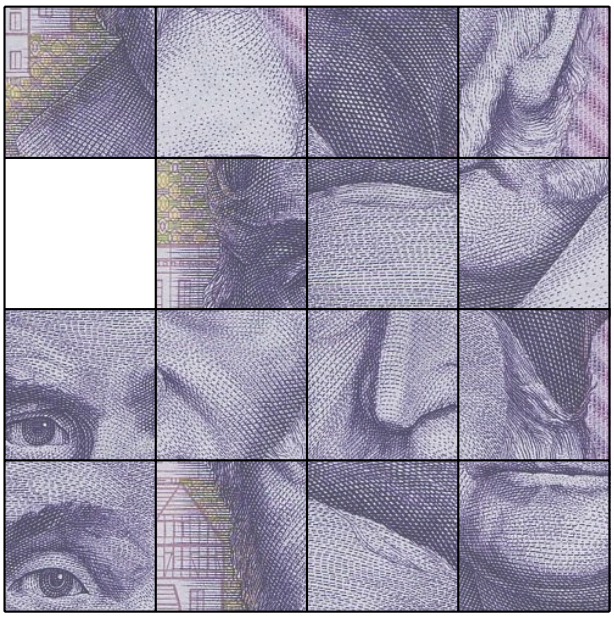
Lineare Algebra für Informatiker

Vorlesungen im Wintersemester 2020/21
und Sommersemester 2021

Benjamin Sambale
Leibniz Universität Hannover

Version: 4. Januar 2025



Inhaltsverzeichnis

Einleitung . . . 3
Vorwort . . . 3
Motivation . . . 4
Ausblick . . . 4
Notation . . . 5
Konventionen . . . 7

Lineare Algebra I **8**
1 Aussagenlogik und Mengenlehre . . . 8
1.1 Aussagen . . . 8
1.2 Mengen . . . 10
1.3 Vollständige Induktion . . . 13
2 Kartesische Produkte und Funktionen . . . 13
2.1 Paare und Tupel . . . 13
2.2 Injektive und surjektive Funktionen . . . 14
3 Körper und Vektorräume . . . 17
3.1 Gruppen und Körper . . . 17
3.2 Vektorräume und Unterräume . . . 19
4 Basen und Dimension . . . 21
4.1 Lineare Unabhängigkeit und Erzeugendensysteme . . . 21
4.2 Charakterisierung und Existenz von Basen . . . 23
4.3 Dimension . . . 25
5 Matrizen . . . 27
5.1 Der Matrizen-Vektorraum . . . 27
5.2 Matrizenmultiplikation . . . 28
5.3 Der Rang einer Matrix . . . 30
6 Der Gauß-Algorithmus . . . 31
6.1 Gleichungssysteme . . . 31
6.2 Elementare Zeilenoperationen . . . 33
6.3 Anwendungen . . . 35
7 Lineare Abbildungen . . . 39
7.1 Definitionen und Beispiele . . . 39
7.2 Darstellungsmatrizen . . . 42
Aufgaben . . . 46

Lineare Algebra II **51**
8 Eigenwerte und Eigenvektoren . . . 51
8.1 Definitionen und Beispiele . . . 51
8.2 Diagonalisierbarkeit . . . 52
9 Determinanten . . . 54
9.1 Rekursive Definition . . . 54
9.2 Eigenschaften . . . 58

9.3 Laplace-Entwicklung . . . 59
9.4 Die Leibniz-Formel . . . 61
10 Polynome . . . 63
10.1 Der Vektorraum der Polynome . . . 63
10.2 Nullstellen . . . 66
10.3 Charakteristische Polynome . . . 68
10.4 Minimalpolynome . . . 71
11 Euklidische Geometrie . . . 74
11.1 Skalarprodukte . . . 74
11.2 Orthonormalbasen . . . 77
11.3 Symmetrische und orthogonale Abbildungen . . . 79
11.4 Komplexe Zahlen . . . 82
11.5 Der Spektralsatz . . . 84
12 Die Jordansche Normalform . . . 86
12.1 Haupträume . . . 86
12.2 Jordanblöcke . . . 89
Aufgaben . . . 93

Stichwortverzeichnis **98**

Einleitung

Vorwort

Dieses Skript entstand aus asynchronen Online-Vorlesungen im Wintersemester 2020/21 und Sommersemester 2021 an der Leibniz Universität Hannover (jeweils 14 Wochen á $2 + 2$ SWS) für Studierende folgender Studiengänge:

- Bachelor of Science Informatik/Meteorologie/Computergestützte Ingenieurwissenschaften
- Bachelor Technical Education Mathematik/Informatik
- Juniorstudium Mathematik/Informatik

Zusätzlich gab es drei Bonusvorlesungen zu folgenden Themen:

- Lights Out, Folien
- Codierungstheorie, Folien
- Lineare Algebra in der Praxis, Folien

Als alternative Quellen empfehle ich die folgenden Bücher:

- A. Beutelspacher, *Lineare Algebra*, Springer Spektrum, Wiesbaden, 8. Auflage, 2014
- E. Weitz, *Konkrete Mathematik (nicht nur) für Informatiker*, Springer Spektrum, Wiesbaden, 2. Auflage, 2021

Beide lassen sich kostenlos (und legal) als E-Book über das LUH-Netz beziehen. Das Buch von Weitz ist äußerst umfangreich, modern und anwendungsorientiert.

Ich bedanke mich für Fehlerhinweise bei Alexander Ivanov, Paulina Ligocka, Christoph Pegel, Dawid Sokolihs und Marcos Soriano Sola.

Motivation

Sie haben zu verschiedenen Zeitpunkten $t_1 = 1, t_2 = 2, \ldots$ durch physikalische Experimente Messdaten $d_1 = -2, d_2 = 3, \ldots$ 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 $i = 1, 2, \ldots$. Dabei hängt f (linear) von unbekannten Parametern $x_1, x_2, \ldots$ ab, zum Beispiel $f(t) = t^2 x_1 - t x_2 + x_3$. Die Bestimmung dieser Parametern 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 wertet man verzerrte Blitzerfotos aus?
- Suchmaschinen: Nach welchen Kriterien bewertet Google Internetseiten?
- Codierungstheorie: Wie erkennt und korrigiert man Fehler bei der Übertragung digitaler Daten?
- 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?

Ausblick

Für den (erfreulichen) Fall, dass Ihr mathematisches Interesse über diese Vorlesung hinaus geht, finden Sie im Folgenden einige weiterführende Anregungen zum Thema *Gleichungen*:

- Durch Mess-Ungenauigkeiten besitzt das System (S) oft keine exakte Lösung. Die *Methode der kleinsten Quadrate* liefert eine Näherungslösung, die die Quadratsumme der Fehler minimiert. Gauß hat damit 1801 die Position des Zwergplaneten Ceres am Himmel vorhergesagt.
- Auf Computern mit Fließkomma-Arithmetik können akkumulierte Rundungsfehler zu falschen Ergebnissen führen. In der *Numerik* untersucht man *stabile* Algorithmen, die solche Effekte minimieren. Zu den bekanntesten zählen die *Cholesky-Zerlegung* und das *QR-Verfahren*.
- In vielen Anwendungen ist man an Parametern interessiert, die ein System von *Ungleichungen* erfüllen. Die Lösungen bilden einen konvexen *Polyeder* und man sucht nur solche, die eine Zielfunktion minimieren oder maximieren (man denke an Produktionskosten eines Unternehmens). Die lineare *Optimierung* stellt für solche Aufgaben das *Simplex-Verfahren* bereit.

- Sind die Parameter $x_1 = 1, x_2 = 7, x_3 = 0$ der obigen Funktion f bekannt, so kann man fragen zu welchem Zeitpunkt t der Wert $f(t) = 5$ angenommen wird. Dies führt auf die *quadratische Gleichung*

$$t^2 - 7t - 5 = 0.$$

  In der Schule haben Sie die *p-q*-Formel zur Lösung solcher Gleichungen kennen gelernt. Durch Einführen von *imaginären* Zahlen lassen sich auch kubische Gleichungen und Gleichungen vierten Grades exakt mit den *Cardanische Formeln* lösen. Zum Beispiel ist $t = \sqrt[3]{1+\sqrt{2}} + \sqrt[3]{1-\sqrt{2}}$ eine Lösung von

$$t^3 + 3t - 2 = 0.$$

  In der *Galois-Theorie* beweist man jedoch, dass es für Gleichungen höheren Grades (d. h. t tritt mindestens in der fünften Potenz auf) keine allgemeinen Lösungsformeln mehr geben kann. Trotzdem ist die Existenz von (komplexen) Lösungen durch den *Fundamentalsatz der Algebra* garantiert. In der *Analysis* entwickelt man das *Newton-Verfahren*, um solche Lösungen iterativ anzunähern.

- Quadratische Gleichungen in mehreren Unbekannten nennt man *Quadriken*, zum Beispiel

$$x^2 + 2y^2 - 3xy + 5x - 1 = 0.$$

  Die Lösungsmenge ist in diesem Fall ein *Kegelschnitt*, also die Schnittmenge eines 3-dimensionalen Kegels mit der xy-Ebene. Mit Hilfe der *Hauptachsentransformation* erhält man daraus Geraden, Ellipsen, Parabeln und Hyperbeln.

- In der *algebraischen Geometrie* studiert man Systeme von Polynomgleichungen. *Hilberts Nullstellensatz* gibt Aufschluss, wann Lösungen existieren. Ggf. bilden diese eine *algebraische Varietät*.

- Manche Vorgänge (zum Beispiel exponentielles Wachstum oder logarithmische Laufzeit von Algorithmen) lassen sich nicht durch (endliche) Polynomgleichungen ausdrücken. Stattdessen benötigt man *transzendenten* Gleichungen. Zum Beispiel ist $\pi = 3.1415\ldots$ eine Lösung von

$$x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \frac{1}{7!}x^7 \pm \ldots = 0.$$

- In der *Zahlentheorie* sucht man ganzzahlige Lösungen von *diophantischen* Gleichungen. *Fermats letzter Satz* besagt zum Beispiel, dass die Gleichung

$$x^n + y^n = z^n$$

  keine ganzzahligen positiven Lösungen x, y, z besitzt, wenn $n \geq 3$ gilt (für $n = 2$ gibt es unendlich viele Lösungen wie $3^2 + 4^3 = 5^2$, die man *pythagoreische Tripel* nennt). Der Beweis dieser Aussage füllt über 100 Seiten und wurde erst Ende des 20. Jahrhunderts vollendet.

Notation

d. h.	das heißt
ggf.	gegebenenfalls
o. B. d. A.	ohne Beschränkung der Allgemeinheit
vgl.	vergleiche
$\mathbf{w}, \mathbf{f}$	wahr, falsch
$\wedge, \vee, \neg, \Rightarrow, \Leftrightarrow, \exists, \forall$	logische Ausdrücke

$:=$	linke Seite wird durch rechte Seite definiert
$\square$	Beweisende
$\varnothing$	leere Menge
$\mathcal{P}(M)$	Potenzmenge von M
$\cup, \dot\cup, \cap, \setminus$	Vereinigung (disjunkt), Durchschnitt, Differenz von Mengen
$\lvert A\rvert$	Mächtigkeit von A (Anzahl der Elemente)
(a, b), $(x_1, \ldots, x_n)$	Paar, n-Tupel
$A_1 \times \ldots \times A_n$	kartesisches Produkt von Mengen $A_1, \ldots, A_n$
$\mathbb{N}, \mathbb{N}_0, \mathbb{Z}$	natürliche Zahlen (ohne und mit 0), ganze Zahlen
$\mathbb{Q}, \mathbb{R}, \mathbb{R}_{>0}, \mathbb{C}$	rationale, reelle (positive) und komplexe Zahlen
$\mathbb{F}_2$	Körper mit zwei Elementen
$K^\times$	$= K \setminus \{0\}$
$f_{\lvert A}, f^{-1}, f \circ g$	Einschränkung, Umkehrfunktion und Komposition von Funktionen
$f(A), f^{-1}(B), \mathrm{Ker}(f)$	Bild, Urbild, Kern von f
$\mathrm{id}, \mathrm{id}_V$	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
$U \leq V$, $U < V$, $U \cong V$	Unterräume (echte), isomorphe Vektorräume
$U \oplus W$	direkte Summe von U und W
$e_1, \ldots, e_n$	Standardbasis von K^n
$\langle S\rangle, \langle s_1, \ldots, s_n\rangle$	Spann von S
$\dim V$	Dimension von V
${}_B[v]$	Koordinatendarstellung von v bzgl. der Basis B
$\mathrm{Hom}(V, W)$	Vektorraum aller linearen Abbildungen $V \to W$
$\mathrm{End}(V)$	$= \mathrm{Hom}(V, V)$
$K^{n\times m}$	Vektorraum der $n \times m$-Matrizen über K
$\mathrm{GL}(V), \mathrm{GL}(n, K)$	allgemeine lineare Gruppe
$\mathrm{O}(V), \mathrm{O}(n, K)$	orthogonale Gruppe
$0_{n\times m}, 1_n$	Nullmatrix, Einheitsmatrix
E_{st}	Standardmatrix mit 1 an Position (s, t)
$A \sim B$	A zeilen-äquivalent zu B
$(A\lvert b)$	erweiterte Koeffizientenmatrix
$A^{\mathrm{t}}, A^{-1}, A^{-\mathrm{t}}$	Transponierte, Inverse, Transponiert-Inverse von A
$\widehat{A}, \widetilde{A}, \overline{A}$	Zeilenstufenform, komplementäre, komplex-konjugierte Matrix von A
$\mathrm{rk}(A), \mathrm{tr}(A), \det(A)$	Rang, Spur, Determinante von A
${}_C[f]_B, [f], {}_C\Delta_B$	Darstellungsmatrix, Basiswechselmatrix
$E_\lambda(f)$, $H_\lambda(f)$	Eigenraum, Hauptraum zum Eigenwert λ von f
S_n	symmetrische Gruppe vom Grad n
$\mathrm{sgn}(\sigma), P_\sigma$	Signum, Permutationsmatrix von σ
$K[X]$	Vektorraum der Polynome mit Koeffizienten in K
$\deg(\alpha)$	Grad von α
$\alpha \mid \beta$	α teilt β
χ_A, μ_A	charakteristisches Polynom, Minimalpolynom von A
$[v, w]$, $\lvert v\rvert$	Skalarprodukt, Norm von Vektoren
$v \perp w$	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

$v \times w$	Kreuzprodukt von v und w
$D(\varphi)$, $S(\varphi)$	Drehung, Spiegelung in $\mathbb{R}^2$
$\mathrm{Re}(z)$, $\mathrm{Im}(z)$, $\overline{z}$	Realteil, Imaginärteil, komplexe Konjugation von z
i	imaginäre Einheit
$J_n(\lambda)$	Jordanblock der Größe $n \times n$

Konventionen

- 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 ($A, B, \ldots, M, \ldots$).
- 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 und Skalare (Körperelemente im Kontext von Vektorräumen) 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.

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 Aussage $A = A(x)$, deren Wahrheitswert von einer Variablen x abhängt. Dann 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**).
- $2^{2^{33}} + 1$ ist eine Primzahl (?[1]).

Keine Aussagen dagegen sind:

- Seien p und q verschiedene Primzahlen. (*Annahme*)
- Dieser Satz ist falsch. (*Paradoxon*)

[1] Ihr Computer kann zur Bestimmung des Wahrheitswerts beitragen, siehe `http://www.prothsearch.com/`.

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. 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* $B \Rightarrow A$.[2]

(c) Für Aussagen $A_1, \ldots, A_n$ definiert man $A_1 \wedge \ldots \wedge A_n$ durch $\forall i : A_i$ und $A_1 \vee \ldots \vee A_n$ 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.[3]

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

(a) Die folgenden Aussagen sind äquivalent zu A:

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

(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 verdienen[4]). Alternativ kann man einige der Behauptungen aus bereits bewiesenen ableiten. So folgt die zweite De Morgansche Regel aus der ersten:

$$\neg(A \vee B) \overset{\text{(a)}}{\Longleftrightarrow} \neg((\neg\neg A) \vee (\neg\neg B)) \Longleftrightarrow \neg(\neg(\neg A \wedge \neg B)) \overset{\text{(a)}}{\Longleftrightarrow} (\neg A \wedge \neg B). \qquad \square^5$$

[2] Man könnte auch $A \Leftarrow B$ schreiben.

[3] Ein *Lemma* ist ein Hilfssatz mit wenig eigener Bedeutung.

[4] Das SAT-*Problem* der theoretischen Informatik ist NP-vollständig. Eines der sieben *Millenniumsprobleme* fragt, ob P = NP.

[5] Diese 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.[6] 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

$$M := \{x : x \notin x\} \qquad \text{(Russellsche } \textit{Antinomie}\text{)}$$

(das Symbol $:=$ besagt, dass die linke Seite durch die rechte Seite festgelegt wird). 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*. 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 zweiten *Unvollständigkeitssatz* ist es unmöglich zu beweisen, dass die Zermelo-Fraenkel-Axiome keine Widersprüche liefern. 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. Das bekannteste Beispiel hierfür ist die *Kontinuumshypothese* (siehe Bemerkung 2.6(e)).

Definition 1.9.

(a) Für Mengen A und B sei

$$\begin{aligned} \varnothing &:= \{\} \qquad (\textit{leere Menge}), \\ A \cup B &:= \{x : x \in A \vee x \in B\} \qquad (\textit{Vereinigung}), \end{aligned}$$

[6] Cantors Wortlaut

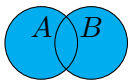
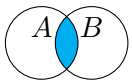
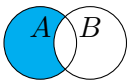
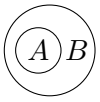
$$A \cap B := \{x : x \in A \wedge x \in B\} \qquad (\textit{Durchschnitt}),^{7}$$
$$A \setminus B := \{x : x \in A \wedge x \notin B\} \qquad (\textit{Differenz}).^{8}$$

(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* Teilmenge[9]). Ist A keine Teilmenge von B, so schreibt man $A \nsubseteq B$.

(c) Man nennt A und B *disjunkt*, falls $A \cap B = \varnothing$. 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:

$A \cup B$:  $A \cap B$:  $A \setminus B$:  $A \subseteq B$: 

Achtung: Sind mehr als drei Mengen im Spiel, so kann die allgemeine Situation nicht mehr durch Kreise dargestellt werden.[10]

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

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

(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* $\mathbb{N} := \{1, 2, 3, \ldots\}$. Wir setzen $\mathbb{N}_0 := \{0, 1, 2 \ldots\} = \mathbb{N} \cup \{0\}$.[11] Achtung: Bei manchen Autoren ist $0 \in \mathbb{N}$.

(b) Die Menge der *ganzen Zahlen* $\mathbb{Z} := \{\ldots, -2, -1, 0, 1, 2, \ldots\}$. 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* $\mathbb{Q} := \{\frac{a}{b} : a, b \in \mathbb{Z},\, b \neq 0\}$.

(d) Die Menge der *reellen Zahlen* $\mathbb{R}$ besteht aus allen *Dezimalbrüchen* wie $2 = 2.0$, $\frac{1}{3} = 0.33\ldots$, $\sqrt{2} = 1.4142\ldots$ oder $\pi = 3.1415\ldots$ (die Dezimalbruchentwicklung kann abbrechend, periodisch oder unperiodisch sein). Auf dem Computer lassen sich reelle Zahlen nur näherungsweise durch *Fließkommazahlen* implementieren. In der Analysis definiert man reelle Zahlen als Grenzwerte von rationalen Cauchy-Folgen. Im Folgenden setzen wir die üblichen Regeln für die Grundrechenarten und die Ordnungsrelation $\leq$ voraus. Auch diese lassen sich streng axiomatisch einführen.

[7] Man beachte die Ähnlichkeit der Symbole $\cup$ und $\vee$ sowie $\cap$ und $\wedge$.
[8] In manchen Bücher schreibt man $A - B$ anstatt $A \setminus B$.
[9] Das Symbol $\subset$ wird in der Literatur leider nicht einheitlich benutzt.
[10] siehe `https://de.wikipedia.org/wiki/Mengendiagramm`
[11] Streng axiomatisch definiert man $0 := \varnothing$, $1 := \{\varnothing\}$ und allgemein $n + 1 := n \cup \{n\}$.

(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).

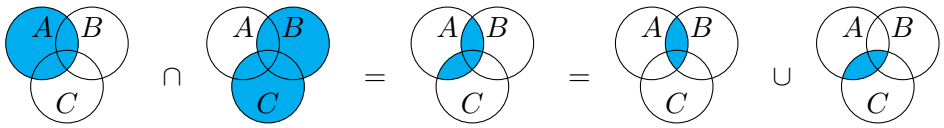
(c) $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$ *und* $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$ (De Morgansche Regeln).

(d) $\boxed{|A \cup B| + |A \cap B| = |A| + |B|}$ *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 $A \cap B = \{x_1, \ldots, x_s\}$, $A = \{x_1, \ldots, x_s, a_1, \ldots, a_t\}$ und $B = \{x_1, \ldots, x_s, b_1, \ldots, b_u\}$. Dann gilt

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

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

1.3 Vollständige Induktion

Satz 1.13 (Prinzip der vollständigen Induktion)**.** *Sei $A(n)$ ein Prädikat für $n \in \mathbb{N}$ mit den Eigenschaften:*

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

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 $A(n-1) \Rightarrow A(n)$. Also gilt $A(n)$ nach Modus ponens. Widerspruch. □

Bemerkung 1.14. Man verwendet oft Varianten der vollständigen Induktion. Zum Beispiel:

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

Beispiel 1.15. Wir beweisen $(1 + 2 + \ldots + n)^2 = 1^3 + 2^3 + \ldots + n^3$ für alle $n \in \mathbb{N}$.

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

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

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

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

(das hat Gauss als 9-Jähriger erkannt[12]). Nach der binomischen Formel gilt nun

$$\begin{aligned} ((1 + 2 + \ldots + n) + (n + 1))^2 &= (1 + 2 + \ldots + n)^2 + 2(1 + 2 + \ldots + n)(n + 1) + (n + 1)^2 \\ &\overset{(*)}{=} 1^3 + 2^3 + \ldots + n^3 + n(n + 1)(n + 1) + (n + 1)^2 \\ &= 1^3 + 2^3 + \ldots + n^3 + (n + 1)^3. \end{aligned}$$

□

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 nun eine geordnete Variante ein.

[12] siehe `https://de.wikipedia.org/wiki/Gau%C3%9Fsche_Summenformel#Herkunft_der_Bezeichnung`

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*[13] (a, b) mit $a \in A$, $b \in B$, sodass gilt

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

  Es gilt $\boxed{|A \times B| = |A||B|,}$ 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* $(a_1, \ldots, a_n)$ für $n \geq 2$. Für Mengen $A_1, \ldots, A_n$ setzt man

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

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

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

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

2.2 Injektive und surjektive Funktionen

Definition 2.4.

- 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.[14] Man schreibt dann[15]

$$f \colon A \to B, \qquad a \mapsto f(a).$$

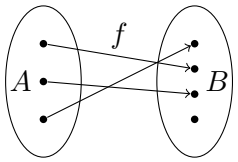
- Man nennt A den *Definitionsbereich* und B den *Wertebereich* von f. Außerdem ist $f(a)$ das *Bild* von a unter f und $f(A) := \{f(a) : a \in A\} \subseteq B$ ist das *Bild* von f. Für $B' \subseteq B$ ist

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

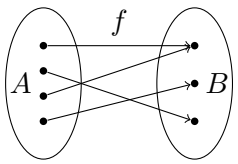
  das *Urbild* von B' unter f.

- Man nennt $f \colon A \to B$

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

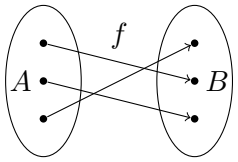


  – *surjektiv*, falls $\forall b \in B : \exists a \in A : f(a) = b$, d. h. $f(A) = B$.



[13] Die formale Definition von Paaren kann man auf Mengen zurückführen: $(a, b) := \{\{a\}, \{a, b\}\}$.
[14] Formal: 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.
[15] Man beachte die unterschiedlichen Pfeile $\to$ und $\mapsto$.

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



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

$$f_{|A'}\colon A' \to B, \qquad a \mapsto f(a).$$

Für eine weitere Funktion $g\colon B \to C$ nennt man die Abbildung

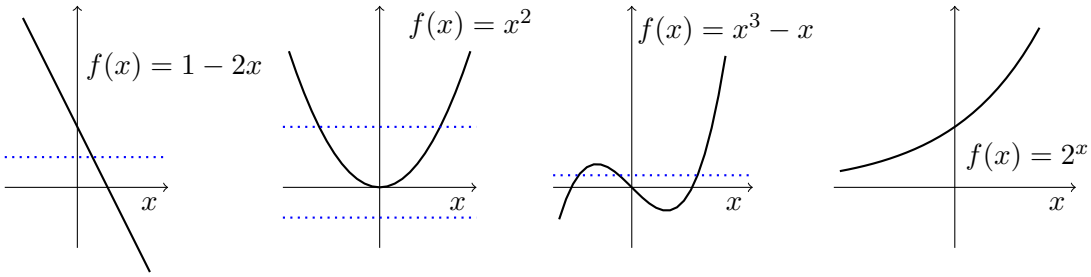
$$g \circ f\colon A \to C, \qquad a \mapsto g(f(a))$$

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

Beispiel 2.5.

(a) Für jede Menge A und $B \subseteq A$ ist $f\colon B \to A$, $b \mapsto b$ eine injektive Funktion, die man *Inklusionsabbildung* nennt. Im Fall $B = A$ ist f sogar bijektiv und man nennt $f = \mathrm{id}_A$ die *Identität* auf A.

(b) Abbildungen $f\colon \mathbb{R} \to \mathbb{R}$ 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 2.6.

(a) Achtung: Injektiv ist nicht das Gegenteil von surjektiv (ein häufiger Anfängerfehler)!

(b) Man kann jede Funktion $f\colon A \to B$ surjektiv machen, indem man auf den Wertebereich auf das Bild einschränkt: $f\colon A \to f(A)$.

