Steffen Kionke

Lineare Algebra

Fakultät für
Mathematik und Informatik

FernUniversität in Hagen

Inhaltsverzeichnis

Einleitung und Studierhinweise **iv**
Einleitung . . . v
Allgemeine Studierhinweise . . . x
Notation . . . xiii
Auszug aus dem griechischen Alphabet . . . xiv

1. Grundbegriffe der Algebra **1**
Studierhinweise . . . 3
1.1. Gruppen . . . 7
1.2. Ringe und Körper . . . 14
I. Definition und Beispiele, 14. II. Einheiten, 18. III. Matrizen über Ringen, 19.
1.3. Restklassenringe und endliche Körper . . . 26
1.4. Polynomringe . . . 34
I. Rechnen im Polynomring, 34. II. Division im Polynomring, 38. III. Nullstellen, 40. IV. Ideale, 49.
1.5. Der Körper der komplexen Zahlen . . . 54
I. Konstruktion der komplexen Zahlen, 54. II. Rechnen mit komplexen Zahlen, 59. III. Der Fundamentalsatz der Algebra, 61.
1.6. Lösungen der Aufgaben in Lektion 1 . . . 67

2. Konzepte der Linearen Algebra **83**
Studierhinweise . . . 85
2.1. Vorwort zur Mengenlehre: das Auswahlaxiom* . . . 89
2.2. Rückblick und Ergänzungen: Vektorräume und Basen . . . 95
I. Vektorräume, 95. II. Basen, 97. III. Lineare Abbildungen, 100.
2.3. Direkte Summe von Vektorräumen . . . 105
I. Definition und Eigenschaften, 105. II. Komplemente, 110.
2.4. Faktorräume . . . 114
I. Konstruktion des Faktorraumes, 115. II. Die Dimension des Faktorraumes, 122.
2.5. Der Dualraum . . . 126
I. Dualraum und Dualbasen, 126. II. Duale Abbildungen, 131.
2.6. Hilfsmittel: Äquivalenzrelationen . . . 134
I. Definition und Beispiele, 134. II. Beispiel: Äquivalente Matrizen, 138.

2.7. Lösungen der Aufgaben in Lektion 2 141

3. Bilinearformen und Hermite'sche Formen **151**
Studierhinweise 153
3.1. Bilinearformen 157
I. Definition und Beispiele, 157. II. Matrixdarstellung, 162. III. Kongruenz von Bilinearformen, 165. IV. Einsetzen, 169. V. Reguläre und ausgeartete Formen, 173.
3.2. Symmetrische Bilinearformen 178
I. Definition und Beispiele, 178. II. Diagonalisierungssatz, 180. III. Reelle symmetrische Bilinearformen, 187.
3.3. Alternierende Bilinearformen 195
3.4. Hermite'sche Formen 200
I. Definition und Beispiele, 200. II. Matrixdarstellung, 202. III. Komplex wird reell, 209. IV. Klassifikation, 212.
3.5. Lösungen der Aufgaben in Lektion 3 219

4. Determinanten **231**
Studierhinweise 233
4.1. Die symmetrischen Gruppen 237
I. Definition, 237. II. Rechnen mit Permutationen, 238. III. Die Signaturabbildung, 242.
4.2. Die Determinante von Matrizen 246
I. Definition, 246. II. Charakterisierungssatz, 252. III. Multiplikationssatz, 257. IV. Matrizen über Körpern, 260.
4.3. Minoren und die Adjunkte 264
I. Minoren, 264. II. Entwicklungssatz, 266. III. Adjunktensatz, 271.
4.4. Determinanten von Endomorphismen 278
4.5. Alternierende Multilinearformen* 280
I. Multilineare Abbildungen, 280. II. Alternierende Multilinearformen, 284. III. Determinante ohne Matrizen, 288.
4.6. Lösungen der Aufgaben in Lektion 4 291

5. Eigenwerte und Eigenvektoren **301**
Studierhinweise 303
5.1. Das Normalformproblem 307
I. Was ist das Normalformproblem?, 307. II. Invariante Unterräume, 309.
5.2. Eigenwerte, Eigenvektoren und Diagonalisierbarkeit 313
I. Eigenwerte und Eigenvektoren, 313. II. Diagonalisierbarkeit, 317.
5.3. Charakteristisches Polynom und Minimalpolynom 323
I. Charakteristisches Polynom, 323. II. Minimalpolynom, 330. III. Der Satz von Cayley-Hamilton, 337.
5.4. Stochastische Matrizen und Anwendungen 342

I. Stochastische Matrizen, 342. II. Google PageRank, 347.
5.5. Lösungen der Aufgaben in Lektion 5 353

6. Die Jordan'sche Normalform 365
Studierhinweise 367
6.1. Invariante Unterräume 371
I. Zerlegung durch das Minimalpolynom, 371. II. Diagonalisierbarkeit, 375.
6.2. Nilpotente Endomorphismen 377
I. Definition und Charakterisierung, 377. II. Zyklische Unterräume, 384. III. Einschub: Partitionen, 388. IV. Normalform, 391.
6.3. Die Jordan'sche Normalform 404
I. Was ist die Jordan'sche Normalform?, 404. II. Beweis des Hauptsatzes, 408. III. Berechnung der Jordan'schen Normalform, 411. IV. Anwendungen, 416.
6.4. Lösungen der Aufgaben in Lektion 6 423

7. Euklidische und unitäre Vektorräume 433
Studierhinweise 435
7.1. Skalarprodukt und Norm 439
I. Definition und Beispiele, 439. II. Die Norm, 443.
7.2. Orthogonalität 448
I. Orthogonale Vektoren, 448. II. Orthonormalbasen, 451. III. Das Verfahren von Gram-Schmidt, 455.
7.3. Adjungierte Abbildungen 459
I. Definition und Existenzsatz, 459. II. Matrixdarstellung, 462.
7.4. Orthogonale und unitäre Abbildungen 466
I. Isometrien, 466. II. Orthogonale und unitäre Matrizen, 470. III. Matrixdarstellung von Isometrien, 476.
7.5. Spektralsätze 478
I. Normale Endomorphismen, 479. II. Unitäre Endomorphismen, 482. III. Selbstadjungierte Endomorphismen, 485.
7.6. Lösungen der Aufgaben in Lektion 7 495

Literaturverzeichnis 509

Symbolverzeichnis 511

Index 513

Einleitung und Studierhinweise

Einleitung

Ich freue mich sehr, dass Sie sich für das Modul „Lineare Algebra“ entschieden haben. Der anstrengende Weg durch dieses Modul lohnt sich sowohl für Studierende der Mathematik als auch der Informatik, denn die Lineare Algebra ist eine der nützlichsten und elegantesten Theorien, die die Mathematik hervorgebracht hat. Vereinfacht gesagt, ist die Lineare Algebra die Theorie der Matrizen, der Vektorräume und der Abbildungen zwischen Vektorräumen; diese Begriffe kennen Sie ja bereits aus den „Mathematischen Grundlagen“. Um genauer zu verstehen, was Lineare Algebra ist und wie man sich ihr nähern kann, müssen wir zuerst einen kurzen (unvollständigen) Blick auf die Geschichte werfen.

Zur Geschichte der Linearen Algebra

Methoden zum Lösen von Gleichungen beschäftigen Menschen schon seit sehr langer Zeit. Lange bevor überhaupt von „Gleichungen“ gesprochen wurde, hatten Menschen Probleme (etwa beim Handel oder bei der Vermessung von Feldern), wobei eine unbekannte Größe aus bekannten Angaben bestimmt werden muss. Verfahren zur Lösung bestimmter Fragestellungen entwickelten schon die Babylonier (1700 v. Chr.). Die systematische Untersuchung solcher Probleme ist der Kern der frühen *Algebra*. Das Wort „Algebra“ geht dabei auf den Titel eines Buches des Universalgelehrten al-Chwarizmi[1] zurück, in dem er systematisch lineare und quadratische Gleichungen untersuchte.

Von besonderem Interesse waren dabei *lineare Gleichungen* und *lineare Gleichungssysteme*. Schon um 200 v. Chr. sind in China Methoden zur Lösung von linearen Systemen mit 5 Gleichungen und 5 Unbekannten dokumentiert. Erst deutlich später taucht dann die Idee der *Determinante* eines Gleichungssystems auf (Seki[2] um

[1] Abu Dscha’far Muhammad ibn Musa AL-CHWARIZMI: choresmischer Universalgelehrter, um 800.
[2] Takakazu SEKI: japanischer Mathematiker, 1642–1708.

1683, Leibniz[3] um 1693). Die Determinante ist zunächst eine Kennzahl, die man berechnen kann, um festzustellen, ob ein Gleichungssystem eine Lösung besitzt. Cramer[4] gelingt es um 1750, mithilfe von Determinanten eine Lösungsformel für lineare Gleichungssysteme anzugeben. Auch die *analytische Geometrie* hat zu dieser Zeit einen großen Einfluss auf die Weiterentwicklung der Linearen Algebra. Insbesondere das Studium der *Kegelschnitte* und der geeigneten Koordinatenwechsel führt um 1800 zur sogenannten „Hauptachsentransformation". Das Bemerkenswerte an diesen Entwicklungen ist, dass Matrizen und Vektorräume (aus heutiger Sicht die zentralen Begriffe) damals noch nicht bekannt waren.

Der Begriff *Matrix* stammt aus dem Jahr 1850 und geht auf Sylvester[5] zurück. Matrizen als algebraische Objekte wurden wahrscheinlich zum ersten Mal systematisch von Cayley[6] untersucht. Man konnte nun mit Matrizen rechnen. Die altbekannten Rechenverfahren bekamen dadurch einen Rahmen, der es ermöglichte, Zusammenhänge klar zu benennen.

Im Jahr 1843 nimmt eine andere entscheidende Entwicklung ihren Lauf. Hamilton[7] konstruiert den Schiefkörper der *Quaternionen*. Dabei handelt es sich um ein vierdimensionales Objekt, das alle Körperaxiome erfüllt, außer dem Kommutativgesetz der Multiplikation. Im Zuge dessen prägt Hamilton den Begriff *Vektor*, beschreibt das *Skalarprodukt* und das *Kreuzprodukt* von Vektoren. Diese Art der Vektorrechnung wird in der Physik direkt aufgegriffen und weiterentwickelt.

Der Begriff des *Vektorraumes* entsteht 1844 in Arbeiten von Graßmann[8]. Er kennt bereits *linear unabhängige* Vektoren und auch den Begriff der *Dimension*. Allerdings sind seine Arbeiten schwer zu lesen und bleiben zunächst ohne große Wirkung. Peano[9] greift die Idee auf und gibt 1888 eine axiomatische Beschreibung von Vektorräumen an. Aber erst als sich die Funktionalanalysis aus Fragestellungen der Physik entwickelt – ab 1903 durch Arbeiten von Fredholm[10], Hilbert[11] und anderen – blüht auch die Theorie der Vektorräume auf.

Um 1920 entsteht dann die *moderne Algebra* – die Theorie der algebraischen Strukturen (Gruppen, Ringe, Körper), deren Eigenschaften man systematisch aus

[3] Gottfried Wilhelm LEIBNIZ: deutscher Universalgelehrter, 1646 – 1716.
[4] Gabriel CRAMER: schweizer Mathematiker, 1704–1752.
[5] James Joseph SYLVESTER: britischer Mathematiker, 1814–1897.
[6] Arthur CAYLEY: englischer Mathematiker, 1821–1895.
[7] William Rowan HAMILTON: irischer Mathematiker, 1805–1865.
[8] Hermann Günther GRASSMANN: deutscher Mathematiker und Sprachwissenschaftler, 1809–1877.
[9] Giuseppe PEANO: italienischer Mathematiker, 1858–1932.
[10] Erik Ivar FREDHOLM: schwedischer Mathematiker, 1866–1927.
[11] David HILBERT: deutscher Mathematiker, 1862–1943

Axiomen („Rechenregeln“) ableitet. Eine prägende Figur in dieser Entwicklung ist Emmy Noether[12]. Im Lehrbuch „Modern Algebra“ von van der Waerden[13] aus dem Jahr 1930 findet man zum ersten Mal die heute gebräuchliche Definition eines Vektorraumes. Auch das französische Autorenkollektiv „Nicolas Bourbaki“ (gegründet 1934) hatte durch seine präzisen Fachbücher großen Einfluss auf die Weiterentwicklung der modernen (Linearen) Algebra und den Blick, den wir heute auf dieses Fach haben. Die Forschung an Linearer Algebra war damit natürlich nicht abgeschlossen. Es gibt auch heute noch offene Fragen!

Die Lineare Algebra, wie sie heute an Universitäten gelehrt wird, ist also das Produkt einer sehr langen Entwicklung. Die Menschheit hat lange gebraucht, bis sie verstanden hat, was ein Vektorraum ist. Es muss Sie also nicht verwundern, wenn Ihnen manche Begriffe nicht sofort einleuchten. Die Beweise und Konzepte wurden seit 1930 durch Generationen von Lehrbüchern geschliffen und verfeinert.

Die zwei Stränge der Linearen Algebra

Der Blick in die Geschichte offenbart uns einen sehr wichtigen Aspekt, den man beim Studium der Linearen Algebra nicht vergessen sollte: *Die Lineare Algebra besteht aus zwei Strängen.*

Der erste Strang besteht aus Rechenmethoden; heute würde man dazu „Matrizenrechnung“ sagen.

- → Wie löse ich ein Gleichungssystem?
- → Wie berechnet man eine Determinante?
- → Wie berechnet man einen Koordinatenwechsel?

Dieser Teil der Linearen Algebra ist sehr alt und wurde meist anhand konkreter Fragestellungen entwickelt. Es geht hier um das Rechnen in Koordinaten. Erfahrungsgemäß fällt Studierenden dieser Aspekt der Linearen Algebra auch etwas leichter, denn man kann diese Rechenverfahren lernen und einüben.

Den zweiten Strang könnte man „abstrakte“ Lineare Algebra nennen. Hier geht es um Objekte, Strukturen und Zusammenhänge. Es geht also darum, etwas zu verstehen.

- → Was ist ein Vektorraum?

[12] Amalie Emmy Noether: deutsche Mathematikerin, 1882–1935.
[13] Bartel Leendert van der Waerden: niederländischer Mathematiker, 1904–1996.

→ Welche Vektorräume gibt es?

→ Was ist eine Bilinearform?

→ Kann man die Bilinearformen klassifizieren?

Hier wird der Einfluss der modernen Algebra sichtbar. Dieser Teil der Linearen Algebra ist gewissermaßen *koordinatenfrei*. Man möchte also abstrakte Ideen und Konzepte verstehen und zueinander in Beziehung setzen. Der „abstrakte" Strang der Linearen Algebra ist schwieriger zu lernen. Nur indem man sich viele Beispiele anschaut, kann man sichergehen, die Begriffe und Ideen wirklich verstanden zu haben.

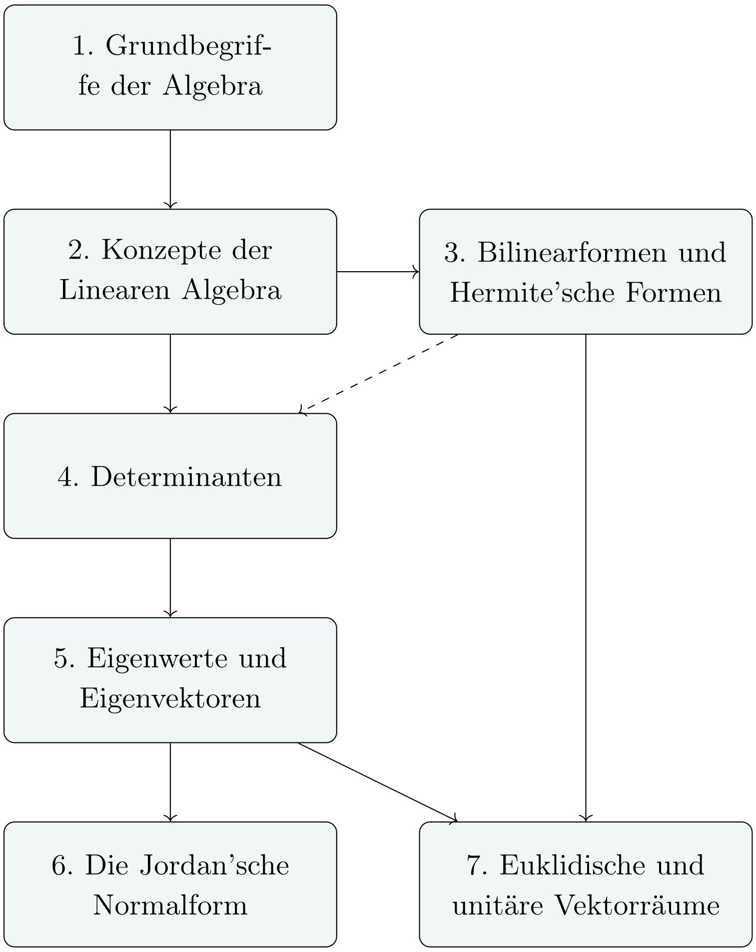
Beide Stränge der Linearen Algebra interagieren miteinander. Einerseits möchten wir Rechnungen interpretieren: Was haben wir eigentlich berechnet? Was bedeutet das? Andererseits möchten wir ausgehend von abstrakten Begriffen in Beispielen konkrete Rechnungen anstellen. Zum Beispiel: Wie kann man eine lineare Abbildung als Matrix darstellen? Kann man die Koordinaten geschickt wählen, sodass man möglichst einfach rechnen kann? Es ist wichtig, dass Sie beide Seiten im Blick behalten. Wenn Sie nur die Rechenverfahren kennen, verstehen Sie nicht, was Sie eigentlich berechnen. Wenn Sie nur die Theorie kennen, werden Sie bei konkreten Beispielen nicht weiterkommen.

Leitfaden

Am Anfang jeder Lektion finden Sie Studierhinweise mit Kommentaren zum Inhalt. Deshalb möchte ich hier nur sehr kurz etwas zum Inhalt der Lektionen sagen.

In der ersten Lektion lernen wir einige Grundbegriffe der Algebra kennen: Gruppen, Ringe, endliche Körper, Polynome und die komplexen Zahlen. Auf diese Begriffe werden wir später immer wieder zurückgreifen. Die zweite Lektion beginnt mit einem Exkurs in die Mengenlehre und einem kurzen Rückblick auf Begriffe, die Sie bereits aus den „Mathematischen Grundlagen" kennen. Dann entwickeln wir wesentliche Konzepte der „abstrakten" Linearen Algebra: direkte Summen, Faktorräume und Dualräume. Diese Begriffe sind entscheidend, um dem Text weiter folgen zu können.

Im dritten Kapitel besprechen wir Bilinearformen und Hermite'sche Formen. Hier werden die zwei Stränge der Linearen Algebra sichtbar. Wir entwickeln die „abstrakte" Theorie der Bilinearformen und lernen parallel, wie wir mithilfe von Matrixdarstellungen konkrete Beispiele bearbeiten können.



Die vierte Lektion behandelt die Determinante. Wie wir beim Blick in die Geschichte gesehen haben, handelt es sich um einen sehr alten Teil der Linearen Algebra. Wir werden zuerst die Determinante definieren und konkrete Beispiele berechnen. Im Anschluss werden wir einige Zeit benötigen, um zu verstehen, was die Determinante aus Sicht der abstrakten Linearen Algebra ist.

Die Theorie der Eigenwerte ist eines der Highlights der Linearen Algebra, weil sie unglaublich viele Anwendungen hat. Wir entwickeln diese Theorie in Lektion 5 und lernen mit dem PageRank-Verfahren von Google eine Anwendung kennen.

Die schwierigste Lektion ist sicherlich die sechste. Hier lernen wir die *Jordan'sche Normalform* kennen und wir lernen, wie man sie berechnet. In diesem Kapitel ist es wirklich entscheidend, dass Sie versuchen, die abstrakte Seite und die konkreten Rechnungen zueinander in Beziehung zu setzen.

Im siebten Kapitel befassen wir uns dann mit Skalarprodukten. Wir werden sehen, dass wir in Vektorräumen mit einem Skalarprodukt Geometrie betreiben können. Hier kratzen wir nur an der Oberfläche, denn die Geometrie kommt in diesem Lehrtext insgesamt etwas zu kurz. Die Spektralsätze bilden einen schönen Schlusspunkt für diese Lehrveranstaltung. Sie werden es uns ermöglichen, viele Ergebnisse aus den vorangegangenen Lektionen besser zu verstehen.

Allgemeine Studierhinweise

Mathematik zu lernen ist anstrengend. Es wird *nicht* ausreichen, wenn Sie diesen Lehrtext einmal wie einen Roman durchlesen. Arbeiten Sie sich langsam, Schritt für Schritt voran. Machen Sie sich eigene Notizen. Die folgenden vier Tipps liegen mir besonders am Herzen:

→ *Bearbeiten Sie Übungsaufgaben!* Das selbstständige Rumprobieren und das Aufschreiben von Argumenten sind der einzige Weg, um mit den Konzepten wirklich vertraut zu werden. Schlagen Sie die Lösung nicht voreilig nach.

→ *Lesen Sie Definitionen genau* und lernen Sie diese auswendig. Suchen Sie Beispiele und Gegenbeispiele.

→ *Nehmen Sie sich die Zeit, um Sätze zu verstehen!* Sehen Sie sich konkrete Beispiele an. Fragen Sie sich: Was sind die Voraussetzungen und was passiert, wenn man diese weglässt?

→ *Stellen Sie Fragen!* Man muss nicht jede Nuss alleine knacken. Ob Sie die Betreuenden oder Ihre Kommilitonen fragen, ist dabei nebensächlich. Wichtig ist, sich überhaupt auszutauschen. Oft lernt man schon durch das Formulieren einer Frage etwas dazu.

Es gibt natürlich kein Patentrezept. Jede und jeder lernt anders. Daher ist es wichtig, dass Sie Ihren Wissensstand ehrlich einschätzen. Wenn Sie große Probleme haben, die Aufgaben im Lehrtext und die Einsendeaufgaben zu bearbeiten, dann müssen Sie etwas an Ihrer Arbeitsweise ändern. Erkundigen Sie sich beispielsweise bei Ihren Kommilitonen, wie sie sich dem Stoff nähern.

In der Moodle-Lernumgebung zum Modul finden Sie übrigens ein Diskussionsforum und Einsendeaufgaben zum Modul. Nutzen Sie auch die Lehrveranstaltungen zum Modul, um Fragen zu stellen und mit anderen Studierenden in Kontakt zu kommen.

Eingangsvoraussetzungen

Der Lehrtext baut auf den ersten drei Lektionen der „Mathematischen Grundlagen“ auf. Mit diesem Teil der Grundlagen müssen Sie gut vertraut sein. In jedem Fall sollten Sie mit Vektorräumen und Linearen Abbildungen umgehen können. Zum Lösen von Aufgaben müssen Sie sicher mit Matrizen rechnen können und in der Lage sein, lineare Gleichungssysteme fehlerfrei zu lösen. Das heißt aber nicht, dass Sie die Inhalte der Grundlagen auswendig kennen müssen. Wenn wir auf Ergebnisse aus den Mathematischen Grundlagen zurückgreifen, verweisen wir auf die Quelle [MG].

Literatur

Der vor Ihnen liegende Lehrtext sollte Ihre Hauptquelle beim Studium der Linearen Algebra sein. Trotzdem kann es gelegentlich helfen, einen Begriff oder Satz in einem anderen Buch nachzuschlagen. Dort findet man vielleicht eine andere Erklärung oder ein hilfreiches Beispiel. Leider deckt kein Lehrbuch den Stoff dieses Lehrtextes vollständig ab.

Exemplarisch habe ich sechs Lehrbücher mit verschiedenen Ansätzen herausgesucht. Diese können Sie über die Bibliothek der FernUniversität als e-Book herunterladen.

→ [Bär]: Christian Bär – *Lineare Algebra und analytische Geometrie.*
Schönes, übersichtliches Lehrbuch. Deckt nicht alles ab, enthält dafür etwas mehr Geometrie als dieser Lehrtext.

→ [Beu]: Albrecht Beutelspacher – *Lineare Algebra.*
Lockerer Umgangston, gut verständlich, deckt aber nur wenige Inhalte des Moduls ab.

→ [Bo]: Siegfried Bosch – *Lineare Algebra.*
Ein Lehrbuch, das kompakt viele wichtige Inhalte behandelt. Oft weicht die Perspektive von diesem Lehrtext ab.

→ [Fi]: Gerd Fischer, Boris Springborn – *Lineare Algebra.*
Ein Standardlehrbuch, das kurz und knapp die wichtigsten Inhalte behandelt.

→ [Gö]: Laurenz Göllmann – *Lineare Algebra.*
Umfangreiches Lehrbuch mit einem etwas anderen, rechenlastigen Ansatz.

→ [KaSt]: Christian Karpfinger, Hellmuth Stachel – *Lineare Algebra.*
Umfangreiches Lehrbuch mit vielen schönen Abbildungen und Hinweisen.

Struktur des Lehrtextes

Der Studienbrief besteht aus sieben Lektionen. Jede Lektion beginnt mit Studierhinweisen, bestehend aus Kommentaren zum Inhalt, Literaturangaben und einem Fahrplan durch die Lektion. Die Unterabschnitte im Lehrtext sind vollständig nummeriert (1.1.1, 1.1.2, ...), um das Suchen im Text und das Fragenstellen zu erleichtern. Die jeweils letzte Unterabschnittsnummer auf einer Seite findet man immer oben rechts in der Kopfzeile, sodass Sie beim Blättern leicht zu den entsprechenden Stellen gelangen.

Der Lehrtext enthält viele Übungsaufgaben, deren Lösungen Sie jeweils am Ende jedes Kapitels finden. Die Übungsaufgaben sind wichtig, um den Stoff zu begreifen. Schlagen Sie die Lösungen nicht voreilig nach. Ein blaues „L“ am rechten Rand dient in der PDF-Version jeweils als Link zur Lösung der Aufgabe. Von der Lösung **L** gelangt man wieder zurück zum entsprechenden Abschnitt im Lehrtext, indem man auf die Nummer klickt.

Damit Sie mathematische Ergebnisse schneller erkennen, sind **Satz**, **Lemma**, **Korollar**, und so weiter farbig hinterlegt. Jede **Definition** wird mit einer Umrandung hervorgehoben. Das Ende eines Beweises ist mit dem Symbol □ markiert.

Sätze und Definitionen, die ich für besonders wichtig halte, sind am äußeren Rand durch einen Balken hervorgehoben.

Das „gefährliche Kurve“-Symbol von Bourbaki wird verwendet, um auf mögliche Missverständnisse hinzuweisen.

Am Ende des Lehrtextes finden Sie einen Index, ein Literaturverzeichnis und ein Symbolverzeichnis. Abbildungen sind in der gedruckten Version dieses Textes schwarz-weiß. In der PDF-Version finden Sie die Abbildungen in Farbe.

Danksagung

Mein herzlicher Dank gilt allen Studierenden, die durch das Aufspüren vieler Fehler und Ungenauigkeiten den vorliegenden Lehrtext deutlich verbessert haben. Insbesondere möchte ich Herrn Sven Schmidt und Herrn Ralph Beckmann für die zahlreichen hilfreichen Rückmeldungen danken.

Notation

Die folgenden Symbole und Bezeichnungen werden im Lehrtext verwendet. Im Studienbrief neu eingeführte Symbole und Schreibweisen findet man im Symbolverzeichnis.

$\mathbb{N}$	Die natürlichen Zahlen ohne Null $\{1, 2, 3, 4, \dots\}$
$\mathbb{N}_0$	Die natürlichen Zahlen mit Null $\{0, 1, 2, 3, 4 \dots\}$
$\mathbb{Q}$	Die rationalen Zahlen.
$\mathbb{R}$	Die reellen Zahlen
$\mathbb{Z}$	Die ganzen Zahlen $\{\dots, -2, -1, 0, 1, 2, \dots\}$
$\mathrm{Abb}(X, Y)$	Die Menge aller Abbildungen $X \to Y$
$f\|_U$	Einschränkung der Abbildung f auf die Teilmenge U
$\subseteq$	Teilmenge
$\bigcup$	Vereinigung
$\bigsqcup$	disjunkte Vereinigung
$\langle S \rangle$	lineare Hülle von $S \subseteq V$ im Vektorraum V

Auszug aus dem griechischen Alphabet

Steffen Kionke

Lineare Algebra

Lektion 1:
Grundbegriffe der Algebra

Fakultät für
Mathematik und Informatik

FernUniversität in Hagen

Studierhinweise zur ersten Lektion

In der ersten Lektion befassen wir uns mit einigen Grundbegriffen der Algebra. Die Algebra ist ein Teilgebiet der Mathematik, das sich ursprünglich mit Methoden zum Lösen von Gleichungen befasst hat. Das Gebiet hat sich vor allem im 19. und 20. Jahrhundert stark weiterentwickelt. Heute versteht man unter dem Begriff „Algebra“ die Theorie der „algebraischen Strukturen“. Vereinfacht gesagt ist eine algebraische Struktur eine Menge zusammen mit Rechenoperationen, die vorgegebenen Axiomen („Rechenregeln“) genügen. Die Algebra bemüht sich, diese Strukturen zu verstehen.

Wir werden in dieser Lektion drei Beispiele algebraischer Strukturen kennenlernen: Gruppen, Ringe und Körper. Diese algebraischen Begriffe ermöglichen es später, Zusammenhänge klar zu benennen. Die sehr umfangreiche Theorie, die man mit diesen Begriffen entwickeln kann, ist für uns nicht wichtig. Wer mehr über Algebra lernen möchte, kann im Anschluss das fortgeschrittene Modul „Algebra“ belegen.

Eine algebraische Struktur kennen Sie bereits aus den „Mathematischen Grundlagen“: die Körper. Zumindest an den Körper der rationalen Zahlen $\mathbb{Q}$ und den Körper der reellen Zahlen $\mathbb{R}$ können Sie sich bestimmt noch erinnern. In dieser Lektion werden wir weitere Beispiele von Körpern kennenlernen: die endlichen Körper $\mathbb{F}_p$ und den Körper der komplexen Zahlen $\mathbb{C}$. Diese Körper werden wir auch später in vielen Beispielen benötigen. Es ist deshalb empfehlenswert, das Rechnen in diesen Körpern zu üben.

Besonders genau werden wir uns mit dem Polynomring $K[X]$ über einem Körper K befassen. Polynome sind ein unerlässliches Hilfsmittel in der Linearen Algebra. In gewissem Sinne versteckt sich hinter der „Jordan’schen Normalform“, die wir später in Lektion 6 herleiten, die Theorie der „Moduln“ über Polynomringen.

Lernziele

Nach dem Durcharbeiten der Lektion sollten Sie

- → die algebraischen Strukturen *Gruppe*, *Ring* und *Körper* kennen und Beispiele benennen können,
- → mit den Rechenregeln für Matrizen über Ringen vertraut sein,
- → die endlichen Körper $\mathbb{F}_p$ kennen und in der Lage sein, darin zu rechnen,
- → mit dem Polynomring $K[X]$ über einem Körper umgehen können,

→ den Zusammenhang zwischen Linearfaktoren und Nullstellen von Polynomen kennen,

→ wissen, was ein Ideal im Polynomring ist und dass alle Ideale von einem Element erzeugt werden,

→ den Körper der komplexen Zahlen kennen und sicher mit komplexen Zahlen rechnen können,

→ den Fundamentalsatz der Algebra kennen und anwenden können.

Literaturhinweise

Falls Sie die Inhalte aus dieser Lektion in anderen Quellen nachschlagen möchten, dann finden Sie wesentliche Aspekte in den meisten Lehrbüchern zur Linearen Algebra. Allerdings ist keines der Lehrbücher deckungsgleich mit diesem Lehrtext. Nachschlagen kann man beispielsweise in

→ [Bär]: 2.5 ($\mathbb{C}$), 3.1 (Gruppen), 3.2 (Ringe), 6.1 (Polynome).

→ [Beu]: Kapitel 6 (Ringe und Polynome), Kapitel 9 (Gruppen).

→ [Bo]: 1.2 (Gruppen), Kapitel 5 (Ringe, Polynome).

→ [Fi]: 2.2 (Gruppen), 2.3 (Ringe, Polynome).

→ [Gö]: Kapitel 1.

→ [KaSt]: Kapitel 2.

Natürlich finden Sie die Themen dieser Lektion auch in Lehrbüchern zur Algebra, die aber allgemeiner geschrieben und dadurch schwerer zu lesen sind.

Fahrplan durch die Lektion

1.1 → Abschnitt 1.1 beginnen wir mit der Definition von *Gruppen* (1.1.1) und einigen Anmerkungen zur Definition und zum Rechen in Gruppen (1.1.2–1.1.7). Danach betrachten wir einige fundamentale Beispiele (1.1.8–1.1.12). Falls Sie sich mit dem Begriff der Gruppe schwertun, dann sollten Sie sich diese Beispiele genau anschauen. Dann wenden wir uns der Frage zu, wie man in Beispielen die Gruppenaxiome nachweisen kann. In Satz 1.1.17 lernen wir dazu eine kürzere äquivalente Definition kennen. Auch der Begriff der *Untergruppe* (1.1.18 – 1.1.20) kann dabei nützlich sein.

Im ersten Teil von Abschnitt 1.2 definieren wir eine weitere algebraische Struktur: die *Ringe* (1.2.1). Das sind Mengen mit zwei Rechenoperationen, die man Addition und Multiplikation nennt. Wir besprechen grundlegende Rechenregeln (1.2.2 und 1.2.5) und Beispiele (1.2.4). Dann definieren wir *Unterringe* (1.2.8) und nutzen den Begriff, um weitere Beispiele von Ringen vorzustellen (1.2.8 und 1.2.12). ← 1.2

Der zweite, sehr kurze Teil dieses Abschnittes befasst sich mit *Einheiten* in Ringen (1.2.13). Mit diesem Begriff definieren wir die *Einheitengruppe* eines Ringes (1.2.18) und erklären, was ein *Körper* ist (1.2.19). Wir lernen, was ein *nullteilerfreier* Ring ist (1.2.21), und zeigen, dass Körper nullteilerfrei sind (1.2.22).

Im dritten Teil studieren wir Matrizen mit Einträgen aus einem Ring R (1.2.24). Dieser Teil ist wichtig, aber nicht schwierig. Wir werden sehen, dass sich die Rechenoperationen und Rechenregeln für Matrizen, die Sie aus den Mathematischen Grundlagen kennen, übertragen lassen (1.2.25, 1.2.26). Damit zeigen wir, dass die Matrizen $\mathrm{M}_{n,n}(R)$ über einem Ring R, ebenfalls einen Ring bilden (1.2.27). Danach stellen wir drei Begriffe aus der Theorie der Matrizen vor, die uns später wieder begegnen werden: die Spur (1.2.29), das Transponieren von Matrizen (1.2.30 – 1.2.33) und Blockdiagonalmatrizen (1.2.35–1.2.39). Wer möchte, kann diese drei Begriffe nur überfliegen und später genauer anschauen, wenn wir wieder darauf zurückkommen.

Im Abschnitt 1.3 konstruieren wir die *Restklassenringe* $\mathbb{Z}/n\mathbb{Z}$. Nach einer kurzen Erinnerung an die Division mit Rest (1.3.1–1.3.3) definieren wir Restklassen (1.3.4) und beweisen einen Satz mit den wichtigsten Eigenschaften (1.3.6). Danach schauen wir, wie man Restklassen addieren und multiplizieren kann (1.3.10), und beweisen, dass die Menge der Restklassen $\mathbb{Z}/n\mathbb{Z}$ ein Ring ist (1.3.11). In Satz (1.3.14) zeigen wir, dass $\mathbb{Z}/n\mathbb{Z}$ genau dann ein Körper ist, wenn n eine Primzahl ist. Damit haben wir wichtige neue Beispiele für Körper, mit denen man sich unbedingt vertraut machen sollte. Zum Abschluss besprechen wir den *Euklidischen Algorithmus* zur Berechnung multiplikativer Inverser (1.3.17). ← 1.3

Abschnitt 1.4 über Polynome ist wahrscheinlich der schwierigste in dieser Lektion. Es ist aber wichtig, dass Sie den Polynomring $K[X]$ gut verstehen. ← 1.4

Im ersten Teil definieren wir *Polynome* (1.4.1) und lernen die Rechenoperationen für Polynome kennen (1.4.5). Insbesondere werden wir zeigen, dass die Polynome einen Ring $K[X]$ bilden (1.4.9). Der *Grad* eines Polynoms ist dabei ein zentraler Begriff (1.4.3) und das unscheinbare Lemma 1.4.7 sollte man nicht übersehen.

Im kurzen, zweiten Teil lernen wir die *Division mit Rest* im Polynomring $K[X]$ kennen (1.4.11). Falls Sie die Polynomdivision schon aus der Schule kennen, lernen

Sie hier wahrscheinlich nicht viel Neues.

Der dritte Teil behandelt ausführlich die Nullstellen von Polynomen und enthält sehr wichtige Begriffe. Zum Einstieg werden einige Vorüberlegungen zum *Einsetzen* in Polynome (1.4.16, 1.4.19) und zum Unterschied zwischen Polynomen und Polynomfunktionen (1.4.17) angestellt. Nach der Definition von *Nullstellen* (1.4.20) beweisen wir den nützlichen Abspaltungssatz (1.4.21) und lernen, dass ein nicht-konstantes Polynom immer nur endlich viele Nullstellen hat (1.4.22). Wir definieren die *Vielfachheit* einer Nullstelle (1.4.24) und erarbeiten den wichtigen Satz über die Nullstellen 1.4.30. Damit erklären wir, was es heißt, dass ein *Polynom in Linearfaktoren zerfällt* (1.4.31), und wie man das nachprüfen kann (1.4.34). Das führt uns zu den *algebraisch abgeschlossenen Körpern* (1.4.40): Hier zerfällt jedes Polynom in Linearfaktoren (1.4.44).

Das Thema im vierten Teil sind *Ideale* (1.4.45) im Polynomring. Der wichtigste Satz ist Satz 1.4.49: Jedes Ideal in $K[X]$ ist ein Hauptideal. Wir besprechen die Summe von Idealen (1.4.51) und definieren damit den *größten gemeinsamen Teiler* zweier Polynome (1.4.53) und *teilerfremde* Polynome (1.4.55).

1.5 $\rightarrow$ In Abschnitt 1.5 lernen wir einen sehr wichtigen Körper kennen: den Körper $\mathbb{C}$ der komplexen Zahlen. Im ersten Teil konstruieren wir die komplexen Zahlen (1.5.1) und zeigen, dass es sich um einen Körper handelt (1.5.2). Der zweite Teil behandelt das Rechnen mit komplexen Zahlen (1.5.15–1.5.24). Bitte üben Sie das Rechnen mit komplexen Zahlen.

Im letzten Teil lernen wir den berühmten Fundamentalsatz der Algebra kennen (1.5.25): $\mathbb{C}$ ist algebraisch abgeschlossen. Diesen Satz werden wir nicht beweisen, aber wir werden zumindest nachweisen, dass jede quadratische Gleichung eine Lösung in $\mathbb{C}$ besitzt (1.5.37). Insbesondere lernen wir Quadratwurzeln komplexer Zahlen kennen (1.5.27–1.5.32).

1.1. Gruppen

In diesem Abschnitt lernen wir eine erste „algebraische Struktur“ kennen: die Gruppen. Eine Gruppe ist eine Menge mit einer Verknüpfung, die drei sehr einfachen Bedingungen – auch genannt Axiome – genügt. Man darf sich von den einfachen Axiomen nicht täuschen lassen: Gruppen sind sehr vielfältige Objekte mit einer reichhaltigen Strukturtheorie. Es handelt sich dabei um ein sehr wichtiges Konzept, denn Gruppen findet man in fast allen Bereichen der Mathematik. Gruppen spielen in der Mathematik und darüber hinaus immer dann eine Rolle, wenn es um Symmetrie geht.

Für die Lineare Algebra müssen wir aber nur sehr wenig über Gruppen und deren Strukturtheorie wissen. Das Ziel in diesem Abschnitt ist es zu verstehen, was eine Gruppe ist und uns mit dem „Rechnen“ in Gruppen vertraut zu machen.

1.1.1 **Definition** Eine *Gruppe* ist ein Paar $(G, *)$, bestehend aus einer Menge G und einer Verknüpfung

$$* : G \times G \to G, \quad (a, b) \mapsto a * b,$$

das folgende Eigenschaften erfüllt:

(G1) *Assoziativgesetz*: $a * (b * c) = (a * b) * c$ für alle $a, b, c \in G$.

(G2) Es gibt ein *neutrales Element* $e \in G$, sodass $a * e = e * a = a$ für alle $a \in G$ gilt.

(G3) Zu jedem $a \in G$ existiert ein *inverses Element* $b \in G$ mit $a * b = b * a = e$.

1.1.2 **Anmerkungen zur Definition**. (a) Die Bedingungen (G1), (G2) und (G3) nennt man die *Gruppenaxiome*.

(b) Die Verknüpfung in einer Gruppe nennt man auch *Gruppenverknüpfung* oder *Gruppenmultiplikation* aber sie wird (je nach Gruppe) durch unterschiedliche Symbole notiert, z.B. $\circ, \cdot, +$ etc. Wenn aus dem Kontext hervorgeht, um welche Verknüpfung es sich handelt, sagt man auch „G ist eine Gruppe“ ohne die Gruppenverknüpfung zu erwähnen.

(c) Ist $(G, *)$ eine Gruppe, dann gibt es in G *genau ein* neutrales Element. Wir können also von *dem* neutralen Element sprechen. Um das zu sehen, nehmen wir einmal an, dass $e, e' \in G$ zwei Elemente sind, die die Eigenschaft (G2) erfüllen. Dann gilt

$$e = e * e' = e',$$

dabei folgt die erste Gleichheit aus der Neutralität von e' und die zweite aus der von e.

Auch für das neutrale Element werden verschiedene Symbole verwendet. Beispielsweise schreibt man 0 in Gruppen, deren Verknüpfung mit $+$ notiert wird. Bevor man mit einer Gruppe arbeitet, sollte man sich immer bewusst machen, welches das neutrale Element ist.

(d) Sei $(G, *)$ eine Gruppe mit neutralem Element e. Das inverse Element von $a \in G$ ist ebenfalls eindeutig. Angenommen $b, b' \in G$ sind zwei inverse Elemente von a, dann folgt aus dem Assoziativgesetz

$$b = b * e = b * (a * b') = (b * a) * b' = e * b' = b'.$$

Das Inverse zu a bezeichnet man mit $-a$, wenn die Verknüpfung mit $+$ notiert wird. Anderenfalls verwendet man die Schreibweise a^{-1}.

1.1.3 **Rechnen mit Inversen.** *Sei $(G, *)$ eine Gruppe. Für alle $a, b \in G$ gilt*

(i) $(a^{-1})^{-1} = a$,

(ii) $(a * b)^{-1} = b^{-1} * a^{-1}$.

Beweis. Ist a^{-1} das Inverse zu a, dann gilt

$$a * a^{-1} = e = a^{-1} * a,$$

und damit ist a auch das Inverse zu a^{-1}, d.h., $(a^{-1})^{-1} = a$.

Ist a^{-1} das Inverse zu a und b^{-1} das Inverse zu b, dann gilt

$$\begin{aligned}(a * b) * (b^{-1} * a^{-1}) &= a * (b * (b^{-1} * a^{-1})) = a * ((b * b^{-1}) * a^{-1}) \\ &= a * (e * a^{-1}) = a * a^{-1} = e.\end{aligned}$$

Mit einer vergleichbaren Rechnung erhält man auch $(b^{-1} * a^{-1}) * (a * b) = e$. Somit ist $b^{-1} * a^{-1}$ das Inverse zu $a * b$. □

1.1.4 **Aufgabe**. Sei G eine Gruppe. Zeigen Sie, dass die Inversionsabbildung $G \to G$ mit $a \mapsto a^{-1}$ bijektiv ist. **L**

1.1.5 **Kürzungsregeln.** *Sei $(G, *)$ eine Gruppe. Für alle $a, b, c \in G$ gelten die Kürzungsregeln*

(i) $a * b = a * c \implies b = c$,

(ii) $b * a = c * a \implies b = c$.

Beweis. Wir nehmen an, dass $a * b = a * c$ gilt. Es sei e das neutrale Element in G und a^{-1} sei das Inverse zu a. Dann erhalten wir

$$b = e * b = (a^{-1} * a) * b = a^{-1} * (a * b) = a^{-1} * (a * c) = (a^{-1} * a) * c = e * c = c.$$

Eine ähnliche Rechnung liefert die andere Kürzungsregel (Übungsaufgabe!). □ **L**

1.1.6 **Aufgabe**. Sei $(G, *)$ eine Gruppe und sei $b \in G$. Zeigen Sie, dass die Abbildung $f\colon G \to G$ mit $f(x) = x * b$ bijektiv ist. **L**

1.1.7 **Kommutativgesetz**. Das Kommutativgesetz gehört nicht zu den Gruppenaxiomen. Gruppen, die es erfüllen, bekommen einen eigenen Namen. Eine Gruppe $(G, *)$ heißt *abelsch*[1] (oder auch *kommutativ*), wenn

$$a * b = b * a \qquad \text{(Kommutativgesetz)}$$

für alle $a, b \in G$ gilt.

In abelschen Gruppen wird für die Verknüpfung oft das Symbol $+$ verwendet.

1.1.8 **Beispiel**. (a) Die ganzen Zahlen $\mathbb{Z}$ bilden zusammen mit der gewöhnlichen Addition $+$ eine abelsche Gruppe. Die Addition ganzer Zahlen erfüllt das Assoziativ- und das Kommutativgesetz. Wegen $0 + n = n + 0 = n$ ist die Zahl 0 das neutrale Element. Das Inverse zu n ist die Zahl $-n$, denn $n + (-n) = 0$.

(b) Die Menge der natürlichen Zahlen $\mathbb{N}$ bildet mit der Addition *keine* Gruppe. Zwar ist die Addition assoziativ, aber es gibt kein neutrales Element.

(c) Die Menge $\mathbb{N}_0$ der nicht-negativen ganzen Zahlen ist mit der Addition *keine* Gruppe. Es gilt das Assoziativgesetz und die Zahl 0 ist ein neutrales Element, aber zu keiner Zahl $n > 0$ gibt es in $\mathbb{N}_0$ ein inverses Element.

(d) Ist K ein Körper, dann ist K mit der Körperaddition eine abelsche Gruppe. Das folgt direkt aus den Körperaxiomen, die Sie aus dem Modul „Mathematische Grundlagen“ [MG, Abschnitt 1.6] kennen. Insbesondere sind also $(\mathbb{Q}, +)$ und $(\mathbb{R}, +)$ abelsche Gruppen.

1.1.9 **Aufgabe**. Zeigen Sie, dass die Menge $\mathbb{R} \times \mathbb{R}$ mit der komponentenweisen Addition **L**

$$(x_1, x_2) + (y_1, y_2) := (x_1 + y_1, x_2 + y_2)$$

eine abelsche Gruppe bildet.

[1] Niels Henrik Abel: norwegischer Mathematiker, 1802–1829.

1.1.10 **Beispiel**. (a) Sei K ein Körper. Jeder K-Vektorraum bildet zusammen mit seiner Addition eine abelsche Gruppe. Das folgt direkt aus den Vektorraumaxiomen [MG, (6.1.1)]. Beispielsweise sind $(\mathbb{Q}^n, +)$ und $(\mathbb{R}^n, +)$ abelsche Gruppen.

(b) Ist K ein Körper, dann ist $K^\times = K \setminus \{0\}$ mit der Multiplikation ebenfalls eine abelsche Gruppe. Das folgt aus den Körperaxiomen, die Sie aus dem Modul „Mathematische Grundlagen“ [MG, Abschnitt 1.6] kennen. Insbesondere sind also $(\mathbb{Q}^\times, \cdot)$ und $(\mathbb{R}^\times, \cdot)$ abelsche Gruppen.

1.1.11 **Aufgabe**. Bildet die Menge $M = \mathbb{Z} \setminus \{0\}$ mit der Multiplikation $\cdot$ eine Gruppe? **L**

1.1.12 **Beispiel** (allgemeine lineare Gruppe). Es sei K ein Körper. Die Menge

$$\mathrm{GL}_n(K) = \{A \in \mathrm{M}_{n,n}(K) \mid A \text{ invertierbar } \}$$

der invertierbaren $(n \times n)$-Matrizen mit Einträgen in K bildet eine Gruppe mit der Matrizenmultiplikation. Dazu erinnern wir uns zunächst, dass das Produkt invertierbarer Matrizen wieder invertierbar ist; siehe „Mathematischen Grundlagen“ [MG, 2.3.12]. Die Matrizenmultiplikation ist also eine Verknüpfung auf $\mathrm{GL}_n(K)$. Dass die Matrizenmultiplikation das Assoziativgesetz erfüllt, ist aus [MG, (2.3.6)] bereits bekannt. Das neutrale Element ist die Einheitsmatrix

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}.$$

Das Inverse zu $A \in \mathrm{GL}_n(K)$ ist genau die inverse Matrix A^{-1}, denn es gilt $AA^{-1} = A^{-1}A = I_n$.

Die Gruppe $\mathrm{GL}_n(K)$ heißt die *allgemeine lineare Gruppe* (engl.: „**g**eneral **l**inear group“).

Für $n \geq 2$ ist die allgemeine lineare Gruppe *nicht* abelsch. Um das zu sehen, muss man zwei Matrizen $A, B \in \mathrm{GL}_n(K)$ angeben, die $AB \neq BA$ erfüllen. Für $n = 2$ kann man zum Beispiel die Matrizen

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

wählen. Man beachte, dass A und B invertierbar sind, denn

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

sind die inversen Matrizen (nachrechnen!). Es gilt jetzt

$$AB = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{und} \quad BA = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

Dabei schreiben wir 2 für das Element $1+1$ in K (im Körper $\mathbb{F}_2$ ist übrigens $2=0$). Per Definition gilt $1 \neq 0$ im Körper K und mit der Kürzungsregel auch $1+1 \neq 1$. Es gilt also $AB \neq BA$.

1.1.13 **Aufgabe**. Sei K ein Körper und sei $n \geq 2$. Zeigen Sie, dass $\mathrm{GL}_n(K)$ nicht abelsch ist. **L**

1.1.14 **Aufgabe**. Bildet die Menge $\mathrm{M}_{2,2}(\mathbb{R})$ eine Gruppe bezüglich der Matrizenmultiplikation? **L**

1.1.15 **Verknüpfungstafeln**. Ist $(G, *)$ eine Gruppe mit einer endlichen Menge G, dann kann man die Gruppe vollständig anhand einer Verknüpfungstafel beschreiben. Dazu schreibt man einfach alle Verknüpfungen $a*b$ in eine Tabelle. Zum Beispiel definiert die folgende Verknüpfungstafel eine Verknüpfung $*$ auf der Menge $\{0,1,2,3\}$:

$*$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

Aus der Verknüpfungstafel kann man zum Beispiel ablesen, dass $2*1=3$ ist. Man sieht auch, dass 0 das neutrale Element ist.

Allerdings ist die Beschreibung durch Verknüpfungstafeln sehr unpraktisch. Will man zum Beispiel zeigen, dass das Assoziativgesetz gilt, dann muss man es anhand der Verknüpfungstafel für alle Tripel $a, b, c \in G$ überprüfen. Im obigen Beispiel sind das schon $4^3 = 64$ Fälle, die betrachtet werden müssen.

1.1.16 **Das Assoziativgesetz: Rechnen ohne Klammern**. Warum schreibt man eigentlich einen Ausdruck wie $3+5+7+12$ ohne Klammern? Der Grund dafür ist das Assoziativgesetz! Es führt dazu, dass beim Rechnen mit einer assoziativen Verknüpfung vollständig auf Klammern verzichtet werden kann. Ist $(M, *)$ eine Menge mit einer assoziativen Verknüpfung $*$, dann gilt beispielsweise

$$a*(b*(c*d)) = a*((b*c)*d) = (a*(b*c))*d = ((a*b)*c)*d = (a*b)*(c*d);$$

d.h., es ist egal wie man die Klammern setzt! Man schreibt daher einfach $a*b*c*d$. Dass dies auch für beliebig lange Ausdrücke gilt, kann man mithilfe von vollständiger Induktion zeigen.

Die Definition 1.1.1 von Gruppen beinhaltet streng genommen „überflüssige“ Axiome. Mit dem nächsten Satz kann man sich etwas Arbeit ersparen, wenn man prüfen will, ob eine Menge mit einer Verknüpfung eine Gruppe ist.

1.1.17 **Satz.** *Es sei* $(G, *)$ *eine Menge mit einer Verknüpfung. Angenommen* $(G, *)$ *erfüllt das Assoziativgesetz (G1) und die beiden Bedingungen*

(G2') *Es gibt ein* links-neutrales Element $e \in G$ *mit* $e * a = a$ *für alle* $a \in G$,

(G3') *Zu jedem* $a \in G$ *existiert ein* links-inverses Element $b \in G$ *mit* $b * a = e$,

dann ist $(G, *)$ *eine Gruppe.*

Beweis. Es sei $a \in G$ und es sei e das links-neutrale Element in G. Es sei b ein links-inverses von a, d.h. $b * a = e$, und es sei c ein links-inverses von b, d.h. $c * b = e$. Dann gilt

$$a * b = e * a * b = (c * b) * a * b = c * (b * a) * b = c * (e * b) = c * b = e.$$

Das links-inverse Element b von a ist also „automatisch“ auch ein rechts-inverses Element. Weiter ist das links-neutrale Element dann auch rechts-neutral, denn

$$a * e = a * (b * a) = (a * b) * a = e * a = a. \qquad \square$$

Manchmal bildet eine Teilmenge $H \subseteq G$ einer Gruppe $(G, *)$ wieder eine Gruppe bezüglich derselben Verknüpfung $*$. Dabei muss man vorsichtig sein, denn für zwei Elemente a, b aus der Teilmenge $H \subseteq G$ kann es passieren, dass $a * b$ *nicht* in H liegt. Dies muss man ausschließen, um überhaupt eine Verknüpfung $*\colon H \times H \to H$ zu erhalten.

1.1.18 **Definition** Es sei $(G, *)$ eine Gruppe. Eine Teilmenge $H \subseteq G$ nennt man *Untergruppe* von G, wenn die folgenden drei Bedingungen erfüllt sind:

(UG1) H ist nicht leer.

(UG2) Für alle $a, b \in H$ gilt $a * b \in H$.

(UG3) Für alle $a \in H$ gilt $a^{-1} \in H$.

Erfüllt eine Teilmenge $H \subseteq G$ die Bedingung (UG2), dann sagt man, dass H *abgeschlossen* unter der Gruppenverknüpfung von G ist.

1.1.19 **Lemma.** *Es sei* $(G, *)$ *eine Gruppe und* $H \subseteq G$ *eine Untergruppe. Dann ist* H *eine Gruppe mit der Verknüpfung* $*$.

Beweis. Zunächst bemerken wir, dass (UG2) impliziert, dass $(a,b) \mapsto a * b$ wirklich eine Abbildung von $H \times H \to H$ definiert, d.h., $*$ schränkt sich zu einer Verknüpfung auf H ein. Wir prüfen, dass die Gruppenaxiome erfüllt sind.

(G1): Da $*$ das Assoziativgesetz für alle Elemente $a, b, c \in G$ erfüllt, gilt es insbesondere für alle $a, b, c \in H \subseteq G$.

(G2): Wegen (UG1) finden wir ein Element $a \in H$ und wegen (UG3) liegt auch das Inverse a^{-1} in H. Wir schreiben das neutrale Element e von G in der Form $e = a * a^{-1}$ und schließen mit (UG2), dass e in H liegt. Dieses Element erfüllt insbesondere $a * e = e * a = a$ für alle $a \in H$.

(G3) Ist $a \in H$ und ist a^{-1} sein Inverses in G, dann gilt nach (UG3) auch $a^{-1} \in H$ und wegen $a^{-1} * a = a * a^{-1} = e$ ist a^{-1} auch das Inverse zu a in H. $\square$

1.1.20 **Beispiel.** (a) Die ganzen Zahlen $\mathbb{Z}$ sind eine Untergruppe von $(\mathbb{R}, +)$. In der Tat, die Menge der ganzen Zahlen ist nicht leer, die Summe zweier ganzer Zahlen ist ganz und für eine ganze Zahl n ist auch die Zahl $-n$ ganz.

(b) Sei K ein Körper und sei V ein K-Vektorraum. Jeder Unterraum $U \subseteq V$ ist auch eine Untergruppe von $(V, +)$; siehe [MG, Proposition 6.2.3].

(c) Es sei $\mathbb{Z}[\sqrt{2}] = \{a + b\sqrt{2} \in \mathbb{R} \mid a, b \in \mathbb{Z}\}$. Dann ist $\mathbb{Z}[\sqrt{2}]$ eine Untergruppe von $(\mathbb{R}, +)$.

(UG1): Das neutrale Element 0 liegt in $\mathbb{Z}[\sqrt{2}]$, denn $0 = 0 + 0\sqrt{2}$. Also ist $\mathbb{Z}[\sqrt{2}]$ nicht leer.

