiederholen kann d höchstens kleiner werden. Nach endlich vielen Wiederholungen ist $a_1 \in \text{gT}(a_2, \dots, a_n)$. Für $i \geq 2$ existieren teilerfremde $b_1, b_i \in \mathbb{Z}$ mit $a_1 b_1 = a_i b_i$. Nach Lemma 20.19 existiert ein $S \in \text{GL}(n, \mathbb{Z})$ mit i -ter Zeile $(b_1, 0, \dots, 0, -b_i, 0, \dots, 0)$. Nachdem wir A durch SA ersetzt haben, gilt $a_i = 0$. Auf diese Weise erreichen wir $a_2 = \dots = a_n = 0$.

In der zweiten Spalte von A können wir analog $a_{22} \geq 0 = a_{32} = \dots = a_{n2}$ erreichen. Im Fall $a_{22} = 0$ gehen wir zur dritten Spalte über. Im Fall $a_{12} < 0$ oder $a_{12} > a_{22}$ addieren bzw. subtrahieren wir a_{22} von a_{12} durch Multiplikation mit einer Elementarmatrix von links (dabei wird die erste Spalte von A nicht verändert). Indem man diese Schritt genügend oft wiederholt, erhält man $0 \leq a_{12} < a_{22}$. Hat man alle Spalten auf diese Weise durchlaufen, so hat A (bzw. H) die gewünschten Eigenschaften.

Eindeutigkeit: Für die Eindeutigkeit von H argumentieren wir wie im Beweis von Satz 6.10. Wir können o. B. d. A. annehmen, dass A selbst die Eigenschaften von H erfüllt. Sei a_i (bzw. s_i) die i -te Spalte von A (bzw. S). Dann ist $h_i = Sa_i$ die i -te Spalte von H . Sei $a_i \in \langle e_1^t, \dots, e_k^t \rangle$. Wir zeigen $a_i = h_i$ und $s_k = e_k^t$ durch Induktion nach k . Im Fall $k = 0$ ist $a_i = 0$ und $h_i = Sa_i = 0$. Sei nun die Behauptung bis $k - 1$ bereits bewiesen. Die erste Spalte (von links) von A , die nicht in $\langle e_1^t, \dots, e_{k-1}^t \rangle$ liegt, hat die Form $a_i = (a_{1i}, \dots, a_{ki}, 0, \dots, 0)^t$ mit $a_{ki} > 0$. Nach Induktionsvoraussetzung gilt $h_{ji} = a_{ji} + s_{jk}a_{ki}$ für $j < k$ und $h_{ki} = s_{kk}a_{ki}$. Da S invertierbar ist, gilt $s_{kk} \neq 0$ und $h_{ki} \neq 0$. Daher muss $i = \tau_k$ und $h_{ki} > 0$ gelten. Es folgt $s_{jk}a_{ki} = h_{ji} = 0$ und $s_{ji} = 0$ für $j > k$. Aus $\det(S) = 1$ ergibt sich $s_{kk} = 1$ und $h_{ki} = a_{ki}$. Für $j < k$ gilt

$$0 \leq h_{ji} = a_{ji} + s_{jk}a_{ki} < h_{ki} = a_{ki}.$$

Aus $0 \leq a_{ji} < a_{ki}$ folgt $s_{jk} = 0$ und $s_k = e_k$. Außerdem ist $a_i = h_i$ wie behauptet. \square

Beispiel 20.22. Hat man ein Gitter Δ mit ganzzahliger Erzeugermatrix A , so kann man mit der Hermite-Normalform eine kanonische Basis bestimmen:

$$A := \begin{pmatrix} 0 & 2 & 6 \\ 3 & 3 & 0 \\ 2 & 2 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & -1 \\ 0 & 2 & 6 \\ 2 & 2 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & -1 \\ 0 & 2 & 6 \\ 0 & 0 & 3 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 2 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Bemerkung 20.23.

- (a) Die Eigenschaften der Hermite-Normalform machen keinen Gebrauch von der Ganzzahligkeit. Man kann daher auch rationale Matrizen in die Hermite-Normalform überführen, indem man zwischenzeitlich mit dem Hauptnenner aller Einträge multipliziert und den Algorithmus auf die entstandene ganzzahlige Matrix anwendet.
- (b) In der Hermite-Normalform operiert man nur auf den Zeilen (also von links) einer gegebenen Matrix. Erlaubt man auch Spaltenoperationen, so gelangt man zu einer eindeutig bestimmten Diagonalmatrix.

Satz 20.24. Für $A \in \mathbb{Z}^{n \times m}$ existieren eindeutig bestimmte nicht-negative Zahlen $d_1 \mid d_2 \mid \dots$ mit

$$SAT = \begin{pmatrix} d_1 & & 0 \\ & d_2 & \\ 0 & & \ddots \end{pmatrix} \quad (\text{SMITH-Normalform})$$

für gewisse $S \in \text{GL}(n, \mathbb{Z})$ und $T \in \text{GL}(m, \mathbb{Z})$. Man nennt $d_1, \dots, d_{\min\{n, m\}}$ Elementarteiler von A .

Beweis. Existenz: Für die Nullmatrix ist die Behauptung klar (beachte $0 \mid 0$). Wie im Beweis von Satz 20.21 kann man $a_{11} > 0 = a_{21} = \dots = a_{n1}$ annehmen. Das gleiche Vorgehen mit den Spalten (also durch Multiplikation mit $T \in \text{GL}(m, \mathbb{Z})$ von rechts) führt zu $A = \text{diag}(d, A_1)$. Ist nicht jeder Eintrag von A_1 durch d teilbar, so addieren man eine entsprechende Spalte von A_1 zur ersten Spalte von A (die entsprechende Elementarmatrix hat Determinante 1). Dann fängt man von vorne an und reduziert d auf diese Weise weiter. Nach endlich vielen Schritten sind alle Einträge von A_1 durch d teilbar. Nun wendet man das Verfahren auf A_1 an. Dabei bleibt die Teilbarkeit durch d in jedem Schritt erhalten. Am Ende erhält man die gesuchte Diagonalmatrix.

Eindeutigkeit: Sei $D_k(A)$ der ggT aller Determinanten von $k \times k$ -Untermatrizen von A . Die Spalten von AT sind ganzzahlige Linearkombinationen der Spalten von A . Dies gilt auch für $k \times k$ -Untermatrizen. Also ist $D_k(A) \mid D_k(AT) \mid D_k(ATT^{-1}) = D_k(A)$. Analog gilt $D_k(A) = D_k(SA)$ und daher $D_k(A) = D_k((d_i \delta_{ij})) = d_1 \dots d_k$. Also sind die Elementarteiler eindeutig durch A bestimmt. \square

Bemerkung 20.25.

- (i) Ist A quadratisch, so ist $|\det A|$ das Produkt der Elementarteiler von A . Der obige Beweis zeigt wie man die Elementarteiler durch Unterdeterminanten charakterisieren kann.
- (ii) Im Gegensatz zum Gauß-Algorithmus ist das Verfahren für die Smith-Normalform wesentlich aufwendiger.
- (iii) Die Smith-Normalform erlaubt es ganzzahlige lineare Gleichungssysteme $Ax = b$ zu lösen. Sei dafür $SAT = D = (d_i \delta_{ij})$ wie in Satz 20.24 mit $d_1, \dots, d_k > 0 = d_{k+1} = \dots = d_{\min\{n,m\}}$. Das einfachere System $Dy = Sb = (b_1, \dots, b_n)^t$ ist genau dann lösbar, wenn $d_i \mid b_i$ für $i = 1, \dots, n$ gilt. Die Lösungen lauten dann $y = (b_1/d_1, \dots, b_k/d_k, *, \dots, *)$, wobei $*$ für eine beliebige ganze Zahl steht. Die Lösungen von $Ax = b$ erhält man durch $x = Ty$.

Beispiel 20.26. Wir suchen alle $x \in \mathbb{Z}^3$ mit

$$Ax := \begin{pmatrix} 21 & -5 & 26 \\ 3 & -1 & 4 \\ -8 & 2 & -10 \end{pmatrix} x = \begin{pmatrix} 47 \\ 7 \\ -18 \end{pmatrix} =: b. \quad (20.1)$$

Der Algorithmus im Beweis von Satz 20.24 (modulo offensichtlicher Vereinfachungen) ergibt:

$$\begin{aligned} \begin{pmatrix} 21 & -5 & 26 \\ 3 & -1 & 4 \\ -8 & 2 & -10 \end{pmatrix} &\sim \begin{pmatrix} -1 & 3 & 4 \\ -5 & 21 & 26 \\ 2 & -8 & -10 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 5 & 6 & 6 \\ -2 & -2 & -2 \end{pmatrix} \\ &\sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 6 & 6 \\ 0 & -2 & -2 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 6 & 6 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} =: D. \end{aligned}$$

Für

$$S := \begin{pmatrix} -5 & -8 & -2 \\ -1 & -1 & 0 \\ 2 & 3 & 1 \end{pmatrix} \in \text{GL}(3, \mathbb{Z}), \quad T := \begin{pmatrix} -1 & 1 & -2 \\ -1 & 0 & -1 \\ 2 & -2 & 3 \end{pmatrix} \in \text{GL}(3, \mathbb{Z})$$

gilt $SAT = A$. Daher ist (20.1) äquivalent zu

$$Dy = S^{-1}b = \begin{pmatrix} -3 \\ -4 \\ 0 \end{pmatrix}$$

mit $y := Tx$. Dies zeigt $y = (-3, -2, a)$ mit $a \in \mathbb{Z}$. Also ist $x = T^{-1}y = (a - 4, 5 - a, 6 - a)$ für $a \in \mathbb{Z}$.

Satz 20.27. Seien $\Lambda \subseteq \Delta \subseteq \mathbb{R}^n$ Gitter. Dann existiert eine Basis b_1, \dots, b_k von Δ und eindeutig bestimmte natürliche Zahlen $d_1 \mid d_2 \mid \dots \mid d_l$, sodass $d_1 b_1, \dots, d_l b_l$ eine Basis von Λ ist.

Beweis. Existenz: Seien a_1, \dots, a_k und c_1, \dots, c_l zunächst beliebige Basen von Δ bzw. Λ . Dabei gilt $l = \text{rk}(\Lambda) \leq \text{rk}(\Delta) = k$. Wir schreiben $c_i = \sum_{j=1}^k x_{ij} a_j$ mit $x_{ij} \in \mathbb{Z}$ für $i = 1, \dots, l$ und setzen $A := (x_{ij}) \in \mathbb{Z}^{l \times k}$. Nach Satz 20.24 existieren $S = (s_{ij}) \in \text{GL}(l, \mathbb{Z})$ und $T = (t_{ij}) \in \text{GL}(k, \mathbb{Z})$ mit $A = S(\delta_{ij} d_i)T$ und $d_1 \mid \dots \mid d_l$. Die Elemente $b_i := \sum_{j=1}^k t_{ij} a_j \in \Delta$ für $i = 1, \dots, k$ bilden eine Basis von Δ . Wegen

$$c_i = \sum_{j=1}^k x_{ij} a_j = \sum_{j=1}^k \sum_{m=1}^l s_{im} d_m t_{mj} a_j = \sum_{m=1}^l s_{im} d_m b_m$$

ist $\{d_1 b_1, \dots, d_l b_l\}$ eine Basis von Λ .

Eindeutigkeit: Sei b'_1, \dots, b'_k eine weitere Basis von Δ und $d'_1 \mid \dots \mid d'_l$, sodass $d'_1 b'_1, \dots, d'_l b'_l$ eine Basis von Λ ist. Dann existieren $S = (s_{ij}) \in \text{GL}(k, \mathbb{Z})$ und $T = (t_{ij}) \in \text{GL}(l, \mathbb{Z})$ mit $b'_i = \sum_{j=1}^k s_{ij} b_j$ und $d'_i b'_i = \sum_{j=1}^l t_{ij} d_j b_j$ für alle i . Ein Koeffizientenvergleich zeigt $d'_i s_{ij} = t_{ij} d_j$ und

$$(\delta_{ij} d'_i) S = T (\delta_{ij} d_i).$$

Aus der Eindeutigkeit der Smith-Normalform folgt $d'_i = d_i$ für $i = 1, \dots, l$. □

20.4 Der LLL-Algorithmus

Bemerkung 20.28. Viele Probleme für Gitter Δ lassen sich algorithmisch lösen, indem man eine Basis aus möglichst „kurzen“, „nahezu orthogonalen“ Vektoren konstruiert. Über $K \in \{\mathbb{R}, \mathbb{C}\}$ kann man mit Gram-Schmidt stets eine Orthonormalbasis konstruieren. Über \mathbb{Z} ist die Situation komplizierter. Im ersten Schritt bringen wir die Gram-Matrix auf eine Blockdiagonalgestalt mit möglichst kleinen Blöcken.

Definition 20.29. Ein Gitter $\Delta \subseteq \mathbb{R}^n$ heißt *zerlegbar*, falls Gitter $\Delta_1, \Delta_2 \subsetneq \Delta$ mit $\Delta = \Delta_1 \oplus \Delta_2$ und $\Delta_1 \subseteq \Delta_2^\perp$ existieren. Ggf. nennt man $\Delta = \Delta_1 \perp \Delta_2$ eine *orthogonale Zerlegung*. Anderenfalls heißt Δ *unzerlegbar*.

Bemerkung 20.30. Ein Gitter Δ ist genau dann zerlegbar, wenn eine Basis B mit ${}_B[\Delta]_B = \text{diag}(A_1, A_2)$ existiert.

Satz 20.31 (EICHLER). *Jedes Gitter Δ besitzt eine orthogonale Zerlegung $\Delta = \Delta_1 \perp \dots \perp \Delta_k$ in bis auf die Reihenfolge eindeutig bestimmte unzerlegbare Gitter $\Delta_1, \dots, \Delta_k$.*

Beweis. Da der Rang nur endlich oft kleiner werden kann, lässt sich Δ stets in unzerlegbare Gitter $\Delta_1, \dots, \Delta_k$ orthogonal zerlegen. Wir nennen $x \in \Delta$ *minimal*, falls x nicht die Summe kürzerer Vektoren ist, d. h. für $y, z \in \Delta$ mit $x = y + z$ gilt $|y| \geq |x|$ oder $|z| \geq |x|$. Da nur endlich viele Vektoren in Δ eine vorgegebene Norm haben (Satz 20.13), lässt sich jedes $x \in \Delta$ als Summe minimaler Elemente schreiben. Wegen $|x_1 + \dots + x_k| = |x_1| + \dots + |x_k|$ für $x_i \in \Delta_i$ liegt jedes minimale Element in einem Δ_i . Sei $x \in \Delta_i$ minimal als Element von Δ_i . Seien $y, z \in \Delta$ mit $x = y + z$. Seien $y = y_1 + \dots + y_k$ und $z = z_1 + \dots + z_k$ die eindeutigen Zerlegungen mit $y_j, z_j \in \Delta_j$ für $j = 1, \dots, k$. Dann gilt $x = y_i + z_i$ und $|y| = |y_1| + \dots + |y_k| \geq |y_i| \geq |x|$ oder $|z| \geq |x|$. Also ist x auch minimal in Δ .