(c) Für $f\colon A \to B$ 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 $\infty \leq \infty$).

(d) Zwei endliche Mengen A und B sind genau dann gleichmächtig, wenn $|A| = |B|$. In diesem Fall sind die Eigenschaften injektiv, surjektiv und bijektiv nach (c) äquivalent. Für unendliche Mengen ist dies im Allgemeinen falsch.

(e) Obwohl $\mathbb{N}$ nur „halb so viele“ Zahlen wie $\mathbb{Z}$ enthält, sind $\mathbb{N}$ und $\mathbb{Z}$ durch die Bijektion

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

gleichmächtig. Nach Cantors *Diagonalisierungsargumenten* ist $\mathbb{N}$ auch zu $\mathbb{Q}$ gleichmächtig, aber nicht zu $\mathbb{R}$, d. h. $\mathbb{R}$ ist „wesentlich größer“ als $\mathbb{N}$ (man sagt: $\mathbb{R}$ ist *überabzählbar*). Es gibt sogar beliebig „große“ Mengen (*Kardinalzahlen*), die man in der Praxis aber selten antrifft (Aufgabe 6). Die (nicht beweisbare) *Kontinuumshypothese* besagt, dass jede unendliche Teilmenge von $\mathbb{R}$ entweder zu $\mathbb{N}$ oder zu $\mathbb{R}$ gleichmächtig ist.

Lemma 2.7. *Seien $f\colon A \to B$, $g\colon B \to C$, $h\colon C \to D$ 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 $g\colon B \to A$ mit $g \circ f = \mathrm{id}_A$ und $f \circ g = \mathrm{id}_B$ 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 $g(B) = C$.

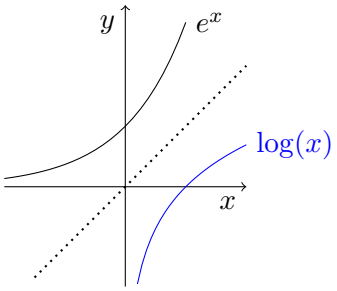
(f) Ist $g \circ f = \mathrm{id}_A$ und $f \circ g = \mathrm{id}_B$, so ist f injektiv nach (d) und surjektiv nach (e), also auch bijektiv. Sei umgekehrt f bijektiv. Für jedes $b \in B$ existiert dann genau ein $g(b) \in A$ mit $f(g(b)) = b$. Daher ist $g\colon B \to A$ die einzige Abbildung mit $f \circ g = \mathrm{id}_B$. Aus $f(a) = f(g(f(a)))$ folgt $g(f(a)) = a$ für alle $a \in A$, da f injektiv ist. Dies zeigt $g \circ f = \mathrm{id}_A$. □

Bemerkung 2.8. Verwechseln Sie die Umkehrfunktion nicht mit dem Urbild. Der Zusammenhang beider Konzepte ist $f^{-1}(\{b\}) = \{f^{-1}(b)\}$ für jede Bijektion $f\colon A \to B$ und $b \in B$.

Beispiel 2.9.

(a) Die Abbildung $f\colon \mathbb{Q} \to \mathbb{Q}$, $x \mapsto 2x + 1$ ist eine Bijektion mit Umkehrabbildung $f^{-1}\colon \mathbb{Q} \to \mathbb{Q}$, $x \mapsto \frac{x-1}{2}$ (nachrechnen).

(b) Die Umkehrabbildung der *Exponentialfunktion* $\exp\colon \mathbb{R} \to \mathbb{R}_{>0}$, $x \mapsto e^x$ ist der *natürliche Logarithmus* $\log\colon \mathbb{R}_{>0} \to \mathbb{R}$. 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 $(\mathbb{Q}, \mathbb{R}, \ldots)$ 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 $G \times G \to G$, $(x, y) \mapsto x \cdot y$. 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*.

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 $x, y \in G$.[16]

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

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 $G_1, \ldots, G_n$ ist auch $G_1 \times \ldots \times G_n$ eine Gruppe mit

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

für $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in G_1 \times \ldots \times G_n$ (Aufgabe 8). Das neutrale Element ist $(1_{G_1}, \ldots, 1_{G_n})$. Man spricht dann vom *direkten Produkt* von $G_1, \ldots, G_n$ (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 geht 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 $\boxed{x \cdot 0 = 0 = 0 \cdot x,}$ 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* $xz = yz \Longrightarrow x = y$, denn

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

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

Beispiel 3.7.

(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 14). 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
0	0	1
1	1	0

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

Auf Computern werden alle Rechnungen in $\mathbb{F}_2$ durchgeführt, indem man 0 und 1 als *Bits* interpretiert. In der Algebra zeigt man, dass für jede Primzahlpotenz q ein Körper mit genau q Elementen existiert (vgl. Aufgabe 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* $K \times V \to V$, $(\lambda, v) \mapsto \lambda \cdot v$ 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*.

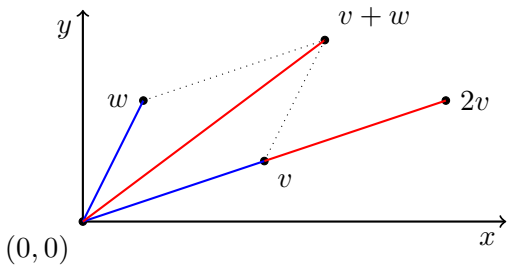
Bemerkung 3.9. Man beachte, dass $+$ sowohl die Addition in K als auch in V bezeichnet. Ebenso steht $\cdot$ für die Multiplikation in K und für die Skalarmultiplikation (das ist ungenau, aber durchaus üblich). Wir werden in beiden Fällen das Symbol $\cdot$ oft einsparen. Falls Missverständnisse ausgeschlossen 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 $V_1, \ldots, V_n$ ist auch das direkte Produkt (bzgl. $+$) $V_1 \times \ldots \times V_n$ ein Vektorraum mit komponentenweiser Skalarmultiplikation: $\lambda(v_1, \ldots, v_n) := (\lambda v_1, \ldots, \lambda v_n)$ 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:



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

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

Sind $v_1, \ldots, v_n$ paarweise verschieden (d. h. $v_i \neq v_j$ für $i \neq j$)[17] 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 nun auch $v_i = \mu_{i1} w_{i1} + \mu_{i2} w_{i2} + \ldots + \mu_{im} w_{im}$ eine Linearkombination für $i = 1, \ldots, n$. 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:

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

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

Definition 3.11. Ein *Unterraum* eines Vektorraums V ist eine Teilmenge $U \subseteq V$, die mit den eingeschränkten Verknüpfungen selbst einen Vektorraum bildet, d. h.

- $0_V \in U$,
- $\forall v, w \in U : v + w \in U$,
- $\forall v \in U, \lambda \in K : \lambda v \in U$.

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

Bemerkung 3.12. Die Bedingungen garantieren, dass U unter Addition und Skalarmultiplikation *abgeschlossen* ist. Somit sind die Verknüpfungen auf U *wohldefiniert*. Die verbleibenden Vektorraumaxiome muss man nicht prüfen, da sie bereits in der größeren Menge V gelten. Man kann die Bedingungen auch 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 10).

[17] „paarweise verschieden“ ist stärker als die Formulierung „nicht alle sind gleich“

Beispiel 3.13.

(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 $U := \big\{(x, 0) : x \in \mathbb{R}\big\} \leq \mathbb{R}^2$ mit Hilfe von Bemerkung 3.12: Wegen $(0, 0) \in U$ ist $U \neq \varnothing$. 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 $U := \big\{(x, y, 0) \in \mathbb{R}^3 : x, y \in \mathbb{R}\big\}$ ein Unterraum von $\mathbb{R}^3$.

(e) Die Teilmenge $U := \big\{(x, x^2) : x \in \mathbb{Q}\big\}$ von $\mathbb{Q}^2$ ist *kein* Unterraum, denn $(1, 1) \in U$, aber $2 \cdot (1, 1) = (2, 2) \notin U$. Wir werden zeigen, dass sich jeder Unterraum durch *lineare* Gleichungen beschreiben lässt.

(f) Offenbar sind $U := \big\{(0, 0), (1, 0)\big\}$ und $W := \big\{(0, 0), (0, 1)\big\}$ Unterräume von $\mathbb{F}_2^2$, aber nicht $U \cup W$ (Warum?)

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.[18] Im Fall $S = \{s_1, \ldots, s_n\}$ schreiben wir auch $\langle s_1, \ldots, s_n \rangle$ 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 *Summe* von U und W. Im Fall $U \cap W = \{0\}$ nennt man die Summe *direkt* und schreibt $U \oplus W$ anstatt $U + W$.[19]

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.*

[18] In manchen Büchern schreibt man $\mathrm{Span}(S)$ anstatt $\langle S \rangle$.
[19] Dies ersetzt die (disjunkte) Vereinigung, siehe Aufgabe 12.

Beweis. Offenbar ist 0 eine Linearkombination von Elementen aus S, d. h. $0 \in \langle S \rangle$ (im Fall $S = \varnothing$ 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 $\langle \varnothing \rangle = \{0\}$, denn die leere Summe ist die einzige Linearkombination aus $\varnothing$.

(b) Für $U \leq W \leq V$ gilt $U + W = W$ und $\langle U \rangle = U = U \oplus \{0\}$.

(c) Für $s_1, \ldots, s_n \in V$ gilt $\langle s_i \rangle = \{\lambda s_i : \lambda \in K\} =: Ks_i$ und $\langle s_1, \ldots, s_n \rangle = Ks_1 + \ldots + Ks_n$. Insbesondere ist $\mathbb{R}^2 = \mathbb{R}(1,0) \oplus \mathbb{R}(0,1)$.

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 $s_1, \ldots, s_n \in S$ und $\lambda_1, \ldots, \lambda_n \in K$ gilt:

$$\sum_{i=1}^{n} \lambda_i s_i = 0 \qquad \Longrightarrow \qquad \lambda_1 = \ldots = \lambda_n = 0.$$

- *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 $s_1, \ldots, s_n$ heißen linear unabhängig (bzw. bilden eine Basis), falls sie paarweise verschieden sind und $\{s_1, \ldots, s_n\}$ linear unabhängig (bzw. eine Basis) ist.

Beispiel 4.7.

(a) Die leere Menge ist stets linear unabhängig und bildet eine Basis des Nullraums.

(b) 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

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

(c) 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.

(d) Jede Teilmenge einer linear unabhängigen Menge ist linear unabhängig.

(e) Für $n \geq 1$ seien

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

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

4.2 Charakterisierung und Existenz von Basen

Satz 4.8. *Sei $b_1, \ldots, b_n$ 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 $\lambda_1, \ldots, \lambda_n \in K$ schreiben. Insbesondere ist die Abbildung*

$$_B[.] \colon V \to K^n, \qquad v \mapsto {}_B[v] := (\lambda_1, \ldots, \lambda_n)$$

eine Bijektion.

Beweis. Wegen $V = \langle b_1, \ldots, b_n \rangle$ ist jedes $v \in V$ eine Linearkombination der angegebenen Form. Seien $\lambda_1, \ldots, \lambda_n, \mu_1, \ldots, \mu_n \in K$ 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 $\{b_1, \ldots, b_n\}$ linear unabhängig ist, folgt $\lambda_i = \mu_i$ für $i = 1, \ldots, n$. □

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*.[20]

(1) ⇒ (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 $\lambda_1, \ldots, \lambda_n \in K$ und $b_1, \ldots, b_n \in B \setminus \{b\}$ 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.

[20] Die zeigt nur die Äquivalenz der Aussagen. Ein *Zirkelschluss* ist die (falsche) Behauptung, dass jede der Aussagen gilt.

(2) $\Rightarrow$ (3): Sei B ein minimales Erzeugendensystem. Sei $\sum_{i=1}^n \lambda_i b_i = 0$ für $\lambda_1, \ldots, \lambda_n \in K$ und paarweise verschiedene $b_1, \ldots, b_n \in B$. 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 $\lambda_1 = \ldots = \lambda_n = 0$ und B ist linear unabhängig. Sei nun $v \in V \setminus B$. Wegen $\langle B \rangle = V$ existieren $\lambda_1, \ldots, \lambda_n \in K$ und $b_1, \ldots, b_n \in B$ 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 $\lambda_1, \ldots, \lambda_n \in K^\times$, $\mu \in K$, $b_1, \ldots, b_n \in B$ 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 $V = \langle B \rangle$. □

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 $E = \{s_1, \ldots, s_n\}$. 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 $E \nsubseteq \langle U \rangle$ 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. □

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

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

Beweis. O. B. d. A. sei E endlich, sagen wir $E = \{s_1, \ldots, s_n\}$. Seien $u_1, \ldots, u_m \in U$ 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 $\lambda_1, \ldots, \lambda_n \in K$ verschwinden. Sei also o. B. d. A. $\lambda_1 \neq 0$ und daher

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

Folglich ist auch $\{u_1, s_2, \ldots, s_n\}$ 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 $\mu_1, \ldots, \mu_n \in K$. 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

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

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

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 $\varnothing$ 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 Austauschsatz gilt $|C| \leq |B| \leq |C|$, also $|C| = |B|$. Da B und C endlich sind, müssen sie nach Bemerkung 2.6(d) gleichmächtig sein. □

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 Austauschsatz 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 $W := \langle D \setminus C \rangle$. □

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 $d = \dim V := |B| \in \mathbb{N}_0$ 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) Sei V ein d-dimensionaler $\mathbb{F}_2$-Vektorraum. Die Koordinatendarstellung bzgl. einer Basis zeigt
$$|V| = |\mathbb{F}_2^d| = |\mathbb{F}_2 \times \ldots \times \mathbb{F}_2| = 2^d.$$

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*

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

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

Beweis. Sei $\{b_1, \ldots, b_n\}$ eine Basis von $U \cap W$. Wir ergänzen zu einer Basis $\{b_1, \ldots, b_n, c_1, \ldots, c_s\}$ von U und einer Basis $\{b_1, \ldots, b_n, d_1, \ldots, d_t\}$ 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 $b_1, \ldots, b_n, c_1, \ldots, c_s, d_1, \ldots, d_t$ erzeugt.

Für die lineare Unabhängigkeit seien $\lambda_1, \ldots, \lambda_n, \mu_1, \ldots, \mu_s, \rho_1, \ldots, \rho_t \in K$ 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 $b_1, \ldots, b_n$ ausdrücken. Andererseits ist die Darstellung von $v + u$ bzgl. der Basis $\{b_1, \ldots, b_n, c_1, \ldots, c_s\}$ eindeutig nach Satz 4.8. Dies zeigt $\mu_1 = \ldots = \mu_s = 0$. Nun ist $v + w = 0$ eine Linearkombination der Basis $\{b_1, \ldots, b_n, d_1, \ldots, d_t\}$. Dies geht nur falls $\lambda_1 = \ldots = \lambda_n = \rho_1 = \ldots = \rho_t = 0$. Daher ist $\{b_1, \ldots, b_n, c_1, \ldots, c_s, d_1, \ldots, d_t\}$ linear unabhängig und folglich eine Basis von $U + W$. Dies zeigt

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

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 $\mathbb{R}^3 = U \oplus W$.

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

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

mit $a_{ij} \in K$ für $i = 1, \ldots, n$ und $j = 1, \ldots, m$.[21] 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 $m \times 1$-Matrizen heißen demnach *Spaltenvektoren*.

(b) Die $n \times m$-*Nullmatrix* $0_{n \times m} := (0)_{i,j} \in K^{n \times m}$ (wie üblich lassen wir den Index $n \times m$ oft 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 $\{e_1, \ldots, e_n\}$ von K^n.

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

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

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

[21] Streng genommen müsste man $a_{i,j}$ anstatt a_{ij} schreiben.

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

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

Offensichtlich ist $(A^{\mathrm{t}})^{\mathrm{t}} = A$. Durch Transponieren werden Zeilenvektoren zur Spaltenvektoren und umgekehrt.

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 $K^{n\times m}$.

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 $e_1, \ldots, e_{nm}$ von K^{nm} entsprechen (bis auf Reihenfolge) genau den Standardmatrizen. □

Beispiel 5.5. Achtung: Nur Matrizen vom gleichen Format können addiert werden. Zum Beispiel:

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

Die Skalarmatrizen sind genau die skalaren Vielfachen der Einheitsmatrix.

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

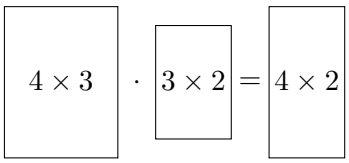
$$\boxed{c_{ij} := \sum_{l=1}^{m} a_{il}b_{lj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{im}b_{mj}.}$$

Bemerkung 5.7.

(a) Merkregel: 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:



(b) Die Multiplikation von Diagonalmatrizen ist einfach:

$$\operatorname{diag}(\lambda_1, \ldots, \lambda_n) \cdot \operatorname{diag}(\mu_1, \ldots, \mu_n) = \operatorname{diag}(\lambda_1\mu_1, \ldots, \lambda_n\mu_n).$$

(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.[22]

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

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

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

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

□

Bemerkung 5.9.

(a) Für $A \in K^{n\times m}$ gilt selbstverständlich $0_{k\times n} \cdot A = 0_{k\times m}$ und $A \cdot 0_{m\times k} = 0_{n\times k}$.

[22] Man nennt diese schwächere Struktur einen *Ring*.

(b) Für $n \geq 2$ ist Matrizenmultiplikation nicht kommutativ:

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = 0_{2\times 2} \neq \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Daher ist $(K^{n\times n}, +, \cdot)$ *kein* Körper. Wir sehen auch, dass nicht jede (von $0_{n\times n}$ 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.[23] Gilt auch $AC = 1_n = CA$, so ist $C = C1_n = C(AB) = (CA)B = 1_nB = B$. Daher ist B eindeutig bestimmt und man schreibt wie bei Gruppen $A^{-1} := B$. 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

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

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

Lemma 5.10. *Die invertierbaren Matrizen in $K^{n\times n}$ bilden eine Gruppe* $\mathrm{GL}(n, K)$ *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. □

Beispiel 5.11. In $\mathrm{GL}(2, \mathbb{F}_2)$ 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 $z_1, \ldots, z_n$ die Zeilen und $s_1, \ldots, s_m$ die Spalten einer Matrix $A \in K^{n\times m}$. Dann gilt* $\dim\langle z_1, \ldots, z_n\rangle = \dim\langle s_1, \ldots, s_m\rangle$.

Beweis. Nach Bemerkung 4.20 existiert $I \subseteq \{1, \ldots, n\}$, sodass $\{z_i : i \in I\}$ eine Basis von $\langle z_1, \ldots, z_n\rangle$ ist. Sei analog $\{s_j : j \in J\}$ eine Basis von $\langle s_1, \ldots, s_m\rangle$ ist. 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 $k \in \{1, \ldots, m\} \setminus J$ 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 $i \in \{1, \ldots, n\}$. Es folgt

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

[23] Invertierbare (bzw. nicht invertierbare) Matrizen nennt man auch *regulär* (bzw. *singulär*).

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 $|I|\le|J|$. Die Zeilen (bzw. Spalten) von A sind die Spalten (bzw. Zeilen) von A^{t}. Benutzt man das obige Argument mit A^{t}, so folgt $|J|\le|I|$. Insgesamt ist $\dim\langle z_1,\dots,z_n\rangle=|I|=|J|=\dim\langle s_1,\dots,s_m\rangle$. □

Definition 5.13. In der Situation von Satz 5.12 nennt man $\mathrm{rk}(A):=\dim\langle z_1,\dots,z_n\rangle=\dim\langle s_1,\dots,s_m\rangle$ den *Rang* von A. Im Fall $\mathrm{rk}(A)=\min\{n,m\}$ 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 $\mathrm{rk}\left(\begin{smallmatrix}1&2\\2&4\end{smallmatrix}\right)=1$, 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 $\mathrm{rk}(A)=\mathrm{rk}(A^{\mathrm{t}})$ und $\mathrm{rk}(A)=0\iff A=0$.

Lemma 5.15.

(a) Für Matrizen A und B von „passenden" Format gilt $\mathrm{rk}(AB)\le\min\{\mathrm{rk}(A),\mathrm{rk}(B)\}$.

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

$$\mathrm{rk}(AB)=\mathrm{rk}((AB)^{\mathrm{t}})=\mathrm{rk}(B^{\mathrm{t}}A^{\mathrm{t}})\le\mathrm{rk}(B^{\mathrm{t}})=\mathrm{rk}(B).$$

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

$$n=\mathrm{rk}(1_n)=\mathrm{rk}(AA^{-1})\overset{\text{(a)}}{\le}\mathrm{rk}(A)\le n,$$

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

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 $L\neq\varnothing$ 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

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

entspricht der Matrixgleichung

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

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

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

Satz 6.4 (Kronecker-Capelli[24])**.** *Genau dann ist das Gleichungssystem $Ax = b$ lösbar, wenn* $\mathrm{rk}(A) = \mathrm{rk}(A|b)$.

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

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

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

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

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 $m - \mathrm{rk}(A)$.

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

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

die Lösungsmenge.

[24] mitunter auch nach Rouché und Frobenius benannt

Beweis.

(a) Wegen $0 \in L_0$ ist $L_0 \neq \varnothing$. 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 3.12). Sei $b_1, \ldots, b_k$ eine Basis von L_0. Wir ergänzen zu einer Basis $b_1, \ldots, b_m$ von $K^{m\times 1}$. Die i-te Spalte von A ist Ae_i^{t} mit dem Standardbasisvektor e_i. Da e_i^{t} eine Linearkombination von $b_1, \ldots, b_m$ ist, liegt jede Spalte von A in $\langle Ab_1, \ldots, Ab_m\rangle$. Insbesondere ist $\mathrm{rk}(A) = \dim\langle Ab_1, \ldots, Ab_m\rangle$. Aus $b_1, \ldots, b_k \in L_0$ folgt $Ab_1 = \ldots = Ab_k = 0$ und

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

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

(b) Es gilt

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

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 $L_0 \to \widetilde{x} + L_0$, $v \mapsto \widetilde{x} + v$ 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

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

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.

(a) 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 $S \in \mathrm{GL}(n, K)$ mit $SA = B$.

(b) 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.

Satz 6.10 (Gauss-Algorithmus[25])**.** *Jede Matrix $A \in K^{n\times m}$ ist zeilen-äquivalent zu genau einer Matrix $\widehat{A}$ in* Zeilenstufenform[26]*, d. h.*

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

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

(1) Setze $z := 1$ (Zeilenindex).

(2) Für $s = 1, \ldots, m$ (Spaltenindex) tue:

- 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 $j = 1, \ldots, z-1, z+1, \ldots, n$ subtrahiere das a_{js}-Fache der z-ten Zeile von der j-ten Zeile. Anschließend gilt $a_{js} = 0$.
  - Erhöhe z um 1.

[25] Auch *Gauß-Elimination* oder *Gauß-Jordan-Algorithmus* genannt

[26] Jede 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.

Für die Eindeutigkeit von $\widehat{A}$ seien B und C Matrizen in Zeilenstufenform mit $A \sim B$ und $A \sim C$. Dann ist auch $B \sim C$ und es existiert $S \in \mathrm{GL}(n, K)$ 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 $e_1, \ldots, e_n$ die Standardbasis von K^n. Sei $b_i \in \langle e_1^{\mathrm{t}}, \ldots, e_k^{\mathrm{t}} \rangle$. Wir zeigen $b_i = c_i$ und $s_k = e_k^{\mathrm{t}}$ durch Induktion nach k. Im Fall $k = 0$ ist $b_i \in \langle \varnothing \rangle = \{0\}$, 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 $\langle e_1^{\mathrm{t}}, \ldots, e_{k-1}^{\mathrm{t}} \rangle$ liegt, ist $b_i = e_k^{\mathrm{t}}$ wegen der Zeilenstufenform. Die Spalten von S sind linear unabhängig, da S invertierbar ist. Dies zeigt

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

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

Beispiel 6.11.

$$\begin{pmatrix} 0 & 1 & 1 & 3 \\ 2 & 1 & 3 & 0 \\ 3 & -1 & 2 & 1 \end{pmatrix} \overset{\leftarrow}{\leftarrow} \sim \begin{pmatrix} 2 & 1 & 3 & 0 \\ 0 & 1 & 1 & 3 \\ 3 & -1 & 2 & 1 \end{pmatrix} \mid :2 \sim \begin{pmatrix} 1 & 1/2 & 3/2 & 0 \\ 0 & 1 & 1 & 3 \\ 3 & -1 & 2 & 1 \end{pmatrix} \overset{-3}{\underset{+}{\leftarrow}}$$

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

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

6.3 Anwendungen

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

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

Bemerkung 6.13. Zur Ermittelung 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 $\mathrm{rk}(A)$) interessiert, so darf man wegen $\mathrm{rk}(A) = \mathrm{rk}(A^{\mathrm{t}})$ 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!

$$\begin{pmatrix} 1 & -1 & 0 \\ 2 & 0 & 2 \\ 3 & -1 & 2 \\ -1 & 2 & 1 \end{pmatrix} \overset{+}{\sim} \begin{pmatrix} 1 & 0 & 0 \\ 2 & 2 & 2 \\ 3 & 2 & 2 \\ -1 & 1 & 1 \end{pmatrix} \overset{-1\ +}{\sim} \begin{pmatrix} 1 & 0 & 0 \\ 2 & 2 & 0 \\ 3 & 2 & 0 \\ -1 & 1 & 0 \end{pmatrix} \implies \operatorname{rk}(A) = 2$$

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

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

Ggf. erhält man die Lösungsmenge wie folgt: Seien $(1, n_1), \ldots, (k, n_k)$ die Positionen der führenden Einsen in $(\widehat{A}|c)$. Die $n-k$ Nullzeilen werden gestrichen. Für alle $i \in \{1, \ldots, m\} \setminus \{n_1, \ldots, n_k\}$ 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 $\widetilde{x}$ von M erfüllt $A\widetilde{x} = b$.