(UG2): Sind $v = a + b\sqrt{2}$ und $w = c + d\sqrt{2}$ in $\mathbb{Z}[\sqrt{2}]$, dann gilt

$$v + w = a + b\sqrt{2} + c + d\sqrt{2} = (a + c) + (b + d)\sqrt{2} \in \mathbb{Z}[\sqrt{2}].$$

Also ist $\mathbb{Z}[\sqrt{2}]$ abgeschlossen unter Addition.

(UG3): Sei $v = a + b\sqrt{2} \in \mathbb{Z}[\sqrt{2}]$. Dann ist $-v = (-a) + (-b)\sqrt{2} \in \mathbb{Z}[\sqrt{2}]$.

1.1.21 **Aufgabe.** Zeigen Sie, dass die Menge $\mathbb{R}_{>0}$ der positiven reellen Zahlen eine Untergruppe von $(\mathbb{R}^{\times}, \cdot)$ ist. **L**

1.2. Ringe und Körper

In diesem Abschnitt befassen wir uns mit Mengen, die mit zwei Rechenoperationen ausgestattet sind: einer Addition und einer Multiplikation.

I. Definition und Beispiele

1.2.1 **Definition** Ein *Ring* $(R, +, \cdot)$ ist eine Menge R mit zwei Verknüpfungen

$$+\colon R \times R \to R \qquad \text{(genannt Addition)}$$
$$\cdot\colon R \times R \to R \qquad \text{(genannt Multiplikation)}$$

die folgende Eigenschaften erfüllen:

(R1) $(R, +)$ ist eine abelsche Gruppe.

(R2) Die Multiplikation erfüllt das *Assoziativgesetz*, d.h., für alle $a, b, c \in R$ gilt $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.

(R3) Es gibt ein neutrales Element $1 \in R$ der Multiplikation, d.h., für alle $a \in R$ gilt $1 \cdot a = a$ und $a \cdot 1 = a$.

(R4) Es gelten die *Distributivgesetze*, d.h.,

$$a \cdot (x + y) = (a \cdot x) + (a \cdot y) \quad \text{und} \quad (x + y) \cdot a = (x \cdot a) + (y \cdot a)$$

für alle $a, x, y \in R$.

1.2.2 **Anmerkungen zur Definition.** (a) Da die Addition und die Multiplikation assoziativ sind, können wir einen Ausdruck, in dem nur die Addition oder nur die Multiplikation vorkommt, ohne Klammern schreiben (vgl. 1.1.16). Um bei Ausdrücken mit Addition und Multiplikation Klammern zu sparen, wendet man die *Punkt-vor-Strich-Regel* an: Multiplikation bindet stärker als Addition. Zum Beispiel schreibt man $a \cdot x + b$ anstelle von $(a \cdot x) + b$.

(b) Ist $b \in R$ ein Ringelement, dann schreibt man $-b$ für das additive Inverse von b. Sind $a, b \in R$, dann schreibt man $a - b$ anstelle von $a + (-b)$.

(c) Beim Rechnen in Ringen schreibt man den Punkt $\cdot$ für die Multiplikation nicht aus, falls dadurch keine Missverständnisse entstehen, d.h., man schreibt ab für $a \cdot b$. Ist $n \in \mathbb{N}$ und a ein Ringelement, dann definiert man

$$a^n := \underbrace{a \cdot a \cdots a}_{n \text{ mal}}.$$

(d) Das neutrale Element der Addition heißt *Null* oder *Nullelement*. Es wird mit dem Symbol 0 notiert. Das neutrale Element 1 der Multiplikation heißt *Einselement*.

(e) Sind $a, x, y, z \in R$, dann kann man mit dem Assoziativgesetz der Addition und den Distributivgesetzen folgende Rechnung durchführen:

$$a \cdot (x + y + z) = a \cdot (x + (y + z)) = ax + a \cdot (y + z) = ax + ay + az.$$

Wendet man dieses Argument induktiv an, dann sieht man, dass für alle $a, x_1, \ldots, x_n \in R$ die allgemeinen Distributivgesetze

$$a \cdot \left(\sum_{i=1}^{n} x_i\right) = \sum_{i=1}^{n} ax_i \qquad \text{und} \qquad \left(\sum_{i=1}^{n} x_i\right) \cdot a = \sum_{i=1}^{n} x_i a$$

gültig sind.

(f) In der Literatur findet man konkurrierende Definitionen für den Begriff „Ring“. Insbesondere ist es verbreitet Ringe ohne Axiom (R3) zu definieren. Ringe mit Axiom (R3) werden dann *unitäre Ringe* genannt. Beim Nachschlagen in anderen Quellen sollte man also aufmerksam sein.

1.2.3 **Definition** Ein *kommutativer Ring* ist ein Ring $(R, +, \cdot)$, dessen Multiplikation das Kommutativgesetz erfüllt, d.h., es gilt $ab = ba$ für alle $a, b \in R$.

1.2.4 **Beispiel**. (a) Die ganzen Zahlen $\mathbb{Z}$ bilden mit der gewöhnlichen Addition und Multiplikation einen kommutativen Ring.

(b) Die reellen Zahlen $\mathbb{R}$ bilden einen kommutativen Ring mit der üblichen Addition und Multiplikation reeller Zahlen.

(c) Sei $n \in \mathbb{N}$ und sei K ein Körper. Die Menge $\mathrm{M}_{n,n}(K)$ aller $(n \times n)$-Matrizen bildet mit der Addition und Matrizenmultiplikation einen Ring. Da $\mathrm{M}_{n,n}(K)$ ein K-Vektorraum ist, ist $(\mathrm{M}_{n,n}(K), +)$ eine abelsche Gruppe; siehe 1.1.10 (a). Aus den „Mathematischen Grundlagen“ [MG, (2.3.6)] ist bekannt, dass die Matrizenmultiplikation assoziativ ist und, dass die Distributivgesetze gelten [MG, (2.4.1)]. Die Einheitsmatrix I_n ist das Einselement.

Für $n \geq 2$ ist $\mathrm{M}_{n,n}(K)$ nicht kommutativ; siehe Aufgabe 1.1.13.

Wir halten einige grundsätzliche Rechenregeln in Ringen fest.

1.2.5 **Lemma.** (Rechenregeln in Ringen) *Sei* $(R, +, \cdot)$ *ein Ring. Für alle* $a, b \in R$ *gelten folgende Rechenregeln:*

(i) $0 \cdot a = 0 = a \cdot 0$.

(ii) $(-1) \cdot a = -a = a \cdot (-1)$.

(iii) $(-a) \cdot b = -(ab) = a \cdot (-b)$.

Beweis. Zu (i): Es gilt

$$0 = 0 \cdot a - 0 \cdot a = (0+0) \cdot a - 0 \cdot a = 0 \cdot a + 0 \cdot a - 0 \cdot a = 0 \cdot a.$$

Genauso sieht man auch $a \cdot 0 = 0$.

Zu (ii): Durch das Distributivgesetz und (i) erhalten wir

$$(-1) \cdot a + a = (-1) \cdot a + 1 \cdot a = (-1+1) \cdot a = 0 \cdot a = 0$$

und damit ist $(-1) \cdot a$ das additive Inverse zu a, d.h., $(-1) \cdot a = -a$. Genauso zeigt man $a \cdot (-1) = -a$.

Zu (iii): Mit (ii) und dem Assoziativgesetz der Multiplikation erhalten wir

$$\begin{aligned}(-a) \cdot b &= ((-1) \cdot a) \cdot b = (-1) \cdot (ab) \\ &= -(ab) = (ab) \cdot (-1) && \text{(mit (ii))} \\ &= a \cdot (b \cdot (-1)) = a \cdot (-b)\end{aligned}$$

und damit sind alle Rechenregeln gezeigt. □

1.2.6 **Aufgabe**. Es sei R ein Ring mit Elementen $a_1, \ldots, a_n, b_1, \ldots, b_m \in R$. Zeigen Sie, dass die Gleichung $\left(\sum_{i=1}^n a_i\right) \cdot \left(\sum_{j=1}^m b_j\right) = \sum_{i=1}^n \sum_{j=1}^m a_i b_j$ gilt. **L**

1.2.7 **Bemerkung: Der Nullring**. Sei R ein Ring. Gilt im Ring R die Gleichung $0 = 1$, dann besteht der Ring nur aus diesem einen Element 0. In der Tat, sei $r \in R$ beliebig. Dann gilt wegen 1.2.5

$$r = 1 \cdot r \overset{1=0}{=} 0 \cdot r = 0.$$

Einen Ring, der nur aus dem Nullelement besteht, nennt man den *Nullring*. Der Nullring ist so uninteressant, dass wir immer annehmen, dass $1 \neq 0$ gilt.

Analog zum Begriff der Untergruppe gibt es auch den Begriff „Unterring".

1.2.8 **Definition** Sei $(R, +, \cdot)$ ein Ring. Eine Teilmenge $T \subseteq R$ heißt *Unterring*, wenn folgende drei Eigenschaften erfüllt sind

(UR1) T ist eine Untergruppe von $(R, +)$,

(UR2) das Einselement von R liegt in T,

(UR3) für alle $a, b \in T$ gilt $a \cdot b \in T$.

1.2.9 **Lemma.** *Sei $(R, +, \cdot)$ ein Ring und $T \subseteq R$ ein Unterring, dann ist T mit der Addition und Multiplikation aus R wieder ein Ring.*

Beweis. Durch die Bedingung (UR1) ist $(T, +)$ eine abelsche Gruppe, denn Untergruppen von abelschen Gruppen sind abelsch. Durch die Bedingung (UR3) lässt sich auch die Multiplikation in R zu einer Verknüpfung $T \times T \to T$ einschränken. Da das Assoziativgesetz und die Distributivgesetze für alle Elemente aus R gültig sind, sind sie insbesondere für alle Elemente aus der Teilmenge T erfüllt. Schließlich sichert (UR2) die Existenz eines neutralen Elementes für die Multiplikation. □

1.2.10 **Bemerkung**. Zu prüfen, dass eine Teilmenge eines Ringes ein Unterring ist, macht im Allgemeinen deutlich weniger Aufwand als von neuem alle Ringaxiome nachzuweisen.

1.2.11 **Beispiel**. (a) Der Ring der ganzen Zahlen $\mathbb{Z}$ ist ein Unterring von $(\mathbb{Q}, +, \cdot)$.

(b) Es sei $\mathbb{Z}[\sqrt{2}] = \{a + b\sqrt{2} \mid a, b \in \mathbb{Z}\}$. Dann ist $\mathbb{Z}[\sqrt{2}]$ ein Unterring von $(\mathbb{R}, +, \cdot)$.

(UR1): In Beispiel 1.1.20 (c) haben wir gesehen, dass $(\mathbb{Z}[\sqrt{2}], +)$ eine Untergruppe von $(\mathbb{R}, +)$ ist.

(UR2): Die Zahl 1 liegt in $\mathbb{Z}[\sqrt{2}]$, denn $1 = 1 + 0 \cdot \sqrt{2}$.

(UR3): Es seien $x = a + b\sqrt{2}$ und $x' = a' + b'\sqrt{2}$ in $\mathbb{Z}[\sqrt{2}]$, d.h., $a, a', b, b' \in \mathbb{Z}$. Dann gilt

$$x \cdot x' = (a + b\sqrt{2}) \cdot (a' + b'\sqrt{2}) = (aa' + 2bb') + (ab' + ba')\sqrt{2}.$$

Da $aa' + 2bb'$ und $ab' + ba'$ in $\mathbb{Z}$ liegen, gilt $x \cdot x' \in \mathbb{Z}[\sqrt{2}]$.

Insgesamt ist also $\mathbb{Z}[\sqrt{2}]$ ein Unterring von $\mathbb{R}$.

1.2.12 **Aufgabe**. Wir betrachten die Teilmenge **L**

$$B = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{R}) \mid c = 0 \right\} \subseteq \mathrm{M}_{2,2}(\mathbb{R}).$$

Zeigen Sie, dass B ein Unterring von $\mathrm{M}_{2,2}(\mathbb{R})$ ist.

II. Einheiten

1.2.13 $\boxed{\textbf{Definition}}$ Sei $(R, +, \cdot)$ ein Ring. Ein Element $a \in R$ heißt *invertierbar*, falls ein $b \in R$ mit $ab = 1$ und $ba = 1$ existiert. Man nennt b das *Inverse* von a und bezeichnet es mit a^{-1}.

Die invertierbaren Elemente nennt man auch *Einheiten* von R. Die Menge aller Einheiten in R bezeichnet man mit $R^\times$.

1.2.14 **Anmerkung zur Definition**. (a) Das Einselement 1 ist immer eine Einheit, denn $1 \cdot 1 = 1$.

(b) Angenommen es gilt $1 \neq 0$ in R, d.h. R ist nicht der Nullring; siehe 1.2.7. Dann ist 0 *keine* Einheit in R. In der Tat, wegen 1.2.5 (i) gilt für alle $b \in R$

$$b \cdot 0 = 0 \neq 1.$$

(c) Wie in 1.1.2 (d) kann man zeigen, dass das Inverse zu $a \in R^\times$ eindeutig ist.

(d) Sei $a \in R$ eine Einheit mit Inversem a^{-1}. Dann ist auch a^{-1} eine Einheit mit Inversem $(a^{-1})^{-1} = a$, denn es gilt $aa^{-1} = a^{-1}a = 1$.

1.2.15 **Beispiel**. (a) Der Ring $\mathbb{Z}$ der ganzen Zahlen hat genau zwei Einheiten: $\mathbb{Z}^\times = \{1, -1\}$.

(b) Sei K ein Körper und sei $n \in \mathbb{N}$. Die Einheiten in $\mathrm{M}_{n,n}(K)$ sind genau die invertierbaren $(n \times n)$-Matrizen, d.h., $\mathrm{M}_{n,n}(K)^\times = \mathrm{GL}_n(K)$.

1.2.16 **Aufgabe**. Zeigen Sie, dass $1 - \sqrt{2}$ eine Einheit in $\mathbb{Z}[\sqrt{2}]$ ist. **L**

1.2.17 **Lemma.** *Sei $(R, +, \cdot)$ ein Ring.*

(i) Sind $a, b \in R^\times$, dann ist $ab \in R^\times$.

(ii) $(R^\times, \cdot)$ ist eine Gruppe.

Beweis. (i): Seien $a, b \in R^\times$. Wir definieren $c := b^{-1}a^{-1}$. Dann gilt

$$c(ab) = b^{-1}a^{-1}ab = b^{-1}b = 1$$

und genauso sieht man $(ab)c = 1$. Damit ist ab eine Einheit.

(ii): Aus (i) folgt, dass die Multiplikation $\cdot$ sich zu einer Verknüpfung auf $R^\times$ einschränkt. Wegen Ringaxiom (R2) gilt das Assoziativgesetz. Das Element $1 \in R$ ist eine Einheit und ist das neutrale Element für die Multiplikation. Die Existenz eines inversen Elementes zu a in $R^\times$ folgt aus 1.2.14 (d). $\square$

1.2.18 $\boxed{\textbf{Definition}}$ Die Gruppe $(R^\times, \cdot)$ heißt *Einheitengruppe* des Ringes R.

Körper kennen Sie bereits aus den Mathematischen Grundlagen [MG, 1.6]. Mit den gerade eingeführten Begriffen können wir eine alternative Definition angeben. Es ist eine gute Übung zu prüfen, dass die folgende Definition zu der bekannten Definition aus den Mathematischen Grundlagen äquivalent ist.

1.2.19 $\boxed{\textbf{Definition}}$ Ein kommutativer Ring $(K, +, \cdot)$ heißt *Körper* (engl.: „field“), wenn

$$K^\times = K \setminus \{0\}$$

gilt.

1.2.20 **Bemerkungen.** (a) In jedem Körper K gilt $0 \neq 1$. In der Tat, das Einselement 1 liegt in $K^\times$, aber das Nullelement 0 liegt wegen Definition 1.2.19 nicht in der Einheitengruppe von K.

(b) Sei $(R, +, \cdot)$ ein Ring und $K \subseteq R$ ein Unterring. Ist K mit der geerbten Ringstruktur ein Körper, dann nennt man K auch *Unterkörper*.

1.2.21 $\boxed{\textbf{Definition}}$ Ein Ring $(R, +, \cdot)$ heißt *nullteilerfrei*, wenn für alle $a, b \in R$ mit $a \neq 0$ und $b \neq 0$ auch $ab \neq 0$ gilt.

1.2.22 **Lemma.** *Jeder Körper ist nullteilerfrei.*

Beweis. Es seien $a, b \in K$ mit $a \neq 0$ und $b \neq 0$ gegeben. Also liegen a, b in $K^\times$ und aus 1.2.17 folgt, dass auch ab in $K^\times = K \setminus \{0\}$ liegt. Insbesondere ist $ab \neq 0$. □

1.2.23 **Kürzungsregel.** *Sei R ein nullteilerfreier Ring. Für alle $a \neq 0$ und $b, c \in R$ gilt die* Kürzungsregel*:*

$$ab = ac \implies b = c.$$

Den Beweis der Kürzungsregel lassen wir als Übungsaufgabe. **L**

III. Matrizen über Ringen

Aus den mathematischen Grundlagen [MG, Kapitel 2] kennen wir bereits Matrizen mit Einträgen aus einem Körper. Wir müssen in diesem Lehrtext manchmal allgemeine Matrizen betrachten, deren Einträge in einem Ring liegen. Die Definitionen und Rechenregeln bleiben dabei unverändert. In diesem Abschnitt werden kurz die wichtigsten Eigenschaften zusammengestellt.

1.2.24 $\boxed{\textbf{Definition}}$ Es seien $(R, +, \cdot)$ ein Ring und $n, m \in \mathbb{N}$. Eine $(m \times n)$-*Matrix* mit Einträgen aus R ist eine rechteckige Anordnung

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & a_{2,3} & \cdots & a_{2,n} \\ \vdots & & & & \vdots \\ a_{m,1} & a_{m,2} & a_{m,3} & \cdots & a_{m,n} \end{pmatrix}$$

von Ringelementen $a_{i,j} \in R$ in m Zeilen und n Spalten. Man schreibt kurz $A = (a_{i,j})_{i,j}$. Dabei nennt man $a_{i,j}$ den *Eintrag* der Matrix an der Stelle (i, j). Wenn keine Verwechselungsgefahr besteht, schreiben wir a_{ij} statt $a_{i,j}$. Die Menge aller $(m \times n)$-Matrizen mit Einträgen aus R bezeichnen wir mit $\mathrm{M}_{m,n}(R)$.

1.2.25 **Rechnen mit Matrizen**. Auf der Menge $\mathrm{M}_{m,n}(R)$ definieren wir die Addition von Matrizen eintragsweise

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ b_{21} & \cdots & b_{2n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix} := \begin{pmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & \cdots & a_{2n} + b_{2n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{pmatrix}.$$

Die Matrix deren Einträge alle 0 sind, nennt man die *Nullmatrix* und schreibt dafür 0. Ist $A = (a_{ij})_{i,j}$ eine Matrix und ist $r \in R$, dann definieren wir $rA = (ra_{ij})_{i,j}$.

Es seien $\ell, m, n \in \mathbb{N}$ gegeben. Für Matrizen $A = (a_{ij})_{i,j} \in \mathrm{M}_{\ell,m}(R)$ und $B = (b_{ij})_{i,j} \in \mathrm{M}_{m,n}(R)$ definieren wir eine Matrix $AB \in \mathrm{M}_{\ell,n}(R)$ mit Einträgen $(c_{ij})_{i,j}$ durch die aus [MG, Abschnitt 2.3] bekannte Formel

$$c_{ij} := \sum_{k=1}^{m} a_{ik} b_{kj}.$$

Die Abbildung $M_{\ell,m}(R) \times M_{m,n}(R) \to M_{\ell,n}(R)$ mit $(A, B) \mapsto AB$ nennt man die *Matrizenmultiplikation*.

Die quadratische Matrix $I_n = (a_{ij})_{i,j} \in \mathrm{M}_{n,n}(R)$ mit Diagonaleinträgen $a_{ii} = 1$ für alle $i \in \{1, \ldots, n\}$ und Einträgen $a_{ij} = 0$ für alle $i \neq j$ nennt man die *Einheitsmatrix*. Für alle $A \in \mathrm{M}_{m,n}(R)$ gilt

$$I_m A = A = A I_n. \tag{1.2.a}$$

1.2.26 **Aufgabe**. Es sei R ein Ring und es seien $m, n, s, t \in \mathbb{N}$ **L**

(i) Zeigen Sie, dass $(\mathrm{M}_{m,n}(R), +)$ eine abelsche Gruppe bildet.

(ii) Seien $A \in \mathrm{M}_{m,n}(R)$, $B \in \mathrm{M}_{n,s}(R)$ und $C \in \mathrm{M}_{s,t}(R)$ gegeben. Zeigen Sie, dass

$$(AB)C = A(BC)$$

gilt.

1.2.27 **Satz.** *Sei R ein Ring und sei $n \in \mathbb{N}$. Dann ist $\mathrm{M}_{n,n}(R)$ mit der Addition und Multiplikation von Matrizen ein Ring mit Einselement I_n.*

Beweis. Wir prüfen die Ringaxiome aus 1.2.1.

(R1): Dass $(\mathrm{M}_{n,n}(R), +)$ eine abelsche Gruppe ist, wurde in Aufgabe 1.2.26 (i) gezeigt.

(R2): Das Assoziativgesetz der Multiplikation in $\mathrm{M}_{n,n}(R)$ folgt aus Aufgabe 1.2.26 (ii).

(R3): Aus Gleichung (1.2.a) folgt, dass I_n das neutrale Element der Multiplikation in $\mathrm{M}_{n,n}(R)$ ist.

(R4): Es seien $A, B, C \in \mathrm{M}_{n,n}(R)$. Wir zeigen, dass das Distributivgesetz

$$A(B + C) = AB + AC$$

erfüllt ist. Es seien $i, j \in \{1, 2, \ldots, n\}$. Wir zeigen, dass die Einträge von $A(B + C)$ und $AB + AC$ an der Stelle (i, j) übereinstimmen. Mit dem Distributivgesetz in R erhalten wir

$$\sum_{k=1}^{n} a_{ik}(b_{kj} + c_{kj}) = \sum_{k=1}^{n} a_{ik}b_{kj} + a_{ik}c_{kj} = \sum_{k=1}^{n} a_{ik}b_{kj} + \sum_{k=1}^{n} a_{ik}c_{kj}.$$

Genauso kann man auch das Distributivgesetz $(B + C)A = BA + CA$ herleiten. □

1.2.28 **Einheiten: Die allgemeine lineare Gruppe.** Die Einheitengruppe von $\mathrm{M}_{n,n}(R)$ nennt man *allgemeine lineare Gruppe* und bezeichnet diese mit $\mathrm{GL}_n(R)$.

1.2.29 **Aufgabe** (Die Spur). Sei R ein kommutativer Ring. Die *Spur* einer quadratischen Matrix $A = (a_{ij})_{i,j} \in \mathrm{M}_{n,n}(R)$ ist definiert als **L**

$$\mathrm{Spur}(A) = \sum_{i=1}^{n} a_{ii},$$

d.h., die Spur ist die Summe aller Diagonaleinträge. Zeigen Sie, dass für alle $A, B \in \mathrm{M}_{n,n}(R)$ und alle $a \in R$ folgende Aussagen gelten:

(a) $\mathrm{Spur}(A + B) = \mathrm{Spur}(A) + \mathrm{Spur}(B)$,

(b) $\mathrm{Spur}(aA) = a\,\mathrm{Spur}(A)$ und

(c) $\mathrm{Spur}(AB) = \mathrm{Spur}(BA)$.

Mit Matrizen kann man noch eine weitere Rechenoperation definieren, die später im Lehrtext noch von Bedeutung sein wird.

1.2.30 **Definition** Sei R ein kommutativer Ring und sei $A = (a_{ij})_{i,j} \in \mathrm{M}_{m,n}(R)$ eine Matrix. Wir definieren die zu A *transponierte Matrix* A^T als

$$A^T := (a_{ji})_{i,j} \in \mathrm{M}_{n,m}(R).$$

Der Eintrag von A^T an der Stelle (i,j) ist also der Eintrag von A an der Stelle (j,i). In anderen Worten: Beim Transponieren wird die i-te Spalte von A zur i-ten Zeile von A^T.

1.2.31 **Beispiel**. Die transponierte Matrix zu

$$A = \begin{pmatrix} 1 & 2 & 7 \\ 3 & 4 & 0 \end{pmatrix} \in \mathrm{M}_{2,3}(\mathbb{Z})$$

ist die Matrix

$$A^T = \begin{pmatrix} 1 & 3 \\ 2 & 4 \\ 7 & 0 \end{pmatrix} \in \mathrm{M}_{3,2}(\mathbb{Z}).$$

Das Transponieren verträgt sich gut mit der Addition und Multiplikation von Matrizen.

1.2.32 **Rechenregeln beim Transponieren.** *Sei R ein kommutativer Ring. Für alle $A, B \in \mathrm{M}_{m,n}(R)$, $C \in \mathrm{M}_{\ell,m}(R)$ und $r \in R$ gelten folgende Rechenregeln:*

(i) $(A+B)^T = A^T + B^T$,

(ii) $(CA)^T = A^T C^T$,

(iii) $(rA)^T = rA^T$,

(iv) $(A^T)^T = A$.

Beweis. Es sei $A = (a_{ij})_{i,j}$, $B = (b_{ij})_{i,j}$ und $C = (c_{ij})_{i,j}$. Dann folgt (i) aus der Rechnung

$$(A+B)^T = (a_{ij}+b_{ij})^T_{i,j} = (a_{ji}+b_{ji})_{i,j} = A^T + B^T.$$

Außerdem gilt (iv), denn

$$(A^T)^T = ((a_{ij})^T_{i,j})^T = (a_{ji})^T_{i,j} = (a_{ij})_{i,j} = A.$$

Genauso einfach kann man auch $(rA)^T = rA^T$ zeigen.

Betrachten wir nun die Matrizenmultiplikation. Die Matrix $CA = (d_{ij})_{i,j}$ ist eine $(\ell \times n)$-Matrix und $(CA)^T$ entsprechend eine $(n \times \ell)$-Matrix. Der Eintrag von $(CA)^T$ an der Stelle (i,j) ist

$$d_{ji} = \sum_{k=1}^{m} c_{jk} a_{ki} = \sum_{k=1}^{m} a_{ki} c_{jk}.$$

Dabei ist der Ausdruck rechts genau der Eintrag von $A^T C^T$ an der Stelle (i,j), denn a_{ki} ist der Eintrag von A^T an der Stelle (i,k) und c_{jk} ist der Eintrag von C^T an der Stelle (k,j). Da i und j beliebig waren, gilt $(CA)^T = A^T C^T$. □

1.2.33 **Aufgabe**. Beweisen oder widerlegen Sie: Für $A \in \mathrm{M}_{n,n}(R)$ gilt immer $A^T A = A A^T$. **L**

1.2.34 **Aufgabe**. Sei K ein Körper und sei $A \in \mathrm{M}_{n,n}(K)$ invertierbar. Zeigen Sie, dass A^T invertierbar ist und, dass $(A^T)^{-1} = (A^{-1})^T$ gilt. **L**

Zum Abschluss besprechen wir noch Blockdiagonalmatrizen. Die Ergebnisse dazu werden aber erst später benötigt. Es genügt, wenn Sie sich zunächst einen groben Überblick verschaffen.

1.2.35 **Definition** Sei $n \in \mathbb{N}$ und seien $p_1, \dots, p_s \in \mathbb{N}$ mit $n = \sum_{r=1}^{s} p_r$ gegeben. Eine Matrix $A = (a_{i,j}) \in \mathrm{M}_{n,n}(R)$ nennt man *Blockdiagonalmatrix* vom Blocktyp $p_1, \dots, p_s$, wenn aus $a_{i,j} \neq 0$ folgt, dass $\sum_{r=1}^{t} p_r < i,j \leq \sum_{r=1}^{t+1} p_r$ für ein $t \in \{0, \dots, s-1\}$ gilt. Das heißt, A hat die Gestalt

$$A = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_s \end{pmatrix},$$

wobei A_t eine $(p_t \times p_t)$-Matrix ist und alle nicht angegebenen Einträge Null sind. Die Matrizen A_t nennt man die Blöcke von A. Wir schreiben in diesem Fall auch

$$A = \operatorname{diag}(A_1, A_2, \dots, A_s)$$

um die Blöcke deutlich zu machen.

1.2.36 Beispiel. Die Matrix

$$A = \begin{pmatrix} 1 & 2 & 0 & 0 & 0 \\ 3 & 4 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 8 & 7 \end{pmatrix} \in \mathrm{M}_{5,5}(\mathbb{R})$$

ist eine Blockdiagonalmatrix vom Blocktyp $2, 1, 2$. Es gilt

$$A = \operatorname{diag}\left(\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, 5, \begin{pmatrix} 6 & 0 \\ 8 & 7 \end{pmatrix} \right).$$

1.2.37 Anmerkung zur Definition. Es sei $A = \operatorname{diag}(A_1, \ldots, A_s) \in \mathrm{M}_{n,n}(R)$ eine Blockdiagonalmatrix vom Typ $p_1, \ldots, p_s$. Was sind die Einträge von A_{t+1}? Sei $P_t = \sum_{r=1}^{t} p_r$. Dann ist der Eintrag von A_{t+1} an der Stelle (k, ℓ) genau der Eintrag von A an der Stelle $(k + P_t, \ell + P_t)$.

1.2.38 Aufgabe. Es seien $A = \operatorname{diag}(A_1, \ldots, A_s)$ und $B = \operatorname{diag}(B_1, \ldots, B_s)$ Blockdiagonalmatrizen in $\mathrm{M}_{n,n}(R)$ vom selben Blocktyp. Beweisen Sie die Gleichung **L**

$$AB = \operatorname{diag}(A_1B_1, A_2B_2, \ldots, A_sB_s).$$

1.2.39 Lemma. *Sei R ein kommutativer Ring und seien $p_1, \ldots, p_s \in \mathbb{N}$ mit $n = \sum_{r=1}^{s} p_r$. Die Menge aller Blockdiagonalmatrizen vom Blocktyp $p_1, \ldots, p_s$ bildet einen Unterring von $\mathrm{M}_{n,n}(R)$.*

Beweis. Es sei T die Menge aller Blockdiagonalmatrizen vom Blocktyp $p_1, \ldots, p_s$. Wir verifizieren die Axiome 1.2.8.

Zu (UR1): Wir prüfen die Untergruppeneigenschaften für T. Die Nullmatrix ist Blockdiagonal für jeden Blocktyp, daher ist T nicht leer.

Es seien $A = (a_{ij})_{i,j}$ und $B = (b_{ij})$ in T. Es gilt $A + B = (a_{ij} + b_{ij})$. Angenommen $a_{ij} + b_{ij} \neq 0$, dann muss eines der beiden Ringelemente a_{ij} oder b_{ij} ungleich Null sein. Sagen wir $a_{ij} \neq 0$. Da A eine Blockdiagonalmatrix vom Typ $p_1, \ldots, p_s$ ist, gibt es ein t mit $\sum_{r=1}^{t} p_r < i, j \leq \sum_{r=1}^{t+1} p_r$. Genauso argumentiert man, falls $b_{ij} \neq 0$ ist. Also ist $A + B$ in T.

Ganz einfach verifiziert man (UG3): $-a_{ij} \neq 0$ genau dann, wenn $a_{ij} \neq 0$ ist. Also ist für $A \in T$ auch das additive Inverse $-A$ in T. Wir schließen, dass T eine Untergruppe von $(R, +)$ ist.

Zu (UR2): Das Einselement von $\mathrm{M}_{n,n}(R)$ ist die Einheitsmatrix I_n und es gilt

$$I_n = \mathrm{diag}(I_{p_1}, \ldots, I_{p_s}) \in T.$$

Das Unterringaxiom (UR3) haben wir bereits in Aufgabe 1.2.38 gezeigt. Damit ist T ein Unterring. $\square$

1.3. Restklassenringe und endliche Körper

In diesem Abschnitt werden wir eine wichtige Familie kommutativer Ringe kennenlernen: die Restklassenringe $\mathbb{Z}/n\mathbb{Z}$. Um diese Ringe zu definieren, benötigen wir einige Begriffe aus der elementaren Zahlentheorie.

1.3.1 **Definition** Es seien $a, b \in \mathbb{Z}$ ganze Zahlen. Man nennt a einen *Teiler* von b, wenn es ein $c \in \mathbb{Z}$ mit $ac = b$ gibt. Man sagt a teilt b und schreibt $a \mid b$.

Aus der Schule kennen Sie sicherlich bereits die *Division mit Rest.*

1.3.2 **Satz zur Division mit Rest.** *Es seien $a, b \in \mathbb{Z}$ mit $a > 0$ gegeben. Dann gibt es eindeutige ganze Zahlen q, r mit $0 \leq r < a$, sodass*

$$b = qa + r$$

gilt. Man nennt q den ganzen Quotienten *und r den* Rest *der Division von b durch a.*[2]

<u>*Beweis*</u>. Zuerst zeigen wir, dass q und r existieren. Die Menge

$$M = \{x \in \mathbb{Z} \mid xa > b\}$$

ist nicht leer und nach unten durch $\frac{b}{a}$ beschränkt, d.h., M besitzt ein Minimum. Es sei $q := \min(M) - 1$. Dann ist $qa \leq b$ aber $(q+1)a > b$ und damit gilt $0 \leq b - qa < a$. Wir definieren also $r = b - qa$ und erhalten

$$b = qa + r.$$

Warum sind q und r mit diesen Eigenschaften eindeutig? Angenommen wir haben eine weitere Darstellung $b = q'a + r'$ mit $0 \leq r' < a$. Ohne Einschränkung können wir $r' \geq r$ annehmen. Dann ist $0 \leq r' - r < a$. Da aber $r' - r = (q - q')a$ ein Vielfaches von a ist, muss $q = q'$ sein. Daraus folgt auch $r' = b - q'a = b - qa = r$. □

1.3.3 **Bemerkungen**. (a) Um den ganzen Quotienten und den Rest der Division von b durch a zu bestimmen, verwendet man das aus der Schule bekannte Divisionsverfahren mit Rest.

(b) Für $a > 0$ gilt: a teilt b genau dann, wenn bei der Division von b durch a kein Rest bleibt.

[2] b heißt *Dividend* und a heißt *Divisor*.

1.3.4 **Definition** Es sei $n \in \mathbb{N}$. Für $k \in \mathbb{Z}$ heißt die Menge

$$k + n\mathbb{Z} = \{k + qn \in \mathbb{Z} \mid q \in \mathbb{Z}\}$$

die *Restklasse von k modulo n.*

Die Menge aller Restklassen modulo n bezeichnen wir mit $\mathbb{Z}/n\mathbb{Z}$.[3]

1.3.5 **Beispiel**. Die Restklasse von 0 modulo 2 besteht genau aus den Zahlen $2q$, d.h., $0 + 2\mathbb{Z}$ ist die Menge der geraden Zahlen. Die Restklasse von 1 modulo 2 besteht genau aus den Zahlen der Form $1 + 2q$, d.h., $1 + 2\mathbb{Z}$ ist die Menge der ungeraden Zahlen.

1.3.6 **Satz von den Restklassen.** *Es seien $n \in \mathbb{N}$ und $k, \ell \in \mathbb{Z}$.*

(a) Die Restklassen $k + n\mathbb{Z}$ und $\ell + n\mathbb{Z}$ sind entweder gleich oder disjunkt.

(b) Die folgenden Aussagen sind äquivalent:

(i) $k + n\mathbb{Z} = \ell + n\mathbb{Z}$.

(ii) Es gibt ein $a \in \mathbb{Z}$ mit $k - \ell = na$.

(iii) Der Rest der Division von k durch n und der Rest der Division von ℓ durch n sind gleich.

(c) Es gibt genau n verschiedene Restklassen modulo n:

$$0 + n\mathbb{Z}, 1 + n\mathbb{Z}, \dots, (n-1) + n\mathbb{Z}.$$

Beweis. (a): Angenommen $k + n\mathbb{Z}$ und $\ell + n\mathbb{Z}$ sind nicht disjunkt, dann gibt es eine Zahl x im Schnitt der Restklassen, d.h.,

$$x = k + na = \ell + nb.$$

Ist $\ell + nq \in \ell + n\mathbb{Z}$ gegeben, dann gilt

$$\ell + nq = \ell + nb + n(q - b) = k + na + n(q - b) = k + n(a + q - b) \in k + n\mathbb{Z}.$$

Daraus folgt $\ell + n\mathbb{Z} \subseteq k + n\mathbb{Z}$. Vertauscht man die Rollen von k und ℓ, folgt außerdem $k + n\mathbb{Z} \subseteq \ell + n\mathbb{Z}$, d.h., die Restklassen sind gleich.

[3] $\mathbb{Z}/n\mathbb{Z}$ ist also eine Menge, deren Elemente selbst Mengen sind.

(b):
„(i) $\implies$ (ii)“: Angenommen es gilt $k + n\mathbb{Z} = \ell + n\mathbb{Z}$. Dann liegt $k \in \ell + n\mathbb{Z}$ und es gibt damit eine Zahl $a \in \mathbb{Z}$ mit $k - \ell = na$.

„(ii) $\implies$ (iii)“: Angenommen $k - \ell = an$ für eine Zahl $a \in \mathbb{Z}$. Durch Division von k durch n mit Rest (siehe 1.3.2) finden wir $q, r \in \mathbb{Z}$ mit $0 \leq r < n$ und $k - qn = r$ und es folgt $\ell = k - an = (q - a)n + r$. Wegen $0 \leq r < n$ ist r der Rest der Division von ℓ durch n.

„(iii) $\implies$ (i)“: Wir nehmen an, dass die Division mit n denselben Rest r liefert, d.h., $k = qn + r$ und $\ell = q'n + r$ für gewisse $q, q' \in \mathbb{Z}$. Dann liegt r im Schnitt $(k + n\mathbb{Z}) \cap (\ell + n\mathbb{Z})$ und aus (a) folgt $k + n\mathbb{Z} = \ell + n\mathbb{Z}$.

(c): Aus (b) folgt, dass die Restklassen

$$0 + n\mathbb{Z}, 1 + n\mathbb{Z}, \ldots, (n-1) + n\mathbb{Z}$$

paarweise verschieden sind.

Es gibt keine anderen Restklassen. Sei $k \in \mathbb{Z}$ beliebig. Durch Division durch n mit Rest erhalten wir $k = qn + r$ mit $0 \leq r < n$ und damit ist $k + n\mathbb{Z} = r + n\mathbb{Z}$. □

1.3.7 **Notation**. Meist legen wir eine natürliche Zahl $n \in \mathbb{N}$ fest und betrachten nur Restklassen modulo n. In diesem Fall verwenden wir die kurze Schreibweise

$$\overline{k} := k + n\mathbb{Z}.$$

1.3.8 **Vertreter**. Sei $n \in \mathbb{N}$ und sei $\overline{k}$ eine Restklasse modulo n. Ein Element $\ell \in \overline{k}$ nennen wir *Vertreter* der Restklasse. Wegen $\ell \in \overline{k}$ sind die Restklassen $\overline{\ell}$ und $\overline{k}$ nicht disjunkt, und nach 1.3.6 (a) gilt damit $\overline{\ell} = \overline{k}$.

1.3.9 **Aufgabe**. Welche der folgenden Restklassen aus $\mathbb{Z}/18\mathbb{Z}$ sind identisch mit $\overline{5} \in \mathbb{Z}/18\mathbb{Z}$? **L**

(a) $\overline{59}$ (b) $\overline{-18}$ (c) $\overline{-157}$ (d) $\overline{91}$

1.3.10 **Rechnen mit Restklassen**. Wir fixieren $n \in \mathbb{N}$. Auf der Menge

$$\mathbb{Z}/n\mathbb{Z} = \{\overline{0}, \overline{1}, \ldots, \overline{n-1}\}$$

der Restklassen modulo n werden wir nun Rechenoperationen $+$ und $\cdot$ definieren.

Sind $\overline{k}, \overline{\ell} \in \mathbb{Z}/n\mathbb{Z}$, dann definieren wir

$$\overline{k} + \overline{\ell} := \overline{k + \ell}$$

und

$$\overline{k} \cdot \overline{\ell} := \overline{k\ell}.$$

An dieser Stelle muss man sehr vorsichtig sein. Hier verwenden wir zwei gewählte Vertreter k und ℓ. Das Ergebnis der Addition und Multiplikation soll aber nur von den Restklassen $\overline{k}$ und $\overline{\ell}$ und nicht von den gewählten Vertretern abhängen. Wir müssen nun zeigen, dass dies der Fall ist; man sagt dann, dass Addition und Multiplikation *wohldefiniert* sind.

Nehmen wir andere Vertreter $k' \in \overline{k}$ und $\ell' \in \overline{\ell}$. Nach Satz 1.3.6 (b) gibt es dann ganze Zahlen $a, b \in \mathbb{Z}$ mit $k' = k + an$ und $\ell' = \ell + bn$. Daraus folgt

$$k' + \ell' = k + \ell + (a + b)n$$

und aus Satz 1.3.6 (b) folgt $\overline{k' + \ell'} = \overline{k + \ell}$. Die Addition ist also wohldefiniert. Genauso erhalten wir

$$k'\ell' = (k + an)(\ell + bn) = k\ell + n(a\ell + kb + nab),$$

d.h., es gilt $\overline{k'\ell'} = \overline{k\ell}$ und die Multiplikation ist ebenfalls wohldefiniert.

1.3.11 **Satz.** *Sei $n \in \mathbb{N}$. Dann ist $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ mit den Rechenoperationen aus 1.3.10 ein kommutativer Ring. Das Nullelement ist $\overline{0}$, das Einselement ist $\overline{1}$.*

Man nennt $\mathbb{Z}/n\mathbb{Z}$ den Restklassenring *von $\mathbb{Z}$ modulo n.*

Beweis. Wir prüfen die Ringaxiome aus 1.2.1. Dabei verwenden wir häufig, dass $(\mathbb{Z}, +, \cdot)$ ein Ring ist. Für $n = 1$ ist $\mathbb{Z}/n\mathbb{Z}$ übrigens der Nullring; siehe 1.2.7.

(R1): Die Addition ist assoziativ, denn für alle $\overline{a}, \overline{b}, \overline{c} \in \mathbb{Z}/n\mathbb{Z}$ gilt

$$\begin{aligned}(\overline{a} + \overline{b}) + \overline{c} &= \overline{(a + b)} + \overline{c} \\ &= \overline{(a + b) + c} = \overline{a + (b + c)} \qquad (+ \text{ in } \mathbb{Z} \text{ ist assoziativ}) \\ &= \overline{a} + \overline{(b + c)} = \overline{a} + (\overline{b} + \overline{c}).\end{aligned}$$

Die Addition ist kommutativ, denn es gilt

$$\overline{a} + \overline{b} = \overline{a + b} = \overline{b + a} = \overline{b} + \overline{a}.$$

Die Restklasse $\overline{0}$ ist das neutrale Element der Addition, denn es gilt $\overline{0} + \overline{a} = \overline{0 + a} = \overline{a}$ und aufgrund der Kommutativität auch $\overline{a} + \overline{0} = \overline{a}$. Das additive Inverse zur Restklasse $\overline{a}$ ist die Restklasse $\overline{-a}$, denn es gilt

$$\overline{a} + \overline{-a} = \overline{a + (-a)} = \overline{0}$$

und wegen der Kommutativität auch $\overline{-a} + \overline{a} = \overline{0}$.

(R2): Die Multiplikation erfüllt das Assoziativgesetz, denn für alle $\overline{a}, \overline{b}, \overline{c} \in \mathbb{Z}/n\mathbb{Z}$ gilt

$$\begin{aligned}(\overline{a} \cdot \overline{b}) \cdot \overline{c} &= \overline{(ab)} \cdot \overline{c} \\ &= \overline{(ab)c} = \overline{a(bc)} && (\cdot \text{ in } \mathbb{Z} \text{ ist assoziativ}) \\ &= \overline{a} \cdot \overline{(bc)} = \overline{a} \cdot (\overline{b} \cdot \overline{c}).\end{aligned}$$

Wir bemerken außerdem, dass die Multiplikation kommutativ ist, denn

$$\overline{a} \cdot \overline{b} = \overline{ab} = \overline{ba} = \overline{b} \cdot \overline{a}$$

gilt für alle $\overline{a}, \overline{b} \in \mathbb{Z}/n\mathbb{Z}$.

(R3): Die Restklasse $\overline{1}$ ist das Einselement, denn es gilt $\overline{1} \cdot \overline{a} = \overline{1a} = \overline{a}$ und genauso $\overline{a} \cdot \overline{1} = \overline{a}$.

(R4): Wir prüfen die Distributivgesetze. Für alle $\overline{a}, \overline{x}, \overline{y} \in \mathbb{Z}/n\mathbb{Z}$ gilt

$$\begin{aligned}\overline{a} \cdot (\overline{x} + \overline{y}) &= \overline{a} \cdot \overline{(x+y)} \\ &= \overline{a(x+y)} = \overline{ax + ay} && (\text{Distributivgesetz in } \mathbb{Z}) \\ &= \overline{ax} + \overline{ay} = \overline{a} \cdot \overline{x} + \overline{a} \cdot \overline{y}.\end{aligned}$$

Da wir bereits wissen, dass die Multiplikation kommutativ ist, folgt daraus unmittelbar das zweite Distributivgesetz, denn

$$(\overline{x} + \overline{y}) \cdot \overline{a} = \overline{a} \cdot (\overline{x} + \overline{y}) = \overline{a} \cdot \overline{x} + \overline{a} \cdot \overline{y} = \overline{x} \cdot \overline{a} + \overline{y} \cdot \overline{a}.$$ □

1.3.12 **Aufgabe**. Berechnen Sie im Ring $\mathbb{Z}/22\mathbb{Z}$: **L**

(a) $\overline{13} \cdot \overline{4} - \overline{20}$ (b) $\overline{87}^{123456789}$

In vielen Teilen der Mathematik – auch in der linearen Algebra – spielen die Primzahlen, die Ihnen bestimmt schon einmal begegnet sind, eine wichtige Rolle.

1.3.13 **Definition** Eine natürliche Zahl $p > 1$ heißt *Primzahl*, wenn für alle $a \in \mathbb{N}$ mit $a \mid p$ entweder $a = 1$ oder $a = p$ gilt.

Eine Primzahl p hat also genau zwei natürliche Teiler: 1 und p. Die Annahme $p > 1$ stellt sicher, dass diese Teiler verschieden sind. Schon seit Euklid (etwa 300 v.Chr.) ist bekannt, dass es unendlich viele Primzahlen gibt. Die Folge der Primzahlen beginnt mit $2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, \ldots$

1.3.14 Satz von den endlichen Primkörpern. *Sei $n \in \mathbb{N}$. Der Restklassenring $\mathbb{Z}/n\mathbb{Z}$ ist genau dann ein Körper, wenn n eine Primzahl ist.*

Beweis. Angenommen n ist keine Primzahl. Ist $n = 1$, dann ist $\mathbb{Z}/n\mathbb{Z}$ der Nullring und somit kein Körper (vgl. 1.2.20). Sei also $n > 1$. Dann hat n einen Teiler $a \in \mathbb{N}$ mit $1 < a < n$, d.h., es gilt $n = ab$ für eine natürliche Zahl $1 < b < n$. Aus Satz 1.3.6 (c) schließen wir, dass in $\mathbb{Z}/n\mathbb{Z}$ sowohl $\overline{a} \neq \overline{0}$ als auch $\overline{b} \neq \overline{0}$ sind. Es gilt aber

$$\overline{a} \cdot \overline{b} = \overline{ab} = \overline{n} = \overline{0}.$$

Damit ist $\mathbb{Z}/n\mathbb{Z}$ *nicht* nullteilerfrei und kann wegen 1.2.22 kein Körper sein.

Nehmen wir nun an, dass $n = p$ eine Primzahl ist. Da $p > 1$ ist, gilt $\overline{0} \neq \overline{1}$ in $\mathbb{Z}/p\mathbb{Z}$ und damit ist $\overline{0}$ keine Einheit; siehe 1.2.14 (b). Es sei nun $a \in \{1, 2, \ldots, p-1\}$. Wir wollen zeigen, dass $\overline{a}$ in $\mathbb{Z}/p\mathbb{Z}$ eine Einheit ist. Dazu werden wir im nächsten Lemma 1.3.16 zeigen, dass es ganze Zahlen $x, y \in \mathbb{Z}$ mit $1 = xa + yp$ gibt. Mit $\overline{p} = \overline{0}$ schließen wir

$$\overline{1} = \overline{x} \cdot \overline{a} + \overline{y} \cdot \overline{p} = \overline{x} \cdot \overline{a}$$

und da $\mathbb{Z}/p\mathbb{Z}$ kommutativ ist, folgt auch $\overline{a} \cdot \overline{x} = \overline{x} \cdot \overline{a} = \overline{1}$. Also ist $\overline{a}$ eine Einheit in $\mathbb{Z}/p\mathbb{Z}$. □

1.3.15 Schreibweise. Ist p eine Primzahl, dann schreibt man oft $\mathbb{F}_p$ anstelle von $\mathbb{Z}/p\mathbb{Z}$ um zu betonen, dass es sich um einen Körper[4] handelt.

1.3.16 Lemma. *Es seien p eine Primzahl und $a \in \mathbb{Z}$. Ist p kein Teiler von a, dann gibt es $x, y \in \mathbb{Z}$ mit*

$$xa + yp = 1.$$

Beweis. Es sei c die kleinste natürliche Zahl, die sich in der Form $xa + yp$ mit $x, y \in \mathbb{Z}$ schreiben lässt. Wir wollen zeigen, dass $c = 1$ ist.

Teilen wir a mit Rest durch p, dann erhalten wir $q \in \mathbb{Z}$ und $r < p$ mit $a = qp + r$. Da p kein Teiler von a ist, geht die Division nicht auf und es gilt $r > 0$. Da wir $r = a - qp$ schreiben können, gilt $c \leq r < p$.

Wir zeigen nun, dass c ein Teiler von p ist. Da p eine Primzahl ist, folgt daraus wegen $c < p$ sofort $c = 1$. Dazu teilen wir p mit Rest durch c und erhalten ganze Zahlen $q' \in \mathbb{Z}$ und $0 \leq r' < c$ mit

$$p = q'c + r'.$$

[4] Der Buchstabe $\mathbb{F}$ leitet sich vom englischen Begriff „field“ ab.

Da $c = xa + yp$ minimal gewählt war und wir r' in der Form $r' = p - q'c = -q'xa + (1 - q'y)p$ schreiben können, kann wegen $r' < c$ nur $r' = 0$ sein, d.h. $c \mid p$. □

1.3.17 **Der erweiterte Euklidische Algorithmus**. Um in den endlichen Primkörpern $\mathbb{F}_p$ zu rechnen, muss man in der Lage sein das multiplikative Inverse einer Restklasse $\overline{a} \neq 0$ zu bestimmen. Ist $1 = xa + yp$ mit $x, y \in \mathbb{Z}$, dann haben wir im Beweis von Satz 1.3.14 gesehen, dass die Restklasse $\overline{x}$ das multiplikative Inverse zu $\overline{a}$ in $\mathbb{F}_p$ ist.

Aus dem Beweis von Lemma 1.3.16 kann man auch ein gutes Verfahren ablesen, mit dem man solche Zahlen $x, y \in \mathbb{Z}$ berechnen kann. Es sei $0 < r_1 < p$ der kleinste positive Vertreter der Restklasse $\overline{a}$. Das heißt, es ist

$$a = q_1 p + r_1$$

mit $0 < r_1 < p$. Falls bereits $0 < a < p$ erfüllt ist, dann setzt man also $r_1 = a$ und $q_1 = 0$. Dann teilt man p mit Rest durch r_1 und erhält

$$p = q_2 r_1 + r_2$$

mit $r_2 < r_1$. So fährt man fort und erhält eine Folge $r_1 > r_2 > r_3 > \dots$

$$\begin{aligned} r_1 &= q_3 r_2 + r_3 \\ r_2 &= q_4 r_3 + r_4 \\ &\vdots \\ r_{k-3} &= q_{k-1} r_{k-2} + r_{k-1} \\ r_{k-2} &= q_k r_{k-1} + r_k \end{aligned}$$

bis schließlich $r_k = 1$ gilt. Nun liest man diese Gleichungskette rückwärts und erhält

$$1 = r_{k-2} - q_k r_{k-1} = r_{k-2} - q_k (r_{k-3} - q_{k-1} r_{k-2}).$$

Dann ersetzt man nach und nach $r_{k-2}, r_{k-3}, \dots, r_1$ bis man einen Ausdruck der Gestalt $1 = xa + yp$ gefunden hat. Dieses Verfahren nenn man den *erweiterten Euklidischen Algorithmus.*

1.3.18 **Beispiel**. Wir betrachten den endlichen Körper $\mathbb{F}_{23}$ und möchten das multiplikative Inverse zu $\overline{9}$ bestimmen. Dazu verwenden wir den Euklidischen Algorithmus mit $p = 23$ und $a = 9$. Es gilt $0 < a < 23$, also setzen wir $r_1 = a$ und erhalten

$$23 = 2 \cdot 9 + 5$$

$$\begin{aligned} 9 &= 1 \cdot 5 + 4 \\ 5 &= 1 \cdot 4 + 1. \end{aligned}$$

Rückwärts Einsetzen liefert nun

$$1 = 5 - 4 = 5 - (9 - 5) = 2 \cdot 5 - 9 = 2 \cdot (23 - 2 \cdot 9) - 9 = -5 \cdot 9 + 2 \cdot 23.$$

Damit ist $\overline{-5} = \overline{18}$ das multiplikative Inverse zu $\overline{9}$. In der Tat, es gilt

$$\overline{9} \cdot \overline{18} = \overline{162} = \overline{1 + 7 \cdot 23} = \overline{1}.$$

1.3.19 **Aufgabe**. Bestimmen Sie das multiplikative Inverse von $\overline{17}$ im Körper $\mathbb{F}_{31}$. **L**

Es ist wichtig, das Rechnen in endlichen Körpern etwas zu üben. Mit den Methoden aus den „Mathematischen Grundlagen“ können Sie beispielsweise lineare Gleichungssysteme über endlichen Körpern lösen.

1.3.20 **Aufgabe**. Bestimmen Sie alle Lösungen des linearen Gleichungssystems **L**

$$\begin{array}{ccccccc} x_1 & + & \overline{2}x_2 & + & \overline{4}x_3 & = & \overline{1} \\ x_1 & & & + & \overline{3}x_3 & = & \overline{1} \\ \overline{2}x_1 & + & \overline{2}x_2 & + & \overline{3}x_3 & = & \overline{1} \end{array}$$

über dem endlichen Körper $\mathbb{F}_5$.

1.4. Polynomringe

Eines der wichtigsten Werkzeuge der Linearen Algebra ist die Theorie der Polynomringe in einer Unbestimmten über einem Körper. In diesem Abschnitt erarbeiten wir die Grundlagen dieser Theorie. Polynome haben Sie schon im Abschnitt 6 der Mathematischen Grundlagen kennengelernt.

I. Rechnen im Polynomring

Weil Polynome zentral für die Lineare Algebra sind, werden wir nochmal mit der Definition anfangen.

1.4.1 **Polynome**. Sei K ein Körper. Ein *Polynom* f über K ist eine Folge

$$f = (a_0, a_1, a_2, a_3, \dots)$$

von Elementen $a_i \in K$ so, dass nur endlich viele Folgenglieder ungleich 0 sind. Die Elemente a_i nennt man *Koeffizienten* des Polynoms. Das Polynom, dessen Koeffizienten alle 0 sind, nennt man das *Nullpolynom* und bezeichnet es mit 0.

Weil die Notation als Folge ziemlich unpraktisch ist, wollen wir für Polynome eine andere Notation verwenden und schreiben

$$f = \sum_i a_i X^i.$$

Dabei ist $X^0, X^1, X^2, X^3, \dots$ eine Folge von Symbolen und die Zahl a_i vor X^i ist genau das i-te Folgenglied. Die Folgenglieder mit $a_i = 0$ lassen wir in dieser Schreibweise weg. Man schreibt also $f = \sum_{i=0}^n a_i X^i$ um klarzumachen, dass alle Koeffizienten a_k mit $k > n$ verschwinden. Auch das Symbol X^0 schreibt man nicht aus, d.h. anstelle von $a_0 X^0$ schreibt man nur a_0.

Die Menge der Polynome über K bezeichnen wir mit $K[X]$. Das Symbol X nennt man die *Unbestimmte*. Grundsätzlich kann man die Unbestimmte natürlich auch mit einem anderen Symbol bezeichnen (z.B. Y oder T). Man nennt $K[X]$ den *Polynomring* (siehe 1.4.9).

1.4.2 **Beispiel**. Wir betrachten Polynome über dem Körper $\mathbb{Q}$.