Sei $M \subseteq \Delta$ die Menge aller minimalen Elemente. Wir definieren eine Äquivalenzrelation auf M durch

$$x \sim y \iff \exists z_1, \dots, z_s \in M : x = z_1, y = z_s, [z_i, z_{i+1}] \neq 0 \text{ für } i = 1, \dots, s-1.$$

Elemente aus unterschiedlichen Äquivalenzklassen sind orthogonal und damit linear unabhängig. Daher kann es nur endlich viele Äquivalenzklassen $M_1, \dots, M_l \subseteq M$ geben. Jedes M_i liegt in einem Δ_j und erzeugt als Gruppe ein Gitter $\Lambda_i \subseteq \Delta_j$ (die Menge aller ganzzahligen Linearkombinationen von Elementen aus M_i). Nach Definition ist

$$\Lambda_1 + \dots + \Lambda_l = \Lambda_1 \perp \dots \perp \Lambda_l.$$

Existiert ein $x \in \Delta_j \setminus \Lambda_i$, so kann man x in minimale Elemente (innerhalb Δ_j) zerlegen, wovon mindestens eins nicht in M_i liegt. Dann müsste Δ_j noch weitere Äquivalenzklassen von M enthalten

und man hätte eine orthogonale Zerlegung $\Delta_j = \Lambda_i \perp \Gamma$. Dies widerspricht der Unzerlegbarkeit von Δ_j . Also ist $\Delta_j = \Lambda_i$. Analog folgt $k = l$ und $\Delta_i = \Lambda_i$ für $i = 1, \dots, k$ nach Ummummerierung. \square

Satz 20.32. *Für jedes unzerlegbare ganze Gitter Δ vom Rang $m \geq 2$ gilt $\min \Delta \geq 2$.*

Beweis. Sei $A = (a_{ij}) \in \mathbb{Z}^{m \times m}$ die Gram-Matrix von Δ bzgl. einer Basis. Angenommen es existiert ein $x \in \mathbb{Z}^m$ mit $xAx^t = 1$. Dann ist $\text{ggT}(x_1, \dots, x_m) = 1$. Nach Lemma 20.19 existiert ein $S \in \text{GL}(m, \mathbb{Z})$ mit erster Spalte x . Indem wir A durch $S^t A S$ ersetzen, erreichen wir $a_{11} = 1$. Nun existiert eine obere Dreiecksmatrix $T \in \text{GL}(m, \mathbb{Z})$ mit Einsen auf der Hauptdiagonale, sodass $T^t A T = \text{diag}(1, A_1)$ mit $A_1 \in \mathbb{Z}^{(m-1) \times (m-1)}$ gilt. Dies widerspricht Bemerkung 20.30. \square

Bemerkung 20.33. Der Beweis zeigt, dass jedes (nicht unbedingt ganze) Gitter eine Basis b_1, \dots, b_k mit $|b_1| = \min \Delta$ besitzt.

Satz 20.34 (HERMITE). *Für jedes Gitter $\Delta \subseteq \mathbb{R}^n$ vom Rang k gilt*

$$\min \Delta \leq \left(\frac{4}{3}\right)^{(k-1)/4} \sqrt[k]{\text{disc}(\Delta)}.$$

Beweis. Sei b_1, \dots, b_k eine Basis von Δ und $G = (g_{ij}) \in \mathbb{R}^{k \times k}$ die entsprechende Gram-Matrix. Nach Bemerkung 20.33 können wir $\min \Delta = |b_1| = \sqrt{g_{11}}$ annehmen. Für $k = 1$ gilt $\sqrt{g_{11}} = \text{disc}(\Delta)$ wie behauptet. Sei $k \geq 2$ und die Behauptung für $k - 1$ bereits bewiesen. Die Vektoren

$$c_i := b_i - \frac{[b_i, b_1]}{|b_1|^2} b_1 = b_i - \frac{g_{1i}}{g_{11}} b_1 \in \mathbb{R}^n$$

für $2 \leq i \leq k$ sind orthogonal zu b_1 . Wir betrachten die Gitter Λ mit Basis c_2, \dots, c_k und Λ_1 mit Basis b_1, c_2, \dots, c_k . Für die Gram-Matrizen gilt $G_{\Lambda_1} = S^t G S = \text{diag}(g_{11}, G_\Lambda)$, wobei

$$S := \begin{pmatrix} 1 & -g_{12}/g_{11} & \cdots & -g_{1k}/g_{11} \\ & \ddots & 0 & 0 \\ & & \ddots & 0 \\ & & & 1 \end{pmatrix} \in \text{GL}(k, \mathbb{Q})$$

Determinante 1 hat. Dies zeigt $\text{disc}(\Delta) = \text{disc}(\Lambda_1) = \sqrt{g_{11}} \text{disc}(\Lambda)$. Nach Induktion existiert ein $y = x_2 c_2 + \dots + x_k c_k \in \Lambda$ mit $x_2, \dots, x_k \in \mathbb{Z}$ und

$$|y| = \min \Lambda \leq \left(\frac{4}{3}\right)^{(k-2)/4} \left(\frac{\text{disc}(\Lambda)}{\sqrt{g_{11}}}\right)^{\frac{1}{k-1}}.$$

Außerdem existiert ein $x_1 \in \mathbb{Z}$ mit

$$\mu := \left| x_1 + \frac{g_{12}}{g_{11}} x_2 + \dots + \frac{g_{1k}}{g_{11}} x_k \right| \leq \frac{1}{2}.$$

Nach rechnet leicht $S^{-1} = 2 \cdot 1_k - S$ nach (die Einträge außerhalb der Hauptdiagonale wechseln das Vorzeichen). Für $v := x_1 b_1 + \dots + x_k b_k \in \Delta$ gilt daher

$$g_{11} = (\min \Delta)^2 \leq |v|^2 = x G x^t = x S^{-t} G_{\Lambda_1} S^{-1} x^t = \mu^2 g_{11} + (x_2, \dots, x_k) G_\Lambda (x_2, \dots, x_k)^t$$

$$\leq \frac{g_{11}}{4} + |y|^2 \leq \frac{g_{11}}{4} + \left(\frac{4}{3}\right)^{(k-2)/2} \left(\frac{\text{disc}(\Delta)^2}{g_{11}}\right)^{\frac{1}{k-1}}.$$

Umstellen ergibt

$$g_{11} \leq \left(\frac{4}{3}\right)^{k/2} \left(\frac{\text{disc}(\Delta)^2}{g_{11}}\right)^{\frac{1}{k-1}}, \quad g_{11}^k \leq \left(\frac{4}{3}\right)^{k(k-1)/2} \text{disc}(\Delta)^2. \quad \square$$

Bemerkung 20.35. Ein Gitter mit Gram-Matrix $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ liefert Gleichheit in Hermites Schranke. Für $k \geq 3$ ist die Abschätzung jedoch nicht optimal. Die kleinste Zahl γ_k mit $\min \Delta \leq \sqrt{\gamma_k} \sqrt[k]{\text{disc}(\Delta)}$ für alle Gitter Δ vom Rang k nennt man die k -te *Hermite-Konstante* (vgl. Bemerkung 20.46). Mit Maßtheorie hat BLICHFELDT

$$\gamma_k \leq \frac{2}{\pi} \Gamma\left(\frac{k}{2} + 2\right)^{2/k} \leq \frac{2}{\pi} \left(\left\lfloor \frac{k+3}{2} \right\rfloor!\right)^{2/k} \leq k$$

bewiesen, wobei Γ die eulersche Gammafunktion ist (vgl. Aufgabe III.28). Bislang kennt man nur die folgenden Werte:

n	1	2	3	4	5	6	7	8	24
γ_n^n	1	$\frac{4}{3}$	2	4	8	$\frac{64}{3}$	64	2^8	2^{48}

Die Werte $\gamma_2^2 = 4/3$ und $\gamma_3^3 = 2$ werden in Satz 20.55 und Satz A.76 hergeleitet. Für $n = 8$ liefert E_8 aus Bemerkung 20.11 den optimalen Wert $\gamma_8 = 2$.

Lemma 20.36. Sei $\Delta \subseteq \mathbb{R}^k$ ein Gitter mit Basis v_1, \dots, v_k . Sei b_1, \dots, b_k die Orthogonalbasis aus dem Gram-Schmidt-Verfahren angewendet auf v_1, \dots, v_k (ohne Normierung). Dann gilt $\text{disc}(\Delta) = |b_1| \dots |b_k|$.

Beweis. Sei $A \in \mathbb{R}^{k \times n}$ die Erzeugermatrix mit den Zeilen v_1, \dots, v_k . Sei B die Matrix mit Zeilen b_1, \dots, b_k . Nach dem Gram-Schmidt-Verfahren existiert eine untere Dreiecksmatrix $R \in \mathbb{R}^{k \times k}$ mit Einsen auf der Hauptdiagonale und $RA = B$. Wegen $\det(R) = 1$ gilt

$$\text{disc}(\Delta) = \sqrt{\det(AA^t)} = \sqrt{\det(BB^t)} = \sqrt{\det(\text{diag}(|b_1|^2, \dots, |b_k|^2))} = |b_1| \dots |b_k|. \quad \square$$

Bemerkung 20.37.

- (a) Im Folgenden sei $[x] \in \mathbb{Z}$ der gerundete Wert von $x \in \mathbb{R}$, d. h. $|x - [x]| \leq \frac{1}{2}$ (es spielt keine Rolle, ob man *.5 auf- oder abrundet).
- (b) Seien $v_1, \dots, v_k \in \mathbb{R}^n$ linear unabhängig und b_1, \dots, b_k die dazu gehörige Gram-Schmidt-Orthogonalbasis. Dann existiert ein $w_k \in \langle b_1, \dots, b_{k-1} \rangle = \langle v_1, \dots, v_{k-1} \rangle$ mit $v_k = b_k + w_k$. Daher ist b_k das Bild der Projektion von v_k auf $\langle b_1, \dots, b_{k-1} \rangle^\perp$. Insbesondere ändert sich b_k nicht, wenn man v_k durch ein Element aus $v_k + \langle v_1, \dots, v_{k-1} \rangle$ ersetzt.

Definition 20.38. Sei $\frac{1}{4} < \delta < 1$. Sei v_1, \dots, v_k eine Basis eines Gitters $\Delta \subseteq \mathbb{R}^n$ und b_1, \dots, b_k die dazu gehörige Gram-Schmidt-Orthogonalbasis. Sei $\mu_{ij} := \frac{|v_i, b_j|}{|b_j|^2}$ für $j < i$. Man nennt v_1, \dots, v_k δ -reduziert, falls die folgenden Bedingungen gelten:

- (Längen-Bedingung) $|\mu_{ij}| \leq \frac{1}{2}$ für alle $1 \leq j < i \leq k$,
- (LOVÁSZ-Bedingung) $|b_i|^2 \geq (\delta - \mu_{i,i-1}^2) |b_{i-1}|^2$ für alle $2 \leq i \leq k$.

Satz 20.39 (LLL-Algorithmus⁶). *Der folgende Algorithmus überführt eine Basis v_1, \dots, v_k eines Gitters $\Delta \subseteq \mathbb{R}^n$ in einer δ -reduzierte Basis.*

(1) *Berechne die Gram-Schmidt-Orthogonalbasis b_1, \dots, b_k aus v_1, \dots, v_k .*

(2) *Berechne μ_{ij} für $1 \leq j < i \leq k$.*

(3) *Setze $i = 2$.*

(4) *Solange $i \leq k$ wiederhole:*

(a) *Für $j = i - 1, i - 2, \dots, 1$:*

- *Ersetze v_i durch $v_i - \lfloor \mu_{ij} \rfloor v_j$.*
- *Aktualisiere μ_{il} für $l = 1, \dots, j$.*

(b) *Wenn die Lovász-Bedingung für v_i gilt: erhöhe i .*

(c) *Anderenfalls: tausche v_i mit v_{i+1} und gehe zu Schritt (1).*

Beweis. Wir zeigen zuerst, dass der Algorithmus terminiert. Die Zahlen

$$d_s := |b_1| \dots |b_s| \qquad D := d_1 \dots d_k$$

ändern sich nur in Schritt (1). Angenommen v_i verletzt die Lovász-Bedingung, d. h. $|b_i|^2 < (\delta - \mu_{i,i-1}^2)|b_{i-1}|^2$. Durch den Tausch $v_i \leftrightarrow v_{i-1}$ bleiben d_1, \dots, d_{i-2} unberührt, während b_{i-1} durch

$$b'_{i-1} = v_i - \sum_{j < i-1} \mu_{ij} b_j = b_i + \mu_{i,i-1} b_{i-1}$$

ersetzt wird. Wegen $[b_{i-1}, b_i] = 0$ ist

$$|b'_{i-1}|^2 = |b_i|^2 + \mu_{i,i-1}^2 |b_{i-1}|^2 < \delta |b_{i-1}|^2.$$

Daher wird d_{i-1} durch

$$|b_1| \dots |b_{i-2}| |b'_{i-1}| < \sqrt{\delta} d_{i-1}$$

ersetzt. Für $j \geq i$ erzeugen v_1, \dots, v_j und $v_1, \dots, v_i, v_{i-1}, \dots, v_j$ offensichtlich das gleiche Gitter Δ_j . Daher ändern sich d_i, \dots, d_k nach Lemma 20.36 nicht. Insgesamt wird D also um mindestens den Faktor $\sqrt{\delta} < 1$ verkleinert. Nach Lemma 20.36 und Hermite gilt

$$d_j = \text{disc}(\Delta_j) \geq \left(\frac{3}{4}\right)^{j(j-1)/4} (\min \Delta_j)^j \geq \left(\frac{3}{4}\right)^{j(j-1)/4} (\min \Delta)^j.$$

Daher ist D nach unten durch eine positive Konstante beschränkt, die nur von Δ abhängt. Dies zeigt, dass die Lovász-Bedingung nur endlich oft verletzt sein kann. Also wird nach endlich vielen Schritten $i = k + 1$ erreicht und der Algorithmus terminiert.

Am Ende des Algorithmus ist die Lovász-Bedingung für v_1, \dots, v_k erfüllt. In Schritt (4) wird v_i verändert. Nach Bemerkung 20.37 bleiben b_1, \dots, b_k davon unberührt. Für $j = i - 1$ ersetzt man μ_{ij} durch

$$\frac{[v_i - \lfloor \mu_{ij} \rfloor v_j, b_j]}{|b_j|^2} = \frac{[v_i, b_j]}{|b_j|^2} - \lfloor \mu_{ij} \rfloor = \mu_{ij} - \lfloor \mu_{ij} \rfloor$$

(beachte $v_j \in b_j + \langle b_1, \dots, b_{j-1} \rangle$). Anschließend gilt $|\mu_{i,i-1}| \leq 1/2$. Für $j = i - 2$ erhält man analog $|\mu_{i,i-2}| \leq 1/2$, während $\mu_{i,i-1}$ nicht mehr verändert wird, da j absteigend durchlaufen wird. Am Ende gilt $|\mu_{ij}| \leq 1/2$ für alle $j < i$, d. h. die Längen-Bedingung ist erfüllt. Insgesamt ist v_1, \dots, v_k δ -reduziert. \square

⁶Gesprochen: Tripel-L-Algorithmus. Benannt man A. LENSTRA, H. LENSTRA und L. LOVÁSZ.