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

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

Beweis. Sei $S \in \mathrm{GL}(n, K)$ 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

$$(\widehat{A}|c) = \left(\begin{array}{ccccc|c} . & 1 & a_1 & . & a_2 & c_1 \\ . & . & . & 1 & a_3 & c_2 \\ . & . & . & . & . & . \end{array}\right) \qquad M = \begin{pmatrix} -1 & . & . & . & . & . \\ . & 1 & a_1 & . & a_2 & c_1 \\ . & . & -1 & . & . & . \\ . & . & . & 1 & a_3 & c_2 \\ . & . & . & . & -1 & . \end{pmatrix}$$

(hier steht . für 0 zur besseren Übersicht). Man sieht leicht, dass $\widehat{A}\widetilde{x} = c$ gilt. Somit gilt auch $A\widetilde{x} = b$. 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 $\dim L_0 = m - \operatorname{rk}(A) = m - \operatorname{rk}(\widehat{A}) = m - k = 3$. Dies zeigt (c) und aus Satz 6.6 folgt (d). □

Beispiel 6.16.

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

$$(A|b) = \left(\begin{array}{ccccc|c} -1 & 3 & 1 & 1 & 0 & 11 \\ -2 & 1 & -3 & 2 & -4 & -6 \\ 0 & 1 & 1 & 0 & 0 & 4 \end{array}\right) \sim \cdots \sim \begin{pmatrix} 1 & 0 & 2 & -1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{pmatrix}$$

$$M = \begin{pmatrix} 1 & 0 & 2 & -1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 4 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{pmatrix}$$

$$L = \widetilde{x} + L_0 = \begin{pmatrix} 1 \\ 4 \\ 0 \\ 0 \\ 2 \end{pmatrix} + \left\langle \begin{pmatrix} 2 \\ 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 0 \\ -1 \\ 0 \end{pmatrix} \right\rangle$$

Kontrolle:

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

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

Beweis. Es gilt

$$A \text{ invertierbar} \overset{5.15}{\iff} \operatorname{rk}(A) = n \overset{6.12}{\iff} \operatorname{rk}(\widehat{A}) = n \iff \widehat{A} = 1_n.$$

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

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

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

Beispiel 6.19.

$$(A|1_3) = \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ -1 & -1 & 1 & 0 & 1 & 0 \\ -1 & 1 & -2 & 0 & 0 & 1 \end{array}\right) \overset{\text{II}+\text{I}}{\underset{\text{III}+\text{I}}{}} \sim \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 & 1 & 0 \\ 0 & 1 & -3 & 1 & 0 & 1 \end{array}\right) \overset{\text{II}}{\underset{\text{III}}{}}$$

$$\sim \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & -3 & 1 & 0 & 1 \\ 0 & -1 & 0 & 1 & 1 & 0 \end{array}\right) \overset{}{}_{\text{III}+\text{II}} \sim \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & -3 & 1 & 0 & 1 \\ 0 & 0 & -3 & 2 & 1 & 1 \end{array}\right) \quad | : (-3)$$

$$\sim \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & -3 & 1 & 0 & 1 \\ 0 & 0 & 1 & -2/3 & -1/3 & -1/3 \end{array}\right) \overset{\text{I}+\text{III}}{\underset{\text{II}+3\,\text{III}}{}} \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1/3 & -1/3 & -1/3 \\ 0 & 1 & 0 & -1 & -1 & 0 \\ 0 & 0 & 1 & -2/3 & -1/3 & -1/3 \end{array}\right)$$

$$A^{-1} = \frac{1}{3}\begin{pmatrix} 1 & -1 & -1 \\ -3 & -3 & 0 \\ -2 & -1 & -1 \end{pmatrix}$$

Bemerkung 6.20. 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.

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*

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

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 *Spalten*operationen 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 $\operatorname{rk}(A) = \dim U + \dim W$. Die Dimensionsformel liefert daher

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

Also ist $\{c_1, \ldots, c_l\}$ 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:

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

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

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 $V \to W$, die Addition und Skalarmultiplikation „respektieren“. Es wird sich zeigen, dass solche Abbildungen durch Matrizen beschrieben werden können.

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

$$\boxed{f(\lambda u + v) = \lambda f(u) + f(v).}$$

Die Menge der linearen Abbildungen $V \to W$ wird mit $\operatorname{Hom}(V, W)$ 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 $f\colon V \to W$ ist genau dann linear, wenn

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

für alle $u, v \in V$ und $\lambda \in K$ gilt (setze $\lambda = 1$ bzw. $v = 0$ in Definition 7.2; vgl. Bemerkung 3.12). Isomorphe Vektorräume unterscheiden sich daher nur durch die Benennung ihrer Elemente.

(b) Für $f \in \operatorname{Hom}(V, W)$ gilt

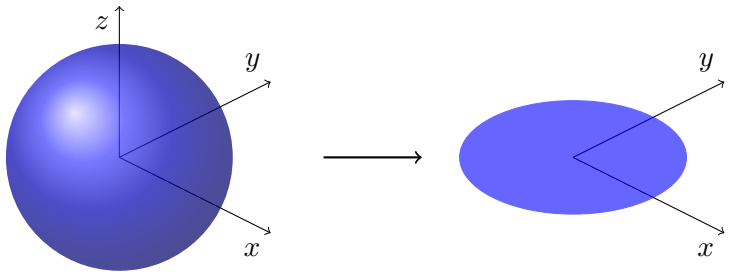
$$f(0_V) = f(0_K \cdot 0_V) = 0_K f(0_V) = 0_W.$$

Beispiel 7.4.

(a) Die *Nullabbildung* $0\colon V \to W$, $v \mapsto 0_W$ ist stets linear. Die Identität $\mathrm{id}_V\colon V \to V$ ist ein Isomorphismus.

(b) Für $f \in \mathrm{Hom}(V,W)$ und $U \leq V$ ist die Einschränkung $f_{|U}$ linear. Insbesondere ist die Inklusionsabbildung $U \to V$ als Einschränkung von id_V linear.

(c) Für $n \leq m$ ist die *Projektion* $K^m \to K^n$, $(x_1,\dots,x_m) \mapsto (x_1,\dots,x_n)$ ein surjektiver Homomorphismus. Die Projektion $\mathbb{R}^3 \to \mathbb{R}^2$ reduziert ein 3-dimensionales Objekt auf seinen „Schatten“:



(d) Sei $f\colon \mathbb{R} \to \mathbb{R}$ linear und $a := f(1)$. Für $x \in \mathbb{R}$ gilt $f(x) = f(x1) = xf(1) = ax$. 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*[27]).

(e) Für $A \in K^{n\times m}$ ist die Abbildung $K^{m\times 1} \to K^{n\times 1}$, $x \mapsto Ax$ nach Lemma 5.8 linear. Wir zeigen in Satz 7.16, dass jede lineare Abbildung (nach Basiswahl) diese Form besitzt.

(f) Die Transposition $K^{n\times m} \to K^{m\times n}$, $A \mapsto A^{\mathrm{t}}$ ist ein Isomorphismus.

Lemma 7.5. *Für $f \in \mathrm{Hom}(V,W)$, $V_1 \leq V$ 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 $\mathrm{Ker}(f) := f^{-1}(\{0\}) \leq V$.*

Beweis. Wegen $0 = f(0) \in f(V_1)$ ist $f(V_1) \neq \varnothing$. 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 3.12). Wegen $0 \in f^{-1}(\{0\}) \subseteq f^{-1}(W_1)$ ist auch $f^{-1}(W_1) \neq \varnothing$. 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 $f^{-1}(W_1) \leq V$. □

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

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

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

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

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

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

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 $f\colon V \to W$ mit $f(b_i) = c_i$ für $i = 1, \dots, n$. Dabei gilt:*

(a) f ist injektiv, genau dann wenn $c_1, \dots, c_n$ linear unabhängig sind.

(b) f ist surjektiv, genau dann wenn $W = \langle c_1, \dots, c_n \rangle$.

(c) f ist ein Isomorphismus, genau dann wenn $c_1, \dots, c_n$ eine Basis von W ist.

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

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

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

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

Aus Lemma 7.7 folgt $\sum_{i=1}^n \lambda_i b_i \in \mathrm{Ker}(f) = \{0\}$. 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 $u := \sum_{i=1}^n \lambda_i b_i \in \mathrm{Ker}(f)$. 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 $\mathrm{Ker}(f) = \{0\}$. 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). □

Bemerkung 7.9. Sei $f \in \mathrm{Hom}(V, W)$. 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 2.6(d)).

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 $f\colon V \to W$ 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 $\{b_1, \ldots, b_n\}$ und $\{c_1, \ldots, c_n\}$ von V bzw. W. Nach Satz 7.8 existiert ein Isomorphismus $f\colon V \to W$ mit $f(b_i) = c_i$ für $i = 1, \ldots, n$. Die zweite Behauptung folgt aus $\dim K^n = n$. Einen expliziten Isomorphismus erhält man durch die Koordinatendarstellung $V \to K^n$, $v \mapsto {}_B[v]$ (sie bildet B auf die Standardbasis von K^n ab). □

Beispiel 7.11. Obwohl $\mathbb{Q}$ und $\mathbb{Q}^2$ gleichmächtig sind, gilt $\mathbb{Q} \ncong \mathbb{Q}^2$ nach Satz 7.10.

7.2 Darstellungsmatrizen

Satz 7.12. *Für K-Vektorräume V und W ist auch* $\mathrm{Hom}(V, W)$ *ein K-Vektorraum via*

$$(f+g)(v) := f(v) + g(v),$$
$$(\lambda f)(v) := \lambda f(v)$$

$f, g \in \mathrm{Hom}(V, W)$, $\lambda \in K$ *und* $v \in V$.

Beweis. Der Beweis benutzt nur die Vektorraumaxiome von W. Seien $f, g, h \in \mathrm{Hom}(V, W)$, $u, v \in V$ und $\lambda, \mu \in K$. Wir müssen zuerst zeigen, dass die angegebenen Verknüpfungen wohldefiniert sind. Wegen

$$(f+g)(\mu u + v) = f(\mu u + v) + g(\mu u + v) = \mu f(u) + f(v) + \mu g(u) + g(v) = \mu(f+g)(u) + (f+g)(v),$$
$$(\lambda f)(\mu u + v) = \lambda f(\mu u + v) = \lambda(\mu f(u) + f(v)) = \mu\lambda f(u) + \lambda f(v) = \mu(\lambda f)(u) + (\lambda f)(v)$$

sind $f + g$ und λf in der Tat linear. Außerdem gilt

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

Dies zeigt $f + (g + h) = (f + g) + h$. Die Nullabbildung ist das neutrale Element bzgl. $+$. Die Abbildung $-f := (-1)f$ ist invers zu f bzgl. $+$. Schließlich gilt offenbar auch $f + g = g + f$. Somit ist $(\mathrm{Hom}(V, W), +)$ eine abelsche Gruppe. Die Regeln der Skalarmultiplikation überprüft man ebenso leicht. □

Definition 7.13. Seien V und W Vektorräume mit Basen $B = \{b_1, \ldots, b_m\}$ bzw. $C = \{c_1, \ldots, c_n\}$. Sei $f \in \mathrm{Hom}(V, W)$ und $f(b_i) = \sum_{j=1}^{n} a_{ji} c_j$ mit $a_{ji} \in K$ für $i = 1, \ldots, m$.

(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 $f = \mathrm{id}_V$ nennt man ${}_C\Delta_B := {}_C[\mathrm{id}_V]_B$ 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 7.14. In der Situation von Definition 7.13 gilt ${}_B\Delta_B = 1_m$.

Beispiel 7.15. Die Abbildung

$$f\colon \mathbb{R}^2 \to \mathbb{R}^3, \qquad (x,y) \mapsto (2x-y, y, -3x)$$

ist linear mit Matrix

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

Offenbar sind $B := \big\{(1,-1),(0,2)\big\}$ und $C := \big\{(1,1,1),(0,-1,1),(1,0,1)\big\}$ 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 7.16. *Sei V ein m-dimensionaler Vektorraum mit Basis B und sei W ein n-dimensionaler Vektorraum mit Basis C. Dann ist die Abbildung*

$${}_C[.]_B\colon \operatorname{Hom}(V,W) \to K^{n\times m}, \qquad f \mapsto {}_C[f]_B$$

ein Isomorphismus mit

$$\boxed{{}_C[f(v)]^{\mathrm{t}} = {}_C[f]_B\,{}_B[v]^{\mathrm{t}}}$$

$$\begin{array}{ccc} V & \xrightarrow{\ f\ } & W \\ {\scriptstyle {}_B[.]}\downarrow & & \downarrow{\scriptstyle {}_C[.]} \\ K^m & \xrightarrow[{}_C[f]_B]{} & K^n \end{array}$$

für alle $v \in V$. Insbesondere ist $\operatorname{rk}(f) = \operatorname{rk}({}_C[f]_B)$.

Beweis. Sei $B = \{b_1,\dots,b_m\}$ und $C = \{c_1,\dots,c_n\}$. Nach Satz 7.8 ist ${}_C[.]_B$ eine Bijektion. Seien $f, g \in \operatorname{Hom}(V,W)$ 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 ${}_C[.]_B$ ein Isomorphismus. Sei $v = \sum_{i=1}^{m} v_i b_i$. Dann ist

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

und es folgt

$${}_C[f]_B\,{}_B[v]^{\mathrm{t}} = \Big(\sum_{i=1}^{m} a_{ji}v_i\Big)_j^{\mathrm{t}} = {}_C[f(v)]^{\mathrm{t}}.$$

Offenbar ist ${}_B[b_i] = e_i$ der i-te Standardbasisvektor. Also ist ${}_C[f(b_i)] = {}_C[f]_{BB}[b_i]^{\mathrm{t}}$ die i-te Spalte von ${}_C[f]_B$. Da ${}_C[.]$ ein Isomorphismus ist, gilt

$$\mathrm{rk}(f) = \dim\langle f(b_1), \ldots, f(b_m)\rangle = \dim\langle {}_C[f(b_1)], \ldots, {}_C[f(b_m)]\rangle = \mathrm{rk}({}_C[f]_B). \qquad \square$$

Bemerkung 7.17. Für $V = W$ und $f = \mathrm{id}_V$ erhält man ${}_C[v]^{\mathrm{t}} = {}_C\Delta_{BB}[v]^{\mathrm{t}}$. Für $f\colon K^n \to K^m$ gilt $f(v)^{\mathrm{t}} = [f]v^{\mathrm{t}}$ bzgl. der Standardbasen.

Beispiel 7.18. Seien f, B und C wie in Beispiel 7.15. Für $v := (2,4) = 2(1,-1) + 3(0,2) \in \mathbb{R}^2$ gilt

$$f(v)^{\mathrm{t}} = [f]v^{\mathrm{t}} = \begin{pmatrix} 2 & -1 \\ 0 & 1 \\ -3 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \\ -6 \end{pmatrix},$$
$${}_C[f(v)]^{\mathrm{t}} = {}_C[f]_{BB}[v]^{\mathrm{t}} = \begin{pmatrix} -7 & 4 \\ -6 & 2 \\ 10 & -6 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} -2 \\ -6 \\ 2 \end{pmatrix}.$$

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

Satz 7.19. *Seien U, V und W Vektorräume mit Basen B, C bzw. D. Seien $f \in \mathrm{Hom}(U,V)$ und $g \in \mathrm{Hom}(V,W)$. Dann ist $g \circ f \in \mathrm{Hom}(U,W)$ mit*

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

Beweis. Für $u, u' \in U$ und $\lambda \in K$ gilt

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

d. h. $g \circ f \in \mathrm{Hom}(U,W)$. Sei $B = \{b_1, \ldots, b_m\}$, $C = \{c_1, \ldots, c_n\}$ und $D = \{d_1, \ldots, d_k\}$. Sei ${}_C[f]_B = (a_{ij})$ und ${}_D[g]_C = (b_{ij})$. Dann gilt

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

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

Bemerkung 7.20. Merkregel: Die Komposition von linearen Abbildungen entspricht der Multiplikation von Matrizen.

Beispiel 7.21. Seien f, B und C wie in Beispiel 7.15. Sei $g \in \mathrm{Hom}(\mathbb{R}^3, \mathbb{R}^2)$ mit Matrix ${}_B[g]_C = -\left(\begin{smallmatrix} 3 & 0 & 2 \\ 1 & 1/2 & 1 \end{smallmatrix}\right)$. Dann gilt

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

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

Folgerung 7.22. *Sei $f\colon V \to W$ ein Isomorphismus zwischen Vektorräumen V und W mit Basen B bzw. C. Dann ist $f^{-1}\colon W \to V$ ein Isomorphismus mit* $\boxed{{}_B[f^{-1}]_C = {}_C[f]_B^{-1}.}$ *Im Fall $V = W$ ist* ${}_C\Delta_B^{-1} = {}_B\Delta_C$.

Beweis. Für $u, w \in W$ und $\lambda \in K$ gilt

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

d. h. $f^{-1} \in \operatorname{Hom}(W, V)$. Aus Satz 7.19 folgt

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

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

Bemerkung 7.23. Die Isomorphismen $f\colon V \to V$ bilden die *allgemeine lineare Gruppe* $\mathrm{GL}(V)$. Sie entsprechen genau den invertierbaren Matrizen, d. h. ${}_B[.]_B\colon \mathrm{GL}(V) \to \mathrm{GL}(n, K)$ ist ein Isomorphismus von Gruppen (anstatt von Vektorräumen).

Folgerung 7.24 (Basiswechsel)**.** *Seien B, B' Basen von V und C, C' Basen von W. Für $f \in \operatorname{Hom}(V, W)$ gilt*

$$\boxed{{}_{C'}[f]_{B'} = {}_{C'}\Delta_C\,{}_C[f]_B\,{}_B\Delta_{B'}.}$$

Im Fall $V = W$ ist

$$\boxed{{}_{B'}[f]_{B'} = {}_{B'}\Delta_B\,{}_B[f]_{B'}\,\Delta_B^{-1}.}$$

Beweis. Aus Satz 7.19 folgt

$${}_{C'}\Delta_C\,{}_C[f]_B\,{}_B\Delta_{B'} = {}_{C'}[\mathrm{id}_W]_C\,{}_C[f]_B\,{}_B[\mathrm{id}_V]_{B'} = {}_{C'}[\mathrm{id}_W]_C\,{}_C[f]_{B'} = {}_{C'}[f]_{B'}.$$

Für $V = W$, $C = B$ und $C' = B'$ erhält man ${}_{B'}[f]_{B'} = {}_{B'}\Delta_B\,{}_B[f]_B\,{}_B\Delta_{B'} = {}_{B'}\Delta_B\,{}_B[f]_{B'}\,\Delta_B^{-1}$ mit Folgerung 7.22. □

Beispiel 7.25. Sei wieder $f\colon \mathbb{R}^2 \to \mathbb{R}^3$ wie in Beispiel 7.15. Wir ersetzen die Basis $B = \big\{(1,-1),(0,2)\big\}$ durch $B' := \big\{(0,1),(1,1)\big\}$. Wegen $(0,1) = 0(1,-1) + \frac{1}{2}(0,2)$ und $(1,1) = (1,-1) + (0,2)$ gilt

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

Bemerkung 7.26. Mit den obigen Sätzen kann abstrakte lineare Abbildungen durch konkrete Matrizen ersetzen, indem man geeignete Basen wählt. Im nächsten Semester konstruieren wir spezielle Basen, sodass die Darstellungsmatrizen möglichst „einfache“ Gestalt haben (zum Beispiel Diagonalmatrizen). Dies beschleunigt Berechnungen.

Satz 7.27 (Rangsatz)**.** *Für $f \in \operatorname{Hom}(V, W)$ gilt*

$$\boxed{\dim V = \operatorname{rk}(f) + \dim \operatorname{Ker}(f).}$$

Beweis. Seien B und C Basen von V bzw. W mit $|B| = m$ und $|C| = n$. Nach Satz 7.16 ist $\operatorname{rk}(f) = \operatorname{rk}({}_C[f]_B)$ und

$$\operatorname{Ker}(f) = \{v \in V : f(v) = 0\} = \{v \in V : {}_C[f]_B [v]^{\mathrm{t}} = 0\} \cong \{x \in K^{m\times 1} : {}_C[f]_B x = 0\}.$$

Aus Satz 6.6 folgt $\dim \operatorname{Ker}(f) = m - \operatorname{rk}({}_C[f]_B) = \dim V - \operatorname{rk}(f)$. □

Definition 7.28. Für $A = (a_{ij})_{i,j} \in K^{n\times n}$ nennen wir $\operatorname{tr}(A) := \sum_{i=1}^{n} a_{ii}$ die *Spur* von A. Dies ist die Summe der Hauptdiagonaleinträge.

Lemma 7.29. *Die Abbildung* $\operatorname{tr}\colon K^{n\times n} \to K$ *ist linear mit* $\operatorname{tr}(A^{\mathrm{t}}) = \operatorname{tr}(A)$ *und* $\operatorname{tr}(AB) = \operatorname{tr}(BA)$ *für* $A, B \in K^{n\times n}$.

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

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

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

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

Bemerkung 7.30. Sei V ein Vektorraum mit Basis B und $f \in \operatorname{Hom}(V, V)$. Dann nennen wir $\operatorname{tr}(f) := \operatorname{tr}({}_B[f]_B)$ die *Spur* von f. Dies hängt nicht von der Wahl von B ab, denn ist auch B' eine Basis von V, so gilt

$$\operatorname{tr}({}_{B'}[f]_{B'}) \overset{7.24}{=} \operatorname{tr}\big(({}_{B'}\Delta_{BB}[f]_B)_{B'}\Delta_B^{-1}\big) \overset{7.29}{=} \operatorname{tr}\big({}_{B'}\Delta_B^{-1}({}_{B'}\Delta_{BB}[f]_B)\big) = \operatorname{tr}({}_B[f]_B).$$

Aufgaben

Die folgenden Aufgaben ergänzen die Vorlesungsinhalte und sind von den Hausübungen unabhängig.

Aufgabe 1. Welche der folgenden Aussagen sind wahr (bezogen auf das Jahr 2021)?

(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 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 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 4. Beweisen Sie mit vollständiger Induktion: Die Summe der ersten n ungeraden Zahlen ist n^2.

Aufgabe 5. Zeigen Sie, dass $\mathbb{N}$ und $\mathbb{N} \times \mathbb{N}$ gleichmächtig sind.

Aufgabe 6. Die *Potenzmenge* einer Menge M ist definiert durch

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

(a) Konstruieren Sie eine injektive Abbildung $M \to \mathcal{P}(M)$.

(b) Zeigen Sie, dass $|\mathcal{P}(M)| = 2^{|M|}$ gilt, falls M endlich ist.

(c) Zeigen Sie, dass M und $\mathcal{P}(M)$ nicht gleichmächtig sind (d. h. $\mathcal{P}(M)$ ist „größer“ als M).
Hinweis: Nehmen Sie an, dass eine Bijektion $f\colon M \to \mathcal{P}(M)$ existiert und betrachten Sie $A := \{x \in M : x \notin f(x)\}$.

Aufgabe 7. Sei $U := \{2z + 1 : z \in \mathbb{Z}\} \cup \{0\}$. Untersuchen Sie, ob $(U, +)$ eine Gruppe ist.

Aufgabe 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 9. Konstruieren Sie einen Körper mit drei Elementen.
Hinweis: Was ist $1 + 1$?

Aufgabe 10. Zeigen Sie, dass 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 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 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 13. 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 14. 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 15. Die direkte Berechnung eines Produkts zweier $n \times n$-Matrizen $A = (a_{ij})$ und $B = (b_{ij})$ erfordert n^3 Multiplikationen von Körperelementen (nämlich $a_{ik}b_{kj}$ für $1 \leq i, j, k \leq n$). STRASSEN fand einen Weg wie man mit weniger Multiplikationen auskommt. Wir betrachten dafür den Fall $n = 2$ und die folgenden sieben Produkte:

$$
\begin{aligned}
&c_1 := (a_{11} + a_{22})(b_{11} + b_{22}), && c_2 := (a_{21} + a_{22})b_{11}, && c_3 := a_{11}(b_{12} - b_{22}),\\
&c_4 := a_{22}(b_{21} - b_{11}), && c_5 := (a_{11} + a_{12})b_{22}, && c_6 := (a_{21} - a_{11})(b_{11} + b_{12}),\\
&c_7 := (a_{12} - a_{22})(b_{21} + b_{22}).
\end{aligned}
$$

(a) Der erste Eintrag von AB lässt sich in der Form

$$a_{11}b_{11} + a_{12}b_{21} = c_1 + c_4 - c_5 + c_7$$

darstellen. Schreiben Sie auch die drei verbleibenden Einträge von AB als Summen/Differenzen von $c_1, \ldots, c_7$. Gegenüber der direkten Matrizenmultiplikation hat man also eine Multiplikation eingespart. Die Addition von Körperelementen ist aus computertheoretischer Sicht vernachlässigbar.

(b) Sei nun $n = 2k$. Führen Sie die Multiplikation AB auf sieben Matrixprodukte vom Format $k \times k$ zurück.

(c) Wie könnte man vorgehen, wenn n ungerade ist?

Optimierungen des Strassen-Algorithmus für $n \times n$-Matrizen benötigen nur noch ca. $n^{2.37286}$ Multiplikationen von Körperelementen.[28]

Aufgabe 16. Seien $A \in K^{n \times m}$, $B \in K^{m \times k}$ und $C \in K^{k \times l}$. Zur Berechnung von ABC kann man entweder $(AB)C$ oder $A(BC)$ benutzen. Bestimmen Sie für beide Klammerungen die Anzahl der benötigten Multiplikationen in Abhängigkeit von n, m, k, l (ohne Berücksichtigung von Aufgabe 15). Welche Variante ist zu bevorzugen, wenn $\frac{1}{n} + \frac{1}{k} < \frac{1}{m} + \frac{1}{l}$?

[28] siehe Quanta Magazine, 23.03.2021

Aufgabe 17. Zeigen Sie:

(a) Das Vertauschen von zwei Zeilen einer Matrix lässt sich auch durch die beiden anderen elementaren Zeilenoperationen realisieren.