(a) Das Polynom $f = (3, -4, 0, \frac{1}{2}, 1, 0, 0, \dots) \in \mathbb{Q}[X]$ schreiben wir als

$$f = X^4 + \frac{1}{2}X^3 - 4X + 3.$$

(b) Zum Polynom $g = 2X^5 - 3X^3 + \frac{7}{9} \in \mathbb{Q}[X]$ gehört die Koeffizientenfolge $(\frac{7}{9}, 0, 0, -3, 0, 2, 0, 0, \dots)$.

1.4.3 **Definition** Es sei $f = \sum_i a_i X^i \in K[X]$. Ist $f \neq 0$, dann ist der *Grad* von f definiert als

$$\mathrm{Grad}(f) = \max\{n \in \mathbb{N}_0 \mid a_n \neq 0\}.$$

Ist $n = \mathrm{Grad}(f)$, dann nennt man a_n den *Leitkoeffizienten* von f. Ein Polynom $f \neq 0$ mit Leitkoeffizient 1 heißt *normiert.*

Den Grad des Nullpolynoms legen wir als $\mathrm{Grad}(0) = -\infty$ fest. Die Polynome vom Grad ≤ 0 nennt man *konstante Polynome*.

1.4.4 **Beispiel**. Das Polynom $g = -4X^6 + X^3 + X + 2 \in \mathbb{Q}[X]$ hat Grad 6 und den Leitkoeffizienten -4.

1.4.5 **Rechnen mit Polynomen**. Auf der Menge $K[X]$ der Polynome definieren wir nun mehrere Rechenoperationen.

Die *Addition* $+\colon K[X] \times K[X] \to K[X]$ von Polynomen ist definiert durch

$$\sum_i a_i X^i + \sum_i b_i X^i \coloneqq \sum_i (a_i + b_i) X^i.$$

Die *Skalarmultiplikation* $K \times K[X] \to K[X]$ ist definiert durch

$$\lambda \sum_i a_i X^i \coloneqq \sum_i \lambda a_i X^i.$$

Die *Multiplikation* $\cdot\colon K[X] \times K[X] \to K[X]$ von Polynomen ist definiert durch

$$\Big(\sum_i a_i X^i\Big) \cdot \Big(\sum_i b_i X^i\Big) \coloneqq \sum_i \Big(\sum_{j=0}^{i} a_j b_{i-j}\Big) X^i.$$

Die Multiplikation ist so definiert, dass $X^i \cdot X^j = X^{i+j}$ gilt. Die Skalarmultiplikation mit λ entspricht der Multiplikation mit dem Polynom λX^0

1.4.6 **Aufgabe**. Berechnen Sie $(X^2 - 6) \cdot (X^2 + X + 1)$ im Polynomring $\mathbb{R}[X]$. L

1.4.7 **Lemma.** (Grad und Multiplikation) *Sei K ein Körper. Für Polynome $f, g \in K[X]$ gilt stets*

$$\mathrm{Grad}(f \cdot g) = \mathrm{Grad}(f) + \mathrm{Grad}(g).$$

Dabei gilt die Vereinbarung, dass die Addition mit $-\infty$ immer $-\infty$ ergibt.

Beweis. Ist $f = 0$ das Nullpolynom, dann gilt $f \cdot g = 0$ und damit $-\infty = -\infty + \operatorname{Grad}(g)$ nach Vereinbarung. Genauso gilt die Gleichung falls $g = 0$ ist.

Es seien nun $f, g \neq 0$ mit $m = \operatorname{Grad}(f)$ und $n = \operatorname{Grad}(g)$. Wir schreiben $f = \sum_{i=0}^{m} a_i X^i$ mit $a_m \neq 0$ und $g = \sum_{j=0}^{n} b_j X^j$ mit $b_n \neq 0$. Nach Definition ist der i-te Koeffizient von $f \cdot g$ genau

$$c_i = \sum_{j=0}^{i} a_j b_{i-j}$$

Ist $i > n + m$, so ist $c_i = 0$, denn für jedes j in der Summe ist entweder $j > m$ (und damit $a_j = 0$) oder $i - j > n$ (und damit $b_{i-j} = 0$). Für $i = n + m$ erhält man aber $c_{n+m} = a_m b_n$. Nach Annahme ist $a_m \neq 0$ und $b_n \neq 0$. Da K ein Körper und damit nullteilerfrei ist (siehe 1.2.22), folgt $c_{n+m} \neq 0$. Damit ist $n + m$ der Grad von $f \cdot g$. □

1.4.8 **Aufgabe.** Es seien $f, g \in K[X]$. Zeigen Sie: $\operatorname{Grad}(f+g) \leq \max(\operatorname{Grad}(f), \operatorname{Grad}(g))$. **L**

1.4.9 **Satz.** *Sei K ein Körper. Dann ist $(K[X], +, \cdot)$ ein kommutativer nullteilerfreier Ring. Die Einheiten in $K[X]$ sind genau die Polynome vom Grad 0.*

Beweis. Wir prüfen die Ringaxiome aus 1.2.1. Wir schreiben dazu $f = \sum_i a_i X^i$, $g = \sum_i b_i X^i$ und $h = \sum_i c_i X^i$.

(R1): Die Addition ist assoziativ, denn für alle $f, g, h \in K[X]$ gilt

$$\begin{aligned} (f+g)+h &= (\sum_i (a_i+b_i)X^i) + h = \sum_i ((a_i+b_i)+c_i)X^i \\ &= \sum_i (a_i + (b_i+c_i))X^i && \text{(Assoz. von + in } K) \\ &= f + \sum_i (b_i+c_i)X^i \\ &= f + (g+h) \end{aligned}$$

Weil die Addition in K kommutativ ist, gilt

$$f + g = \sum_i (a_i+b_i)X^i = \sum_i (b_i+a_i)X^i = g + f$$

und damit ist die Addition von Polynomen kommutativ.

Das Nullpolynom ist sicherlich ein neutrales Element bezüglich der Addition. Das additive Inverse zu f ist das Polynom $-f = (-1)f$, denn es gilt

$$f + (-1) \cdot f = \sum_i (a_i + (-a_i))X^i = 0.$$

(R2): Wir rechnen nach, dass die Multiplikation assoziativ ist. Es gilt

$$\begin{aligned}
(f\cdot g)\cdot h &= \left(\sum_i (\sum_{j=0}^{i} a_j b_{i-j}) X^i\right)\cdot h \\
&= \sum_k (\sum_{i=0}^{k} (\sum_{j=0}^{i} a_j b_{i-j}) c_{k-i}) X^k \\
&= \sum_k (\sum_{j=0}^{k} \sum_{i=j}^{k} a_j b_{i-j} c_{k-i}) X^k \\
&= \sum_k (\sum_{j=0}^{k} a_j (\sum_{i'=0}^{k-j} b_{i'} c_{k-j-i'})) X^k && \text{(Setze } i' = i-j) \\
&= f\cdot \sum_\ell (\sum_{i'=0}^{\ell} b_{i'} c_{\ell-i'}) X^\ell = f\cdot(g\cdot h)
\end{aligned}$$

für alle f, g, h.

Die Multiplikation ist auch kommutativ, denn es ist

$$\begin{aligned}
f\cdot g &= \sum_i (\sum_{j=0}^{i} a_j b_{i-j}) X^i \\
&= \sum_i (\sum_{j'=0}^{i} b_{j'} a_{i-j'}) X^i = g\cdot f. && \text{(Setze } j' = i-j)
\end{aligned}$$

(R3): Das Polynom $1 := X^0$ ist ein neutrales Element der Multiplikation, denn es gilt $1\cdot f = f = f\cdot 1$.

(R4): Übungsaufgabe! **L**

Wir wissen nun, dass $K[X]$ ein kommutativer Ring ist. Warum ist der Polynomring nullteilerfrei? Es seien $f \neq 0$ und $g \neq 0$ in $K[X]$. Es gilt also $\mathrm{Grad}(f) \geq 0$ und $\mathrm{Grad}(g) \geq 0$. Dann folgt aus Lemma 1.4.7, dass

$$\mathrm{Grad}(f\cdot g) = \mathrm{Grad}(f) + \mathrm{Grad}(g) \geq 0$$

gilt. Also ist $f\cdot g \neq 0$.

Schließlich zeigen wir, dass die Einheiten in $K[X]$ genau die Polynome vom Grad 0 sind. Die Polynome vom Grad 0 sind von der Form $a = aX^0$ für $a \in K$ mit $a \neq 0$.

In diesem Fall ist a eine Einheit in K und es gilt $(aX^0) \cdot (a^{-1}X^0) = (aa^{-1}X^0) = 1$. Also ist aX^0 eine Einheit in $K[X]$.

Sei umgekehrt f eine Einheit in $K[X]$, dann gibt es ein Polynom $g \in K[X]$ mit $f \cdot g = 1$. Aus Lemma 1.4.7 folgt dann

$$0 = \mathrm{Grad}(1) = \mathrm{Grad}(f) + \mathrm{Grad}(g)$$

und damit muss $\mathrm{Grad}(f) = 0$ sein. □

1.4.10 **K als Unterring von $K[X]$.** Die Menge aller konstanten Polynome in $K[X]$ ist ein Unterring von $K[X]$ (siehe 1.2.8). Um einzusehen, dass die Unterringaxiome erfüllt sind, halten wir folgende Beobachtungen fest: (UR1) Die Addition konstanter Polynome liefert wieder ein konstantes Polynom und auch das additive Inverse $-a$ eines konstanten Polynoms a ist wieder ein konstantes Polynom. (UR2) Das Einselement in $K[X]$ ist das konstante Polynom 1. (UR3) Die Multiplikation konstanter Polynome ergibt ein konstantes Polynom (siehe 1.4.7).

Den Unterring der konstanten Polynome in $K[X]$ kann man problemlos mit dem Körper K identifizieren. Jedem Element $a \in K$ entspricht dabei das konstante Polynom $a = aX^0$ in $K[X]$.

II. Division im Polynomring

Wir werden nun sehen, dass es im Polynomring eine Division mit Rest gibt. Das Ergebnis ist sehr ähnlich zur Division im Ring der ganzen Zahlen (vgl. 1.3.2). Dies ist kein Zufall, der Ring der ganzen Zahlen und der Polynomring über einem Körper gehören zur Klasse der „Euklidischen Ringe“, die immer eine Division mit Rest ermöglichen.

1.4.11 **Satz zur Polynomdivision mit Rest.** *Sei K ein Körper und es seien Polynome $f, g \in K[X]$ mit $f \neq 0$ gegeben. Dann gibt es eindeutige Polynome $h, r \in K[X]$ mit $\mathrm{Grad}(r) < \mathrm{Grad}(f)$, sodass*

$$g = hf + r$$

gilt. Man nennt h den Quotienten *und r den* Rest *der Division von g durch f.*

Beweis. Im ersten Schritt zeigen wir, dass h und r mit diesen Eigenschaften existieren. Wir wählen zunächst das Polynom $h \in K[X]$ dergestalt, dass $\mathrm{Grad}(g - hf)$ so klein wie möglich ist. Wir setzen $r = g - hf$; es gilt also $g = hf + r$.

Wir behaupten, dass $\mathrm{Grad}(r) < \mathrm{Grad}(f)$ ist. Für einen Widerspruch nehmen wir $\mathrm{Grad}(r) \geq \mathrm{Grad}(f)$ an. Es sei $r = a_n X^n + a_{n-1}X^{n-1} + \cdots + a_0$, wobei $n = \mathrm{Grad}(r)$ und a_n der Leitkoeffizient von r ist. Sei $f = b_m X^m + b_{m-1}X^{m-1} + \cdots + b_0$, wobei $m = \mathrm{Grad}(f)$ und b_m der Leitkoeffizient von f ist. Dann ist

$$\frac{a_n}{b_m} X^{n-m} f = a_n X^n + \frac{b_{m-1}a_n}{b_m} X^{n-1} + \cdots + \frac{b_0 a_n}{b_m} X^{n-m}$$

und damit haben r und $\frac{a_n}{b_m} X^{n-m} f$ denselben Grad (nämlich n) und denselben Leitkoeffizienten a_n. Bei der Subtraktion hebt sich der Term $a_n X^n$ weg, d.h.,

$$\mathrm{Grad}(r - \frac{a_n}{b_m} X^{n-m} f) < \mathrm{Grad}(r).$$

Wir setzen $h' = h + \frac{a_n}{b_m} X^{n-m}$, dann erhalten wir

$$g - h'f = g - hf - \frac{a_n}{b_m} X^{n-m} f = r - \frac{a_n}{b_m} X^{n-m} f.$$

Damit ist $\mathrm{Grad}(g - hf) > \mathrm{Grad}(g - h'f)$. Dies ist ein Widerspruch, denn wir haben h so gewählt, dass der Grad von $g - hf$ minimal ist.

Warum sind r und h mit diesen Eigenschaften eindeutig? Angenommen es gilt

$$g = hf + r = h'f + r'$$

mit $\mathrm{Grad}(r) < \mathrm{Grad}(f)$ und $\mathrm{Grad}(r') < \mathrm{Grad}(f)$. Dann gilt

$$(h - h')f = r' - r.$$

Aus Lemma 1.4.7 schließen wir

$$\mathrm{Grad}(h - h') + \mathrm{Grad}(f) = \mathrm{Grad}(r' - r) < \mathrm{Grad}(f)$$

und damit muss $\mathrm{Grad}(h - h') = -\infty$, also $h = h'$, sein. Dann folgt aber auch $r' - r = 0$ und somit $r = r'$. □

1.4.12 **Polynomdivision**. Im Beweis versteckt sich auch ein Verfahren um die Polynomdivision durchzuführen und den Quotienten und Rest zu berechnen. Es ist das naheliegende Verfahren, dass Sie vielleicht schon aus der Schule kennen.

Zum Einstieg setzen wir $h = 0$. Es seien $g = \sum_{i=0}^{n} a_i X^i$ und $f = \sum_{j=0}^{m} b_j X^j$ gegeben. Falls $n = \mathrm{Grad}(g) \geq m = \mathrm{Grad}(f)$ ist, dann ersetzen wir g durch

$$g - \frac{a_n}{b_m} X^{n-m} f$$

und addieren $\frac{a_n}{b_m} X^{n-m}$ zu h. Dadurch sinkt der Grad von g. Irgendwann gilt $\mathrm{Grad}(g) < \mathrm{Grad}(f)$. Dann ist h der Quotient und der Rest der Division ist g.

1.4.13 **Beispiel**. Gegeben sind $f = X^2 - 2 \in \mathbb{Q}[X]$ und $g = X^5 - X^3 + 2X^2 - 1 \in \mathbb{Q}[X]$. Wir führen die Polynomdivision durch:

$$
\begin{array}{rrrrr}
(X^5 & -X^3 & +2X^2 & & -1) : (X^2 - 2) = X^3 + X + 2 \\
-(X^5 & -2X^3) & & & \\
\hline
& X^3 & +2X^2 & & -1 \\
& -(X^3 & & -2X) & \\
\hline
& & 2X^2 & +2X & -1 \\
& & -(2X^2 & & -4) \\
\hline
& & & 2X & +3
\end{array}
$$

Es gilt also

$$g = (X^3 + X + 2)f + (2X + 3);$$

dabei ist $X^3 + X + 2$ der Quotient und $2X + 3$ der Rest der Division.

1.4.14 **Aufgabe**. Bestimmen Sie den Quotienten und der Rest der Division von $g = X^6 + X^4 + X + \overline{1} \in \mathbb{F}_2[X]$ durch $f = X^2 + X + \overline{1} \in \mathbb{F}_2[X]$. **L**

Ist der Rest der Division von g durch f gleich 0, dann sagt man, dass f ein Teiler von g ist.

1.4.15 **Definition** Sei K ein Körper und seien $f, g \in K[X]$. Man nennt f einen *Teiler* von g, wenn $g = hf$ für ein Polynom $h \in K[X]$ gilt. Wir schreiben, dann $f \mid g$ und sagen f teilt g.

III. Nullstellen

1.4.16 **Einsetzen in Polynome: Polynomfunktionen**. Sei K ein Körper und sei $f = \sum_{i=0} a_i X^i \in K[X]$ ein Polynom. Ist $x \in K$, dann kann man x *in* f *einsetzen*, d.h., man ersetzt X im Polynom durch x und wertet den Ausdruck in K aus:

$$f(x) = \sum_i a_i x^i \in K.$$

Da fast alle Koeffizienten $a_i = 0$ sind, handelt es sich dabei um eine endliche Summe. Die so definierte Funktion $K \to K$ mit $x \mapsto f(x)$ nennt man die zu f gehörige *Polynomfunktion*.

1.4.17 **Polynome vs. Polynomfunktionen**. Es ist wichtig Polynomfunktionen und Polynome sorgfältig auseinander zu halten. Das Polynom $f \in K[X]$ kann im Allgemeinen *nicht* eindeutig durch die Polynomfunktion $x \mapsto f(x)$ auf K bestimmt werden. Beispielsweise ist die Polynomfunktion zum Polynom $f = X^2 + X \in \mathbb{F}_2[X]$ die Nullfunktion, denn es gilt $f(\overline{0}) = \overline{0}$ und $f(\overline{1}) = \overline{0}$. Zum Polynom $X^2+X \in \mathbb{F}_2[X]$ und zum Nullpolynom gehört also dieselbe Polynomfunktion, obwohl die Polynome verschieden sind.

Auch mit dem Begriff „konstantes Polynom“ muss man vorsichtig sein. Die zu einem konstanten Polynom gehörige Polynomfunktion ist konstant. Doch auch ein nicht-konstantes Polynom kann eine konstante Polynomfunktion definieren.

1.4.18 **Aufgabe**. Zeigen Sie, dass die Polynome $f = X \in \mathbb{F}_3[X]$ und $g = X^4 - X^2 + X \in \mathbb{F}_3[X]$ dieselbe Polynomfunktion definieren. **L**

Das Einsetzen in Polynome ist mit den Rechenoperationen im Polynomring verträglich.

1.4.19 **Lemma.** *Sei K ein Körper. Für alle $x \in K$ und alle $f, g \in K[X]$ gelten*

$$(f+g)(x) = f(x) + g(x) \quad \textit{und} \quad (f \cdot g)(x) = f(x)g(x).$$

Beweis. Wir schreiben $f = \sum_i a_i X^i$ und $g = \sum_i b_i X^i$ mit Koeffizienten $a_i, b_i \in K$. Dann erhält man

$$\begin{aligned}(f+g)(x) &= \sum_i (a_i + b_i)x^i = \sum_i a_i x^i + b_i x^i \\ &= \sum_i a_i x^i + \sum_i b_i x^i = f(x) + g(x)\end{aligned}$$

und

$$\begin{aligned}(f \cdot g)(x) &= \sum_i \sum_{j=0}^{i} (a_j b_{i-j}) x^i = \sum_i \sum_{j=0}^{i} (a_j x^j)(b_{i-j} x^{i-j}) \\ &= \sum_{i',j} (a_j x^j)(b_{i'} x^{i'}) && \text{(setze } i' = i - j) \\ &= \Big(\sum_i a_i x^i\Big)\Big(\sum_i b_i x^i\Big) = f(x)g(x). && \square\end{aligned}$$

1.4.20 $\boxed{\textbf{Definition}}$ Ein Element $a \in K$ mit der Eigenschaft $f(a) = 0$ nennt man *Nullstelle* des Polynoms.

Ein normiertes Polynom von Grad 1 hat die Form $X - a \in K[X]$ mit $a \in K$. Diese Polynome nennt man auch *Linearfaktoren*. Der folgende nützliche Satz erlaubt es für jede Nullstelle eines Polynoms einen Linearfaktor „abzuspalten“.

1.4.21 **Abspaltungssatz.** *Es sei $f \in K[X]$ ein Polynom. Das Element $a \in K$ ist genau dann eine Nullstelle von f, wenn $X - a$ ein Teiler von f ist.*

Beweis. Nehmen wir zunächst an, dass $X - a$ ein Teiler von f ist. Dann gibt es ein Polynom $h \in K[X]$ mit $f = (X - a)h$. Wegen Lemma 1.4.19 erhalten wir

$$f(a) = (a - a)h(a) = 0$$

und a ist eine Nullstelle von f.

Umgekehrt nehmen wir nun an, dass a eine Nullstelle von f ist, d.h., $f(a) = 0$. Wir verwenden die Division mit Rest 1.4.11 und teilen f durch $X - a$. Wir erhalten Polynome $h, r \in K[X]$ mit $\mathrm{Grad}(r) < 1$ und

$$f = (X - a)h + r.$$

Mit Lemma 1.4.19 folgt

$$0 = f(a) = (a - a)h(a) + r(a) = r(a).$$

Da $\mathrm{Grad}(r) < 1$ ist, d.h., r ist ein konstantes Polynom, muss r das Nullpolynom sein. Wir schließen $f = (X - a)h$ und damit ist $X - a$ ein Teiler von f. □

Daraus folgt nun insbesondere, dass ein Polynom $f \neq 0$ immer nur endlich viele Nullstellen besitzt.

1.4.22 **Korollar.** *Sei $f \in K[X]$ ein Polynom von Grad $n \geq 0$. Dann hat f höchstens n Nullstellen in K.*

Beweis. Wir führen den Beweis durch vollständige Induktion nach n. Beim Induktionsanfang $n = 0$ ist f ein konstantes Polynom, aber nicht das Nullpolynom (denn das hat Grad $-\infty$). Also hat f keine Nullstelle.

Für den Induktionsschritt sei $n \geq 1$. Wir nehmen an, dass die Aussage für alle Polynome vom Grad $n - 1$ richtig ist. Falls f keine Nullstelle besitzt, ist nichts weiter zu tun. Wir können also annehmen, dass f eine Nullstelle $a \in K$ besitzt. Aus dem Abspaltungssatz 1.4.21 folgt, dass f in der Form

$$f = (X - a)h$$

für ein Polynom $h \in K[X]$ geschrieben werden kann. Wir bemerken, dass h den Grad $n-1$ hat; siehe 1.4.7. Ist $b \neq a$ eine weitere Nullstelle von f, dann gilt mit Lemma 1.4.19

$$0 = f(b) = (b-a)h(b).$$

Da K nullteilerfrei ist (siehe 1.2.22) und $b - a \neq 0$ ist, schließen wir $h(b) = 0$. Alle Nullstellen $b \neq a$ sind also Nullstellen von h. Nach Induktionsvoraussetzung hat h höchstens $n-1$ Nullstellen und damit hat f höchstens n verschiedene Nullstellen. □

1.4.23 **Beispiel**. Ein Polynom kann auch gar keine Nullstellen haben. Zum Beispiel hat das reelle Polynom $X^2 + 1 \in \mathbb{R}[X]$ keine Nullstelle. In der Tat, in $\mathbb{R}$ erfüllt jedes Quadrat $x^2 \geq 0$. Also ist $x^2 + 1 \geq 1 > 0$ für alle reellen Zahlen $x \in \mathbb{R}$.

1.4.24 **Definition** Sei $f \in K[X]$ ein Polynom ungleich dem Nullpolynom. Sei $a \in K$. Die größte Zahl $m \in \mathbb{N}_0$, sodass $(X-a)^m$ ein Teiler von f ist, nennt man die *Vielfachheit* von $X - a$ in f. Ist a eine Nullstelle von f, dann nennt man m auch die Vielfachheit der Nullstelle a in f.

1.4.25 **Vielfachheit 0?** Für jedes Element $a \in K$ gilt $(X-a)^0 = 1$ und dies ist immer ein Teiler von f. Ist 0 die Vielfachheit von $X - a$ in f, dann ist $(X - a)$ kein Teiler von f und wegen 1.4.21 ist a dann auch keine Nullstelle von f.

In anderen Worten: Die Vielfachheit von $(X - a)$ in f ist genau dann positiv, wenn a eine Nullstelle von f ist.

1.4.26 **Lemma.** *Sei $f \neq 0$ in $K[X]$ und sei $a \in K$. Die Zahl $m \in \mathbb{N}_0$ ist genau dann die Vielfachheit von $(X - a)$ in f, wenn sich f in der Form*

$$f = (X-a)^m h$$

mit $h(a) \neq 0$ schreiben lässt.

Beweis. Ist m die Vielfachheit von $X - a$ in f, dann lässt sich f in der Form

$$f = (X-a)^m h$$

schreiben, denn $(X-a)^m$ teilt f. Wir zeigen $h(a) \neq 0$. Nehmen wir für einen Widerspruch an, dass $h(a) = 0$ ist. Mit dem Abspaltungssatz schreiben wir h in der Form $h = (X-a)h'$ und erhalten $(X-a)^{m+1}h' = f$. Damit ist $(X-a)^{m+1}$ ein Teiler von f und wir erhalten einen Widerspruch zur Maximalität von m.

Nehmen wir umgekehrt an, dass sich f in der Form

$$f = (X-a)^n h$$

mit $h(a) \neq 0$ schreiben lässt. Sei m die Vielfachheit von $X-a$ in f; wir können f also in der Form $f = (X-a)^m g$ schreiben. Aus der Definition der Vielfachheit folgt $m \geq n$. Es gilt also

$$(X-a)^n h = f = (X-a)^m g = (X-a)^n (X-a)^{m-n} g.$$

Da der Polynomring nullteilerfrei ist, folgt aus der Kürzungsregel 1.2.23 nun

$$h = (X-a)^{m-n} g.$$

Da h aber keine Nullstelle bei a hat, muss $m-n=0$ sein, d.h., n ist die Vielfachheit von $X-a$ in f. □

1.4.27 **Bemerkung**. Die Vielfachheit m einer Nullstelle a von f erfüllt immer $m \leq \mathrm{Grad}(f)$, denn aus $(X-a)^m h = f$ folgt mit 1.4.7 auch $m + \mathrm{Grad}(h) = \mathrm{Grad}(f)$.

1.4.28 **Beispiel**. Wir betrachten das Polynom $f = X^3 - 4X^2 + 4X \in \mathbb{Q}[X]$. Dann ist 2 eine Nullstelle, denn $f(2) = 8 - 16 + 8 = 0$.

Wir wollen die Vielfachheit der Nullstelle 2 bestimmen. Dazu spalten wir $(X-2)$ mithilfe der Polynomdivision ab. Es gilt

$$f = (X-2) f_1$$

mit $f_1 = X^2 - 2X$. Wir prüfen, ob 2 auch eine Nullstelle von f_1 ist. Es gilt: $f_1(2) = 4 - 4 = 0$. Wir spalten wieder den Faktor $(X-2)$ aus $X^2 - 2X$ ab und erhalten

$$f = (X-2)^2 f_2$$

mit $f_2 = X$. Weil das Polynom f_2 keine Nullstelle bei 2 hat, schließen wir mit Lemma 1.4.26, das 2 die Vielfachheit der Nullstelle 2 im Polynom f ist.

1.4.29 **Aufgabe**. Bestimmen Sie die Vielfachheit der Nullstelle 1 im Polynom $f = X^6 - 2X^5 + X^4 - X^2 + 2X - 1 \in \mathbb{R}[X]$. **L**

1.4.30 **Satz über die Nullstellen.** *Sei K ein Körper und sei $f \in K[X]$ ein Polynom mit $f \neq 0$. Dann lässt sich f in der Form*

$$f = (X-a_1)^{m_1}(X-a_2)^{m_2} \cdots (X-a_r)^{m_r} h \qquad (1.4.a)$$

mit paarweise verschiedenen $a_1, \ldots, a_r$ und einem nullstellenfreien Polynom h darstellen.

Diese Darstellung ist bis auf die Reihenfolge der Linearfaktoren eindeutig. Dabei sind $a_1, \ldots, a_r \in K$ stets die paarweise verschiedenen Nullstellen von f und m_i ist die Vielfachheit der Nullstelle a_i in f.

Insbesondere ist die Summe über die Vielfachheiten aller Nullstellen durch den Grad des Polynoms beschränkt.

Beweis. Angenommen, wir haben f in der Form (1.4.a) geschrieben. Dann sind $a_1, \ldots, a_r$ Nullstellen von f. Da h keine Nullstellen hat, kann f keine weiteren Nullstellen besitzen.

Wir zeigen nun, dass m_i die Vielfachheit von a_i ist. Daraus folgt unmittelbar, dass diese Darstellung von f bis auf Reihenfolge der Faktoren eindeutig ist. Es ist

$$f = (X - a_i)^{m_i} \underbrace{h \cdot \prod_{j \neq i} (X - a_j)^{m_j}}_{w}$$

und dabei gilt $w(a_i) = h(a_i) \prod_{j \neq i} (a_i - a_j)^{m_j} \neq 0$, weil h keine Nullstellen besitzt und für $i \neq j$ auch $a_i \neq a_j$ gilt. Daraus folgt mit Lemma 1.4.26, dass m_i die Vielfachheit von a_i in f ist. Insbesondere gilt dann $\mathrm{Grad}(f) = \mathrm{Grad}(h) + \sum_{i=1}^r m_i \geq \sum_{i=1}^r m_i$.

Wir beweisen die Existenz der Produktschreibweise (1.4.a) durch vollständige Induktion über die Anzahl r der Nullstellen von f. Ist $r = 0$, so hat f keine Nullstelle und wir setzen $f = h$.

Sei nun $r > 0$. Als Induktionsannahme setzen wir voraus, dass sich jedes Polynom mit weniger als r Nullstellen in dieser Form schreiben lässt. Es sei a_1 eine Nullstelle von f mit der Vielfachheit m_1. Wir wenden Lemma 1.4.26 an; wir schreiben

$$f = (X - a_1)^{m_1} g$$

für ein Polynom $g \in K[X]$ mit $g(a_1) \neq 0$. Jede Nullstelle von g ist eine Nullstelle von f. Umgekehrt muss jede Nullstelle $b \neq a_1$ von f eine Nullstelle von g sein, denn

$$0 = f(b) = (b - a_1)^{m_1} g(b)$$

und $a_1 - b \neq 0$. Das Polynom g hat also $(r-1)$ verschiedene Nullstellen, nennen wir diese $a_2, \ldots, a_r$. Aus der Induktionsvoraussetzung folgt, dass wir g in der Form

$$g = (X - a_2)^{m_2} \cdots (X - a_r)^{m_r} h$$

mit einem Polynom h ohne Nullstellen schreiben können. Wir erhalten

$$f = (X - a_1)^{m_1} g = (X - a_1)^{m_1} (X - a_2)^{m_2} \cdots (X - a_r)^{m_r} h. \qquad \square$$

Die folgende Definition ist für die Lineare Algebra besonders wichtig.

1.4.31 **Definition** Sei $f \in K[X]$ ein normiertes Polynom mit $\mathrm{Grad}(f) \geq 1$. Man sagt f *zerfällt in Linearfaktoren*, wenn sich f als Produkt von Linearfaktoren schreiben lässt.

Aus dem Satz über die Nullstellen schließen wir nun:

1.4.32 **Korollar.** *Sei $f \in K[X]$ ein normiertes, nicht-konstantes Polynom und seien $a_1, \dots, a_r \in K$ die paarweise verschiedenen Nullstellen von f. Falls f in Linearfaktoren zerfällt, gilt*

$$f = (X - a_1)^{m_1} \cdots (X - a_r)^{m_r},$$

wobei m_i die Vielfachheit der Nullstelle a_i in f ist.

Bis auf die Reihenfolge der Linearfaktoren ist dies die einzige Möglichkeit f als Produkt von Linearfaktoren zu schreiben.

1.4.33 **Beispiel**. (a) Das Polynom $X^2 + X - 20 \in \mathbb{R}[X]$ zerfällt in Linearfaktoren, denn es gilt

$$X^2 + X - 20 = (X + 5)(X - 4).$$

(b) Das Polynom $X^2 + 1 \in \mathbb{R}[X]$ zerfällt nicht in Linearfaktoren, denn es hat keine Nullstellen; siehe 1.4.23.

1.4.34 **Wie prüft man, ob ein Polynom in Linearfaktoren zerfällt?** Um zu prüfen, ob ein normiertes Polynom f in Linearfaktoren zerfällt, muss man Nullstellen von f suchen und mit Polynomdivision abspalten. Diesen Vorgang wiederholt man solange bis man entweder f als Produkt von Linearfaktoren geschrieben hat oder auf ein Polynom stößt, das keine Nullstellen besitzt.

Dieses Verfahren hat aber einen Haken: Es kann sehr schwierig sein die Nullstellen eines Polynoms zu finden. Eine Lösungsformel wie für quadratische Gleichungen gibt es im Allgemeinen nicht.

Auf Verfahren zum Finden von Nullstellen werden wir hier nicht eingehen. Bei Polynomen vom Grad ≥ 3 bleibt uns hier nur die Möglichkeit, Nullstellen durch Ausprobieren zu finden. Zumindest in endlichen Körpern mit wenigen Elementen funktioniert diese Methode gut, weil wir die Möglichkeit haben, alle Elemente in das Polynom einzusetzen.

1.4.35 **Beispiel**. Wir betrachten das Polynom $f = X^5 - X^4 - X + \overline{1} \in \mathbb{F}_3[X]$ und wollen prüfen, ob das Polynom in Linearfaktoren zerfällt.

Dazu suchen wir eine Nullstelle und probieren $a = \overline{1}$. Es gilt $f(\overline{1}) = \overline{1} - \overline{1} - \overline{1} + \overline{1} = \overline{0}$. Wir spalten die Nullstelle ab und erhalten

$$f = (X - \overline{1})(X^4 - \overline{1}).$$

Nun suchen wir eine Nullstelle von $f_1 = X^4 - \overline{1}$. Wir probieren $a = \overline{1}$. Es gilt $f_1(\overline{1}) = \overline{1} - \overline{1} = \overline{0}$. Wir spalten die Nullstelle ab und erhalten

$$X^4 - \overline{1} = (X - \overline{1})(X^3 + X^2 + X + \overline{1}).$$

Wir definieren $f_2 = X^3 + X^2 + X + \overline{1}$ und suchen wieder eine Nullstelle. Wir probieren $a = \overline{1}$. Es gilt $f_2(\overline{1}) = \overline{1} + \overline{1} + \overline{1} + \overline{1} = \overline{1} \neq \overline{0}$. Dann probieren wir $a = \overline{2}$ und erhalten

$$f_2(\overline{2}) = \overline{8} + \overline{4} + \overline{2} + \overline{1} = \overline{15} = \overline{0}.$$

Wir spalten $(X - \overline{2})$ von f_2 ab und erhalten

$$f_2 = (X - \overline{2})(X^2 + \overline{1}).$$

Wir setzen $f_3 = X^2 + \overline{1}$ und suchen nach Nullstellen. Wir probieren alle Elemente aus $\mathbb{F}_3$ aus und finden

$$f_3(\overline{0}) = \overline{1}, \quad f_3(\overline{1}) = \overline{2}, \quad f_3(\overline{2}) = \overline{2}.$$

Das Polynom f_3 hat also keine Nullstellen. Damit gilt

$$f = (X - \overline{1})^2(X - \overline{2})(X^2 + \overline{1})$$

und wir stellen fest, dass f nicht in Linearfaktoren zerfällt.

1.4.36 **Aufgabe**. Zerfällt das Polynom $f = X^4 + X^2 \in \mathbb{F}_2[X]$ in Linearfaktoren? **L**

Zerfällt ein Polynom in Linearfaktoren, dann ist es sehr einfach alle normierten Teiler des Polynoms zu beschreiben.

1.4.37 **Lemma.** *Die normierten Teiler von $f = (X - a_1)^{m_1} \cdots (X - a_r)^{m_r}$ sind genau die Polynome der Form*

$$g = (X - a_1)^{k_1} \cdots (X - a_r)^{k_r}$$

mit $0 \leq k_i \leq m_i$ für alle $i \in \{1, 2, \dots, r\}$.

Beweis. Ist $g = (X - a_1)^{k_1} \cdots (X - a_r)^{k_r}$ mit $0 \le k_i \le m_i$ gegeben, dann ist g ein Teiler von f, denn $f = g \cdot (X - a_1)^{m_1 - k_1} \cdots (X - a_r)^{m_r - k_r}$.

Sei umgekehrt g ein normiertes Polynom, das f teilt, d.h., $f = gh$. Jede Nullstelle von g (bzw. h) ist auch eine Nullstelle von f, also liegen die Nullstellen von g (bzw. h) in $\{a_1, \ldots, a_r\}$. Sei k_i die Vielfachheit der Nullstelle a_i in g und sei k_i' die Vielfachheit der Nullstelle a_i in h. Wir schreiben nun $g = (X - a_1)^{k_1} \cdots (X - a_r)^{k_r} \cdot g'$ und $h = (X - a_1)^{k_1'} \cdots (X - a_r)^{k_r'} \cdot h'$, wobei g' und h' wie in Satz 1.4.30 nullstellenfrei sind. Dann besitzt auch $g'h'$ keine Nullstellen und es gilt

$$f = (X - a_1)^{k_1 + k_1'} \cdots (X - a_r)^{k_r + k_r'} g'h'.$$

Die Eindeutigkeitsaussage im Satz über die Nullstellen zeigt nun, dass $k_i + k_i' = m_i$ für alle i gilt und, dass $g'h' = 1$ ist. Also sind g', h' Polynome vom Grad 0; siehe 1.4.9. Da g normiert ist, muss $g' = 1$ sein und damit gilt $g = (X - a_1)^{k_1} \cdots (X - a_r)^{k_r}$. □

1.4.38 **Beispiel**. Das reelle Polynom $(X-1)(X-5)^2 \in \mathbb{R}[X]$ hat also genau sechs normierte Teiler: 1, $(X - 1)$, $(X - 5)$, $(X - 5)^2$, $(X - 1)(X - 5)$ und $(X - 1)(X - 5)^2$.

Insbesondere liefert uns Lemma 1.4.37 folgende nützliche Beobachtung.

1.4.39 **Korollar.** *Sei $f \in K[X]$ ein normiertes Polynom von positivem Grad, das in Linearfaktoren zerfällt. Dann zerfällt auch jeder normierte nicht-konstante Teiler von f in Linearfaktoren.*

Jetzt kommen wir zu einem sehr wichtigen Begriff der Algebra, den wir hier aber nur sehr oberflächlich diskutieren können. Sie können im Modul „Algebra“ mehr über algebraisch abgeschlossene Körper lernen.

1.4.40 **Definition** Ein Körper K heißt *algebraisch abgeschlossen*, wenn jedes nicht-konstante Polynom $f \in K[X]$ eine Nullstelle in K besitzt.

1.4.41 **Beispiel**. (a) Der Körper der reellen Zahlen $\mathbb{R}$ ist *nicht* algebraisch abgeschlossen. Das Polynom $X^2 + 1 \in \mathbb{R}[X]$ hat keine Nullstellen; siehe 1.4.23.

(b) Der Körper $\mathbb{F}_2$ ist nicht algebraisch abgeschlossen. Das Polynom $f = X^2 + X + \overline{1} \in \mathbb{F}_2[X]$ hat keine Nullstellen, denn es gilt

$$f(\overline{0}) = \overline{1} \quad \text{und} \quad f(\overline{1}) = \overline{1}.$$

1.4.42 **Anmerkung zur Definition**. *Keiner* der Körper, die wir bisher kennengelernt haben, ist algebraisch abgeschlossen. Es in der Tat anspruchsvoll, die Existenz solcher Körper zu zeigen. Im Abschnitt 1.5 werden wir aber ein konkretes Beispiel kennenlernen: den Körper der komplexen Zahlen.

1.4.43 **Aufgabe**. Zeigen Sie, dass die Körper $\mathbb{F}_3$ und $\mathbb{F}_5$ nicht algebraisch abgeschlossen sind. **L**

1.4.44 **Satz.** *Sei K ein Körper. Folgende Aussagen sind äquivalent:*

(i) K ist algebraisch abgeschlossen.

(ii) Jedes normierte, nicht-konstante Polynom $f \in K[X]$ zerfällt in Linearfaktoren.

Beweis. (i) $\implies$ (ii): Nehmen wir zunächst an, dass K algebraisch abgeschlossen ist. Sei f ein normiertes, nicht-konstantes Polynom. Es seien $a_1, \dots, a_r \in K$ die paarweise verschiedenen Nullstellen von f. Mit dem Satz über die Nullstellen 1.4.30 schreiben wir f in der Form

$$f = (X - a_1)^{m_1}(X - a_2)^{m_2} \cdots (X - a_r)^{m_r} h$$

mit einem Polynom $h \in K[X]$ ohne Nullstellen. Da K algebraisch abgeschlossen ist, muss h konstant sein, d.h., $h = \lambda \in K^\times$. Da f normiert ist, muss $\lambda = 1$ sein. Damit zerfällt f in Linearfaktoren.

(ii) $\implies$ (i): Sei $f \in K[X]$ ein beliebiges nicht-konstantes Polynom mit Leitkoeffizient $c \in K^\times$. Dann ist $c^{-1}f$ ein normiertes, nicht-konstantes Polynom. Nach Voraussetzung zerfällt $c^{-1}f$ in Linearfaktoren:

$$c^{-1}f = (X - a_1)^{m_1}(X - a_2)^{m_2} \cdots (X - a_r)^{m_r};$$

insbesondere hat $c^{-1}f$ (und damit auch f) mindestens eine Nullstelle. □

IV. Ideale

In diesem Abschnitt lernen wir einen Begriff aus der Ringtheorie - einem Teilgebiet der Algebra - kennen, der später im Lehrtext nützlich wird: den Begriff des *Ideals*. Ideale sind zentrale Objekte, wenn man Ringe und ihre Eigenschaften verstehen möchte. Hier wird die allgemeine Theorie aber nicht benötigt und wir werden Ideale nur für Polynomringe einführen. Wer mehr über Ringe und Ideale wissen möchte, wird in fast jedem Lehrbuch zur Algebra fündig (z.B. in [Fis17]).

1.4.45 **Definition** Sei K ein Körper. Ein *Ideal* in $K[X]$ ist eine nicht-leere Teilmenge $I \subseteq K[X]$ mit folgenden Eigenschaften:

(I1) für alle $f, g \in I$ ist auch $f + g \in I$,

(I2) für alle $f \in I$ und $h \in K[X]$ ist auch $hf \in I$.

1.4.46 **Anmerkungen zur Definition**. (a) Ein Ideal $I \subseteq K[X]$ ist auch unter der Skalarmultiplikation abgeschlossen. Das heißt, für alle $\lambda \in K$ und $f \in I$ gilt $\lambda f \in I$, denn die Skalarmultiplikation mit λ entspricht der Multiplikation mit dem konstanten Polynom λ.

(b) Ein Ideal $I \subseteq K[X]$ ist stets eine Untergruppe der additiven Gruppe $(K[X], +)$. Nach Definition ist I nicht leer und aus (I1) folgt 1.1.18 (UG2). Wegen Lemma 1.2.5 gilt $-f = (-1) \cdot f$. Liegt also $f \in I$, dann liegt wegen (I2) auch das additive Inverse $-f$ in I. Die Bedingung 1.1.18 (UG3) ist also auch erfüllt.

1.4.47 **Beispiele**. (a) $K[X]$ ist offensichtlich ein Ideal in $K[X]$.

(b) $I = \{0\}$ ist ein Ideal in $K[X]$, denn $0 + 0 = 0$ und $h \cdot 0 = 0$ für alle $h \in K[X]$.

(c) Sei $f \in K[X]$. Wir definieren

$$fK[X] := \{fh \in K[X] \mid h \in K[X]\}.$$

Dann ist $fK[X]$ ein Ideal in $K[X]$. Man nennt $fK[X]$ das *von f erzeugte Hauptideal*.

1.4.48 **Aufgabe**. Es sei $f \in K[X]$. Beweisen Sie, dass $fK[X]$ ein Ideal ist. **L**

Wir werden nun zeigen, dass wir damit schon alle Ideale kennen.

1.4.49 **Satz.** *Sei K ein Körper und $I \subseteq K[X]$ sei ein Ideal $I \neq \{0\}$. Dann gibt es ein eindeutiges normiertes Polynom $f \in K[X]$ mit $I = fK[X]$.*

Das Polynom f nennen wir den normierten Erzeuger *von I.*

Beweis. Da wir annehmen, dass $I \neq \{0\}$ ist, enthält I ein Polynom $\neq 0$. Es sei n der kleinste Grad von allen Polynomen $\neq 0$ in I. Wir wählen ein Polynom $f' \in I$ mit $\mathrm{Grad}(f') = n$. Ist a_n der Leitkoeffizient von f', dann liegt auch das normierte Polynom

$$f = a_n^{-1} f'$$

in I und es gilt $\mathrm{Grad}(f) = \mathrm{Grad}(f') = n$.

Wir zeigen, dass $I = fK[X]$ gilt. Zuerst stellen wir fest, dass wegen (I2) alle Elemente der Form fh in I liegen. Es gilt also $fK[X] \subseteq I$.

Für die umgekehrte Inklusion sei $g \in I$ beliebig gegeben. Die Division mit Rest 1.4.11 von g durch f liefert Polynome $h, r \in K[X]$ mit $\mathrm{Grad}(r) < n$ und

$$g = hf + r.$$

Da f und g im Ideal I liegen, liegen auch hf und $r = g - hf$ im Ideal I. Da n der kleinste Grad $\neq -\infty$ unter allen Elementen von I ist und $\mathrm{Grad}(r) < n$ gilt, schließen wir $\mathrm{Grad}(r) = -\infty$ und damit $r = 0$. Es gilt also $g = hf \in fK[X]$. Da $g \in I$ beliebig war, folgt $I \subseteq fK[X]$.

Warum ist f eindeutig? Angenommen $\tilde{f}$ ist ein weiteres normiertes Polynom mit $\tilde{f}K[X] = I$. Wegen $f \in I$ gilt $f = h\tilde{f}$ für ein $h \in K[X]$. Aus der Gradformel 1.4.7 folgt

$$\mathrm{Grad}(f) = \mathrm{Grad}(h) + \mathrm{Grad}(\tilde{f}).$$

Da $f \neq 0$ ist, sind sowohl h als auch $\tilde{f}$ ungleich 0 und damit ist $\mathrm{Grad}(\tilde{f}) \leq \mathrm{Grad}(f) = n$. Da n minimal gewählt war, folgt $\mathrm{Grad}(f) = \mathrm{Grad}(\tilde{f}) = n$ und $\mathrm{Grad}(h) = 0$. Damit ist h ein Polynom der Form $\lambda \in K^\times$. Das Polynom $\tilde{f}$ ist normiert (d.h., der Leitkoeffizient ist 1) und damit ist der Leitkoeffizient von $f = \lambda f'$ genau λ. Da aber auch f normiert ist, folgt $\lambda = 1$ und damit $f = \tilde{f}$. □

1.4.50 **Lemma.** *Sei $I \subseteq K[X]$ ein Ideal. Es gilt $I = K[X]$, genau dann, wenn I ein Polynom vom Grad 0 enthält.*

Beweis. Ist $I = K[X]$, dann enthält I das Polynom 1 vom Grad 0. Sei umgekehrt $a \in K^\times$ in I. Für alle $f \in K[X]$ gilt

$$f = (a^{-1}f) \cdot a \in I$$

und damit ist $I = K[X]$. □

1.4.51 **Summe von Idealen.** Sei K ein Körper und seien $I_1, I_2 \subseteq K[X]$ zwei Ideale. Dann definieren wir

$$I_1 + I_2 := \{f_1 + f_2 \in K[X] \mid f_1 \in I_1, f_2 \in I_2\}.$$

Dies ist ein Ideal in $K[X]$. Um das zu sehen prüft man die Eigenschaften (I1) und (I2) aus 1.4.45 nach.

Zu (I1): Seien $f, g \in I_1 + I_2$. Wir schreiben $f = f_1 + f_2$ und $g = g_1 + g_2$ mit $f_i, g_i \in I_i$ (für $i \in \{1, 2\}$). Dann gilt

$$f + g = f_1 + f_2 + g_1 + g_2 = \underbrace{(f_1 + g_1)}_{\in I_1} + \underbrace{(f_2 + g_2)}_{\in I_2} \in I_1 + I_2,$$

weil I_1 und I_2 die Eigenschaft (I1) haben.

Zu (I2): Sei $f = f_1 + f_2 \in I_1 + I_2$ und sei $h \in K[X]$. Dann gilt

$$hf = \underbrace{hf_1}_{\in I_1} + \underbrace{hf_2}_{\in I_2} \in I_1 + I_2,$$

weil I_1 und I_2 die Eigenschaft (I2) erfüllen.

1.4.52 **Aufgabe**. Es seien $I_1, I_2, \dots I_n$ Ideale in $K[X]$. Wir setzen **L**

$$\sum_{i=1}^{n} I_i := \Big\{\sum_{i=1}^{n} f_i \mid f_i \in I_i\Big\}.$$

Zeigen Sie, dass $\sum_{i=1}^{n} I_i$ ein Ideal ist.

1.4.53 **Definition** Es seien $f, g \in K[X]$ zwei Polynome, die nicht beide das Nullpolynom sind. Den normierten Erzeuger des Ideals $fK[X] + gK[X]$ nennt man den *größten gemeinsamen Teiler* von f und g. Wir notieren den größten gemeinsamen Teiler von f und g mit $\mathrm{ggT}(f, g)$.

Zur Vollständigkeit definieren wir auch $\mathrm{ggT}(0, 0) = 0$ und bemerken, dass 0 das Ideal $\{0\} + \{0\} = \{0\}$ erzeugt.

Der folgende Satz erklärt, weshalb die Bezeichnung größter gemeinsamer Teiler gerechtfertigt ist.

1.4.54 **Satz.** *Es seien $f_1, f_2 \in K[X]$ zwei Polynome. Dann gelten folgende Aussagen:*

(a) $\mathrm{ggT}(f_1, f_2)$ ist ein gemeinsamer Teiler von f_1 und f_2.

(b) Ist h ein gemeinsamer Teiler von f_1 und f_2, dann ist h auch ein Teiler von $\mathrm{ggT}(f_1, f_2)$

(c) Es gibt Polynome $g_1, g_2 \in K[X]$ mit $\mathrm{ggT}(f_1, f_2) = f_1g_1 + f_2g_2$.

Beweis. Zu (a): Es gilt $f_1 = f_1 + 0 \cdot f_2 \in f_1K[X] + f_2K[X]$. Da $\mathrm{ggT}(f_1, f_2)$ dieses Ideal erzeugt, gibt es ein Polynom $g \in K[X]$ mit $f_1 = g \cdot \mathrm{ggT}(f_1, f_2)$, d.h., $\mathrm{ggT}(f_1, f_2) \mid f_1$. Genauso zeigt man die Aussage für f_2.

Zu (c): Dies folgt aus der Definition, denn $\mathrm{ggT}(f_1, f_2)$ liegt im Ideal $f_1K[X]+f_2K[X]$ und kann damit in der Form

$$\mathrm{ggT}(f_1, f_2) = f_1g_1 + f_2g_2$$

mit $g_1, g_2 \in K[X]$ geschrieben werden.

Zu (b): Übungsaufgabe! □ **L**

1.4.55 $\boxed{\textbf{Definition}}$ Sei K ein Körper. Zwei Polynome $f, g \in K[X]$ heißen *teilerfremd*, wenn $\mathrm{ggT}(f, g) = 1$ gilt.

Da 1 der normierte Erzeuger des Ideals $K[X]$ ist, schließen wir: f und g sind genau dann teilerfremd, wenn $fK[X] + gK[X] = K[X]$ ist.

1.4.56 **Beispiel**. Sei K ein Körper und seien $a \neq b$ in K. Die Linearfaktoren $X - a$ und $X - b$ sind teilerfremd, denn es gilt

$$(X - a) - (X - b) = b - a \in K^{\times}.$$

Das Ideal $(X - a)K[X] + (X - b)K[X]$ enthält also ein Polynom vom Grad 0 und aus Lemma 1.4.50 folgt $(X - a)K[X] + (X - b)K[X] = K[X]$.

1.4.57 **Aufgabe**. Seien $f, g_1, g_2 \in K[X]$. Wir nehmen an, dass f und g_i für alle $i \in \{1, 2\}$ teilerfremd sind. Zeigen Sie: f und $g_1 g_2$ sind teilerfremd. **L**

1.4.58 **Aufgabe**. Seien $f, g \in K[X]$ und $n, m \in \mathbb{N}$. Wir nehmen an, dass f und g teilerfremd sind. Zeigen Sie, dass auch f^n und g^m teilerfremd sind. **L**

1.5. Der Körper der komplexen Zahlen

Im Laufe Ihrer Schulausbildung und Ihres Studiums haben Sie gelernt mit immer neuen Arten von Zahlen zu rechnen: Angefangen von den natürlichen Zahlen $\mathbb{N}$, über die ganzen Zahlen $\mathbb{Z}$ und rationalen Zahlen $\mathbb{Q}$ bis zu den reellen Zahlen $\mathbb{R}$.

Der jeweils neue, größere Zahlbereich behebt dabei ein Problem, dass beim Rechnen in den zuvor gebräuchlichen Zahlen auftritt. In den natürlichen Zahlen $\mathbb{N}$ kann man nicht unbeschwert subtrahieren, also fügt man die negativen Zahlen dazu und geht damit zu den ganzen Zahlen $\mathbb{Z}$ über, wo die Subtraktion immer definiert ist. In den ganzen Zahlen geht nicht jede Division auf, also geht man zu den rationalen Zahlen über, wo das Teilen unproblematisch ist (solange man nicht durch 0 teilt). Von den rationalen Zahlen gelangt man zu den reellen Zahlen, weil man feststellt, dass der Zahlenstrahl sonst zu viele „Löcher“ enthält. In anderen Worten: die Grenzwertbildung in $\mathbb{Q}$ ist schwierig, weil auch Cauchy-Folgen (also Folgen, die so aussehen, als ob sie konvergieren) im Allgemeinen keinen Grenzwert haben. Also geht man zu den reellen Zahlen über, wo Grenzwerte von Cauchy-Folgen immer existieren.

In diesem Abschnitt werden wir die reellen Zahlen $\mathbb{R}$ zu den *komplexen Zahlen* $\mathbb{C}$ erweitern und dabei wieder ein Problem beheben, das Sie vom Rechnen in den reellen Zahlen kennen. Betrachten wir dazu die Gleichung

$$x^2 + 1 = 0.$$

In den reellen Zahlen hat diese polynomielle Gleichung keine Lösung; siehe 1.4.23. Die gesuchte Lösung wäre eine Quadratwurzel aus -1. Die komplexen Zahlen entstehen aus den reellen Zahlen, indem man eine Wurzel aus -1 zu den reellen Zahlen „hinzufügt“. Die formale, saubere Konstruktion werden wir im Abschnitt I durchführen. Wir werden sehen, dass die komplexen Zahlen wieder einen Körper bilden. Das Bemerkenswerte ist nun, dass dieser Körper algebraisch abgeschlossen ist, d.h., jedes nicht-konstante Polynom hat eine Nullstelle!

I. Konstruktion der komplexen Zahlen

1.5.1 **Konstruktion der komplexen Zahlen**. Wir definieren zunächst die Menge

$$\mathbb{C} = \mathbb{R} \times \mathbb{R} = \{(x, y) \mid x, y \in \mathbb{R}\}$$

von Paaren reeller Zahlen. Darauf definieren wir zwei Verknüpfungen: eine Addition $+$ und eine Multiplikation $\cdot$. Die Addition in $\mathbb{C}$ ist definiert als

$$(x, y) + (x', y') := (x + x', y + y')$$

und die Multiplikation in $\mathbb{C}$ ist definiert als

$$(x, y) \cdot (x', y') := (xx' - yy', xy' + yx').$$

1.5.2 Satz. $(\mathbb{C}, +, \cdot)$ *ist ein Körper.*

Beweis. Wir zeigen zunächst, dass $(\mathbb{C}, +, \cdot)$ ein kommutativer Ring ist.

Zu (R1): Wir wissen aus Aufgabe 1.1.9, dass $\mathbb{R} \times \mathbb{R}$ mit der komponentenweisen Addition eine abelsche Gruppe ist. Das Nullelement ist $(0, 0)$.

Zu (R2): Wir verifizieren das Assoziativgesetz der Multiplikation. Für alle (x, y), (x', y'), (x'', y'') in $\mathbb{C}$ gilt

$$\begin{aligned}
\Big((x, y) \cdot (x', y')\Big) \cdot (x'', y'') &= (xx' - yy', xy' + yx') \cdot (x'', y'') \\
&= ((xx' - yy')x'' - (xy' + yx')y'', (xx' - yy')y'' + (xy' + yx')x'') \\
&= (xx'x'' - yy'x'' - xy'y'' - yx'y'', xx'y'' - yy'y'' + xy'x'' + yx'x'') \\
&= (x(x'x'' - y'y'') - y(y'x'' + x'y''), x(x'y'' + y'x'') + y(x'x'' - y'y'')) \\
&= (x, y) \cdot (x'x'' - y'y'', x'y'' + y'x'') \\
&= (x, y) \cdot \Big((x', y') \cdot (x'', y'')\Big).
\end{aligned}$$

Zu (R3): Das Element $(1, 0) \in \mathbb{C}$ ist neutral bzgl. der Multiplikation, denn für alle $(x, y) \in \mathbb{C}$ gilt

$$(x, y) \cdot (1, 0) = (x, y) = (1, 0) \cdot (x, y).$$

Wir zeigen, dass die Multiplikation auch kommutativ ist. Seien $(x, y), (x', y') \in \mathbb{C}$, dann gilt

$$(x, y) \cdot (x', y') = (xx' - yy', xy' + yx') = (x'x - y'y, x'y + y'x) = (x', y') \cdot (x, y)$$

Zu (R4): Wir zeigen, dass die Distributivgesetze gelten. Für $(a, b), (x, y), (x', y') \in \mathbb{C}$ gilt

$$\begin{aligned}
(a, b) \cdot \Big((x, y) + (x', y')\Big) &= (a, b) \cdot (x + x', y + y') \\
&= (a(x + x') - b(y + y'), a(y + y') + b(x + x')) \\
&= (ax + ax' - by - by', ay + ay' + bx + bx') \\
&= (ax - by, ay + bx) + (ax' - by', ay' + bx')
\end{aligned}$$

$$= (a,b) \cdot (x,y) + (a,b) \cdot (x',y').$$

Da wir bereits wissen, dass die Multiplikation kommutativ ist, ist das andere Distributivgesetz ebenfalls erfüllt.