Bemerkung 20.40. Man kann zeigen, dass der LLL-Algorithmus polynomiale Laufzeit in k besitzt. Größere Werte von δ (d. h. nahe bei 1) liefern kürzere Basisvektoren bei längerer Laufzeit. In vielen Implementierungen wird $\delta = 3/4$ als Kompromiss gewählt. Tatsächlich terminiert der Algorithmus auch für $\delta = 1$, aber nicht unbedingt in polynomialer Laufzeit (ohne Beweis).

Satz 20.41. Sei v_1, \dots, v_k eine δ -reduzierte Basis eines Gitters $\Delta \subseteq \mathbb{R}^n$. Sei $\rho := \frac{2}{\sqrt{4\delta-1}}$. Dann gilt

$$|v_1| \leq \rho^{k-1} \min \Delta, \quad |v_1| \dots |v_k| \leq \rho^{k(k-1)/2} \text{disc}(\Delta).$$

Beweis. Sei b_1, \dots, b_k die Gram-Schmidt-Orthogonalbasis zu v_1, \dots, v_k . Es gilt $\sigma := \rho^2 = \frac{1}{\delta-1/4} > \frac{4}{3}$. Aus der δ -Reduktion folgt

$$|b_i|^2 \geq (\delta - \mu_{i,i-1}^2) |b_{i-1}|^2 \geq \sigma^{-1} |b_{i-1}|^2 \geq \dots \geq \sigma^{j-i} |b_j|^2$$

für $j < i$. Sei $v = \lambda_1 v_1 + \dots + \lambda_i v_i = \eta_1 b_1 + \dots + \eta_i b_i \in \Delta$ ein kürzester Vektor mit $\lambda_1, \dots, \lambda_i \in \mathbb{Z}$, $\eta_1, \dots, \eta_i \in \mathbb{R}$ und $\lambda_i \neq 0$. Nach Konstruktion von b_1, \dots, b_k gilt $\eta_i = \lambda_i$. Insbesondere ist $|\eta_i| \geq 1$. Da b_1, \dots, b_k paarweise orthogonal sind, folgt

$$(\min \Delta)^2 = |v|^2 = \eta_1^2 |b_1|^2 + \dots + \eta_i^2 |b_i|^2 \geq |b_i|^2 \geq \sigma^{1-i} |b_1|^2 = \sigma^{1-i} |v_1|^2 \geq \sigma^{1-k} |v_1|^2$$

und $|v_1| \leq \rho^{k-1} \min \Delta$. Wir zeigen

$$1 + \frac{1}{4} \sum_{j=1}^{i-1} \sigma^j \leq \sigma^{i-1}$$

durch Induktion nach i . Für $i = 1$ ist $1 = \sigma^0$. Sei $i \geq 2$ und die Behauptung für $i - 1$ bereits bewiesen. Dann ist

$$1 + \frac{1}{4} \sum_{j=1}^{i-1} \sigma^j \leq \sigma^{i-2} + \frac{1}{4} \sigma^{i-1} = \sigma^{i-2} \left(1 + \frac{1}{4} \sigma \right) \leq \sigma^{i-1}.$$

Wie oben ist daher

$$|v_i|^2 = \left| b_i + \sum_{j=1}^{i-1} \mu_{ij} b_j \right|^2 = |b_i|^2 + \sum_{j=1}^{i-1} \mu_{ij}^2 |b_j|^2 \leq |b_i|^2 + \frac{1}{4} \sum_{j=1}^{i-1} \sigma^{i-j} |b_i|^2 \leq \sigma^{i-1} |b_i|^2.$$

Es folgt

$$|v_1| \dots |v_k| \leq \rho^{k(k-1)/2} |b_1| \dots |b_k| \stackrel{20.36}{=} \rho^{k(k-1)/2} \text{disc}(\Delta). \quad \square$$

Bemerkung 20.42. Für $\delta = 3/4$ erhält man $\rho = \sqrt{2} \approx 1.41$. Mit $\delta \rightarrow 1$ nähert sich ρ den Faktor $2/\sqrt{3} \approx 1.15$ aus der Hermite-Schrank an. Darauf kommen wir in Bemerkung 20.63 zurück. Die Abschätzung für $|v_1|$ suggeriert, dass v_1 unter den Basisvektoren der kürzeste ist. Dies ist jedoch nicht immer der Fall wie das folgende Beispiel zeigt.

Beispiel 20.43. Wir betrachten ein Gitter Δ mit Gram-Matrix

$$\begin{pmatrix} 14 & 7 & 5 & 3 & -2 \\ 7 & 11 & 3 & 6 & 3 \\ 5 & 3 & 17 & -5 & 4 \\ 3 & 6 & -5 & 11 & 3 \\ -2 & 3 & 4 & 3 & 9 \end{pmatrix}$$

und $\text{disc}(\Delta) = 169$. Hermite liefert $\min \Delta \leq \frac{4}{3} \sqrt[5]{169} \approx 3.72$. Der LLL-Algorithmus mit $\delta = 3/4$ bzw. $\delta = 9/10$ liefert reduzierte Basen mit Gram-Matrizen

$$\begin{pmatrix} 14 & 7 & 4 & -4 & 5 \\ 7 & 11 & 1 & -5 & -1 \\ 4 & 1 & 7 & -1 & 3 \\ -4 & -5 & -1 & 10 & 1 \\ 5 & -1 & 3 & 1 & 10 \end{pmatrix}, \quad \begin{pmatrix} 7 & 1 & -1 & 3 & 0 \\ 1 & 11 & -5 & -4 & 3 \\ -1 & -5 & 10 & 1 & 0 \\ 3 & -4 & 1 & 11 & -5 \\ 0 & 3 & 0 & -5 & 9 \end{pmatrix}.$$

Es gibt bis auf Vorzeichen nur einen kürzesten Vektor und $\min \Delta = \sqrt{7} \approx 2.65$.

20.5 Quadratische Formen

Bemerkung 20.44. Wir verändern unseren Blickwinkel auf Gitter, indem wir die Erzeugermatrix vernachlässigen und stattdessen nur noch Gram-Matrizen betrachten. Wir kennen dann zwar den Rang, aber nicht mehr den übergeordneten euklidischen Raum. Die Situation ist hier ähnlich wie mit Bilinearformen und deren Gram-Matrizen. In Bemerkung 12.4 hatten wir quadratische Formen über Bilinearformen eingeführt. Wir wiederholen die Definition mit \mathbb{Z} anstelle eines Körpers.

Definition 20.45.

- Für eine symmetrische Matrix $C \in \mathbb{R}^{n \times n}$ nennt man die Abbildung

$$q = q_C: \mathbb{Z}^n \rightarrow \mathbb{R}, \quad x \mapsto xCx^t = \sum_{i,j=1}^n c_{ij}x_i x_j$$

eine *quadratische Form* vom Rang n mit *Gram-Matrix* C .

- Man nennt q_C
 - *nicht-ausgeartet*, falls $C \in \text{GL}(n, \mathbb{R})$.
 - *positiv* (bzw. *negativ*), falls C positiv (bzw. negativ) definit ist, d. h. $q(x) > 0$ (bzw. $q(x) < 0$) für alle $x \neq 0$. Ist q positiv, so nennt man

$$\min q := \min\{|q(x)| : x \in \mathbb{Z}^n \setminus \{0\}\}$$

das *Minimum* von q .

- *ganz*, falls $C \in \mathbb{Z}^{n \times n}$.
- Man nennt $\det(q) := \det(C)$ die *Determinante* von q .
- Zwei quadratische Formen q_A und q_B heißen *äquivalent*, falls ein $S \in \text{GL}(n, \mathbb{Z})$ mit $S^t A S = B$ existiert.

Bemerkung 20.46.

- (a) Die Symmetrie von C ist unwesentlich, denn für eine beliebige Matrix $A \in \mathbb{R}^{n \times n}$ ist $C := \frac{1}{2}(A + A^t)$ symmetrisch mit $q_A = q_C$. Ist A ganzzahlig, so muss C jedoch nicht ganzzahlig sein. Manche Autoren definieren ganze quadratische Formen daher durch die Bedingung $q_C(x) \in \mathbb{Z}$ für alle $x \in \mathbb{Z}^n$. Dies bedeutet $c_{ii} \in \mathbb{Z}$ und $2c_{ij} \in \mathbb{Z}$ für alle $i \neq j$.

- (b) Ist q ausgeartet, so kann man durch Übergang zu einer äquivalenten Form eine Variable eliminieren und den Rang reduzieren. Wir sind hauptsächlich an positiven quadratischen Formen interessiert (vgl. Aufgabe III.30). Ist q negativ, so kann man zur positiven Form $-q$ wechseln und Resultate übertragen.
- (c) Jedes Gitter Δ mit Gram-Matrix G definiert eine positive quadratische Form q_G mit $|x|^2 = q_G(x_1, \dots, x_k)$ für alle $x \in \Delta$, wobei x_1, \dots, x_k die Koeffizienten bzgl. der gewählten Basis sind. Ein Basiswechsel von Δ entspricht nach Lemma 20.7 dem Übergang zu einer äquivalenten quadratischen Form. Umgekehrt haben wir in Beispiel 20.4 gesehen, wie man aus einer positiv definiten Matrix ein Gitter konstruiert. Auf diese Weise kann man zwischen Gittern und positiven quadratischen Formen hin- und herwechseln. Allerdings gibt es viele Gitter mit der gleichen Gram-Matrix und beliebigem Rang. Man beachte, dass $\min q_G = (\min \Delta)^2$ und $\det(q) = \text{disc}(\Delta)^2$ gilt. Die Hermite-Schranke hat deshalb die Form

$$\min q \leq \left(\frac{4}{3}\right)^{(k-1)/2} \sqrt[k]{\det(q)}$$

- (d) Äquivalente quadratische Formen q_A und q_B haben die gleiche Determinante und nehmen die gleichen Werte an. Insbesondere ist $\min q_A = \min q_B$, falls die Formen positiv sind. So wie wir eine Gitterbasis mit dem LLL-Algorithmus reduziert haben, werden wir eine quadratische Form durch eine möglichst „einfache“ äquivalente Form ersetzen.
- (e) Wie bei Gittern gibt es für ein vorgegebenes $d > 0$ nur endlich viele $x \in \mathbb{Z}^n$ mit $q(x) \leq d$ (Bemerkung 20.11). Außerdem existiert nach Bemerkung 20.33 eine äquivalente Form q_C mit $c_{11} = \min q$.

Satz 20.47. *Jede ganze quadratische Form q ist zu einer Form q_C mit*

$$C = \begin{pmatrix} * & * & & 0 \\ * & \ddots & \ddots & \\ & \ddots & \ddots & * \\ 0 & & * & * \end{pmatrix} \geq 0$$

äquivalent.

Beweis. Der Beweis verläuft ähnlich wie in Satz 20.24. Sei $q = q_A$. Sei (a_1, \dots, a_n) die erste Zeile/Spalte von A und $b_2, \dots, b_n \in \mathbb{Z}$ mit

$$d := \text{ggT}(a_2, \dots, a_n) = a_2 b_2 + \dots + a_n b_n.$$

Nach Lemma 20.19 existiert ein $U \in \text{GL}(n-1, \mathbb{Z})$ mit erster Spalte (b_2, \dots, b_n) . Sei $S := \text{diag}(1, U) \in \text{GL}(n, \mathbb{Z})$. Dann hat AS an Position $(1, 2)$ den Eintrag d . Wiederholt man diesen Schritt, so teilt a_2 schließlich jede der Zahlen a_2, \dots, a_n . Wie üblich erreicht man $a_i = 0$ für $i = 3, \dots, n$. Durch die Multiplikation von S^t von links an AS wird die erste Zeile nicht verändert. Da $S^t AS$ symmetrisch ist, hat die erste Spalte von A nun die gleiche Gestalt. Man verfährt mit den Zeilen $2, \dots, n-1$ analog. Am Ende hat man A in die gesuchte Matrix C überführt. \square

Beispiel 20.48.

$$\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \sim \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & -1 \\ 1 & 1 & 1 \end{pmatrix} \sim \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \sim \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}$$

Bemerkung 20.49.

- (a) Satz 20.47 eignet sich nicht, um Äquivalenzklassen quadratischer Formen zu zählen, denn unendlich viele solcher *Tridiagonalmatrizen* sind zueinander äquivalent. Für *binäre* quadratische Formen (d. h. $n = 2$) liefert Satz 20.47 lediglich eine Vorzeichen-Reduktion. Diesen Fall behandeln wir in Satz 20.55.
- (b) Eine ganze positive quadratische Form q heißt *unimodular*, falls $\det(q) = 1$. Sei q_C unimodular mit $n \leq 5$. Wegen $(4/3)^2 < 2$ ist $\min q = 1$ nach Hermite. Wie in Satz 20.32 konstruiert man eine äquivalente Form mit $C = \text{diag}(1, C_1)$ und $\det(C_1) = 1$. Man kann nun das gleiche Argument mit q_{C_1} wiederholen. Schließlich ist q zu q_{1_n} äquivalent. Insbesondere ist q bis auf Äquivalenz die einzige unimodulare quadratische Form vom Rang n . Nach einem Satz von MORDELL gilt dies auch für $n \in \{6, 7\}$. Für $n = 8$ bestimmt das in Bemerkung 20.11 eingeführte E_8 -Gitter eine unimodulare Form q mit $\min q = 2$. Insbesondere ist q nicht zu q_{1_8} äquivalent.

Satz 20.50. *Für alle $n \in \mathbb{N}$ und $d \in \mathbb{R}$ existieren bis auf Äquivalenz nur endlich viele positive quadratische Formen q mit Rang n und $\det(q) \leq d$.*

Beweis. Sei q eine quadratische Form vom Rang n mit $\det(q) \leq d$. Sei Δ ein entsprechendes Gitter (Bemerkung 20.46). Sei v_1, \dots, v_n eine δ -reduzierte Basis von Δ (zum Beispiel für $\delta = 3/4$). Durch Übergang zu einer äquivalenten quadratischen Form kann man q_C mit $c_{ii} = |v_i|^2$ für $i = 1, \dots, n$ annehmen. Nach Satz 20.41 ist $c_{11} \dots c_{nn}$ durch eine Funktion in d beschränkt. Da C positiv definit ist, gilt außerdem $c_{ii}c_{jj} - c_{ij}^2 > 0$, d. h. $|c_{ij}| < \sqrt{c_{ii}c_{jj}}$ für $i \neq j$. Daher gibt es nur endlich viele Möglichkeiten für C . □

Definition 20.51. Eine positive quadratische Form q mit Gram-Matrix $C = (c_{ij}) \in \mathbb{R}^{n \times n}$ heißt (*Minkowski*)-*reduziert*, falls für $i = 1, \dots, n$ gilt:

- $c_{ii} \leq q(x)$ für alle $x \in \mathbb{Z}^n$ mit $\text{ggT}(x_i, x_{i+1}, \dots, x_n) = 1$,
- $c_{i,i+1} \geq 0$ falls $i < n$.