(b) Durch elementare Zeilen- und Spaltenoperationen lässt sich jede Matrix $A \in K^{n\times m}$ in die Form

$$\begin{pmatrix} 1_r & 0_{r\times(m-r)} \\ 0_{(n-r)\times r} & 0_{(n-r)\times(m-r)} \end{pmatrix}.$$

überführen. Dabei ist $r = \operatorname{rk}(A)$.

(c) 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 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 19. Seien $f\colon U \to V$, $g\colon V \to W$ und $h\colon W \to X$ lineare Abbildungen.

(a) Zeigen Sie $\operatorname{rk}(g \circ f) + \operatorname{rk}(h \circ g) \leq \operatorname{rk}(g) + \operatorname{rk}(h \circ g \circ f)$ (Frobenius-Ungleichung).
Hinweis: Für $g(V) = g(f(U)) \oplus Y$ gilt $h(g(V)) = h\big(g(f(U))\big) + h(Y)$.

(b) Folgern Sie Lemma 5.15(a) aus Teil (a).

(c) Zeigen Sie $\operatorname{rk}(A) + \operatorname{rk}(B) \leq \operatorname{rk}(AB) + n$ für $A \in K^{m\times n}$ und $B \in K^{n\times k}$ (Sylvester-Ungleichung).

Aufgabe 20. Zeigen Sie:

(a) Die Summe und das Produkt von oberen (bzw. unteren) Dreiecksmatrizen sind wieder obere (bzw. untere) Dreiecksmatrizen (siehe Definition 8.14).

(b) Die Menge der invertierbaren oberen (bzw. unteren) Dreiecksmatrizen ist eine Gruppe bzgl. Multiplikation.

Aufgabe 21. Zeigen Sie, dass $A \in K^{n\times n}$ genau dann eine Skalarmatrix ist, wenn $AB = BA$ für alle $B \in K^{n\times n}$ gilt.

Aufgabe 22. Sei $A \in \mathrm{GL}(n, K)$. Zeigen Sie, dass man die Matrix $\left(\begin{smallmatrix} A \\ 1_n \end{smallmatrix}\right) \in K^{2n\times n}$ durch elementare Spaltenoperationen in die Form $\left(\begin{smallmatrix} 1_n \\ B \end{smallmatrix}\right)$ überführen kann. Dabei ist $B = A^{-1}$.

Aufgabe 23. Seien U und W Unterräume eines Vektorraums V mit $U \cap W = \{0\}$. Konstruieren Sie einen Isomorphismus $U \times W \to U \oplus W$.

Aufgabe 24. Seien V und W Vektorräume und $f \in \operatorname{Hom}(V, W)$ injektiv. Zeigen Sie, dass eine Abbildung $g \in \operatorname{Hom}(W, V)$ mit $g \circ f = \mathrm{id}_V$ existiert. Ist g eindeutig bestimmt?

Aufgabe 25. Sei V ein K-Vektorraum mit Basis $b_1, \ldots, b_n$. Man nennt $V^* := \mathrm{Hom}(V, K)$ den *Dualraum* von V. Für $i = 1, \ldots, n$ sei $b_i^* \in V^*$ mit $b_i^*(b_j) = \delta_{ij}$. Zeigen Sie:

(a) $b_1^*, \ldots, b_n^*$ ist eine Basis von V^*. Man sie *duale* Basis bzgl. $b_1, \ldots, b_n$.

(b) Für $f \in \mathrm{Hom}(V, W)$ sei $f^* \colon W^* \to V^*$, $\varphi \mapsto \varphi \circ f$ die zu f *duale* Abbildung. Dann gilt $f^* \in \mathrm{Hom}(W^*, V^*)$ und ${}_{B^*}[f^*]_{C^*} = {}_C[f]_B^{\mathrm{t}}$, wobei B^* und C^* die zu B bzw. C dualen Basen sind. *Bemerkung:* Dies ist eine mögliche Interpretation der transponierten Matrix.

Lineare Algebra II

8 Eigenwerte und Eigenvektoren

8.1 Definitionen und Beispiele

Bemerkung 8.1. Im vergangenen Semester haben wir lineare Abbildungen $f\colon V \to W$ zwischen endlichdimensionalen K-Vektorräumen V und W studiert. In diesem Semester nehmen wir stets $V = W$ an. Man spricht dann von *Endomorphismen* anstatt Homomorphismen und schreibt $\operatorname{End}(V) := \operatorname{Hom}(V, V)$. 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 $f \in \operatorname{End}(V)$. Man nennt $\lambda \in K$ einen *Eigenwert* 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

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

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

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

$$f(v) = \lambda v \overset{7.16}{\iff} Ax = \lambda x \iff (A - \lambda 1_n)x = 0.$$

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.

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

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

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

[1] Man findet stets Basen B, C von V, sodass ${}_C[f]_B = \operatorname{diag}(1, \dots, 1, 0, \dots, 0)$ gilt. Das Produkt solcher Matrizen lässt sich allerdings nicht sinnvoll interpretieren.

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 ${}_B[f]_B = \operatorname{diag}(1, 1, 4)$. Damit berechnet man leicht ${}_B[f \circ f \circ f]_B = {}_B[f]_B^3 = \operatorname{diag}(1, 1, 4)^3 = \operatorname{diag}(1, 1, 64)$.

8.2 Diagonalisierbarkeit

Definition 8.5. Man nennt $f \in \operatorname{End}(V)$ *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 eine Matrix $S \in \operatorname{GL}(n, K)$ existiert, sodass $S^{-1}AS$ eine Diagonalmatrix ist.

Bemerkung 8.6. Offenbar ist $f \in \operatorname{End}(V)$ genau dann diagonalisierbar, wenn V eine Basis aus Eigenvektoren von f besitzt. Eine Matrix A ist genau dann diagonalisierbar, wenn die entsprechende lineare Abbildung $f\colon K^{n\times 1} \to K^{n\times 1}$, $x \mapsto Ax$ diagonalisierbar ist (Folgerung 7.24).

Beispiel 8.7.

(a) Die Abbildung aus Beispiel 8.4 ist diagonalisierbar.

(b) Diagonalmatrizen sind offensichtlich diagonalisierbar (wähle $S = 1_n$).

(c) Sei $A := \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right) \in K^{2\times 2}$. 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 $A := \left(\begin{smallmatrix} 0 & 2 \\ 1 & 0 \end{smallmatrix}\right) \in \mathbb{Q}^{2\times 2}$ und $\lambda \in \mathbb{Q}$. Dann ist

$$A - \lambda 1_2 = \begin{pmatrix} -\lambda & 2 \\ 1 & -\lambda \end{pmatrix} \begin{matrix} \leftarrow\!\urcorner^{+} \\ \lrcorner_{\lambda} \end{matrix} \sim \begin{pmatrix} 0 & 2 - \lambda^2 \\ 1 & -\lambda \end{pmatrix}.$$

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 $U_1, \ldots, U_n \leq V$ definiert man induktiv

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

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

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

(1) $U_1 + \ldots + U_n = U_1 \oplus \ldots \oplus U_n$.

(2) $\dim(U_1 + \ldots + U_n) = \dim(U_1) + \ldots + \dim(U_n)$.

(3) Jedes Element $u \in U_1 + \ldots + U_n$ lässt sich eindeutig in der Form $u = u_1 + \ldots + u_n$ mit $u_i \in U_i$ für $i = 1, \ldots, n$ schreiben.

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

Beweis.

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

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

(2)⇒(3): Sei B_i eine Basis von U_i für $i = 1, \ldots, n$. Dann ist $B := B_1 \cup \ldots \cup B_n$ ein Erzeugendensystem von $U_1 + \ldots + U_n$. Wegen

$$|B| \leq |B_1| + \ldots + |B_n| = \dim(U_1) + \ldots + \dim(U_n) \overset{(2)}{=} \dim(U_1 + \ldots + U_n) \leq |B|$$

muss B auch eine Basis von $U_1 + \ldots + U_n$ sein. Insbesondere ist jedes $u \in U_1 + \ldots + U_n$ eine eindeutige Linearkombination von Elementen aus B. Daraus folgt auch die Eindeutigkeit der Zerlegung $u = u_1 + \ldots + u_n$ mit $u_i \in U_i$.

(3)⇒(4): Die beiden Zerlegungen des Nullvektors $u_1 + \ldots + u_n = 0 + \ldots + 0$ müssen nach (3) identisch sein, d. h. $u_1 = \ldots = u_n = 0$.

(4)⇒(1): Für $n = 1$ ist nichts zu zeigen. Induktiv können wir $U_1 + \ldots + U_{n-1} = U_1 \oplus \ldots \oplus U_{n-1}$ annehmen, denn die Voraussetzung (4) überträgt sich auf $U_1, \ldots, U_{n-1}$. Sei nun $u = u_1 + \ldots + u_{n-1} \in (U_1 + \ldots + U_{n-1}) \cap U_n$. Dann ist

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

und (4) zeigt $u = 0$. Also gilt (1). □

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

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

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 $v_1 + \ldots + v_k = 0$. Dann gilt

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

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

Bemerkung 8.11. Merkregel: 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 $\lambda_1, \ldots, \lambda_n$ von A gilt

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

nach Satz 8.10. Dies zeigt $E_{\lambda_1}(A) \oplus \ldots \oplus E_{\lambda_n}(A) = K^{n \times 1}$. 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 10.28).

(b) Die Matrix $A := \left(\begin{smallmatrix} 0 & 2 \\ 1 & 0 \end{smallmatrix}\right) \in \mathbb{R}^{2\times 2}$ 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 10.26).

Definition 8.14. Man nennt $A = (a_{ij}) \in K^{n\times n}$ eine (*obere*[2]) *Dreiecksmatrix*, falls alle Einträge unterhalb der Hauptdiagonalen verschwinden, d. h. $a_{ij} = 0$ für alle $i > j$:

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

Beispiel 8.15. Sei $A = (a_{ij}) \in K^{n\times n}$ eine obere Dreiecksmatrix. Für $\lambda \in \{a_{11}, \ldots, a_{nn}\}$ hat $A - \lambda 1_n$ nicht vollen Rang, denn beim Gauß-Algorithmus tritt ein Versatz der Zeilen auf (Bemerkung 6.20). Ist andererseits $\lambda \notin \{a_{11}, \ldots, a_{nn}\}$, 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 Hauptdiagonalen sind. Insbesondere ist A diagonalisierbar, wenn $a_{11}, \ldots, a_{nn}$ paarweise verschieden sind.

9 Determinanten

9.1 Rekursive Definition

Bemerkung 9.1.

(a) Mathematiker versuchen oft komplizierte Objekte (wie $f \in \mathrm{End}(V)$) durch einfachere (wie $\mathrm{rk}(f)$ oder $\mathrm{tr}(f)$) zu ersetzen, um wesentliche Informationen sichtbar zu machen. So haben wir in Lemma 5.15 gesehen, dass $\mathrm{rk}(f)$ 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 $\mathrm{rk}(f)$ hängt diese Kriterium nicht mehr von $\dim(V)$ ab.[3])

(b) In der Maßtheorie versucht man möglichst vielen Mengen $S \subseteq \mathbb{R}^n$ ein „Volumen“ $\mathrm{vol}(S) \in \mathbb{R}_{\geq 0}$ zuzuordnen. Sei $f\colon \mathbb{R}^n \to \mathbb{R}^n$ linear. Die in (a) beschriebene Zahl $\det(f)$ misst wie sehr sich das Volumen durch Anwenden von f verändert, d. h. es gilt $\mathrm{vol}(f(S)) = |\det(f)|\,\mathrm{vol}(S)$ sofern $\mathrm{vol}(S)$ definiert ist. Dem n-dimensionalen *Hyperwürfel*

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

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

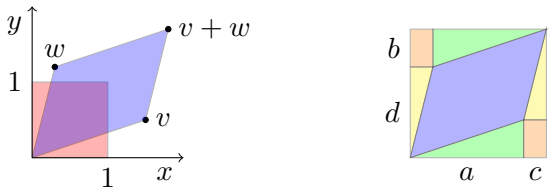
$$|\det(f)| = \mathrm{vol}(f(H)).$$

[2] Analog definiert man *untere* Dreiecksmatrizen.

[3] Aus 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.

Das Vorzeichen von $\det(f)$ beschreibt, ob f orientierungserhaltend (Beispiel: Drehung) oder orientierungsumkehrend (Beispiel: Spiegelung) ist. Mehr dazu in Beispiel 11.21.

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



Die Fläche von $f(H)$ ist

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

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

Definition 9.3. Sei $A = (a_{ij}) \in K^{n\times n}$ und $1 \le s, t \le n$. 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*[4] von A ist rekursiv definiert:

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

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, sollten wir untersuchen wie sich die Determinante bei elementaren Zeilenoperationen verändert.

[4] In manchen Büchern schreibt man $|A|$ anstelle von $\det(A)$. Das kann aber mit einer Matrixnorm verwechselt werden.

Lemma 9.5. *Die Abbildung* $\det\colon K^{n\times n} \to K$ *ist linear in jeder Zeile, d. h. für* $a_1, \dots, a_n, b \in K^n$, $\lambda \in K$ *und* $1 \le k \le n$ *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

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

wie behauptet. Sei nun $n \ge 2$ 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

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

□

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 $n \ge 2$. 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 $a_s \leftrightarrow a_{s+1} \leftrightarrow a_{s+2} \leftrightarrow \ldots \leftrightarrow a_{t-1}$:

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

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

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

Es genügt also $\det(A) = 0$ zu zeigen, falls A zwei identische Zeilen besitzt.[5] 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

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

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

Beispiel 9.7.

$$\det\begin{pmatrix}-4 & -2 & -2\\ 6 & 3 & 2\\ 8 & 7 & 6\end{pmatrix} \mid :(-2) = -2\det\begin{pmatrix}2 & 1 & 1\\ 6 & 3 & 2\\ 8 & 7 & 6\end{pmatrix} \begin{matrix}\text{—}\!\!\rfloor^{-3}\\ \leftarrow\!\!\rfloor_{+}\\ \leftarrow\text{———}\!\!\rfloor_{+}\end{matrix}^{-4} = -2\det\begin{pmatrix}2 & 1 & 1\\ 0 & 0 & -1\\ 0 & 3 & 2\end{pmatrix} \begin{matrix}\leftarrow\\ \leftarrow\!\!\rfloor\end{matrix}$$

$$= 2\det\begin{pmatrix}2 & 1 & 1\\ 0 & 3 & 2\\ 0 & 0 & -1\end{pmatrix} \overset{9.4}{=} -12$$

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.

9.2 Eigenschaften

Satz 9.9. *Für $A \in K^{n\times n}$ gilt*

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

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

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

Satz 9.10 (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 9.9. Wir können also $A \in \mathrm{GL}(n, K)$ 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

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

nach Satz 9.6. Insgesamt folgt

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

Folgerung 9.11.

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

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

Beweis.

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

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

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

Bemerkung 9.12. Wegen $\det(A^{\mathrm{t}}) = \det(A)$ darf man bei der Berechnung von $\det(A)$ auch elementare Spaltenoperationen benutzen.

Definition 9.13. Sei B eine Basis des K-Vektorraums V. Für $f \in \operatorname{End}(V)$ definieren wir die *Determinante* von f durch $\det(f) := \det({}_B[f]_B)$. Dies hängt nicht von der Wahl von B ab, denn ist auch B' eine Basis von V, so gilt

$$\det({}_{B'}[f]_{B'}) \overset{7.24}{=} \det({}_{B'}\Delta_B \, {}_B[f]_B \, {}_{B'}\Delta_B{}^{-1}) = \det({}_{B'}\Delta_B)\det({}_B[f]_B)\det({}_{B'}\Delta_B)^{-1} = \det({}_B[f]_B).$$

9.3 Laplace-Entwicklung

Satz 9.14 (Laplace-Entwicklung)**.** *Sei $n \geq 2$ und $A \in K^{n\times n}$. Für $1 \leq k \leq n$ gilt*

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

Beweis. Seien $a_1, \dots, a_n$ die Spalten von A. Nach Bemerkung 9.12 gilt

$$\begin{aligned}\det\big(\cdots \overset{\downarrow}{a_{k-1}} \overset{\downarrow}{a_k} \cdots\big) &= -\det\big(\cdots \overset{\downarrow}{a_{k-2}} \overset{\downarrow}{a_k} a_{k-1} \cdots\big) = \dots \\ &= (-1)^{k-1} \det\big(a_k \;\; a_1 \;\; \cdots \;\; a_{k-1} \;\; a_{k+1} \;\; \cdots\big) = (-1)^{k-1} \sum_{i=1}^{n} (-1)^{i+1} a_{ik} \det(A_{ik}).\end{aligned}$$

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

Bemerkung 9.15. Die Gleichungen in Satz 9.14 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 9.16. Wie die meisten rekursiven Verfahren ist auch die Laplace-Entwicklung in der Regel ineffizient. Sie eignet sich jedoch für sogenannte *dünnbesetzte* Matrizen, d. h. viele Einträge sind 0. Wir entwickeln zuerst nach der dritten Zeile und anschließend nach der zweiten Spalte:

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

Definition 9.17. Für $A \in K^{n\times n}$ nennt man

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

die zu A *komplementäre* Matrix.[6]

Satz 9.18. *Für alle $A \in K^{n\times n}$ gilt* $\boxed{A\widetilde{A} = \det(A)1_n = \widetilde{A}A.}$ *Insbesondere ist* $\boxed{A^{-1} = \dfrac{1}{\det(A)}\widetilde{A},}$ *falls $A \in \mathrm{GL}(n, K)$.*

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 $A\widetilde{A} = (c_{ij})$. 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 $A\widetilde{A} = \det(A)1_n$. Die Gleichung $\widetilde{A}A = \det(A)1_n$ zeigt man analog durch Entwicklung nach einer Spalte. □

Beispiel 9.19. Für jede invertierbare 2×2-Matrix $A := \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ erhält man

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

Bemerkung 9.20. Die Formel $A^{-1} = \frac{1}{\det(A)}\widetilde{A}$ hat für „große“ n 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 $\det(A)A^{-1} = \widetilde{A} \in \mathbb{Z}^{n\times n}$. 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 9.21 (Cramersche Regel)**.** *Sei $A \in \mathrm{GL}(n, K)$ und $b \in K^{n\times 1}$. Für $k = 1, \ldots, n$ sei A_k die Matrix, die aus A entsteht, indem man die k-te Spalte durch b ersetzt. Für die eindeutige Lösung $x = (x_1, \ldots, x_n)^{\mathrm{t}}$ des Gleichungssystems $Ax = b$ gilt dann $x_k = \frac{\det A_k}{\det A}$ für $k = 1, \ldots, n$.*

Beweis. Wir benutzen Satz 9.18 und entwickeln A_k nach der k-ten Spalte:

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

[6] auch *Adjunkte* genannt; Verwechslungsgefahr mit der *adjungierten Matrix*

9.4 Die Leibniz-Formel

Bemerkung 9.22. Führt man die Laplace-Entwicklung für $n \times n$-Matrizen bis auf 1×1-Matrizen zurück, so erhält man die Determinante als Summe von $n(n-1) \cdot \ldots \cdot 2 \cdot 1 = n!$ Termen. Wir bestimmen diese Terme explizit.

Definition 9.23. Sei $n \in \mathbb{N}$ und $N := \{1, \ldots, n\}$. Eine Bijektion der Form $N \to N$ heißt *Permutation* von N. Die Menge der Permutationen von N wird mit S_n bezeichnet. 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 $\operatorname{sgn}(\sigma) := \det(P_\sigma)$ das *Signum* oder *Vorzeichen* von σ.

Bemerkung 9.24.

(a) Analog zu $\mathrm{GL}(V)$ ist auch S_n eine Gruppe bzgl. Komposition von Abbildungen. Man nennt S_n die *symmetrische Gruppe von Grad n*.

(b) Sei $\sigma \in S_n$. Für die Wahl von $\sigma(1) \in \{1, \ldots, n\}$ 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. $|S_n| = n!$.

(c) Die Permutationsmatrix von σ entsteht, indem man die Zeilen der Einheitsmatrix (also die Standardbasis $e_1, \ldots, e_n$) 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 also $\operatorname{sgn}(\sigma) \in \{\pm 1\}$ tatsächlich ein „Vorzeichen“.

Beispiel 9.25. Wir geben die $3! = 6$ Elemente aus S_3 in der Form $\sigma(1)\sigma(2)\sigma(3)$ an:

σ	id = 123	132	213	231	312	321
P_σ	1_3	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$
$\operatorname{sgn}(\sigma)$	1	-1	-1	1	1	-1

Satz 9.26. *Für $\sigma, \tau \in S_n$ gilt $P_{\sigma\circ\tau} = P_\sigma P_\tau$ und* $\boxed{\operatorname{sgn}(\sigma \circ \tau) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau).}$

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. □

Beispiel 9.27. Eine Permutation, die nur zwei Ziffern vertauscht und alle anderen festlässt, nennt man *Transposition*. Die entsprechende Permutationsmatrix ist genau die Elementarmatrix zur Vertauschung von Zeilen einer Matrix. Insbesondere hat jede Transposition Signum -1. Nach dem Gauß-Algorithmus ist jede Permutation ein Produkt von Transpositionen, die allerdings nicht eindeutig bestimmt sind. Nach Satz 9.26 ist das Produkt einer geraden Anzahl an Transpositionen niemals ein Produkt einer ungeraden Anzahl an Transpositionen.

Satz 9.28 (LEIBNIZ-Formel). *Für $A = (a_{ij}) \in K^{n\times n}$ gilt*

$$\boxed{\det(A) = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \dots a_{n\sigma(n)}.}$$

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

$$\begin{aligned}\det(A) &= \sum_{i_1=1}^n a_{1i_1} \det \begin{pmatrix} e_{i_1} \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \sum_{i_1=1}^n a_{1i_1} \sum_{i_2=1}^n a_{2i_2} \det \begin{pmatrix} e_{i_1} \\ e_{i_2} \\ a_3 \\ \vdots \end{pmatrix} = \dots \\ &= \sum_{1\le i_1,\dots,i_n\le n} a_{1i_1} \dots a_{ni_n} \det \begin{pmatrix} e_{i_1} \\ \vdots \\ e_{i_n} \end{pmatrix}.\end{aligned}$$

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

$$\det(A) = \sum_{\sigma\in S_n} a_{1\sigma(1)} \dots a_{n\sigma(n)} \det P_\sigma = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \dots a_{n\sigma(n)}. \qquad \square$$

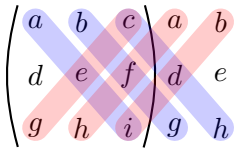
Folgerung 9.29 (SARRUS-Regel). *Für 3×3-Matrizen gilt*

$$\boxed{\det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = aei + bfg + cdh - gec - hfa - idb.}$$

Beweis. Man benutze die Leibniz-Formel mit Beispiel 9.25. $\square$

Bemerkung 9.30.

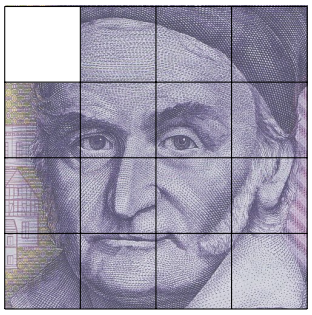
(a) Man kann sich die Sarrus-Regel mit folgendem Schema merken:

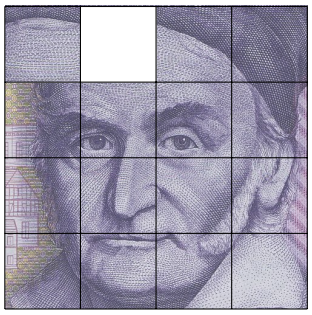


(b) Achtung: Die Sarrus-Regel gilt *nur* für 3×3-Matrizen (für 4×4-Matrizen braucht man $4! = 24$ Summanden).

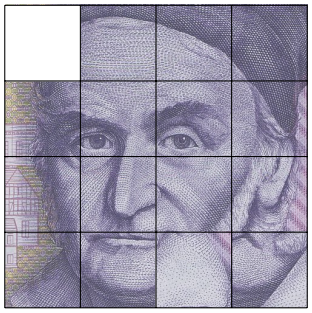
(c) Sei $A \in K^{n\times n}$ mit Eigenwert λ. Dann ist $E_\lambda(A) = \operatorname{Ker}(A - \lambda 1_n) \neq \{0\}$ und $\det(A - \lambda 1_n) = 0$. Nach der Leibniz-Formel ist $\det(A - \lambda 1_n)$ ein Polynom in λ. Auf diese Weise werden wir alle Eigenwerte von A berechnen.

Beispiel 9.31. 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 Beispiel 9.27 ist diese Konfiguration also unlösbar und Loyd musste das Preisgeld nie auszahlen.

10 Polynome

10.1 Der Vektorraum der Polynome

Definition 10.1. Ein *Polynom* über einem Körper K in der *Variablen* X ist eine Summe der Form

$$\alpha = \sum_{k=0}^{d} a_k X^k = a_0 + a_1 X + \ldots + a_d X^d$$

mit *Koeffizienten* $a_0, \ldots, a_d \in K$.[8]

- Man nennt a_0 das *Absolutglied* von α.

[7] siehe [Slocum und Sonneveld, *The 15 puzzle*, The Slocum Puzzle Foundation, Beverly Hills, 2006]
[8] Formal: Ein Polynom ist eine Abbildung $\mathbb{N}_0 \to K$, $k \mapsto a_k$ mit $|\{k \in \mathbb{N}_0 : a_k \neq 0\}| < \infty$.

- 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.[9] 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 10.2.

(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 10.3. Das Polynom $\alpha = X^2 - 3X + 1 \in \mathbb{Q}[X]$ ist normiert vom Grad 2 mit Absolutglied 1.