Also ist $(\mathbb{C}, +, \cdot)$ ein kommutativer Ring. Damit $\mathbb{C}$ ein Körper ist, müssen wir noch nachweisen, dass alle Elemente außer dem Nullelement $(0,0)$ Einheiten sind. Es sei $(x,y) \in \mathbb{C}$, dabei soll entweder x oder y ungleich 0 sein. Insbesondere ist das Element $x^2 + y^2 \in \mathbb{R}$ ungleich Null. Wir behaupten, dass $(\frac{x}{x^2+y^2}, -\frac{y}{x^2+y^2})$ das Inverse zu (x,y) ist. Wir rechnen nach:

$$(x,y) \cdot (\frac{x}{x^2+y^2}, -\frac{y}{x^2+y^2}) = (\frac{x^2}{x^2+y^2} + \frac{y^2}{x^2+y^2}, \frac{xy}{x^2+y^2} - \frac{xy}{x^2+y^2}) = (1,0).$$

Da die Multiplikation kommutativ ist, gilt auch

$$(\frac{x}{x^2+y^2}, -\frac{y}{x^2+y^2}) \cdot (x,y) = (1,0)$$

Wir haben bereits gesehen, dass $(1,0)$ das Einselement in $\mathbb{C}$ ist, damit ist (x,y) eine Einheit. □

1.5.3 **Definition** Der Körper $\mathbb{C}$ heißt *Körper der komplexen Zahlen.*

1.5.4 **Wieso enthalten die komplexen Zahlen die reellen Zahlen**? Man kann sich überlegen, dass die Teilmenge

$$R = \{(x,0) \in \mathbb{C}\} \subseteq \mathbb{C}$$

ein Unterkörper von $\mathbb{C}$ ist. Sehen wir uns einmal die Addition und Multiplikation in dieser Teilmenge genauer an. Für alle $(x,0), (x',0) \in \mathbb{C}$ gelten

$$(x,0) + (x',0) = (x+x',0)$$
$$(x,0) \cdot (x',0) = (xx',0).$$

In der zweiten Komponente passiert also gar nichts und in der ersten Komponente sehen wir genau die Addition und Multiplikation von reellen Zahlen. Wir können also die Menge R mit den reellen Zahlen identifizieren und schreiben daher einfach a anstelle von $(a,0)$.

Das passt gut zusammen, denn damit gilt

$$a \cdot (x,y) = (a,0) \cdot (x,y) = (ax, ay).$$

1.5.5 **Die imaginäre Einheit**. Das Element $i := (0,1)$ nennt man die *imaginäre Einheit*. In den komplexen Zahlen gilt nun

$$i^2 = (0,1) \cdot (0,1) = (-1,0) = -1.$$

Die imaginäre Einheit i ist also eine Quadratwurzel von -1.

1.5.6 **Notation für komplexe Zahlen**. Mit den gerade festgelegten Schreibweisen erhalten wir

$$(x,y) = (x,0) + (0,y) = x + yi.$$

Wir werden von nun an komplexe Zahlen nur noch in der Form $x + yi$ schreiben, weil dies zum Rechnen praktischer ist als die Schreibweise als Paare. Auch den Punkt $\cdot$ schreiben wir für die Multiplikation nun nicht mehr aus.

1.5.7 **Definition** Ist $z = x + yi \in \mathbb{C}$ eine komplexe Zahl, so nennt man

$$\mathrm{Re}(z) = x$$

den *Realteil* von z und

$$\mathrm{Im}(z) = y$$

den *Imaginärteil* von z.

1.5.8 **Aufgabe**. Zeigen Sie: Für alle $z \in \mathbb{C}$ gilt $\mathrm{Im}(z) = \mathrm{Re}(-iz)$. **L**

1.5.9 **Aufgabe**. Es seien $z, w \in \mathbb{C}$ und $a \in \mathbb{R}$. Zeigen Sie: **L**

$$\mathrm{Re}(az + w) = a\mathrm{Re}(z) + \mathrm{Re}(w).$$

1.5.10 **Die komplexe Zahlenebene**. Während man sich die reellen Zahlen als Punkte auf einer Geraden vorstellen kann, kann man sich die komplexen Zahlen als Punkte einer Ebene denken. Dabei trägt man den Realteil auf der x-Achse und den Imaginärteil auf der y-Achse ab. Die x-Achse nennt man die *reelle Gerade*.

1.5.11 **Definition** Sei $z = x + yi \in \mathbb{C}$ eine komplexe Zahl mit $x, y \in \mathbb{R}$. Man definiert

$$|z| := \sqrt{x^2 + y^2} \in \mathbb{R}.$$

Die reelle Zahl $|z|$ nennt man den *Betrag* von z.

Es gilt $|z|^2 = x^2 + y^2$, also ist der Betrag durch den Satz von Pythagoras genau der Abstand des Punktes z vom Ursprung in der komplexen Zahlenebene. Insbesondere gilt $|z| = 0$ genau dann, wenn $z = 0$ ist.

1.5.12 **Aufgabe**. Es sei $z \in \mathbb{C}$. Zeigen Sie: $\mathrm{Re}(z) \leq |z|$. **L**

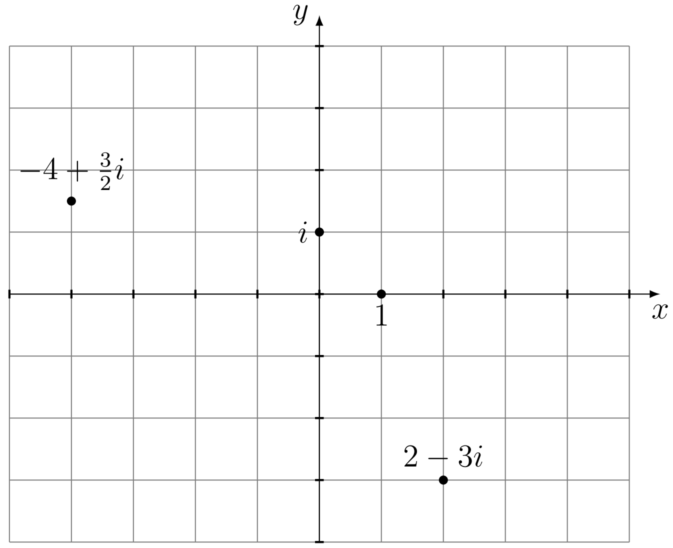


Abbildung 1.1.: Die komplexe Zahlenebene

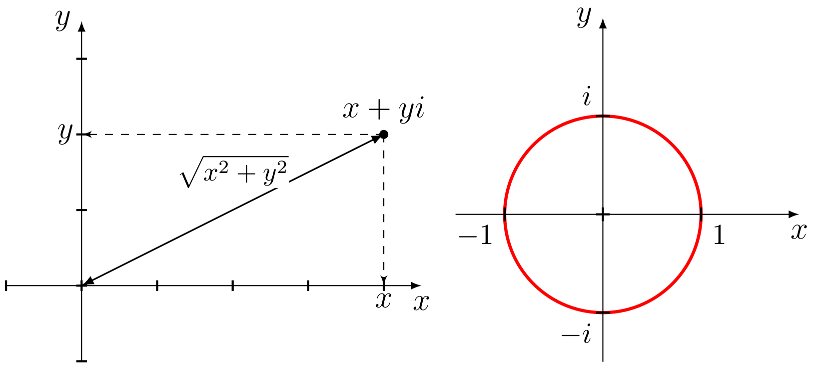


Abbildung 1.2.: Links: Betrag einer komplexen Zahl | Rechts: Einheitskreis

1.5.13 $\boxed{\textbf{Definition}}$ Die Menge

$$S^1 = \{z \in \mathbb{C} \mid \ |z| = 1\}$$

der komplexen Zahlen mit Betrag 1 nennt man den *Einheitskreis*.

Die Zahlen 1, -1, i und $-i$ liegen auf dem Einheitskreis.

1.5.14 **Aufgabe**. Geben Sie eine komplexe Zahl z mit $\mathrm{Re}(z) \neq 0$ und $\mathrm{Im}(z) \neq 0$ an, die auf dem Einheitskreis liegt. **L**

II. Rechnen mit komplexen Zahlen

Das Rechnen mit komplexen Zahlen der Form $x + yi$ ist eigentlich ganz einfach. Man muss sich nur etwas daran gewöhnen.

1.5.15 **Beispiel** (Addition komplexer Zahlen). Bei der Addition komplexer Zahlen muss man jeweils nur die Real- und Imaginärteile addieren. Einige Beispiele:

- $(5 + 3i) + (1 - 4i) = 6 - i$
- $(\sqrt{2} - 8i) + (-\sqrt{2} + 3i) = -5i$

Dasselbe gilt für die Subtraktion komplexer Zahlen.

1.5.16 **Beispiel** (Multiplikation komplexer Zahlen). Bei der Multiplikation komplexer Zahlen kann man wie gewohnt ausmultiplizieren und muss sich nur merken, dass $i^2 = -1$ gilt. Einige Beispiele:

- $(5 + 3i)(-2 + 4i) = -10 - 6i + 20i + 12i^2 = -22 + 14i$
- $(1 - i)(-2 + 8i) = -2 + 2i + 8i - 8i^2 = 6 + 10i$

1.5.17 **Aufgabe**. Berechnen Sie $(4 + 2i) - (1 - i)$ und $(2 - 5i)(3 + i)$. **L**

Beim Rechnen mit komplexen Zahlen ist die folgende Rechenoperation hilfreich.

1.5.18 **Definition** Sei $z = x + yi$ eine komplexe Zahl mit $x, y \in \mathbb{R}$. Man nennt

$$\overline{z} = x - yi$$

die zu z *komplex-konjugierte Zahl*.

(Anschaulich: Die Abbildung $z \mapsto \overline{z}$ ist die Spiegelung an der reellen Geraden in der komplexen Zahlenebene.)

Das folgende Lemma zeigt, dass die komplexe Konjugation sehr gut mit den Rechenoperationen in den komplexen Zahlen verträglich ist.

1.5.19 **Lemma.** (Rechenregeln für die komplexe Konjugation) *Für alle komplexen Zahlen $z, w \in \mathbb{C}$ gelten folgende Aussagen:*

(a) $\overline{z + w} = \overline{z} + \overline{w}$

(b) $\overline{z \cdot w} = \overline{z} \cdot \overline{w}$

(c) $z\,\overline{z} = |z|^2$

(d) $\overline{\overline{z}} = z$

(e) *$z \in \mathbb{R}$ genau dann, wenn $z = \overline{z}$ gilt.*

Beweis. Es sei $z = x + yi$ und $w = u + vi$ mit $x, y, u, v \in \mathbb{R}$.

Zu (a): Es gilt

$$\overline{z + w} = \overline{(x + u) + (y + v)i} = (x + u) - (y + v)i = x - yi + u - vi = \overline{z} + \overline{w}.$$

Zu (b): Es gilt

$$\overline{z \cdot w} = \overline{(xu - yv) + (xv + yu)i} = (xu - yv) - (xv + yu)i = (x - yi) \cdot (u - vi) = \overline{z} \cdot \overline{w}.$$

Zu (c): Es gilt

$$z\,\overline{z} = (x + yi)(x - yi) = x^2 - y^2 i^2 = x^2 + y^2 = |z|^2.$$

Zu (d): Es gilt

$$\overline{\overline{z}} = \overline{x - yi} = x + yi = z.$$

Zu (e): Es gilt $z = \overline{z}$ genau dann, wenn $x + yi = x - yi$ ist. Durch Umformen ist dies genau dann der Fall, wenn $y = 0$ ist; d.h., genau dann, wenn $z \in \mathbb{R}$ ist. □

Durch Umstellen von (c) erhält man einen Ausdruck für das multiplikative Inverse einer komplexen Zahl, den man sich gut merken kann.

1.5.20 **Korollar.** *Sei $z \neq 0$ eine komplexe Zahl. Dann gilt*

$$z^{-1} = \frac{1}{|z|^2}\overline{z}.$$

1.5.21 **Aufgabe**. Schreiben Sie die komplexen Zahlen $(1+i)^{-1}$ und $(2+3i)^{-1}$ in der Form $a + bi$ mit $a, b \in \mathbb{R}$. **L**

1.5.22 **Beispiel** (Division von komplexen Zahlen). Sind $z, w \in \mathbb{C}$ komplexe Zahlen mit $w \neq 0$, so ist $\frac{z}{w} = zw^{-1}$ wieder eine komplexe Zahl. Aber wie kann man diese in der Form $a + bi$ mit $a, b \in \mathbb{R}$ schreiben? Ein nützlicher Trick ist es, den Bruch mit dem komplex Konjugierten $\overline{w}$ des Nenners zu erweitern:

$$\frac{z}{w} = \frac{z \cdot \overline{w}}{w \cdot \overline{w}} = \frac{z \cdot \overline{w}}{|w|^2}.$$

Hier sind zwei Beispiele:

(a) $$\frac{2+i}{1-3i} = \frac{(2+i)(1+3i)}{1^2+3^2} = \frac{-1+7i}{10} = -\frac{1}{10} + \frac{7}{10}i$$

(b) $$\frac{6-2i}{1+i} = \frac{(6-2i)(1-i)}{1^2+1^2} = \frac{4-8i}{2} = 2 - 4i$$

1.5.23 **Aufgabe**. Schreiben Sie die komplexen Zahlen $\frac{11-2i}{2+i}$ und $\frac{13i}{-3+2i}$ in der Form $a+bi$ mit $a, b \in \mathbb{R}$. **L**

Es ist wichtig, das Rechnen mit komplexen Zahlen zu üben. Mit diesen einfachen Rechenregeln kann man schon sehr gut arbeiten. Zum Beispiel sind wir damit in der Lage, lineare Gleichungssysteme über den komplexen Zahlen zu lösen (mit den Methoden, die Sie aus den Mathematischen Grundlagen kennen).

1.5.24 **Aufgabe**. Lösen Sie das folgende lineare Gleichungssystem in den komplexen Unbekannten $x_1, x_2, x_3 \in \mathbb{C}$: **L**

$$\begin{array}{rcrcrcl} x_1 & & & + & ix_3 & = & 1 \\ & & (1+i)x_2 & + & 2x_3 & = & 1+3i \\ 2x_1 & + & 2x_2 & - & ix_3 & = & 6+2i \end{array}$$

III. Der Fundamentalsatz der Algebra

Die komplexen Zahlen gehen wahrscheinlich auf Cardano[5] zurück, der Ausdrücke mit Wurzeln aus negativen Zahlen verwendete, um Gleichungen zu lösen. Im 17. und 18. Jahrhundert kamen verschiedene Mathematiker auf die Idee, dass möglicherweise *alle* Polynomgleichungen eine komplexe Lösung besitzen; also, dass der Körper der komplexen Zahlen *algebraisch abgeschlossen* im Sinne von 1.4.40 ist. Das wurde nach Vorarbeiten von d'Alembert[6] dann erstmals 1799 von C. F. Gauß[7] bewiesen.

1.5.25 **Fundamentalsatz der Algebra.** *Der Körper der komplexen Zahlen $\mathbb{C}$ ist algebraisch abgeschlossen.*

Insbesondere folgt aus dem Fundamentalsatz und Satz 1.4.44:

1.5.26 **Korollar.** *Jedes normierte, nicht-konstante Polynom in $\mathbb{C}[X]$ zerfällt in Linearfaktoren.*

Der Beweis des Fundamentalsatzes ist nicht ganz einfach und seltsamerweise basieren die meisten bekannten Beweise auf Methoden der Analysis, der Funktionentheorie oder der Topologie und haben mit Algebra ziemlich wenig zu tun. Wir werden in diesem Studienbrief daher keinen Beweis des Fundamentalsatzes geben.

[5] Gerolamo CARDANO: italienischer Mathematiker, 1501–1576.
[6] Jean-Baptiste le Rond D'ALEMBERT: französischer Mathematiker, 1717–1783.
[7] Carl Friedrich GAUSS: deutscher Mathematiker, 1777–1855.

Stattdessen wollen wir uns speziell Polynome in $\mathbb{C}[X]$ vom Grad 2 ansehen und zeigen, dass diese immer in Linearfaktoren zerfallen. Aus der Schule wissen Sie, dass die Lösungen einer quadratischen Gleichung

$$aX^2 + bX + c = 0$$

mit $a, b, c \in \mathbb{R}$ und $a \neq 0$ durch die Formel

$$\frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

beschrieben werden, solange die Quadratwurzel definiert ist, d.h., solange die *Diskriminante* $b^2 - 4ac$ nicht negativ ist. Die gute Nachricht ist: diese Formel gilt auch, wenn a, b, c komplexe Zahlen sind. Wir müssen uns nur überlegen, was $\pm\sqrt{b^2 - 4ac}$ in diesem Fall bedeuten soll.

1.5.27 **Definition** Sei K ein Körper und sei $x \in K$. Ein Element $w \in K$ mit $w^2 = x$ nennt man *Quadratwurzel* von x.

1.5.28 **Beispiel**. In den komplexen Zahlen gilt $i^2 = -1$ und damit ist i eine Quadratwurzel von -1.

1.5.29 **Aufgabe**. Welche Elemente im Körper $\mathbb{F}_5$ besitzen eine Quadratwurzel? **L**

1.5.30 **Beispiel**. Betrachten wir die reellen Zahlen. Eine reelle Zahl $x \in \mathbb{R}$ hat genau dann eine Quadratwurzel, wenn $x \geq 0$ ist. Ist $x > 0$, so hat x genau zwei verschiedene Quadratwurzeln, denn für eine Quadratwurzel w von x ist auch $-w$ eine Quadratwurzel von x:

$$(-w)^2 = (-1)^2 w^2 = x.$$

Eine der beiden Quadratwurzeln ist also stets größer oder gleich 0. In den reellen Zahlen gibt es die Konvention diese eindeutige nicht-negative Quadratwurzel mit $\sqrt{x}$ zu bezeichnen.

1.5.31 **Vorsicht mit dem Wurzelzeichen $\sqrt{x}$.** Ist K ein Körper und ist $w \in K$ eine Quadratwurzel von x, dann ist immer auch $-w$ eine Quadratwurzel von x. Im Allgemeinen gibt es keine sinnvolle Festlegung auf eine dieser Quadratwurzeln, daher ist die Schreibweise $\sqrt{x}$ außerhalb der reellen Zahlen unüblich. Wir verwenden die Notation $\sqrt{x}$ nur, wenn x eine reelle Zahl ist.

1.5.32 **Satz.** *Es sei $z = x + yi$ eine komplexe Zahl. Dann besitzt z genau die Quadratwurzeln*

$$w_{1,2} = \pm\left(\sqrt{\frac{|z| + x}{2}} + \operatorname{sign}(y)\sqrt{\frac{|z| - x}{2}}\, i\right), \tag{1.5.a}$$

dabei bezeichnet $\operatorname{sign}(y)$ *das Vorzeichen von* y.[8] *Die beiden Quadratwurzeln sind genau dann verschieden, wenn* $z \neq 0$ *gilt.*

Hinweis: Da $|z| \geq |x|$ ist (vgl. Aufgabe 1.5.12), sind die Terme unter den Wurzeln immer nicht-negative reelle Zahlen!

Beweis. Wir weisen nach, dass w_i (für $i \in \{1,2\}$) eine Quadratwurzel von z ist:

$$\begin{aligned} w_i^2 &= (\pm 1)^2 \left(\sqrt{\frac{|z|+x}{2}}^2 - \left(\operatorname{sign}(y) \sqrt{\frac{|z|-x}{2}} \right)^2 + 2 \operatorname{sign}(y) \sqrt{\frac{|z|+x}{2} \cdot \frac{|z|-x}{2}}\, i \right) \\ &= \frac{|z|+x}{2} - \operatorname{sign}(y)^2 \frac{|z|-x}{2} + 2 \operatorname{sign}(y) \sqrt{\frac{|z|^2-x^2}{4}} i \\ &= x + 2 \operatorname{sign}(y) \sqrt{\frac{y^2}{4}} i \\ &= x + \operatorname{sign}(y)|y| i = x + yi. \end{aligned}$$

Für $z = 0$ gilt offensichtlich $w_1 = w_2 = 0$. Ist $z \neq 0$, so folgt aus $w_i^2 = z$ auch $w_i \neq 0$. Insbesondere gilt dann $w_1 \neq w_2$ (Beachte: $w_2 = -w_1$).

Warum sind dies die beiden einzigen Quadratwurzeln? Jede Quadratwurzel ist eine Nullstelle des Polynoms $X^2 - z \in \mathbb{C}[X]$. Wir wissen, dass ein Polynom vom Grad 2 höchstens 2 Nullstellen besitzt; siehe 1.4.22. Für $z \neq 0$ kann es also nur die beiden Quadratwurzeln w_1 und w_2 geben. Ist $z = 0$, dann folgt aus $w^2 = 0$ schon $w = 0$, denn $\mathbb{C}$ ist ein Körper und ist damit nullteilerfrei 1.2.22. Es gibt also nur die Quadratwurzel $w = 0$. □

1.5.33 **Beispiel.** (a) Ist $z = x \in \mathbb{R}$ mit $x > 0$, dann ist $|z| = x$ und man erhält mit Formel (1.5.a) die beiden Quadratwurzeln

$$w_{1,2} = \pm \left(\sqrt{\frac{x+x}{2}} + \sqrt{\frac{x-x}{2}}\, i \right) = \pm \sqrt{x}.$$

(b) Ist $z = x \in \mathbb{R}$ mit $x < 0$ (also $|z| = -x$), dann liefert (1.5.a) die beiden komplexen Quadratwurzeln

$$w_{1,2} = \pm \left(\sqrt{\frac{-x+x}{2}} + \sqrt{\frac{-x-x}{2}}\, i \right) = \pm \sqrt{|x|}\, i.$$

[8] $\operatorname{sign}(y)$ nimmt die Werte 1 oder -1 an, abhängig davon, ob $y \geq 0$ oder $y < 0$ ist.

(c) Die beiden Quadratwurzeln von $i \in \mathbb{C}$ sind

$$w_{1,2} = \pm\left(\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}\, i\right).$$

(d) Sei $z = -3 + 4i \in \mathbb{C}$. Mit Formel (1.5.a) findet man die zwei Quadratwurzeln

$$w_{1,2} = \pm(1 + 2i).$$

1.5.34 **Aufgabe**. Bestimmen Sie die beiden Quadratwurzeln der komplexen Zahl $z = 8 - 6i$. **L**

1.5.35 **Satz.** *Es seien $a, b, c \in \mathbb{C}$ mit $a \neq 0$. Die Gleichung*

$$az^2 + bz + c = 0$$

hat die komplexen Lösungen

$$z_{1,2} = \frac{-b \pm w}{2a},$$

wobei w eine Quadratwurzel der Diskriminante $b^2 - 4ac$ ist.

Beweis. Wir multiplizieren die ganze Gleichung mit $4a$. Wegen der Annahme $a \neq 0$, bleiben die Nullstellen dadurch unverändert. Dann setzen wir $\frac{-b \pm w}{2a}$ in die Gleichung ein und erhalten:

$$\begin{aligned} 4a^2\left(\frac{-b \pm w}{2a}\right)^2 + 4ab\frac{-b \pm w}{2a} + 4ac &= (-b \pm w)^2 - 2b^2 \pm 2bw + 4ac \\ &= b^2 \mp 2bw + \underbrace{w^2}_{=b^2-4ac} -2b^2 \pm 2bw + 4ac \\ &= b^2 + b^2 - 4ac - 2b^2 + 4ac = 0. \end{aligned}$$ □

1.5.36 **Aufgabe**. Finden Sie die Lösungen $z \in \mathbb{C}$ der quadratischen Gleichung **L**

$$z^2 + (1 + 2i)z + 3 + 3i = 0.$$

Damit können wir nun einen speziellen Fall von Korollar 1.5.26 beweisen.

1.5.37 **Korollar.** *Jedes Polynom der Form $X^2 + bX + c \in \mathbb{C}[X]$ zerfällt in Linearfaktoren.*

<u>*Beweis*</u>. Es seien
$$z_{1,2} = \frac{-b \pm w}{2}$$
die komplexen Lösungen der Gleichung $z^2 + bz + c$ aus Satz 1.5.35. Dann kann man nachrechnen, dass
$$X^2 + bX + c = (X - z_1)(X - z_2)$$
gilt. □

Auch mit diesen Überlegungen sind wir von einem Beweis des Fundamentalsatzes noch weit entfernt. Aber zumindest haben wir den Fall von Polynomen vom Grad 2 nun so gut verstanden, dass wir viele Beispiele damit behandeln können.

1.6. Lösungen der Aufgaben in Lektion 1

L1.1.4 **Lösung.** Sei G eine Gruppe. Um zu zeigen, dass die Inversionsabbildung $i\colon G \to G$ mit $i(a) = a^{-1}$ bijektiv ist, geben wir eine Umkehrabbildung an. Nach 1.1.3 gilt $i(i(a)) = (a^{-1})^{-1} = a$ für alle $a \in G$. Damit ist i selbst die Umkehrabbildung zu i.

L1.1.5 **Lösung.** Nehmen wir an, es sei $b * a = c * a$. Dann erhalten wir

$$b = b * e = b * (a * a^{-1}) = (b * a) * a^{-1} = (c * a) * a^{-1} = c * (a * a^{-1}) = c.$$

L1.1.6 **Lösung.** Um zu zeigen, dass die Abbildung $f\colon G \to G$ mit $f(x) = x * b$ bijektiv ist, geben wir eine Umkehrabbildung an. Wir definieren $h\colon G \to G$ durch $h(x) = x * b^{-1}$. Dann gilt

$$f(h(x)) = h(x) * b = (x * b^{-1}) * b = x * (b^{-1} * b) = x$$

und genauso

$$h(f(x)) = f(x) * b^{-1} = (x * b) * b^{-1} = x * (b * b^{-1}) = x.$$

Also ist h die Umkehrabbildung zu f und f ist bijektiv.

L1.1.9 **Lösung.** Es soll gezeigt werden, dass die Menge $\mathbb{R} \times \mathbb{R}$ mit der komponentenweisen Addition

$$(x_1, x_2) + (y_1, y_2) := (x_1 + y_1, x_2 + y_2)$$

eine abelsche Gruppe bildet. Dazu prüfen wir nach, dass die Gruppenaxiome erfüllt sind:

(G1): Es seien $(x_1, x_2), (y_1, y_2), (z_1, z_2) \in \mathbb{R} \times \mathbb{R}$. Dann gilt

$$\begin{aligned}
\big((x_1, x_2) + (y_1, y_2)\big) + (z_1, z_2) &= (x_1 + y_1, x_2 + y_2) + (z_1, z_2) \\
&= ((x_1 + y_1) + z_1, (x_2 + y_2) + z_2) \\
&= (x_1 + (y_1 + z_1), x_2 + (y_2 + z_2)) \qquad \text{(Assoz. in } \mathbb{R}) \\
&= (x_1, x_2) + (y_1 + z_1, y_2 + z_2) \\
&= (x_1, x_2) + \big((y_1, y_2) + (z_1, z_2)\big).
\end{aligned}$$

Kommutativgesetz: Für alle $(x_1, x_2), (y_1, y_2) \in \mathbb{R} \times \mathbb{R}$ gilt

$$\begin{aligned}
(x_1, x_2) + (y_1, y_2) &= (x_1 + y_1, x_2 + y_2) \\
&= (y_1 + x_1, y_2 + x_2) \qquad \text{(Kommutativgesetz in } \mathbb{R})
\end{aligned}$$

$$= (y_1, y_2) + (x_1, x_2).$$

(G2): Das neutrale Element ist das Paar $(0,0)$, denn für alle $(x_1, x_2) \in \mathbb{R} \times \mathbb{R}$ gilt

$$(0,0) + (x_1, x_2) = (0 + x_1, 0 + x_2) = (x_1, x_2).$$

Da wir bereits wissen, dass das Kommutativgesetz gilt, folgt auch $(x_1, x_2) + (0,0) = (0,0) + (x_1, x_2) = (x_1, x_2)$.

(G3): Das Inverse zu $(x_1, x_2) \in \mathbb{R} \times \mathbb{R}$ ist $(-x_1, -x_2)$, denn es gilt

$$(x_1, x_2) + (-x_1, -x_2) = (x_1 + (-x_1), x_2 + (-x_2)) = (0,0).$$

Das Kommutativgesetz liefert uns nun auch $(-x_1, -x_2) + (x_1, x_2) = (x_1, x_2) + (-x_1, -x_2) = (0,0)$.

L1.1.11 Lösung. Die Menge $M = \mathbb{Z} \setminus \{0\}$ bildet mit der Multiplikation $\cdot$ keine Gruppe. Die Multiplikation ist zwar assoziativ und die Zahl 1 ist das neutrale Element, aber nur für die Zahlen $\pm 1 \in M$ findet man auch inverse Elemente. Zum Beispiel wählen wir $2 \in M$. Dann gilt für alle $n \in M$

$$(2n)^2 = 4n^2 \geq 4 > 1$$

und damit ist $2n \neq 1$, d.h., 2 besitzt kein Inverses bezüglich der Multiplikation in M.

L1.1.13 Lösung. Sei K ein Körper und sei $n \geq 2$. Wir wollen zeigen, dass $\mathrm{GL}_n(K)$ nicht abelsch ist. Für $n = 2$ wissen wir das aus 1.1.12. Hier gehen wir ganz ähnlich vor und definieren die Blockdiagonalmatrizen

$$A = \left(\begin{array}{cc|c} 1 & 1 & \\ 0 & 1 & \\ \hline & & I_{n-2} \end{array}\right) \quad \text{und} \quad B = \left(\begin{array}{cc|c} 1 & 0 & \\ 1 & 1 & \\ \hline & & I_{n-2} \end{array}\right)$$

(dabei sind in den Blöcken links unten und rechts oben alle Einträge 0). Dass diese Matrizen invertierbar sind, sieht man wie in 1.1.12. Jetzt gilt

$$AB = \left(\begin{array}{cc|c} 2 & 1 & \\ 1 & 1 & \\ \hline & & I_{n-2} \end{array}\right) \quad \text{und} \quad BA = \left(\begin{array}{cc|c} 1 & 1 & \\ 1 & 2 & \\ \hline & & I_{n-2} \end{array}\right).$$

Wegen $1 + 1 \neq 1$ erhalten wir $AB \neq BA$, d.h., $\mathrm{GL}_n(K)$ ist nicht abelsch.

L1.1.14 Lösung. Die Menge $\mathrm{M}_{2,2}(\mathbb{R})$ bildet keine Gruppe bezüglich der Matrizenmultiplikation. Die Matrizenmultiplikation ist assoziativ und die Einheitsmatrix I_2 ist das neutrale Element, aber es gibt Matrizen, die nicht invertierbar sind. Zum Beispiel hat die Nullmatrix 0 kein Inverses, denn $0B = 0 \neq I_2$ für alle $B \in \mathrm{M}_{2,2}(\mathbb{R})$.

L1.1.21 Lösung. Wir zeigen, dass die Menge $\mathbb{R}_{>0}$ der positiven reellen Zahlen eine Untergruppe der Gruppe $(\mathbb{R}^\times, \cdot)$ ist. Dazu prüfen wir die Bedingungen aus Definition 1.1.18.

(UG1): Die Menge der positiven reellen Zahlen ist nicht leer, z.B. ist $1 \in \mathbb{R}_{>0}$.

(UG2): Das Produkt zweier positiver Zahlen ist positiv.

(UG3): Ist $x \in \mathbb{R}_{>0}$ beliebig, so ist das Inverse $x^{-1} = \frac{1}{x}$ ebenfalls positiv, d.h., $x^{-1} \in \mathbb{R}_{>0}$.

L1.2.6 Lösung. Wir leiten die Gleichung $\left(\sum_{i=1}^n a_i\right) \cdot \left(\sum_{j=1}^m b_j\right) = \sum_{i=1}^n \sum_{j=1}^m a_i b_j$ mit vollständiger Induktion nach m her. Der Induktionsanfang mit $m = 1$ wurde in 1.2.2 (e) gemacht. Für den Induktionsschritt nehmen wir an, dass die Gleichung für m bekannt ist und betrachten $m + 1$. Mit dem Distributivgesetz erhalten wir

$$\begin{aligned}
\Big(\sum_{i=1}^n a_i\Big) \cdot \Big(\sum_{j=1}^{m+1} b_j\Big) &= \Big(\sum_{i=1}^n a_i\Big) \cdot \Big(\sum_{j=1}^{m} b_j + b_{m+1}\Big) \\
&= \Big(\sum_{i=1}^n a_i\Big) \cdot \Big(\sum_{j=1}^{m} b_j\Big) + \Big(\sum_{i=1}^n a_i\Big) \cdot b_{m+1} && \text{(Distributivgesetz)} \\
&= \sum_{i=1}^n \sum_{j=1}^m a_i b_j + \sum_{i=1}^n a_i b_{m+1} && \text{(I.V.)} \\
&= \sum_{i=1}^n \sum_{j=1}^{m+1} a_i b_j .
\end{aligned}$$

Damit ist die Gleichung gezeigt.

L1.2.12 Lösung. (UR2): Die Menge B enthält die Einheitsmatrix

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix};$$

dies ist das Einselement des Ringes $\mathrm{M}_{2,2}(\mathbb{R})$.

(UR1): Wir zeigen, dass $B \subseteq \mathrm{M}_{2,2}(\mathbb{R})$ eine Untergruppe ist, indem wir nachweisen, dass die Untergruppenaxiome aus 1.1.18 erfüllt sind.

- (UG1) Wegen $I_2 \in B$ ist B nicht leer.
- (UG2) B ist unter der Addition von Matrizen abgeschlossen, denn es gilt

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix} + \begin{pmatrix} a' & b' \\ 0 & d' \end{pmatrix} = \begin{pmatrix} a+a' & b+b' \\ 0 & d+d' \end{pmatrix}$$

für alle $a, a', b, b', d, d' \in \mathbb{R}$.

- (UG3) Sei $A = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in B$. Das additive Inverse zu A ist

$$-A = \begin{pmatrix} -a & -b \\ 0 & -d \end{pmatrix}$$

und dieses Element liegt ebenfalls in B.

(UR3): B ist unter der Matrizenmultiplikation abgeschlossen, denn es gilt

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \begin{pmatrix} a' & b' \\ 0 & d' \end{pmatrix} = \begin{pmatrix} aa' & ab'+bd' \\ 0 & dd' \end{pmatrix}$$

für alle $a, a', b, b', d, d' \in \mathbb{R}$.

L1.2.16 Lösung. Das neutrale Element der Multiplikation in $\mathbb{Z}[\sqrt{2}]$ ist die Zahl 1. Mit der dritten binomischen Formel, gilt $(-1-\sqrt{2})(1-\sqrt{2}) = -1+2 = 1$. Da $\mathbb{Z}[\sqrt{2}]$ ein kommutativer Ring ist, gilt genauso $(1-\sqrt{2})(-1-\sqrt{2}) = 1$. Also ist $1-\sqrt{2}$ eine Einheit mit dem Inversen $-1-\sqrt{2}$.

Das Inverse kann man beispielsweise mit einem Ansatz $(a+b\sqrt{2})(1-\sqrt{2}) = 1$ finden, wenn man nach $a, b \in \mathbb{Z}$ auflöst.

L1.2.23 Lösung. Es seien $a, b, c \in R$ mit $a \neq 0$ gegeben. Gilt $ab = ac$, dann folgt

$$a(b-c) = 0.$$

Da R nullteilerfrei ist und $a \neq 0$ angenommen wurde, gilt $b-c = 0$, d.h., $b = c$.

L1.2.26 Lösung. Es sei R ein Ring und es seien $m, n, s, t \in \mathbb{N}$.

Zu (i): Es soll gezeigt werden, dass $(\mathrm{M}_{m,n}(R), +)$ eine abelsche Gruppe bildet. Dazu verwenden wir die Gruppenaxiome aus 1.1.1. Es seien $A = (a_{ij})_{i,j}, B = (b_{ij})_{i,j}, C = (c_{ij})_{i,j} \in \mathrm{M}_{m,n}(R)$. Zuerst rechnen wir nach, dass das Assoziativgesetz gilt. Es ist

$$(A+B)+C = (a_{ij}+b_{ij})_{i,j}+C = (a_{ij}+b_{ij}+c_{ij})_{i,j} = A+(b_{ij}+c_{ij})_{i,j} = A+(B+C).$$

Die Addition von Matrizen ist also assoziativ.

Da $0 \in R$ das neutrale Element der Addition in R ist, erfüllt die Nullmatrix sicherlich $0 + A = A$ und $A + 0 = A$. Die Nullmatrix ist also das neutrale Element der Addition.

Die Addition von Matrizen erfüllt das Kommutativgesetz, denn die Addition in R ist kommutativ; genauer gilt

$$A + B = (a_{ij} + b_{ij})_{i,j} = (b_{ij} + a_{ij})_{i,j} = B + A.$$

Das Inverse zu $A \in \mathrm{M}_{m,n}(R)$ ist die Matrix $-A = (-a_{ij})_{i,j}$, denn es gilt

$$A + (-A) = (a_{ij} - a_{ij})_{i,j} = (0)_{i,j} = 0.$$

und aufgrund der Kommutativität gilt auch $(-A) + A = 0$. Damit ist $(\mathrm{M}_{m,n}(R), +)$ eine abelsche Gruppe.

Zu (ii): Seien $A \in \mathrm{M}_{m,n}(R)$, $B \in \mathrm{M}_{n,s}(R)$ und $C \in \mathrm{M}_{s,t}(R)$ gegeben. Wir wollen zeigen, dass

$$(AB)C = A(BC)$$

gilt. Wir definieren $AB = (d_{ij})_{i,j}$ mit

$$d_{ij} = \sum_{\ell=1}^{n} a_{i,\ell} b_{\ell,j}$$

für alle $1 \le i \le m$ und $1 \le j \le s$; genauso setzen wir $BC = (e_{ij})_{i,j}$ mit

$$e_{ij} = \sum_{k=1}^{s} b_{i,k} c_{k,j}$$

für alle $1 \le i \le n$ und $1 \le j \le t$.

Es seien nun $i \in \{1, 2, \ldots, m\}$ und $j \in \{1, 2, \ldots, t\}$ gegeben. Der Eintrag an der Stelle (i, j) von $(AB)C$ ist

$$\begin{aligned} \sum_{k=1}^{s} d_{ik} c_{kj} &= \sum_{k=1}^{s} \left(\sum_{\ell=1}^{n} a_{i,\ell} b_{\ell k} \right) c_{kj} = \sum_{k=1}^{s} \sum_{\ell=1}^{n} (a_{i\ell} b_{\ell k}) c_{kj} \\ &= \sum_{\ell=1}^{n} \sum_{k=1}^{s} a_{i\ell} (b_{\ell k} c_{kj}) = \sum_{\ell=1}^{n} a_{i\ell} \left(\sum_{k=1}^{s} b_{\ell,k} c_{k,j} \right) = \sum_{\ell=1}^{n} a_{i\ell} e_{\ell j}. \end{aligned}$$

Dabei verwenden wir die Distributivgesetze in R und die Assoziativität der Multiplikation in R. Der letzte Ausdruck ist der Eintrag von $A(BC)$ an der Stelle (i, j). Da (i, j) beliebig war, haben $(AB)C$ und $A(BC)$ dieselben Einträge, d.h., $(AB)C = A(BC)$.

L1.2.29 Lösung. Sei R ein kommutativer Ring. Seien $A = (a_{ij})_{i,j} \in \mathrm{M}_{n,n}(R)$ und $B = (b_{ij})_{i,j} \in \mathrm{M}_{n,n}(R)$.

Zu (a): Es gilt

$$\operatorname{Spur}(A+B) = \sum_{i=1}^{n}(a_{ii}+b_{ii}) = \sum_{i=1}^{n} a_{ii} + \sum_{i=1}^{n} b_{ii} = \operatorname{Spur}(A) + \operatorname{Spur}(B).$$

Zu (b): Für alle $a \in R$ ist

$$\operatorname{Spur}(aA) = \sum_{i=1}^{n} a a_{ii} = a \sum_{i=1}^{n} a_{ii} = a \operatorname{Spur}(A).$$

Zu (c) Es sei $C = AB$ und es sei $D = BA$. Dann erhalten wir

$$\begin{aligned}
\operatorname{Spur}(AB) = \operatorname{Spur}(C) &= \sum_{i=1}^{n} c_{ii} \\
&= \sum_{i=1}^{n}\sum_{k=1}^{n} a_{ik}b_{ki} && \text{(Def. } AB\text{)} \\
&= \sum_{k=1}^{n}\sum_{i=1}^{n} b_{ki}a_{ik} && \text{(Summen tauschen, } R \text{ kommutativ)} \\
&= \sum_{k=1}^{n} d_{kk} = \operatorname{Spur}(D) = \operatorname{Spur}(BA). && \text{(Def. } BA\text{)}
\end{aligned}$$

L1.2.33 Lösung. Diese Aussage ist bereits für $n = 2$ falsch, falls in R die natürliche Bedingung $1 \neq 0$ erfüllt ist, d.h. wenn R nicht der Nullring ist; siehe 1.2.7. Wir setzen

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

Dann ist

$$A^T A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

aber

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

Also gilt $A^T A \neq AA^T$.

L1.2.34 Lösung. Sei $I_n \in \mathrm{M}_{n,n}(K)$ die Einheitsmatrix. Wir stellen fest, dass $I_n^T = I_n$ gilt. Es sei A^{-1} die Inverse zu $A \in \mathrm{M}_{n,n}(K)$. Mit den Rechenregeln aus 1.2.32 erhalten wir nun

$$A^T(A^{-1})^T = (A^{-1}A)^T = I_n^T = I_n$$

und genauso

$$(A^{-1})^T A^T = (AA^{-1})^T = I_n^T = I_n.$$

Also ist A^T invertierbar und $(A^{-1})^T$ ist die zugehörige Inverse, d.h., $(A^T)^{-1} = (A^{-1})^T$.

L1.2.38 Lösung. Es seien $A = \mathrm{diag}(A_1, \ldots, A_s)$ und $B = \mathrm{diag}(B_1, \ldots, B_s)$ Blockdiagonalmatrizen in $\mathrm{M}_{n,n}(R)$ vom Blocktyp $p_1, p_2, \ldots, p_s$. Die Einträge von A bezeichnen wir mit a_{ij}, die Einträge von B mit b_{ij} und die Einträge von $C = AB$ mit c_{ij}. Wir definieren $P_0 = 0$ und $P_t = \sum_{r=1}^{t} p_r$ für alle $t \in \{1, 2, \ldots, s\}$. Wir bemerken, dass $P_s = n$ gilt. Die Formel zur Matrizenmultiplikation (siehe 1.2.25) liefert uns

$$\begin{aligned} c_{ij} &= \sum_{k=1}^{n} a_{ik}b_{kj} \\ &= \sum_{t=0}^{s-1} \sum_{k=P_t+1}^{P_{t+1}} a_{ik}b_{kj}. \end{aligned}$$

Betrachten wir nun ein festes k mit $P_t < k \leq P_{t+1}$. Da A eine Blockdiagonalmatrix vom Blocktyp $p_1, \ldots, p_s$ ist, gilt $a_{ik} = 0$, falls i nicht ebenfalls $P_t < i \leq P_{t+1}$ erfüllt. Genauso gilt $b_{kj} = 0$, falls j nicht ebenfalls $P_t < j \leq P_{t+1}$ erfüllt. Also folgt aus $c_{ij} \neq 0$, die Bedingung $P_t < i, j \leq P_{t+1}$ für ein $t \in \{0, 1, \ldots, s-1\}$; d.h., $C = AB$ ist eine Blockdiagonalmatrix.

Falls nun $P_t < i, j \leq P_{t+1}$ für ein t gilt, dann erhalten wir aus der obigen Formel

$$\begin{aligned} c_{ij} &= \sum_{k=1}^{m} a_{ik}b_{kj} \\ &= \sum_{k=P_t+1}^{P_{t+1}} a_{ik}b_{kj} \end{aligned}$$

und das ist der Eintrag der Matrix $A_{t+1}B_{t+1}$ an der Stelle $(i - P_t, j - P_t)$. Das heißt, es gilt

$$AB = \mathrm{diag}(A_1B_1, \ldots, A_sB_s).$$

L1.3.9 **Lösung.** Wir betrachten Restklassen in $\mathbb{Z}/18\mathbb{Z}$.

(a) Es gilt $\overline{59} = \overline{5}$, denn $59 - 5 = 54 = 3 \cdot 18$; vgl. 1.3.6 (b).

(b) Es gilt $\overline{-18} = \overline{0}$ und damit $\overline{-18} \neq \overline{5}$; vgl. 1.3.6 (c).

(c) Es gilt $\overline{-157} = \overline{5}$, denn $-157 - 5 = -162 = -9 \cdot 18$.

(d) Teilt man 91 durch 18, dann erhält man den Rest 1. Es gilt also $\overline{91} = \overline{1}$ und $\overline{91} \neq \overline{5}$.

L1.3.12 **Lösung.** Im Ring $\mathbb{Z}/22\mathbb{Z}$ gelten

$$\overline{13} \cdot \overline{4} - \overline{20} = \overline{13 \cdot 4} - \overline{20} = \overline{52 - 20} = \overline{32} = \overline{10}$$

und

$$\overline{87}^{123456789} = \overline{(-1)}^{123456789} = \overline{(-1)^{123456789}} = \overline{-1} = \overline{21}.$$

L1.3.19 **Lösung.** Wir verwenden den Euklidischen Algorithmus mit $p = 31$ und $a = 17$.

$$\begin{aligned} 31 &= 17 + 14 \\ 17 &= 14 + 3 \\ 14 &= 4 \cdot 3 + 2 \\ 3 &= 2 + 1. \end{aligned}$$

Nun setzen wir ein und erhalten

$$\begin{aligned} 1 = 3 - 2 = 3 - (14 - 4 \cdot 3) &= 5 \cdot 3 - 14 \\ &= 5 \cdot (17 - 14) - 14 = 5 \cdot 17 - 6 \cdot 14 \\ &= 5 \cdot 17 - 6 \cdot (31 - 17) = 11 \cdot 17 - 6 \cdot 31. \end{aligned}$$

Damit ist $\overline{11}$ das multiplikative Inverse zu $\overline{17}$ in $\mathbb{F}_{31}$.

L1.3.20 **Lösung.** Wir überführen das Gleichungssystem in eine erweiterte Koeffizientenmatrix

$$A = \left(\begin{array}{ccc|c} \overline{1} & \overline{2} & \overline{4} & \overline{1} \\ \overline{1} & \overline{0} & \overline{3} & \overline{1} \\ \overline{2} & \overline{2} & \overline{3} & \overline{1} \end{array}\right)$$

und arbeiten mit den Gaußalgorithmus. Wir ziehen die oberste Zeile einmal von der zweiten und zweimal von der dritten Zeile ab; dann erhalten wir

$$\left(\begin{array}{ccc|c} \overline{1} & \overline{2} & \overline{4} & \overline{1} \\ \overline{0} & \overline{-2} & \overline{-1} & \overline{0} \\ \overline{0} & \overline{-2} & \overline{0} & \overline{-1} \end{array}\right).$$

Es gilt $\overline{-2} = \overline{3}$ in $\mathbb{F}_5$. Das multiplikative Inverse von $\overline{3}$ ist $\overline{2}$. Nach Vertauschen der letzten beiden Zeilen und Multiplikation mit $\overline{2}$ erhalten wir

$$\left(\begin{array}{ccc|c} \overline{1} & \overline{2} & \overline{4} & \overline{1} \\ \overline{0} & \overline{1} & \overline{0} & \overline{3} \\ \overline{0} & \overline{1} & \overline{3} & \overline{0} \end{array}\right).$$

Wir ziehen die zweite von der letzten Zeile ab und gelangen schließlich zur Matrix

$$\left(\begin{array}{ccc|c} \overline{1} & \overline{2} & \overline{4} & \overline{1} \\ \overline{0} & \overline{1} & \overline{0} & \overline{3} \\ \overline{0} & \overline{0} & \overline{3} & \overline{2} \end{array}\right).$$

Nun können wir das Gleichungssystem von unten nach oben auflösen und finden genau eine Lösung

$$x_1 = \overline{4}, \quad x_2 = \overline{3}, \quad x_3 = \overline{4}.$$

L1.4.6 **Lösung.** Man kann wie gewohnt ausmultiplizieren und erhält

$$(X^2-6)\cdot(X^2+X+1) = X^4+X^3+X^2-6X^2-6X-6 = X^4+X^3-5X^2-6X-6.$$

L1.4.8 **Lösung.** Es seien $f = \sum_i a_i X^i$ und $g = \sum_i b_i X^i$ mit $m = \operatorname{Grad}(f)$ und $n = \operatorname{Grad}(g)$. Ohne Beschränkung der Allgemeinheit ist $m \geq n$, d.h., für alle $i > m$ ist $a_i = b_i = 0$. Dann ist auch $a_i + b_i = 0$ für alle $i > m$ und damit hat $f + g = \sum_{i=0}(a_i + b_i)X^i$ einen Grad kleiner oder gleich m. Das heißt, $\operatorname{Grad}(f+g) \leq m = \max(\operatorname{Grad}(f), \operatorname{Grad}(g))$.

(Eine echte Ungleichung tritt auf, wenn f und g denselben Grad haben und sich die Leitkoeffizienten bei der Addition wegheben.)

L1.4.9 **Lösung.** Da wir bereits wissen, dass die Multiplikation in $K[X]$ kommutativ ist, genügt es, eines der beiden Distributivgesetze zu prüfen. Für alle $f, g, h \in K[X]$ folgt aus den Distributivgesetzen in K, dass

$$\begin{aligned} f\cdot(g+h) &= \sum_i \Big(\sum_{j=0}^{i} a_j(b_{i-j}+c_{i-j})\Big)X^i = \sum_i \Big(\sum_{j=0}^{i} a_j b_{i-j} + a_j c_{i-j}\Big)X^i \\ &= \sum_i \Big(\sum_{j=0}^{i} a_j b_{i-j}\Big)X^i + \sum_i \Big(\sum_{j=0}^{i} a_j c_{i-j}\Big)X^i = (f\cdot g) + (f\cdot h) \end{aligned}$$

gilt, d.h., die Distributivgesetze in $K[X]$ sind erfüllt.

L1.4.14 Lösung. Gegeben sind $g = X^6 + X^4 + X + \overline{1} \in \mathbb{F}_2[X]$ und $f = X^2 + X + \overline{1} \in \mathbb{F}_2[X]$. Wir führen die Polynomdivision durch:

$$\begin{array}{llllll}
(X^6 & & +X^4 & & & +X+\overline{1}) : (X^2 + X + \overline{1}) = X^4 + X^3 + X^2 + \overline{1} \\
-(X^6 & +X^5 & +X^4) & & & \\
\hline
 & X^5 & & & & +X+\overline{1} \\
 & -(X^5 & +X^4 & +X^3) & & \\
\hline
 & & X^4 & +X^3 & & +X+\overline{1} \\
 & & -(X^4 & +X^3 & +X^2) & \\
\hline
 & & & & X^2 +X & +\overline{1} \\
 & & & & -(X^2 +X & +\overline{1}) \\
\hline
 & & & & & 0
\end{array}$$

Die Division geht also auf, d.h., der Rest der Division ist 0. Der Quotient ist $X^4 + X^3 + X^2 + \overline{1}$ und es gilt damit

$$g = (X^4 + X^3 + X^2 + \overline{1})f.$$

L1.4.18 Lösung. Wir setzen alle Werte $\overline{0}, \overline{1}, \overline{2} \in \mathbb{F}_3$ in $f = X$ und $g = X^4 - X^2 + X$ ein.

$$\begin{aligned}
g(\overline{0}) &= \overline{0} - \overline{0} + \overline{0} && = \overline{0} = f(\overline{0}) \\
g(\overline{1}) &= \overline{1}^4 - \overline{1}^2 + \overline{1} && = \overline{1} = f(\overline{1}) \\
g(\overline{2}) &= \overline{2}^4 - \overline{2}^2 + \overline{2} = \overline{1} - \overline{1} + \overline{2} && = \overline{2} = f(\overline{2})
\end{aligned}$$

Beide Polynome definieren also dieselbe Funktion.

L1.4.29 Lösung. Gesucht ist die Vielfachheit der Nullstelle 1 im Polynom $X^6 - 2X^5 + X^4 - X^2 + 2X - 1 \in \mathbb{R}[X]$.

Wir spalten mit der Polynomdivision einen Faktor $(X - 1)$ ab und erhalten

$$f = (X - 1)f_1$$

mit $f_1 = X^5 - X^4 - X + 1$. Es ist $f_1(1) = 1 - 1 - 1 + 1 = 0$, also sind wir noch nicht fertig. Wir spalten aus f_1 wieder einen Faktor $(X - 1)$ ab und erhalten

$$f = (X - 1)^2 f_2$$

mit $f_2 = X^4 - 1$. Es gilt $f_2(1) = 1 - 1 = 0$. Wir spalten daher wieder einen Faktor $(X - 1)$ aus f_2 ab und erhalten

$$f = (X-1)^3 f_3$$

mit $f_3 = X^3 + X^2 + X + 1$. Da $f_3(1) = 4 \neq 0$ ist, können wir den Linearfaktor $(X - 1)$ nicht weiter abspalten. Die Vielfachheit der Nullstelle 1 in f beträgt also 3.

L1.4.36 **Lösung.** Wir untersuchen, ob $f = X^4 + X^2 \in \mathbb{F}_2[X]$ in Linearfaktoren zerfällt. Wir schreiben $f = X^2(X^2 + \overline{1})$ und haben damit schon den Linearfaktor $X = (X - \overline{0})$ doppelt abgespaltet.

Wir setzen $f_1 = X^2 + \overline{1}$ und suchen eine Nullstelle. Wir testen, ob $\overline{1}$ eine Nullstelle ist. Es gilt $f_1(\overline{1}) = \overline{1} + \overline{1} = \overline{0}$. Wir spalten diese Nullstelle von f_1 ab und erhalten $(X + \overline{1})(X + \overline{1}) = X^2 + \overline{1}$. Daraus folgt

$$f = X^2(X + \overline{1})^2$$

und f zerfällt somit in Linearfaktoren.

L1.4.43 **Lösung.** Um zu zeigen, dass diese Körper nicht algebraisch abgeschlossen sind, müssen wir jeweils ein nicht-konstantes Polynom ohne Nullstellen angeben.

In $\mathbb{F}_3$: Wir wollen zeigen, dass $f = X^3 - X + \overline{1} \in \mathbb{F}_3[X]$ keine Nullstelle hat. Das rechnen wir einfach nach:

$$\begin{aligned} f(\overline{0}) &= \overline{0} - \overline{0} + \overline{1} = \overline{1} \\ f(\overline{1}) &= \overline{1} - \overline{1} + \overline{1} = \overline{1} \\ f(\overline{2}) &= \overline{8} - \overline{2} + \overline{1} = \overline{1} \end{aligned}$$

In $\mathbb{F}_5$: Wir behaupten, dass das Polynom $f = X^2 + \overline{2}$ keine Nullstelle besitzt. Um das zu zeigen, setzen wir alle Elemente aus $\mathbb{F}_5$ in das Polynom ein:

$$\begin{aligned} f(\overline{0}) &= \overline{0} + \overline{2} = \overline{2} \\ f(\overline{1}) &= \overline{1} + \overline{2} = \overline{3} \\ f(\overline{2}) &= \overline{4} + \overline{2} = \overline{1} \\ f(\overline{3}) &= \overline{9} + \overline{2} = \overline{1} \\ f(\overline{4}) &= \overline{16} + \overline{2} = \overline{3} \end{aligned}$$

L1.4.48 Lösung. Wir zeigen, dass

$$fK[X] = \{fh \in K[X] \mid h \in K[X]\}$$

ein Ideal ist, indem wir die Bedingungen 1.4.45 (I1) und (I2) prüfen.

(I1): Es seien $fh_1, fh_2 \in fK[X]$ mit $h_1, h_2 \in K[X]$. Dann gilt

$$fh_1 + fh_2 = f(h_1 + h_2) \in fK[X].$$

(I2): Es seien $fh \in fK[X]$ und $g \in K[X]$ gegeben. Dann gilt $g(fh) = f(gh) \in fK[X]$.