Satz 20.52 (MINKOWSKI). *Jede positive quadratische Form ist zu einer reduzierten Form äquivalent.*

Beweis. Wir reduzieren eine positive quadratische Form q mit Gram-Matrix C schrittweise. Sei $1 \leq i \leq n$. Unter den Vektoren $x \in \mathbb{Z}^n$ mit $\text{ggT}(x_i, \dots, x_n) = 1$ können wir nach Bemerkung 20.46 einen mit minimalem $q(x)$ wählen. Nach Lemma 20.19 existiert ein $T \in \text{GL}(n - i + 1, \mathbb{Z})$ mit erster Spalte (x_i, \dots, x_n) . Dann ist

$$S := \begin{pmatrix} 1 & & x_1 & 0 \\ & \ddots & \vdots & \vdots \\ & & 1 & x_{i-1} & 0 \\ & & & T & \end{pmatrix} \in \text{GL}(n, \mathbb{Z}).$$

Ersetzt man C durch $S^t C S$, so bleiben $c_{11}, \dots, c_{i-1, i-1}$ unverändert, während c_{ii} jetzt die erste Minkowski-Bedingung erfüllt. Auf diese Weise erreicht man, dass die Bedingung für c_{11}, \dots, c_{nn} gilt. Die Vorzeichen von $c_{i,i+1}$ lassen sich durch $S = \text{diag}(\epsilon_1, \dots, \epsilon_n)$ mit $\epsilon_1, \dots, \epsilon_n \in \{\pm 1\}$ verändern (c_{11}, \dots, c_{nn} bleiben unverändert). □

Beispiel 20.53. Sei q die quadratische Form zu dem Gitter $\Delta \subseteq \mathbb{R}^5$ aus Beispiel 20.43. Die Gram-Matrix

$$\begin{pmatrix} 7 & 0 & 3 & -1 & 3 \\ 0 & 9 & 4 & 0 & 2 \\ 3 & 4 & 10 & 1 & -2 \\ -1 & 0 & 1 & 10 & 3 \\ 3 & 2 & -2 & 3 & 13 \end{pmatrix}$$

definiert eine äquivalente reduzierte Form. Im Vergleich zur δ -Reduktion sind die Diagonalelemente noch kleiner geworden.

Bemerkung 20.54.

- (a) Sei q_C reduziert. Sei $x \in \mathbb{Z}^n \setminus \{0\}$, $g := \text{ggT}(x_1, \dots, x_n)$ und $y_i := x_i/g$. Dann gilt $c_{11} \leq q(y_1, \dots, y_n) = g^2 q(x)$. Dies zeigt $c_{11} = \min q$. Außerdem ist $c_{11} \leq q(e_2) = c_{22} \leq q(e_3) = c_{33} \leq \dots \leq c_{nn}$. Für $i < j$ gilt außerdem

$$c_{jj} \leq q(e_i \pm e_j) = c_{ii} \pm 2c_{ij} + c_{jj},$$

d. h. $2|c_{ij}| \leq c_{ii}$.

- (b) Man kann zeigen, dass q_C bereits dann reduziert ist, wenn endlich viele Ungleichungen der Form $c_{ii} \leq q(x)$ gelten (zusammen mit $c_{i,i+1} \geq 0$). Allerdings wächst die Zahl dieser Ungleichungen exponentiell mit n . Für $n \geq 10$ ist es sehr aufwendig eine gegebene quadratische Form zu reduzieren. Die Reduktion nach KORKINE-ZOLOTAREV liefert einen Kompromiss zwischen der Effizienz des LLL-Algorithmus und der Güte der Minkowski-Reduktion. Darauf gehen wir nicht im Detail ein.
- (c) Die Diagonaleinträge c_{ii} einer reduzierten quadratischen Form q_C sind über die Minkowski-Bedingung eindeutig bestimmt. Außerdem existieren nur endlich viele Vektoren $x \in \mathbb{Z}^n$ mit $q_C(x) = c_{ii}$. Daher existieren nur endlich viele Matrizen $S \in \text{GL}(n, \mathbb{Z})$, sodass auch die äquivalente Form $q_{S^t C S}$ reduziert ist. Dies zeigt, dass jede positive quadratische Form zu höchstens endlich vielen reduzierten Formen äquivalent ist. Unter diesen kann man einen kanonischen Repräsentanten wählen, indem man eine Ordnung auf den Matrixeinträgen festlegt (z.B. unter allen äquivalenten Formen sei c_{12} minimal etc.). Für $n \leq 2$ gibt es generell nur eine reduzierte Form pro Äquivalenzklasse.

Satz 20.55 (GAUSS). Eine positive quadratische Form $q = q_C$ ist genau dann reduziert, wenn $0 \leq 2c_{12} \leq c_{11} \leq c_{22}$. Ggf. ist $c_{11}c_{22} \leq \frac{4}{3} \det(q)$. Ist q_D reduziert und zu q äquivalent, so gilt $C = D$.

Beweis. Sei $a := c_{11}$, $b := c_{12}$ und $c := c_{22}$. Ist q reduziert, so gilt $0 \leq 2b \leq a \leq c$ nach Bemerkung 20.54. Sei umgekehrt $0 \leq 2b \leq a \leq c$ gegeben. Für $(x, y) \in \mathbb{Z}^2 \setminus \{0\}$ gilt

$$\begin{aligned} q(x, y) &= ax^2 + 2bxy + cy^2 \geq a(x^2 + y^2) + 2bxy \\ &\geq a(x^2 + y^2) - a|xy| = a(|x| - |y|)^2 + a|xy| \geq a \end{aligned} \tag{20.2}$$

mit Gleichheit, falls $(x, y) = (1, 0)$. Daher ist $a = \min q \leq q(x_1, x_2)$ für $\text{ggT}(x_1, x_2) = 1$. Außerdem ist $c = q(0, 1) \leq q(x_1, x_2)$ für $\text{ggT}(x_2) = 1$. Also ist q reduziert.

Für die Eindeutigkeit sei $S = \begin{pmatrix} x & u \\ y & v \end{pmatrix} \in \text{GL}(2, \mathbb{Z})$ mit $S^t C S = \begin{pmatrix} a & b' \\ b' & c' \end{pmatrix}$ und $0 \leq 2b' \leq a \leq c'$ (beachte $a = \min q$). O. B. d. A. sei $c' \leq c$. Für (x, y) gilt Gleichheit in (20.2). Dafür gibt es zwei Fälle:

Fall 1: $y \neq 0$.

Aus (20.2) folgt $c = a \leq c' \leq c$. Also ist $c = c'$. Aus $a^2 - b^2 = \det(C) = a^2 - (b')^2$ und $b' \geq 0$ erhält man $b = b'$.

Fall 2: $(x, y) = (\pm 1, 0)$.

Indem man S durch $-S$ ersetzt, kann man $x = 1$ annehmen. Dann ist $v = \det(S) = \pm 1$. Aus

$$ua - \frac{a}{2} \leq ua \pm b = b' \leq \frac{a}{2}$$

folgt $u \leq 1$. Im Fall $u < 0$ wäre $b' < 0$. Für $u = 0$ ist $S = 1_2$. Sei also $u = 1$. Dann ist $b = a/2 = b'$. Aus $ac - b^2 = \det(C) = ac' - b^2$ folgt $c = c'$. \square

Bemerkung 20.56.

(a) Die Schranke $c_{11}c_{22} \leq \frac{4}{3} \det(q)$ entspricht der Abschätzung aus Satz 20.41 für $\delta \rightarrow 1$.

(b) Sei Δ ein Gitter mit δ -reduzierter Basis v_1, v_2 und Gram-Matrix $C = (c_{ij})$. Dann gilt

$$\frac{|c_{12}|}{c_{11}} = \frac{|[v_2, v_1]|}{|v_1|^2} = |\mu_{21}| \leq \frac{1}{2},$$

d. h. $2|c_{12}| \leq c_{11}$. Die (einzige) Lovász-Bedingung lautet

$$(\delta - \mu_{21}^2)|v_1|^2 \leq |b_2|^2 = |v_2 - \mu_{21}v_1|^2 = |v_2|^2 - 2\mu_{21}[v_1, v_2] + \mu_{21}^2|v_1|^2 = |v_2|^2 - \mu_{21}^2|v_1|^2,$$

d. h. $\delta c_{11} \geq c_{22}$. Für $\delta = 1$ ist q_C reduziert, wenn man das Vorzeichen von c_{12} ignoriert.

(c) Der folgende Algorithmus reduziert eine positive binäre quadratische Form q_C :

(1) Falls $c_{11} > c_{22}$, tausche c_{11} und c_{22} .

(2) Sei $\lambda := \lfloor c_{12}/c_{11} \rfloor$. Subtrahiere das λ -fache der ersten Zeile/Spalte von der zweiten Zeile/Spalte von C . Anschließend gilt $2|c_{12}| \leq c_{11}$.

(3) Gilt $c_{11} > c_{22}$, so gehe zu (1).

(4) Ersetze c_{12} durch $|c_{12}|$.

(d) Die Äquivalenzklassen der positiven quadratischen Formen q_C mit Rang 2 und $\det(q) \leq 10$ sind durch folgende Parameter gegeben:

$\det(q)$	1	2	3	3	4	4	5	5	6	6	7	7	8	8	8	9	9	9	10	10
c_{11}	1	1	1	2	1	2	2	1	2	1	2	1	3	2	1	3	2	1	2	1
c_{12}	0	0	0	1	0	0	1	0	0	0	1	0	1	0	0	0	1	0	0	0
c_{22}	1	2	3	2	4	2	3	5	3	6	4	7	3	4	8	3	5	9	5	10

Es gibt solche Tabellen für $n \leq 5$ im Internet (vgl. Aufgabe III.32).⁷

(e) Für $n \geq 3$ existieren äquivalente positive reduzierte quadratische Formen mit unterschiedlicher Gram-Matrix. Betrachten wir dazu $q_k := q_{C_k}$ mit

$$C_k := \begin{pmatrix} 2 & 1 & k \\ 1 & 2 & 1 \\ k & 1 & 2 \end{pmatrix}$$

für $k \in \{0, 1\}$. Nach Beispiel 20.48 sind q_0 und q_1 äquivalent. Nach Beispiel 12.35 ist q_0 positiv und $\min q_0 = 2$. Wegen $c_{ii} = 2 = \min q_k$ für $i = 1, 2, 3$ und $c_{12} = c_{23} = 1$ sind q_0 und q_1 reduziert.

⁷<https://www.math.rwth-aachen.de/homes/Gabriele.Nebe/LATTICES/>

20.6 Sukzessive Minima

Definition 20.57. Sei $q: \mathbb{Z}^n \rightarrow \mathbb{R}$ eine positive quadratische Form. In Anlehnung an Courant-Fischer definieren wir die *sukzessiven Minima*

$$\mu_i(q) := \min_{\substack{s_1, \dots, s_i \in \mathbb{Z} \\ \text{linear unabhängig}}} \max\{q(s_1), \dots, q(s_i)\}.$$

Bemerkung 20.58. Offenbar ist $\min q = \mu_1(q) \leq \dots \leq \mu_n(q)$. Ist q_C reduziert, so gilt

$$\mu_i(q) \leq \max\{q(e_1), \dots, q(e_i)\} = c_{ii}$$

für $i = 1, \dots, n$. Im Allgemeinen gilt aber nicht unbedingt Gleichheit.

Beispiel 20.59.

(a) Sei

$$C = \begin{pmatrix} 1 & \cdot & \cdot & \cdot & 1/2 \\ \cdot & 1 & \cdot & \cdot & 1/2 \\ \cdot & \cdot & 1 & \cdot & 1/2 \\ \cdot & \cdot & \cdot & 1 & 1/2 \\ 1/2 & 1/2 & 1/2 & 1/2 & 5/4 \end{pmatrix}.$$

Wegen

$$q(x) = x_1^2 + \dots + x_4^2 + \frac{5}{4}x_5^2 + x_1x_5 + \dots + x_4x_5 = \sum_{i=1}^4 \left(x_i + \frac{1}{2}x_5\right)^2 + \frac{1}{4}x_5^2 \geq 1$$

für $x \in \mathbb{Z}^n \setminus \{0\}$ ist q positiv und reduziert ist. Andererseits ist

$$B := \{e_1, \dots, e_4, (-1, -1, -1, -1, 2)\}$$

linear unabhängig mit $q(b) = 1$ für $b \in B$. Dies zeigt $\mu_5(q) = 1 < \frac{5}{4} = c_{55}$.

(b) Die sukzessiven Minima der Form q aus Beispiel 20.53 sind 7, 9, 10, 10, 11. Auch hier gilt $\mu_5(q) = 11 < 13 = c_{55}$.

Lemma 20.60. Für jede positive quadratische Form q vom Rang n existieren linear unabhängige Vektoren $s_1, \dots, s_n \in \mathbb{Z}^n$ mit $q(s_i) = \mu_i(q)$ für $i = 1, \dots, n$.

Beweis. Offenbar existiert $s_1 \in \mathbb{Z}^n$ mit $q(s_1) = \mu_1(q)$. Seien s_1, \dots, s_k mit $q(s_i) = \mu_i(q)$ für $i = 1, \dots, k$ bereits gewählt. Nach Definition existieren linear unabhängige t_1, \dots, t_{k+1} mit $q(t_i) \leq \mu_{k+1}(q)$ für $i = 1, \dots, k+1$. Mindestens einer der t_i , sagen wir t_1 muss linear unabhängig zu s_1, \dots, s_k sein. Sei $i \leq k+1$ minimal mit $\mu_i(q) = \mu_{k+1}(q)$. Aus

$$\mu_i(q) \leq \max\{q(s_1), \dots, q(s_{i-1}), q(t_1)\} \leq \mu_{k+1}(q)$$

folgt $q(t_1) = \mu_{k+1}(q)$. Wir können daher $s_{k+1} := t_1$ setzen. Die Behauptung folgt nun induktiv. \square

Bemerkung 20.61. Der folgende Satz verbessert Hermites Schranke (Bemerkung 20.46).

Satz 20.62 (MINKOWSKI). Für jede positive quadratische Form q vom Rang n gilt

$$\mu_1(q) \cdots \mu_n(q) \leq \left(\frac{4}{3}\right)^{n(n-1)/2} \det(q).$$

Beweis. Sei $S \in \mathbb{Z}^{n \times n}$ mit Spalten s_1, \dots, s_n wie in Lemma 20.60 (Achtung: S liegt nicht unbedingt in $\text{GL}(n, \mathbb{Z})$). Sei $q = q_C$ und $S^t C S = R^t R$ die Cholesky-Zerlegung von $S^t C S$. Sei

$$D := \text{diag}(\mu_1(q), \dots, \mu_n(q)), \quad C_1 := S^{-t} R^t D^{-1} R S^{-1}$$

und $q_1 := q_{C_1}$. Dann gilt

$$\det(q_1) = \det(D)^{-1} \det(S^{-t} R^t R S^{-1}) = \frac{\det(q)}{\mu_1(q) \cdots \mu_n(q)}.$$

Sei $x \in \mathbb{Z}^{n \times 1} \setminus \{0\}$, $y := S^{-1}x$ und $z := Ry$. Sei $k := \max\{1 \leq i \leq n : y_i \neq 0\}$. Dann ist $x = Sy = y_1 s_1 + \dots + y_k s_k$ linear unabhängig zu s_1, \dots, s_{k-1} . Insbesondere ist $q(x) \geq \mu_k(q)$. Da R eine obere Dreiecksmatrix ist, gilt außerdem $z_{k+1} = \dots = z_n = 0$. Daher ist

$$q_1(x) = y^t R^t D^{-1} R y = z^t D^{-1} z = \frac{z_1^2}{\mu_1(q)} + \dots + \frac{z_k^2}{\mu_k(q)} \geq \frac{1}{\mu_k(q)} |z|^2 = \frac{1}{\mu_k(q)} y^t R^t R y = \frac{1}{\mu_k(q)} q(x) \geq 1.$$

Insbesondere ist q_1 positiv und

$$1 \leq (\min q_1)^n \leq \left(\frac{4}{3}\right)^{n(n-1)/2} \det(q_1) = \left(\frac{4}{3}\right)^{n(n-1)/2} \frac{\det(q)}{\mu_1(q) \cdots \mu_n(q)}$$

nach Hermite (Bemerkung 20.46). □

Bemerkung 20.63.