Satz 10.4. *Mit den Verknüpfungen*

$$\sum a_k X^k + \sum b_k X^k := \sum (a_k + b_k) X^k,$$
$$\lambda \sum a_k X^k := \sum (\lambda a_k) X^k$$

wird $K[X]$ ein unendlich-dimensionaler K-Vektorraum mit Basis $1, X, X^2, \ldots$.

Beweis. Seien $\alpha = \sum a_k X^k$ und $\beta = \sum b_k X^k$ mit $d := \deg(\alpha) \geq \deg(\beta)$. Man kann $(a_0, \ldots, a_d)$ und $(b_0, \ldots, b_d)$ 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 $1, X, X^2, \ldots$, d. h. $K[X] = \langle 1, X, X^2, \ldots \rangle$. Aus der Eindeutigkeit der Koeffizienten (Bemerkung 10.2) folgt die lineare Unabhängigkeit von $\{1, X, X^2, \ldots\}$. □

Bemerkung 10.5.

(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 $1, X, \ldots, X^{d-1}$.

(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.

[9] auch *Leitkoeffizient* genannt

Satz 10.6. *Für Polynome* $\alpha = \sum a_k X^k$ $\beta = \sum b_k X^k$ *ist*

$$\alpha \cdot \beta := \sum_{k=0}^{\infty} \Big(\sum_{l=0}^{k} a_l b_{k-l} \Big) X^k$$

ein Polynom vom Grad $\deg(\alpha) + \deg(\beta)$. *Es gelten folgende Rechenregeln:*

$$\alpha\beta = \beta\alpha, \qquad \alpha(\beta+\gamma) = \alpha\beta + \alpha\gamma, \qquad \alpha(\beta\gamma) = (\alpha\beta)\gamma.$$

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

$$\alpha\beta = \sum_{k=0}^{\infty} \Big(\sum_{l=0}^{k} a_l b_{k-l} \Big) X^k = \sum_{k=0}^{\infty} \Big(\sum_{l=0}^{k} b_l a_{k-l} \Big) X^k = \beta\alpha$$

$$\alpha(\beta+\gamma) = \sum_{k=0}^{\infty} \Big(\sum_{l=0}^{k} a_l (b_{k-l} + c_{k-l}) \Big) X^k = \sum_{k=0}^{\infty} \Big(\sum_{l=0}^{k} a_l b_{k-l} \Big) X^k + \sum_{k=0}^{\infty} \Big(\sum_{l=0}^{k} a_l c_{k-l} \Big) X^k = \alpha\beta + \alpha\gamma.$$

Der Koeffizient von X^k in $\alpha(\beta\gamma)$ ist

$$\sum_{l=0}^{k} a_l \sum_{m=0}^{k-l} b_m c_{k-l-m} = \sum_{\substack{r,s,t \in \mathbb{N}_0 \\ r+s+t=k}} a_r b_s c_t = \sum_{l=0}^{k} \Big(\sum_{m=0}^{l} a_m b_{l-m} \Big) c_{k-l}.$$

Dies ist auch der Koeffizient von X^k in $(\alpha\beta)\gamma$. Also ist $\alpha(\beta\gamma) = (\alpha\beta)\gamma$. □

Bemerkung 10.7. 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 10.8 (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

$$\delta := \alpha - \beta\gamma = a_d X^d + a_{d-1} X^{d-1} + \ldots + a_0$$

möglichst kleinen Grad $d \in \mathbb{N}_0 \cup \{-\infty\}$ hat. Sei $\beta = b_e X^e + \ldots + b_0$ und $e := \deg \beta$. Gilt $d \geq e$, so ist

$$a_d b_e^{-1} X^{d-e} \beta = a_d b_e^{-1} (b_e X^d + b_{e-1} X^{d-1} + \ldots + b_0 X^{d-e}) = a_d X^d + \ldots$$

und es folgt

$$\deg\big(\alpha - \beta(\gamma + a_d b_e^{-1} X^{d-e})\big) = \deg\big(\delta - a_d b_e^{-1} X^{d-e} \beta\big) < d.$$

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 10.6 ist

$$e + \deg(\tilde{\gamma} - \gamma) = \deg(\beta) + \deg(\tilde{\gamma} - \gamma) = \deg(\beta(\tilde{\gamma} - \gamma)) = \deg(\delta - \tilde{\delta}) \leq \max\{\deg(\delta), \deg(\tilde{\delta})\} < e$$

und es folgt $\deg(\tilde{\gamma} - \gamma) = -\infty = \deg(\delta - \tilde{\delta})$. Dies zeigt $\tilde{\gamma} = \gamma$ und $\tilde{\delta} = \delta$. □

Definition 10.9. In der Situation von Satz 10.8 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 10.10.

$$\begin{array}{rrrrl} (2X^3 & -X^2 & +5X & +1) & : (X^2+3) = 2X - 1 =: \gamma \\ -(2X^3 & & +6X) & & \\ & -X^2 & -X & +1 & \\ & -(-X^2 & & -3) & \\ & & -X & +4 =: \delta & \end{array}$$

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 10.11. Die Division durch normierte Polynome vom Grad 1 lässt sich mit dem HORNER-*Schema*[10] effizient gestalten. Sei dazu $\alpha = a_n X^n + a_{n-1}X^{n-1} + \ldots + a_0$ und $\beta = X - b$. Wir berechnen $c_n := 0$, $c_k := a_{k+1} + bc_{k+1}$ für $k = n-1, \ldots, 0$ und $d := a_0 + bc_0$:

$$\begin{array}{rccccc} & a_n & a_{n-1} & a_{n-2} & \cdots & a_0 \\ + & 0 & bc_{n-1} & bc_{n-2} & \cdots & bc_0 \\ \hline & c_{n-1} & c_{n-2} & \cdots & c_0 & d \end{array}$$

Für $\gamma := c_{n-1}X^{n-1} + \ldots + c_0$ gilt nun

$$\begin{aligned} \beta\gamma + d &= c_{n-1}X^n + (c_{n-2} - bc_{n-1})X^{n-1} + \ldots + (c_0 - bc_1)X - bc_0 + d \\ &= a_n X^n + a_{n-1}X^{n-1} + \ldots + a_0 = \alpha. \end{aligned}$$

Beispiel 10.12. 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 10.13. 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 $\alpha(x) = 0$.

[10] Auch *Ruffinis Regel* genannt

Lemma 10.14. *Für $\alpha, \beta \in K[X]$ und $x \in K$ gilt*

$$(\alpha + \beta)(x) = \alpha(x) + \beta(x),$$
$$(\alpha\beta)(x) = \alpha(x)\beta(x).$$

Beweis. Seien $\alpha = \sum a_k X^k$ und $\beta = \sum b_k X^k$. Dann gilt

$$(\alpha + \beta)(x) = \sum (a_k + b_k)x^k = \sum a_k x^k + \sum b_k x^k = \alpha(x) + \beta(x),$$
$$(\alpha\beta)(x) = \sum_{n=0}^{\infty} \sum_{k=0}^{n} a_k b_{n-k} x^n = \sum_{n=0}^{\infty} \sum_{k=0}^{n} (a_k x^k)(b_{n-k} x^{n-k}) = \sum a_k x^k \sum b_k x^k = \alpha(x)\beta(x). \qquad \square$$

Bemerkung 10.15. Merkregel: 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 $x, y \in K$.

Lemma 10.16. *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

$$\delta = \delta(x) = (\alpha - (X - x)\gamma)(x) \overset{10.14}{=} \alpha(x) - (x - x)\gamma(x) = \alpha(x). \qquad \square$$

Definition 10.17. 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.

Lemma 10.18. *Seien $x_1, \ldots, x_n \in K$. Dann hat jeder normierte Teiler von $(X - x_1) \ldots (X - x_n) \in K[X]$ die Form $(X - x_{i_1}) \ldots (X - x_{i_k})$ mit $1 \leq i_1 < \ldots < i_k \leq n$.*

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 $(X - x_1) \ldots (X - x_n) = \alpha\beta$. Dann ist $\alpha(x_n)\beta(x_n) = (\alpha\beta)(x_n) = 0$, o. B. d. A. sei $\alpha(x_n) = 0$. Nach Lemma 10.16 gilt $\alpha = (X - x_n)\gamma$ für ein $\gamma \in K[X]$. Nach Bemerkung 10.7 darf man $X - x_n$ kürzen und erhält $(X - x_1) \ldots (X - x_{n-1}) = \gamma\beta$. Die Behauptung folgt nun durch Induktion. $\square$

Beispiel 10.19.

(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 Grades) 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:

$$x_2 = \frac{1}{2}\big(-2 + \sqrt{4+12}\big) = 1, \qquad x_3 = \frac{1}{2}\big(-2 - \sqrt{4+12}\big) = -3.$$

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 11.26).

(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.

Bemerkung 10.20. In der Analysis identifiziert man Polynome $\alpha \in \mathbb{R}[X]$ mit ihren Funktionen $\mathbb{R} \to \mathbb{R}$, $x \mapsto \alpha(x)$ (Aufgabe 33). Über endlichen Körpern K (anstelle von $\mathbb{R}$) würde man dabei Information verlieren, denn es gibt nur endlich viele Abbildungen $K \to K$, aber unendlich viele Polynome. Zum Beispiel entsprechen die Polynome $X, X^2, \ldots \in \mathbb{F}_2[X]$ alle der Identität $\mathrm{id}_{\mathbb{F}_2}$.

10.3 Charakteristische Polynome

Bemerkung 10.21. Im Folgenden betrachten wir Matrizen mit Einträgen in $K[X]$. Aufgrund der Rechenregeln für Polynome (Satz 10.6) 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 9.18 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.

Definition 10.22. 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.[11]

Bemerkung 10.23. Sei V ein n-dimensionaler K-Vektorraum mit Basis B. Für $f \in \mathrm{End}(V)$ definiert man $\chi_f := \det(X1_n - {}_B[f]_B)$. Dies hängt wegen

$$\begin{aligned} \det(X1_n - {}_{B'}[f]_{B'}) &= \det(X1_n - {}_{B'}\Delta_B {}_B[f]_B {}_{B'}\Delta_B{}^{-1}) = \det\big({}_{B'}\Delta_B (X1_n - {}_B[f]_B) {}_{B'}\Delta_B{}^{-1}\big) \\ &= \det({}_{B'}\Delta_B)\det(X1_n - {}_B[f]_B)\det({}_{B'}\Delta_B)^{-1} = \det(X1_n - {}_B[f]_B) \end{aligned}$$

[11] In 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.

nicht von der Wahl von B ab. In den nachfolgenden Sätzen kann man Matrizen also durch Endomorphismen ersetzen (und umgekehrt).

Beispiel 10.24. Das charakteristische Polynom von $A := \left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right) \in \mathbb{Q}[X]$ ist

$$\chi_A = \det \begin{pmatrix} X-1 & -2 \\ -3 & X-4 \end{pmatrix} = (X-1)(X-4) - (-2)(-3) = X^2 - 5X - 2 = X^2 - \operatorname{tr}(A)X + \det(A).$$

Lemma 10.25. *Für $A \in K^{n\times n}$ gilt $\chi_A = X^n - \operatorname{tr}(A)X^{n-1} + \ldots + (-1)^n \det A$. Insbesondere ist χ_A normiert vom Grad n.*

Beweis. Sei $A = (a_{ij})$. Nach der Leibniz-Formel gilt

$$\begin{aligned}\chi_A = \det(X1_n - A) &= (X - a_{11})(X - a_{22}) \ldots (X - a_{nn}) \\ &+ \sum_{\sigma \in S_n \setminus \{\mathrm{id}\}} \operatorname{sgn}(\sigma)(\delta_{1\sigma(1)}X - a_{1\sigma(1)}) \ldots (\delta_{n\sigma(n)}X - a_{n\sigma(n)}).\end{aligned}$$

Für $\sigma \in S_n \setminus \{\mathrm{id}\}$ existiert ein $k \in \{1, \ldots, n\}$ 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

$$(\delta_{1\sigma(1)}X - a_{1\sigma(1)}) \ldots (\delta_{n\sigma(n)}X - a_{n\sigma(n)})$$

ist ein Polynom vom Grad $\leq n-2$. Insgesamt ist

$$\chi_A = (X - a_{11})(X - a_{22}) \ldots (X - a_{nn}) + \alpha$$

mit $\deg(\alpha) \leq n-2$. Ausmultiplizieren zeigt $\chi_A = X^n - (a_{11} + \ldots + a_{nn})X^{n-1} + \ldots = X^n - \operatorname{tr}(A)X^{n-1} + \ldots$.

Zur Berechnung des Absolutglieds setzt man $X = 0$ und erhält $\chi_A(0) \overset{10.14}{=} \det(-A) = (-1)^n \det A$ aus Bemerkung 9.8. □

Satz 10.26. *Die Eigenwerte von $A \in K^{n\times n}$ sind die Nullstellen von χ_A.*

Beweis. Es gilt

$$\operatorname{Ker}(A - \lambda 1_n) \neq \{0\} \iff \det(A - \lambda 1_n) = 0 \overset{9.8}{\iff} \det(\lambda 1_n - A) = 0 \overset{10.14}{\iff} \chi_A(\lambda) = 0. \qquad \square$$

Lemma 10.27. *Sei $\lambda \in K$ ein Eigenwert von $f \in \operatorname{End}(V)$. 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 $b_1, \ldots, b_e$ von $E_\lambda(f)$ zu einer Basis $B := \{b_1, \ldots, b_n\}$ von V. Dann gilt

$$\chi_f = \det(X1_n - {}_B[f]_B) = \det \begin{pmatrix} (X-\lambda)1_e & * \\ 0 & * \end{pmatrix}.$$

Wiederholte Laplace-Entwicklung nach der ersten Spalte zeigt

$$\chi_f = (X - \lambda) \det \begin{pmatrix} (X-\lambda)1_{e-1} & * \\ 0 & * \end{pmatrix} = \ldots = (X - \lambda)^e \beta$$

für ein $\beta \in K[X]$. Also ist die algebraische Vielfachheit von λ als Nullstelle von χ_f mindestens e. □

Satz 10.28. *Genau dann ist $f \in \mathrm{End}(V)$ diagonalisierbar, wenn χ_f in Linearfaktoren zerfällt und für jede Nullstelle von χ_f die algebraische Vielfachheit mit der geometrischen Vielfachheit übereinstimmt.*

Beweis. Seien $\lambda_1, \ldots, \lambda_k \in K$ die verschiedenen Nullstellen von χ_f mit algebraischen Vielfachheiten $m_1, \ldots, m_k$. Dann existiert ein $\alpha \in K[X]$ mit $\chi_f = (X-\lambda_1)^{m_1} \ldots (X-\lambda_k)^{m_k}\alpha$. Sei m_i' die geometrische Vielfachheit von λ_i als Eigenwert von f. Nach Lemma 10.27 gilt

$$m_1' + \ldots + m_k' \leq m_1 + \ldots + m_k + \deg(\alpha) \overset{10.6}{=} \deg\big((X-\lambda_1)^{m_1} \ldots (X-\lambda_k)^{m_k}\alpha\big) = \deg(\chi_f) = \dim V.$$

Nach Satz 8.10 besitzt V genau dann eine Basis aus Eigenvektoren, wenn $m_1' + \ldots + m_k' = \dim V$. 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 $i = 1, \ldots, k$). □

Bemerkung 10.29. Zerfällt χ_A in Linearfaktoren, so gilt

$$\chi_A = (X-\lambda_1)\ldots(X-\lambda_n) = X^n - (\lambda_1 + \ldots + \lambda_n)X^{n-1} + \ldots + (-1)^n\lambda_1 \ldots \lambda_n.$$

Ein Vergleich mit Lemma 10.25 zeigt:

$$\boxed{\begin{aligned} \mathrm{tr}(A) &= \lambda_1 + \ldots + \lambda_n, \\ \det(A) &= \lambda_1 \ldots \lambda_n, \end{aligned}}$$

d. h. die Spur ist die Summe der Eigenwerte und die Determinante ist das Produkt der Eigenwerte (sofern diese existieren). Hat man bereits $\lambda_1, \ldots, \lambda_{n-1}$ bestimmt, so erhält man $\lambda_n = \mathrm{tr}(A) - \lambda_1 - \ldots - \lambda_{n-1}$.

Beispiel 10.30. Die FIBONACCI-*Zahlen* F_k sind rekursiv definiert:

$$F_k := k \quad (k = 0, 1) \qquad F_{k+1} := F_k + F_{k-1} \quad (k \geq 1).$$

n	0	1	2	3	4	5	6	7	8	9	10
F_n	0	1	1	2	3	5	8	13	21	34	55

Wir suchen eine explizite Formel für F_k. Für $A := \left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right) \in \mathbb{R}^{2\times 2}$ gilt

$$\begin{pmatrix} F_{k+1} \\ F_k \end{pmatrix} = A \begin{pmatrix} F_k \\ F_{k-1} \end{pmatrix} = A^2 \begin{pmatrix} F_{k-1} \\ F_{k-2} \end{pmatrix} = \ldots = A^k \begin{pmatrix} F_1 \\ F_0 \end{pmatrix} = A^k \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

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

$$E_\varphi(A) = \mathrm{Ker}(A - \varphi 1_2) = \langle \begin{pmatrix} \varphi \\ 1 \end{pmatrix} \rangle,$$
$$E_\psi(A) = \langle \begin{pmatrix} \psi \\ 1 \end{pmatrix} \rangle.$$

Für $S := \begin{pmatrix} \varphi & \psi \\ 1 & 1 \end{pmatrix}$ gilt also $S^{-1}AS = \mathrm{diag}(\varphi, \psi)$ und

$$A^k = (S\,\mathrm{diag}(\varphi,\psi)S^{-1})^k = S\,\mathrm{diag}(\varphi,\psi)^k S^{-1} = S\,\mathrm{diag}(\varphi^k, \psi^k)S^{-1}.$$

Nach Beispiel 9.19 ist

$$S^{-1} = \frac{1}{\det(S)}\widetilde{S} = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 & -\psi \\ -1 & \varphi \end{pmatrix}.$$

Insgesamt erhält man

$$A^k = S\,\mathrm{diag}(\varphi^k, \psi^k)S^{-1} = \frac{1}{\sqrt{5}}\begin{pmatrix} \varphi^{k+1} & \psi^{k+1} \\ \varphi^k & \psi^k \end{pmatrix}\begin{pmatrix} 1 & -\psi \\ -1 & \varphi \end{pmatrix} = \frac{1}{\sqrt{5}}\begin{pmatrix} * & * \\ \varphi^k - \psi^k & * \end{pmatrix}$$

und

$$\boxed{F_k = \frac{1}{\sqrt{5}}(\varphi^k - \psi^k)} \qquad \text{(Binet-}\textit{Formel}^{12}\text{)}$$

Wegen $|\psi^k| \approx 0.618^k \to 0$ gilt $F_k \approx \frac{1}{\sqrt{5}}\varphi^k$, d. h. F_k wächst exponentiell.

10.4 Minimalpolynome

Bemerkung 10.31. 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 10.14 gelten auch in dieser Allgemeinheit.

Satz 10.32. *Für $A \in K^{n\times n}$ existiert genau ein normiertes Polynom $\mu_A \in K[X]\setminus\{0\}$ mit $\mu_A(A) = 0_{n\times n}$ und $\deg(\mu_A)$ minimal.*

Beweis. Wegen $\dim K^{n\times n} = n^2$ (Lemma 5.4) sind die Potenzen $1_n = A^0, A, A^2, \ldots, A^{n^2}$ linear abhängig in $K^{n\times n}$. Also existieren $a_0, \ldots, a_{n^2} \in K$ (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. □

Definition 10.33. Man nennt μ_A das *Minimalpolynom* von A.

Beispiel 10.34. Sei $A := \left(\begin{smallmatrix} 1 & -1 \\ 2 & 0 \end{smallmatrix}\right) \in \mathbb{Q}^{2\times 2}$. 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 10.35. *Sei $A \in K^{n\times n}$ und $\alpha \in K[X]$ mit $\alpha(A) = 0$. Dann gilt $\mu_A \mid \alpha$.*

[12] Man kann die Formel auch durch Induktion beweisen, sofern man sie zuvor erraten hat.

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 $\mu_A \mid \alpha$. □

Bemerkung 10.36.

(a) Sei V ein K-Vektorraum mit Basis B. Für $f \in \mathrm{End}(V)$ sei wie üblich $\mu_f := \mu_{{}_B[f]_B}$. Wegen

$$\sum a_k {}_B[f]_B^k = 0 \iff {}_{B'}\Delta_B \Big(\sum a_k {}_B[f]_B^k\Big) {}_{B'}\Delta_B{}^{-1} = 0 \iff \sum a_k {}_{B'}\Delta_{BB}[f]_{BB'}^k \Delta_B{}^{-1} = 0$$
$$\iff \sum a_k \big({}_{B'}\Delta_{BB}[f]_{BB'}\Delta_B{}^{-1}\big)^k = 0 \iff \sum a_k {}_{B'}[f]_{B'}^k = 0$$

hängt μ_f nicht von der Wahl von B ab. Die folgenden Sätze über Matrizen gelten sinngemäß auch für Endomorphismen.

(b) Aus dem Beweis von Satz 10.32 erhält man $\deg(\mu_A) \le n^2$. Der nächste Satz impliziert $\deg(\mu_A) \le n$.

Satz 10.37 (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 $\widetilde{B} \in K[X]^{n\times n}$ die zu B komplementäre Matrix. Aus jedem Eintrag von $\widetilde{B}$ extrahieren wir den Koeffizienten von X^k und bilden daraus die Matrix $B_k \in K^{n\times n}$. Es gilt nun

$$\widetilde{B} = \sum_{k=0}^{\infty} B_k X^k,$$

wobei nur endlich viele der B_k ungleich 0 sind. Sei $\chi_A = \sum a_k X^k$. Nach Satz 9.18 gilt

$$\sum_{k=0}^{\infty} a_k 1_n X^k = \chi_A 1_n = \det(B)1_n = \widetilde{B}B = \sum_{k=0}^{\infty} B_k X^k (X1_n - A) = \sum_{k=0}^{\infty} (B_{k-1} - B_k A)X^k,$$

wobei $B_{-1} := 0_{n\times n}$. Ein Koeffizientenvergleich ergibt $a_k 1_n = B_{k-1} - B_k A$ für $k = 0, 1, \ldots$. 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 10.35. □

Beispiel 10.38.

(a) Für $A \in K^{2\times 2}$ gilt $A^2 - \mathrm{tr}(A)A + \det(A)1_2 = 0$ nach Lemma 10.25.

(b) Sei $A \in \mathrm{GL}(n, K)$ mit $\chi_A = \mu_A\gamma$ für ein $\gamma \in K[X]$. Nach Lemma 10.25 gilt

$$\mu_A(0)\gamma(0) = \chi_A(0) = \det(A) \neq 0.$$

Also hat μ_A die Form $\mu_A = X^d + a_{d-1}X^{d-1} + \ldots + a_0$ mit $a_0 \neq 0$. Man kann nun die Gleichung $A^d + a_{d-1}A^{d-1} + \ldots + a_0 1_n = 0$ auf beiden Seiten mit A^{-1} multiplizieren und erhält $A^{d-1} + a_{d-1}A^{d-2} + \ldots + a_1 1_n + a_0 A^{-1} = 0$. Dies liefert eine Formel für die Inverse

$$A^{-1} = -\frac{1}{a_0}\big(A^{d-1} + a_{d-1}A^{d-2} + \ldots + a_1 1_n\big).$$

Speziell für $n = 2$:

$$A^{-1} \overset{\text{(a)}}{=} \frac{1}{\det A}\big(\mathrm{tr}(A)1_2 - A\big) = \frac{1}{\det A}\widetilde{A}$$

(vgl. Beispiel 9.19).

Satz 10.39. *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 10.26). Sei umgekehrt $\lambda \in K$ ein Eigenwert von A mit Eigenvektor $v \in K^{n\times 1}$. Für $k \in \mathbb{N}_0$ ist $A^k v = A^{k-1}\lambda v = \ldots = \lambda^k v$. 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 10.40. Wegen $\deg \mu_A \leq \deg \chi_A$ vereinfacht Satz 10.39 die Berechnung der Eigenwerte. Andererseits ist nicht klar, wie man μ_A effizient berechnet. Der nächste Satz verbessert Folgerung 8.12.

Satz 10.41. *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 $S \in \mathrm{GL}(n, K)$ mit $S^{-1}AS = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Die λ_i lassen sich sortieren, indem man die Spalten von S (d. h. die Eigenvektoren von A) entsprechend anordnet. Wie in Bemerkung 10.36 zeigt man außerdem $\mu_{S^{-1}AS} = \mu_A$. Wir können also

$$A = \begin{pmatrix} \lambda_1 1_{n_1} & & 0 \\ & \ddots & \\ 0 & & \lambda_k 1_{n_k} \end{pmatrix}$$

annehmen, wobei $n = n_1 + \ldots + n_k$ und $\lambda_i \neq \lambda_j$ für $i \neq j$. Dann gilt $(A - \lambda_1 1_n) \ldots (A - \lambda_k 1_n) = 0$, denn

$$\mathrm{diag}(\underbrace{0, \ldots, 0}_{n_1}, *, \ldots, *)\, \mathrm{diag}(*, \ldots, *, \underbrace{0, \ldots, 0}_{n_2}, *, \ldots, *) \ldots \mathrm{diag}(*, \ldots, *, \underbrace{0, \ldots, 0}_{n_k}) = 0$$

Nach Lemma 10.35 ist μ_A ein Teiler von $(X - \lambda_1) \ldots (X - \lambda_k)$. Andererseits sind $\lambda_1, \ldots, \lambda_k$ Eigenwerte und damit Nullstellen von μ_A. Dies zeigt $\mu_A = (X - \lambda_1) \ldots (X - \lambda_k)$.