Also ist $fK[X]$ ein Ideal.

L1.4.52 Lösung. Man kann hier den Beweis von 1.4.51 nehmen und für viele Faktoren umschreiben.

Wir wollen einen anderen Weg gehen und die Aussage per Induktion über n aus 1.4.51 herleiten. Der Induktionsanfang $n = 1$ ist offensichtlich.

Es seien nun $n+1$ Ideale $I_1, \ldots, I_{n+1}$ in $K[X]$ gegeben. Wir stellen fest, dass gilt

$$\sum_{i=1}^{n+1} I_i = \Big(\sum_{i=1}^{n} I_i\Big) + I_{n+1}.$$

Also folgt die Aussage aus der Induktionsvorraussetzung und 1.4.51.

L1.4.54 Lösung. Es sei h ein gemeinsamer Teiler von f_1 und f_2. Wir schreiben $f_i = hu_i$ mit $u_i \in K[X]$. Mit (c) erhalten wir

$$\mathrm{ggT}(f_1, f_2) = f_1 g_1 + f_2 g_2 = h(u_1 g_1 + u_2 g_2),$$

d.h. h teilt $\mathrm{ggT}(f_1, f_2)$.

L1.4.57 Lösung. Da f und g_i teilerfremd sind, gibt es nach Satz 1.4.54 (c) Polynome $p_i, q_i \in K[X]$ mit

$$p_i f + q_i g_i = 1$$

für $i \in \{1, 2\}$. Wir multiplizieren diese Ausdrücke und erhalten

$$\begin{aligned} 1 &= (p_1 f + q_1 g_1)(p_2 f + q_2 g_2) \\ &= f^2 p_1 p_2 + f p_1 q_2 g_2 + f p_2 q_1 g_1 + q_1 q_2 g_1 g_2 \\ &= f(p_1 p_2 f + p_1 q_2 g_2 + p_2 q_1 g_1) + g_1 g_2 q_1 q_2. \end{aligned}$$

Daraus folgt $1 \in fK[X] + g_1 g_2 K[X]$ und, da alle Vielfachen von 1 ebenfalls im Ideal liegen, erhalten wir $K[X] = fK[X] + g_1 g_2 K[X]$. Damit sind f und $g_1 g_2$ teilerfremd.

L1.4.58 Lösung. Seien $f, g \in K[X]$ und $n, m \in \mathbb{N}$. Wir nehmen an, dass f und g teilerfremd sind, und wollen zeigen, dass auch f^n und g^m teilerfremd sind. Dazu zeigen wir zuerst mit vollständiger Induktion nach m, dass f und g^m teilerfremd sind. Für $m = 1$ ist das genau unsere Voraussetzung. Sei nun $m > 1$. Im Induktionsschritt verwenden wir Aufgabe 1.4.57. Sind f und g^{m-1} teilerfremd, dann folgt aus Aufgabe 1.4.57, dass auch f und $g^m = g^{m-1}g$ teilerfremd sind.

Sei nun $m \in \mathbb{N}$ fest. Wir wissen nun, dass f und g^m teilerfremd sind. Mit demselben Induktionsargument folgt nun, dass auch f^n und g^m teilerfremd sind für alle $n \in \mathbb{N}$.

L1.5.8 Lösung. Sei $z \in \mathbb{C}$. Wir schreiben z in der Form $z = x + yi$ mit $x, y \in \mathbb{R}$. Es ist $(-i)i = -i^2 = 1$ und wir erhalten

$$\mathrm{Re}(-iz) = \mathrm{Re}(-ix + y(-i)i) = \mathrm{Re}(y - ix) = y = \mathrm{Im}(z).$$

L1.5.9 Lösung. Es seien $z, w \in \mathbb{C}$ und $a \in \mathbb{R}$. Wir schreiben $z = x + yi$ und $w = u + vi$ mit $x, y, u, v \in \mathbb{R}$. Dann ist $x = \mathrm{Re}(z)$ und $u = \mathrm{Re}(w)$. Damit erhalten wir

$$\mathrm{Re}(az + w) = \mathrm{Re}((ax + u) + (ay + v)i) = ax + u = a\mathrm{Re}(z) + \mathrm{Re}(w).$$

L1.5.12 Lösung. Es sei $z \in \mathbb{C}$. Wir schreiben $z = a + bi$ mit $a, b \in \mathbb{R}$. Dabei ist $a = \mathrm{Re}(z)$. Dann gilt

$$a \leq |a| = \sqrt{a^2} \leq \sqrt{a^2 + b^2} = |z|.$$

L1.5.14 Lösung. Um komplexe Zahlen $z = x + yi$ auf dem Einheitskreis zu finden, muss man x und y mit $x^2 + y^2 = 1$ wählen. Zum Beispiel liegen die Zahlen

$$z_1 = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}i, \quad z_2 = \frac{3}{5} - \frac{4}{5}i, \quad z_3 = \frac{1}{2} + \frac{\sqrt{3}}{2}i$$

auf dem Einheitskreis.

L1.5.17 Lösung. Man erhält:

- $(4 + 2i) - (1 - i) = 3 + 3i$
- $(2 - 5i)(3 + i) = 6 - 15i + 2i - 5i^2 = 11 - 13i.$

L1.5.21 Lösung. Wir verwenden die Formel aus Korollar 1.5.20.

$$(1+i)^{-1} = \frac{1}{1^2+1^2}(1-i) = \frac{1}{2} - \frac{1}{2}i$$
$$(2+3i)^{-1} = \frac{1}{2^2+3^2}(2-3i) = \frac{2}{13} - \frac{3}{13}i$$

L1.5.23 Lösung. Wir verwenden den Trick aus 1.5.22 und erhalten

(a) $$\frac{11-2i}{2+i} = \frac{(11-2i)(2-i)}{4+1} = \frac{20-15i}{5} = 4-3i$$

(b) $$\frac{13i}{-3+2i} = \frac{13i(-3-2i)}{9+4} = \frac{13(2-3i)}{13} = 2-3i$$

L1.5.24 Lösung. Wir erstellen die erweiterte Koeffizientenmatrix

$$A = \left(\begin{array}{ccc|c} 1 & 0 & i & 1 \\ 0 & 1+i & 2 & 1+3i \\ 2 & 2 & -i & 6+2i \end{array}\right)$$

und arbeiten mit dem Gaußalgorithmus. Wir ziehen die oberste Zeile zweimal von der untersten ab und erhalten

$$\left(\begin{array}{ccc|c} 1 & 0 & i & 1 \\ 0 & 1+i & 2 & 1+3i \\ 0 & 2 & -3i & 4+2i \end{array}\right).$$

Es gilt $(1+i)(1-i) = 2$. Also ziehen wir das $(1-i)$-fache der zweiten Zeile von der letzten ab und erhalten (mit der Nebenrechnung $(1+3i)(1-i) = 4+2i$)

$$\left(\begin{array}{ccc|c} 1 & 0 & i & 1 \\ 0 & 1+i & 2 & 1+3i \\ 0 & 0 & -2-i & 0 \end{array}\right).$$

Damit können wir die Lösung ablesen. Es muss $x_3 = 0$ sein. Damit erhalten wir

$$x_2 = \frac{1+3i}{1+i} = \frac{(1+3i)(1-i)}{1^2+1^2} = \frac{4+2i}{2} = 2+i.$$

Und schließlich sehen wir, dass $x_1 = 1$ sein muss.

L1.5.29 Lösung. Wir gehen die Elemente im Körper $\mathbb{F}_5$ durch und bestimmen die Quadrate:

$$\overline{0}^2 = \overline{0}, \overline{1}^2 = \overline{1}, \overline{2}^2 = \overline{4}, \overline{3}^2 = \overline{4}, \overline{4}^2 = \overline{1}.$$

Aus dieser Liste sehen wir:

- $\overline{0}$ hat genau eine Quadratwurzel in $\mathbb{F}_5$, nämlich $\overline{0}$.
- $\overline{1}$ hat genau zwei Quadratwurzeln, nämlich $\overline{1}$ und $\overline{4} = -\overline{1}$.
- $\overline{4}$ hat genau zwei Quadratwurzeln, nämlich $\overline{2}$ und $\overline{3} = -\overline{2}$.
- $\overline{2}$ und $\overline{3}$ haben keine Quadratwurzeln in $\mathbb{F}_5$.

L1.5.34 Lösung. Gegeben ist $z = 8 - 6i$. Es gilt

$$|z| = \sqrt{64 + 36} = \sqrt{100} = 10$$

und $\operatorname{sign}(-6) = -1$. Mit Formel (1.5.a) erhalten wir die Quadratwurzeln

$$w_{1,2} = \pm\left(\sqrt{\frac{10+8}{2}} - \sqrt{\frac{10-8}{2}}\, i\right) = \pm(3 - i).$$

L1.5.36 Lösung. Wir betrachten die Gleichung

$$z^2 + (1 + 2i)z + 3 + 3i = 0;$$

es gilt also $a = 1$, $b = 1 + 2i$ und $c = 3 + 3i$. Die Diskriminante ist

$$\Delta := b^2 - 4ac = -3 + 4i - 12 - 12i = -15 - 8i.$$

Wir bestimmen die Quadratwurzeln der Diskriminante mit Formel (1.5.a). Es gilt $|\Delta| = \sqrt{225 + 64} = \sqrt{289} = 17$. Damit sind die Quadratwurzeln von Δ

$$w_{1,2} = \pm\left(\sqrt{\frac{17-15}{2}} - \sqrt{\frac{17+15}{2}}\, i\right) = \pm(1 - 4i).$$

Die Lösungen der quadratischen Gleichung ergeben sich aus Satz 1.5.35 als

$$z_1 = -3i \quad \text{und} \quad z_2 = -1 + i.$$

Steffen Kionke

Lineare Algebra

Lektion 2:
Konzepte der Linearen Algebra

Fakultät für
Mathematik und Informatik

FernUniversität in Hagen

Studierhinweise zur zweiten Lektion

In der zweiten Lektion wenden wir uns der Linearen Algebra zu – also der Lehre von den Vektorräumen und den linearen Abbildungen. Hier werden Begriffe und Konzepte bereitgestellt, die wir im weiteren Verlauf einsetzen.

Die Lektion beginnt mit einem kleinen Exkurs über das Auswahlaxiom – also mit einem Ausflug in die Mengenlehre. Ich hoffe, dass Sie daraus etwas Hintergrundwissen mitnehmen, obwohl das Auswahlaxiom in der Linearen Algebra nur eine Nebenrolle spielt. (Das Auswahlaxiom verwenden wir in dieser Lektion zweimal und dann nie wieder.)

Als nächstes werden wichtige Begriffe aus den Mathematischen Grundlagen wiederholt und teilweise ergänzt oder zumindest aus einer anderen Perspektive präsentiert. Das ist eine gute Gelegenheit um die eigene Erinnerung aufzufrischen und in den Mathematischen Grundlagen nachzuschlagen. Versuchen Sie Lücken im Verständnis dieser Grundbegriffe schnell zu schließen, denn diese werden Ihnen später im Studium auf die Füße fallen!

Im Anschluss lernen wir drei fundamentale Konzepte der Linearen Algebra kennen: die *direkte Summe*, den *Faktorraum* und den *Dualraum*. Die direkte Summe ist eine Zerlegung eines Vektorraumes in kleinere Unterräume. Sie wird ein ständiger Begleiter sein und es ist wichtig, dass Sie sich damit gründlich auseinandersetzen. Der Faktorraum und der Dualraum sind für uns zunächst zwei „abstrakte“ Konstruktionen um aus einem Vektorraum neue Vektorräume zu bauen. Um ein gutes Verständnis dieser Konstruktionen zu erhalten, sollten Sie sich möglichst viele Beispiele anschauen.

Zum Abschluss führen wir noch einen nützlichen Begriff ein: *Äquivalenzrelationen*. Das Konzept der Äquivalenzrelation dient dazu „Ähnlichkeit“ zwischen Objekten zu formalisieren und man trifft wahrscheinlich in allen Gebieten der Mathematik darauf. Auf den ersten Blick wirken Äquivalenzrelationen wie abstrakter Unfug, aber wenn man das Konzept einmal verstanden hat, ist es wirklich hilfreich um Zusammenhänge zu benennen.

Lernziele

Nach dem Durcharbeiten der Lektion sollten Sie

- → wissen, was das *Lemma von Zorn* ist,
- → den *Basisergänzungssatz* kennen,

→ verstehen, was es heißt, dass ein Vektorraum die *direkte Summe* von Unterräumen ist,

→ den *Satz vom Komplement* verstehen,

→ mit der Konstruktion des *Faktorraumes* vertraut sein und die *Dimensionsformel* für Faktorräume kennen,

→ die Definition des *Dualraumes* kennen und erklären können, was eine *Dualbasis* und eine *duale Abbildung* ist,

→ wissen, was eine *Äquivalenzrelation* ist und Beispiele benennen können.

Literaturhinweise

Die Themen dieser Lektion finden Sie in den meisten Lehrbüchern zur Linearen Algebra, allerdings nicht immer in der Ausführlichkeit, die diese Themen aus meiner Sicht verdienen. Nachschlagen können Sie z.B. in

→ [Beu]: 1.2 (Äquivalenzrelationen), 3.3.4 (Faktorräume), 5.4 (Dualraum).

→ [Bo]: 1.6 (direkte Summe), 2.2 (Faktorraum, Äquivalenzrelationen), 2.3 (Dualraum).

→ [Fi]: 2.1 (Äquivalenzrelationen), 2.6 (direkte Summe), 3.2 (Faktorraum), 7.1 (Dualraum).

→ [Gö]: 3.5 & 4.2 (Faktorraum), 4.11 (Dualraum)

→ [KaSt]: 1.4 (Äquivalenzrelationen, Lemma von Zorn), 4.5 (direkte Summe, Faktorraum), 6.9 (Dualraum).

Der Faktorraum wird in [Bo, Fi, Gö] (und einigen anderen Quellen) „Quotientenvektorraum“ genannt.

Fahrplan durch die Lektion

2.1 → Diese Lektion beginnt in Abschnitt 2.1 mit einem Exkurs in die Mengenlehre. Für die Lineare Algebra müssen Sie hier nicht alle Details verstehen, aber ich hoffe, dass dieses Wissen Ihnen später im Studium weiterhilft. Jedenfalls ist dieser Abschnitt nicht klausurrelevant.

Wir erklären zuerst das *Auswahlaxiom* 2.1.2 und besprechen dann *partielle Ordnungen* (2.1.4) und zugehörige Begriffe (2.1.6–2.1.12). Damit formulieren wir das

Lemma von Zorn (2.1.13) und zeigen, wie man damit das Auswahlaxiom beweisen kann (2.1.14).

In Abschnitt 2.2 blicken wir auf zentrale Begriffe aus den Mathematischen Grundlagen zurück, mit denen Sie gut vertraut sein müssen, um dieses Modul zu meistern. Diesen Abschnitt können Sie wahrscheinlich zügig bearbeiten, wenn Sie des Modul „Mathematische Grundlagen“ erst vor kurzem absolviert haben. Im ersten Teil wiederholen wir die Grundbegriffe *Vektorraum* (2.2.1) und *Unterraum* (2.2.4). Im zweiten Teil befassen wir uns kurz mit linearer Unabhängigkeit (2.2.7), Basen (2.2.10) und der Dimension (2.2.14). Wir beweisen den wichtigen Basisergänzungssatz (2.2.11 und 2.2.12) in einer Version, die auch unendlich-dimensionale Vektorräume zulässt (dafür benötigen wir das Lemma von Zorn!). ← 2.2

Was eine *lineare Abbildung* (2.2.17) ist, wiederholen wir im dritten Teil. Insbesondere fassen wir nochmal zusammen, wie man mit linearen Abbildungen in Koordinaten (2.2.22) rechnen kann. Besonders wichtig sind *Matrixdarstellungen* (2.2.23), *Transformationsmatrizen* (2.2.26) und die *Transformationsformel* für den Basiswechsel (2.2.28)

In Abschnitt 2.3 lernen wir die *direkte Summe* kennen. Zuerst klären wir, was die Summe einer Familie von Unterräumen ist (2.3.1–2.3.4) und definieren dann, wann man eine Summe *direkt* nennt (2.3.5). Satz 2.3.6 enthält nützliche Charakterisierungen der direkten Summe. Eine Dimensionsformel behandeln wir in Satz 2.3.12. Im zweiten Teil untersuchen wir den Fall von zwei direkten Summanden genauer. Das führt uns zum Begriff des *Komplementes* (2.3.14). Wir beweisen, dass es zu jedem Unterraum ein Komplement gibt (2.3.17). ← 2.3

Im Abschnitt 2.4 lernen wir die Faktorräume kennen. Dabei konstruiert man aus einem Vektorraum V und einem Unterraum U einen neuen Vektorraum V/U: den *Faktorraum*. Die Konstruktion ist schwer bekömmlich, weil die Elemente von V/U selbst Teilmengen von V sind. Studierende, denen das Kopfzerbrechen bereitet, können sich am Beispiel der *Parallelenschar* festhalten, das wir immer wieder aufgreifen (2.4.1, 2.4.12, 2.4.15, 2.4.17). ← 2.4

Im ersten Teil definieren wir *affine Unterräume* (2.4.2) und zeigen, dass man diese addieren (2.4.8) und mit Skalaren multiplizieren kann (2.4.9). Wir zeigen, dass die Menge V/U der affinen Unterräume mit diesen Rechenoperationen ein Vektorraum ist (2.4.11) und, dass es eine kanonische lineare Abbildung $\pi\colon V \to V/U$ gibt (2.4.13). Im zweiten Teil überlegen wir, wie man Basen von V/U bestimmen kann (2.4.18). Damit leiten wir eine Formel für die Dimension von V/U her (2.4.22).

2.5 → Im Abschnitt 2.5 lernen wir den Dualraum kennen. Der Dualraum V^* eines Vektorraumes V ist der Vektorraum aller Linearformen auf V. Auch diese Konstruktion ist zunächst gewöhnungsbedürftig, aber wenn Sie sich trauen, damit zu arbeiten, werden Sie merken, dass hier nichts Tiefgründiges passiert.

Im ersten Teil definieren wir den Dualraum V^* (2.5.3). Wir besprechen, wie man aus einer Basis von V eine *duale Basis* von V^* erhält (2.5.10) und wie man Koordinaten der Dualbasis berechnen kann (2.5.13–2.5.15).

Die Konstruktion des Dualraumes ist sogar ein „Funktor", d.h., man erhält aus Abbildungen zwischen Vektorräumen automatisch Abbildungen zwischen den Dualräumen: die sogenannten *dualen Abbildungen*. Diese definieren wir im zweiten Teil (2.5.16) und zeigen, dass es sich um lineare Abbildungen handelt (2.5.18). Zum Abschluss beschreiben wir einen Zusammenhang zwischen den Matrixdarstellungen einer Abbildung und ihrer dualen Abbildung (2.5.20).

2.6 → Im letzten Abschnitt 2.6 führen wir *Äquivalenzrelationen* ein (2.6.1) und lernen verschiedene Beispiele kennen (2.6.2–2.6.5). Wir definieren *Äquivalenzklassen* (2.6.6) und stellen fest, dass je zwei Äquivalenzklassen entweder gleich oder disjunkt sind (2.6.8). Im zweiten Teil besprechen wir ein Beispiel aus der Linearen Algebra, das Sie im Wesentlichen schon aus den Mathematischen Grundlagen kennen.

2.1. Vorwort zur Mengenlehre: das Auswahlaxiom*

Dieser Studienbrief behandelt in erster Linie die Theorie der endlich-dimensionalen Vektorräume. Einige fundamentale Ergebnisse sind aber auch für Vektorräume von unendlicher Dimension gültig und können fast auf dieselbe Weise bewiesen werden. Dazu benötigen wir aber ein Hilfsmittel aus der Mengenlehre: das Auswahlaxiom. Diesen Zusammenhang zur Mengenlehre werden wir hier kurz vorstellen, weil Sie dem Auswahlaxiom im Laufe Ihres Studiums immer wieder begegnen werden. In der Klausur gibt es aber keine Fragen zum Auswahlaxiom.

In den Mathematischen Grundlagen haben Sie bereits die Grundbegriffe der Mengenlehre kennengelernt. Die Grundlage ist dabei ein „naiver" Mengenbegriff, wie ihn Georg Cantor[1] bereits 1895 in [Can95] formulierte:

> Unter einer Menge verstehen wir jede Zusammenfassung M von bestimmten wohlunterschiedenen Objekten m unserer Anschauung [...] zu einem Ganzen.

Natürlich ist diese Definition für die moderne Mathematik etwas zu vage. Außerdem hat sich bald herausgestellt, dass ein solcher „naiver" Mengenbegriff in sich widersprüchlich ist. Das bekannteste Beispiel dafür ist das Paradoxon von Bertrand Russell[2].

2.1.1 **Russell'sches Paradoxon**. Es sei

$$R = \{X \mid X \notin X\}$$

die Menge aller Mengen, die sich nicht selbst enthalten. Ist R ein Element von R?

Wäre $R \in R$ dann gilt nach Definition von R aber $R \notin R$. Widerspruch! Aber auch umgekehrt erhält man einen Widerspruch: Ist $R \notin R$, dann müsste R nach Definition ein Element von R sein.

Das Russell'sche Paradoxon wurde dadurch aufgelöst, dass die Mengenlehre *axiomatisiert* wurde. Das heißt, man hat genauer definiert, was eine Menge ist und welche Operationen mit Mengen erlaubt sind. Mit diesen erlaubten Operationen kann das seltsame Objekt R von Russell nicht mehr definiert werden und das Russell'sche Paradoxon verschwindet.

Die heute gebräuchlichen Axiome der Mengenlehre gehen auf Zermelo[3] und Fraenkel[4]

[1] Georg CANTOR: deutscher Mathematiker, 1845–1918. Begründer der Mengenlehre.
[2] Bertrand RUSSELL: britischer Philosoph und Mathematiker, 1872–1970.
[3] Ernst ZERMELO: deutscher Mathematiker und Logiker, 1871–1953.
[4] Abraham FRAENKEL: deutsch-israelischer Mathematiker, 1891–1965.

zurück. Es ist für unsere Zwecke nicht nötig, diese Axiome im Detail anzugeben; wer es genau wissen möchte, kann einen Blick in [Jec03] werfen. Die meisten Axiome der Zermelo-Fraekel Mengenlehre sind so naheliegend, dass man darüber wenig sagen kann (z.B. zwei Mengen sind gleich, wenn sie dieselben Elemente haben). Es gibt darunter aber ein Axiom, das man sich einmal genauer ansehen muss, weil es einige überraschende Konsequenzen hat.

2.1.2 **Auswahlaxiom.** Ist $f\colon X \to Y$ eine surjektive Abbildung zwischen Mengen, dann existiert eine Abbildung $g\colon Y \to X$ mit $f \circ g = \mathrm{id}_Y$.

Man sagt, die Abbildung g ist *rechtsinvers* zu f. Man kann g auch eine *Auswahlfunktion* nennen: Für jedes $y \in Y$ ist $g(y)$ ein Element im Urbild $f^{-1}(y)$. Da f surjektiv ist, ist das Urbild $f^{-1}(y) \neq \emptyset$ für alle $y \in Y$. Die Funktion g *wählt* also aus jeder der Mengen $f^{-1}(y)$ ein Element aus.

Die Aussage des Auswahlaxioms scheint auf den ersten Blick offensichtlich zu sein. Natürlich ist es möglich aus *einer* nicht-leeren Menge ein Element zu wählen. Auch endlich viele Wahlen kann man problemlos treffen. Ist also Y eine endliche Menge, dann kann man die Existenz einer Auswahlfunktion auch ohne das Auswahlaxiom beweisen. Das Seltsame ist, dass man das Auswahlaxiom wirklich benötigt[5], wenn Y eine unendliche Menge ist, d.h., wenn man unendliche viele Wahlen treffen muss.

Das Auswahlaxiom wird häufig verwendet, um die Existenz von Objekten zu beweisen, die man nicht explizit angeben kann (z.B. eine Basis eines Vektorraumes von unendlicher Dimension). Dadurch wurde das Auswahlaxiom von Verfechtern einer „konstruktiven“ Mathematik manchmal kritisch beäugt.

Es gibt viele nützliche Resultate, die äquivalent zum Auswahlaxiom sind, d.h., man kann diese mit dem Auswahlaxiom beweisen und umgekehrt das Auswahlaxiom auch wieder daraus herleiten. Hier wollen wir mit dem *Lemma von Zorn* eine solche Aussage vorstellen. Dazu benötigen wir einige Begriffe.

2.1.3 **Relationen**. Erinnern wir uns zunächst, dass eine binäre *Relation* R auf einer Menge M eine Teilmenge $R \subseteq M \times M$ ist. Zwei Elemente $x, y \in M$ weisen die „Beziehung“ R auf, wenn (x, y) in R liegt. Wir schreiben dann xRy anstelle von $(x, y) \in R$. Meistens verwenden wir daher für binäre Relationen keine Buchstaben, sondern Symbole wie $\leq, <, \preceq, \sim, \approx, \ldots$.

2.1.4 **Definition** Sei M eine Menge. Eine binäre Relation $\preceq$ auf M heißt *partielle Ordnung*, wenn folgende drei Eigenschaften erfüllt sind:

[5] Genauer heißt das: Das Auswahlaxiom lässt sich nicht aus den anderen Axiomen der Mengenlehre herleiten (zumindest sofern nicht die anderen Axiome in sich widerspüchlich sind).

(PO1) $\preceq$ ist *reflexiv*, d.h., für alle x in M gilt

$$x \preceq x.$$

(PO2) $\preceq$ ist *antisymmetrisch*, d.h., für alle $x, y \in M$ gilt

$$x \preceq y \text{ und } y \preceq x \quad \Longrightarrow \quad x = y.$$

(PO3) $\preceq$ ist *transitiv*, d.h., für alle $x, y, z \in M$ gilt

$$x \preceq y \text{ und } y \preceq z \quad \Longrightarrow \quad x \preceq z.$$

Eine partielle Ordnung $\preceq$ nennt man *Totalordnung*, wenn für alle $x, y \in M$ immer $x \preceq y$ oder $y \preceq x$ gilt.

2.1.5 **Beispiel**. (a) Die Relation $\leq$ auf den reellen Zahlen ist eine Totalordnung; siehe Abschnitt 12.1 in [MG].

(b) Es sei X eine Menge. Die Menge aller Teilmengen von X nennt man die *Potenzmenge* $\mathcal{P}(X)$ von X. Die Enthaltungsrelation $\subseteq$ ist eine partielle Ordnung auf der Potenzmenge. Hat X mindestens zwei Elemente $a \neq b$, dann ist dies keine Totalordnung, denn keine der beiden Mengen $\{a\}$ und $\{b\}$ enthält die andere.

(c) Für zwei natürliche Zahlen $m, n \in \mathbb{N}$ scheiben wir $m \mid n$, wenn m die Zahl n teilt. Die Teilerrelation „|“ ist eine partielle Ordnung auf $\mathbb{N}$. Die Teilerrelation ist keine Totalordnung, z.B., teilt 2 nicht die Zahl 3 und 3 teilt nicht die Zahl 2.

2.1.6 **Definition** Es sei M eine Menge mit einer partiellen Ordnung $\preceq$. Ein Element $m \in M$ heißt *maximal*, wenn aus $m \preceq x$ bereits $m = x$ folgt für alle $x \in M$. Sei $N \subseteq M$ eine Teilmenge. Eine *obere Schranke* für N ist ein Element $s \in M$ mit der Eigenschaft

$$x \preceq s$$

für alle $x \in N$.

2.1.7 **Vorsicht mit dem Begriff „maximal“**. Den Begriff „maximal“ muss man sich hier genau anschauen. Ein Element m ist maximal, wenn es *keine* Elemente gibt, die echt größer als m sind. Ein maximales Element muss *nicht* größer als alle anderen Elemente sein. Ist $\preceq$ die partielle Ordnung auf der Menge $M = \{x, y\}$, sodass weder $x \preceq y$ noch $y \preceq x$ gelten, dann sind beide Elemente x und y maximal. Insbesondere kann eine partiell geordnete Menge mehrere verschiedene maximale Elemente besitzen.

2.1.8 **Aufgabe**. Es sei M eine Menge mit einer Totalordnung $\preceq$. Zeigen Sie, dass für beliebige $x_1, \ldots, x_k \in M$ ein $j \in \{1, \ldots, k\}$ existiert, sodass **L**

$$x_i \preceq x_j$$

für alle $i \in \{1, 2, \ldots, k\}$ gilt.

2.1.9 **Beispiel**. (a) In den reellen Zahlen mit der Relation $\leq$ gibt es keine maximalen Elemente. Die Zahl 5 ist eine obere Schranke für das Intervall $(-2, 3]$. Auch 3 und 500 sind obere Schranken für $(-2, 3]$.

(b) Es sei X eine Menge. Die Menge X ist ein maximales Element bezüglich Inklusion in der Potenzmenge $\mathcal{P}(X)$.

(c) In den natürlichen Zahlen mit der Teilerrelation „|“ gibt es keine maximalen Elemente. Die Zahl 30 ist eine obere Schranke an die Menge $\{2, 3, 5\}$.

2.1.10 **Induzierte Ordnungen**. Sei M eine Menge mit einer partiellen Ordnung $\preceq$ und sei $N \subseteq M$ eine Teilmenge. Schränkt man $\preceq$ zu einer Relation auf N ein, so erhält man eine partielle Ordnung auf N. In der Tat, sind die Eigenschaften aus Definition 2.1.4 für alle $x, y, z \in M$ erfüllt, dann natürlich auch für alle x, y, z in der Teilmenge N. Man sagt, die partielle Ordnung auf N wurde von der Ordnung auf M *induziert*.

Ist $\preceq$ eine Totalordnung, dann ist die induzierte Ordnung auf einer Teilmenge N ebenfalls eine Totalordnung.

2.1.11 $\boxed{\textbf{Definition}}$ Sei M eine Menge mit einer partiellen Ordnung $\preceq$. Eine *Kette* K in M ist eine Teilmenge $K \subseteq M$, auf der die induzierte Ordnung eine Totalordnung ist. Das heißt, für alle $x, y \in K$ gilt $x \preceq y$ oder $y \preceq x$.

2.1.12 **Beispiel**. (a) In einer totalgeordneten Menge ist jede Teilmenge eine Kette.

(b) In der Potenzmenge $\mathcal{P}(\mathbb{N})$ ist die Teilmenge

$$K_1 = \Big\{\{1\}, \{1, 3\}, \{1, 3, 6, 400\}, \{1, 3, 6, 7, 400\}\Big\}$$

eine Kette bzgl. der Inklusion $\subseteq$. Genauso ist auch

$$K_2 = \Big\{\{n^2 \mid n \in \mathbb{N}, n \leq m\} \mid m \in \mathbb{N}\Big\}$$

eine Kette bzgl. der Inklusion.

Mit diesen Begriffen können wir das Lemma von Zorn formulieren.

2.1.13 **Lemma von Zorn.** *Sei M eine nicht-leere Menge mit einer partiellen Ordnung $\preceq$. Besitzt jede Kette in M eine obere Schranke in M, dann hat M (mindestens) ein maximales Element.*

Wie beweist man das Lemma von Zorn? Die Grundidee ist einfach: Wenn M keine maximalen Elemente hat, dann sollte man in der Lage sein eine Kette zu konstruieren, die keine obere Schranke besitzt. Um diese Kette zu finden, verwendet man das Auswahlaxiom. Die kurzen Beweise verwenden dabei transfinite Methoden, die wir hier nicht zur Verfügung haben. Es gibt auch elementare Beweise, aber auch diese Beweise erfordern zusätzliche Terminologie, weshalb wir hier darauf verzichten. Die interessierten Leserinnen und Leser finden machbare Beweise des Lemmas in [Kne50] oder [Lew91].

Zum Abschluss wollen wir stattdessen erklären, wie man umgekehrt das Auswahlaxiom aus dem Lemma von Zorn herleiten kann. Anhand des Beweises kann man gut verstehen, wie das Lemma eingesetzt wird.

2.1.14 **Beweis des Auswahlaxioms mit dem Lemma von Zorn**.

Beweis. Sei $f\colon X \to Y$ eine surjektive Abbildung. Wir definieren die Menge M als Menge aller Paare (U, g), wobei $U \subseteq Y$ eine Teilmenge von Y ist und $g\colon U \to X$ eine Abbildung mit $f \circ g = \mathrm{id}_U$ ist. Die Menge M ist nicht leer, denn für ein $y \in Y$ und ein Urbild $x \in f^{-1}(y)$ ist $(\{y\}, y \mapsto x)$ ein Element von M.

Auf M definieren wir nun eine partielle Ordnung $\preceq$. Wir setzen $(U, g) \preceq (U', g')$ genau dann, wenn

$$U \subseteq U' \quad \text{und} \quad g'|_U = g$$

erfüllt sind. Wir behaupten, dass jede Kette K in M eine obere Schranke besitzt. Sei K eine Kette. Dann definieren wir eine Teilmenge $V \subseteq Y$ durch

$$V = \bigcup_{(U,g)\in K} U.$$

Auf V definieren wir die Abbildung $h\colon V \to X$ wie folgt: Für jedes $v \in V$ gibt es ein $(U, g) \in K$ mit $v \in U$ und wir setzen $h(v) := g(v)$. Weil K eine Kette ist, muss man dazu keine Wahl treffen: Ist (U', g') in K mit $v \in U'$, dann gilt $(U, g) \preceq (U', g')$ oder $(U', g') \preceq (U, g)$; also $g(v) = g'(v)$. Es gilt nun $f \circ h = \mathrm{id}_V$, denn für $v \in U$ mit $(U, g) \in K$ gilt $f(h(v)) = f(g(v)) = v$. Also ist $(V, h) \in M$ eine obere Schranke an K.

Durch das Lemma von Zorn muss es in M ein maximales Element (U, g) geben. Wir behaupten, dass $U = Y$ gilt und g damit die gesuchte Auswahlfunktion ist. Nehmen wir für einen Widerspruch an, es sei $U \neq Y$. Dann gibt es ein Element $y \in Y \setminus U$. Wir wählen ein Urbild $x \in f^{-1}(y)$ und definieren $U' = U \cup \{y\}$ und $g' \colon U' \to X$ durch $g'(u) = g(u)$ für alle $u \in U$ und $g'(y) = x$. Dann gilt $(U, g) \preceq (U', g')$ und $(U, g) \neq (U', g')$. Dies ist ein Widerspruch zur Maximalität von (U, g). $\square$

Das beendet unseren kurzen Exkurs in die Mengenlehre. Lassen Sie sich von diesem knappen Überblick nicht entmutigen. Das Lemma von Zorn ist auf den ersten Blick nur schwer zu fassen und es wird auch nur in wenigen Beweisen in dieser Lektion eine Rolle spielen.

2.2. Rückblick und Ergänzungen: Vektorräume und Basen

I. Vektorräume

Der Begriff des Vektorraumes, den Sie aus dem Modul „Mathematische Grundlagen“ kennen, ist für uns von zentraler Bedeutung. Deshalb wiederholen wir hier einige Begriffe und erinnern an wichtige Eigenschaften. Alle Details finden Sie in den „Mathematischen Grundlagen“ [MG, Kapitel 6].

2.2.1 **Vektorräume**. Sei K ein Körper. Ein K-*Vektorraum* V ist eine Menge mit zwei Verknüpfungen: einer Addition

$$V \times V \to V \quad \text{mit} \quad (v, w) \mapsto v + w$$

und einer Skalarmultiplikation

$$K \times V \to V \quad \text{mit} \quad (a, v) \mapsto av.$$

Dabei sollen folgende Axiome erfüllt sein:

(i) $(V, +)$ ist eine abelsche Gruppe.

(ii) Für alle $a, b \in K$ und $v \in V$ gilt $(ab)v = a(bv)$. Ist 1 das Einselement in K, dann gilt $1v = v$ für alle $v \in V$.

(iii) Es gelten die Distributivgesetze

$$\begin{aligned} (a + b)v &= av + bv \\ a(v + w) &= av + aw \end{aligned}$$

für alle $a, b \in K$ und $v, w \in V$.

Das neutrale Element der Addition 0 heißt *Nullvektor*.

2.2.2 **Beispiel**. Das wichtigste Beispiel eines Vektorraumes ist der Raum K^n der Spaltenvektoren mit n Einträgen aus K. Bei der Addition in K^n werden die Einträge komponentenweise addiert:

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ \vdots \\ v_n + w_n \end{pmatrix}.$$

Bei der Skalarmultiplikation mit $a \in K$ wird jeder Eintrag im Vektor v mit a multipliziert:

$$a \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} av_1 \\ av_2 \\ \vdots \\ av_n \end{pmatrix}.$$

2.2.3 **Beispiel**. Sei V ein K-Vektorraum und sei X eine Menge. Es sei $\mathrm{Abb}(X,V)$ die Menge aller Abbildungen von X nach V. Auf $\mathrm{Abb}(X,V)$ definieren wir eine Addition „punktweise", das heißt

$$(f+g)(x) \coloneqq f(x) + g(x)$$

für alle $x \in X$. Dabei verwenden wir auf der rechten Seite die Addition im Vektorraum V. Genauso kann man eine Skalarmultiplikation $K \times \mathrm{Abb}(X,V) \to \mathrm{Abb}(X,V)$ „punktweise" definieren durch

$$(cf)(x) \coloneqq cf(x).$$

für alle $x \in X$ und $c \in K$. Bezüglich dieser Rechenoperationen ist $\mathrm{Abb}(X,V)$ ein K-Vektorraum. Durch die punktweise Definition lassen sich alle Axiome direkt aus den Axiomen von V herleiten. Zur Illustration schauen wir uns das Assoziativgesetz der Addition an. Für alle $f,g,h \in \mathrm{Abb}(X,V)$ und $x \in X$ gilt

$$\begin{aligned} ((f+g)+h)(x) &= (f+g)(x) + h(x) \\ &= (f(x)+g(x)) + h(x) \\ &= f(x) + (g(x)+h(x)) && \text{(Assoz. von + in V)} \\ &= f(x) + (g+h)(x) = (f+(g+h))(x). \end{aligned}$$

Weil diese Gleichung für alle $x \in X$ gilt, folgt daraus $(f+g)+h = f+(g+h)$. Alle anderen Axiome kann man auf dieselbe Weise herleiten.

2.2.4 **Unterräume**. Sei K ein Körper und sei V ein K-Vektorraum. Eine Teilmenge $U \subseteq V$ heißt *Unterraum* von V, wenn folgende drei Bedingungen erfüllt sind:

(a) Der Nullvektor liegt in U.

(b) Für alle $u, u' \in U$ gilt $u + u' \in U$.

(c) Für alle $u \in U$ und $a \in K$ gilt $au \in U$.

In diesem Fall kann man die Addition und die Skalarmultiplikation von V auf U einschränken und U ist damit ebenfalls ein K-Vektorraum; siehe [MG, 6.2.3].

2.2.5 **Beispiel**. Im reellen Vektorraum $\mathbb{R}^2$ ist

$$U = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mid x_1 + x_2 = 0 \right\} \subseteq \mathbb{R}^2$$

ein Unterraum.

Die folgende Aufgabe enthält einen kleinen Trick, wie man schnell nachprüfen kann, ob eine nicht-leere Teilmenge ein Unterraum ist.

2.2.6 **Aufgabe**. Sei V ein K-Vektorraum und sei $U \subseteq V$ eine nicht-leere Teilmenge. Zeigen Sie: U ist genau dann ein Unterraum, wenn **L**

$$au + u' \in U$$

für alle $a \in K$ und $u, u' \in U$ gilt.

II. Basen

2.2.7 $\boxed{\textbf{Definition}}$ Sei V ein K-Vektorraum. Eine Teilmenge $S \subseteq V$ heißt *linear unabhängig*, wenn jede endliche Teilmenge linear unabhängig ist. Das heißt: sind $v_1, \ldots, v_n \in S$ paarweise verschieden und erfüllen $a_1, \ldots, a_n \in K$ die Gleichung

$$\sum_{i=1}^{n} a_i v_i = 0,$$

dann gilt $a_i = 0$ für alle i.

2.2.8 **Wie prüft man lineare Unabhängigkeit**? Wie kann man für gegebene Vektoren $v_1, \ldots, v_k \in K^n$ entscheiden, ob sie linear unabhängig sind? Dazu kann man die Vektoren als Spalten in eine Matrix A schreiben und den Rang mit dem Gauß-Verfahren berechnen. Dann gilt:

$$v_1, \ldots, v_k \text{ linear unabhängig} \quad \Leftrightarrow \quad \operatorname{Rg}(A) = k$$

Das folgt aus [MG, Korollar 9.2.5].

Tipp zum Rechnen: Wir werden am Ende dieser Lektion sehen, dass $\operatorname{Rg}(A^T) = \operatorname{Rg}(A)$ gilt (2.6.16).

2.2.9 **Aufgabe**. Sind die Vektoren **L**

$$v_1 = \begin{pmatrix} \overline{1} \\ \overline{2} \\ \overline{6} \\ \overline{2} \end{pmatrix}, \quad v_2 = \begin{pmatrix} \overline{3} \\ \overline{2} \\ \overline{1} \\ \overline{6} \end{pmatrix}, \quad v_3 = \begin{pmatrix} \overline{4} \\ \overline{3} \\ \overline{1} \\ \overline{1} \end{pmatrix}$$

linear unabhängig in $\mathbb{F}_7^4$?

2.2.10 **Definition** Sei V ein K-Vektorraum. Ein linear unabhängiges Erzeugendensystem $\mathcal{B} \subseteq V$ nennt man *Basis* von V.

Aus den Mathematischen Grundlagen wissen Sie bereits, dass jeder endlich erzeugte Vektorraum eine Basis besitzt. Mit Lemma von Zorn (und damit indirekt mit dem Auswahlaxiom) können wir dieses Resultat nun auch für beliebige Vektorräume zeigen. Wir formulieren zuerst ein allgemeines Ergebnis.

2.2.11 **Satz.** *Sei V ein K-Vektorraum mit Teilmengen $S \subseteq E$. Ist S linear unabhängig und E ein Erzeugendensystem von V, dann gibt es eine Basis $\mathcal{B}$ von V mit*

$$S \subseteq \mathcal{B} \subseteq E.$$

Beweis. Wir wollen mit dem Lemma von Zorn 2.1.13 argumentieren und definieren dazu die Menge M aller linear unabhängigen Teilmengen $C \subseteq E$, die S enthalten. Wegen $S \in M$ ist M nicht leer. Die Inklusionsrelation „$\subseteq$" induziert eine partielle Ordnung auf der Menge M.

Wir behaupten, dass jede Kette in M eine obere Schranke in M besitzt. Sei dazu $\mathcal{C}$ eine Kette in M. Ist $\mathcal{C}$ leer, so ist $S \in M$ eine obere Schranke. Wir nehmen nun an, dass $\mathcal{C}$ nicht leer ist und definieren D als die Vereinigung aller Kettenglieder, d.h.,

$$D = \bigcup_{C \in \mathcal{C}} C.$$

Wir wollen zeigen, dass D eine obere Schranke für $\mathcal{C}$ ist. Dazu müssen wir zeigen, dass D in M liegt. Da alle $C \in \mathcal{C}$ in E enthalten sind, ist sicherlich auch D eine Teilmenge von E. Da $\mathcal{C}$ nicht leer ist, gilt auch $S \subseteq D$. Wir prüfen nun, dass D linear unabhängig ist. Seien $v_1, \ldots, v_n \in D$ paarweise verschieden und seien $a_1, \ldots, a_n \in K$ mit

$$\sum_{i=1}^{n} a_i v_i = 0.$$

Da v_i in D liegt, gibt es eine Menge $C_i \in \mathcal{C}$ mit $v_i \in C_i$. Da $\mathcal{C}$ total geordnet ist, gibt es ein j mit $C_i \subseteq C_j$ für alle $i \in \{1, 2 \ldots, n\}$; siehe Aufgabe 2.1.8. Da C_j linear unabhängig ist, schließen wir $a_i = 0$ für alle $i \in \{1, 2, \ldots, n\}$. Insgesamt ist also auch D linear unabhängig.

Nach dem Lemma von Zorn 2.1.13 gibt es (mindestens) ein maximales Element $\mathcal{B} \in M$. Wir behaupten, dass $\mathcal{B}$ eine Basis von V ist. Nach Definition von M ist $\mathcal{B}$ linear unabhängig. Wir müssen also nachweisen, dass $\mathcal{B}$ auch ein Erzeugendensystem von V ist. Es sei $U = \langle \mathcal{B} \rangle$, der von $\mathcal{B}$ erzeugte Unterraum von V. Für einen

Widerspruch nehmen wir $U \neq V$ an. Da E ein Erzeugendensystem von V ist, liegt E nicht im Unterraum U (sonst wäre $\langle E \rangle \subseteq U$). Wir finden also ein Element $v \in E$ mit $v \notin U$. Wir definieren

$$C = \mathcal{B} \cup \{v\}$$

und behaupten, dass C linear unabhängig ist. Dies liefert den gewünschten Widerspruch zur Maximalität von $\mathcal{B}$, denn dann wäre $C \in M$ und $\mathcal{B} \subsetneq C$.

Es seien $v_1, \ldots, v_n \in \mathcal{B}$ paarweise verschieden und $a_1, \ldots, a_n, b \in K$ mit

$$bv + \sum_{i=1}^{n} a_i v_i = 0$$

gegeben. Wir zeigen, dass alle Koeffizienten a_i und b verschwinden. Es ist $b = 0$, denn andernfalls könnten wir die Gleichung nach v auflösen und damit wäre

$$v = -\sum_{i=1}^{n} b^{-1} a_i v_i$$

ein Element von U; wegen $v \notin U$ kann das nicht sein. Aus der verbliebenen Gleichung $\sum_{i=1}^{n} a_i v_i = 0$ folgt aber $a_1 = \cdots = a_n = 0$, denn $\mathcal{B}$ ist linear unabhängig. Damit ist C linear unabhängig und wir haben den gewünschten Widerspruch erzeugt. □

Aus diesem allgemeinen Ergebnis können wir nun den nützlichen Basisergänzungssatz ableiten.

2.2.12 Basisergänzungssatz. *Sei V ein K-Vektorraum.*

(a) Jede linear unabhängige Teilmenge $S \subseteq V$ kann zu einer Basis von V ergänzt werden.

(b) Jedes Erzeugendensystem $E \subseteq V$ enthält eine Basis von V.

Beweis. Beide Aussagen folgen direkt aus Satz 2.2.11. Für Aussage (a) setzt man $E = V$, für Aussage (b) setzt man $S = \emptyset$. □

Insbesondere folgt aus (b) mit $E = V$ folgendes wichtiges Ergebnis.

2.2.13 Korollar. *Jeder Vektorraum besitzt eine Basis.*

Man kann zeigen, dass je zwei Basen eines Vektorraumes immer dieselbe Anzahl an Elementen haben; man sagt auch, sie haben dieselbe *Kardinalität*. Für endlich erzeugte Vektorräume kennen Sie diese Aussage aus 7.3.6 in [MG]. Deshalb ist folgende Definition sinnvoll.

2.2.14 **Definition** Sei V ein K-Vektorraum. Die Kardinalität einer Basis $\mathcal{B}$ nennt man die *Dimension von V*;

$$\dim_K(V) = |\mathcal{B}|.$$

Einen Vektorraum, der eine endliche Basis besitzt, nennt man *endlich-dimensional*.

2.2.15 **Bemerkung zu unendlich-dimensionalen Vektorräumen**. Besitzt der Vektorraum V keine endliche Basis, dann sagt man, dass V *unendlich-dimensonal* ist. Wir schreiben schlicht $\dim_K(V) = \infty$. Das ist eigentlich etwas ungenau, denn man kann in der Mengenlehre verschiedene Abstufungen von „unendlich" definieren und die Dimension unendlich-dimensionaler Vektorräume genauer messen.

Bei endlich-dimensionalen Vektorräumen wird es oft nötig sein den Elementen einer Basis eine feste Reihenfolge zuzuweisen. Wir sprechen dann von einer geordneten Basis.

2.2.16 **Definition** Sei V ein endlich-dimensionaler K-Vektorraum. Ein Tupel $\mathcal{B} = (v_1, \dots, v_n)$ heißt *geordnete Basis*, wenn die Einträge paarweise verschieden sind und die Menge $\{v_1, \dots, v_n\}$ eine Basis von V ist.

III. Lineare Abbildungen

2.2.17 **Definition** Seien U und V zwei K-Vektorräume. Eine Abbildung $\varphi\colon U \to V$ nennt man *K-linear*, wenn folgende Bedingungen erfüllt sind:

(L1) $\varphi(u_1 + u_2) = \varphi(u_1) + \varphi(u_2)$ für alle $u_1, u_2 \in U$

(L2) $\varphi(au) = a\varphi(u)$ für alle $a \in K$ und $u \in U$

Wenn der zugrundeliegende Körper K aus dem Kontext klar hervorgeht, sagen wir kurz, dass φ linear ist.

2.2.18 **Aufgabe**. Sei $\varphi\colon U \to V$ eine Abbildung zwischen zwei K-Vektorräumen. Zeigen Sie, dass φ genau dann linear ist, wenn für alle $u_1, u_2 \in U$ und alle $a \in K$ die Gleichung L

$$\varphi(au_1 + u_2) = a\varphi(u_1) + \varphi(u_2) \qquad (2.2.a)$$

gilt.

2.2.19 **Beispiel**. Wir betrachten die Vektorräume K^n und K^m. Ist $A \in \mathrm{M}_{m,n}(K)$ eine Matrix, dann ist

$$f_A\colon K^n \to K^m \quad \text{mit} \quad f_A(x) = Ax$$

eine K-lineare Abbildung. Jede lineare Abbildung zwischen K^n und K^m ist von dieser Form; vgl. [MG, Satz 9.1.9].

2.2.20 **Der Kern einer linearen Abbildung**. Sei $\varphi\colon U \to V$ eine K-lineare Abbildung. Man nennt

$$\mathrm{Ker}(\varphi) = \{u \in U \mid \varphi(u) = 0\}$$

den *Kern* von φ (engl.: *kernel*). Der Kern einer linearen Abbildung ist ein Unterraum. Die Abbildung φ ist genau dann injektiv, wenn $\mathrm{Ker}(\varphi) = \{0\}$ gilt. Das sieht man so: Es gilt $\varphi(u) = \varphi(u')$ mit $u, u' \in U$ genau dann, wenn $\varphi(u - u') = 0$ ist; d.h., wenn $u - u'$ im Kern liegt (vgl. [MG, 8.3.10]).

Ist $A \in \mathrm{M}_{m,n}(K)$, dann schreiben wir kurz $\mathrm{Ker}(A) := \mathrm{Ker}(f_A)$, d.h.,

$$\mathrm{Ker}(A) = \{x \in K^n \mid Ax = 0\}.$$

Der Unterraum $\mathrm{Ker}(A)$ wird manchmal auch *Nullraum* der Matrix A genannt.

2.2.21 $\boxed{\textbf{Definition}}$ Eine bijektive K-lineare Abbildung $\varphi\colon U \to V$ nennt man *Isomorphismus*. Die Umkehrabbildung φ^{-1} ist ebenfalls wieder linear; siehe [MG, 8.1.8].

Zwei Vektorräume U und V nennt man *isomorph*, wenn es einen Isomorphismus zwischen U und V gibt; man schreibt dann $U \cong V$.

Ist $U \cong V$ und $V \cong W$, dann auch $U \cong W$, denn die Verkettung von Isomorphismen ist wieder ein Isomorphismus; siehe [MG, 8.1.10].

2.2.22 **Koordinaten**. Sei U ein endlich-dimensionaler K-Vektorraum und sei $\mathcal{B} = (u_1, \ldots, u_n)$ eine geordnete Basis. Jeder Vektor $v \in U$ lässt sich eindeutig als Linearkombination dieser Basisvektoren schreiben:

$$v = \sum_{i=1}^{n} a_i u_i$$

mit $a_i \in K$. Man nennt $\kappa_{\mathcal{B}}(v) = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ den *Koordinatenvektor* von v bzgl. $\mathcal{B}$. Die Abbildung $\kappa_{\mathcal{B}}\colon U \to K^n$ mit $v \mapsto \kappa_{\mathcal{B}}(v)$ ist ein Isomorphismus; siehe [MG, 8.2.2]

2.2.23 **Matrixdarstellung einer linearen Abbildung**. Es seien U, V zwei K-Vektorräume von endlicher Dimension und es sei $\varphi\colon U \to V$ eine K-lineare Abbildung. Sei $\mathcal{B} = (u_1, \ldots, u_n)$ eine geordnete Basis von U und sei $\mathcal{C} = (v_1, \ldots, v_m)$ eine geordnete Basis von V. Für jedes $j \in \{1, 2, \ldots, n\}$ kann man den Vektor $\varphi(u_j)$ als Linearkombination in der Basis $\mathcal{C}$ schreiben:

$$\varphi(u_j) = \sum_{i=1}^{m} a_{ij} v_i.$$

Die $(m \times n)$-Matrix ${}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\varphi) = (a_{ij})_{i,j}$ nennt man die *Matrixdarstellung* von φ bezüglich $\mathcal{B}$ und $\mathcal{C}$.

Die lineare Abbildung φ ist eindeutig durch die Bilder der Basisvektoren aus $\mathcal{B}$ festgelegt; siehe [MG, 8.4.1]. Weil man aus den Einträgen in der Matrix ${}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\varphi)$ ablesen kann, wohin φ die Basisvektoren abbildet, ist die lineare Abbildung φ eindeutig durch ihre Matrixdarstellung bestimmt.

Für alle $u \in U$ gilt die Beziehung (siehe [MG, 9.2.1])

$$ {}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\varphi)\ \kappa_{\mathcal{B}}(u) = \kappa_{\mathcal{C}}(\varphi(u)); \tag{2.2.b}$$

das heißt, die Matrixdarstellung beschreibt die Abbildung φ in den Koordinaten bzgl. $\mathcal{B}$ und $\mathcal{C}$.

2.2.24 **Beispiel**. Sei K ein Körper. Sei e_i der Vektor in K^n, dessen Einträge alle verschwinden außer einem Eintrag 1 an der Stelle i. Die geordnete Basis $(e_1, e_2, \ldots, e_n)$ nennen wir die *Standardbasis* von K^n. Für $A \in \mathrm{M}_{m,n}(K)$ betrachten wir die zugehörige lineare Abbildung $f_A\colon K^n \to K^m$ mit $f_A(x) = Ax$; siehe 2.2.19. Ist $\mathcal{B}$ die Standardbasis von K^n und $\mathcal{C}$ die Standardbasis von K^m, dann gilt

$${}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(f_A) = A,$$

denn $f_A(e_j) = Ae_j = \sum_{i=1}^{m} a_{ij} e_i$.

Das folgende nützliche Lemma versteckt sich in den Ergebnissen aus [MG].

2.2.25 **Lemma.** *Sei $\varphi\colon U \to V$ eine lineare Abbildung zwischen zwei endlich-dimensionalen Vektorräumen und seien $\mathcal{B}$ bzw. $\mathcal{C}$ geordnete Basen von U bzw. V.*

Dann ist φ genau dann ein Isomorphismus, wenn ${}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\varphi)$ eine invertierbare Matrix ist.

Beweis. Sei φ ein Isomorphismus mit Umkehrabbildung φ^{-1}. Dann gilt nach [MG, Aufgabe 9.3.2] ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{C}}(\varphi^{-1}) = {}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\varphi)^{-1}$; insbesondere ist ${}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\varphi)$ invertierbar.

Nehmen wir umgekehrt an, dass ${}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\varphi)$ invertierbar ist, dann handelt es sich um eine quadratische $(n \times n)$-Matrix mit $n = \dim_K(U) = \dim_K(V)$. Nach [MG, 4.5.4] hat die Matrix ${}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\varphi)$ den Rang n. Aus [MG, 9.2.4] folgt dann, dass auch $\mathrm{Rg}(\varphi) = n$ gilt, d.h., dass φ surjektiv ist. Da U und V dieselbe Dimension haben, ist φ damit schon ein Isomorphismus; siehe [MG, 8.3.17] □

2.2.26 Transformationsmatrizen. Es sei U ein K-Vektorraum. Sind $\mathcal{B}$ und $\mathcal{B}'$ geordnete Basen von U, dann nennt man

$$ {}_{\mathcal{B}'}\mathrm{M}_{\mathcal{B}}(\mathrm{id}_U) $$

die *Transformationsmatrix* zum Basiswechsel von $\mathcal{B}$ nach $\mathcal{B}'$. Für alle $u \in U$ gilt

$$ {}_{\mathcal{B}'}\mathrm{M}_{\mathcal{B}}(\mathrm{id}_U)\kappa_{\mathcal{B}}(u) = \kappa_{\mathcal{B}'}(u). $$

Die Transformationsmatrix rechnet also die Koordinaten bzgl. $\mathcal{B}$ in die Koordinaten bzgl. $\mathcal{B}'$ um. Aus [MG, 9.3.2] ist bekannt, dass ${}_{\mathcal{B}'}\mathrm{M}_{\mathcal{B}}(\mathrm{id}_U)^{-1} = {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}'}(\mathrm{id}_U)$ gilt.