- (a) Der Beweis zeigt, dass man $(4/3)^{n(n-1)/2}$ in Minkowskis Schranke durch die n -te Hermite-Konstante γ_n ersetzen darf (Bemerkung 20.35).
- (b) Für eine δ -reduzierte Gitterbasis v_1, \dots, v_n gilt nach Satz 20.41 mit $\delta \rightarrow 1$ asymptotisch die gleiche Schranke $|v_1|^2 \cdots |v_n|^2 \leq \left(\frac{4}{3}\right)^{n(n-1)/2} \det(q)$.
- (c) Eine ganze positive quadratische Form q heißt *universell*, falls $q: \mathbb{Z}^n \rightarrow \mathbb{N}_0$ surjektiv ist. Ein Satz von Lagrange aus der Zahlentheorie besagt, dass die quadratische Form

$$q(x) = x_1^2 + x_2^2 + x_3^2 + x_4^2$$

universell ist, d. h. jede natürliche Zahl ist die Summe von vier Quadraten.⁸ Der 15-Satz von CONWAY besagt, dass eine ganze positive quadratische Form bereits dann universell ist, wenn sie die Werte 1, 2, 3, 5, 6, 7, 10, 14, und 15 annimmt.

⁸siehe Zahlentheorie-Skript

Aufgaben

Aufgabe III.1 (Binäre Exponentiation). Sei K ein Körper und $A \in K^{n \times n}$.

- (a) Entwerfen Sie einen Algorithmus zur effizienten Berechnung von A^k durch iteriertes Quadrieren.
- (b) Zeigen Sie, dass man auf diese Weise mit höchstens $2 \lfloor \log_2(k) \rfloor$ Matrixmultiplikationen auskommt (dabei ist $\lfloor \log_2(k) \rfloor$ die größte ganze Zahl z mit $2^z \leq k$).

Aufgabe III.2 (HERON-Verfahren). Sei $a \in \mathbb{R}$ mit $a \geq 1$. Zeigen Sie, dass die Folge

$$x_0 := 1, \quad x_{n+1} := \frac{1}{2} \left(x_n + \frac{a}{x_n} \right)$$

quadratisch gegen \sqrt{a} konvergiert, d. h. $\lim_{n \rightarrow \infty} x_n = \sqrt{a}$ und $|x_{n+1} - \sqrt{a}| \leq \frac{1}{2} |x_n - \sqrt{a}|^2$ für $n \in \mathbb{N}$. Begründen Sie, warum sich die Anzahl der korrekten Nachkommastellen von x_n in jedem Iterationsschritt verdoppelt.

Aufgabe III.3. Sei K ein Körper und $A \in K^{n \times n}$ vom Rang k . Zeigen Sie, dass $B, C \in K^{n \times k}$ mit $A = BC^t$ existieren. Wie kann man damit Berechnungen beschleunigen und Speicherbedarf reduzieren, wenn $k \ll n$.

Aufgabe III.4. Beweisen oder widerlegen Sie: Ähnliche Matrizen in $\mathbb{C}^{n \times n}$ haben die gleiche Konditionszahl.

Aufgabe III.5. Zeigen Sie, dass $A \in \mathbb{C}^{n \times m}$ den gleichen Rang wie die Pseudoinverse A^+ hat.

Aufgabe III.6. Beweisen oder widerlegen Sie: $(AB)^+ = B^+A^+$ für alle $A \in \mathbb{C}^{n \times m}$ und $B \in \mathbb{C}^{m \times k}$.

Aufgabe III.7. Zeigen Sie: Hat das System $Ax = b$ mit $A \in \mathbb{C}^{n \times m}$ mindestens eine Lösung, so haben alle Lösungen die Form $A^+b + (1_m - A^+A)y$ mit $y \in \mathbb{C}^m$.

Aufgabe III.8. Bestimmen Sie die größte (endliche) Zahl, die man mit dem Datentyp `float` (Bemerkung 17.29) darstellen kann.

Aufgabe III.9. Sei $A \in K^{n \times n}$ mit Hauptminoren $\det(A_k) \neq 0$ für $k = 1, \dots, n-1$ (siehe Bemerkung 12.43). Zeigen Sie, dass eine eindeutige LR-Zerlegung in der Form $A = LR$ existiert (siehe Satz 17.35).

Aufgabe III.10 (BRUHAT-Zerlegung). Zeigen Sie, dass für jede Matrix $A \in \text{GL}(n, K)$ Permutationsmatrizen P, Q , untere Dreiecksmatrizen L_1, L_2 und obere Dreiecksmatrizen R_1, R_2 mit $A = L_1 P L_2 = R_1 Q R_2$ existieren.

Aufgabe III.11 (Polardarstellung). Zeigen Sie, dass für jede Matrix $A \in \mathbb{C}^{n \times n}$ eindeutig bestimmte Matrizen $U \in U(n, \mathbb{C})$ und $P \in \mathbb{C}^{n \times n}$ mit $A = UP$ existieren, wobei P positiv semidefinit ist.

Aufgabe III.12 (HADAMARD-Ungleichung). Sei $A \in \mathbb{C}^{n \times n}$ mit Spalten s_1, \dots, s_n . Zeigen Sie $|\det(A)| \leq |s_1| \dots |s_n|$ mit Gleichheit genau dann, wenn s_1, \dots, s_n paarweise orthogonal sind.

Hinweis: QR-Zerlegung.

Aufgabe III.13. Seien $p, q > 1$ mit $\frac{1}{p} + \frac{1}{q} = 1$, $s, x, y \in \mathbb{R}_{\geq 0}$ und $v, w \in \mathbb{C}^n$. Zeigen Sie:

(a) (BERNOULLI-Ungleichung) $1 + sx \leq (1 + x)^s$ für $s \geq 1$.

Hinweis: Man braucht die Stetigkeit der Potenzfunktion.

(b) (YOUNG-Ungleichung) $xy \leq \frac{x^p}{p} + \frac{y^q}{q}$.

(c) (HÖLDER-Ungleichung) $\sum_{i=1}^n |v_i w_i| \leq \|v\|_p \|w\|_q$.

Hinweis: Dies verallgemeinert die Cauchy-Schwarz-Ungleichung.

(d) (MINKOWSKI-Ungleichung) $\|v + w\|_p \leq \|v\|_p + \|w\|_p$.

(e) Die p -Norm ist tatsächlich eine Norm auf \mathbb{C}^n .

Aufgabe III.14. Sei $\|\cdot\|$ eine Norm, die auf $\mathbb{C}^{n \times 1}$ für alle n definiert ist. Zeigen Sie, dass

$$\|A\| = \max_{0 \neq x \in \mathbb{C}^{m \times 1}} \frac{\|Ax\|}{\|x\|} = \max_{\|x\|=1} \|Ax\|$$

eine Matrixnorm auf $\mathbb{C}^{n \times m}$ definiert.

Aufgabe III.15. Formulieren und beweisen Sie das Orthonormalisierungsverfahren mit Householder-Transformationen und Givens-Rotationen aus Bemerkung 17.87 für $A \in GL(n, \mathbb{C})$.

Aufgabe III.16. Zeigen Sie $\lim_{k \rightarrow \infty} (1_n + \frac{1}{k}A)^k = \exp(A)$ für alle $A \in \mathbb{C}^{n \times n}$.

Aufgabe III.17. Zeigen Sie, dass $\exp(A)$ positiv definit ist, falls $A \in \mathbb{C}^{n \times n}$ hermitesch ist.

Aufgabe III.18. Zeigen Sie:

(a) Die Anzahl der Permutationen in S_n mit genau k Zyklen (einschließlich 1-Zyklen) beträgt

$$\left[\begin{matrix} n \\ k \end{matrix} \right] := \frac{n!}{k!} \sum_{\substack{1 \leq l_1, \dots, l_k \leq n \\ l_1 + \dots + l_k = n}} \frac{1}{l_1 \dots l_k}.$$

Hinweis: Man nennt $\left[\begin{matrix} n \\ k \end{matrix} \right]$ Stirling-Zahl der ersten Art.

(b) Für $n \geq 2$ gibt es genauso viele Permutationen in S_n mit einer geraden Zyklenzahl wie mit einer ungeraden Zyklenzahl, d. h. $\sum_{k=1}^n (-1)^k \left[\begin{matrix} n \\ k \end{matrix} \right] = 0$.

(c) Sei $N \in \mathbb{C}^{n \times n}$ nilpotent und

$$A := - \sum_{k=1}^n \frac{1}{k} N^k \in \mathbb{C}^{n \times n}.$$

Dann gilt $\exp(A) = 1_n - N$.

Aufgabe III.19. Zeigen Sie, dass für $A \in \text{SL}(n, \mathbb{C})$ ein $B \in \mathbb{C}^{n \times n}$ mit $\exp(B) = A$ und $\text{tr}(B) = 0$ existiert.

Aufgabe III.20. Seien $A, B \in \mathbb{R}^{n \times n}$ stochastische Matrizen. Zeigen Sie, dass AB stochastisch ist.

Aufgabe III.21. Sind $A \in \mathbb{R}^{n \times n}$ und A^{-1} stochastisch, so ist A eine Permutationsmatrix.

Aufgabe III.22. Sei $0 < p, q < 1$ und $A := \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}$. Berechnen Sie $\lim_{k \rightarrow \infty} A^k$.

Aufgabe III.23. Sei $A \in \mathbb{R}^{n \times n}$ nicht-negativ und unzerlegbar. Zeigen Sie $(1_n + A)^{n-1} > 0$.

Aufgabe III.24. Sei V ein K -Vektorraum der Dimension n . Man nennt $A \subseteq V$ *affin abhängig*, falls paarweise verschiedene Elemente $a_1, \dots, a_k \in A$ und $\lambda_1, \dots, \lambda_k \in K^\times$ mit $\lambda_1 a_1 + \dots + \lambda_k a_k = 0$ und $\lambda_1 + \dots + \lambda_k = 0$ existieren. Anderenfalls heißt A *affin unabhängig*. Zeigen Sie:

- (a) Es existiert eine affin unabhängige Menge mit $n + 1$ Elementen.
- (b) Jede Menge aus mindestens $n + 2$ Elementen ist affin abhängig.
- (c) Genau dann ist $A \subseteq V$ affin abhängig, wenn $\{v - w : w \in A \setminus \{v\}\}$ für ein $v \in A$ linear abhängig ist.

Aufgabe III.25. Sei V ein K -Vektorraum. Eine *affine* Linearkombination von $v_1, \dots, v_k \in V$ ist eine Summe der Form $\lambda_1 v_1 + \dots + \lambda_k v_k$ mit $\lambda_1, \dots, \lambda_k \in K$ und $\lambda_1 + \dots + \lambda_k = 1$.⁹ Die Menge aller affinen Linearkombinationen von Elementen aus $\Delta \subseteq K^n$ nennt man die *affine Hülle* $\text{aff}(\Delta)$.

- (a) Beschreiben Sie die affine Hülle zweier Punkte in \mathbb{R}^n .
- (b) Zeigen Sie, dass eine Menge $\Delta \subseteq V$ genau dann unter affinen Linearkombinationen abgeschlossen ist (d. h. $\text{aff}(\Delta) = \Delta$), wenn ein $x \in \Delta$ existiert, sodass $\Delta - x$ ein Unterraum von V ist.

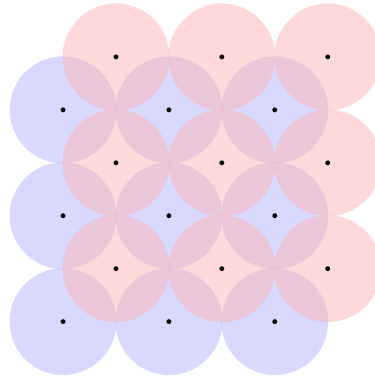
Aufgabe III.26. Seien $A \in \mathbb{R}^{n \times m}$ und $b \in \mathbb{R}^{n \times 1}$. Zeigen Sie folgende Varianten von Folgerung 19.13:

- (a) Genau dann existiert ein x mit $Ax \leq b$, wenn $y^t b \geq 0$ für alle $y \geq 0$ mit $y^t A = 0$ gilt.
Hinweis: Betrachten $(1_n, A, -A)$.
- (b) Genau dann existiert ein $x \geq 0$ mit $Ax \leq b$, wenn $y^t b \geq 0$ für alle $y \geq 0$ mit $y^t A \geq 0$ gilt.

⁹Im Gegensatz zu Konvexkombinationen sind die λ_i nicht beschränkt.

Aufgabe III.27.

- (a) Bestimmen Sie das Gitter $\Delta \subseteq \mathbb{R}^3$ zu der quadratischen Anordnung von Einheitskugeln (die Kugeln in jeder Ebene berühren Kugeln in der darunter liegenden Ebene, siehe Bemerkung 20.11):



- (b) Berechnen Sie $\min \Delta$ und die Dichte $\rho(\Delta)$.
 (c) Begründen Sie, warum diese Anordnung kein Gegenispiel zu Keplers Vermutung ist.

Aufgabe III.28. Sei $\Delta \subseteq \mathbb{R}^n$ ein Gitter mit vollem Rang und $\delta := \frac{1}{\sqrt{n}} \min \Delta$. Sei $F \subseteq \mathbb{R}^n$ die Fundamentalmasche von Δ (Beispiel 20.9). Sei $W := (-\delta/2, \delta/2)^n \subseteq \mathbb{R}^n$ der offene Würfel mit Mittelpunkt 0 und Seitenlänge δ . Zeigen Sie:

- (a) Für verschiedene $x, y \in \Delta$ ist $(x + W) \cap (y + W) = \emptyset$.
 (b) Es existieren $x_1, \dots, x_k \in \Delta$ mit $W \subseteq \bigcup_{i=1}^k (x_i + F)$.
 (c) Für das Volumen gilt

$$\delta^n = \text{vol}(W) = \text{vol}(F \cap ((W - x_1) \cup \dots \cup (W - x_k))) \leq \text{vol}(F) = \text{disc}(\Delta),$$

d. h. $\min \Delta \leq \sqrt{n} \sqrt[n]{\text{disc}(\Delta)}$.

- (d) Für das duale Gitter Δ^* gilt $(\min \Delta)(\min \Delta^*) \leq n$.
 (e) Für jede positive quadratische Form q vom Rang n gilt $\min q \leq n \det(q)$.