Nehmen wir nun umgekehrt $\mu_A = (X - \lambda_1) \ldots (X - \lambda_k)$ mit paarweise verschiedenen $\lambda_1, \ldots, \lambda_k$ an. Sei P_k der k-dimensionale Vektorraum aller Polynome vom Grad kleiner k (Bemerkung 10.5). Für $i = 1, \ldots, k$ sei

$$\gamma_i := (X - \lambda_1) \ldots (X - \lambda_{i-1})(X - \lambda_{i+1}) \ldots (X - \lambda_k) \in P_k.$$

Seien $a_1, \ldots, a_k \in K$ mit $a_1\gamma_1 + \ldots + a_k\gamma_k = 0$. Dann gilt

$$a_i(\lambda_i - \lambda_1) \ldots (\lambda_i - \lambda_{i-1})(\lambda_i - \lambda_{i+1}) \ldots (\lambda_i - \lambda_n) = a_i\gamma_i(\lambda_i) = (a_1\gamma_1 + \ldots + a_k\gamma_k)(\lambda_i) = 0$$

und es folgt $a_i = 0$ für $i = 1, \ldots, k$. Also sind $\gamma_1, \ldots, \gamma_k$ linear unabhängig in P_k. Wegen $\dim P_k = k$ bilden sie sogar eine Basis. Insbesondere existieren $b_1, \ldots, b_k \in K$ mit $\gamma := b_1\gamma_1 + \ldots + b_k\gamma_k = X^0 = 1$. 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

$$v = 1_n v = A^0 v = \gamma(A)(v) = b_1\gamma_1(A)(v) + \ldots + b_k\gamma_k(A)(v).$$

Dies zeigt $K^{n\times 1} = E_{\lambda_1}(A) + \ldots + E_{\lambda_k}(A)$. Also ist A diagonalisierbar nach Bemerkung 8.6. □

Beispiel 10.42. 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 10.41 ist A nicht diagonalisierbar.

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 $V \times V \to \mathbb{R}$, $(v, w) \mapsto [v, w]$ heißt *Skalarprodukt*[13], 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 $V \to \mathbb{R}$, $v \mapsto [v, x]$ und $V \to \mathbb{R}$, $v \mapsto [x, v]$ linear (dies erklärt den Begriff *bilinear*). Insbesondere ist $[v, 0] = 0 = [0, v]$ für alle $v \in V$. Dennoch ist die Abbildung $V \times V \to \mathbb{R}$, $(v, w) \mapsto [v, w]$ *nicht* linear (z. B. $[v, v] > 0 = [0, v] + [v, 0]$ für $v \neq 0$).

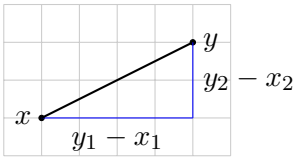
(b) Jeder Unterraum eines euklidischen Raums ist selbst ein euklidischer Raum mit dem eingeschränkten Skalarprodukt.

Beispiel 11.4. Das wichtigste Beispiel eines euklidischen Raums ist $V = \mathbb{R}^n$ mit dem *Standardskalarprodukt*

$$\boxed{[x, y] := xy^{\mathrm{t}} = \sum_{i=1}^{n} x_i y_i} \qquad (x, y \in \mathbb{R}^n).$$

[13] Die Schreibweise $[v, w]$ ist in der Literatur nicht einheitlich. Man findet auch $\langle v, w \rangle$ (Verwechselung mit Spann), $(v \mid w)$ u. ä. Ebenso findet man $\|v\|$ anstatt $|v|$.

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 Pythagoras[14] entspricht die Norm $|y - x|$ im $\mathbb{R}^2$ dem geometrischen Abstand zwischen x und y:

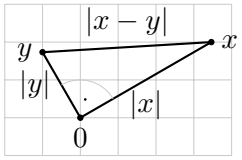


$$|y - x| = \sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2}$$

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*:

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

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) $|[v, w]| \leq |v||w|$ (Cauchy-Schwarz-Ungleichung) *mit Gleichheit genau dann, wenn v und w linear abhängig sind.*

(c) $||v| - |w|| \leq |v + w| \leq |v| + |w|$ (Dreiecksungleichung).

Beweis.

(a) $|\lambda v| = \sqrt{[\lambda v, \lambda v]} = \sqrt{\lambda^2 [v, v]} = \sqrt{\lambda^2}\sqrt{[v, v]} = |\lambda||v|$.

(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 $|[v, w]| = |\mu||w|^2 \overset{(a)}{=} |\mu w||w| = |v||w|$.

[14] Man könnte auch den Abstand zwischen x und y durch $|y - x|$ *definieren* und damit den Satz des Pythagoras *beweisen*.

(c) Zunächst ist

$$|v+w|^2 = [v+w, v+w] = [v,v] + 2[v,w] + [w,w] \overset{\text{(b)}}{\leq} |v|^2 + 2|v||w| + |w|^2 = (|v| + |w|)^2$$

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 $||v| - |w|| \leq |v+w|$. □

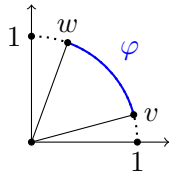
Bemerkung 11.6.

(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:

$$\frac{[\lambda v, \mu w]}{|\lambda v||\mu w|} = \frac{\lambda\mu[v,w]}{|\lambda||\mu||v||w|} = \frac{[v,w]}{|v||w|} \qquad (\lambda, \mu > 0)$$

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 Einheitskreis[15] zwischen v und w:



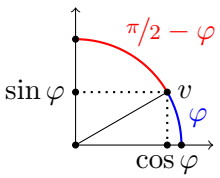
Die Länge des Halbkreisbogens nennt man π und berechnet $\pi \approx 3.14$ (Aufgabe 46). Der *Kosinus* von φ wird durch $\cos\varphi := [v, w]$ definiert.[16] Es gilt

$$\cos 0 = [e_1, e_1] = 1, \qquad \cos(\pi/2) = [e_1, e_2] = 0, \qquad \cos\pi = [e_1, -e_1] = -1.$$

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

$$\sin\varphi := \cos(\varphi - \pi/2) = \cos(\pi/2 - \varphi)$$

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 $y = [v, e_2] = \cos(\pi/2 - \varphi) = \sin\varphi$:



[15] in der Ebene $\langle v, w\rangle$

[16] Man kann zeigen, dass diese Definition mit der analytischen Definition als Potenzreihe übereinstimmt.

11.2 Orthonormalbasen

Definition 11.7. Sei V ein n-dimensionaler euklidischer Raum. Vektoren $b_1, \ldots, b_n$ 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 $b_1, \ldots, b_n$ von V ist tatsächlich eine Basis. Dafür genügt es die lineare Unabhängigkeit zu prüfen. Seien $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ mit $\lambda_1 b_1 + \ldots + \lambda_n b_n = 0$. Dann gilt

$$\lambda_i = \sum_{j=1}^{n} \lambda_j [b_j, b_i] = \Big[\sum_{j=1}^{n} \lambda_j b_j, b_i\Big] = [0, b_i] = 0$$

für $i = 1, \ldots, n$.

Beispiel 11.9. Die Standardbasis $e_1, \ldots, e_n \in \mathbb{R}^n$ ist eine Orthonormalbasis bzgl. des Standardskalarprodukts. Jede Permutation einer Orthonormalbasis ist wieder eine Orthonormalbasis.

Satz 11.10 (Gram-Schmidt-Verfahren)**.** *Seien $v_1, \ldots, v_k \in V$ linear unabhängig. Wir definieren rekursiv:*

$$\boxed{b_s := v_s - \sum_{i=1}^{s-1} \frac{[v_s, b_i]}{[b_i, b_i]} b_i} \qquad (s = 1, \ldots, k).$$

Dann sind $b_1, \ldots, b_k$ paarweise orthogonal mit $\langle v_1, \ldots, v_k \rangle = \langle b_1, \ldots, b_k \rangle$. Folglich ist $\frac{1}{|b_1|} b_1, \ldots, \frac{1}{|b_k|} b_k$ eine Orthonormalbasis von $\langle v_1, \ldots, v_k \rangle$.

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. $\langle v_1, \ldots, v_{k-1} \rangle = \langle b_1, \ldots, b_{k-1} \rangle$ und $[b_i, b_j] = 0$ für $1 \leq i < j \leq k - 1$. Wegen $\sum_{i=1}^{k-1} \frac{[v_k, b_i]}{[b_i, b_i]} b_i \in \langle b_1, \ldots, b_{k-1} \rangle$ gilt

$$\langle v_1, \ldots, v_k \rangle = \langle b_1, \ldots, b_{k-1}, v_k \rangle = \langle b_1, \ldots, b_k \rangle.$$

Für $i = 1, \ldots, k - 1$ 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. □

Folgerung 11.11. *Jeder euklidische Raum besitzt (mindestens) eine Orthonormalbasis.*

Beweis. Man wende das Gram-Schmidt-Verfahren auf eine beliebige Basis an. □

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

$$\begin{aligned} b_1 &:= v_1 = (1, 0, 1), \\ b_2 &:= v_2 - \frac{[v_2, b_1]}{[b_1, b_1]} b_1 = (0, 1, 1) - \frac{1}{2}(1, 0, 1) = \frac{1}{2}(-1, 2, 1), \end{aligned}$$

$$b_3 := v_3 - \frac{[v_3,b_1]}{[b_1,b_1]}b_1 - \frac{[v_3,b_2]}{[b_2,b_2]}b_2 = (-1,2,0) + \frac{1}{2}(1,0,1) - \frac{5}{6}(-1,2,1) = \frac{1}{3}(1,1,-1).$$

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

$$S^\perp := \big\{v \in V : \forall s \in S : [v,s] = 0\big\}$$

das *orthogonale Komplement* von S in V.

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

$$n = k + (n-k) \le \dim U + \dim U^\perp \overset{4.21}{=} \dim(U \oplus U^\perp) \le \dim V = n.$$

Dies zeigt $V = U \oplus U^\perp$.

(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

$$U \subseteq W \implies W^\perp \subseteq U^\perp \overset{\text{(b)}}{\implies} U = (U^\perp)^\perp \subseteq (W^\perp)^\perp = W. \qquad \square$$

Beispiel 11.16.

(a) In $V = \mathbb{R}^n$ lässt sich ein orthogonales Komplement von $U \le V$ 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 $\dim L_0 = n - \operatorname{rk}(A) = n - \dim U = \dim U^\perp$.

(b) Für $v = (x,y) \in \mathbb{R}^2 \setminus \{0\}$ gilt $\langle v\rangle^\perp = \langle (y,-x)\rangle$.

(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 $Ax^{\mathrm{t}} = e_1^{\mathrm{t}}$ und zwar

$$x^{\mathrm{t}} = A^{-1} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \overset{9.18}{=} \widetilde{A} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \det(A_{11}) \\ -\det(A_{12}) \\ \det(A_{13}) \end{pmatrix} = \begin{pmatrix} v_2w_3 - v_3w_2 \\ v_3w_1 - v_1w_3 \\ v_1w_2 - v_2w_1 \end{pmatrix} =: v \times w.$$

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 $f \in \mathrm{End}(V)$. Man nennt f

- *symmetrisch*,[17] falls $[f(v), w] = [v, f(w)]$ für alle $v, w \in V$.
- *orthogonal*,[18] falls $[f(v), f(w)] = [v, w]$ für alle $v, w \in V$.

Bemerkung 11.18.

(a) Die Nullabbildung ist symmetrisch. Sind $f, g \in \mathrm{End}(V)$ symmetrisch und $\lambda \in \mathbb{R}$, dann ist offenbar auch $\lambda f + g$ symmetrisch. Die symmetrischen Abbildungen bilden also einen Unterraum von $\mathrm{End}(V)$.

(b) Orthogonale Abbildungen $f \in \mathrm{End}(V)$ sind Isomorphismen, denn aus $v \in \mathrm{Ker}(f)$ 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 Gruppe $\mathrm{O}(V)$ bzgl. Komposition (aber keinen Unterraum, da die Nullabbildung nicht orthogonal ist). Man nennt $\mathrm{O}(V)$ die *orthogonale Gruppe* von V.

Lemma 11.19. *Sei V euklidisch mit Orthonormalbasis B, $f \in \mathrm{End}(V)$ und $A := {}_B[f]_B$. Dann gilt:*

(a) f symmetrisch $\iff A^{\mathrm{t}} = A$.

(b) f orthogonal $\iff A^{\mathrm{t}} = A^{-1}$.

[17] oder *selbstadjungiert*
[18] oder *isometrisch*

Beweis. Sei $B = \{b_1, \dots, b_n\}$ und $A = (a_{ij})$. Dann ist $f(b_i) = \sum_{j=1}^n a_{ji} b_j$ für $i = 1, \dots, n$.

(a) Ist f symmetrisch, so gilt $a_{ji} = [f(b_i), b_j] = [b_i, f(b_j)] = a_{ij}$ für $1 \le i, j \le n$. Also ist $A = A^{\mathrm{t}}$. Sei umgekehrt $A = A^{\mathrm{t}}$. Dann folgt $[f(b_i), b_j] = [b_i, f(b_j)]$ für $1 \le i, j \le n$. Für beliebige $v = \sum \lambda_i b_i$ und $w = \sum \mu_i b_i$ 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 $A^{\mathrm{t}} A = 1_n$ und $A^{\mathrm{t}} = A^{-1}$. Ist umgekehrt $A^{\mathrm{t}} A = 1_n$, so gilt $[f(b_i), f(b_j)] = \delta_{ij} = [b_i, b_j]$. Wie in (a) folgt $[f(v), f(w)] = [v, w]$ für alle $v, w \in V$, d. h. f ist orthogonal. □

Bemerkung 11.20.

(a) Quadratische Matrizen A mit $A^{\mathrm{t}} = A$ (bzw. $A^{\mathrm{t}} = A^{-1}$) heißen *symmetrisch* (bzw. *orthogonal*).

(b) Für orthogonale Matrizen A gilt $A^{\mathrm{t}} A = 1_n = A A^{\mathrm{t}}$. Diese bedeutet, dass die Spalten (bzw. Zeilen) von A eine Orthonormalbasis von $\mathbb{R}^n$ bzgl. des Standardskalarprodukts bilden. Wegen

$$\det(A)^2 \overset{9.11}{=} \det(A)\det(A^{\mathrm{t}}) = \det(A A^{\mathrm{t}}) = \det(1_n) = 1$$

gilt $\det(A) = \pm 1$. Sei v ein Eigenvektor von A zum Eigenwert $\lambda \in \mathbb{R}$. Dann gilt

$$|v|^2 = v^{\mathrm{t}} v = v^{\mathrm{t}} A^{\mathrm{t}} A v = (Av)^{\mathrm{t}} (Av) = \lambda^2 v^{\mathrm{t}} v = \lambda^2 |v|^2$$

und $\lambda = \pm 1$.

(c) Die orthogonalen Matrizen bilden die Gruppe $\mathrm{O}(n, \mathbb{R})$ bzgl. Matrizenmultiplikation. Sie entspricht $\mathrm{O}(V)$ durch Basiswahl (so wie sich $\mathrm{GL}(V)$ und $\mathrm{GL}(n, K)$ entsprechen).

Beispiel 11.21.

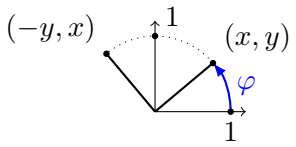
(a) Jede Permutationsmatrix ist orthogonal, denn die Zeilen bilden eine Orthonormalbasis (nämlich eine Permutation der Standardbasis).

(b) Für $A \in \mathrm{O}(2, \mathbb{R})$ 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

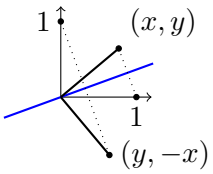
$$A = D(\varphi) := \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}.$$

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 45):

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

$$D(\pi/12) = D(\pi/3 - \pi/4) = D(\pi/3)D(\pi/4)^{-1} = D(\pi/3)D(\pi/4)^{\mathrm{t}}$$

berechnen.

Bemerkung 11.22. Um zu zeigen, dass symmetrische Endomorphismen diagonalisierbar sind, müssen wir die reellen Zahlen vorübergehend verlassen.

11.4 Komplexe Zahlen

Lemma 11.23. *Der $\mathbb{R}$-Vektorraum $\mathbb{C} := \mathbb{R}^2$ wird durch die Multiplikation*

$$(a,b)\cdot(c,d) := (ac-bd, ad+bc) \qquad (a,b,c,d\in\mathbb{R})$$

zu einem Körper.

Beweis. Als Vektorraum ist $(\mathbb{C},+)$ bereits eine abelsche Gruppe. Die Multiplikation in $\mathbb{C}$ führen wir auf die Matrizenmultiplikation zurück.[19] Dafür betrachten wir die injektive Abbildung

$$P\colon \mathbb{C}\to\mathbb{R}^{2\times 2},\quad z=(a,b)\mapsto P(z)=\begin{pmatrix} a & -b\\ b & a\end{pmatrix}.$$

Für $x,y,z\in\mathbb{C}$ gilt $P(x)+P(y)=P(x+y)$ und $P(x)P(y)=P(x\cdot y)$ (nachrechnen). Aus Lemma 5.8 folgt

$$P(x(yz))=P(x)P(yz)=P(x)(P(y)P(z))=(P(x)P(y))P(z)=P(xy)P(z)=P((xy)z)$$

und $x(yz)=(xy)z$, da P injektiv ist. Analog beweist man das Kommutativ- und Distributivgesetz. Wegen $P(1,0)=1_2$ ist $(1,0)$ das Einselement in $\mathbb{C}$. Es bleibt zu zeigen, dass $z\neq 0$ invertierbar ist. Es gilt $\det(P(z))=a^2+b^2>0$ und

$$P(z)^{-1}\overset{9.19}{=}\frac{1}{\det(P(z))}\widetilde{P(z)}=\frac{1}{a^2+b^2}\begin{pmatrix} a & b\\ -b & a\end{pmatrix}=P\Big(\frac{a}{a^2+b^2},\frac{-b}{a^2+b^2}\Big).$$ □

Definition 11.24. Man nennt $\mathbb{C}$ den Körper der *komplexen Zahlen*.

- Mittels der Abbildung $\mathbb{R}\to\mathbb{C}$, $a\mapsto(a,0)$ 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 $\mathrm{i}:=(0,1)\in\mathbb{C}\setminus\mathbb{R}$ die *imaginäre Einheit*. Es gilt $\mathrm{i}^2=(-1,0)=-1$. Da $1,\mathrm{i}$ eine Basis von $\mathbb{C}$ bilden, lässt sich jede komplexe Zahl eindeutig in der Form $z=a+b\mathrm{i}$ schreiben, wobei $\mathrm{Re}(z):=a\in\mathbb{R}$ der *Realteil* und $\mathrm{Im}(z):=b\in\mathbb{R}$ der *Imaginärteil* von z ist.
- Man nennt $|z|:=\sqrt{\mathrm{Re}(z)^2+\mathrm{Im}(z)^2}\geq 0$ den *Betrag* von z (dies entspricht der Norm in $\mathbb{R}^2$).

Bemerkung 11.25. Im Gegensatz zu $\mathbb{R}$ lassen sich komplexe Zahlen nicht sinnvoll anordnen, denn wegen $\mathrm{i}^2=(-\mathrm{i})^2=-1<0$ kann weder $\mathrm{i}\geq 0$ noch $-\mathrm{i}\geq 0$ gelten.[20]

Lemma 11.26. *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 $\zeta_1^n=\ldots=\zeta_n^n=z$.*

Beweis. Sei $z_0:=\frac{1}{|z|}z\in\mathbb{C}$ und $P(z_0)=\left(\begin{smallmatrix} a & -b\\ b & a\end{smallmatrix}\right)\in\mathbb{R}^{2\times 2}$ wie im Beweis von Lemma 11.23. Wegen $a^2+b^2=|z_0|^2=1$ ist $P(z_0)=D(\varphi)\in\mathrm{O}(2,\mathbb{R})$ 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

$$D(\varphi_k)^n=D(n\varphi_k)=D(\varphi+2k\pi)=D(\varphi)=P(z_0).$$

[19] Man kann die Axiome auch direkt mit der Definition nachrechnen.
[20] „sinnvoll“ wäre die Eigenschaft: $a,b\geq 0\Longrightarrow ab\geq 0$

Die Zahlen $z_k := \cos\varphi_k + \mathrm{i}\sin\varphi_k \in \mathbb{C}$ 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 $k = 1, \ldots, n$. Sei nun $\zeta_k = \zeta_l$ für $1 \leq k \leq l \leq n$. 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 $\zeta_1, \ldots, \zeta_n$ paarweise verschieden. □

Beispiel 11.27. 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, $-\mathrm{i}$.

Folgerung 11.28. *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 10.35. Die Nullstellen von $X^k - X$ sind die $(k-1)$-ten Einheitswurzeln und 0. Nach Lemma 10.18 zerfällt μ_A in paarweise verschiedene Linearfaktoren. Die Behauptung folgt nun aus Satz 10.41. □

Beispiel 11.29. 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 11.30. Die Abbildung $\mathbb{C} \to \mathbb{C}$, $a + b\mathrm{i} \mapsto a - b\mathrm{i} =: \overline{a + b\mathrm{i}}$ heißt *komplexe Konjugation*.

Lemma 11.31. *Für $z, w \in \mathbb{C}$ gilt $|z|^2 = z\overline{z}$, $\overline{z+w} = \overline{z} + \overline{w}$ und $\overline{zw} = \overline{z} \cdot \overline{w}$.*

Beweis. Für $z = a + b\mathrm{i}$ und $w = c + d\mathrm{i}$ gilt

$$\begin{aligned} |z|^2 &= a^2 + b^2 = (a + b\mathrm{i})(a - b\mathrm{i}) = z\overline{z}, \\ \overline{z+w} &= a + c - (b+d)\mathrm{i} = a - b\mathrm{i} + c - d\mathrm{i} = \overline{z} + \overline{w}, \\ \overline{zw} &= ac - bd - (ad + bc)\mathrm{i} = (a - b\mathrm{i})(c - d\mathrm{i}) = \overline{z} \cdot \overline{w}. \end{aligned}$$ □

Bemerkung 11.32.

(a) Für Matrizen $A = (a_{ij}) \in \mathbb{C}^{n\times m}$ definieren wir $\overline{A} := (\overline{a_{ij}}) \in \mathbb{C}^{n\times m}$. Für $B = (b_{ij}) \in \mathbb{C}^{m\times k}$ gilt

$$\overline{AB} = \Big(\overline{\sum_{k=1}^{m} a_{ik}b_{kj}}\Big)_{ij} \overset{11.31}{=} \Big(\sum_{k=1}^{m} \overline{a_{ik}b_{kj}}\Big)_{ij} \overset{11.31}{=} \Big(\sum_{k=1}^{m} \overline{a_{ik}}\,\overline{b_{kj}}\Big)_{ij} = \overline{A} \cdot \overline{B}.$$

(b) Nach Lemma 11.26 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 aber sogar jedes nicht-konstante Polynom in $\mathbb{C}[X]$ eine Nullstelle.[21] Der Beweis benutzt Analysis (genauer die Vollständigkeit von $\mathbb{R}$) und ist für diese Vorlesung zu schwierig.

Satz 11.33 (Fundamentalsatz der Algebra)**.** *Jedes Polynom $\alpha \in \mathbb{C}[X] \setminus \mathbb{C}$ besitzt eine Nullstelle in $\mathbb{C}$.*

Folgerung 11.34. *Jedes normierte Polynom $\alpha \in \mathbb{C}[X] \setminus \mathbb{C}$ zerfällt in Linearfaktoren.*

[21] Man sagt: $\mathbb{C}$ ist *algebraisch abgeschlossen*.

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 10.16 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. □

Beispiel 11.35. 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 $\lim_{x\to\pm\infty} \alpha(x) = \pm\infty$. Nach dem Zwischenwertsatz der Analysis besitzt die stetige Funktion $\mathbb{R} \to \mathbb{R}$, $x \mapsto \alpha(x)$ 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

$$x_{n+1} := x_n - \frac{\alpha(x_n)}{\alpha'(x_n)} \qquad (n \geq 0),$$

also $x_1 = 0.50847\ldots$, $x_2 = 0.50849948\ldots$ 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.

Bemerkung 11.36. Sei $x \in \mathbb{C}$ eine Nullstelle von $\alpha = \sum a_k X^k \in \mathbb{C}[X]$. Wir definieren $\overline{\alpha} := \sum \overline{a_k} X^k \in \mathbb{C}[X]$. Es gilt dann

$$\overline{\alpha}(\overline{x}) = \sum \overline{a_k x^k} = \overline{\alpha(x)} = \overline{0} = 0,$$

d. h. $\overline{x}$ ist eine Nullstelle von $\overline{\alpha}$. Im Fall $\alpha \in \mathbb{R}[X]$ gilt also: $\alpha(x) = 0 \iff \alpha(\overline{x}) = 0$. Ggf. ist

$$(X - x)(X - \overline{x}) = X^2 - (x + \overline{x})X + x\overline{x} = X^2 - 2\mathrm{Re}(x)X + |x|^2 \in \mathbb{R}[X]$$

ein Teiler von α.

11.5 Der Spektralsatz

Satz 11.37 (Spektralsatz[22])**.** *Sei V ein euklidischer Raum und $f \in \mathrm{End}(V)$. 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 11.19 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 $v = (v_1, \ldots, v_n)^{\mathrm{t}} \in \mathbb{C}^{n\times 1}$ ein entsprechender Eigenvektor. Wegen

$$\lambda \sum_{i=1}^{n} |v_i|^2 = \lambda \overline{v}^{\mathrm{t}} v = \overline{v}^{\mathrm{t}} A v = (\overline{v}^{\mathrm{t}} A v)^{\mathrm{t}} = v^{\mathrm{t}} A^{\mathrm{t}} \overline{v} = v^{\mathrm{t}} A \overline{v} = v^{\mathrm{t}} \overline{Av} = \overline{\lambda} v^{\mathrm{t}} \overline{v} = \overline{\lambda} \sum_{i=1}^{n} |v_i|^2$$

ist $\lambda = \overline{\lambda} \in \mathbb{R}$. 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 $U := \langle b_1 \rangle^{\perp}$ gilt $V = \langle b_1 \rangle \oplus U$ nach Lemma 11.15.