2.2.27 Beispiel. Wir betrachten den Vektorraum $\mathbb{R}^2$ mit den geordneten Basen

$$ \mathcal{B} = \left(\begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \end{pmatrix} \right) \quad \text{und} \quad \mathcal{B}' = \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right). $$

Um die Transformationsmatrix ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}'}(\mathrm{id}_{\mathbb{R}^2})$ zu bestimmen, schreiben wir die Vektoren der Basis $\mathcal{B}'$ als Linearkombination der Vektoren aus $\mathcal{B}$ (dazu muss man ein lineares Gleichungssystem lösen). Wir erhalten

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

und damit

$$ {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}'}(\mathrm{id}_{\mathbb{R}^2}) = \begin{pmatrix} -1 & 3 \\ 1 & -2 \end{pmatrix}. $$

Da $\mathcal{B}'$ die Standardbasis ist, können wir direkt ablesen, wie die Vektoren aus $\mathcal{B}$ als Linearkombination in der Standardbasis geschrieben werden. Die Transformationsmatrix ${}_{\mathcal{B}'}\mathrm{M}_{\mathcal{B}}(\mathrm{id}_{\mathbb{R}^2})$ erhält man dann einfach, indem man die Vektoren aus $\mathcal{B}$ als Spalten in eine Matrix schreibt:

$$ {}_{\mathcal{B}'}\mathrm{M}_{\mathcal{B}}(\mathrm{id}_{\mathbb{R}^2}) = \begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix}. $$

Man kann leicht nachrechnen, dass diese Matrizen invers zueinander sind. Alternativ hätten wir ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}'}(\mathrm{id}_{\mathbb{R}^2})$ bestimmen können, indem wir die Matrix ${}_{\mathcal{B}'}\mathrm{M}_{\mathcal{B}}(\mathrm{id}_{\mathbb{R}^2})$ bestimmen und die Inverse berechnen.

Die Matrixdarstellung ${}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\varphi)$ hängt von den gewählten Basen ab. Was passiert mit der Matrixdarstellung, wenn man andere Basen wählt?

2.2.28 **Transformationsformel für lineare Abbildungen.** *Es seien U, V zwei K-Vektorräume und sei $\varphi\colon U \to V$ eine K-lineare Abbildung.*

Sind $\mathcal{B}$, $\mathcal{B}'$ geordnete Basen von U und $\mathcal{C}$, $\mathcal{C}'$ geordnete Basen von V, dann gilt

$$ {}_{\mathcal{C}'}\mathrm{M}_{\mathcal{B}'}(\varphi) = {}_{\mathcal{C}'}\mathrm{M}_{\mathcal{C}}(\mathrm{id}_V)\ {}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\varphi)\ {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}'}(\mathrm{id}_U). $$

Beweis. Das folgt aus [MG, 9.3.1] indem man $\varphi = \mathrm{id}_V \circ \varphi \circ \mathrm{id}_U$ schreibt. □

2.3. Direkte Summe von Vektorräumen

In diesem Abschnitt lernen wir ein wichtiges Hilfsmittel für das Arbeiten mit Vektorräumen kennen: die direkte Summe. Die direkte Summe erlaubt es - grob gesagt - Vektorräume zu „zerlegen“. Wir werden dies häufig verwenden um Probleme und Fragestellungen zu Vektorräumen auf kleinere Unterräume zu reduzieren.

I. Definition und Eigenschaften

Zunächst beginnen wir mit Summen von Unterräumen. Aus den Mathematischen Grundlagen kennen Sie folgende Aussage: Sind $U_1, U_2 \subseteq V$ zwei Unterräume eines Vektorraumes V, dann ist auch

$$U_1 + U_2 = \{u_1 + u_2 \mid u_1 \in U_1,\ u_2 \ \in U_2\}$$

ein Unterraum von V. Diese Konstruktion funktioniert natürlich auch für 3, 4, 5 oder mehr Unterräume.

2.3.1 **Definition und Satz.** *Sei V ein K-Vektorraum. Sind $U_1, \ldots, U_k$ Unterräume von V, so definieren wir*

$$U_1 + U_2 + \cdots + U_k := \big\{u_1 + u_2 + \cdots + u_k \mid u_i \in U_i \text{ für alle } 1 \leq i \leq k\big\}.$$

Man nennt $U_1 + U_2 + \cdots + U_k$ die Summe der Unterräume *$U_1, \ldots, U_k$ und schreibt dafür auch $\sum_{i=1}^k U_i$.*

Die Summe $\sum_{i=1}^k U_i$ ist wieder ein Unterraum von V.

Beweis. Wir führen den Beweis mit vollständiger Induktion nach k. Für $k = 1$ ist nichts zu tun und für $k = 2$ kennen wir die Aussage aus den Mathematischen Grundlagen [MG, Prop. 6.2.11].

Als Induktionsvoraussetzung können wir also annehmen, dass $\sum_{i=1}^{k-1} U_i$ ein Unterraum ist. Dann sehen wir aus der Definition der Summe, dass

$$\sum_{i=1}^{k} U_i = \left(\sum_{i=1}^{k-1} U_i\right) + U_k$$

gilt. Mit der Aussage für zwei Unterräume schließen wir, dass $\sum_{i=1}^k U_i$ ein Unterraum ist. $\square$

2.3.2 **Beispiel**. Die lineare Hülle einer Menge von Vektoren ist ein Beispiel einer Summe. Sei V ein K-Vektorraum. Sind $v_1, \dots, v_k \in V$, dann gilt

$$\langle v_1, v_2, \dots, v_k \rangle = \langle v_1 \rangle + \langle v_2 \rangle + \cdots + \langle v_k \rangle.$$

Um das zu sehen, verwenden die Definition der linearen Hülle aus [MG, 6.3.2]. Es ist $\langle v \rangle = \{av \mid a \in K\}$. Also gilt

$$\langle v_1 \rangle + \langle v_2 \rangle + \cdots + \langle v_k \rangle = \Big\{ \sum_{i=1}^{k} a_i v_i \mid a_i \in K \Big\}$$

und dies ist per Definition die lineare Hülle der Vektoren $v_1, \dots, v_k$.

2.3.3 **Aufgabe**. Es sei V ein K-Vektorraum und es seien $S, T \subseteq V$ zwei Teilmengen. Zeigen Sie: $\langle S \cup T \rangle = \langle S \rangle + \langle T \rangle$. **L**

Im Allgemeinen kann man ein Element v in der Summe $U_1 + U_2 + \cdots + U_k$ auf viele verschiedene Weisen in der Form

$$v = u_1 + u_2 + \cdots + u_k$$

mit $u_i \in U_i$ schreiben.

2.3.4 **Beispiel**. Wir betrachten den Vektorraum $\mathbb{R}^3$ und die Unterräume

$$U_1 = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mid x_1 = 0 \right\}$$

und

$$U_2 = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mid x_2 + x_3 = 0 \right\}.$$

Dann gilt $U_1 + U_2 = \mathbb{R}^3$, denn die Standardbasisvektoren e_2 und e_3 liegen in U_1 und der Standardbasisvektor e_1 liegt in U_2, d.h.,

$$U_1 + U_2 \supseteq \langle e_2, e_3 \rangle + \langle e_1 \rangle = \langle e_1, e_2, e_3 \rangle = \mathbb{R}^3.$$

Die Vektoren im $\mathbb{R}^3$ kann man nun aber auf unendlich viele verschiedene Weisen in der Form $u_1 + u_2$ mit $u_1 \in U_1$ und $u_2 \in U_2$ schreiben. Beispielsweise gilt für alle $t \in \mathbb{R}$:

$$0 = \underbrace{\begin{pmatrix} 0 \\ t \\ -t \end{pmatrix}}_{\in U_1} + \underbrace{\begin{pmatrix} 0 \\ -t \\ t \end{pmatrix}}_{\in U_2}.$$

Besonders nützlich ist daher folgende spezielle Situation:

2.3.5 **Definition** Sei V ein Vektorraum und seien $U_1, \ldots, U_k$ Unterräume von V. Man sagt, die Summe $\sum_{i=1}^k U_i$ ist *direkt*, wenn jeder Vektor $v \in \sum_{i=1}^k U_i$ auf genau eine Weise in der Form

$$v = \sum_{i=1}^{k} u_i$$

mit $u_i \in U_i$ für alle $i \in \{1, 2, \ldots, k\}$ geschrieben werden kann.

Man notiert die Summe durch $U_1 \oplus U_2 \oplus \cdots \oplus U_k$ (oder kürzer $\bigoplus_{i=1}^k U_i$) um auszudrücken, dass die Summe dieser Unterräume eine direkte Summe ist.

Diese Definition der direkten Summe ist nützlich, ist aber nicht gut geeignet um in konkreten Fällen nachzuweisen, dass eine Summe direkt ist. Bevor wir Beispiele besprechen, leiten wir eine alternative Charakerisierung der direkten Summe her.

2.3.6 **Satz.** *Sei V ein K-Vektorraum und seien $U_1, U_2, \ldots, U_k$ Unterräume. Die folgenden Aussagen sind äquivalent.*

(i) Die Summe $\sum_{i=1}^k U_i$ ist direkt.

(ii) Für alle $j \in \{1, 2, \ldots, k\}$ gilt

$$U_j \cap \sum_{\substack{i=1 \\ i \neq j}}^{k} U_i = \{0\}.$$

(iii) Aus $\sum_{i=1}^k u_i = 0$ mit $u_i \in U_i$ für alle $i \leq k$, folgt $u_i = 0$ für alle $i \leq k$.

Beweis. (i) $\implies$ (ii): Wir nehmen an, dass die Summe direkt ist. Weil jeder Unterraum den Nullvektor enthält, ist $0 = \sum_{i=1}^k 0$ eine Schreibweise des Nullvektors. Sei $j \in \{1, 2, \ldots, k\}$ und sei

$$v \in U_j \cap \sum_{\substack{i=1 \\ i \neq j}}^{k} U_i.$$

Wir schreiben $v = \sum_{i \neq j} u_i$ mit $u_i \in U_i$ für alle $i \neq j$. Es sei $u_j = -v \in U_j$. Dann ist

$$0 = \sum_{i=1}^{k} u_i$$

und weil wir schon eine Summendarstellung für den Nullvektor kennen und diese Darstellung nach Definition der direkten Summe eindeutig ist, folgt $u_i = 0$ für alle i. Insbesondere gilt $v = 0$ und damit

$$U_j \cap \sum_{\substack{i=1\\i\neq j}}^{k} U_i = \{0\}.$$

(ii) $\implies$ (iii): Es sei $\sum_{i=1}^{k} u_i = 0$. Sei $j \in \{1, 2, \dots, k\}$. Dann ist

$$-u_j = \sum_{\substack{i=1\\i\neq j}}^{k} u_i \in U_j \cap \sum_{\substack{i=1\\i\neq j}}^{k} U_i = \{0\}.$$

Also ist $u_j = 0$ für alle $j \in \{1, 2, \dots, k\}$.

(iii) $\implies$ (i): Es sei $v \in \sum_{i=1}^{k} U_i$. Sind

$$v = \sum_{i=1}^{k} u_i = \sum_{i=1}^{k} u_i'$$

zwei Summendarstellungen von v mit $u_i, u_i' \in U_i$, dann gilt $\sum_{i=1}^{k}(u_i - u_i') = 0$. Da $u_i - u_i'$ in U_i liegt (U_i ist ein Unterraum), folgt aus (iii), dass $u_i - u_i' = 0$ für alle $i \leq k$ ist. Das heißt, die beiden Summendarstellungen von v sind identisch. □

Für die direkte Summe zweier Unterräume erhalten wir eine besonders einfache Beschreibung.

2.3.7 **Korollar.** *Es seien $U_1, U_2 \subseteq V$ Unterräume. Die Summe $U_1 + U_2$ ist genau dann direkt, wenn $U_1 \cap U_2 = \{0\}$ gilt.*

Beweis. Das folgt aus 2.3.6 (ii), denn sowohl für $j = 1$ als auch für $j = 2$ erhält man jeweils die Bedingung $U_1 \cap U_2 = \{0\}$. □

2.3.8 **Beispiel.** Wir betrachten den Vektorraum $\mathbb{R}^3$ mit den Unterräumen

$$U_1 = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mid x_1 = 0 \right\} \quad \text{und} \quad U_2 = \langle \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \rangle.$$

Die Summe $U_1 + U_2$ ist direkt. In der Tat, es gilt $U_1 \cap U_2 = \{0\}$, denn ein Vektor $a \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ mit $a \in \mathbb{R}$ liegt genau dann in U_1, wenn $a = 0$ ist.

2.3.9 **Beispiel**. Sei V ein K-Vektorraum. Es seien $v_1, \ldots, v_k \in V$ Vektoren mit $v_i \neq 0$ für alle i. Die Summe

$$\langle v_1 \rangle + \cdots + \langle v_k \rangle$$

ist genau dann direkt, wenn die Vektoren $v_1, \ldots, v_k$ linear unabhängig sind. Das folgt direkt aus der Bedingung 2.3.6 (iii).

2.3.10 **Aufgabe**. Wir betrachten den Vektorraum $\mathbb{Q}^4$ mit den Unterräumen **L**

$$U_1 = \left\{ x \in \mathbb{Q}^4 \mid x_1 + x_2 = 0 \text{ und } x_3 + x_4 = 0 \right\}$$

und

$$U_2 = \left\{ x \in \mathbb{Q}^4 \mid x_1 = 0 \text{ und } x_3 = 0 \right\}.$$

Zeigen Sie, dass die Summe $U_1 + U_2$ direkt ist.

2.3.11 **Aufgabe**. Es sei V ein K-Vektorraum mit einer Basis $\mathcal{B}$. Wir nehmen an, dass $\mathcal{B} = \bigcup_{i=1}^k \mathcal{B}_i$ die disjunkte Vereinigung von Teilmengen $\mathcal{B}_i \subseteq \mathcal{B}$ ist. Zeigen Sie, dass **L**

$$V = \bigoplus_{i=1}^{k} \langle \mathcal{B}_i \rangle$$

gilt.

2.3.12 **Satz.** *Sei V ein K-Vektorraum. Wir nehmen an, dass $V = U_1 \oplus \cdots \oplus U_k$ die direkte Summe von Unterräumen $U_1, \ldots, U_k$ ist. Ist $\mathcal{B}_i$ eine Basis von U_i für alle $i \in \{1, 2, \ldots, k\}$, dann ist $\mathcal{B} = \bigcup_{i=1}^k \mathcal{B}_i$ eine Basis von V.*

Ist V endlich-dimensional, dann gilt $\dim_K(V) = \sum_{i=1}^k \dim_K(U_i)$.

Beweis. Wir zeigen zuerst, dass $\mathcal{B}$ ein Erzeugendensystem ist. Indem wir Aufgabe 2.3.3 wiederholt anwenden, erhalten wir

$$\langle \mathcal{B} \rangle = \langle \bigcup_{i=1}^{k} \mathcal{B}_i \rangle = \sum_{i=1}^{k} \langle \mathcal{B}_i \rangle = \sum_{i=1}^{k} U_i = V,$$

d.h., $\mathcal{B}$ erzeugt V.

Als nächstes prüfen wir, ob $\mathcal{B}$ linear unabhängig ist. Für jedes $i \in \{1, 2, \ldots, k\}$ seien $v_1^{(i)}, \ldots, v_{r_i}^{(i)} \in \mathcal{B}_i$ paarweise verschiedene Vektoren (mit $r_i \in \mathbb{N}_0$). Seien $a_1^{(i)}, \ldots, a_{r_i}^{(i)} \in K$ mit

$$\sum_{i=1}^{k} \sum_{j=1}^{r_i} a_j^{(i)} v_j^{(i)} = 0.$$

Da $\sum_{j=1}^{r_i} a_j^{(i)} v_j^{(i)} \in U_i$ ist und die Summe direkt ist, folgt aus 2.3.6, dass

$$\sum_{j=1}^{r_i} a_j^{(i)} v_j^{(i)} = 0$$

für alle $i \in \{1, 2, \ldots, k\}$ gilt. Da die Vektoren $v_1^{(i)}, \ldots, v_{r_i}^{(i)}$ paarweise verschiedene Basisvektoren (und damit linear unabhängig) sind, folgt $a_j^{(i)} = 0$ für alle $j \leq r_i$. Die Menge $\mathcal{B}$ ist also linear unabhängig und damit eine Basis.

Ist V endlich-dimensional, dann gilt

$$\dim_K(V) = |\mathcal{B}| = \left|\bigcup_{i=1}^{k} \mathcal{B}_i\right| = \sum_{i=1}^{k} |\mathcal{B}_i| = \sum_{i=1}^{k} \dim_K(U_i). \qquad \square$$

2.3.13 **Aufgabe**. Es seien V, W zwei K-Vektorräume. Wir nehmen an, dass $V = \bigoplus_{i=1}^k V_i$ und $W = \bigoplus_{i=1}^k W_i$ jeweils direkte Summen von k Unterräumen sind. Es sei $\varphi\colon V \to W$ eine lineare Abbildung mit der Eigenschaft $\varphi(V_i) \subseteq W_i$. Zeigen Sie, dass dann gilt **L**

$$\operatorname{Ker}(\varphi) = \bigoplus_{i=1}^{k} \operatorname{Ker}(\varphi|_{V_i}).$$

Dabei bezeichnet $\varphi|_{V_i}$ die Einschränkung von φ auf V_i.

II. Komplemente

2.3.14 $\boxed{\textbf{Definition}}$ Sei V ein Vektorraum und sei $U \subseteq V$ ein Unterraum. Ein Unterraum $W \subseteq V$ heißt *Komplement* von U in V, wenn

$$V = U \oplus W$$

gilt. Das heißt, es gelten $U + W = V$ und $U \cap W = \{0\}$.

2.3.15 **Anmerkungen zur Definition**. (a) Ist W ein Komplement zu U in V, dann ist auch U ein Komplement zu W in V. Man sagt daher auch, dass U und W zueinander komplementäre Unterräume von V sind.

(b) Sei V ein endlich-dimensionaler Vektorraum. Sind $U, W \subseteq V$ komplementäre Unterräume, dann gilt mit 2.3.12

$$\dim(U) + \dim(W) = \dim(V).$$

2.3.16 **Beispiel**. Wir betrachten den Vektorraum $\mathbb{R}^3$. Es sei U ein 2-dimensionaler Unterraum, d.h., eine Ebene durch den Ursprung. Da $\mathbb{R}^3$ die Dimension 3 hat und U die Dimension 2, schließen wir aus Satz 2.3.12, dass jedes Komplement zu U ein-dimensional sein muss.

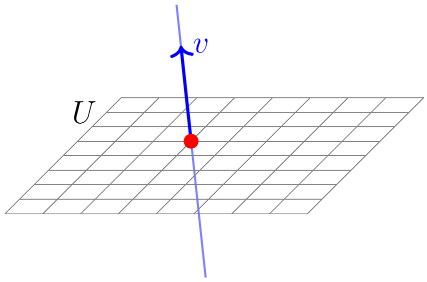


Abbildung 2.1.: Zwei komplementäre Unterräume im $\mathbb{R}^3$.

Es sei $v \in \mathbb{R}^3$. Wann ist die Gerade $\langle v \rangle$ ein Komplement zu U? Falls $v \in U$ liegt, gilt $U + \langle v \rangle = U$, also ist $\langle v \rangle$ kein Komplement. Sei also $v \notin U$. Dann liegt auch av mit $a \neq 0$ nicht in U und damit gilt

$$\langle v \rangle \cap U = \{0\}.$$

Insbesondere ist die Summe direkt und mit 2.3.12 folgt

$$\dim_{\mathbb{R}}(\langle v \rangle \oplus U) = 1 + 2 = 3.$$

Folglich ist $\langle v \rangle \oplus U = \mathbb{R}^3$ und $\langle v \rangle$ ist ein Komplement zu U.

Im Beispiel haben wir gesehen, dass es zu einem Unterraum im Allgemeinen viele verschiedene Komplemente geben kann. Der nächste Satz stellt sicher, dass zu jedem Unterraum immer mindestens ein Komplement existiert.

2.3.17 **Satz vom Komplement.** *Sei V ein Vektorraum. Zu jedem Unterraum $U \subseteq V$ existiert ein Komplement.*

Beweis. Wir werden den Satz mit dem Lemma von Zorn beweisen, weil dies ein schönes Beispiel dafür ist, wie das Lemma eingesetzt wird[6]. Der Satz vom Komplement benötigt also das Auswahlaxiom.

Wir definieren die Menge M aller Unterräume $W \subseteq V$ mit $U \cap W = \{0\}$. Die Menge ist bezüglich der Inklusion $\subseteq$ partiell geordnet. Wir behaupten, dass jede Kette $\mathcal{K}$ in M eine obere Schranke in M besitzt.

[6] Alternativ könnte man auch mit dem Basisergänzungssatz argumentieren.

Sei $\mathcal{K}$ eine Kette. Wir definieren

$$W' = \bigcup_{W \in \mathcal{K}} W$$

als Vereinigung aller Elemente von $\mathcal{K}$. Dann ist W' wieder ein Unterraum von V. Es seien $w_1, w_2 \in W'$ und $a \in K$. Da w_i in der Vereinigung liegt, gibt es jeweils ein $W_i \in \mathcal{K}$ mit $w_i \in W_i$. Da $\mathcal{K}$ eine Kette ist, gilt $W_1 \subseteq W_2$ oder $W_2 \subseteq W_1$. Sagen wir (o.B.d.A.) es gilt $W_1 \subseteq W_2$, also $w_1, w_2 \in W_2$. Da W_2 ein Unterraum ist, schließen wir

$$aw_1 + w_2 \in W_2 \subseteq W'.$$

Also ist W' nach 2.2.6 ein Unterraum von V. Es gilt auch $U \cap W' = \bigcup_{W \in \mathcal{K}} U \cap W = \{0\}$. Also ist $W' \in M$ eine obere Schranke für $\mathcal{K}$.

Das Lemma von Zorn 2.1.13 liefert uns nun ein maximales Element W in M. Wir behaupten, dass W ein Komplement zu U ist. Nach Definition von M gilt $U \cap W = \{0\}$. Wir zeigen nun, dass $U + W = V$ ist. Für einen Widerspruch nehmen wir an, dass es einen Vektor $v \in V$ gibt, der nicht in $U + W$ liegt. Wir definieren $Z = W + \langle v \rangle$. Wir behaupten, dass $Z \cap U = \{0\}$ ist und Z damit in M liegt.

Sei $u \in U \cap Z$. Wir schreiben $u = w + av$ mit $w \in W$ und $a \in K$. Dann erhalten wir

$$av = u - w \in U + W$$

und damit $a = 0$, denn v liegt nicht in $U + W$. Daraus folgt nun, dass $u = w$ in $U \cap W = \{0\}$ liegt, d.h., $U \cap Z = \{0\}$.

Daraus folgt $Z \in M$ und wegen $W \subsetneq Z$ ist dies ein Widerspruch zur Maximalität von W. Es muss also $U + W = V$ sein und W ist das gesuchte Komplement. □

2.3.18 **Aufgabe**. Beweisen Sie mit dem Satz vom Komplement das „lineare Auswahlaxiom“: Zu jeder surjektiven K-linearen Abbildung existiert ein K-lineares Rechtsinverses. **L**

Zum Abschluss überlegen wir uns, wie man ein Komplement bestimmen kann. Dazu muss man nur eine Basis von U zu einer Basis von V ergänzen.

2.3.19 **Lemma.** *Sei V ein endlich-dimensionaler K-Vektorraum und sei U ein Unterraum. Ist $\mathcal{B} = (v_1, \ldots, v_n)$ eine geordnete Basis mit $U = \langle v_1, \ldots, v_k \rangle$, dann ist*

$$W = \langle v_{k+1}, \ldots, v_n \rangle$$

ein Komplement von U in V.

Beweis. Aus 2.3.3 schließen wir

$$U + W = \langle v_1, \dots, v_k \rangle + \langle v_{k+1}, \dots, v_n \rangle = \langle v_1, \dots v_n \rangle = V.$$

Warum gilt $U \cap W = \{0\}$? Sei $v \in U \cap W$. Wir können also v auf zwei Weisen als Linearkombination schreiben:

$$v = \sum_{i=1}^{k} a_i v_i = \sum_{j=k+1}^{n} a_j v_j$$

mit Koeffizienten $a_1, \dots, a_n \in K$. Damit erhalten wir

$$0 = \sum_{i=1}^{k} a_i v_i - \sum_{j=k+1}^{n} a_j v_j.$$

Da die Vektoren $v_1, \dots, v_n$ linear unabhängig sind, folgt $a_i = 0$ für alle $i \in \{1, 2, \dots, n\}$ und damit $v = 0$. $\square$

2.3.20 **Beispiel**. Wir betrachten den Raum $\mathbb{F}_2^4$ mit dem Unterraum $U = \langle v_1, v_2 \rangle$ mit

$$v_1 = \begin{pmatrix} \overline{1} \\ \overline{0} \\ \overline{1} \\ \overline{0} \end{pmatrix} \quad \text{und} \quad v_2 = \begin{pmatrix} \overline{0} \\ \overline{1} \\ \overline{0} \\ \overline{1} \end{pmatrix}$$

Wir ergänzen die beiden Vektoren mit den Vektoren

$$v_3 = \begin{pmatrix} \overline{1} \\ \overline{0} \\ \overline{0} \\ \overline{0} \end{pmatrix} \quad \text{und} \quad v_4 = \begin{pmatrix} \overline{0} \\ \overline{1} \\ \overline{0} \\ \overline{0} \end{pmatrix}$$

zu einer Basis von $\mathbb{F}_2^4$. Hier muss man einmal prüfen, dass diese Vektoren wirklich linear unabhängig sind, d.h., dass die Matrix

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

vollen Rang hat. Also ist $W = \langle v_3, v_4 \rangle$ ein Komplement zu U in $\mathbb{F}_2^4$.

2.4. Faktorräume

Hat man einen Vektorraum V und einen Unterraum U, so kann man damit immer einen neuen Vektorraum konstruieren: den sogenannten *Faktorraum* V/U. Diese Konstruktion werden wir in diesem Abschnitt kennenlernen. Die Vektoren im Faktorraum sind selbst Mengen (genauer: „affine Unterräume“ von V). Bei der ersten Begegnung wirkt die Konstruktion des Faktorraumes daher etwas abstrakt und unintuitiv. Sobald man sich aber an das Arbeiten mit Faktorräumen gewöhnt hat, wird man feststellen, dass diese Objekte sehr vielseitige Hilfsmittel sind. Der Faktorraum ist ein wunderbares Beispiel für viele weiterführende Ideen der Algebra, wo ähnliche „Faktorobjekte“ (Faktorgruppen, Faktorringe, Faktormoduln...) eine zentrale Bedeutung haben.

Wir beginnen mit einem motivierenden Beispiel. Wir verzichten dabei auf einige Details, denn wir werden danach in einer viel allgemeineren Situation alle Beweise ausführen.

2.4.1 **Die Parallelenschar als Vektorraum.** Wir betrachten den zwei-dimensionalen reellen Vektorraum $\mathbb{R}^2$ mit dem ein-dimensionalen Unterraum

$$U = \langle \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rangle = \left\{ \begin{pmatrix} t \\ t \end{pmatrix} \mid t \in \mathbb{R} \right\}.$$

Die Punkte im Unterraum U bilden eine Gerade im $\mathbb{R}^2$. Es sei $\mathcal{G}$ die Menge aller

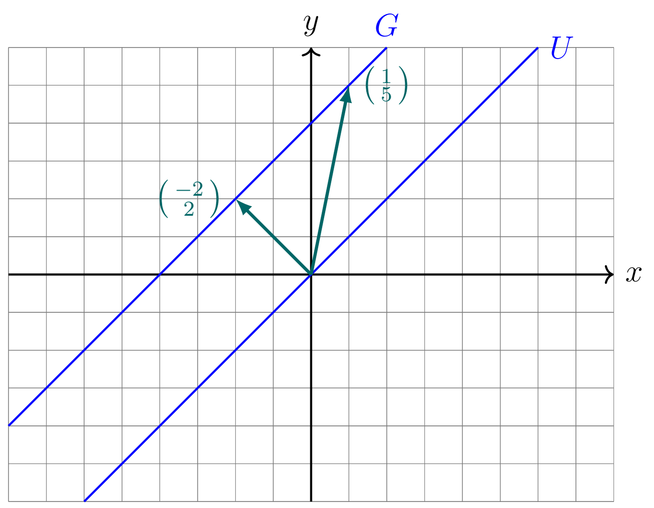


Abbildung 2.2.: Die Gerade U und die parallele Gerade G im $\mathbb{R}^2$.

zu U parallelen Geraden; man spricht auch von der zugehörigen *Parallelenschar*. Eine parallele Gerade erhält man, indem man alle Punkte auf U durch Addition mit einem Vektor verschiebt. Zum Beispiel ist

$$G = \begin{pmatrix} -2 \\ 2 \end{pmatrix} + U := \left\{ \begin{pmatrix} -2+t \\ 2+t \end{pmatrix} \mid t \in \mathbb{R} \right\}$$

die zu U parallele Gerade durch den Punkt $\left(\begin{smallmatrix} -2 \\ 2 \end{smallmatrix}\right)$. Verschiedene „Verschiebungsvektoren" können dabei dieselbe parallele Gerade erzeugen. Zum Beispiel gilt

$$G = \begin{pmatrix} -2 \\ 2 \end{pmatrix} + U = \begin{pmatrix} 1 \\ 5 \end{pmatrix} + U$$

Zwei zu U parallele Geraden $v + U$ und $w + U$ sind dabei entweder gleich oder disjunkt.

Wir definieren nun eine Addition auf der Parallelenschar. Sind $G_1, G_2 \in \mathcal{G}$ dann definiert man

$$G_1 + G_2 := \{w_1 + w_2 \mid w_1 \in G_1, w_2 \in G_2\}.$$

Jetzt kommt die spannende Beobachtung: $G_1 + G_2$ ist ebenfalls eine zu U parallele Gerade! Genauer gilt: $(v_1 + U) + (v_2 + U) = (v_1 + v_2) + U$. Es ist hilfreich, das an einem Beispiel zu überprüfen: Berechnen Sie $(\left(\begin{smallmatrix} -2 \\ 2 \end{smallmatrix}\right) + U) + (\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) + U)$. Wir haben also eine Addition auf der Parallelenschar $\mathcal{G}$ definiert. Die Addition ist kommutativ, weil die Addition in $\mathbb{R}^2$ kommutativ ist. Es gibt sogar ein neutrales Element zu dieser Addition: die Gerade U. Es gilt $G_1 + U = G_1 = U + G_1$. **L**

Als nächstes definieren wir eine Skalarmultiplikation auf $\mathcal{G}$. Ist $G_1 \in \mathcal{G}$ und $a \in \mathbb{R}$ mit $a \neq 0$ gegeben, dann setzen wir

$$aG_1 = \{av \mid v \in G_1\}.$$

Wieder kann man zeigen: aG_1 ist ebenfalls eine zu U parallele Gerade! Genauer gilt: $a(v + U) = av + U$. Zur Vollständigkeit definiert man $0 \cdot G = 0 \cdot v + U = U$; man verwendet also dieselbe Formel.

Man kann nun beweisen, dass die Parallelenschar $\mathcal{G}$ mit dieser Addition und dieser Skalarmultiplikation wieder ein $\mathbb{R}$-Vektorraum ist.

I. Konstruktion des Faktorraumes

Alles was wir gerade im Beispiel 2.4.1 gesehen haben, funktioniert ganz allgemein über jedem Körper und mit jedem Unterraum U eines Vektorraumes. Welche Objekte ersetzen die zu U parallelen Geraden? Mengen von Punkten, die man durch „Verschiebung" von U erhält.

2.4.2 **Definition** Sei V ein Vektorraum. Ist $U \subseteq V$ ein Unterraum und $v \in V$, dann nennt man

$$v + U = \{v + u \in V \mid u \in U\}$$

den *affinen Unterraum* zu U durch den Punkt v. Der Vektor v wird *Stützvektor* genannt.

2.4.3 **Affine Unterräume vs. Unterräume**. Ein affiner Unterraum ist im Allgemeinen *kein* Unterraum im Sinne von 2.2.4. Die affinen Unterräume zu U nennt man manchmal auch *Nebenklassen* von U.

2.4.4 **Beispiel**. In den Mathematischen Grundlagen sind Sie affinen Unterräumen schon beim Lösen von linearen Gleichungssystemen begegnet. Wir betrachten ein lineares Gleichungssystem $Ax = b$ mit $A \in \mathrm{M}_{m,n}(K)$ und $b \in K^m$. Besitzt dieses Gleichungssystem eine Lösung $x_0 \in K^n$, dann hat die Lösungsmenge $\mathcal{L}$ die Form

$$\mathcal{L} = x_0 + U$$

wobei U der Lösungsraum des zugehörigen homogenen Gleichungssystems $Ax = 0$ ist, d.h., $U = \mathrm{Ker}(A)$. Die Lösungsmenge ist also ein affiner Unterraum von K^n.

Im nächsten Lemma sehen wir, dass je zwei affine Unterräume zum selben Unterraum U gewissermaßen „parallel“ sind.

2.4.5 **Lemma.** *Sei V ein K-Vektorraum und sei U ein Unterraum. Je zwei affine Unterräume zu U sind entweder gleich oder disjunkt.*

Die affinen Unterräume $v + U$ und $w + U$ sind genau dann gleich, wenn $v - w$ in U liegt.

Beweis. Es seien $v_1 + U$ und $v_2 + U$ zwei affine Unterräume zu U durch die Punkte $v_1, v_2 \in V$. Angenommen $v_1 + U$ und $v_2 + U$ sind nicht disjunkt, dann gibt es einen Punkt $x \in (v_1 + U) \cap (v_2 + U)$. Es gibt also Vektoren $u_1, u_2 \in U$ mit

$$x = v_1 + u_1 = v_2 + u_2.$$

Dann gilt $v_1 = v_2 + (u_2 - u_1)$. Ist $y = v_1 + u$ mit $u \in U$ nun ein beliebiger Punkt im affinen Unterraum $v_1 + U$, dann erhält man

$$y = v_1 + u = v_2 + \underbrace{(u_2 - u_1 + u)}_{\in U} \in v_2 + U.$$

Das heißt, $v_1 + U \subseteq v_2 + U$. Dasselbe Argument mit vertauschten Rollen liefert $v_2 + U \subseteq v_1 + U$; also gilt $v_1 + U = v_2 + U$.

Kommen wir nun zur zweiten Aussage. Nehmen wir zuerst $v + U = w + U$ an. Dann ist $v \in v + U = w + U$ und man kann v in der Form $v = w + u$ mit $u \in U$ schreiben. Dann gilt aber auch $v - w = u \in U$.

Gilt umgekehrt $v - w \in U$, dann ist $v = w + (v - w) \in w + U$. Damit sind $v + U$ und $w + U$ nicht disjunkt und somit gilt $v + U = w + U$. □

2.4.6 **Aufgabe**. Wir betrachten den reellen Vektorraum $\mathbb{R}^3$ und den Unterraum **L**

$$U = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mid 2x_1 - x_2 - x_3 = 0 \right\}.$$

Welche der folgenden affinen Unterräume sind gleich?

$$A = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + U, \quad B = \begin{pmatrix} 0 \\ 1 \\ -4 \end{pmatrix} + U, \quad C = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} + U$$

2.4.7 $\boxed{\textbf{Definition}}$ Sei V ein Vektorraum mit einem Unterraum U. Dann bezeichnet

$$V/U = \{v + U \mid v \in V\}$$

die Menge der affinen Unterräume zu U in V. Man spricht V/U als „V modulo U“.

2.4.8 **Addition affiner Unterräume**. Sei V ein Vektorraum mit einem Unterraum U. Auf der Menge V/U definieren wir nun eine Addition. Sind $A, B \in V/U$, dann setzen wir

$$A + B = \{x + y \in V \mid x \in A, y \in B\}.$$

Wir zeigen, dass $A + B$ wieder ein affiner Unterraum von V ist. Es seien $A = v + U$ und $B = w + U$ mit $v, w \in V$. Sind $x = v + u_1 \in A$ und $y = w + u_2 \in B$ gegeben (dabei sind $u_1, u_2 \in U$), dann gilt

$$x + y = v + w + \underbrace{(u_1 + u_2)}_{\in U} \in v + w + U;$$

also ist $A + B \subseteq v + w + U$. Umgekehrt durchlaufen u_1, u_2 (und damit $u_1 + u_2$) alle Elemente von U, wenn x und y die Elemente von A bzw. B durchlaufen, also gilt $A + B = v + w + U$. Damit erhalten wir die eingängliche Formel

$$(v + U) + (w + U) = v + w + U. \tag{2.4.a}$$

2.4.9 **Skalarmultiplikation mit affinen Unterräumen**. Sei V ein Vektorraum mit einem Unterraum U. Auf der Menge V/U definieren wir nun eine Skalarmultiplikation $K \times V/U \to V/U$. Für $a \in K^\times$ und $A \in V/U$ setzen wir

$$aA = \{ax \in V \mid x \in A\}.$$

Wir zeigen, dass aA wieder ein affiner Unterraum von V ist. Es sei $A = v + U$ mit $v \in V$. Ist $x = v + u \in A$ mit $u \in U$, dann gilt

$$ax = a(v+u) = av + \underbrace{au}_{\in U} \in av + U;$$

also ist $aA \subseteq av + U$. Mit $x \in A$ durchläuft u (und damit au) alle Elemente von U, also gilt $aA = av + U$. Daraus erhalten wir die Formel

$$a(v+U) = av + U. \tag{2.4.b}$$

Die Skalarmultiplikation mit 0 definieren wir einfach mit derselben Formel, d.h.,

$$0 \cdot (v+U) = 0 \cdot v + U = 0 + U = U.$$

2.4.10 **Aufgabe**. Wir betrachten die Parallelenschar $\mathbb{R}^2/U$ aus Beispiel 2.4.1. Berechnen Sie in $\mathbb{R}^2 + U$: $2 \cdot \left(\begin{pmatrix} 3 \\ -5 \end{pmatrix} + U\right) + \left(\begin{pmatrix} -1 \\ 1 \end{pmatrix} + U\right)$. **L**

2.4.11 **Satz vom Faktorraum.** *Sei V ein K-Vektorraum und sei U ein Unterraum. Die Menge V/U bildet mit den Rechenoperationen aus 2.4.8 und 2.4.9 einen K-Vektorraum.*

Man nennt V/U den Faktorraum *von V nach U.*

Beweis. Wir verifizieren die Vektorraumaxiome aus 2.2.1.

2.2.1 (i):
Wir werden die Gruppenaxiome aus 1.1.1 mithilfe der Formel (2.4.a) herleiten. Zuerst halten wir fest, dass die Addition kommutativ ist; es gilt

$$(v+U) + (w+U) = v + w + U = w + v + U = (w+U) + (v+U)$$

für alle $v, w \in V$, denn die Addition in V ist kommutativ.

(G1): Für alle $v_1, v_2, v_3 \in V$ gilt

$$\big((v_1+U) + (v_2+U)\big) + (v_3+U) = (v_1+v_2+U) + (v_3+U)$$

$$\begin{aligned} &= (v_1 + v_2) + v_3 + U && \text{(Assoz. in } V\text{)} \\ &= v_1 + (v_2 + v_3) + U \\ &= (v_1 + U) + (v_2 + v_3 + U) \\ &= (v_1 + U) + \big((v_2 + U) + (v_3 + U)\big). \end{aligned}$$

Das Assoziativgesetz ist somit erfüllt.

(G2): $U = 0 + U$ ist das neutrale Element, denn mit (2.4.a) gilt

$$(v + U) + U = v + U = U + (v + U).$$

für alle $v \in V$.

(G3): Das Inverse zu $v + U$ ist $-v + U$, denn mit (2.4.a) erhält man

$$(v + U) + (-v + U) = v + (-v) + U = 0 + U = U$$

und die Kommutativität liefert $(-v + U) + (v + U) = 0 + U = U$.

2.2.1 (ii):
Es seien $a, b \in K$ und $v \in V$. Da V das Axiom 2.2.1 (ii) erfüllt, folgt aus (2.4.b)

$$(ab)(v + U) = (ab)v + U = a(bv) + U = a(bv + U) = a(b(v + U)).$$

Genauso einfach erhalten wir $1(v + U) = 1v + U = v + U$ für alle $v \in V$.

2.2.1 (iii):
Die Distributivgesetze kann man mithilfe der Formeln (2.4.a) und (2.4.b) aus den Distributivgesetzen für V herleiten (Übungsaufgabe!). □ **L**

Natürlich gibt es eine Verbindung zwischen dem Vektorraum V und dem Faktorraum V/U: die sogenannte „kanonische Projektion". Schauen wir uns dazu nochmal die Parallelenschar aus 2.4.1 an.

2.4.12 Die kanonische Projektion auf die Parallelenschar. Wir betrachten wieder den zwei-dimensionalen reellen Vektorraum $\mathbb{R}^2$ mit dem ein-dimensionalen Unterraum

$$U = \left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle$$

und der zugehörigen Parallelenschar $\mathcal{G}$. Wie wir gerade gesehen haben, ist $\mathcal{G} = \mathbb{R}^2/U$ der Faktorraum von $\mathbb{R}^2$ nach U.

Jeder Punkt $v \in \mathbb{R}^2$ liegt auf *genau einer* der Parallelen G in $\mathcal{G}$ – nämlich genau auf der Geraden $G = v + U$. Ordnet man nun jedem Punkt $v \in \mathbb{R}^2$ die parallele Gerade $v + U$ in $\mathcal{G}$ zu, dann erhält man eine Abbildung

$$\pi\colon \mathbb{R}^2 \to \mathcal{G} \quad \text{mit} \quad \pi(v) = v + U.$$

Die Abbildung π ist surjektiv, denn auf jeder Geraden liegt mindestens ein Punkt. Man nennt π die *kanonische Projektion* auf $\mathcal{G}$. Wir werden nun allgemein sehen, dass π sogar eine lineare Abbildung ist!

2.4.13 **Definition und Satz.** *Sei V ein K-Vektorraum und sei $U \subseteq V$ ein Unterraum. Die Abbildung*

$$\pi\colon V \to V/U \quad \textit{mit} \quad \pi(v) = v + U$$

ist surjektiv und K-linear. Der Kern von π ist $\operatorname{Ker}(\pi) = U$.

Man nennt π die kanonische Projektion *auf V/U oder auch* Faktorabbildung.

Beweis. Sei $v + U \in V/U$ beliebig mit einem $v \in V$ gegeben. Dann liegt $v + U$ im Bild von π, denn es gilt $\pi(v) = v + U$. Damit ist die kanonische Projektion surjektiv.

Wir rechnen nun nach, dass π linear ist; siehe [MG, 8.1]. Es seien $v, w \in V$. Aus (2.4.a) folgt nun

$$\pi(v + w) = v + w + U = (v + U) + (w + U) = \pi(v) + \pi(w).$$

Sei $a \in K$. Dann gilt mit (2.4.b)

$$\pi(av) = av + U = a(v + U) = a\pi(v).$$

Wir bestimmen noch den Kern von π. Der Nullvektor in V/U ist $U = 0 + U$. Wann gilt $\pi(v) = 0 + U$? Nach 2.4.5 ist dies genau dann der Fall, wenn $v - 0 = v$ in U liegt, d.h., $\operatorname{Ker}(\pi) = U$. $\square$

Das folgende Lemma enthält ein Hilfsresultat, dass wir erst später benötigen. Der Beweis gibt aber ein gutes Beispiel ab, wie man mit Faktorräumen arbeitet.

2.4.14 **Lemma.** *Es seien V ein K-Vektorraum und $U, W_1, \dots, W_s \subseteq V$ Unterräume. Es sei $\pi\colon V \to V/U$ die kanonische Projektion. Angenommen es gelten $U = \bigoplus_{i=1}^{s} (U \cap W_i)$ und $V/U = \bigoplus_{i=1}^{s} \pi(W_i)$, dann gilt auch $V = \bigoplus_{i=1}^{s} W_i$.*

<u>*Beweis*</u>. *Schritt 1:* $V = \sum_{i=1}^{s} W_i$.

Sei $v \in V$ beliebig. Wir bilden v mit der Projektion π nach V/U ab. Da $V/U = \sum_{i=1}^{s} \pi(W_i)$ gilt, finden wir Vektoren $w_i \in W_i$ für alle $i \in \{1, 2 \ldots, s\}$, sodass

$$\pi(v) = \sum_{i=1}^{s} \pi(w_i) = \pi\left(\sum_{i=1}^{s} w_i\right)$$

gilt. Durch Umstellen dieser Gleichung folgt, dass der Vektor $u = v - \sum_{i=1}^{s} w_i$ in $U = \operatorname{Ker}(\pi)$ liegt; siehe 2.4.13. Da außerdem $U = \sum_{i=1}^{s}(U \cap W_i)$ gilt finden wir Vektoren $u_i \in U \cap W_i$, sodass $u = \sum_{i=1}^{s} u_i$ gilt. Da W_i ein Unterraum ist, liegt auch $u_i + w_i$ in W_i. Insgesamt erhalten wir damit

$$v = u + \sum_{i=1}^{s} w_i = \sum_{i=1}^{s}(u_i + w_i) \in \sum_{i=1}^{s} W_i.$$

Da $v \in V$ beliebig war, ist V die Summe der Unterräume $W_1, \ldots, W_s$.

Schritt 2: $V = \bigoplus_{i=1}^{s} W_i$.

Wir zeigen, dass die Summe direkt ist, indem wir Bedingung (iii) aus Satz 2.3.6 nachweisen. Angenommen, es sind Vektoren $w_i \in W_i$ mit $\sum_{i=1}^{s} w_i = 0$ gegeben. Wir wenden die kanonische Projektion π an und erhalten

$$0 = \pi(0) = \pi\left(\sum_{i=1}^{s} w_i\right) = \sum_{i=1}^{s} \underbrace{\pi(w_i)}_{\in \pi(W_i)}.$$

Da die Summe der Unterräume $\pi(W_1), \ldots, \pi(W_s)$ von V/U direkt ist, gilt nach Satz 2.3.6 $\pi(w_i) = 0$ für alle $i \leq s$. Das heißt, $w_i \in \operatorname{Ker}(\pi) \cap W_i = U \cap W_i$. Nach Voraussetzung ist auch die Summe der Unterräume $U \cap W_1, \ldots, U \cap W_s$ direkt und aus der Gleichung $\sum_{i=1}^{s} w_i = 0$ und Satz 2.3.6 folgt somit $w_i = 0$ für alle $i \in \{1, 2, \ldots, s\}$. □

Bei der Arbeit mit Faktorräumen muss man aufpassen, wenn man eine lineare Abbildung $\varphi\colon V/U \to W$ definieren möchte. Es ist verlockend das Bild $\varphi(v + U)$ mithilfe des Stützvektors v zu definieren. Da der Stützvektor nicht eindeutig ist, muss man sicherstellen, dass die Wahl des Stützvektors irrelevant ist; d.h., dass die Abbildung *wohldefiniert* ist. Sehen wir uns dazu wieder die Parallelenschar an.

2.4.15 **Eine Abbildung auf der Parallelenschar**. Im reellen Vektorraum $\mathbb{R}^2$ betrachten wir den ein-dimensionalen Unterraum $U = \langle \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rangle$ mit der zugehörigen Parallelenschar $\mathcal{G} = \mathbb{R}^2/U$.

Können wir eine Abbildung $\varphi\colon \mathcal{G} \to \mathbb{R}$ mit der Abbildungsvorschrift

$$\varphi\colon \begin{pmatrix} x \\ y \end{pmatrix} + U \mapsto x - y$$

definieren? Wir müssen prüfen, dass die Vorschrift nicht vom Stützvektor abhängt. Angenommen es sei $\begin{pmatrix} x \\ y \end{pmatrix} + U = \begin{pmatrix} x' \\ y' \end{pmatrix} + U$; das heißt, es gilt

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x' + t \\ y' + t \end{pmatrix}$$

für ein $t \in \mathbb{R}$. Dann ist aber auch

$$x - y = x' + t - (y' + t) = x' - y'$$

und φ ist wohldefiniert.

Man kann anschaulich verstehen, was diese Abbildung φ ist. Die Parallelen in $\mathcal{G}$ sind genau die Niveaulinien von $x - y$, d.h., für jedes $c \in \mathbb{R}$ beschreibt die Gleichung

$$x - y = c$$

ein Element der Parallelenschar $\mathcal{G}$. Die Abbildung φ ordnet jedem Element der Parallelenschar den entsprechenden Wert c zu.

2.4.16 **Aufgabe**. Sei $\mathcal{G} = \mathbb{R}^2/U$ die Paralellenschar aus 2.4.15. Definiert die Vorschrift **L**

$$\begin{pmatrix} x \\ y \end{pmatrix} + U \mapsto 3x - 2y$$

eine Abbildung $\mathcal{G} \to \mathbb{R}$?

II. Die Dimension des Faktorraumes

Schließlich wollen wir uns überlegen, wie V/U „aussieht“. Welche Dimension hat V/U? Wie findet man eine Basis von V/U? Sehen wir uns dazu wieder die Parallelenschar an.

2.4.17 **Die Dimension der Parallelenschar**. Im zwei-dimensionalen reellen Vektorraum $\mathbb{R}^2$ betrachten wir den ein-dimensionalen Unterraum

$$U = \langle \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rangle$$

mit der zugehörigen Parallelenschar $\mathcal{G} = \mathbb{R}^2/U$.

Nun benötigen wir ein Komplement zu U in $\mathbb{R}^2$. In diesem Fall ist ein Komplement W von U einfach ein von U verschiedener ein-dimensionaler Unterraum. Konkret wählen wir

$$W = \langle \begin{pmatrix} -1 \\ 2 \end{pmatrix} \rangle.$$

Das Komplement W schneidet jede Gerade $G \in \mathcal{G}$ in genau einem Punkt! Und umgekehrt liegt jeder Punkt des Unterraumes W auf genau einer Geraden aus $\mathcal{G}$. Diese Beobachtung liefert uns eine Bijektion zwischen W und $\mathcal{G}$.

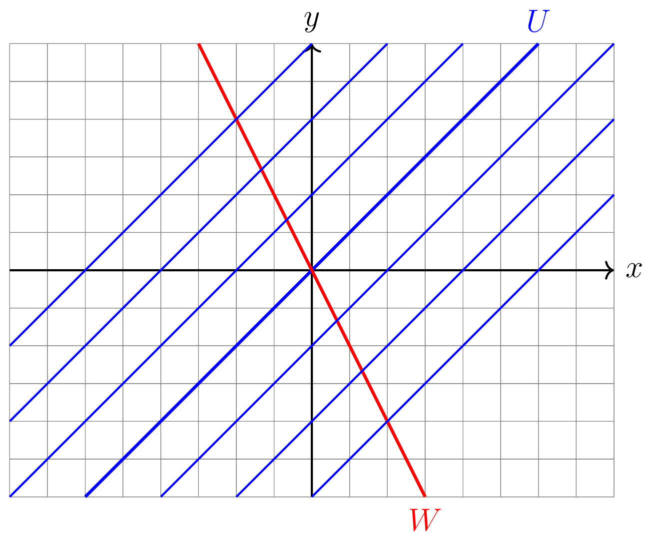


Abbildung 2.3.: Die komplementäre Gerade W schneidet jedes Element von $\mathcal{G}$ in genau einem Punkt.

Wir werden nun sehen, dass diese Bijektion wieder eine K-lineare Abbildung ist. Damit sind $\mathcal{G}$ und die komplementäre Gerade W isomorph. Insbesondere hat der Faktorraum $\mathcal{G} = \mathbb{R}^2/U$ die Dimension 1.

2.4.18 **Satz.** *Sei V ein K-Vektorraum und sei $U \subseteq V$ ein Unterraum. Ist W ein Komplement zu U in V, dann ist die Einschränkung der kanonischen Projektion auf W*

$$\pi|_W : W \to V/U$$

ein Isomorphismus von K-Vektorräumen.

Insbesondere gilt: Ist $\mathcal{B}$ eine Basis von W, dann ist $\pi(\mathcal{B})$ eine Basis von V/U.

Beweis. Wir wissen bereits, dass π linear ist. Wir zeigen nun, dass $\pi|_W$ bijektiv ist.

Surjektiv: Sei $v \in V$. Es gilt $V = U \oplus W$, also können wir $v = u + w$ mit $u \in U$ und $w \in W$ schreiben. Dann gilt $v - w \in U$ und mit Lemma 2.4.5 folgt

$$\pi|_W(w) = \pi(w) = w + U = v + U.$$

Da v beliebig war, ist $\pi|_W$ surjektiv.

Injektiv: Da W ein Komplement zu U ist und $\mathrm{Ker}(\pi) = U$ ist (siehe 2.4.13), gilt

$$\mathrm{Ker}(\pi|_W) = \mathrm{Ker}(\pi) \cap W = U \cap W = \{0\}.$$

Wegen [MG, Prop. 8.3.10] ist $\pi|_W$ damit injektiv.

Der Zusatz folgt aus der Feststellung, dass jeder Isomorphismus von Vektorräumen Basen auf Basen abbildet; vgl. [MG, Kor. 8.3.18]. □

2.4.19 **Beispiel**. Wir betrachten den Vektorraum $\mathbb{C}^3$ mit dem ein-dimensionalen Unterraum

$$U = \langle \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix} \rangle.$$

Wir wollen eine Basis des Faktorraumes $\mathbb{C}^3/U$ bestimmen.

Wir suchen ein Komplement von U in $\mathbb{C}^3$. Dazu ergänzen wir $v_1 = \left(\begin{smallmatrix} 1 \\ i \\ 0 \end{smallmatrix}\right)$ zu einer Basis von $\mathbb{C}^3$; vgl. Beispiel 2.3.16. Zum Beispiel wählen wir $v_2 = \left(\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right)$ und $v_3 = \left(\begin{smallmatrix} 0 \\ 0 \\ 1 \end{smallmatrix}\right)$ (Hier muss man einmal nachprüfen, dass diese Vektoren linear unabhängig sind). Dann ist $W = \langle v_2, v_3 \rangle$ ein Komplement und $\{v_2, v_3\}$ ist eine Basis von W. Also bilden die Elemente

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + U \quad \text{und} \quad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + U$$

eine Basis des Faktorraumes $\mathbb{C}^3/U$.

2.4.20 **Aufgabe**. Es seien V ein K-Vektorraum und $U \subseteq V$ ein Unterraum. Gegeben sind zwei linear unabhängige Vektoren $v_1, v_2 \in V$. Zeigen Sie: $v_1 + U$ und $v_2 + U$ sind linear unabhängig in V/U genau dann, wenn $\langle v_1, v_2 \rangle \cap U = \{0\}$ gilt. **L**

Aus Satz 2.4.18 wollen wir noch einige nützliche Folgerungen ableiten.

2.4.21 **Korollar.** *Sei V ein Vektorraum und sei U ein Unterraum. Sind W_1, W_2 Komplemente zu U in V, dann gilt $W_1 \cong W_2$.*

Beweis. Aus Satz 2.4.18 folgt $W_1 \cong V/U \cong W_2$. □

2.4.22 **Korollar.** *Sei V ein endlich-dimensionaler K-Vektorraum und sei $U \subseteq V$ ein Unterraum. Dann gilt die Dimensionsformel*

$$\dim_K(V) = \dim_K(U) + \dim_K(V/U).$$

Beweis. Nach dem Satz vom Komplement 2.3.17 gibt es ein Komplement W zu U in V und es gilt

$$\dim_K(V) = \dim_K(U) + \dim_K(W);$$

siehe 2.3.15. Da W und V/U isomorph sind, gilt $\dim_K(W) = \dim_K(V/U)$. □

2.5. Der Dualraum

In diesem Abschnitt lernen wir eine weitere wichtige Konstruktion kennen, die es erlaubt aus einem Vektorraum V einen neuen Vektorraum V^* zu erhalten - den sogenannten „Dualraum“ zu V. Der Dualraum besteht aus allen K-linearen Abbildungen von V in den Körper K.

2.5.1 **Erinnerung: Vektorräume von linearen Abbildungen**. Es sei K ein Körper. Sind V, W zwei Vektorräume über K, dann bezeichnet $\mathrm{Hom}_K(V, W)$ die Menge der K-linearen Abbildungen von V nach W. Die Bezeichung Hom ist gebräuchlich, da man lineare Abbildungen auch *Homomorphismen* nennt.

Aus den Mathematischen Grundlagen ist bekannt, dass $\mathrm{Hom}_K(V, W)$ selbst wieder ein K-Vektorraum ist; siehe [MG, 9.1.2]. Die Addition zweier K-linearer Abbildungen $\varphi, \psi\colon V \to W$ ist dabei punktweise definiert, d.h.,

$$(\varphi + \psi)(v) := \varphi(v) + \psi(v)$$

für alle $v \in V$. Ebenso ist die Skalarmultiplikation punktweise definiert, d.h.

$$(a\varphi)(v) := a\varphi(v)$$

für alle $a \in K$ und $\varphi \in \mathrm{Hom}_K(V, W)$.