Aufgabe III.29. Sei q_C eine reduzierte positive quadratische Form vom Rang 2. Zeigen Sie für das sukzessive Minimum $\mu_2(q) = c_{22}$.

Aufgabe III.30. Sei q eine *indefinite* quadratische Form vom Rang 2, d. h. es existieren $x, y \in \mathbb{Z}^2$ mit $q(x) < 0 < q(y)$. Sei $d := -\det(q)$ keine Quadratzahl. Zeigen Sie:

- (a) $d > 0$.
 (b) q ist zu einer Form q_C mit $0 < c_{12} < \sqrt{d} < |c_{11}| + c_{12} \leq |c_{22}| + c_{12}$ äquivalent.
 (c) Bis auf Äquivalenz gibt es nur endlich viele ganze indefinite quadratische Formen vom Rang 2 und Determinante $-d$ (auch wenn d ein Quadrat ist).
 (d) Bestimmen Sie bis auf Äquivalenz alle ganzen indefiniten quadratischen Formen q vom Rang 2 mit $\det(q) \geq -4$.

Bemerkung: Im Gegensatz zu positiven binären quadratischen Formen gibt es für indefinite Formen keine kanonische Normalform.

Aufgabe III.31. Zeigen Sie, dass eine positive quadratische Form q_C vom Rang 3 genau dann reduziert ist, wenn gilt:

$$\begin{aligned} 0 \leq 2c_{12} \leq c_{11} \leq c_{22}, & & 2|c_{13}| \leq c_{11}, \\ 0 \leq 2c_{23} \leq c_{22} \leq c_{33}, & & 2(c_{12} + c_{23} - c_{13}) \leq c_{11} + c_{22}. \end{aligned}$$

Aufgabe III.32. Bestimmen Sie alle reduzierten ganzen quadratischen Formen vom Rang 3 mit Determinante ≤ 5 . Welche davon sind äquivalent?

Aufgabe III.33. Für $A \in K^{n \times m}$ und $B \in K^{s \times t}$ definieren wir das *Kronecker-Produkt*¹⁰

$$A \otimes B := \begin{pmatrix} a_{11}B & \cdots & a_{1m}B \\ \vdots & & \vdots \\ a_{n1}B & \cdots & a_{nm}B \end{pmatrix} \in K^{ns \times mt}$$

als Blockmatrix. Zeigen Sie für Matrizen A, B, C, D mit geeignetem Format und $\lambda \in K$:

- (a) $A \otimes (B \otimes C) = (A \otimes B) \otimes C$.
- (b) $(A \otimes B)^t = A^t \otimes B^t$.
- (c) $A \otimes (B + C) = A \otimes B + A \otimes C$.
- (d) $(A + B) \otimes C = A \otimes C + B \otimes C$.
- (e) $\lambda(A \otimes B) = (\lambda A) \otimes B = A \otimes (\lambda B)$.
- (f) $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$.
- (g) $\text{tr}(A \otimes B) = \text{tr}(A) \text{tr}(B)$.
- (h) $\text{rk}(A \otimes B) = \text{rk}(A) \text{rk}(B)$.
- (i) $\det(A \otimes B) = \det(A)^m \det(B)^n$ falls $A \in K^{n \times n}$ und $B \in K^{m \times m}$.

Aufgabe III.34. Sei $A \in \mathbb{R}^{n \times n}$ und $B \in \mathbb{R}^{m \times m}$. Zeigen Sie:

- (a) Sind A und B orthogonal, so auch $A \otimes B$.
- (b) Sind A und B positiv definit, so auch $A \otimes B$.
- (c) Sind A und B stochastisch, so auch $A \otimes B$.

Aufgabe III.35. Seien $A, B \in \mathbb{R}^{n \times n}$ positiv definit. Zeigen Sie $\min q_{A \otimes B} \leq (\min q_A)(\min q_B)$.

Bemerkung: KITAOKA hat bewiesen, dass für quadratische Formen mit Rang ≤ 43 Gleichheit gilt. Für Rang 292 existieren Beispiele für die strikte Ungleichung.

¹⁰auch *Tensorprodukt* genannt

Anhang

Konvexe Optimierung

Bemerkung A.64. Wir untersuchen in diesem Abschnitt Optimierungsprobleme mit einer konvexen (oder konkaven) Zielfunktion unter linearen Nebenbedingungen. Da in diesem Konzept der Simplex-Algorithmus nicht funktioniert, gibt es keinen Grund die Nebenbedingungen durch Einführen von Schlupfvariablen auf die Standardform zu bringen. Wir betrachten daher konvexe Mengen der Form

$$M := \{x \in \mathbb{R}^m : Ax \leq b\}$$

mit $A \in \mathbb{R}^{n \times m}$, $\text{rk}(A) = m < n$ und $b \in \mathbb{R}^n$ (Bemerkung 19.3). Wie üblich heißt $x \in M$ *Ecke*, falls $M \setminus \{x\}$ konvex ist. Für $I \subseteq \{1, \dots, n\}$ sei $A_I := (a_{ij} : i \in I, j = 1, \dots, m)$. Wir nennen I eine *Basismenge*, falls A_I invertierbar ist. Ggf. ist $|A| = m$.

Satz A.65. Sei $M := \{x \in \mathbb{R}^m : Ax \leq b\}$ mit $A \in \mathbb{R}^{n \times m}$ und $b \in \mathbb{R}^n$. Genau dann ist $x \in M$ eine Ecke, wenn eine Basismenge I mit $A_I x = b_I$ existiert.

Beweis. Sei I eine Basismenge mit $A_I x = b_I$. Seien $y, z \in M$ und $0 \leq \lambda \leq 1$ mit $x = \lambda y + (1 - \lambda)z$. Aus

$$b_I = \lambda A_I y + (1 - \lambda)A_I z \leq \lambda b_I + (1 - \lambda)b_I = b_I$$

folgt $A_I y = b_I$ und $A_I z = b_I$. Da A_I invertierbar ist, gilt $y = x = z$. Also ist x eine Ecke.

Ein umgekehrt x eine Ecke. Sei I die Menge der Indizes i mit $\sum_{j=1}^m a_{ij} x_j = b_i$. Nehmen wir $\text{rk}(A_I) < m$ an. Dann existiert ein $y \in \mathbb{R}^m \setminus \{0\}$ mit $A_I y = 0$. Für $i \notin I$ gilt $\sum_{j=1}^m a_{ij} x_j < b_i$. Daher existiert ein $\epsilon > 0$ mit $x \pm \epsilon y \in M$. Wegen $x = \frac{1}{2}(x + \epsilon y) + \frac{1}{2}(x - \epsilon y)$ kann x keine Ecke sein. Also ist $\text{rk}(A_I) = m$. Daher existiert eine Basismenge $J \subseteq I$ mit $A_J x = b_J$. \square

Definition A.66. Sei $M \subseteq \mathbb{R}^n$ konvex. Eine Funktion $f: M \rightarrow \mathbb{R}$ heißt

- *konvex*, falls

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$

für alle $x, y \in M$ und $0 < \lambda < 1$ gilt. Gilt im Fall $x \neq y$ die echte Ungleichung, so nennt man f *strikt konvex*.

- (*strikt*) *konkav*, falls $-f$ (*strikt*) konvex ist, d. h.

$$f(\lambda x + (1 - \lambda)y) \geq \lambda f(x) + (1 - \lambda)f(y).$$

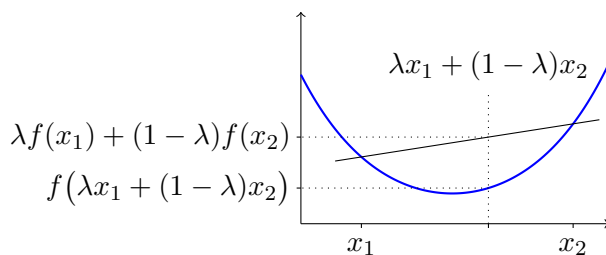
Bemerkung A.67.

(a) Jede lineare Funktion ist konvex und konkav.

(b) Ist $f: M \rightarrow \mathbb{R}$ konvex, so ist $N := \{(x, y) \in M \times \mathbb{R} : f(x) \leq y\} \subseteq \mathbb{R}^{n+1}$ konvex, denn für $f(x_i) \leq y_i$ und $0 \leq \lambda \leq 1$ gilt

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2) \leq \lambda y_1 + (1 - \lambda)y_2.$$

- (c) In der Analysis zeigt man, dass f genau dann (strikt) konvex ist, wenn die Hesse-Matrix der zweiten Ableitungen ($\frac{\partial^2 f}{\partial x_i \partial x_j}$) positiv semidefinit (definit) ist. Für $n = 1$ bedeutet dies $f''(x) \geq 0$ für alle $x \in M$, d. h. die Steigung des Graphen nimmt kontinuierlich zu.



Satz A.68. Sei $M = \{x \in \mathbb{R}^m : Ax \leq b\} \neq \emptyset$ konvex und $f: M \rightarrow \mathbb{R}$ konvex (bzw. konkav). Besitzt f ein Maximum (bzw. Minimum) auf M , so wird dieses an einer Ecke angenommen.

Beweis. Wir argumentieren wie im Beweis von Satz 19.24. Sei $x \in M$ ein Maximum von f . Sei $I(x) = I$ die Menge der Indizes i mit $\sum_{j=1}^m a_{ij}x_j = b_i$. Nehmen wir $\text{rk}(A_I) < m$ an. Wie im Beweis von Satz A.65 existieren $y \in \mathbb{R}^m$ und $\epsilon > 0$ mit $x_{\pm} := x \pm \epsilon y \in M$. Da f konvex ist, gilt

$$f(x) = f\left(\frac{1}{2}x_+ + \frac{1}{2}x_-\right) \leq \frac{1}{2}f(x_+) + \frac{1}{2}f(x_-) \leq \max\{f(x_+), f(x_-)\} \leq f(x)$$

und $f(x_+) = f(x) = f(x_-)$. Da wir stets annehmen, dass A vollen Rang hat, gilt $Ay \neq 0$. Wir können daher ϵ so wählen, dass $I(x_+) \supsetneq I$ oder $I(x_-) \supsetneq I$ gilt. Anschließend ersetzen wir x durch x_+ (bzw. x_-) und wiederholen das Argument. Nach endlich vielen Schritten erreicht man $\text{rk}(A_I) = m$. Dann ist x eine Ecke. Der Beweis für konkave Funktionen verläuft analog. \square

Bemerkung A.69. Ist f strikt konvex (bzw. konkav), so kann das Maximum (bzw. Minimum) nur an Ecken angenommen werden (dafür braucht man Satz A.65 nicht).

Satz A.70 (Ungleichung vom harmonischen, geometrischen und arithmetischen Mittel). Für alle positiven $x_1, \dots, x_n \in \mathbb{R}$ gilt

$$\frac{n}{\frac{1}{x_1} + \dots + \frac{1}{x_n}} \leq \sqrt[n]{x_1 \dots x_n} \leq \frac{x_1 + \dots + x_n}{n}$$

mit Gleichheit genau dann, wenn $x_1 = \dots = x_n$.

Beweis (CAUCHY). Die erste Ungleichung folgt aus der zweiten, indem man x_i durch $1/x_i$ ersetzt. Wir zeigen die zweite Ungleichung durch eine ungewöhnliche Induktion nach n . Für $n = 1$ gilt Gleichheit. Für $n = 2$ ist

$$\frac{(x_1 + x_2)^2}{4} - x_1x_2 = \frac{(x_1 - x_2)^2}{4} \geq 0$$

mit Gleichheit genau dann, wenn $x_1 = x_2$. Sei nun die Aussage für n erfüllt. Dann ist

$$x_1 \dots x_{2n} \leq \left(\sum_{i=1}^n \frac{x_i}{n} \sum_{i=n+1}^{2n} \frac{x_i}{n} \right)^n \leq \left(\sum_{i=1}^{2n} \frac{x_i}{2n} \right)^{2n}$$

mit Gleichheit genau dann, wenn $x_1 = \dots = x_{2n}$. Also gilt die Behauptung für $2n$. Sei nun $A := \sum_{i=1}^{n-1} \frac{x_i}{n-1}$. Dann ist

$$x_1 \dots x_{n-1} A \leq \left(\sum_{i=1}^{n-1} \frac{x_i}{n} + \frac{A}{n} \right)^n = \left(\frac{A(n-1)}{n} + \frac{A}{n} \right)^n = A^n$$

und $x_1 \dots x_{n-1} \leq A^{n-1}$ mit Gleichheit genau dann, wenn $x_1 = \dots = x_{n-1}$. Also gilt die Behauptung auch für $n-1$. \square

Satz A.71. Für jede positiv semidefinite Matrix $A \in \mathbb{C}^{n \times n}$ gilt $\sqrt[n]{\det(A)} \leq \operatorname{tr}(A)/n$ mit Gleichheit genau dann, wenn A eine Skalarmatrix ist.

Beweis. Nach Aufgabe II.18 hat A nicht-negative Eigenwerte $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. Nach Bemerkung 10.35 gilt

$$\sqrt[n]{\det(A)} = \sqrt[n]{\lambda_1 \dots \lambda_n} \leq \frac{1}{n}(\lambda_1 + \dots + \lambda_n) = \frac{1}{n} \operatorname{tr}(A)$$

mit Gleichheit genau dann, wenn $\lambda_1 = \dots = \lambda_n$. Da A nach dem Spektralsatz diagonalisierbar ist, muss A ggf. eine Skalarmatrix sein. \square

Lemma A.72. Für $x_1, \dots, x_n, y_1, \dots, y_n \geq 0$ gilt

$$\sqrt[n]{(x_1 + y_1) \dots (x_n + y_n)} \geq \sqrt[n]{x_1 \dots x_n} + \sqrt[n]{y_1 \dots y_n}$$

mit Gleichheit genau dann, wenn (x_1, \dots, x_n) und (y_1, \dots, y_n) linear abhängig sind.

Beweis. O. B. d. A. sei $x_i + y_i > 0$ für $i = 1, \dots, n$. Nach der Ungleichung zwischen dem arithmetischen und geometrischen Mittel gilt

$$1 = \frac{1}{n} \sum_{i=1}^n \frac{x_i}{x_i + y_i} + \frac{1}{n} \sum_{i=1}^n \frac{y_i}{x_i + y_i} \geq \sqrt[n]{\prod_{i=1}^n \frac{x_i}{x_i + y_i}} + \sqrt[n]{\prod_{i=1}^n \frac{y_i}{x_i + y_i}} = \frac{\sqrt[n]{x_1 \dots x_n} + \sqrt[n]{y_1 \dots y_n}}{\sqrt[n]{(x_1 + y_1) \dots (x_n + y_n)}}$$

mit Gleichheit genau dann, wenn $\frac{x_i}{x_i + y_i} = \frac{x_j}{x_j + y_j}$ und $\frac{y_i}{x_i + y_i} = \frac{y_j}{x_j + y_j}$ für alle i, j . Dies bedeutet $x_1 y_i = x_i y_1$ für $i = 1, \dots, n$. Ist x oder y der Nullvektor, so sind die Vektoren linear abhängig. Anderenfalls gilt o. B. d. A. $x_1 \neq 0$ und $y_i = \frac{y_1}{x_1} x_i$ für $i = 1, \dots, n$. Also sind x und y linear abhängig. \square

Lemma A.73 (Simultane Diagonalisierung). Sei $A \in \mathbb{C}^{n \times n}$ positiv definit und $B \in \mathbb{C}^{n \times n}$ positiv semidefinit. Dann existiert ein $S \in \operatorname{GL}(n, \mathbb{C})$ mit $S^* A S = 1_n$ und $S^* B S = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$.