[22] Die Menge der Eigenwerte einer Matrix bezeichnet man als *Spektrum*.

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 $\mathrm{End}(U)$. Offenbar ist g auch symmetrisch. Wegen $\dim U = n - 1$ besitzt U nach Induktion eine Orthonormalbasis $b_2, \ldots, b_n \in U$ aus Eigenvektoren von g. Wegen $f(b_i) = g(b_i)$ sind $b_2, \ldots, b_n$ auch Eigenvektoren von f. Insgesamt ist $b_1, \ldots, b_n$ eine Orthonormalbasis von V aus Eigenvektoren von f. □

Bemerkung 11.38.

(a) Die Matrix-Version des Spektralsatzes lautet: Genau dann ist $A \in \mathbb{R}^{n\times n}$ symmetrisch, wenn eine Matrix $S \in \mathrm{O}(n, \mathbb{R})$ mit $S^{\mathrm{t}}AS = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ existiert. Eine komplexe Version des Spektralsatzes besagt, dass Matrizen $A \in \mathbb{C}^{n\times n}$ mit $\overline{A}A^{\mathrm{t}} = A^{\mathrm{t}}\overline{A}$ diagonalisierbar sind (ohne Beweis).

(b) Sei $f \in \mathrm{End}(V)$ 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 11.39. In Beispiel 8.4 hatten wir die Abbildung $f \in \mathrm{End}(\mathbb{R}^3)$ mit symmetrischer Matrix

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

untersucht. Es gilt

$$E_1(A) = \left\langle \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right\rangle, \qquad E_4(A) = \left\langle \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\rangle$$

(beachte: $\mathrm{tr}(A) = 6 = 1 + 1 + 4$). Das Gram-Schmidt-Verfahren für $E_1(A)$ liefert

$$b_1 := (1, 0, -1), \qquad b_2 := (0, 1, -1) - \frac{1}{2}(1, 0, -1) = \frac{1}{2}(-1, 2, -1).$$

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 $S^{\mathrm{t}}AS = S^{-1}AS = \mathrm{diag}(1, 1, 4)$. Man kann das Ergebnis nutzen, um Wurzeln von Matrizen zu ziehen. Für $\sqrt{A} := S\,\mathrm{diag}(1, 1, 2)S^{\mathrm{t}}$ gilt nämlich

$$\sqrt{A}^2 = S\,\mathrm{diag}(1, 1, 2)S^{\mathrm{t}}S\,\mathrm{diag}(1, 1, 2)S^{\mathrm{t}} = S\,\mathrm{diag}(1, 1, 2)^2S^{\mathrm{t}} = S\,\mathrm{diag}(1, 1, 4)S^{\mathrm{t}} = A.$$

Satz 11.40 (Euler)**.** *Sei V ein 3-dimensionaler euklidischer Raum und $f \in \mathrm{O}(V)$. 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 11.35. Nach Bemerkung 11.20 ist $\lambda \in \{\pm 1\}$. Sei b_1 ein entsprechender normierter Eigenvektor und $U := \langle b_1 \rangle^\perp$. 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 $\mathrm{O}(U)$. Für eine Orthonormalbasis $C := \{b_2, b_3\}$ und U ist ${}_C[g]_C \in \mathrm{O}(2, \mathbb{R})$ nach Lemma 11.19. Im Fall $\det(g) = 1$ ist ${}_C[g]_C = D(\varphi)$ nach Beispiel 11.21 und die Behauptung folgt mit $B := \{b_1, b_2, b_3\}$. Sei nun $\det(g) = -1$. Nach Beispiel 11.21 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. □

Bemerkung 11.41. Die Matrizen in Satz 11.40 mit Determinante 1 (also von der Form $\mathrm{diag}(1, D(\varphi))$) 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. Man kann zeigen, dass im n-dimensionalen Raum jede orthogonale Abbildung eine Komposition von höchstens n Spiegelungen (an Hyperebenen) ist.

Beispiel 11.42. 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.

12 Die Jordansche Normalform

12.1 Haupträume

Bemerkung 12.1. Nach Satz 10.28 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 $K = \mathbb{C}$ tritt das erste Problem nach dem Fundamentalsatz der Algebra nicht auf. Um das zweite Problem zu umgehen, ersetzen wir Eigenräume durch größere Unterräume.

Definition 12.2. Sei V ein n-dimensionaler K-Vektorraum und $f \in \mathrm{End}(V)$.

- Ein Unterraum $U \leq V$ heißt *f-invariant*, falls $f(U) \subseteq U$ gilt.

- Für einen Eigenwert $\lambda \in K$ von f nennt man

$$H_\lambda(f) := \operatorname{Ker}\big((f - \lambda \operatorname{id}_V)^n\big) \leq V$$

den *Hauptraum* von f zu λ. Ist λ Eigenwert einer Matrix $A \in K^{n\times n}$, so definiert man analog $H_\lambda(A) := \operatorname{Ker}\big((A - \lambda 1_n)^n\big)$.

Beispiel 12.3.

(a) Für $f \in \operatorname{End}(V)$ sind $\operatorname{Ker}(f)$ und $f(V)$ stets f-invariante Unterräume, denn $f(\operatorname{Ker}(f)) = \{0\} \subseteq \operatorname{Ker}(f)$ und $f(f(V)) \subseteq f(V)$.

(b) Sei $f \in \operatorname{End}(K^2)$ mit $A := [f] = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$. 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 12.4.

(a) Sei $U \leq V$ ein f-invarianter Unterraum. Wir ergänzen eine Basis von U zu einer Basis B von V. Dann ist ${}_B[f]_B = \left(\begin{smallmatrix} A & C \\ 0 & D \end{smallmatrix}\right)$, wobei A die Darstellungsmatrix der Einschränkung $f_{|U}$ ist. 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.

(b) Für $v \in E_\lambda(f)$ gilt

$$(f - \lambda \operatorname{id}_V)^n(v) = (f - \lambda \operatorname{id}_V)^{n-1}\big((f - \lambda \operatorname{id}_V)(v)\big) = (f - \lambda \operatorname{id}_V)^{n-1}(0) = 0.$$

Dies zeigt $E_\lambda(f) \subseteq H_\lambda(f)$. Offenbar ist f mit $f - \lambda \operatorname{id}$ und $(f - \lambda \operatorname{id})^n$ vertauschbar. Für $v \in H_\lambda(f)$ gilt daher

$$(f - \lambda \operatorname{id}_V)^n(f(v)) = \big((f - \lambda \operatorname{id}_V)^n \circ f\big)(v) = f\big((f - \lambda \operatorname{id}_V)^n(v)\big) = f(0) = 0.$$

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 12.5 (Fitting). *Sei V ein n-dimensionaler K-Vektorraum und $f \in \operatorname{End}(V)$. Dann ist*

$$\boxed{V = f^n(V) \oplus \operatorname{Ker}(f^n)}$$

eine Zerlegung in f-invariante Unterräume.

Beweis. Für $v \in \operatorname{Ker}(f^k)$ gilt $f^{k+1}(v) = f(f^k(v)) = f(0) = 0$. Dies zeigt $\operatorname{Ker}(f) \leq \operatorname{Ker}(f^2) \leq \ldots$. Da die Dimension dieser Unterräume durch n beschränkt ist, existiert ein $k \leq n$ mit $\operatorname{Ker}(f^k) = \operatorname{Ker}(f^{k+1})$. Für $v \in \operatorname{Ker}(f^{k+2})$ gilt $f^{k+1}(f(v)) = f^{k+2}(v) = 0$, also $f(v) \in \operatorname{Ker}(f^{k+1}) = \operatorname{Ker}(f^k)$. Dies zeigt $f^{k+1}(v) = 0$ und $v \in \operatorname{Ker}(f^{k+1})$. Induktiv erhält man

$$\operatorname{Ker}(f^k) = \operatorname{Ker}(f^{k+1}) = \ldots = \operatorname{Ker}(f^n) = \operatorname{Ker}(f^{n+1}) = \ldots.$$

Für $v \in \operatorname{Ker}(f^n) \cap f^n(V)$ existiert ein $w \in V$ mit $v = f^n(w)$. Wegen $f^{2n}(w) = f^n(v) = 0$ ist $w \in \operatorname{Ker}(f^{2n}) = \operatorname{Ker}(f^n)$ und es folgt $v = f^n(w) = 0$. Also ist $\operatorname{Ker}(f^n) \cap f^n(V) = \{0\}$ und der Rangsatz zeigt $V = \operatorname{Ker}(f^n) \oplus f^n(V)$. Wegen $f(\operatorname{Ker}(f^n)) \subseteq \operatorname{Ker}(f^{n-1}) \subseteq \operatorname{Ker}(f^n)$ und $f(f^n(V)) = f^n(f(V)) \subseteq f^n(V)$ sind beide Unterräume f-invariant. □

Beispiel 12.6. Seien $\lambda, \lambda' \in K$ verschiedene Eigenwerte von $f \in \mathrm{End}(V)$. Für $U := H_\lambda(f) \cap H_{\lambda'}(f)$ gilt $(f - \lambda \,\mathrm{id})^n(U) = \{0\} = (f - \lambda' \,\mathrm{id})^n(U)$. Das Minimalpolynom der Einschränkung $g := f_{|U} \in \mathrm{End}(U)$ teilt daher $(X - \lambda)^n$ und $(X - \lambda')^n$ nach Lemma 10.35. Aus Lemma 10.18 folgt $\mu_g = 1$ und $U = \{0\}$.

Satz 12.7 (Hauptraumzerlegung)**.** *Sei V ein K-Vektorraum und $f \in \mathrm{End}(V)$, sodass μ_f (oder χ_f) in Linearfaktoren zerfällt. Dann gilt*

$$V = H_{\lambda_1}(f) \oplus \ldots \oplus H_{\lambda_k}(f)$$

für die verschiedenen Eigenwerte $\lambda_1, \ldots, \lambda_k \in K$ von f.

Beweis. Zerfällt χ_f in Linearfaktoren, so auch μ_f nach Cayley-Hamilton und Lemma 10.18. 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 10.39. Für die Abbildung $g := f - \lambda \,\mathrm{id}_V$ gilt

$$V = g^n(V) \oplus \mathrm{Ker}(g^n) = g^n(V) \oplus H_\lambda(f)$$

nach Fitting. Für $U := g^n(V)$ gilt $f(U) = (g + \lambda \,\mathrm{id})(U) \subseteq g(U) + U \subseteq U$, d. h. U ist f-invariant. Wir betrachten nun die Einschränkung $h := f_{|U} \in \mathrm{End}(U)$. Nach Lemma 10.35 ist $\mu_h \mid \mu_f$ und μ_h zerfällt in Linearfaktoren (Lemma 10.18). Wegen $\dim H_\lambda(f) \geq \dim E_\lambda(f) \geq 1$ ist $\dim U < n$. Nach Induktion gilt daher

$$U = H_{\lambda_2}(h) \oplus \ldots \oplus H_{\lambda_k}(h)$$

für die verschiedenen Eigenwerte $\lambda_2, \ldots, \lambda_k$ von h. Wir wollen $H_{\lambda_i}(h) = H_{\lambda_i}(f)$ für $i = 2, \ldots, k$ zeigen. Wegen $E_{\lambda_i}(h) \subseteq E_{\lambda_i}(f)$ sind $\lambda_2, \ldots, \lambda_k$ auch Eigenwerte von f. Wegen $E_{\lambda_i}(h) \cap H_\lambda(f) \leq U \cap H_\lambda(f) = \{0\}$ ist $\lambda \neq \lambda_i$ für $i = 2, \ldots, k$. Sei $v \in H_{\lambda_i}(f)$ für ein $i \geq 2$. Dann existieren $u \in U$ und $w \in H_\lambda(f)$ mit $v = u + w$. Es folgt

$$0 = (f - \lambda_i \,\mathrm{id})^n(v) = (f - \lambda_i \,\mathrm{id})^n(u) + (f - \lambda_i \,\mathrm{id})^n(w) \in U \oplus H_\lambda(f).$$

Aus Lemma 8.9 erhält man $(f - \lambda_i \,\mathrm{id})^n(u) = 0 = (f - \lambda_i \,\mathrm{id})^n(w)$. Nach Beispiel 12.6 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

$$V = H_\lambda(f) \oplus U = H_\lambda(f) \oplus H_{\lambda_2}(h) \oplus \ldots \oplus H_{\lambda_k}(h) = H_{\lambda_1}(f) \oplus \ldots \oplus H_{\lambda_k}(f). \qquad \square$$

Beispiel 12.8. Sei

$$A := \begin{pmatrix} . & 1 & 1 & 1 \\ 1 & 1 & 1 & . \\ 1 & 1 & . & 1 \\ . & . & 1 & 1 \end{pmatrix} \in \mathbb{F}_2^{4\times 4}.$$

Da 0 und 1 die einzig möglichen Eigenwerte sind, berechnen wir probeweise

$$A^2 = \begin{pmatrix} . & . & . & . \\ . & 1 & . & . \\ 1 & . & 1 & . \\ 1 & 1 & 1 & . \end{pmatrix}, \qquad (A - 1_4)^2 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & . & 1 & . \\ 1 & 1 & 1 & 1 \\ . & . & 1 & . \end{pmatrix}^2 = \begin{pmatrix} 1 & . & . & . \\ . & . & . & . \\ 1 & . & . & . \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

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

$$\mathbb{F}_2^4 = H_0(A) \oplus H_1(A) = \langle (0,0,0,1), (1,0,1,0) \rangle \oplus \langle (0,1,1,0), (0,1,0,1) \rangle.$$

12.2 Jordanblöcke

Bemerkung 12.9. Nach der Hauptraumzerlegung interessieren wir uns für Basen von Haupträumen.

Definition 12.10. Für $\lambda \in K$ und $n \geq 1$ nennt man

$$J_n(\lambda) := \begin{pmatrix} \lambda & & & 0 \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ 0 & & 1 & \lambda \end{pmatrix} \in K^{n \times n}$$

einen *Jordanblock*.[23]

Bemerkung 12.11. Offenbar hat $A := J_n(\lambda)$ das charakteristisches 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 10.28). Für den Standardbasisvektor $e_1 \in K^n$ gilt $Ae_1^{\mathrm{t}} = (*, 1, 0, \ldots, 0)^{\mathrm{t}}$ und induktiv

$$A^k e_1^{\mathrm{t}} = A(A^{k-1}e_1^{\mathrm{t}}) = A(\underbrace{*, \ldots, *}_{k-1}, 1, 0, \ldots, 0)^{\mathrm{t}} = (\underbrace{*, \ldots, *}_{k}, 1, 0, \ldots, 0)^{\mathrm{t}}.$$

Also sind $e_1^{\mathrm{t}}, Ae_1^{\mathrm{t}}, \ldots, A^{n-1}e_1^{\mathrm{t}}$ linear unabhängig. Daher müssen auch die Matrizen $1_n, A, \ldots, A^{n-1}$ linear unabhängig sein. Dies zeigt $\deg \mu_A \geq n$. Aus Cayley-Hamilton folgt $\mu_A = \chi_A = (X - \lambda)^n$.

Satz 12.12. *Sei V ein K-Vektorraum und $f \in \mathrm{End}(V)$ mit $f^m = 0$ für ein $m \in \mathbb{N}$.*[24] *Dann existiert eine Basis B von V, sodass*

$${}_B[f]_B = \begin{pmatrix} J_{n_1}(0) & & 0 \\ & \ddots & \\ 0 & & J_{n_s}(0) \end{pmatrix}.$$

Die Zahlen $n_1 \geq \ldots \geq n_s \geq 1$ sind durch die Gleichungen

$$\boxed{\left|\{1 \leq i \leq s : n_i \geq k\}\right| = \mathrm{rk}(f^{k-1}) - \mathrm{rk}(f^k)} \qquad (k = 1, \ldots, m) \tag{12.1}$$

eindeutig bestimmt. Insbesondere ist $s = \dim \mathrm{Ker}(f)$.

Beweis. O. B. d. A. sei $f^{m-1} \neq 0$. Für $k \in \mathbb{N}$ gilt $f(\mathrm{Ker}(f^k)) \subseteq \mathrm{Ker}(f^{k-1})$. Nach Folgerung 4.16 existieren Unterräume $U_1, \ldots, U_m$ mit

$$\begin{aligned} V = \mathrm{Ker}(f^m) &= \mathrm{Ker}(f^{m-1}) \oplus U_1, \\ \mathrm{Ker}(f^{m-1}) &= \big(\mathrm{Ker}(f^{m-2}) + f(U_1)\big) \oplus U_2, \\ &\vdots \\ \mathrm{Ker}(f) &= \big(f^{m-1}(U_1) + \ldots + f(U_{m-1})\big) \oplus U_m. \end{aligned}$$

Wir zeigen, dass alle Summen direkt sind. Dies ist für die erste Summe gegeben. Sei induktiv bereits gezeigt:

$$\mathrm{Ker}(f^{m-k+1}) = \mathrm{Ker}(f^{m-k}) \oplus f^{k-1}(U_1) \oplus f^{k-2}(U_2) \oplus \ldots \oplus f(U_{k-1}) \oplus U_k. \tag{12.2}$$

[23] In manchen Büchern benutzt man $J_n(\lambda)^{\mathrm{t}}$. Das macht keinen wesentlichen Unterschied (Aufgabe 54).
[24] Abbildungen mit dieser Eigenschaft nennt man *nilpotent*.

Sei $w + f^k(u_1) + \ldots + f(u_k) = 0$ mit $w \in \operatorname{Ker}(f^{m-k-1})$ und $u_i \in U_i$ für $i = 1, \ldots, k$. Dann folgt

$$\begin{aligned} f^{m-k}\big(f^{k-1}(u_1) + \ldots + u_k\big) &= f^{m-k-1}\big(f^k(u_1) + \ldots + f(u_k)\big) \\ &= f^{m-k-1}\big(w + f^k(u_1) + \ldots + f(u_k)\big) = f^{m-k-1}(0) = 0 \end{aligned}$$

und

$$f^{k-1}(u_1) + \ldots + u_k \in \operatorname{Ker}(f^{m-k}) \cap \big(f^{k-1}(U_1) \oplus \ldots \oplus f(U_{k-1}) \oplus U_k\big) \overset{(12.2)}{=} \{0\}.$$

Dies zeigt $f^{k-1}(u_1) = \ldots = u_k = 0$ (Lemma 8.9) und es folgt $w = 0$. Also ist

$$\operatorname{Ker}(f^{m-k}) = \operatorname{Ker}(f^{m-k-1}) \oplus f^k(U_1) \oplus f^{k-1}(U_2) \oplus \ldots \oplus f(U_k) \oplus U_{k+1}$$

wie gewünscht.

Insgesamt ist

$$V = U_1 \oplus f(U_1) \oplus \ldots \oplus f^{m-1}(U_1) \oplus U_2 \oplus f(U_2) \oplus \ldots \oplus f^{m-2}(U_2) \oplus \ldots \oplus U_m.$$

Sei $b_{i1}, \ldots, b_{ik_i}$ eine Basis von U_i für $i = 1, \ldots, m$ (der Fall $U_i = \{0\}$ mit $k_i = 0$ ist zugelassen). Wegen $\operatorname{Ker}(f^j) \cap U_i = \{0\}$ ist die Einschränkung $f^j_{|U_i}$ für $j = 1, \ldots, m - i$ injektiv (Lemma 7.7). Insbesondere ist $f^j(b_{i1}), \ldots, f^j(b_{ik_i})$ eine Basis von $f^j(U_i)$. Also ist

$$B := \bigcup_{i=1}^{m} \bigcup_{j=1}^{k_i} \big\{b_{ij}, f(b_{ij}), \ldots, f^{m-i}(b_{ij})\big\}$$

eine Basis von V. Wegen $U_i \subseteq \operatorname{Ker}(f^{m-i+1})$ ist $f^{m-i+1}(b_{ij}) = 0$. Somit entsprechen die Elemente $b_{ij}, f(b_{ij}), \ldots, f^{m-i}(b_{ij})$ dem Jordanblock $J_{m-i+1}(0)$ in ${}_B[f]_B$. Insgesamt hat ${}_B[f]_B$ nun die gewünschte Form. Außerdem ist

$$\begin{aligned} k_1 + \ldots + k_l &= \dim\big(f^{l-1}(U_1) \oplus f^{l-2}(U_2) \oplus \ldots \oplus U_l\big) = \dim \operatorname{Ker}(f^{m-l+1}) - \dim \operatorname{Ker}(f^{m-l}) \\ &\overset{7.27}{=} \operatorname{rk}(f^{m-l}) - \operatorname{rk}(f^{m-l+1}) \end{aligned}$$

die Anzahl der Jordanblöcke $J_{n_i}(0)$ mit $n_i \geq m - l + 1$. Indem man $m - l + 1$ durch k ersetzt, folgt (12.1). Die letzte Behauptung erhält man mit $k = 1$ in (12.1). □

Bemerkung 12.13. In der Situation von Satz 12.12 gilt

$$\begin{aligned} \big|\{1 \leq i \leq s : n_i = k\}\big| &= \big|\{1 \leq i \leq s : n_i \geq k\}\big| - \big|\{1 \leq i \leq s : n_i \geq k+1\}\big| \\ &= \operatorname{rk}(f^{k-1}) + \operatorname{rk}(f^{k+1}) - 2\operatorname{rk}(f^k) \end{aligned}$$

für $k = 1, \ldots, n$. Insbesondere ist $2\operatorname{rk}(f^k) \leq \operatorname{rk}(f^{k+1}) + \operatorname{rk}(f^{k-1})$, d. h. die Folge $\operatorname{rk}(f)$, $\operatorname{rk}(f^2), \ldots$ kann nicht zu „schnell“ fallen.

Beispiel 12.14. Sei $f \in \operatorname{End}(\mathbb{R}^5)$ 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

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

Also ist $\operatorname{rk}(f^2) = 1$. Wie im Beweis von Satz 12.12 können wir $b_1 := e_1 \notin \operatorname{Ker}(f^2)$ und $U_1 := \langle b_1 \rangle$ mit $\mathbb{R}^5 = \operatorname{Ker}(f^2) \oplus \langle b_1 \rangle$ wählen. Weiter ist

$$A \sim \begin{pmatrix} 1 & -2 & 0 & 0 & 1 \\ 0 & 1 & -1 & -1 & 0 \\ 0 & 2 & -1 & -1 & 0 \\ 0 & -2 & 1 & 1 & 0 \\ 0 & -1 & 1 & 1 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -2 & -2 & 1 \\ 0 & 1 & -1 & -1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\operatorname{Ker}(f) \oplus f(U_1) = \left\langle \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle.$$

Mit $b_2 := e_3 \in \operatorname{Ker}(f^2) \setminus \big(\operatorname{Ker}(f) \oplus f(U_1)\big)$ und $U_2 := \langle b_2 \rangle$ gilt $\operatorname{Ker}(f^2) = \operatorname{Ker}(f) \oplus f(U_1) \oplus U_2$. Schließlich ist $\operatorname{Ker}(f) = f^2(U_1) \oplus f(U_2)$ (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 $\operatorname{diag}\big(J_3(0), J_2(0)\big)$.

Satz 12.15 (JORDANsche Normalform)**.** *Sei V ein $\mathbb{C}$-Vektorraum und $f \in \operatorname{End}(V)$. Dann existiert eine Basis B von V mit*

$${}_B[f]_B = \begin{pmatrix} J_{n_1}(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J_{n_s}(\lambda_s) \end{pmatrix},$$

wobei $\lambda_1, \ldots, \lambda_s$ die Eigenwerte von f sind (möglicherweise mit Vielfachheiten). Die Jordanblöcke $J_{n_i}(\lambda_i)$ sind bis auf die Reihenfolge eindeutig bestimmt.[25]

Beweis. Wir fügen alle Puzzleteile zusammen: Nach Folgerung 11.34 zerfällt μ_f in Linearfaktoren. Nach Satz 12.7 gilt

$$V = H_{\lambda_1}(f) \oplus \ldots \oplus H_{\lambda_k}(f),$$

wobei $\lambda_1, \ldots, \lambda_k$ die verschiedenen Eigenwerte von f sind. Sei $H_i := H_{\lambda_i}(f)$ für $i = 1, \ldots, k$. Für $g_i := (f - \lambda_i \operatorname{id}_V)_{|H_i}$ gilt $g_i^n = 0$. Nach Satz 12.12 existiert eine Basis B_i von H_i, sodass ${}_{B_i}[g_i]_{B_i} = \operatorname{diag}\big(J_{m_1}(0), \ldots, J_{m_t}(0)\big)$ gilt. Nach Satz 7.16 ist

$$\begin{aligned} {}_{B_i}[f_{|H_i}]_{B_i} &= {}_{B_i}[g_i + \lambda_i \operatorname{id}_{H_i}]_{B_i} = {}_{B_i}[g_i]_{B_i} + \lambda_i {}_{B_i}[\operatorname{id}_{H_i}]_{B_i} = \operatorname{diag}\big(J_{m_1}(0), \ldots, J_{m_t}(0)\big) + \lambda_i 1 \\ &= \operatorname{diag}\big(J_{m_1}(\lambda_i), \ldots, J_{m_t}(\lambda_i)\big). \end{aligned}$$

Also ist $B := B_1 \cup \ldots \cup B_k$ 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 12.13 angewendet auf g_i. □

[25] Anders als in Satz 12.12 lässt sich keine *kanonische* (d. h. naheliegende) Reihenfolge der Jordanblöcke angeben, da man $\mathbb{C}$ nicht sinnvoll ordnen kann (Bemerkung 11.25).

Bemerkung 12.16.

(a) Die Matrix-Version Satz 12.15 lautet: Für jede Matrix $A \in \mathbb{C}^{n\times n}$ existiert ein $S \in \mathrm{GL}(n,\mathbb{C})$ mit $S^{-1}AS = \operatorname{diag}\big(J_{n_1}(\lambda_1),\dots,J_{n_s}(\lambda_s)\big)$.