Sind V und W endlich-dimensional, sagen wir $\dim_K(V) = n$ und $\dim_K(W) = m$, dann ist $\mathrm{Hom}_K(V, W)$ isomorph zum Vektorraum $\mathrm{M}_{m,n}(K)$ aller $m \times n$-Matrizen; siehe [MG, 9.1.9]. Insbesondere gilt dann

$$\dim_K(\mathrm{Hom}_K(V, W)) = mn. \tag{2.5.a}$$

Die Menge $\mathrm{Hom}_K(V, W)$ besteht aus Abbildungen von V nach W. Es stellt sich heraus, dass dies sogar ein Unterraum von $\mathrm{Abb}(V, W)$ ist.

2.5.2 **Aufgabe**. Zeigen Sie, dass $\mathrm{Hom}_K(V, W)$ ein Unterraum von $\mathrm{Abb}(V, W)$ ist. **L**

I. Dualraum und Dualbasen

2.5.3 **Definition** Sei K ein Körper und sei V ein K-Vektorraum. Man nennt

$$V^* := \mathrm{Hom}_K(V, K)$$

den *Dualraum* zu V. Die Elemente im Dualraum nennt man auch *Linearformen*.

2.5.4 **Bemerkung**. (a) Ist V ein endlich-dimensionaler K-Vektorraum der Dimension n, dann folgt aus (2.5.a), dass der Dualraum V^* ebenfalls Dimension n hat. Die Vektorräume V und V^* sind also isomorph; siehe [MG, 8.3.19]. Dennoch sollte man sich V und V^* nicht als „gleich" vorstellen, denn es gibt keinen *kanonischen* Isomorphismus zwischen diesen Räumen. Ein Isomorphismus ist „kanonisch", wenn man ihn ohne eine Auswahl zu treffen (z.B. die Wahl einer Basis) angeben kann.

(b) Ist V ein unendlich-dimensionaler Vektorraum, dann sind V und V^* *nicht* isomorph. Diese Aussage ist nicht offensichtlich und hat damit zu tun, dass es unendliche Mengen verschiedener Mächtigkeit gibt.

Die Elemente im Dualraum sind Linearformen, also lineare Abbildungen. Das macht es etwas schwierig, sich den Dualraum vorzustellen. Wir wollen es im nächsten Beispiel dennoch versuchen.

2.5.5 **Beispiel**. Sei K ein Körper. Wir betrachten den Vektorraum $V = K^n$ aller Spaltenvektoren der Länge n. Es sei $e_1, e_2, \dots, e_n$ die Standardbasis. Eine Linearform $\alpha \in V^*$ ist eindeutig durch die Werte

$$\alpha(e_1), \alpha(e_2), \dots, \alpha(e_n)$$

auf der Standardbasis bestimmt. Es sei $\mathrm{M}_{1,n}(K)$ die Menge der Zeilenvektoren der Länge n. Wir definieren die K-lineare Abbildung

$$E\colon V^* \to \mathrm{M}_{1,n}(K) \quad \text{durch} \quad \alpha \mapsto (\alpha(e_1), \alpha(e_2), \dots, \alpha(e_n)).$$

Wegen Satz 9.1.9 [MG] ist E ein Isomorphismus von Vektorräumen. Unter Verwendung dieses Isomorphismus kann man sich den Dualraum zum Raum der Spaltenvektoren K^n also als den Raum der Zeilenvektoren vorstellen.

Aber weshalb verwendet man hier überhaupt Zeilenvektoren? Eine Linearform α ist eine lineare Abbildung und $E(\alpha) = {}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\alpha)$ ist die Matrixdarstellung von α zur Standardbasis $\mathcal{B} = (e_1, \dots, e_n)$ von K^n und der Basis $\mathcal{C} = (1)$ von K. Das bedeutet, es gilt mit (2.2.b)

$$\alpha(v) = {}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\alpha)v = E(\alpha)v.$$

Der Zeilenvektor $E(\alpha)$ liefert also mithilfe der Matrix-Vektor-Multiplikation die Linearform α auf K^n.

Im folgenden Lemma wollen wir uns davon überzeugen, dass der Dualraum immer ausreichend viele Linearformen enthält, sodass jeder Vektor (außer dem Nullvektor) auf einen Wert ungleich Null abgebildet werden kann. Dazu verwenden wir den Satz vom Komplement.

2.5.6 **Lemma.** *Sei V ein K-Vektorraum. Für alle $v \neq 0$ in V gibt es eine Linearform $\alpha \in V^*$ mit $\alpha(v) \neq 0$.*

Beweis. Es sei $U = \langle v \rangle$ der von v aufgespannte Unterraum. Nach dem Satz vom Komplement 2.3.17 finden wir ein Komplement $W \subseteq V$, d.h., $U \oplus W = V$. Jeder Vektor $x \in V$ lässt sich also eindeutig in der Form

$$x = cv + w$$

mit $c \in K$ und $w \in W$ schreiben. Damit definieren wir eine Linearform α auf V, indem wir $\alpha(cv + w) = c$ definieren, wenn $w \in W$ und $c \in K$ ist. Das ist eine lineare Abbildung, denn ist $x = cv + w$ und $x' = c'v + w'$ mit $c, c' \in K$ und $w, w' \in W$, dann gilt

$$\alpha(ax + x') = \alpha((ac + c')v + \underbrace{aw + w'}_{\in W}) = ac + c' = a\alpha(x) + \alpha(x')$$

für alle $a \in K$. Aus Aufgabe 2.2.18 folgt nun, dass α linear ist, d.h., $\alpha \in V^*$. Es gilt $\alpha(v) = 1$ und damit ist das Lemma bewiesen. □

2.5.7 **Aufgabe**. Wir betrachten den Vektorraum $\mathbb{F}_7^3$ und den Vektor L

$$v = \begin{pmatrix} \overline{0} \\ \overline{3} \\ \overline{2} \end{pmatrix}.$$

Geben Sie eine Linearform $\alpha \in (\mathbb{F}_7^3)^*$ an, die $\alpha(v) \neq 0$ erfüllt.

2.5.8 **Notation**. Sei V ein endlich-dimensionaler K-Vektorraum. Ist $v_1, \ldots, v_n$ eine Basis von V, dann definieren wir Linearformen $v_1^*, \ldots, v_n^* \in V^*$ durch

$$v_i^*(v_j) = \delta_{i,j};$$

dabei bezeichnet δ das *Kronecker-Symbol*

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

wobei mit 1 und 0 immer die Elemente im Körper K gemeint sind. Als lineare Abbildung ist v_i^* eindeutig durch die Werte auf der Basis $v_1, \ldots, v_n$ definiert.

2.5.9 **Vorsicht mit der Notation v_i^*.** Die Notation v_i^* ist gefährlich, denn die Linearform v_i^* hängt nicht nur von v_i sondern von der ganzen Basis $v_1, \ldots, v_n$ ab. Wir verwenden die Notation daher nur, wenn vorher eine Basis fest gewählt wurde.

2.5.10 **Definition und Satz.** *Sei V ein endlich-dimensionaler K-Vektorraum. Sei $\mathcal{B} = \{v_1, v_2, \dots, v_n\}$ eine Basis von V. Dann ist $\mathcal{B}^* = \{v_1^*, v_2^*, \dots, v_n^*\}$ eine Basis von V^*.*

Man nennt $\mathcal{B}^*$ die Dualbasis *zu $\mathcal{B}$.*

Beweis. Wir wissen bereits aus (2.5.a), dass V und V^* dieselbe Dimension haben, d.h., $\dim_K(V^*) = n$. Es ist daher ausreichend zu zeigen, dass die Linearformen $v_1^*, \dots, v_n^*$ linear unabhängig sind; siehe [MG, Prop. 7.4.4].

Dazu seien $a_1, a_2, \dots, a_n \in K$ mit

$$\sum_{i=1}^{n} a_i v_i^* = 0$$

gegeben. Was bedeutet es, dass eine Linearform gleich 0 ist? Das heißt, das man beim Einsetzen eines beliebigen Vektors immer das Ergebnis 0 erhält. Setzen wir nun den Basisvektor v_j ein, dann erhalten wir

$$0 = \sum_{i=1}^{n} a_i v_i^*(v_j) = \sum_{i=1}^{n} a_i \delta_{i,j} = a_j.$$

Da $j \in \{1, 2, \dots, n\}$ beliebig war, sind die Linearformen $v_1^*, \dots, v_n^*$ damit linear unabhängig. □

2.5.11 **Bemerkung**. Ist $\mathcal{B} = (v_1, \dots, v_n)$ eine geordnete Basis von V, dann nennen wir $\mathcal{B}^* := (v_1^*, \dots, v_n^*)$ die *geordnete Dualbasis* zu $\mathcal{B}$.

2.5.12 **Beispiel**. Wir betrachten den reellen Vektorraum $\mathbb{R}^2$. Mithilfe des Isomorphismus $E\colon V^* \to \mathrm{M}_{1,2}(\mathbb{R})$ aus 2.5.5 stellen wir uns den Dualraum $(\mathbb{R}^2)^*$ als Raum der Zeilenvektoren der Länge 2 vor.

Gegeben ist die Basis

$$v_1 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 3 \end{pmatrix} \in \mathbb{R}^2.$$

Wir suchen die Dualbasis v_1^*, v_2^* oder genauer die zugehörigen Zeilenvektoren $u_1 = E(v_1^*) = (u_{1,1}, u_{1,2})$ und $u_2 = E(v_2^*) = (u_{2,1}, u_{2,2})$. Diese Zeilenvektoren sind durch die Gleichungen

$$u_1 v_1 = 1, \quad u_1 v_2 = 0, \quad u_2 v_1 = 0, \quad u_2 v_2 = 1 \tag{2.5.b}$$

eindeutig bestimmt. Dies führt zu einem linearen Gleichungssystem

$$\begin{array}{rcrcl} 2u_{1,1} & - & u_{1,2} & = & 1 \\ u_{1,1} & + & 3u_{1,2} & = & 0 \\ 2u_{2,1} & - & u_{2,2} & = & 0 \\ u_{2,1} & + & 3u_{2,2} & = & 1 \end{array}$$

für das man mit etwas Rechnen die Lösung

$$u_1 = (\frac{3}{7}, -\frac{1}{7}), \quad u_2 = (\frac{1}{7}, \frac{2}{7})$$

findet.

Man kann die Lösung im Beispiel 2.5.12 auch etwas eleganter herleiten, wenn man die folgende Beobachtung einsetzt.

2.5.13 **Die Dualbasis ausrechnen**. Es sei $v_1, \dots, v_n$ eine Basis der Raumes K^n. Es sei $v_1^*, \dots, v_n^*$ die Dualbasis von $(K^n)^*$. Die zugehörigen Zeilenvektoren nennen wir $u_i \coloneqq E(v_i^*)$ für $1 \leq i \leq n$. Nach Definition der Dualbasis gilt

$$u_i v_j = \delta_{i,j}$$

für alle $i, j \in \{1, 2, \dots, n\}$. Ist A die Matrix deren Spalten $v_1, \dots, v_n$ sind und ist B die Matrix, deren Zeilen $u_1, \dots, u_n$ sind, dann gilt also $BA = I_n$. Damit ist A invertierbar und es gilt $B = A^{-1}$; siehe [MG, 4.5.6].

Wir halten fest: Schreibt man die Basisvektoren $v_1, \dots, v_n$ in eine Matrix A, dann sind die Zeilen der inversen Matrix A^{-1} genau die Zeilenvektoren $E(v_1^*), E(v_2^*), \dots, E(v_n^*)$, die die Dualbasis beschreiben.

2.5.14 **Beispiel**. Versuchen wir das nochmal im Beispiel 2.5.12. Sei

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

die Matrix mit den Spalten v_1, v_2. Die inverse Matrix ist

$$A^{-1} = \frac{1}{7} \begin{pmatrix} 3 & -1 \\ 1 & 2 \end{pmatrix}.$$

Wir lesen die Dualbasis aus den Zeilen der inversen Matrix A^{-1} ab und erhalten

$$u_1 = (\frac{3}{7}, -\frac{1}{7}) \quad \text{und} \quad u_2 = (\frac{1}{7}, \frac{2}{7}).$$

2.5.15 **Aufgabe**. Wir betrachten den Vektorraum $\mathbb{R}^3$ mit der Basis **L**

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ -1 \\ 4 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \in \mathbb{R}^3.$$

Geben Sie die Dualbasis v_1^*, v_2^*, v_3^* in Form von Zeilenvektoren an.

II. Duale Abbildungen

Wie wir gerade gesehen haben, kann man zu jedem Vektorraum V seinen Dualraum V^* definieren. Hier werden wir uns jetzt überlegen, wie sich diese Konstruktion mit linearen Abbildungen zwischen Vektorräumen verträgt.

Es sei $\varphi\colon V \to W$ eine lineare Abbildung und es sei $\alpha\colon W \to K$ eine Linearform auf W. Die Komposition linearer Abbildungen ist wieder eine lineare Abbildung (siehe [MG, 8.1.9]). Also ist dann

$$\alpha \circ \varphi\colon V \to K$$

eine Linearform auf V. Dies definiert eine Abbildung vom Dualraum W^* in den Dualraum V^*.

2.5.16 **Definition** Sei $\varphi\colon V \to W$ eine lineare Abbildung zwischen zwei Vektorräumen V und W. Man nennt die Abbildung

$$\varphi^*\colon W^* \to V^* \quad \text{mit} \quad \varphi^*(\alpha) := \alpha \circ \varphi$$

die *duale Abbildung* zu φ.

2.5.17 **Beispiel**. Wir betrachten die identische Abbildung $\mathrm{id}_V\colon V \to V$. Die duale Abbildung dazu ist die identische Abbildung auf V^*, denn es gilt

$$\mathrm{id}_V^*(\alpha) = \alpha \circ \mathrm{id}_V = \alpha$$

für alle $\alpha \in V^*$.

2.5.18 **Satz.** *Sei $\varphi\colon V \to W$ eine K-lineare Abbildung zwischen zwei K-Vektorräumen V und W. Die duale Abbildung $\varphi^*\colon W^* \to V^*$ ist K-linear.*

Beweis. Es seien $\alpha, \beta \in W^*$. Dann gilt für alle $v \in V$

$$\begin{aligned}
\varphi^*(\alpha+\beta)(v) &= (\alpha+\beta)(\varphi(v)) && \text{(Def. von } \varphi^*\text{)}\\
&= \alpha(\varphi(v)) + \beta(\varphi(v)) && \text{(Def. von + in } W^*\text{)}\\
&= \varphi^*(\alpha)(v) + \varphi^*(\beta)(v) && \text{(Def. von } \varphi^*\text{)}\\
&= \big(\varphi^*(\alpha) + \varphi^*(\beta)\big)(v). && \text{(Def. von + in } V^*\text{)}
\end{aligned}$$

Da $v \in V$ beliebig war, schließen wir $\varphi^*(\alpha+\beta) = \varphi^*(\alpha) + \varphi^*(\beta)$.

Sei nun $a \in K$ beliebig. Dann gilt für alle $v \in V$

$$\varphi^*(a\alpha)(v) = (a\alpha)(\varphi(v)) = a\alpha(\varphi(v)) = a\varphi^*(\alpha)(v) = (a\varphi^*(\alpha))(v).$$

Daraus folgt $\varphi^*(a\alpha) = a\varphi^*(\alpha)$, das v beliebig war. Damit haben wir gezeigt, dass φ^* eine K-lineare Abbildung ist. □

Der folgende Satz beschreibt, wie sich die Komposition von Abbildungen und die Dualisierung miteinander vertragen. In der Sprache der Kategorientheorie würde man sagen, dass die Dualisierung einen *kontravarianten Funktor* auf der Kategorie der K-Vektorräume definiert.

2.5.19 **Satz.** *Seien U, V, W drei K-Vektorräume mit K-linearen Abbildungen $\psi\colon U \to V$ und $\varphi\colon V \to W$. Es gilt*

$$(\varphi \circ \psi)^* = \psi^* \circ \varphi^*.$$

Beweis. Für alle $\alpha \in W^*$ gilt

$$(\varphi\circ\psi)^*(\alpha) = \alpha\circ(\varphi\circ\psi) = (\alpha\circ\varphi)\circ\psi = \psi^*(\alpha\circ\varphi) = \psi^*(\varphi^*(\alpha)) = (\psi^*\circ\varphi^*)(\alpha);$$

dabei verwenden wir, dass die Komposition von Abbildungen assoziativ ist. Da α beliebig war, folgt daraus $(\varphi\circ\psi)^* = \psi^*\circ\varphi^*$. □

Wir wissen, dass man lineare Abbildungen zwischen endlich-dimensionalen Vektorräumen mithilfe der Matrixdarstellung beschreiben kann, sobald man geordnete Basen der Räume gewählt hat. Wir werden uns nun überlegen, wie die Darstellungsmatrix der dualen Abbildung bezüglich der entsprechenden Dualbasen aussieht.

2.5.20 **Satz.** *Sei $\varphi\colon V \to W$ eine lineare Abbildung zwischen zwei endlich-dimensionalen K-Vektorräumen. Es sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine geordnete Basis von V und es sei $\mathcal{C} = (w_1, \ldots, w_m)$ eine geordnete Basis von W.*

Sind $\mathcal{B}^*$ und $\mathcal{C}^*$ die geordneten Dualbasen zu $\mathcal{B}$ und $\mathcal{C}$, dann gilt

$${}_{\mathcal{B}^*}\mathrm{M}_{\mathcal{C}^*}(\varphi^*) = {}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\varphi)^T.$$

Beweis. Sei ${}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\varphi) = (a_{ij})_{\substack{i \le m \\ j \le n}}$, d.h., $a_{ij} \in K$ ist der Eintrag von ${}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\varphi)$ an der Stelle (i, j). Es gilt also

$$\varphi(v_j) = \sum_{i=1}^{m} a_{ij} w_i \tag{2.5.c}$$

für alle $j \in \{1, \dots, n\}$.

Genauso sei ${}_{\mathcal{B}^*}\mathrm{M}_{\mathcal{C}^*}(\varphi^*) = (b_{jk})_{\substack{j \le n \\ k \le m}}$. Es gilt also

$$\varphi^*(w_k^*) = \sum_{j=1}^{n} b_{jk} v_j^* \tag{2.5.d}$$

für alle $k \in \{1, \dots, m\}$. Für alle $\ell \in \{1, \dots, n\}$ und $k \in \{1, \dots, m\}$ erhalten wir

$$\begin{aligned}
b_{\ell k} &= \sum_{j=1}^{n} b_{jk} v_j^*(v_\ell) \\
&= \varphi^*(w_k^*)(v_\ell) && \text{(2.5.d)} \\
&= (w_k^* \circ \varphi)(v_\ell) = w_k^*(\varphi(v_\ell)) && \text{(Def. duale Abb.)} \\
&= w_k^*\Big(\sum_{i=1}^{m} a_{i\ell} w_i\Big) && \text{(2.5.c)} \\
&= \sum_{i=1}^{m} a_{i\ell} w_k^*(w_i) = a_{k\ell}.
\end{aligned}$$

Aus der Gleichung $b_{\ell k} = a_{k\ell}$ für alle $k \in \{1, \dots, m\}$, $\ell \in \{1, \dots, n\}$ folgt nun

$${}_{\mathcal{B}^*}\mathrm{M}_{\mathcal{C}^*}(\varphi^*) = {}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\varphi)^T. \qquad \square$$

2.6. Hilfsmittel: Äquivalenzrelationen

In diesem Abschnitt lernen wir ein grundlegendes Konzept kennen, das in zahlreichen Ausprägungen in allen Teilen der Mathematik vorkommt: *Äquivalenzrelationen*. Dieses Konzept ist immer dann nützlich, wenn man auf den Elementen einer Menge M einen Begriff von „Ähnlichkeit“ definieren möchte. Zum Beispiel könnte man die Menge aller Autos betrachten und zwei Autos „ähnlich“ nennen, wenn sie dieselbe Farbe haben. Die Menge aller Autos zerfällt dann offensichtlich in disjunkte Teilmengen: die Menge der roten Autos, der schwarzen Autos, der blauen Autos, usw. Alternativ könnte man zwei Autos „ähnlich“ nennen, wenn Sie vom selben Hersteller sind. Wieder zerfällt die Gesamtheit aller Autos in Teilmengen: die Menge aller Autos von BMW, die Menge aller Autos von Volkswagen, die Menge aller Autos von Opel, usw. Man nennt diese Teilmengen die „Äquivalenzklassen“. Wir werden sehen, dass Äquivalenzrelationen auch in der Linearen Algebra ein hilfreicher Begriff sind, weil wir damit in der Lage sind die wesentlichen Fragestellungen aus verschiedenen Blickwinkeln zu beschreiben.

I. Definition und Beispiele

Wir beginnen mit einer mathematischen Formulierung des Begriffes „Äquivlenzrelation“. Die Definition wirkt auf den ersten Blick sehr abstrakt, aber davon sollte man sich nicht abschrecken lassen. Wenn man sich mit einigen konkreten Beispielen beschäftigt, wird die Definition greifbar. Zur Definition benötigen wir binäre Relationen (siehe 2.1.3).

2.6.1 **Definition** Sei M eine Menge. Eine binäre Relation $\sim$ auf M ist eine *Äquivalenzrelation*, wenn sie die folgenden drei Eigenschaften hat:

(Ä1) $\sim$ ist *reflexiv*: Für alle $x \in M$ gilt

$$x \sim x.$$

(Ä2) $\sim$ ist *symmetrisch*: Für alle $x, y \in M$ gilt

$$x \sim y \quad \Longrightarrow \quad y \sim x.$$

(Ä3) $\sim$ ist *transitiv*: Für alle $x, y, z \in M$ gilt

$$x \sim y \text{ und } y \sim z \quad \Longrightarrow \quad x \sim z.$$

2.6.2 **Beispiel**. Es sei M die Menge aller Autos. Wir schreiben $x \sim y$, wenn Auto x und Auto y dieselbe Farbe haben. Dies ist eine Äquivalenzrelation. Die Relation $\sim$ ist reflexiv, denn jedes Auto x hat dieselbe Farbe wie Auto x. Die Relation ist auch symmetrisch, denn wenn x dieselbe Farbe wie y hat, dann hat auch y dieselbe Farbe wie x. Auch die Transitivität der Relation ist erfüllt, denn wenn x dieselbe Farbe wie y hat und y dieselbe Farbe wie z, dann haben auch x und z dieselbe Farbe.

Dieses einfache Beispiel ist ein Spezialfall, der folgenden allgemeineren Konstruktion.

2.6.3 **Beispiel** (Faserrelationen). Es seien M und N zwei Mengen und $f\colon M \to N$ sei eine Abbildung. Es seien $x, y \in M$. Wir schreiben $x \sim_f y$ genau dann, wenn $f(x) = f(y)$ ist. Dann ist $\sim_f$ eine binäre Relation auf M; man nennt diese Relation, die *Faserrelation* zu f. Die Faserrelation ist eine Äquivalenzrelation.

reflexiv: Für alle $x \in M$ gilt $f(x) = f(x)$ und damit $x \sim_f x$.

symmetrisch: Seien $x, y \in M$ mit $x \sim_f y$. Dann gilt $f(x) = f(y)$ und damit auch $f(y) = f(x)$, d.h., $y \sim_f x$.

transitiv: Seien $x, y, z \in M$. Es gelte $x \sim_f y$ und $y \sim_f z$. Dann ist $f(x) = f(y) = f(z)$ und damit gilt auch $x \sim_f z$.

2.6.4 **Beispiel**. Es sei $n \in \mathbb{N}$ eine natürliche Zahl. Auf der Menge der ganzen Zahlen definieren wir eine Relation $\equiv_n$: Für $a, b \in \mathbb{Z}$ schreiben wir $a \equiv_n b$ genau dann, wenn n ein Teiler von $a - b$ ist. Die Relation $\equiv_n$ ist eine Äquivalenzrelation auf der Menge $\mathbb{Z}$.

reflexiv: Für alle $a \in \mathbb{Z}$ ist n ein Teiler von $0 = a - a$, denn $0 \cdot n = 0$. Also gilt $a \equiv_n a$.

symmetrisch: Seien $a, b \in \mathbb{Z}$ mit $a \equiv_n b$ gegeben. Es gilt also $a - b = kn$ für ein $k \in \mathbb{Z}$. Also ist $b - a = (-k)n$ und es gilt auch $b \equiv_n a$.

transitiv: Seien $a, b, c \in \mathbb{Z}$. Wir nehmen an, dass $a \equiv_n b$ und $b \equiv_n c$ erfüllt sind. Es gibt also ganze Zahlen $k, \ell \in \mathbb{Z}$ mit $a - b = kn$ und $b - c = \ell n$. Dann gilt auch

$$a - c = a - b + b - c = kn + \ell n = (k + \ell)n,$$

d.h. n teilt $a - c$ und damit gilt $a \equiv_n c$.

2.6.5 **Aufgabe**. Auf der Menge $\mathbb{C}^\times$ der komplexen Zahlen ohne 0 definieren wir eine binäre Relation $\smile$. Für $z, w \in \mathbb{C}^\times$ schreiben wir $z \smile w$, wenn es eine reelle Zahl $r \in \mathbb{R}^\times$ mit $rz = w$ gibt. Zeigen Sie, dass $\smile$ eine Äquivalenzrelation ist. **L**

2.6.6 $\boxed{\textbf{Definition}}$ Sei M eine Menge mit einer Äquivalenzrelation $\sim$. Für $x \in M$ nennen wir die Menge

$$[x]_\sim = \{y \in M \mid x \sim y\}$$

die *Äquivalenzklasse von* x bzgl. $\sim$.

2.6.7 **Beispiel**. Seien M, N Mengen und sei $f\colon M \to N$ eine Abbildung. Was ist die Äquivalenzklasse von $x \in M$ bezüglich der Faserrelation $\sim_f$? Per Definition gilt $x \sim_f y$ genau dann, wenn $f(x) = f(y)$ ist. Damit ist

$$[x]_{\sim_f} = \{y \in M \mid f(y) = f(x)\} = f^{-1}(f(x))$$

genau das Urbild von $f(x)$ bzgl. f.

2.6.8 **Satz.** *Sei M eine Menge mit einer Äquivalenzrelation $\sim$. Je zwei Äquivalenzklassen $[x]_\sim$ und $[y]_\sim$ sind entweder gleich oder disjunkt.*

Beweis. Seien $x, y \in M$ und nehmen wir an, die Äquivalenzklassen $[x]_\sim$ und $[y]_\sim$ sind nicht disjunkt. Es gibt also ein Element $z \in [x]_\sim \cap [y]_\sim$. Wir möchten nun zeigen, dass die Klassen dann bereits gleich sind. Dazu leiten wir die Inklusion $[x]_\sim \subseteq [y]_\sim$ her. Wiederholt man dieses Argument mit vertauschten Rollen, so folgt daraus die Gleichheit der Äquivalenzklassen.

Sei $a \in [x]_\sim$ beliebig. Es gilt also $x \sim a$. Es gilt $z \in [x]_\sim$, also $x \sim z$ und durch Symmetrie auch $z \sim x$. Aus der Transitivität von $\sim$ schließen wir $z \sim a$. Weiter ist auch $z \in [y]_\sim$, d.h., es ist $y \sim z$. Wir verwenden nochmal, dass $\sim$ transitiv ist und erhalten $y \sim a$. Damit liegt a in der Äquivalenzklasse $[y]_\sim$. Da $a \in [x]_\sim$ beliebig war, erhalten wir $[x]_\sim \subseteq [y]_\sim$. □

2.6.9 **Aufgabe**. Sei K ein Körper und sei V ein K-Vektorraum. Sei $U \subseteq V$ ein Unterraum. Für $x, y \in V$ definieren wir $x \sim_U y$ genau dann, wenn $x - y$ in U liegt. **L**

Zeigen Sie, dass $\sim_U$ eine Äquivalenzrelation auf V ist und bestimmen Sie die Äquivalenzklassen.

2.6.10 $\boxed{\textbf{Definition}}$ Sei M eine Menge mit einer Äquivalenzrelation $\sim$. Eine Teilmenge $R \subseteq M$ heißt *Vertretersystem* der Äquivalenzklassen, wenn jede Äquivalenzklasse bzgl. $\sim$ genau ein Element aus R enthält.

In anderen Worten: Ist R ein Vertretersystem, dann gibt es für jedes $x \in M$ ein *eindeutiges* Element $r \in R$ mit $x \sim r$.

2.6.11 **Beispiel**. Es sei $n \in \mathbb{N}$. Wir betrachten nochmal die Äquivalenzrelation $\equiv_n$ auf der Menge der ganzen Zahlen $\mathbb{Z}$ aus Beispiel 2.6.4. Die Äquivalenzklasse von $k \in \mathbb{Z}$ bzgl.

$\equiv_n$ besteht aus allen Zahlen $b \in \mathbb{Z}$ mit der Eigenschaft, dass $k - b$ ein Vielfaches von n ist, d.h.,

$$[k]_{\equiv_n} = \{k + qn \in \mathbb{Z} \mid q \in \mathbb{Z}\} = k + n\mathbb{Z}.$$

Die Äquivalenzklassen sind also genau die Restklassen modulo n, die wir in 1.3.4 kennengelernt haben. Im Satz 1.3.6 haben wir auch schon gesehen, dass je zwei Restklassen entweder gleich oder disjunkt sind. Aus dem Satz wissen wir auch, dass $\{0, 1, 2, \ldots, n-1\}$ ein Vertretersystem der Restklassen ist.

2.6.12 **Beispiel**. Sei K ein Körper. Für Matrizen $A, B \in \mathrm{M}_{n,n}(K)$ schreiben wir $A \sim B$, wenn es eine invertierbare Matrix $S \in \mathrm{GL}_n(K)$ gibt, sodass

$$A = S^T BS$$

gilt. Wir sagen dann, dass A und B *kongruent* sind. Dann ist $\sim$ eine Äquivalenzrelation.

reflexiv: Für alle $A \in \mathrm{M}_{n,n}(K)$ gilt $A \sim A$, denn die Einheitsmatrix I_n ist invertierbar und es gilt $A = I_n A I_n = I_n^T A I_n$.

symmetrisch: Seien $A, B \in \mathrm{M}_{n,n}(K)$ mit $A \sim B$ gegeben. Es gibt also $S \in \mathrm{GL}_n(K)$ mit $S^T BS = A$. Multipliziert man diese Gleichung von links mit $(S^{-1})^T$ und von rechts mit S^{-1} dann erhält man

$$B = (SS^{-1})^T B = (S^{-1})^T S^T B(SS^{-1}) = (S^{-1})^T AS^{-1}.$$

Da die Matrix S^{-1} ebenfalls invertierbar ist, schließen wir daraus, dass $B \sim A$ gilt.

transitiv: Seien $A, B, C \in \mathrm{M}_{n,n}(K)$ mit $A \sim B$ und $B \sim C$ gegeben. Es gibt also invertierbare Matrizen $S_1, S_2 \in \mathrm{GL}_n(K)$ mit

$$A = S_1^T BS_1 \quad \text{und} \quad B = S_2^T CS_2.$$

Daraus folgt

$$A = S_1^T BS_1 = S_1^T S_2^T CS_2 S_1 = (S_2 S_1)^T C(S_2 S_1).$$

Da $\mathrm{GL}_n(K)$ mit der Matrizenmultiplikation eine Gruppe ist (siehe 1.1.12), liegt auch $S_2 S_1$ in $\mathrm{GL}_n(K)$ (ist also invertierbar). Nach Definition von Kongruenz gilt damit $A \sim C$.

Die Äquivalenzklassen der Kongruenzrelation auf der Menge $\mathrm{M}_{n,n}(K)$ nennt man *Kongruenzklassen*. Wir werden in der nächsten Lektion sehen, warum es nützlich ist, die Kongruenzklassen von Matrizen zu studieren.

II. Beispiel: Äquivalente Matrizen

In diesem Abschnitt werden wir uns ein weiteres, sehr instruktives Beispiel ansehen. Dabei lernen wir einige Konzepte aus den Mathematischen Grundlagen aus einem neuen Blickwinkel kennen.

2.6.13 **Definition** Sei K ein Körper und seien $m, n \in \mathbb{N}$. Wir nennen Matrizen $A, B \in \mathrm{M}_{m,n}(K)$ *äquivalent*, wenn es invertierbare Matrizen $S \in \mathrm{GL}_m(K)$ und $T \in \mathrm{GL}_n(K)$ gibt, sodass

$$A = SBT.$$

Wir schreiben $A \approx B$, wenn A und B äquivalent sind.

2.6.14 **Lemma.** *Sei K ein Körper und seien $m, n \in \mathbb{N}$. Die Relation $\approx$ ist eine Äquivalenzrelation auf $\mathrm{M}_{m,n}(K)$.*

Beweis. *reflexiv:* Für alle $A \in \mathrm{M}_{m,n}(K)$ gilt $A \approx A$, denn die Einheitsmatrizen I_m, I_n sind invertierbar und es gilt $A = I_m A I_n$.

symmetrisch: Seien $A, B \in \mathrm{M}_{m,n}(K)$ mit $A \approx B$ gegeben. Es gibt also invertierbare Matrizen $S \in \mathrm{GL}_m(K)$ und $T \in \mathrm{GL}_n(K)$ mit $SBT = A$. Multipliziert man diese Gleichung von links mit S^{-1} und von rechts mit T^{-1}, dann erhalten wir

$$B = (S^{-1}S)B(TT^{-1}) = S^{-1}AT^{-1}.$$

Da die Matrizen $S^{-1} \in \mathrm{GL}_m(K)$ und $T^{-1} \in \mathrm{GL}_n(K)$ ebenfalls invertierbar sind, folgt daraus $B \approx A$.

transitiv: Seien $A, B, C \in \mathrm{M}_{m,n}(K)$ mit $A \approx B$ und $B \approx C$ gegeben. Es gibt also invertierbare Matrizen $S_1, S_2 \in \mathrm{GL}_m(K)$ und $T_1, T_2 \in \mathrm{GL}_n(K)$ mit

$$A = S_1BT_1 \quad \text{und} \quad B = S_2CT_2.$$

Daraus folgt

$$A = S_1BT_1 = S_1S_2CT_2T_1.$$

Da $\mathrm{GL}_m(K)$ und $\mathrm{GL}_n(K)$ mit der Matrizenmultiplikation Gruppen sind (siehe 1.1.12), gilt S_1S_2 in $\mathrm{GL}_m(K)$ und $T_1T_2 \in \mathrm{GL}_n(K)$. Damit sind A und C äquivalent. □

Was sind die Äquivalenzklassen dieser Relation? Wie können wir ein Vertretersystem angeben? Wenn man auf eine solche Problemstellung trifft, ist es meistens hilfreich Eigenschaften zu suchen, die *invariant* unter dem Begriff der Äquivalenz sind. Das heißt, wir suchen Eigenschaften von Matrizen, die sich nicht ändern, wenn wir von rechts oder links mit invertierbaren Matrizen multiplizieren. Aus den Mathematischen Grundlagen kennen wir eine solche Eigenschaft: den Rang einer Matrix. Die verblüffende Beobachtung ist nun, dass der Rang die Äquivalenz von Matrizen vollständig beschreibt.

2.6.15 **Satz.** *Seien* $A, B \in \mathrm{M}_{m,n}(K)$*. Dann gilt*

$$A \approx B \quad \Leftrightarrow \quad \mathrm{Rg}(A) = \mathrm{Rg}(B).$$

In anderen Worten: Die Äquivalenz von Matrizen in $\mathrm{M}_{m,n}(K)$ ist die Faserrelation der Rangabbildung $\mathrm{Rg}\colon \mathrm{M}_{m,n}(K) \to \mathbb{N}_0$.

Beweis. „$\Rightarrow$“: Wir nehmen an, dass A und B äquivalent sind. Es gibt also Matrizen $S \in \mathrm{GL}_m(K)$ und $T \in \mathrm{GL}_n(K)$ mit $A = SBT$. Aus [MG, 4.5.5 (a) und (c)] erhalten wir

$$\mathrm{Rg}(A) = \mathrm{Rg}(SBT) = \mathrm{Rg}(BT) = \mathrm{Rg}(B).$$

„$\Leftarrow$“: Wir nehmen an, dass $\mathrm{Rg}(A) = r$ ist und definieren die Matrix

$$I(r, m, n) = \begin{pmatrix} I_r & 0_{r,n-r} \\ 0_{m-r,r} & 0_{m-r,n-r} \end{pmatrix} \in \mathrm{M}_{m,n}(K).$$

Das heißt, $I(r, m, n) = \sum_{i=1}^{r} E_{ii}$, wobei E_{ii} die Elementarmatrix mit Eintrag 1 an der Stelle (i, i) bezeichnet [MG, 3.2.1]. Wir behaupten, dass $A \approx I(r, m, n)$ ist. Daraus folgt dann $A \approx B$, denn wenn B eine andere Matrix mit $\mathrm{Rg}(B) = r$ ist, so erhalten wir aus dem analogen Argument $I(r, m, n) \approx B$. Da $\approx$ eine Äquivalenzrelation ist, folgt $A \approx B$ aus der Transitivität .

Sei T die Treppennormalform zu A mit Pivot-Positionen $(1, j_1), (2, j_2), \ldots, (r, j_r)$. Es gibt also Elementarmatrizen $E_1, \ldots, E_s$ mit

$$T = E_1 \cdots E_s A.$$

Da Elementarmatrizen invertierbar sind (siehe [MG, 3.2.11]), ist $S = E_1 \cdots E_s$ eine invertierbare Matrix.

Die Transponierte T^T hat nun Einträge 1 an den Stellen $(j_1, 1), \ldots, (j_r, r)$. Durch elementare Zeilenumformungen können wir nun die übrigen Einträge von T^T an den Stellen (s, i) mit $s > j_i$ und $i \in \{1, \ldots, r\}$ zu Null machen. Damit ist die Treppennormalform von T^T die Matrix $I(r, n, m)$. Es gibt also wieder eine invertierbare

Matrix $Q \in \mathrm{GL}_n(K)$ (ein Produkt von Elementarmatrizen) mit $QT^T = I(r,n,m)$. Durch Transponieren erhalten wir

$$SAQ^T = TQ^T = (QT^T)^T = I(r,m,n),$$

denn $I(r,n,m)^T = I(r,m,n)$. Da auch Q^T invertierbar ist (siehe Aufgabe 1.2.34) sind A und $I(r,m,n)$ äquivalent. □

Dieser Satz gestattet uns eine wichtige Beobachtung.

2.6.16 **Korollar.** *Sei* $A \in \mathrm{M}_{m,n}(K)$. *Dann gilt* $\mathrm{Rg}(A) = \mathrm{Rg}(A^T)$.

Beweis. Sei $r = \mathrm{Rg}(A)$. Da auch die Matrix $I(r,m,n) = \sum_{i=1}^r E_{ii} \in \mathrm{M}_{m,n}(K)$ den Rang r hat, gilt $A \approx I(r,m,n)$. Es gibt also invertierbare Matrizen $S \in \mathrm{GL}_m(K)$ und $T \in \mathrm{GL}_n(K)$ mit $A = SI(r,m,n)T$. Daraus folgt

$$A^T = T^T I(r,n,m) S^T.$$

Das heißt, $A^T \approx I(r,n,m)$, denn S^T und T^T sind invertierbar (siehe Aufgabe 1.2.34). Da $I(r,n,m)$ den Rang r hat, folgt $\mathrm{Rg}(A^T) = r = \mathrm{Rg}(A)$. □

2.6.17 **(Zeilen)rang und Spaltenrang**. In den mathematischen Grundlagen haben Sie die elementaren Zeilenumformungen auf Matrizen kennengelernt und ausführlich studiert. Genauso könnte man sich aber auch mit den elementaren *Spaltenumformungen* auf Matrizen befassen. Hier kann man im Wesentlichen genauso vorgehen und sich überlegen, dass man jede Matrix mittels elementarer Spaltenumformungen auf „Spalten-Treppennormalform“ bringen kann. Mit dieser Spalten-Treppennormalform könnte man dann den „Spaltenrang“ einer Matrix definieren.

Da Spaltenumformungen in der Matrix A gerade den Zeilenumformungen in der Matrix A^T entsprechen, gibt uns Korollar 2.6.16 eine beruhigende Antwort: Der Spaltenrang einer Matrix ist der Rang. Wir müssen uns also nicht die Mühe machen, die Überlegungen aus dem Mathematischen Grundlagen für Spaltenumformungen zu wiederholen.

2.7. Lösungen der Aufgaben in Lektion 2

L2.1.8 **Lösung.** Es sei M eine Menge mit einer Totalordnung $\preceq$. Wir zeigen die Aussage mit vollständiger Induktion nach k. Für den Induktionsanfang $k = 1$ setzen wir $j = 1$ und es gilt $x_1 \preceq x_1$.

Sei nun $k > 1$ und seien $x_1, \ldots, x_k \in M$. Nach Induktionsvoraussetzung gibt es einen Index $j_1 \leq k - 1$, sodass

$$x_i \preceq x_{j_1} \tag{2.7.a}$$

für alle $i \in \{1, 2, \ldots, k - 1\}$ gilt. Falls $x_k \preceq x_{j_1}$ ist, setzen wir $j = j_1$ und sind fertig. Anderenfalls muss $x_{j_1} \preceq x_k$ sein, denn $\preceq$ ist eine Totalordnung. In diesem Fall definieren wir $j = k$. Für $i \leq k - 1$ erhalten wir

$$x_i \preceq x_{j_1} \preceq x_k$$

aus (2.7.a) durch Transitivität und $x_k \preceq x_k$ folgt aus der Reflexivität.

L2.2.6 **Lösung.** Sei V ein K-Vektorraum und sei $U \subseteq V$ eine nicht-leere Teilmenge.

Behauptung: U ist genau dann ein Unterraum, wenn

$$au + u' \in U \tag{2.7.b}$$

für alle $a \in K$ und $u, u' \in U$ gilt.

Nehmen wir zunächst an, dass U ein Unterraum ist. Dann gilt für alle $a \in K$ und $u, u' \in U$ auch $au \in U$ (wegen 2.2.4 (c)) und durch 2.2.4 (b) auch $au + u' \in U$.

Umgekehrt nehmen wir nun an, dass U die Bedingung (2.7.b) erfüllt. Wir überlegen uns, dass U den Nullvektor enthält. Da U nicht-leer ist, gibt es mindestens einen Vektor $w \in U$. Also gilt

$$0 = -w + w = (-1)w + w \in U.$$

Mit $a = 1$ folgt aus (2.7.b), dass U unter Addition abgeschlossen ist, d.h. 2.2.4 (b) ist erfüllt. Mit $u' = 0$ und a beliebig, sehen wir, dass U unter Skalarmultiplikation abgeschlossen ist, d.h., 2.2.4 (c) ist erfüllt.

L2.2.9 **Lösung.** Wir schreiben die Vektoren als Spalten in eine Matrix

$$A = \begin{pmatrix} \overline{1} & \overline{3} & \overline{4} \\ \overline{2} & \overline{2} & \overline{3} \\ \overline{6} & \overline{1} & \overline{1} \\ \overline{2} & \overline{6} & \overline{1} \end{pmatrix} \in \mathrm{M}_{4,3}(\mathbb{F}_7)$$

und berechnen dann den Rang mit dem Gauß-Verfahren.

Wir ziehen das Zweifache der ersten Zeile jeweils von der zweiten und vierten Zeile ab und wir addieren die erste zur dritten Zeile. Dann erhalten wir

$$A \rightsquigarrow \begin{pmatrix} \overline{1} & \overline{3} & \overline{4} \\ \overline{0} & \overline{3} & \overline{2} \\ \overline{0} & \overline{4} & \overline{5} \\ \overline{0} & \overline{0} & \overline{0} \end{pmatrix}$$

Addiert man die zweite Zeile zur dritten, dann gelangt man zur Matrix

$$\begin{pmatrix} \overline{1} & \overline{3} & \overline{4} \\ \overline{0} & \overline{3} & \overline{2} \\ \overline{0} & \overline{0} & \overline{0} \\ \overline{0} & \overline{0} & \overline{0} \end{pmatrix}.$$

Hier ist abzusehen, dass die Treppennormalform zwei nicht-verschwindende Zeilen besitzt, d.h., es ist $\operatorname{Rg}(A) = 2$. Die Vektoren sind also linear abhängig.

L2.2.18 **Lösung.** Sei $\varphi\colon U \to V$ eine Abbildung zwischen zwei K-Vektorräumen.

Wir nehmen an, dass für alle $u_1, u_2 \in U$ und alle $a \in K$ die Gleichung (2.2.a) gilt. Setzt man $a = 1$, dann erhält man (L1). Außerdem gilt damit auch $\varphi(0) = \varphi(0 + 0) = \varphi(0) + \varphi(0)$ und damit $\varphi(0) = 0$. Mit allgemeinem $a \in K$ und $u_2 = 0$ erhält man daraus (L2). Also ist φ linear.

Nehmen wir umgekehrt an, dass φ linear ist. Mit (L1) und (L2) erhalten wir für alle $u_1, u_2 \in U$ und $a \in K$

$$\varphi(au_1 + u_2) = \varphi(au_1) + \varphi(u_2) = a\varphi(u_1) + \varphi(u_2).$$

L2.3.3 **Lösung.** *Behauptung*: $\langle S \cup T\rangle = \langle S\rangle + \langle T\rangle$.

Wir zeigen die Gleichheit, indem wir zeigen, dass diese Unterräume sich gegenseitig enthalten.

Alle Linearkombinationen von Elementen aus S liegen auch in $\langle S \cup T\rangle$, d.h., $\langle S\rangle \subseteq \langle S \cup T\rangle$. Genauso gilt auch $\langle T\rangle \subseteq \langle S \cup T\rangle$. Da $\langle S \cup T\rangle$ ein Unterraum ist (also unter Addition abgeschlossen ist), folgt

$$\langle S\rangle + \langle T\rangle \subseteq \langle S \cup T\rangle.$$

Sei umgekehrt $v \in \langle S \cup T \rangle$. Das heißt, es gibt Elemente $v_1, \dots, v_n \in S \cup T$ und Skalare $a_1, \dots, a_n \in K$ mit

$$v = \sum_{i=1}^{n} a_i v_i.$$

Durch Umsortieren der Vektoren können wir annehmen, dass $v_1, \dots, v_k \in S$ und $v_{k+1}, \dots, v_n \in T$ für ein $0 \le k \le n$ gilt. Daraus schließen wir

$$v = \sum_{i=1}^{n} a_i v_i = \sum_{i=1}^{k} a_i v_i + \sum_{i=k+1}^{n} a_i v_i \in \langle S \rangle + \langle T \rangle;$$

also $\langle S \cup T \rangle \subseteq \langle S \rangle + \langle T \rangle$.

L2.3.10 **Lösung.** Wir zeigen, $U_1 \cap U_2 = \{0\}$. Es sei $x \in U_1 \cap U_2$. Da $x \in U_2$ ist, gilt $x_1 = 0 = x_3$. Daraus folgt

$$x_2 = -x_1 = 0 \quad \text{und} \quad x_4 = -x_3 = 0$$

und insgesamt $x = 0$.

L2.3.11 **Lösung.** Es sei V ein K-Vektorraum mit einer Basis $\mathcal{B}$. Wir nehmen an, dass $\mathcal{B} = \bigcup_{i=1}^{k} \mathcal{B}_i$ die disjunkte Vereinigung von Teilmengen $\mathcal{B}_i \subseteq \mathcal{B}$ ist und wollen zeigen, dass

$$V = \bigoplus_{i=1}^{k} \langle \mathcal{B}_i \rangle$$

gilt. Da $\mathcal{B}$ ein Erzeugendensystem ist, sehen wir mit Aufgabe 2.3.3, dass

$$\sum_{i=1}^{k} \langle \mathcal{B}_i \rangle = \langle \bigcup_{i=1}^{k} \mathcal{B}_i \rangle = \langle \mathcal{B} \rangle = V$$

ist. Wir prüfen mit 2.3.6 (iii), dass die Summe direkt ist. Es seien dazu Vektoren $u_i \in \langle \mathcal{B}_i \rangle$ mit $\sum_{i=1}^{k} u_i = 0$ gegeben. Für jedes $i \in \{1, \dots, k\}$ schreiben wir

$$u_i = \sum_{j=1}^{m_i} a_{ij} v_{ij}$$

als Linearkombination paarweise verschiedener Basisvektoren $v_{ij} \in \mathcal{B}_i$. Dann gilt

$$0 = \sum_{i=1}^{k} u_i = \sum_{i=1}^{k} \sum_{j=1}^{m_i} a_{ij} v_{ij}.$$

Da die Vereinigung der $\mathcal{B}_i$ disjunkt ist, kommt jeder Basisvektor aus $\mathcal{B}$ höchstens einmal in der Summe vor. Da $\mathcal{B}$ linear unabhängig ist, folgt nun $a_{ij} = 0$ für alle i, j und damit $u_i = 0$ für alle i. Damit ist die Summe direkt.

L2.3.13 Lösung. Gegeben sind $V = \bigoplus_{i=1}^k V_i$ und $W = \bigoplus_{i=1}^k W_i$ und eine lineare Abbildung $\varphi\colon V \to W$ mit der Eigenschaft $\varphi(V_i) \subseteq W_i$. Wir bemerken zuerst, dass $\mathrm{Ker}(\varphi|V_i) \subseteq \mathrm{Ker}(\varphi)$ gilt. Insbesondere liegt die Summe $\sum_{i=1}^k \mathrm{Ker}(\varphi|_{V_i})$ in $\mathrm{Ker}(\varphi)$.

Wir behaupten nun , dass

$$\mathrm{Ker}(\varphi) = \bigoplus_{i=1}^k \mathrm{Ker}(\varphi|_{V_i})$$

gilt.

Sei dazu $v = \sum_{i=1}^k v_i \in \mathrm{Ker}(\varphi)$ mit $v_i \in V_i$. Es gilt

$$0 = \varphi(v) = \sum_{i=1}^k \underbrace{\varphi(v_i)}_{\in W_i}.$$

Da W die direkte Summe der Unterräume W_i ist, gilt somit $\varphi(v_i) = 0$ für alle $i \in \{1, \dots, k\}$. Damit ist $\mathrm{Ker}(\varphi) = \sum_{i=1}^k \mathrm{Ker}(\varphi|_{V_i})$.

Warum ist die Summe direkt? Da die Summe der Unterräume V_i direkt ist, gilt mit Satz 2.3.6 für alle $i \in \{1, 2, \dots, k\}$

$$\mathrm{Ker}(\varphi|_{V_i}) \cap \sum_{i' \neq i} \mathrm{Ker}(\varphi|_{V_{i'}}) \subseteq V_i \cap \sum_{i' \neq i} V_{i'} = \{0\}$$

und aus Satz 2.3.6 folgt wieder, dass auch die Summe der Kerne direkt ist.

L2.3.18 Lösung. Es seien V_1, V_2 zwei K-Vektorräume und es sei $\varphi\colon V_1 \to V_2$ eine surjektive K-lineare Abbildung. Wir wählen ein Komplement W zu $\mathrm{Ker}(\varphi)$ in V_1, d.h.,

$$V_1 = W \oplus \mathrm{Ker}(\varphi).$$

Wir behaupten, dass $\varphi|_W$ – die Einschränkung von φ auf W – ein Isomorphismus von W und V_2 ist. Wegen $\mathrm{Ker}(\varphi) \cap W = \{0\}$ ist $\varphi|_W$ injektiv. Es sei $v_2 \in V_2$ beliebig. Da φ surjektiv ist, finden wir ein $v_1 \in V_1$ mit $\varphi(v_1) = v_2$. Wir schreiben nun

$$v_1 = w + u$$

mit $w \in W$ und $u \in \mathrm{Ker}(\varphi)$. Dann erhalten wir

$$v_2 = \varphi(v_1) = \varphi(w) + \underbrace{\varphi(u)}_{=0} = \varphi(w);$$

d.h., $\varphi|_W$ ist surjektiv.

Insgesamt ist $\varphi|_W$ also bijektiv und K-linear und damit ein Isomorphismus von Vektorräumen. Nach 8.1.8 in [MG] ist auch die Umkehrabbildung $\psi := (\varphi|_W)^{-1}$ ein Isomorphismus. Fassen wir ψ als Abbildung nach V_1 auf, so gilt

$$\varphi \circ \psi = \varphi \circ (\varphi|_W)^{-1} = \varphi|_W \circ (\varphi|_W)^{-1} = \mathrm{id}_{V_2},$$

weil $(\varphi|_W)^{-1}$ nur Werte in W annimmt. Damit ist ψ die gesuchte rechtsinverse Abbildung zu φ.

L2.4.1 **Lösung.** Gegeben sind die beiden Geraden

$$\begin{aligned} G_1 &= \begin{pmatrix} -2 \\ 2 \end{pmatrix} + U = \left\{ \begin{pmatrix} -2+s \\ 2+s \end{pmatrix} \mid s \in \mathbb{R} \right\} \\ G_2 &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} + U = \left\{ \begin{pmatrix} 1+t \\ t \end{pmatrix} \mid t \in \mathbb{R} \right\}. \end{aligned}$$

Addieren wir zwei beliebige Punkte aus G_1 und G_2 erhalten wir

$$\begin{pmatrix} -2+s \\ 2+s \end{pmatrix} + \begin{pmatrix} 1+t \\ t \end{pmatrix} = \begin{pmatrix} -1+(s+t) \\ 2+(s+t) \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix} + (s+t)\begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Da $s+t$ jeden Wert aus $\mathbb{R}$ annimmt erhalten wir

$$G_1 + G_2 = \begin{pmatrix} -1 \\ 2 \end{pmatrix} + U.$$

L2.4.6 **Lösung.** Es gilt $A \neq B$, denn $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \\ -4 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 7 \end{pmatrix} \notin U$.

Es gilt $A = C$, denn $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} - \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} \in U$. Folglich gilt auch $C \neq B$.

L2.4.10 **Lösung.** Man erhält:

$$2 \cdot (\begin{pmatrix} 3 \\ -5 \end{pmatrix} + U) + (\begin{pmatrix} -1 \\ 1 \end{pmatrix} + U) = (\begin{pmatrix} 6 \\ -10 \end{pmatrix} + U) + (\begin{pmatrix} -1 \\ 1 \end{pmatrix} + U) = \begin{pmatrix} 5 \\ 9 \end{pmatrix} + U.$$

L2.4.11 Lösung. Für alle $a, b \in K$ und $v, w \in V$ gilt

$$\begin{aligned}(a+b)(v+U) &= (a+b)v + U = av + bv + U \\ &= (av+U) + (bv+U) = a(v+U) + b(v+U)\end{aligned}$$

und analog gilt

$$\begin{aligned}a\big((v+U)+(w+U)\big) &= a(v+w+U) = a(v+w)+U \\ &= av + aw + U = (av+U)+(aw+U) \\ &= a(v+U) + a(w+U).\end{aligned}$$

L2.4.16 Lösung. Die Vorschrift

$$\begin{pmatrix} x \\ y \end{pmatrix} + U \mapsto 3x - 2y$$

definiert *keine* Abbildung $\mathcal{G} \to \mathbb{R}$, denn das Ergebnis hängt wirklich vom Stützvektor v und nicht nur vom affinen Unterraum $v + U$ ab. Zum Beispiel gilt $\begin{pmatrix} 0 \\ 0 \end{pmatrix} + U = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + U$, aber

$$3 \cdot 0 - 2 \cdot 0 = 0 \neq 1 = 3 \cdot 1 - 2 \cdot 1.$$

L2.4.20 Lösung. Es seien V ein K-Vektorraum und $U \subseteq V$ ein Unterraum. Gegeben sind zwei linear unabhängige Vektoren $v_1, v_2 \in V$.

Wir beweisen die Aussage über die Kontraposition. Angenommen, $v_1 + U$ und $v_2 + U$ sind linear abhängig. Dann gibt es $\lambda_1, \lambda_2 \in K$ (nicht beide Null) dergestalt, dass $\lambda_1 v_1 + U + \lambda_2 v_2 + U = 0 + U$ gilt. Das heißt, $\lambda_1 v_1 + \lambda_2 v_2 \in U$. Da v_1, v_2 linear unabhängig in V sind, folgt $\lambda_1 v_1 + \lambda_2 v_2 \neq 0$ und damit $\langle v_1, v_2 \rangle \cap U \neq \{0\}$.

Umgekehrt nehmen wir an, dass $\langle v_1, v_2 \rangle \cap U \neq \{0\}$ ist. Es gilt also $u = \lambda_1 v_1 + \lambda_2 v_2$ für ein $u \in U$ mit $u \neq 0$. Insbesondere ist mindestens ein λ_i ungleich 0. Damit schließen wir

$$\lambda_1 v_1 + U + \lambda_2 v_2 + U = u + U = 0 + U,$$

also sind $v_1 + U$ und $v_2 + U$ linear abhängig in V/U.