Beweis. Nach Aufgabe II.18 existiert eine hermitesche Wurzel \sqrt{A} von A . Offenbar ist auch $C := \sqrt{A}^{-1} B \sqrt{A}^{-1}$ hermitesch. Nach dem Spektralsatz existiert ein $U \in \operatorname{U}(n, \mathbb{C})$ mit $U^* C U = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$. Die Behauptung gilt nun für $S := \sqrt{A}^{-1} U$. \square

Satz A.74 (MINKOWSKI-UNGLEICHUNG). Sei $A \in \mathbb{C}^{n \times n}$ positiv definit und $B \in \mathbb{C}^{n \times n}$ positiv semidefinit. Dann gilt

$$\sqrt[n]{\det(A+B)} \geq \sqrt[n]{\det(A)} + \sqrt[n]{\det(B)}$$

mit Gleichheit genau dann, wenn $B = \lambda A$ für ein $\lambda \geq 0$. Insbesondere ist $\det(A+B) \geq \det(A) + \det(B)$.

Beweis. Nach Lemma A.73 existiert ein $S \in \text{GL}(n, \mathbb{C})$ mit $S^*AS = 1_n$ und

$$D := S^*BS = \text{diag}(\lambda_1, \dots, \lambda_n).$$

Indem wir von links und rechts mit $\sqrt[n]{\det(S^*)}$ bzw. $\sqrt[n]{\det(S)}$ multiplizieren, können wir $A = 1_n$ und $B = D$ annehmen. Nach Voraussetzung ist $\lambda_1, \dots, \lambda_n \geq 0$. Aus Lemma A.72 folgt

$$\sqrt[n]{\det(1_n + D)} = \sqrt[n]{(1 + \lambda_1) \dots (1 + \lambda_n)} \geq 1 + \sqrt[n]{\lambda_1 \dots \lambda_n} = \sqrt[n]{\det(1_n)} + \sqrt[n]{\det(D)}$$

mit Gleichheit genau dann, wenn $\lambda := \lambda_1 = \dots = \lambda_n$. Für die Ausgangsmatrix bedeutet dies $B = S^{-*}DS^{-1} = \lambda A$.

Die zweite Behauptung folgt aus

$$\begin{aligned} \det(A + B) &= \sqrt[n]{\det(A + B)^n} \geq \left(\sqrt[n]{\det(A)} + \sqrt[n]{\det(B)} \right)^n \\ &= \sum_{k=1}^n \binom{n}{k} \sqrt[n]{\det(A)^k} \sqrt[n]{\det(B)^{n-k}} \geq \det(A) + \det(B). \end{aligned} \quad \square$$

Satz A.75. Seien $v \in \mathbb{R}^n$. Die Menge $\mathcal{M} \subseteq \mathbb{R}^{n \times n}$ aller positiv definiten Matrizen mit Hauptdiagonale v ist konvex. Die Abbildung $\mathcal{M} \rightarrow \mathbb{R}$, $A \mapsto \sqrt[n]{\det(A)}$ ist strikt konkav.

Beweis. Seien $A, B \in \mathcal{M}$ und $0 < \lambda < 1$. Offenbar hat $C := \lambda A + (1 - \lambda)B$ Hauptdiagonale v . Für $x \in \mathbb{R}^n \setminus \{0\}$ gilt

$$xCx^t = \lambda xAx^t + (1 - \lambda)xBx^t > 0.$$

Dies zeigt $C \in \mathcal{M}$. Nach der Minkowski-Ungleichung gilt

$$\sqrt[n]{\det(C)} \geq \sqrt[n]{\det(\lambda A)} + \sqrt[n]{\det((1 - \lambda)B)} = \lambda \sqrt[n]{\det(A)} + (1 - \lambda) \sqrt[n]{\det(B)}$$

mit Gleichheit genau dann, wenn $\lambda \mu A = (1 - \lambda)B$ für ein $\mu > 0$ gilt. Da A und B die gleichen Hauptdiagonalen haben, folgt $\lambda \mu = (1 - \lambda)$ und $A = B$. Also ist $A \mapsto \sqrt[n]{\det(A)}$ strikt konkav. \square

Satz A.76 (OPPENHEIM). Sei q_C eine reduzierte positive quadratische Form vom Rang 3. Dann gilt

$$c_{11}c_{22}c_{33} \leq c_{11}c_{22}c_{33} + \frac{1}{2}c_{11}c_{22}(c_{33} - c_{22}) + \frac{1}{2}c_{11}c_{33}(c_{22} - c_{11}) \leq 2 \det(q).$$

Gilt $c_{11}c_{22}c_{33} = 2 \det(q)$, so ist q äquivalent zu der quadratischen Form mit Gram-Matrix

$$\frac{c_{11}}{2} \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}.$$

Beweis (BARNES). Sei $\mathcal{M} \subseteq \mathbb{R}^{3 \times 3}$ die Menge der reduzierten positiv definiten Matrizen mit Hauptdiagonale (a, b, c) , wobei $a \leq b \leq c$. Für

$$C := \begin{pmatrix} a & x & z \\ x & b & y \\ z & y & c \end{pmatrix} \in \mathcal{M}$$

gilt $0 \leq x \leq a/2$, $0 \leq y \leq b/2$, $z \leq a/2$, $-z \leq a/2$ und $x + y - z \leq (a + b)/2$ nach Aufgabe III.31. Man kann daher \mathcal{M} mit der konvexen Menge $M \subseteq \mathbb{R}^3$ aller (x, y, z) mit

$$A = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -1 \\ 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \leq \frac{1}{2} \begin{pmatrix} a \\ b \\ 0 \\ a \\ a \\ a+b \end{pmatrix} =: v$$

identifizieren. Nach der Sarrus-Regel gilt

$$\det(C) = abc + 2xyz - ay^2 - bz^2 - cx^2.$$

Nach Satz A.75 und Bemerkung A.69 wird das Minimum von $C \mapsto \sqrt[3]{\det(C)}$ (und damit auch das Minimum von \det) auf M nur an Ecken angenommen. Nach Satz A.65 ergeben sich die Ecken, indem man drei linear unabhängig Zeilen von A wählt und das entsprechende Gleichungssystem löst. Es gibt acht Möglichkeiten aus den ersten sechs Zeilen von A drei linear unabhängig auszuwählen. Die Wahl $(x, y, z) = (a, b, -a)$ erfüllt jedoch nicht die letzte Ungleichung. Wählt man die siebte Zeile als Gleichheit, so ist $z = \frac{1}{2}(a + b) - x - y$. Dies liefert zwei weitere Ecken:

I	$2(x, y, z)$	$\det(C)$	I	$2(x, y, z)$	$\det(C)$
$\{1, 3, 5\}$	(a, b, a)	$abc - \frac{1}{4}(ab^2 + a^2c)$	$\{1, 3, 7\}$	$(a, b, 0)$	$abc - \frac{1}{4}(ab^2 + a^2c)$
$\{1, 4, 5\}$	$(a, 0, a)$	$abc - \frac{1}{4}(ab^2 + a^2c)$	$\{1, 4, 7\}$	$(a, 0, -b)$	wie $I = \{1, 4, 6\}$ mit $a = b$
$\{1, 4, 6\}$	$(a, 0, -a)$	$abc - \frac{1}{4}(ab^2 + a^2c)$	$\{1, 6, 7\}$	$(a, b - a, -a)$	$abc - \frac{1}{4}(ab^2 + a^2c)$
$\{2, 3, 5\}$	$(0, b, a)$	$abc - \frac{1}{4}(ab^2 + a^2b)$	$\{2, 3, 7\}$	$(0, b, -a)$	wie $I = \{2, 3, 6\}$
$\{2, 3, 6\}$	$(0, b, -a)$	$abc - \frac{1}{4}(ab^2 + a^2b)$	$\{2, 6, 7\}$	$(0, b, -a)$	
$\{2, 4, 5\}$	$(0, 0, a)$	$abc - \frac{1}{4}a^2b$	$\{3, 6, 7\}$	$(0, b, -a)$	
$\{2, 4, 6\}$	$(0, 0, -a)$	$abc - \frac{1}{4}a^2b$	$\{4, 6, 7\}$	$(b, 0, -a)$	wie $I = \{1, 4, 6\}$ mit $a = b$

Offenbar ist $\det(C) \geq abc - \frac{1}{4}(ab^2 + a^2c)$ und

$$abc \leq abc + \frac{1}{2}ab(c - b) + \frac{1}{2}ac(b - a) \leq 2 \det(C).$$

Gleichheit gilt nur, wenn $a = b = c$ und $2(x, y, z) \in \{(a, a, a), (a, a, 0), (a, 0, \pm a), (0, a, \pm a)\}$. Nach Beispiel 20.48 gilt $(a, a, a) \sim (a, a, 0)$. Durch Permutation von x, y, z und Vorzeichenwechsel sieht man $(a, a, 0) \sim (a, 0, \pm a) \sim (0, a, \pm a)$. \square

Satz A.77 (GAUSS). Für jedes Gitter $\Delta \subseteq \mathbb{R}^3$ gilt $\rho(\Delta) \leq \frac{\pi}{3\sqrt{2}}$. Gilt Gleichheit, so hat die Gram-Matrix von Δ die Form

$$c \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

bei geeigneter Basiswahl.

Beweis. Sei C die Gram-Matrix von Δ . Durch Basiswechsel können wir annehmen, dass q_C reduziert ist. Wie in Bemerkung 20.11 ordnen wir Einheitskugeln mit Mittelpunkten in Δ an. Da sich die Kugeln nicht überschneiden, gilt $c_{11} = \min(\Delta)^2 \geq 4$. Aus Oppenheim folgt $\det(C) \geq 2^5$. Dies zeigt

$$\rho(\Delta) = \frac{4\pi}{3 \operatorname{disc}(\Delta)} = \frac{4\pi}{3\sqrt{\det(C)}} \leq \frac{\pi}{3\sqrt{2}}.$$

Gilt Gleichheit, so kann man nach Oppenheim eine Basis von Δ konstruieren, sodass C die angegebene Form hat. \square

Satz A.78 (FISHER-Ungleichung). Sei $M = \begin{pmatrix} A & B \\ B^* & C \end{pmatrix} \in \mathbb{C}^{n \times n}$ positiv definit. Dann gilt $\det(M) \leq \det(A) \det(C)$ mit Gleichheit genau dann, wenn $B = 0$.

Beweis. Nach dem Sylvester-Kriterium sind A, C sowie A^{-1}, C^{-1} positiv definit. Mit Aufgabe II.18 gilt

$$D := \text{diag}(\sqrt{A}, \sqrt{C})^{-1} M \text{diag}(\sqrt{A}, \sqrt{C})^{-1} = \begin{pmatrix} 1 & \sqrt{A}^{-1} B \sqrt{C}^{-1} \\ \sqrt{C}^{-1} B^* \sqrt{A}^{-1} & 1 \end{pmatrix}.$$

Aus Satz A.71 folgt

$$\det(A)^{-1} \det(C)^{-1} \det(M) = \det(D) \leq (\text{tr}(D)/n)^n = 1,$$

also $\det(M) \leq \det(A) \det(C)$. Gleichheit gilt genau dann, wenn D eine Skalarmatrix ist, d. h. $\sqrt{A}^{-1} B \sqrt{C}^{-1} = 0 = B$. \square

Folgerung A.79. Sei $A = (a_{ij}) \in \mathbb{C}^{n \times n}$ positiv definit und $A^{-1} = (b_{ij})$. Dann ist $a_{ii} b_{ii} \geq 1$ für $i = 1, \dots, n$.

Beweis. Nach der Formel für die komplementäre Matrix \tilde{A} ist $b_{ii} = \det(A_{ii}) / \det(A)$. Durch Vertauschen der ersten und i -ten Zeile/Spalte ist A ähnlich zu $\begin{pmatrix} a_{ii} & B \\ B^* & A_{ii} \end{pmatrix}$. Aus Fisher folgt $\det(A) \leq a_{ii} \det(A_{ii}) = a_{ii} b_{ii} \det(A)$. \square

Stichwortverzeichnis

- A**
- Abbildung, 19
 - adjungierte, 127
 - affine, 51
 - bijektive, 19
 - diagonalisierbare, 64
 - simultan, 135
 - duale, 62
 - halbeinfache, 160
 - hermitesche, 127
 - injektive, 19
 - konkave, 249
 - strikt, 249
 - konvexe, 249
 - strikt, 249
 - lineare, 50
 - nilpotente, 138
 - normale, 127
 - orthogonale, 104
 - separable, 160
 - surjektive, 19
 - symmetrische, 104
 - trigonalisierbare, 134
 - simultan, 135
 - unitäre, 127
 - Ableitung, 158
 - Absolutglied, 86
 - Additionstheoreme, 106
 - Adjunkte, 74
 - affin (un)abhängig, 245
 - affine Hülle, 245
 - affiner Raum, 28
 - allgemeiner Entwicklungssatz, 83
 - AlphaEvolve, 175
 - Äquivalenzklasse, 18
 - Äquivalenzrelation, 18
 - Assoziativgesetz, 21, 23
 - Auslöschung, 179
 - Aussage, 11
 - äquivalente, 11
 - Austauschsatz, 32
 - Auswahlaxiom, 13, 17
 - Axiom, 13
- B**
- Banach-Tarski-Paradoxon, 13
 - Banachraum, 185
 - Banachs Fixpunktsatz, 186
 - Barnes, 252
 - Basis
 - duale, 59
 - eines Gitter, 222
 - eines Vektorraums, 30
 - Basisergänzungssatz, 32
 - Basismenge, 217, 249
 - zulässige, 217
 - Basistransposition, 83
 - Basiswechsel, 57
 - Basiswechselmatrix, 54
 - Bauer-Fike, 191
 - Begleitmatrix, 148
 - Bernoulli-Ungleichung, 244
 - Betrag, 99
 - komplexer Zahl, 108
 - Bidualraum, 59
 - Bijektion, 19
 - Bild, 19
 - Bilinearform, 113
 - alternierende, 113
 - antisymmetrische, 113
 - ausgeartete, 113
 - indefinite, 121
 - Index, 119
 - negativ (semi)definite, 120
 - positiv (semi)definite, 120
 - symmetrische, 113
 - Binet-Formel, 94
 - binäre Exponentiation, 243
 - Blichfeldt, 233
 - Blockdiagonalmatrix, 38
 - Bruhat-Zerlegung, 243
 - Bubblesort, 83
 - Bézout, 146
- C**
- Cantor, 13, 21
 - Carathéodory, 214
 - Cartan-Dieudonné, 112
 - Cauchy, 250
 - Cauchy-Binet-Formel, 72
 - Cauchy-Schwarz-Ungleichung, 100, 125
 - Cauchys Reißverschluss-Satz, 199
 - Cayley-Hamilton, 96
 - Ceres, 183