(b) Sei λ ein Eigenwert von f und seien $n_1 \geq \ldots \geq n_k$ die Größen der Jordanblöcke $J_{n_i}(\lambda_i)$ mit $\lambda_i = \lambda$ in der Jordanschen Normalform von f. Nach Satz 12.12 (angewendet auf g_i im obigen Beweis) ist k die geometrische Vielfachheit von λ. Da die Jordansche Normalform eine untere Dreiecksmatrix ist, ist $n_1 + \ldots + n_k$ die algebraische Vielfachheit von λ als Nullstelle von χ_f. Nach Satz 10.28 ist f also genau dann diagonalisierbar, wenn die Jordansche Normalform eine Diagonalmatrix ist (d. h. $n_1 = \ldots = n_s = 1$). Da μ_f nicht von der Basiswahl abhängt, gilt

$$0 = \mu_f\big(\operatorname{diag}(J_{n_1}(\lambda_1),\dots,J_{n_s}(\lambda_s))\big) = \operatorname{diag}\big(\mu_f(J_{n_1}(\lambda_1)),\dots,\mu_f(J_{n_s}(\lambda_s))\big).$$

Daher ist μ_f durch das Minimalpolynom von $J_{n_i}(\lambda_i)$ für $i = 1,\dots,s$ teilbar. Nach Bemerkung 12.11 ist n_1 die algebraische Vielfachheit von λ als Nullstelle von μ_f.

Beispiel 12.17. Sei $f \in \mathrm{End}(\mathbb{C}^3)$ mit

$$A := [f] = \begin{pmatrix} 5 & 0 & 1 \\ -5-\mathrm{i} & -\mathrm{i} & -1 \\ -9 & 0 & -1 \end{pmatrix}.$$

Durch Laplace-Entwicklung nach der zweiten Spalte erhält man

$$\chi_f = \chi_A = (X+\mathrm{i})\big((X-5)(X+1)+9\big) = (X+\mathrm{i})(X^2-4X+4) = (X+\mathrm{i})(X-2)^2.$$

Also sind $\lambda_1 = 2$ und $\lambda_2 = -\mathrm{i}$ die Eigenwerte von f. Offensichtlich ist $b_3 := e_2$ ein Eigenvektor zu λ_2. Wegen

$$A - 2\cdot 1_3 = \begin{pmatrix} 3 & 0 & 1 \\ -5-\mathrm{i} & -2-\mathrm{i} & -1 \\ -9 & 0 & -3 \end{pmatrix} \sim \begin{pmatrix} 3 & 0 & 1 \\ -5-\mathrm{i} & -2-\mathrm{i} & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

ist $\operatorname{rk}(f - 2\,\mathrm{id}) = 2$, d. h. $\lambda_1 = 2$ hat geometrische Vielfachheit 1. Also ist f nicht diagonalisierbar und die Jordansche Normalform von f muss $\operatorname{diag}\big(J_2(2), J_1(-\mathrm{i})\big)$ sein. Wegen

$$(A - 2\cdot 1_3)^2 = \begin{pmatrix} 0 & 0 & 0 \\ 3+4\mathrm{i} & 3+4\mathrm{i} & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

können wir $b_1 := e_3 \in \operatorname{Ker}\big((f-2\,\mathrm{id})^2\big) \setminus \operatorname{Ker}(f - 2\,\mathrm{id})$ wie in Beispiel 12.14 wählen. Für

$$b_2 := (f - 2\,\mathrm{id})(b_1) = (1,-1,-3)$$

gilt $f(b_1) = 2b_1 + b_2$ und $f(b_2) = (f-2\,\mathrm{id})(b_2) + 2b_2 = 2b_2$. Für die Basis $B := \{b_1, b_2, b_3\}$ erhält man

$${}_B[f]_B = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & -\mathrm{i} \end{pmatrix} = \operatorname{diag}\big(J_2(2), J_1(-\mathrm{i})\big).$$

Bemerkung 12.18. Über beliebigen Körpern K zerfällt das Minimalpolynom μ_f nicht unbedingt in Linearfaktoren. Man kann μ_f aber „so weit wie möglich“ faktorisieren, sagen wir $\mu_f = \gamma_1^{a_1} \ldots \gamma_k^{a_k}$ mit paarweise verschiedenen normierten Polynomen $\gamma_1,\dots,\gamma_k \in K[X]$. Es lässt sich zeigen, dass

die Faktoren $\gamma_1, \ldots, \gamma_k$ dabei bis auf die Reihenfolge eindeutig bestimmt sind (man hat also eine „Primfaktorzerlegung“ in $K[X]$). Analog zur Hauptraumzerlegung existiert eine f-invariante Zerlegung

$$V = \operatorname{Ker}\big(\gamma_1^{a_1}(f)\big) \oplus \ldots \oplus \operatorname{Ker}\big(\gamma_k^{a_k}(f)\big).$$

Es existiert nun eine Basis B von V, sodass

$${}_B[f]_B = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_l \end{pmatrix} \quad \text{mit} \quad A_i = \begin{pmatrix} 0 & & 0 & * \\ 1 & \ddots & & \vdots \\ & \ddots & 0 & * \\ 0 & & 1 & * \end{pmatrix} \quad \text{für } i = 1, \ldots, l$$

(vgl. Aufgabe 40). Im schlechtesten Fall ($l = 1$) sind wie bei der Jordanschen Normalform nur $2n - 1$ Matrixeinträge von ${}_B[f]_B$ „belegt“.

Beispiel 12.19. Sei $f \in \operatorname{End}(\mathbb{F}_2^3)$ mit

$$A := [f] = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \in \mathbb{F}_2^{3\times 3}.$$

Mit der Sarrus-Regel berechnet man $\chi_f = \chi_A = X(X+1)(X+1)+X+X+1 = X^3+X+1$. Da χ_f keine Nullstellen besitzt (nur 0 und 1 kommen in Frage), kann χ_f auch keine „echten“ Teiler besitzen. Nach Cayley-Hamilton muss also $\mu_f = \chi_f$ gelten. Offenbar sind e_1, $f(e_1) = e_2$ und $f^2(e_1) = f(e_2) = (1, 1, 1)$ linear unabhängig. Wegen $f(1, 1, 1) = (1, 1, 0) = e_1 + e_2$ erhält man

$${}_B[f]_B = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

wobei $B = \big\{e_1, e_2, (1, 1, 1)\big\}$.

Aufgaben

Aufgabe 26. 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 27. 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 28. Sei $A = \left(\begin{smallmatrix} A_1 & A_2 \\ 0 & A_3 \end{smallmatrix}\right) \in K^{(n+m)\times(n+m)}$ mit $A_1 \in K^{n\times n}$, $A_2 \in K^{n\times m}$ und $A_3 \in K^{m\times m}$. Zeigen Sie $\det(A) = \det(A_1)\det(A_3)$.

Aufgabe 29. Ist das Schiebepuzzle auf dem Cover lösbar?

Aufgabe 30. Eine Transposition $\sigma \in S_n$ heißt *Basistransposition*, wenn σ zwei aufeinanderfolgende Zahlen vertauscht. Zeigen Sie, dass jede Permutation ein Produkt von Basistransformationen ist.
Bemerkung: Dies ist die Grundlage von *Bubblesort*.

Aufgabe 31. Zeigen Sie, dass die symmetrische Gruppe S_n für $n \geq 2$ genau so viele Permutationen mit Signum 1 wie mit Signum -1 besitzt.
Hinweis: Wenden Sie die Leibniz-Formel auf die Matrix $(1)_{i,j=1}^n \in \mathbb{Q}^{n\times n}$ an.

Aufgabe 32. Seien $\sigma, \tau \in S_n$ mit $\operatorname{sgn}(\sigma) \neq \operatorname{sgn}(\tau)$. Zeigen Sie $\det(P_\sigma + P_\tau) = 0$.
Hinweis: $P_\sigma + P_\tau = P_\sigma(P_\sigma^{-1} + P_\tau^{-1})P_\tau$.

Aufgabe 33.

(a) Sei $\alpha \in K[X]$ vom Grad $d \geq 0$. Zeigen Sie, dass α höchstens d Nullstellen besitzt.

(b) Sei $|K| = \infty$ und $\alpha, \beta \in K[X]$ mit $\alpha(x) = \beta(x)$ für alle $x \in K$. Zeigen Sie $\alpha = \beta$.

Aufgabe 34. 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 35. Für paarweise verschiedene $x_1, \ldots, x_n \in K$ 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*. Zeigen Sie:

(a) $\det(A) = \prod_{1\leq i<j\leq n}(x_j - x_i) \neq 0$.[26]
Hinweis: Gauß-Algorithmus mit Spalten und Induktion nach n.

(b) Für beliebige $y_1, \ldots, y_n \in K$ existiert genau ein Polynom $\alpha \in K[X]$ mit $\deg(\alpha) < n$ und $\alpha(x_i) = y_i$ für $i = 1, \ldots, n$.
Hinweis: Die Koeffizienten von α bilden die Lösung eines Gleichungssystems.

(c) Das LAGRANGE-*Polynom*

$$\alpha := \sum_{i=1}^n y_i \prod_{j\neq i} \frac{X - x_j}{x_i - x_j} \in K[X]$$

erfüllt die Bedingungen aus (b).

(d) Geben Sie einen weiteren Beweis für die Eindeutigkeit von α in (b) mit Hilfe von Aufgabe 33.

Aufgabe 36. Zeigen Sie, dass $A \in \mathbb{F}_2^{n\times n}$ genau dann diagonalisierbar ist, wenn $A^2 = A$ gilt.
Hinweis: Satz 10.41.

Aufgabe 37. Zeigen Sie $\chi_A = \chi_{A^\mathrm{t}}$ und $\mu_A = \mu_{A^\mathrm{t}}$ für $A \in K^{n\times n}$. Wie lassen sich $\chi_{A^{-1}}$ und $\mu_{A^{-1}}$ aus χ_A und μ_A berechnen, falls A invertierbar ist?

Aufgabe 38. Zeigen Sie $\chi_{AB} = \chi_{BA}$ für $A, B \in K^{n\times n}$. Insbesondere haben AB und BA die gleichen Eigenwerte. Konstruieren Sie Matrizen A und B mit $\mu_{AB} \neq \mu_{BA}$.
Hinweis: Berechnen Sie folgende Determinante auf zwei Weisen:

$$\det\left(\begin{pmatrix} X1_n & A \\ B & 1_n \end{pmatrix}\begin{pmatrix} 1_n & 0 \\ -B & X1_n \end{pmatrix}\right).$$

[26] Hierbei ist $\prod$ das Produktzeichen, d. h. $\prod_{i=1}^n a_i = a_1 a_2 \ldots a_n$.

Aufgabe 39. Zeigen Sie $\chi_A = X^3 - \operatorname{tr}(A)X^2 + \frac{1}{2}\big(\operatorname{tr}(A)^2 - \operatorname{tr}(A^2)\big)X - \det(A)$ für alle $A \in K^{3\times 3}$.

Aufgabe 40. Sei

$$A := \begin{pmatrix} 0 & & & -a_0 \\ 1 & \ddots & & \vdots \\ & \ddots & 0 & -a_{n-2} \\ 0 & & 1 & -a_{n-1} \end{pmatrix} \in K^{n\times n}.$$

(a) Berechnen Sie χ_A (man nennt A die *Begleitmatrix* von χ_A).

(b) Folgern Sie, dass jedes normierte Polynom in $K[X]$ als charakteristisches Polynom einer Matrix auftritt.

(c) Zeigen Sie $\chi_A = \mu_A$.
Hinweis: Es genügt zu zeigen, dass $e_1, Ae_1, \ldots, A^{n-1}e_1$ linear unabhängig sind.

Aufgabe 41. Sei V ein euklidischer Raum und $U, W \leq V$. Zeigen Sie:

(a) $(U + W)^\perp = U^\perp \cap W^\perp$.

(b) $(U \cap W)^\perp = U^\perp + W^\perp$.

Aufgabe 42. Sei V ein euklidischer Raum und $f\colon V \to V$ mit $[f(v), f(w)] = [v, w]$ für alle $v, w \in V$. Zeigen Sie, dass f linear ist (und damit orthogonal).

Aufgabe 43. Sei V ein euklidischer Raum mit Orthonormalbasis B und $f \in \operatorname{End}(V)$. Wir definieren die zu f *adjungierte Abbildung* $f^* \in \operatorname{End}(V)$ durch $f^*(b_i) := \sum_{k=1}^n [f(b_k), b_i] b_k$ für $i = 1, \ldots, n$. Zeigen Sie:

(a) $[f(v), w] = [v, f^*(w)]$ für alle $v, w \in V$.

(b) ${}_B[f^*]_B = {}_B[f]_B^{\mathrm{t}}$.

Aufgabe 44. 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 45. Für $\zeta := \cos(\pi/5) + \mathrm{i}\sin(\pi/5) \in \mathbb{C}$ gilt $\zeta^5 = -1$ (siehe Beweis von Lemma 11.26). Sei $\omega := \zeta + \zeta^{-1} = 2\mathrm{Re}(\zeta) \in \mathbb{R}$. 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)

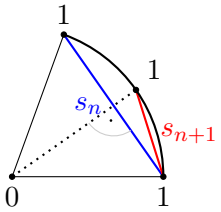
$$D(\pi/5) = \frac{1}{4}\begin{pmatrix} \sqrt{5}+1 & -\sqrt{10-2\sqrt{5}} \\ \sqrt{10-2\sqrt{5}} & \sqrt{5}+1 \end{pmatrix}.$$

Hinweis: $\cos(\varphi)^2 + \sin(\varphi)^2 = 1$.

Aufgabe 46. Wir wollen die Halbkreisbogenlänge π durch den halben Umfang eines regelmäßigen 2^n-Ecks mit „Radius“ 1 approximieren. 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

$$\underbrace{2^n \sqrt{2 - \sqrt{2 + \sqrt{2 + \ldots \sqrt{2}}}}}_{n \text{ Wurzeln}} = \lim_{n\to\infty} 2^n s_{n+1} = \pi.$$

Aufgabe 47. Sei $v \in \mathbb{R}^{n\times 1}$ normiert. Zeigen Sie, dass die Matrix $1_n - 2vv^{\mathrm{t}}$ eine Spiegelung an der Hyperebene $\langle v\rangle^{\perp}$ beschreibt.
Bemerkung: Solche Matrizen nennt man *Householder-Transformationen.*

Aufgabe 48. Sei $\alpha \in \mathbb{R}[X]$. Zeigen Sie, dass Polynome $\gamma_1, \ldots, \gamma_k \in \mathbb{R}[X]$ mit $\alpha = \gamma_1 \ldots \gamma_k$ und $\deg(\gamma_i) \leq 2$ für $i = 1, \ldots, k$ existieren.
Hinweis: Bemerkung 11.36.

Aufgabe 49.

(a) Sei V ein euklidischer Raum mit Basis $b_1, \ldots, b_n$. Sei $A := \big([b_i, b_j]\big)_{ij} \in \mathbb{R}^{n\times n}$. Zeigen Sie, dass A symmetrisch ist und nur positive Eigenwerte besitzt.

(b) Sei umgekehrt $A \in \mathbb{R}^{n\times n}$ symmetrisch mit Eigenwerten $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ und $\chi_A = X^n + a_{n-1}X^{n-1} + \ldots + a_0 \in \mathbb{R}[X]$. Zeigen Sie, dass die folgenden Aussagen äquivalent sind:

(1) Die Abbildung

$$\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \qquad (v, w) \mapsto vAw^{\mathrm{t}}$$

definiert ein Skalarprodukt.

(2) $\lambda_1, \ldots, \lambda_n > 0$.

(3) $(-1)^k a_{n-k} > 0$ für $k = 1, \ldots, n$.

Ggf. nennt man A *positiv definit.*
Hinweis: Benutzen Sie den Spektralsatz und $\chi_A = (X - \lambda_1) \ldots (X - \lambda_n)$.

Aufgabe 50. Seien $n, k \in \mathbb{N}$ und $\lambda \in K$. Zeigen Sie

$$J_n(\lambda)^k = \begin{pmatrix} \lambda^k & & & & 0 \\ k\lambda^{k-1} & \ddots & & & \\ \binom{k}{2}\lambda^{k-2} & \ddots & \ddots & & \\ \vdots & \ddots & \ddots & \ddots & \\ \binom{k}{n-1}\lambda^{k-n+1} & \dots & \binom{k}{2}\lambda^{k-2} & k\lambda^{k-1} & \lambda^k \end{pmatrix},$$

wobei $\binom{k}{l} = \frac{k(k-1)\dots(k-l+1)}{l!}$ für $l = 1, \dots, n-1$.

Aufgabe 51. Sei V ein K-Vektorraum und $f \in \mathrm{End}(V)$. Zeigen Sie, dass die folgenden Aussagen äquivalent sind:

(a) χ_f zerfällt in Linearfaktoren.

(b) μ_f zerfällt in Linearfaktoren.

(c) Es existiert eine Basis B von V, sodass ${}_B[f]_B$ eine untere Dreiecksmatrix ist.

Hinweis: Der Beweis von Satz 12.15 benötigt nur (b).

Aufgabe 52. Zeigen Sie, dass für $A, B \in \mathbb{C}^{3\times 3}$ die folgenden Aussagen äquivalent sind:

(a) Es existiert $S \in \mathrm{GL}(3, \mathbb{C})$ mit $A = SBS^{-1}$.

(b) Es gilt $\chi_A = \chi_B$ und $\mu_A = \mu_B$.

Zeigen Sie, dass die Aussage für 4×4-Matrizen falsch ist.
Hinweis: Jordansche Normalform.

Aufgabe 53. Sei $n \leq 6$ und $A, B \in \mathbb{C}^{n\times n}$. Zeigen Sie, dass die folgenden Aussagen äquivalent sind:

(a) Es existiert $S \in \mathrm{GL}(n, \mathbb{C})$ mit $A = SBS^{-1}$.

(b) Es gilt $\chi_A = \chi_B$, $\mu_A = \mu_B$ und die geometrischen Vielfachheiten der Eigenwerte von A stimmen mit denen von B überein.

Zeigen Sie, dass die Aussage für 7×7-Matrizen falsch ist.

Aufgabe 54. Sei $A \in \mathbb{C}^{n\times n}$. Zeigen Sie, dass eine Matrix $S \in \mathrm{GL}(n, \mathbb{C})$ mit $A^{\mathrm{t}} = SAS^{-1}$ existiert.
Hinweis: Betrachten Sie $PJ_k(\lambda)P^{-1}$ für eine geeignete Permutationsmatrix P.
Bemerkung: Die Aussage gilt über beliebigen Körpern, ist dort aber schwerer zu beweisen.

Stichwortverzeichnis

A
Abbildung, 14
    adjungierte, 95
    affine, 40
    bijektive, 15
    diagonalisierbare, 52
    duale, 50
    injektive, 14
    lineare, 39
    orthogonale, 79
    surjektive, 14
    symmetrische, 79
Absolutglied, 63
Additionstheoreme, 81
Adjunkte, 60
Assoziativgesetz, 16, 17
Aussage, 8
    äquivalente, 8
Austauschsatz, 24
Auswahlaxiom, 10
Axiom, 10

B
Banach-Tarski-Paradoxon, 10
Basis, 22
    duale, 50
Basisergänzungssatz, 24
Basistransposition, 93
Basiswechsel, 45
Basiswechselmatrix, 42
Begleitmatrix, 95
Betrag, 74
    komplexer Zahl, 82
Bijektion, 15
Bild, 14
Binet-Formel, 71
Bubblesort, 93

C
Cantor, 10
Cardanische Formeln, 5
Cauchy-Schwarz-Ungleichung, 75
Cayley-Hamilton, 72
Cholesky-Zerlegung, 4
Cramersche Regel, 60

D
Darstellungsmatrix, 42
De Morgansche Regeln, 9, 12
Dedekind-Identität, 47
Definitionsbereich, 14
Determinante
    einer Abbildung, 59
    einer Matrix, 55
Determinantensatz, 58
Dezimalbruch, 11
Diagonalisierungsargument, 16
Diagonalmatrix, 27
Differenz, 11
Dimension, 25
Dimensionsformel, 26
Direktes Produkt, 18
Distributivgesetz, 9, 12, 18
Division mit Rest, 65
Drehspiegelung, 86
Drehung, 80
Dreiecksmatrix, 54
Dreiecksungleichung, 75
Dualraum, 50
Durchschnitt, 11

E
Eigenraum, 51
Eigenvektor, 51
Eigenwert, 51
Einheitsmatrix, 27
Einheitswurzeln, 83
Einschränkung, 15
Elementarmatrix, 33
Endomorphismus, 51
Erzeugendensystem, 22
euklidischer Raum, 74
Euler, 85

F
Fermats letzter Satz, 5
Fibonacci-Zahlen, 70
Fitting, 87
Fließkommazahl, 11
Froebnius-Ungleichung, 49
Fundamentalsatz der Algebra, 83
Funktion, *siehe* Abbildung
Funktionalanalysis, 26

G
Gauß-Algorithmus, 34

Gleichungssystem, 31
homogen, 32
inhomogen, 32
lösbar, 31
unterbestimmt, 32
überbestimmt, 33
goldener Schnitt, 70
Gram-Schmidt-Verfahren, 77
Gruppe, 17
abelsche, 17
allgemeine lineare, 30
orthogonale, 79
Gödels Unvollständigkeitssätze, 10

H
Hauptachsentransformation, 5
Hauptdiagonale, 27
Hauptraum, 87
Hauptraumzerlegung, 88
Hilberts Nullstellensatz, 5
Hintereinanderausführung, 15
Homogenität, 75
Homomorphismus, 39
Horner-Schema, 66
Householder-Transformation, 96
Hyperwürfel, 54

I
Identität, 15
Imaginärteil, 82
Inklusionsabbildung, 15
Invariante, 54
inverses Element, 17
Isomorphismus, 39

J
Jordanblock, 89
Jordansche Normalform, 91

K
Kardinalzahl, 16
kartesisches Produkt, 14
Kegelschnitt, 5
Kern, 40
Koeffizient, 63
führender, 64
Koeffizientenmatrix, 31
erweiterte, 32
Kommutativgesetz, 17
Komplement, 25
orthogonales, 78
komplexe Konjugation, 83
Komposition, 15
Kontinuumshypothese, 10, 16
Kontraposition, 9
Koordinatendarstellung, 23
Kosinus, 76
Kosinussatz, 95
Kreuzprodukt, 79
Kronecker-Capelli, 32
Kronecker-Delta, 27
Körper, 18

L
Lagrange-Polynom, 94
Laplace-Entwicklung, 59
Leibniz-Formel, 62
Leitkoeffizient, 64
linear (un)abhängig, 22
Linearfaktor, 67
Linearkombination, 20
Lösungsmenge, 31

M
Matrix, 27
diagonalisierbare, 52
dünnbesetze, 60
invertierbare, 30
komplementäre, 60
orthogonale, 80
positiv definite, 96
quadratische, 27
reguläre, 30
singuläre, 30
symmetrische, 80
transponierte, 28
zeilen-äquivalente, 34
Menge
endliche, 10
gleichmächtige, 15
unendliche, 10
überabzählbare, 16
Methode der kleinsten Quadrate, 4
Millenniumsproblem, 9
Minimalpolynom, 71
Modus ponens, 9

N
neutrales Element, 17
Newton-Verfahren, 84
Norm, 74
Nullmatrix, 27
Nullpolynom, 64
Nullraum, 19
Nullstelle, 66
doppelte, 68
Nullvektor, 19

O
Orthonormalbasis, 77

P
Paar, 14
Parallelogrammgleichung, 75

Permutation, 61
Permutationsmatrix, 61
Polynom, 63
    konstantes, 64
    normiertes, 64
Potenzmenge, 47
Projektion, 40
Prädikat, 8
Pythagoras, 75
    trigonometrischer, 95

Q
QR-Verfahren, 4
Quadrik, 5

R
Rang
    einer Abbildung, 40
    einer Matrix, 31
Rangsatz, 45
Realteil, 82
Rechte-Hand-Regel, 79
Rest, 66
Ring, 29
Ruffinis Regel, 66
Russellsche Antinomie, 10

S
Sarrus-Regel, 62
SAT-Problem, 9
Satz vom ausgeschlossenen Dritten, 9
Satz vom Widerspruch, 9
Selectionsort, 61
Signum, 61
Simplex-Verfahren, 4
Sinus, 76
Sinussatz, 95
Skalar, 19
Skalarmatrix, 27
Skalarprodukt, 74
Spaltenoperation, 34
Spaltenvektor, 27
Spann, 21
Spektralsatz, 84
Spektrum, 84
Spiegelung, 81
Spur
    einer Abbildung, 46
    einer Matrix, 46
Standardbasis, 23
Standardmatrix, 28
Standardskalarprodukt, 74
Steinitz, 24
Strassen-Algorithmus, 48
Summe von Unterräumen, 21, 52
    direkte, 21
Sylvester-Ungleichung, 49
symmetrische Gruppe, 61

T
Teiler, 66
Teilmenge, 11
    echte, 11
Transitivität, 9
Transposition, 61
Tripel, 14
Tupel, 14

U
Umkehrfunktion, 16
Umkehrung, 9
Unterraum, 20
    f-invarianter, 86
Urbild, 14

V
Vandermonde-Matrix, 94
Variable, 63
Vektor, 19
    normierter, 74
Vektorraum, 19
    endlich erzeugter, 22
    endlich-dimensionaler, 25
    euklidischer, 74
    isomorph, 39
Venn-Diagramm, 11
Vereinigung, 10
Verkettung, 15
Vielfachheit
    algebraische, 67
    geometrische, 51
Vollständige Induktion, 13
Vorzeichen, 61

W
Wertebereich, 14

Z
Zahlen
    ganze, 11
    komplexe, 82
    natürliche, 11
    rationale, 11
    reelle, 11
Zassenhaus-Algorithmus, 38
Zeilenoperation, 33
Zeilenstufenform, 34
Zeilenvektor, 27
Zermelo-Fraenkel-System, 10
Zirkelschluss, 23
Zorns Lemma, 25