Chan-Li, 132
 charakteristisches Polynom
 einer Abbildung, 92
 einer Matrix, 92
 Chinesischer Restsatz, 158
 Cholesky-Verfahren, 194
 Cholesky-Zerlegung, 183
 Collatz-Wielandt, 208
 Conway, 242
 Courant-Fischer, 199
 Cramersche Regel, 75

D
 Darstellungsmatrix, 54
 De Morgansche Regeln, 12, 15
 Dedekind-Identität, 81
 Definitionsbereich, 19
 Determinante
 einer Abbildung, 92
 einer Matrix, 68
 quadratische Form, 236
 Determinantensatz, 71
 Dezimalbruch, 15
 Diagonalisierbarkeit, 64
 simultane, 135
 für Bilinearformen, 251
 Diagonalisierungsargument, 21
 Diagonalmatrix, 36
 Differentialgleichung, 205
 Differenz, 14
 Dimension, 33
 Dimensionsformel, 34
 direktes Produkt, 24
 Diskriminante, 224
 Distributivgesetz, 12, 15, 24
 Division mit Rest, 88
 Drehspiegelung, 112
 Drehung, 106
 Dreiecksmatrix, 66
 strikte, 66
 Dreiecksungleichung, 100, 126, 185
 Dualitätssatz, 216
 Dualraum, 59
 Durchschnitt, 14

E
 Ecke, 217, 249
 Eichler, 231
 Eigenraum, 63
 Eigenvektor, 63
 Eigenwert, 63
 Einheitsmatrix, 35
 Einheitswurzeln, 108
 Einschränkung, 19
 Einwegfunktion, 22
 Elementarmatrix, 43

Elementarteiler, 229
 Endomorphismus, 63
 Erzeugendensystem, 30
 Erzeugermatrix, 222
 euklidischer Algorithmus, 145
 euklidischer Raum, 99
 Euler, 111
 Exponentialfunktion, 22, 202

F
 Faktorraum, 28
 Faltungssatz, 172
 Farkas' Lemma, 215
 FFT, *siehe* Fourier-Transformation
 Fibonacci-Zahlen, 94
 Fillmore, 58
 Fisher-Ungleichung, 254
 Fitting, 136
 Fourier-Matrix, 172
 Fourier-Transformation
 diskrete, 172
 kontinuierliche, 173
 schnelle, 173
 Francis-Algorithmus, 193
 Frobenius, 154
 Frobenius-Norm, 188
 Frobenius-Normalform, 149
 Froebnius-Ungleichung, 82
 Fundamentalmasche, 224
 Fundamentalsatz der Algebra, 109
 Funktion, *siehe* Abbildung
 Funktional, 59
 Funktionalanalysis, 33
 Funktionalgleichung, 203

G
 Gauß, 239, 253
 Gauß-Algorithmus, 44
 mit Pivotisierung, 180
 Gauß-Seidel-Verfahren, 186
 geometrische Reihe, 200
 Gershgorin, 195
 Gitter, 222
 Basis
 δ -reduziert, 233
 duales, 223
 E_8 , 226
 ganzes, 224
 selbstduales, 223
 (un)zerlegbares, 231
 Givens-Rotation, 197
 Gleichungssystem, 41
 (in)homogenes, 41
 lösbares, 41
 unterbestimmtes, 42
 überbestimmtes, 42

Gleitkommazahlen, 179
 Goldbachs Vermutung, 11
 Golden-Thompson-Ungleichung, 204
 goldener Schnitt, 94
 Google-Matrix, 210
 Gram-Matrix
 einer Bilinearform, 114
 einer quadratischen Form, 236
 eines Gitters, 222
 Gram-Schmidt-Verfahren, 102
 modifiziertes, 196
 Gruppe, 23
 abelsche, 23
 affine, 82
 allgemeine lineare, 39
 alternierende, 77
 orthogonale, 104
 spezielle lineare, 71
 spezielle orthogonale, 105
 spezielle unitäre, 128
 symmetrische, 75
 unitäre, 127, 128
 Gödels Unvollständigkeitssätze, 13

H

Hadamard-Ungleichung, 244
 Hales, 225
 Hamiltonscher Schiefkörper, 169
 Harriot, 225
 Harvey-van der Hoeven, 173
 Hauptachsensatz, 110
 Hauptdiagonale, 36
 Hauptminor, 123
 Hauptraum, 136
 Hauptraumzerlegung, 137
 Hauptsatz
 lineare Optimierung, 219
 Hermite, 232
 Hermite-Konstante, 233
 Hermite-Normalform, 228
 Heron-Verfahren, 243
 Hesse-Matrix, 121
 Hessenberg-Matrix, 193
 Hilbert-Matrix, 176
 Hintereinanderausführung, 19
 Hölder-Ungleichung, 244
 Homogenität, 100, 125, 185
 Homomorphiesatz, 54
 Homomorphismus, 50
 Homöomorphismus, 50
 Horner-Schema, 88
 Householder-Transformation, 166, 197
 Hyperebene, 33
 Hyperwürfel, 67

I

Identität, 20

Imaginärteil, 108
 Index, 119
 Inklusionsabbildung, 20
 Interpolation, 89
 Invariante, 67
 inverses Element, 23
 Isomorphiesatz
 erster, 53
 zweiter, 53
 Isomorphismus, 50

J

Jacob, 150
 Jacobi, 203
 Jordan-Chevalley-Zerlegung, 162
 Jordan-Normalform, 140
 Jordanblock, 138
 verallgemeinerter, 161

K

Karatsuba-Algorithmus, 171
 Kardinalzahl, 21
 kartesisches Produkt, 17
 Kepler, 225
 Kern, 51
 Kitaoka, 247
 Koeffizient, 86
 führender, 86
 Koeffizientenmatrix, 41
 erweiterte, 41
 Kommutativgesetz, 23
 Komplement, 32
 duales, 60
 orthogonales, 103, 116
 komplexe Konjugation, 108
 Komposition, 19
 Konditionszahl, 176, 189
 Kongruenz
 von Matrizen, 116
 von Polynomen, 157
 Kontinuumshypothese, 13, 21
 Kontraktion, 186
 Kontraposition, 12
 konvexe Hülle, 214
 Konvexkombination, 214
 Koordinatendarstellung, 31
 Korkine-Zolotarev, 239
 Korrespondenzsatz, 81
 Kosinus, 101
 Kosinussatz, 166
 Kreuzprodukt, 104
 Kronecker-Capelli, 41
 Kronecker-Delta, 35
 Kronecker-Produkt, 247
 Körper, 24
 angeordneter, 108

der komplexen Zahlen, 108
der rationalen Funktionen, 159
mit drei Elementen, 81
mit vier Elementen, 156
mit zwei Elementen, 24

L

Lagrange-Polynom, 90

Länge

eines Zyklus, 75

Laplace-Entwicklung, 72

Large-Language-Model, 175

Leibniz-Formel, 78

Leitkoeffizient, 86

Lights Out, 84

linear (un)abhängig, 30

lineares Programm, 213

duales, 216

lösbares, 213

unbeschränktes, 213

unzulässiges, 213

Linearfaktor, 90

Linearkombination, 26

affine, 245

konvexe, 214

LLL-Algorithmus, 234

Logarithmentafel, 172

Logarithmus, 22

Lovász-Bedingung, 233

LR-Zerlegung, 181

LWE, 224

Lösungsmenge, 41

M

Mantisse, 179

Markov-Ketten, 205

Matrix, 35

adjungierte, 128

ähnliche, 57

antisymmetrische, 114

äquivalente, 47

Block-, 38

diagonal-dominante, 195

diagonalisierbare, 64

simultan, 167

dünnbesetzte, 73

gut/schlecht konditionierte, 176

hermitesche

positiv (semi)definite, 167

hermitsche, 128

inverse, 38

invertierbare, 38

komplementäre, 74

kongruente, 116

konvergiert, 190

konvergiert quasi, 192

nicht-negative, 205

nilpotente, 138

normale, 128

orthogonale, 105

positive, 205

quadratische, 35

reguläre, 38

schiefsymmetrische, 114

singuläre, 38

stochastische, 205

symmetrische, 36

transponierte, 36

trigonalisierbare, 130, 134

simultan, 167

unitäre, 128

(un)zerlegbare, 205

vertauschbare, 38

zeilen-äquivalente, 43

Matrixnorm, 188

induzierte, 188

submultiplikative, 188

Matrizeninversion, 47

Mazur-Ulam, 105

Menge

abzählbare, 21

disjunkte, 14

gleichmächtige, 19

konvexe, 214

leere, 14

(un)endliche, 13

überabzählbare, 21

Mercator-Reihe, 205

Merkregel, 37, 53, 54, 56, 65, 89, 111

Methode der kleinsten Quadrate, 183

Millenniumsproblem, 12

Min-Max-Satz, 199

Minimal-Norm, 224

Minimalpolynom, 95

Minkowski, 29, 238, 242

Minkowski-Raum, 119

Minkowski-Ungleichung, 244, 251

Mirsky, 93

Modul, 222

Modus ponens, 12

Moore-Penrose, 178

Mordell, 238

Multilinearform, 113

N

Neumann-Reihe, 200

neutrales Element, 23

Newton-Verfahren, 109

Norm, 99, 125, 184

äquivalente, 185

Normalform

Frobenius, 149

Hermite, 228
 Jordan, 140
 Smith, 229
 Weierstraß, 149
 Normalgleichungssystem, 183
 Nullmatrix, 35
 Nullpolynom, 86
 Nullraum, 25
 Nullstelle, 89
 doppelte, 91
 einfache/mehrfache, 90
 Nullvektor, 25

O
 Oppenheim, 252
 Ordnungsrelation, 18
 lexikografische, 141
 totale, 18
 Orthogonalbasis
 bzgl. Bilinearform, 117
 orthogonale Zerlegung, 231
 Orthonormalbasis, 102
 bzgl. Bilinearform, 117

P
 Page-Rang, 210
 Paradoxon, 12
 Parallelogrammgleichung, 100
 Partition, 14, 138
 Permutation, 75
 Permutationsmatrix, 77
 Perron, 206
 Perron-Frobenius, 207
 Piazzini, 183
 Pivot, 180
 Polardarstellung, 182, 244
 Polarisierung, 114
 Polyeder, 217
 Polynom, 86
 Ableitung, 158
 Begleitmatrix, 148
 irreduzibles, 145
 kongruente, 157
 konstantes, 86
 normiertes, 86
 separables, 158
 teilerfremde, 145
 Polytop, 218
 Potenzmenge, 15
 Potenzmethode, 192
 Primfaktorzerlegung in $K[X]$, 146
 Primärzerlegung, 147
 Produktregel, 159
 Projektion, 51
 Prädikat, 11
 Pseudoinverse, 178

Pythagoras, 100
 trigonometrischer, 166

Q

QR-Verfahren, 193
 QR-Zerlegung, 182
 quadratische Form, 114, 236
 äquivalente, 236
 binäre, 238
 ganze, 236
 indefinite, 246
 Minimum, 236
 nicht-ausgeartete, 236
 positive, 236
 reduzierte, 238
 unimodulare, 238
 universelle, 242
 Quantenmechanik, 204

R

Rang
 einer Abbildung, 51
 einer Matrix, 39
 eines Gitters, 222
 Rayleigh-Quotient, 199
 Realteil, 108
 Rechte-Hand-Regel, 104
 Relation, 17
 antisymmetrische, 18
 asymmetrische, 18
 reflexive, 17
 symmetrische, 17
 transitive, 18
 triviale, 18
 Repräsentantensystem, 18
 Rest, 88
 Ring, 37
 RSA, 225
 Ruffinis Regel, 88
 Russellsche Antinomie, 13
 Rückwärtssubstitution, 184

S

Sarrus-Regel, 78
 SAT-Problem, 12
 Satz vom ausgeschlossenen Dritten, 12
 Satz vom Widerspruch, 12
 Schatzman, 194
 Schlupfvariable, 213
 Schur-Horn, 132
 Schur-Zerlegung, 130
 reelle, 167
 Schurs Lemma, 155
 Schönhage-Strassen-Algorithmus, 173
 Selectionsort, 77
 Sesquilinearform, 125
 Signum, 77, 83

Simplex-Algorithmus, 221
 Simplex-Kriterium, 219
 Singulärwertzerlegung, 176
 Sinus, 102
 Sinussatz, 166
 Skalar, 25
 Skalarmatrix, 36
 Skalarprodukt, 99, 125
 Smith-Normalform, 229
 Spaltenoperation, 43
 Spaltenvektor, 35
 Spann, 29
 Spektralradius, 200
 Spektralsatz, 129
 Spektrum, 129
 Spiegelung
 in \mathbb{R}^2 , 106
 in \mathbb{R}^n , 112
 Spur
 einer Abbildung, 58
 einer Matrix, 58
 Standard-Simplex, 214
 Standardbasis, 30
 Standardmatrix, 36
 Standardskalarprodukt, 100, 125
 Steinitz, 32
 Stirling-Zahl, 244
 Strassen-Algorithmus, 174
 sukzessive Minima, 241
 Summe von Unterräumen, 29, 64
 direkte, 29
 Summenregel, 159
 Sylvester-Kriterium, 122
 Sylvester-Ungleichung, 82
 Sylvesters Determinanten-Formel, 165
 Sylvesters Trägheitssatz
 für hermitesche Matrizen, 131
 für symmetrische Bilinearformen, 118
 symplektischer Raum, 120

T

Taussky, 168
 Teiler, 88
 gemeinsamer, 227
 größter, 227
 Teilmenge, 14
 echte, 14
 Tensorprodukt, 247
 Transitivität, 12
 Translation, 82
 Transposition, 76
 Tridiagonalmatrix, 238
 Tripel, 17
 Träger, 217
 Tupel, 17

U

Übergangsmatrix, 205
 Umkehrfunktion, 21
 Umkehrung, 12
 Ungleichung harm., geom., arithm. Mittel, 250
 unitärer Raum, 125
 Untergruppe, 26
 echte, 27
 Unterraum, 27
 echter, 27
 F-invarianter, 169
 f-invarianter, 134
 zyklischer, 148
 Urbild, 19

V

Vandermonde-Matrix, 73
 Variable, 86
 Vektor, 25
 kurzer, 224
 normierter, 99, 125
 orthogonale, 99, 125
 Vektorraum, 25
 endlich erzeugter, 30
 endlich-dimensionaler, 33
 euklidischer, 99
 isomorph, 50
 unitärer, 125
 Venn-Diagramm, 14
 Vereinigung, 14
 disjunkte, 14
 Verkettung, 19
 Viazovska, 226
 Vielfachheit
 algebraische, 90
 geometrische, 63
 vollständige Induktion, 16
 von Mises, 208
 Vorwärtssubstitution, 184
 Vorzeichen, 77

W

Weierstraß-Normalform, 149
 Wertebereich, 19
 Wielandt, 209
 Wilkinson, 193

Y

Young-Ungleichung, 244

Z

Zahlen
 ganze, 14
 komplexe, 108
 natürliche, 14
 rationale, 14
 reelle, 15

Zassenhaus-Algorithmus, 48
Zeilenoperation, 43
Zeilenstufenform, 44
Zeilenvektor, 35
Zentralisator, 153, 169
Zerfällungskörper, 156
Zermelo-Fraenkel-System, 13
Zirkelschluss, 31
Zorns Lemma, 33
Zyklus, 75
 disjunkte, 76