L2.5.2 Lösung. Wir zeigen, dass $\mathrm{Hom}_K(V, W)$ ein Unterraum von $\mathrm{Abb}(V, W)$ ist. Wir bemerken zunächst, dass die Addition und Skalarmultiplikation auf beiden Räumen jeweils punktweise definiert wurden und es sich damit um dieselben Verknüpfungen handelt.

Nun verwenden wir das Kriterium aus Aufgabe 2.2.6. Sind $\varphi, \psi \in \mathrm{Hom}_K(V, W)$ und $b \in K$, dann prüfen wir nach, dass auch $\varphi + b\psi$ linear ist. Zuerst stellen wir fest, dass für alle $v_1, v_2 \in V$ gilt

$$\begin{aligned}(\varphi + b\psi)(v_1 + v_2) &= \varphi(v_1 + v_2) + b\psi(v_1 + v_2) \\ &= \varphi(v_1) + \varphi(v_2) + b\psi(v_1) + b\psi(v_2) && (\varphi, \psi \text{ linear}) \\ &= (\varphi + b\psi)(v_1) + (\varphi + b\psi)(v_2).\end{aligned}$$

Für alle $v \in V$ und $a \in K$ gilt außerdem

$$(\varphi + b\psi)(av) = \varphi(av) + b\psi(av) = a\varphi(v) + ab\psi(v) = a(\varphi + b\psi)(v).$$

Also ist $\varphi+b\psi \in \mathrm{Hom}_K(V, W)$ und $\mathrm{Hom}_K(V, W)$ ist ein Unterraum von $\mathrm{Abb}(V, W)$.

L2.5.7 Lösung. In der Praxis ist es ganz einfach, auch wenn es im Beweis von Lemma 2.5.6 kompliziert scheint, eine entsprechende Linearform zu finden. Der zweite Eintrag des Vektors $v = \begin{pmatrix} 0 \\ \overline{3} \\ \overline{2} \end{pmatrix} \in (\mathbb{F}_7^3)^*$ ist $\overline{3} \neq 0$. Daher hat die Projektion auf die zweite Koordinate, also die Linearform $\alpha\colon \mathbb{F}_7^3 \to \mathbb{F}_7$ mit

$$\alpha \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_2,$$

die Eigenschaft $\alpha(v) = \overline{3} \neq 0$.

L2.5.15 Lösung. Zur Basis

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ -1 \\ 4 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \in \mathbb{R}^3$$

wollen wir die Dualbasis v_1^*, v_2^*, v_3^* durch die Zeilenvektoren $E(v_1^*), E(v_2^*), E(v_3^*)$ beschreiben. Wir verwenden das Verfahren aus 2.5.13. Dazu bilden wir die Matrix

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

mit den Spalten v_1, v_2, v_3 und berechnen die inverse Matrix A^{-1} mit dem bekannten Verfahren aus [MG, Abschnitt 4.3, 4.5.7]. Wir schreiben

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 \\ 1 & 4 & 1 & 0 & 0 & 1 \end{array}\right)$$

und ziehen die erste von der dritten Zeile ab:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 4 & 1 & -1 & 0 & 1 \end{array}\right).$$

Dann addieren wir 4 mal die zweite zur letzten Zeile und erhalten

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 4 & 1 \end{array}\right).$$

Schließlich multiplizieren wir die zweite Zeile mit -1 und haben die Inverse bestimmt:

$$A^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ -1 & 4 & 1 \end{pmatrix}.$$

Die Bilder der Dualbasis $E(v_1^*), E(v_2^*), E(v_3^*)$ sind nun genau die Zeilen der Matrix A^{-1}; also

$$E(v_1^*) = (1,0,0), \quad E(v_2^*) = (0,-1,0), \quad E(v_3^*) = (-1,4,1).$$

L2.6.5 **Lösung.** Für $z, w \in \mathbb{C}^\times$ schreiben wir $z \smile w$, wenn es eine reelle Zahl $r \in \mathbb{R}^\times$ mit $rz = w$ gibt. Dies definiert eine Äquivalenzrelation:

reflexiv: Sei $z \in \mathbb{C}^\times$. Für die reelle Zahl $r = 1$ gilt $1 \cdot z = z$ und damit $z \smile z$.

symmetrisch: Seien $z, w \in \mathbb{C}^\times$ mit $z \smile w$. Es gibt also eine Zahl $r \in \mathbb{R}^\times$ mit $rz = w$. Da $r \neq 0$ gilt, besitzt r ein multiplikatives Inverses r^{-1}. Es gilt $z = r^{-1}w$ und damit $w \smile z$.

transitiv: Seien $z, w, s \in \mathbb{C}^\times$ mit $z \smile w$ und $w \smile s$ gegeben. Es gibt also reelle Zahlen $r_1, r_2 \in \mathbb{R}^\times$ mit $r_1 z = w$ und $r_2 w = s$. Es gilt also auch $r_2 r_1 z = r_2 w = s$ Da $\mathbb{R}^\times$ eine Gruppe bzgl. Multiplikation ist (siehe 1.1.10), gilt $r = r_2 r_1 \in \mathbb{R}^\times$ und damit $z \smile s$.

L2.6.9 **Lösung.** Sei V ein K-Vektorraum und sei $U \subseteq V$ ein Unterraum. Für $x, y \in V$ schreiben wir $x \sim_U y$, wenn $x - y$ in U liegt. Wir zeigen nun, dass dies eine Äquivalenzrelation ist.

reflexiv: Der Nullvektor 0 liegt in U. Für alle $x \in V$ gilt $0 = x - x$ und somit $x \sim_U x$.

symmetrisch: Seien $x, y \in V$ mit $x \sim_U y$ gegeben. Es gilt damit $x - y = u \in U$. Da U ein Unterraum ist, liegt auch $-u \in U$. Das heißt, $y - x = -u \in U$ und damit gilt $y \sim_U x$.

transitiv: Seien $x, y, z \in V$ mit $x \sim_U y$ und $y \sim_U z$ gegeben. Es gilt also $x-y = u \in U$ und $y - z = u' \in U$. Dann gilt auch

$$x - z = x - y + y - z = u + u' \in U,$$

denn U ist ein Unterraum und damit unter Addition abgeschlossen.

Was sind die Äquivalenzklassen dieser Äquivalenzrelation? Es gilt

$$[x]_{\sim_U} = \{x + u \mid u \in U\} = x + U.$$

Die Äquivalenzklasse von x ist also genau der affine Unterraum zu U durch den Punkt x; siehe 2.4.2.

Steffen Kionke

Lineare Algebra

Lektion 3:
Bilinearformen und Hermite'sche Formen

Fakultät für
Mathematik und Informatik

FernUniversität in Hagen

Studierhinweise zur dritten Lektion

In dieser umfangreichen dritten Lektion wenden wir uns den Bilinearformen und den (eng verwandten) Hermite'schen Formen zu. Eine Bilinearform $\beta \colon U \times U \to K$ ist eine Abbildung vom kartesischen Produkt eines Vektorraumes in den Körper K. Es handelt sich also um eine Abbildung, in die man zwei Argumente einsetzt. Dabei soll die Abbildung „linear in jeder Komponente" sein, d.h., jede Komponente soll mit Addition und Skalarmultiplikation verträglich sein.

Bilinearformen sind ein abstraktes Konzept, das aber viele Anwendungen in der Mathematik und darüber hinaus hat. Beispielsweise benötigt man Bilinearformen in der Differentialgeometrie um (pseudo-)Riemann'sche Metriken auf Mannigfaltigkeiten zu definieren. Spezielle Beispiele sind Lorentzmetriken, die eine wesentliche Zutat der allgemeinen Relativitätstheorie sind. Aber auch in der Informatik trifft man Bilinearformen (unter anderem in der Theorie der fehlerkorrigierenden Codes). Auf Anwendungen können wir hier leider nicht eingehen, weil dies den Umfang dieses Moduls sprengen würde. Unser Ziel ist es, solide Grundlagen zu legen, sodass Sie keine Angst zu haben brauchen, wenn Sie wieder auf Bilinearformen treffen.

Das Hauptziel in diesem Abschnitt ist es, Bilinearformen und Hermite'sche Formen zu klassifizieren. Dazu führen wir mit der *Kongruenz* von Formen eine Äquivalenzrelation ein, sodass kongruente Formen dieselben Eigenschaften haben. Im Allgemeinen ist die Klassifikation ein schwieriges Problem, aber in speziellen Fällen (alternierende Bilinearformen, reelle symmetrische Bilinearformen, Hermite'sche Formen) können wir wirklich alle Kongruenzklassen von Formen vollständig beschreiben. Wir verstehen in diesen Fällen also genau, welche Formen es gibt.

Ich halte die Theorie der Bilinearformen für wichtig: Sie ist nicht nur nützlich in Anwendungen, sondern auch wirklich schön. Es lohnt sich, Zeit in das Studium dieser Lektion zu investieren, denn man lernt vieles über mathematisches Denken. Auf einige Ergebnisse aus dieser Lektion werden wir in Kapitel 7 übrigens wieder zurückgreifen. Ich hoffe deshalb, dass Sie dranbleiben, auch wenn diese Lektion anspruchsvoll ist.

Lernziele

Nach dem Durcharbeiten der Lektion sollten Sie

→ wissen, was eine *Bilinearform* ist und mit Matrixdarstellungen rechnen können,

→ mit dem Begriff der *Kongruenz* von Bilinearformen umgehen können,

→ wissen, was *reguläre* Unterräume sind und den *Zerlegungssatz* kennen,

→ den *Diagonalisierungssatz* für symmetrische Bilinearformen kennen und anwenden können,

→ den *Trägheitssatz* und die *Klassifikation* der reellen symmetrischen Bilinearformen kennen,

→ die *Normalform* einer alternierenden Bilinearform berechnen können,

→ mit *Hermite'schen Formen* und ihren Matrixdarstellungen umgehen können,

→ die *Klassifikation* der Hermite'schen Formen kennen und anwenden können.

Literaturhinweise

Bilinearformen und Hermite'sche Formen finden Sie wahrscheinlich in allen Lehrbüchern zur Linearen Algebra. In vielen Texten finde ich die Diskussion zu oberflächlich. Teilweise beschränken sich die Lehrbücher auf reelle symmetrische Bilinearformen um direkt zu den Skalarprodukten und Spektralsätzen zu kommen, die wir erst in Lektion 7 untersuchen.

→ [Bär]: 7.1 (Bilinearformen).

→ [Beu]: 10.2 (Bilinearformen).

→ [Bo]: 7.1 und 7.3 (behandelt reelle Bilinearformen und Hermite'sche Formen gleichzeitig).

→ [Fi]: 6.4 (Bilinearformen).

→ [Gö]: 4.8–4.10 (Bilinearformen).

→ [KaSt]: 10.1 (symmetrische Bilinearformen), 10.2 (Hermite'sche Formen).

Achtung: Bei den Hermite'schen Formen verwenden [Bo, Fi, KaSt] und einige andere Bücher eine andere Konvention als wir!

Fahrplan durch die Lektion

Abschnitt 3.1 behandelt die grundlegende Theorie der *Bilinearformen*. Bitte lesen Sie diesen Abschnitt sehr genau. ← 3.1

Der erste Teil beginnt mit der Definition (3.1.1) und einigen Beispielen (3.1.4–3.1.8). Im zweiten Teil besprechen wir die *Matrixdarstellung* von Bilinearformen (3.1.10). Wir lernen, dass Bilinearformen durch eine Matrixdarstellung bestimmt sind (3.1.14) und berechnen die Dimension des Vektorraumes aller Bilinearformen (3.1.16). Entscheidend für alles Weitere ist die *Transformationsformel* (3.1.17).

Im dritten Teil definieren wir *Kongruenz* von Bilinearformen (3.1.22) und beschreiben Kongruenz mithilfe der Matrixdarstellungen (3.1.23 und 3.1.19). Damit formulieren wir die zentralen Fragestellungen dieser Lektion (3.1.21, 3.1.25). Es ist wichtig, dass Sie diesen Abschnitt genau verstehen, damit Sie wissen, was wir in dieser Lektion erreichen möchten.

Der vierte Teil besteht aus einigen formalen Überlegungen über das feste Einsetzen eines Vektors in eine Komponente einer Bilinearform. Die Beobachtungen erlauben es uns, einen anderen Blick auf Bilinearformen zu werfen: Eine Bilinearform β auf U kann man als lineare Abbildung $\beta^{(1)}\colon U \to U^*$ verstehen (3.1.32). Damit haben wir Zugriff auf unser Wissen über lineare Abbildungen! In diesem technischen Abschnitt passiert nichts Tiefgründiges. Lassen Sie sich von den abstrakten Rechnungen nicht verwirren.

Im letzten Teil definieren wir *reguläre* und *ausgeartete* Bilinearformen (3.1.37) und geben mehrere alternative Beschreibungen (3.1.38). Wir beweisen dann den wichtigen *Zerlegungssatz* (3.1.43), der unser wichtigstes Werkzeug in dieser Lektion sein wird. Wir definieren schließlich den Rang (3.1.44) und den Nullraum (3.1.45) und beweisen eine Dimensionsformel (3.1.47).

In Abschnitt 3.2 studieren wir die *symmetrischen* Bilinearformen (3.2.1). Das sind Bilinearformen, die eine symmetrische Matrixdarstellung besitzen (3.2.4, 3.2.6). ← 3.2

Im zweiten Teil beweisen wir mit dem *Diagonalisierungssatz* (3.2.10) ein zentrales Ergebnis über symmetrische Bilinearformen. Dabei betrachten wir nur Körper mit der Eigenschaft $2 \neq 0$ (3.2.8), weil wir die Polarisationsformel (3.2.9) benötigen. Wir besprechen kurz, wie man eine diagonale Matrixdarstellung berechnen kann (3.2.12). Der Diagonalisierungssatz ist im Allgemeinen nur ein erster Schritt hin zur Klassifikation von Bilinearformen. Über dem Körper der komplexen Zahlen lässt sich daraus direkt eine Klassifikation ableiten (3.2.16, 3.2.18).

Im dritten Teil studieren wir dann symmetrische Bilinearformen über dem Körper der reellen Zahlen. Der *Trägheitssatz* (3.2.24) liefert die entscheidende Beobachtung zur Klassifikation: Die Anzahl der positiven und negativen Diagonaleinträge hängt nicht von der gewählten diagonalen Matrixdarstellung ab. Damit definieren wir den *Typ* reeller symmetrischer Bilinearformen (3.2.25) und leiten daraus die Klassifikation ab (3.2.33). Dazu verwenden wir Matrixdarstellungen in *Normalform* (3.2.29).

3.3 → In Abschnitt 3.3 klassifizieren wir die *alternierenden* Bilinearformen. Wir beginnen mit der Definition alternierender Bilinearformen (3.3.1) und beschreiben diese anhand einer Matrixdarstellung (3.3.5, 3.3.7). Wir leiten eine Matrixdarstellung in *Normalform* her (3.3.9) und besprechen, wie man eine solche Darstellung berechnen kann (3.3.11). Damit können wir zeigen, dass alternierende Bilinearformen durch den Rang klassifiziert sind (3.3.12).

3.4 → In Abschnitt 3.4 wenden wir uns den *Hermite'schen Formen* zu (3.4.1). Diese Formen studieren wir nur über den komplexen Zahlen $\mathbb{C}$. Es handelt sich dabei um Abbildungen mit zwei Komponenten, die linear in der zweiten aber *konjugiert-linear* in der ersten Komponente sind (3.4.2). Die ganze Theorie ist sehr ähnlich zur Theorie der reellen symmetrischen Bilinearformen und lässt sich sogar darauf zurückführen.

Nachdem wir im ersten Teil die Grundbegriffe geklärt haben, diskutieren wir im zweiten Teil die *Matrixdarstellung* von Hermite'schen Formen (3.4.20). Dazu benötigen wir einige kleine Überlegungen zu komplexen Matrizen (3.4.11–3.4.18). Wir zeigen, dass eine Matrixdarstellung die Hermite'sche Form eindeutig bestimmt (3.4.23) und leiten eine Transformationsformel her (3.4.25). Schließlich definieren wir *Hermite'sche Kongruenz* von Matrizen (3.4.26) und die *Kongruenz* von Hermite'schen Formen (3.4.31). Damit beschreiben wir die wesentlichen Fragestellungen.

Im dritten Teil lernen wir, wie man aus einer Hermite'schen Form eine reelle symmetrische Bilinearform bekommt (3.4.39). Zuerst müssen wir uns klarmachen, dass jeder $\mathbb{C}$-Vektorraum auch ein $\mathbb{R}$-Vektorraum ist (3.4.34–3.4.38). Damit können wir einen *Zerlegungssatz* für Hermite'sche Formen (3.4.42) vom Zerlegungssatz für reelle symmetrische Bilinearformen ableiten.

Im vierten Teil klassifizieren wir die Hermite'schen Formen. Das Fundament bildet wieder ein Trägheitssatz 3.4.43 und der Begriff des Typs (3.4.44). Damit sind wir in der Lage eine Matrixdarstellung in Normalform anzugeben (3.4.45, 3.4.46) und die Klassifikation abzuschließen (3.4.50).

3.1. Bilinearformen

Aus den Mathematischen Grundlagen kennen Sie bereits die linearen Abbildungen zwischen Vektorräumen. Wir erinnern uns, dass eine Abbildung $\varphi\colon U \to V$ genau dann linear ist, wenn $\varphi(au + u') = a\varphi(u) + \varphi(u')$ für alle $a \in K$ und $u, u' \in U$ gilt; siehe Aufgabe 2.2.18. Eine lineare Abbildung $U \to K$ nennen wir Linearform.

In diesem Abschnitt werden wir ein wichtiges, verwandtes Konzept einführen: *Bilinearformen*. Dabei handelt es sich um Abbildungen $U \times U \to K$ – also um Abbildungen, in die man zwei Argumente einsetzt – wobei jede Komponente der Abbildung linear ist. Hier werden wir uns Beispiele anschauen und die Grundlagen dieser Theorie erarbeiten. Insbesondere werden wir lernen, wie man mit Bilinearformen in Koordinaten rechnen kann.

I. Definition und Beispiele

3.1.1 **Definition** Sei K ein Körper und sei U ein K-Vektorraum. Eine Abbildung

$$\beta\colon U \times U \to K$$

nennt man *Bilinearform* über K, wenn die Abbildung linear in beiden Komponenten ist. Das heißt, die folgenden Bedingungen sind erfüllt:

(a) Für alle $u, u', v \in U$ und $a \in K$ gilt

$$\beta(au + u', v) = a\beta(u, v) + \beta(u', v).$$

(b) Für alle $u, v, v' \in U$ und $a \in K$ gilt

$$\beta(u, av + v') = a\beta(u, v) + \beta(u, v').$$

Die Menge aller Bilinearformen auf U bezeichnen wir mit $\mathrm{Bil}_K(U)$ oder kürzer $\mathrm{Bil}(U)$, wenn der Körper K fest ist.

3.1.2 **Aufgabe**. Sei $\beta\colon U \times U \to K$ eine Bilinearform. Zeigen Sie, dass $\beta(u, 0) = 0 = \beta(0, u)$ für alle $u \in U$ gilt. **L**

3.1.3 **Anmerkung zur Definition**. Sei $\beta\colon U \times U \to K$ eine Bilinearform. Aus der Linearität in beiden Komponenten erhalten wir insbesondere die Rechenregeln

- $\beta(u + u', v) = \beta(u, v) + \beta(u', v)$ aus (a) mit $a = 1$,

- $\beta(u, v+v') = \beta(u,v) + \beta(u,v')$ aus (b) mit $a = 1$.
- $\beta(au, v) = a\beta(u,v) = \beta(u, av)$ aus (a) mit $u' = 0$ und (b) mit $v' = 0$ mit Aufgabe 3.1.2,

Umgekehrt kann man aus diesen drei Eigenschaften wieder die Eigenschaften (a) und (b) herleiten (genau wie in Aufgabe 2.2.18). In manchen Büchern werden daher diese drei Bedingungen zur Definition von Bilinearformen verwendet. Sind $u = \sum_{i=1}^n a_i u_i$ und $v = \sum_{j=1}^n b_j v_j$ als Linearkombinationen gegeben, dann kann man den Ausdruck $\beta(u,v)$ mit diesen Regeln berechnen und erhält

$$\beta(u,v) = \beta\Big(\sum_{i=1}^n a_i u_i, \sum_{j=1}^n b_j v_j\Big) = \sum_{i=1}^n \sum_{j=1}^n a_i b_j \beta(u_i, v_j).$$

Diesen Ausdruck kann man sich wie das Ausmultiplizieren von Klammerausdrücken merken: jeder mit jedem.

3.1.4 **Beispiel**. (a) Ist U ein K-Vektorraum, dann ist die Nullabbildung $0\colon U \times U \to K$ mit $0(u,v) = 0$ für alle $u, v \in U$ eine Bilinearform.

(b) Die Abbildung $\beta\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ mit

$$\beta\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = 2x_1y_1 - x_1y_2 + x_2y_2$$

ist eine Bilinearform. Um das einzusehen, überprüfen wir die zwei definierenden Eigenschaften. Für alle $x, x', y \in \mathbb{R}^2$ und $a \in \mathbb{R}$ gilt

$$\begin{aligned} \beta(ax + x', y) &= 2(ax_1 + x_1')y_1 - (ax_1 + x_1')y_2 + (ax_2 + x_2')y_2 \\ &= 2ax_1y_1 + 2x_1'y_1 - ax_1y_2 - x_1'y_2 + ax_2y_2 + x_2'y_2 \\ &= a(2x_1y_1 - x_1y_2 + x_2y_2) + (2x_1'y_1 - x_1'y_2 + x_2'y_2) \\ &= a\beta(x,y) + \beta(x', y). \end{aligned}$$

Für alle $x, y, y' \in \mathbb{R}^2$ und $a \in \mathbb{R}$ gilt

$$\begin{aligned} \beta(x, ay + y') &= 2x_1(ay_1 + y_1') - x_1(ay_2 + y_2') + x_2(ay_2 + y_2') \\ &= 2ax_1y_1 + 2x_1y_1' - ax_1y_2 - x_1y_2' + ax_2y_2 + x_2y_2' \\ &= a(2x_1y_1 - x_1y_2 + x_2y_2) + (2x_1y_1' - x_1y_2' + x_2y_2') \\ &= a\beta(x,y) + \beta(x, y'). \end{aligned}$$

Also ist β eine Bilinearform.

(c) Wir betrachten den Vektorraum $U = \mathrm{M}_{n,n}(K)$ der $(n \times n)$-Matrizen und definieren

$$\beta(A, B) = \operatorname{Spur}(AB);$$

die Spur haben wir in Aufgabe 1.2.29 definiert. Dann ist β eine Bilinearform auf $\mathrm{M}_{n,n}(K)$. Für alle $a \in K,\ A, A', B \in \mathrm{M}_{n,n}(K)$ gilt

$$\begin{aligned}\beta(aA + A', B) &= \operatorname{Spur}((aA + A')B) = \operatorname{Spur}(aAB + A'B) \\ &= a \operatorname{Spur}(AB) + \operatorname{Spur}(A'B) && (1.2.29) \\ &= a\beta(A, B) + \beta(A', B).\end{aligned}$$

Genauso gilt für alle $A, B, B' \in \mathrm{M}_{n,n}(K)$

$$\begin{aligned}\beta(A, aB + B') &= \operatorname{Spur}(A(aB + B')) = \operatorname{Spur}(aAB + AB') \\ &= a \operatorname{Spur}(AB) + \operatorname{Spur}(AB') && (1.2.29) \\ &= a\beta(A, B) + \beta(A, B').\end{aligned}$$

Damit ist β eine Bilinearform.

3.1.5 **Beispiel**. Wir betrachten den $\mathbb{R}$-Vektorraum $\mathbb{R}[X]$ aller Polynome mit reellen Koeffizienten. Für Polynome $f, g \in \mathbb{R}[X]$ definieren wir

$$\beta(f, g) = \int_0^1 f(x)g(x)\mathrm{d}x$$

über das Riemann-Integral. Dann ist $\beta\colon \mathbb{R}[X] \times \mathbb{R}[X] \to \mathbb{R}$ eine Bilinearform. In der Tat, für alle $f_1, f_2, g \in \mathbb{R}[X]$ und $a \in \mathbb{R}$ gilt

$$\begin{aligned}\beta(af_1 + f_2, g) &= \int_0^1 (af_1 + f_2)(x)g(x)\ \mathrm{d}x \\ &= \int_0^1 af_1(x)g(x) + f_2(x)g(x)\ \mathrm{d}x \\ &= a \int_0^1 f_1(x)g(x)\ \mathrm{d}x + \int_0^1 f_2(x)g(x)\ \mathrm{d}x && ([\text{MG}, 20.1.16]) \\ &= a\beta(f_1, g) + \beta(f_2, g).\end{aligned}$$

Da für alle $f, g \in \mathbb{R}[X]$ die Gleichung $\beta(f, g) = \beta(g, f)$ gilt, ist β auch in der zweiten Komponente linear. Mit derselben Formel kann man natürlich auch auf anderen Funktionenräumen eine Bilinearform definieren, sofern das Integral definiert ist. Zum Beispiel könnte man auch mit dem Vektorraum $C([0, 1])$ aller stetigen Funktionen auf dem Interval $[0, 1]$ arbeiten.

3.1.6 **Beispiel**. Es sei U ein K-Vektorraum. Für zwei Linearformen $\alpha_1, \alpha_2 \in U^*$ definieren wir eine Abbildung $\alpha_1 \otimes \alpha_2 \colon U \times U \to K$ durch[1]

$$(\alpha_1 \otimes \alpha_2)(u, v) \coloneqq \alpha_1(u)\alpha_2(v)$$

für alle $u, v \in U$. Dann ist $\alpha_1 \otimes \alpha_2$ eine Bilinearform. Für alle $u, u', v \in U$ und $a \in K$ gilt

$$\begin{aligned}(\alpha_1 \otimes \alpha_2)(au + u', v) &= \alpha_1(au + u')\alpha_2(v) \\ &= (a\alpha_1(u) + \alpha_1(u'))\alpha_2(v) && (\alpha_1 \text{ linear}) \\ &= a\alpha_1(u)\alpha_2(v) + \alpha_1(u')\alpha_2(v) \\ &= a(\alpha_1 \otimes \alpha_2)(u, v) + (\alpha_1 \otimes \alpha_2)(u', v).\end{aligned}$$

Analog zeigt man $(\alpha_1 \otimes \alpha_2)(u, av + v') = a(\alpha_1 \otimes \alpha_2)(u, v) + (\alpha_1 \otimes \alpha_2)(u, v')$ für alle $u, v, v' \in U$ und $a \in K$ (Übungsaufgabe!). **L**

3.1.7 **Beispiel**. Es sei $U = K^n$ und es sei $A \in \mathrm{M}_{n,n}(K)$ eine $(n \times n)$-Matrix. Dann können wir durch A eine Bilinearform auf K^n definieren. Wir setzen

$$\beta(u, v) = u^T Av$$

für alle $u, v \in K^n$. Hier bezeichnet u^T den transponierten Vektor. Warum ist dieser Ausdruck sinnvoll? Hier multipliziert man eine $(1 \times n)$-Matrix (Zeilenvektor) mit einer $(n \times n)$-Matrix und einer $(n \times 1)$-Matrix (Spaltenvektor); das Ergebnis ist also ein Element aus K (umständlich ausgedrückt: eine (1×1)-Matrix).

Wir prüfen, dass dies eine Bilinearform ist. Für alle $u, u', v \in K^n$ und $a \in K$ gilt

$$\begin{aligned}\beta(au + u', v) &= (au + u')^T Av \\ &= (au^T + (u')^T)Av && (\text{mit } 1.2.32) \\ &= au^T Av + (u')^T Av = a\beta(u, v) + \beta(u', v).\end{aligned}$$

Für alle $u, v, v' \in K^n$ und $a \in K$ gilt analog

$$\begin{aligned}\beta(u', av + v') &= u^T A(av + v') \\ &= u^T A(av) + u^T Av' && (\text{Distributivgesetz}) \\ &= au^T Av + u^T Av' = a\beta(u, v) + \beta(u, v').\end{aligned}$$

Wir werden bald sehen, dass auf dem K^n alle Bilinearformen von dieser Gestalt sind.

[1] Beim Lesen kann man Symbol $\otimes$ als „Tensorprodukt“ oder kurz „Tensor“ aussprechen.

3.1.8 **Bemerkung: Einschränkung von Bilinearformen**. Sei U ein K-Vektorrraum und sei $\beta \in \mathrm{Bil}_K(U)$ eine Bilinearform. Ist $W \subseteq U$ ein Unterraum, dann können wir die Bilinearform β auf W *einschränken* und erhalten eine Bilinearform auf W. Das heißt, wir definieren $\beta|_W : W \times W \to K$ durch

$$\beta|_W(w_1, w_2) = \beta(w_1, w_2)$$

für alle $w_1, w_2 \in W$. Da W in U enthalten ist, bleiben die definierenden Bedingungen aus 3.1.1 erhalten.

3.1.9 **Satz.** *Sei K ein Körper und sei U ein K-Vektorraum. Dann ist $\mathrm{Bil}_K(U)$ ein Unterraum des Vektorraumes $\mathrm{Abb}(U \times U, K)$ aller Abbildungen von $U \times U$ nach K.*

Beweis. Die Addition und Skalarmultiplikation auf $\mathrm{Abb}(U \times U, K)$ wurden punktweise definiert; siehe 2.2.3. Weil die Nullabbildung in $\mathrm{Bil}_K(U)$ liegt, müssen wir nur zeigen, dass $\mathrm{Bil}_K(U)$ unter der Addition und Skalarmultiplikation abgeschlossen ist.

Es seien $\beta, \gamma \in \mathrm{Bil}_K(U)$. Wir zeigen $\beta + \gamma \in \mathrm{Bil}_K(U)$. Für alle $u, u', v \in U$ und $a \in K$ gilt

$$\begin{aligned}(\beta+\gamma)(au+u', v) &= \beta(au+u', v) + \gamma(au+u', v)\\ &= a\beta(u,v) + \beta(u',v) + a\gamma(u,v) + \gamma(u',v)\\ &= a(\beta+\gamma)(u,v) + (\beta+\gamma)(u',v).\end{aligned}$$

Genauso gilt für alle $u, v, v' \in U$ und $a \in K$ auch

$$\begin{aligned}(\beta+\gamma)(u, av+v') &= \beta(u, av+v') + \gamma(u, av+v')\\ &= a\beta(u,v) + \beta(u,v') + a\gamma(u,v) + \gamma(u,v')\\ &= a(\beta+\gamma)(u,v) + (\beta+\gamma)(u,v').\end{aligned}$$

Insgesamt ist $\beta + \gamma$ wieder eine Bilinearform.

Als nächstes zeigen wir, dass $\mathrm{Bil}_K(U)$ unter der Multiplikation mit Skalaren abgeschlossen ist. Es sei dazu $c \in K$. Wir zeigen, dass $c\beta$ wieder eine Bilinearform ist. Für alle $u, u', v \in U$ und $a \in K$ gilt

$$\begin{aligned}(c\beta)(au+u', v) = c\beta(au+u', v) &= ac\beta(u,v) + c\beta(u',v)\\ &= a(c\beta)(u,v) + (c\beta)(u',v).\end{aligned}$$

Führt man dieselbe Rechnung in der zweiten Komponente aus, folgt analog

$$(c\beta)(u, av+v') = a(c\beta)(u,v) + (c\beta)(u,v')$$

für alle $u, v, v' \in U$ und $a \in K$. □

II. Matrixdarstellung

Bekanntlich können wir lineare Abbildungen zwischen Vektorräumen nach Wahl von geordneten Basen durch Matrizen darstellen. Diese Schreibweise durch Matrizen ist wesentlich um mit linearen Abbildungen zu rechnen. Jetzt werden wir in ähnlicher Weise Bilinearformen durch Matrizen darstellen. Auch dazu muss man wieder eine geordnete Basis wählen. Um mit diesen Matrixdarstellungen sinnvoll rechnen zu können, müssen wir uns im Anschluss überlegen was bei einem Basiswechsel passiert.

3.1.10 **Definition** Sei U ein endlich-dimensionaler K-Vektorraum und sei $\beta\colon U \times U \to K$ eine Bilinearform. Ist $\mathcal{B} = (u_1, u_2, \ldots, u_n)$ eine geordnete Basis von U, dann nennen wir

$$\mathrm{M}_{\mathcal{B}}(\beta) = (\beta(u_i, u_j))_{i,j} \in \mathrm{M}_{n,n}(K)$$

die *Matrixdarstellung von β* zur Basis $\mathcal{B}$.

Die Matrixdarstellung von β zur Basis $\mathcal{B}$ erhält man also, indem man alle Kombinationen von Basisvektoren in die Bilinearform einsetzt.

3.1.11 **Beispiel**. Wir betrachten nochmal die Bilinearform $\beta\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ mit

$$\beta\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = 2x_1y_1 - x_1y_2 + x_2y_2$$

aus Beispiel 3.1.4. Es sei $\mathcal{B} = (e_1, e_2)$ die Standardbasis von $\mathbb{R}^2$. Wir bestimmen die Matrixdarstellung von β zur Standardbasis. Wir setzen die Basisvektoren in die Bilinearform ein und erhalten

$$\begin{aligned}
\beta\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) &= 2, \\
\beta\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) &= -1, \\
\beta\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) &= 0, \\
\beta\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) &= 1.
\end{aligned}$$

Damit ist

$$\mathrm{M}_{\mathcal{B}}(\beta) = \begin{pmatrix} 2 & -1 \\ 0 & 1 \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{R})$$

die Matrixdarstellung von β zur Standardbasis.

Die Matrixdarstellung hängt von der gewählten Basis ab. Wir setzen $v_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ und betrachten nun die geordnete Basis $\mathcal{C} = (v_1, e_2)$ von $\mathbb{R}^2$. Dann erhalten wir

$$\mathrm{M}_{\mathcal{C}}(\beta) = \begin{pmatrix} 7 & -1 \\ 1 & 1 \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{R})$$

als Matrixdarstellung von β zur Basis $\mathcal{C}$.

3.1.12 **Aufgabe**. Es sei $A \in \mathrm{M}_{n,n}(K)$. Wir betrachten die Bilinearform $\beta(u, v) = u^T Av$ aus Beispiel 3.1.7. **L**

Zeigen Sie: Ist $\mathcal{B}$ die Standardbasis von K^n, dann ist $\mathrm{M}_{\mathcal{B}}(\beta) = A$.

3.1.13 **Aufgabe**. Wir betrachten den Vektorraum $\mathbb{R}[X]_{\leq 2}$ der reellen Polynom vom Grad höchstens 2 mit der geordneten Basis $\mathcal{B} = (1, X, X^2)$. Es sei **L**

$$\beta(f, g) = \int_0^1 f(x)g(x) \,\mathrm{d}x$$

die Einschränkung der Bilinearform aus Beispiel 3.1.5 auf $\mathbb{R}[X]_{\leq 2}$. Bestimmen Sie $\mathrm{M}_{\mathcal{B}}(\beta)$.

3.1.14 **Satz.** *Es sei U ein endlich-dimensionaler K-Vektorraum mit einer geordneten Basis $\mathcal{B}$ und es sei $\beta \in \mathrm{Bil}_K(U)$. Für alle $u, v \in U$ gilt*

$$\beta(u, v) = \kappa_{\mathcal{B}}(u)^T \mathrm{M}_{\mathcal{B}}(\beta) \kappa_{\mathcal{B}}(v)$$

dabei ist $\kappa_{\mathcal{B}}$ die Koordinatenabbildung aus 2.2.22.

Insbesondere ist β vollständig durch die Matrixdarstellung $\mathrm{M}_{\mathcal{B}}(\beta)$ bestimmt.

Beweis. Es sei $n = \dim_K U$ und $\mathcal{B} = (u_1, \ldots, u_n)$. Sind $u, v \in U$ gegeben, dann schreiben wir die Vektoren als Linearkombination in der Basis $\mathcal{B}$, d.h.,

$$u = \sum_{i=1}^n a_i u_i \quad \text{und} \quad v = \sum_{i=1}^n b_i u_i$$

mit $a_1, \ldots, a_n, b_1, \ldots, b_n \in K$. Dann sind $\kappa_{\mathcal{B}}(u) = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ und $\kappa_{\mathcal{B}}(v) = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ die zugehörigen Koordinatenvektoren. Mit den Rechenregeln aus 3.1.3 erhalten wir

$$\beta(u, v) = \beta\Big(\sum_{i=1}^n a_i u_i, \sum_{j=1}^n b_j u_j\Big)$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} a_i \beta(u_i, u_j) b_j = \kappa_{\mathcal{B}}(u)^T \mathrm{M}_{\mathcal{B}}(\beta) \kappa_{\mathcal{B}}(v). \qquad \square$$

3.1.15 **Beispiel**. Mit Satz 3.1.14 kann man nun einsehen, dass jede Bilinearform auf K^n aus der Konstruktion in Beispiel 3.1.7 entsteht. In der Tat, sei $\beta \in \mathrm{Bil}(K^n)$ beliebig. Ist $A = \mathrm{M}_{\mathcal{B}}(\beta)$ die Matrixdarstellung von β zur Standardbasis $\mathcal{B} = (e_1, e_2, \ldots, e_n)$, dann gilt $\kappa_{\mathcal{B}}(u) = u$ für alle $u \in K^n$ und damit

$$\beta(u, v) = \kappa_{\mathcal{B}}(u)^T A \kappa_{\mathcal{B}}(v) = u^T A v.$$

3.1.16 **Korollar.** *Sei U ein n-dimensionaler Vektorraum mit einer geordneten Basis $\mathcal{B}$. Die Abbildung $\mathrm{M}_{\mathcal{B}} \colon \mathrm{Bil}_K(U) \to \mathrm{M}_{n,n}(K)$ mit $\beta \mapsto \mathrm{M}_{\mathcal{B}}(\beta)$ ist ein Isomorphismus von Vektorräumen. Insbesondere gilt $\dim_K \mathrm{Bil}_K(U) = n^2$.*

Beweis. Sei $\mathcal{B} = (u_1, \ldots, u_n)$. Wir weisen zuerst nach, dass die Abbildung $\mathrm{M}_{\mathcal{B}}$ linear ist. Für alle $\beta, \gamma \in \mathrm{Bil}(U)$ und $c \in K$ gilt dann

$$\begin{aligned} \mathrm{M}_{\mathcal{B}}(c\beta + \gamma) &= \big((c\beta + \gamma)(u_i, u_j)\big)_{i,j} \\ &= c\big(\beta(u_i, u_j)\big)_{i,j} + \big(\gamma(u_i, u_j)\big)_{i,j} = c\mathrm{M}_{\mathcal{B}}(\beta) + \mathrm{M}_{\mathcal{B}}(\gamma). \end{aligned}$$

Also ist $\mathrm{M}_{\mathcal{B}}$ linear; siehe 2.2.18. Aus Satz 3.1.14 wissen wir, dass eine Bilinearform durch eine Matrixdarstellung vollständig bestimmt ist, d.h., $\mathrm{M}_{\mathcal{B}}$ ist injektiv.

Schließlich müssen wir noch zeigen, dass $\mathrm{M}_{\mathcal{B}}$ surjektiv ist. Sei $A \in \mathrm{M}_{n,n}(K)$ gegeben. Wir definieren eine Bilinearform $\beta \colon U \times U \to K$ durch

$$\beta(u, v) = \kappa_{\mathcal{B}}(u)^T A \kappa_{\mathcal{B}}(v).$$

Wie in Beispiel 3.1.7 kann man nachweisen, dass dies eine Bilinearform ist. Wir behaupten, dass $\mathrm{M}_{\mathcal{B}}(\beta) = A$ gilt. Da $\kappa_{\mathcal{B}}(u_i) = e_i$ der i-te Standardbasisvektor ist, erhalten wir

$$\beta(u_i, u_j) = e_i^T A e_j = a_{i,j}$$

wie in Aufgabe 3.1.12 und damit $\mathrm{M}_{\mathcal{B}}(\beta) = A$. Daraus folgt, dass $\mathrm{M}_{\mathcal{B}}$ surjektiv ist.

Da isomorphe Vektorräume die gleiche Dimension haben, schließen wir $\dim_K \mathrm{Bil}(U) = n^2$ (siehe [MG, 8.3.19]). $\square$

In welcher Beziehung stehen verschiedene Matrixdarstellungen derselben Bilinearform zueinander? Die Antwort gibt der folgende wichtige Satz über den Basiswechsel.

3.1.17 **Transformationsformel für Bilinearformen.** *Sei U ein K-Vektorraum von endlicher Dimension mit zwei geordneten Basen $\mathcal{B}$ und $\mathcal{C}$. Für jede Bilinearform $\beta \in \mathrm{Bil}(U)$ gilt die Transformationsformel*

$$\mathrm{M}_{\mathcal{C}}(\beta) = S^T \mathrm{M}_{\mathcal{B}}(\beta) S$$

wobei $S = {}_{\mathcal{B}}\mathrm{M}_{\mathcal{C}}(\mathrm{id}_U)$ die Transformationsmatrix von $\mathcal{C}$ nach $\mathcal{B}$ ist.

Beweis. Es seien $\mathcal{B} = (u_1, \ldots, u_n)$ und $\mathcal{C} = (v_1, \ldots, v_n)$. Ist $S = (s_{ij})_{i,j} \in \mathrm{M}_{n,n}(K)$ die Transformationsmatrix ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{C}}(\mathrm{id}_U)$, dann gilt (vgl. Definition 2.2.23)

$$v_k = \sum_{i=1}^{n} s_{ik} u_i$$

für alle $1 \leq k \leq n$. Es seien nun $\ell, k \in \{1, \ldots, n\}$ beliebig. Mithilfe dieser Formel berechnen wir $\beta(v_k, v_\ell)$ und erhalten

$$\begin{aligned}\beta(v_k, v_\ell) &= \beta\Big(\sum_{i=1}^{n} s_{ik} u_i, \sum_{j=1}^{n} s_{j\ell} u_j\Big) = \sum_{i=1}^{n}\sum_{j=1}^{n} s_{ik}\beta(u_i, u_j) s_{j\ell} \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} s'_{ki}\beta(u_i, u_j) s_{j\ell}\end{aligned}$$

wobei $s'_{ki} = s_{ik}$ der Eintrag von S^T an der Stelle (k, i) ist. Nach Definition der Matrizenmultiplikation ist $\beta(v_k, v_\ell)$ damit genau der Eintrag von $S^T \mathrm{M}_{\mathcal{B}}(\beta) S$ an der Stelle (k, ℓ). Da k und ℓ beliebig waren, folgt die Aussage des Satzes. $\square$

3.1.18 **Aufgabe**. Berechnen Sie für die Bilinearform β aus Beispiel 3.1.11 die Matrixdarstellung $\mathrm{M}_{\mathcal{C}}(\beta)$ nochmal mit der Transformationsformel. **L**

III. Kongruenz von Bilinearformen

In diesem Abschnitt werden wir klären, was wir über Bilinearformen eigentlich wissen möchten und welche Fragen wir in dieser Lektion untersuchen werden. Dazu beginnen wir mit einer wichtigen Beobachtung. Aus der Transformationsformel 3.1.17 sehen wir, dass je zwei Matrixdarstellungen einer Bilinearform kongruent im Sinne von 2.6.12 sind. Auch umgekehrt gilt: Ist A kongruent zu $\mathrm{M}_{\mathcal{B}}(\beta)$, dann ist A ebenfalls eine Matrixdarstellung von β. Das folgt aus dem nächsten Lemma.

3.1.19 **Lemma.** *Es sei β eine Bilinearform auf dem endlich-dimensionalen K-Vektorraum U. Dann ist die Menge*

$$\{\ \mathrm{M}_{\mathcal{B}}(\beta) \mid \mathcal{B} \text{ geordnete Basis von } U\ \}$$

aller Matrixdarstellungen von β eine Kongruenzklasse von Matrizen.

Beweis. Sind $\mathrm{M}_{\mathcal{B}}(\beta)$ und $\mathrm{M}_{\mathcal{C}}(\beta)$ zwei Matrixdarstellungen von β, dann folgt aus der Transformationsformel 3.1.17, dass diese Matrizen kongruent sind. Die Menge aller Matrixdarstellungen von β ist also in einer Kongruenzklasse enthalten.

Sei nun umgekehrt A kongruent zu $\mathrm{M}_{\mathcal{B}}(\beta)$ für eine geordnete Basis $\mathcal{B} = (u_1, \dots, u_n)$. Es gibt also eine invertierbare Matrix $S = (s_{ij})_{i,j} \in \mathrm{GL}_n(K)$, sodass $A = S^T \mathrm{M}_{\mathcal{B}}(\beta) S$ gilt. Für $j \in \{1, \dots, n\}$ definieren wir nun

$$v_j = \sum_{i=1}^{n} s_{ij} u_i.$$

Da S invertierbar ist, ist dann auch $\mathcal{C} = (v_1, \dots, v_n)$ eine Basis von U und die Matrix S ist genau die Transformationsmatrix ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{C}}(\mathrm{id}_U)$. Aus der Transformationsformel 3.1.17 erhalten wir nun

$$A = S^T \mathrm{M}_{\mathcal{B}}(\beta) S = {}_{\mathcal{B}}\mathrm{M}_{\mathcal{C}}(\mathrm{id}_U)^T \mathrm{M}_{\mathcal{B}}(\beta) {}_{\mathcal{B}}\mathrm{M}_{\mathcal{C}}(\mathrm{id}_U) = \mathrm{M}_{\mathcal{C}}(\beta).$$

Also ist A ebenfalls eine Matrixdarstellung von β. Da A beliebig war, ist die Menge der Matrixdarstellungen von β genau eine Kongruenzklasse. □

3.1.20 **Beispiel**. Auf dem $\mathbb{C}^2$ betrachten wir die Bilinearform $\beta(u, v) = u^T A v$ mit der Matrix

$$A = \begin{pmatrix} 5 & 1+3i \\ 1+3i & -1+2i \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{C}).$$

Die Matrix A ist die Matrixdarstellung von β zur Standardbasis $\mathcal{B} = (e_1, e_2)$; siehe 3.1.12. Es gilt $A = S^T I_2 S$ mit

$$S = \begin{pmatrix} 1 & 1+i \\ 2 & i \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{C}).$$

Damit sind A und die Einheitsmatrix I_2 kongruent, d.h., I_2 ist eine andere Matrixdarstellung von β.

3.1.21 **Frage: Hat jede Bilinearform eine „einfache" Matrixdarstellung**? Wir wissen bereits, dass eine Bilinearform auf einem endlich-dimensionalen Vektorraum verschiedene Matrixdarstellungen hat. Wie wir gerade im Beispiel 3.1.20 gesehen haben, kann es passieren, dass eine Matrixdarstellung besonders „einfach" erscheint, während eine andere relativ „kompliziert" aussieht.

Ist also eine Bilinearform β gegeben, dann drängen sich folgende Fragen auf:

- Gibt es eine geordnete Basis $\mathcal{B}$, sodass $\mathrm{M}_{\mathcal{B}}(\beta)$ besonders „einfach" ist?

- Falls ja: Wie kann man eine solche Basis bestimmen?

Welche Matrizen wir als „einfach“ betrachten, werden wir später klären. Um mit dieser „einfachen“ Matrixdarstellung unkompliziert rechnen zu können, sollten jedenfalls viele Einträge 0 in der Matrix zu finden sein.

Da die Menge aller Matrixdarstellungen einer Bilinearform genau eine Kongruenzklasse von Matrizen ist (Lemma 3.1.19), kann man diese Fragen natürlich auch als Fragen über Matrizen formulieren. Gegeben eine quadratische Matrix $A \in \mathrm{M}_{n,n}(K)$ möchten wir also folgendes wissen:

- Gibt es in der Kongruenzklasse von A eine besonders „einfache“ Matrix?

- Falls ja: Wie kann man einen solchen Vertreter der Kongruenzklasse bestimmen?

In dieser Lektion werden wir diese Fragen nur für spezielle Bilinearformen (symmetrische oder alternierende) beantworten. Im Fall symmetrischer Bilinearformen benötigen wir dazu noch eine kleine Einschränkung an den Körper K.

Wir wollen diese Fragen noch etwas genauer ausdrücken. Dazu wechseln wir einmal die Perspektive. Eine Matrixdarstellung einer Bilinearform sagt uns, wie wir in gegebenen Koordinaten mit einer Bilinearform rechnen können. Aus der Transformationsformel wissen wir außerdem, wie man von einem Koordinatensystem in ein anderes wechselt. Die wesentlichen Eigenschaften einer Bilinearform sollten aber nicht davon abhängen, in welchen Koordinaten man mit dieser Form rechnet. In anderen Worten: Haben zwei Bilinearformen dieselben Matrixdarstellungen (bezüglich verschiedener Basen), dann sind diese Bilinearformen „sehr ähnlich“ und haben dieselben Eigenschaften. Das führt uns zur folgenden Definition.

3.1.22 **Definition** Zwei Bilinearformen β_1, β_2 auf einem endlich-dimensionalen K-Vektorraum U nennen wir *kongruent*, wenn es geordnete Basen $\mathcal{B}_1$ und $\mathcal{B}_2$ von U gibt dergestalt, dass gilt

$$\mathrm{M}_{\mathcal{B}_1}(\beta_1) = \mathrm{M}_{\mathcal{B}_2}(\beta_2).$$

3.1.23 **Lemma.** *Bilinearformen β_1, β_2 auf einem endlich-dimensionalen K-Vektorraum U sind genau dann kongruent, wenn für ein Paar (und dann für alle Paare) von geordneten Basen $\mathcal{C}_1$ und $\mathcal{C}_2$ von U die Matrixdarstellungen $\mathrm{M}_{\mathcal{C}_1}(\beta_1)$ und $\mathrm{M}_{\mathcal{C}_2}(\beta_2)$ kongruent sind.*

Beweis. Angenommen β_1 und β_2 sind kongruent. Es gilt also $\mathrm{M}_{\mathcal{B}_1}(\beta_1) = \mathrm{M}_{\mathcal{B}_2}(\beta_2)$ für geeignete geordnete Basen $\mathcal{B}_1, \mathcal{B}_2$ von U. Es seien $\mathcal{C}_1, \mathcal{C}_2$ beliebige geordnete Basen von U. Aus Lemma 3.1.19 folgt dann, dass $\mathrm{M}_{\mathcal{B}_i}(\beta_i)$ und $\mathrm{M}_{\mathcal{C}_i}(\beta_i)$ für $i \in \{1, 2\}$ kongruent sind. Da Kongruenz eine Äquivalenzrelation ist, schließen wir, dass auch $\mathrm{M}_{\mathcal{C}_1}(\beta_1)$ und $\mathrm{M}_{\mathcal{C}_2}(\beta_2)$ kongruent sind.

Nehmen wir umgekehrt an, dass $\mathrm{M}_{\mathcal{C}_1}(\beta_1)$ und $\mathrm{M}_{\mathcal{C}_2}(\beta_2)$ für irgendwelche geordnete Basen $\mathcal{C}_1$ und $\mathcal{C}_2$ kongruent sind. Da zwei Kongruenzklassen entweder disjunkt oder gleich sind, folgt aus Lemma 3.1.19

$$\{ \mathrm{M}_{\mathcal{B}}(\beta_1) \mid \mathcal{B} \text{ geordnete Basis von } U \} = \{ \mathrm{M}_{\mathcal{B}}(\beta_2) \mid \mathcal{B} \text{ geordnete Basis von } U \}$$

und damit gibt es geordnete Basen von U, sodass die Matrixdarstellungen von β_1 und β_2 gleich sind. □

3.1.24 **Korollar.** *Sei U ein endlich-dimensionaler K-Vektorraum. Die Kongruenz von Bilinearformen ist eine Äquivalenzrelation auf* $\mathrm{Bil}_K(U)$

Beweis. Die Kongruenz von Bilinearformen ist reflexiv, denn mit $\mathcal{B} = \mathcal{B}_1 = \mathcal{B}_2$ gilt $\mathrm{M}_{\mathcal{B}_1}(\beta) = \mathrm{M}_{\mathcal{B}_2}(\beta)$ für jede Bilinearform β. Die Relation ist auch symmetrisch, denn aus $\mathrm{M}_{\mathcal{B}_1}(\beta_1) = \mathrm{M}_{\mathcal{B}_2}(\beta_2)$ folgt sicherlich auch $\mathrm{M}_{\mathcal{B}_2}(\beta_2) = \mathrm{M}_{\mathcal{B}_1}(\beta_1)$.

Wir prüfen nun, dass die Relation auch transitiv ist. Es seien dazu $\beta_1, \beta_2, \beta_3 \in \mathrm{Bil}_K(U)$. Wir nehmen an, dass β_1 und β_2, sowie β_2 und β_3 kongruent sind. Sei $\mathcal{B}$ eine geordnete Basis. Nach Lemma 3.1.23 sind $\mathrm{M}_{\mathcal{B}}(\beta_1)$ und $\mathrm{M}_{\mathcal{B}}(\beta_2)$, sowie $\mathrm{M}_{\mathcal{B}}(\beta_2)$ und $\mathrm{M}_{\mathcal{B}}(\beta_3)$ kongruent. Da Kongruenz von Matrizen eine Äquivalenzrelation ist (siehe 2.6.12), sind damit auch $\mathrm{M}_{\mathcal{B}}(\beta_1)$ und $\mathrm{M}_{\mathcal{B}}(\beta_3)$ kongruent und nach Lemma 3.1.23 sind damit auch β_1 und β_3 kongruent. □

3.1.25 **Klassifikation und Normalformproblem für Bilinearformen**. Unser Ziel beim Studium der Bilinearformen ist es, die Bilinearformen bis auf Kongruenz zu *klassifizieren*. Aber was heißt das? Die Menge der Bilinearformen $\mathrm{Bil}_K(U)$ auf einem endlich-dimensionalen Vektorraum zerfällt in verschiedene Kongruenzklassen – also in Äquivalenzklassen bzgl. der Kongruenz von Bilinearformen. Das Ziel der Untersuchung von Bilinearformen ist es, diese Kongruenzklassen vollständig zu beschreiben.

Es gibt verschiedene Möglichkeiten eine solche Klassifikation anzugehen. Beispielsweise kann man versuchen, die Kongruenz von Matrizen als Faserrelation einer Abbildung zu identifizieren. Auf diese Weise haben wir in Satz 2.6.15 bereits die Äquivalenzklassen von Matrizen klassifiziert. Ein anderer Lösungsansatz besteht darin, ein Vertretersystem der Kongruenzklassen zu bestimmen, d.h., für jede Kongruenzklasse einen eindeutigen Vertreter dieser Klasse anzugeben.

Verbinden wir das mit der Frage 3.1.21 nach „einfachen“ Matrixdarstellungen, dann erhalten wir das folgende *Normalformproblem*:

> Man gebe aus jeder Kongruenzklasse von Matrizen einen eindeutigen „einfachen“ Vertreter an.

Hat man solche eindeutige „einfache“ Vertreter gefunden und festgelegt, dann nennt man den Vertreter einer Klasse oft die *Normalform*.

Die Klassifikation der Bilinearformen werden wir nur für alternierende Bilinearformen und für symmetrische Bilinearformen über $\mathbb{R}$ und $\mathbb{C}$ vollständig behandeln. Bevor wir uns diesen Fällen zuwenden, benötigen wir noch weitere Hilfsmittel.

IV. Einsetzen

Bilinearformen sind in beiden Komponenten linear. Das kann man verwenden um durch Einsetzen eines Vektors in eine der Komponenten eine Linearform zu bekommen. Diese Operation wollen wir uns in diesem Abschnitt anschauen.

3.1.26 **Einsetzen**. Sei K ein Körper und sei U ein Vektorraum über K. Es sei $\beta\colon U \times U \to K$ eine Bilinearform. Für $w \in U$ definieren wir zwei Linearformen $\iota^1_w(\beta), \iota^2_w(\beta) \in U^*$ durch die Abbildungsvorschriften

$$\iota^1_w(\beta)(u) = \beta(w,u) \quad \text{und} \quad \iota^2_w(\beta)(u) = \beta(u,w).$$

In anderen Worten: $\iota^k_w(\beta)$ setzt den Vektor w fest in die k-te Komponente von β ein. Dass $\iota^1_w(\beta)$ eine Linearform ist, folgt aus der Bilinearität von β, denn für alle $u, u' \in U$ gilt

$$\iota^1_w(\beta)(u+u') = \beta(w,u+u') = \beta(w,u) + \beta(w,u') = \iota^1_w(\beta)(u) + \iota^1_w(\beta)(u')$$

und für alle $u \in U$ und $a \in K$ gilt außerdem

$$\iota^1_w(\beta)(au) = \beta(w,au) = a\beta(w,u) = a\iota^1_w(\beta)(u).$$

Analog sieht man, dass $\iota^2_w(\beta)$ eine Linearform ist.

Man kann zeigen, dass

$$\iota^k_w\colon \operatorname{Bil}_K(U) \to U^*$$

für alle $w \in U$ und $k \in \{1,2\}$ eine K-lineare Abbildung ist. Wir werden das etwas allgemeiner in 4.5.6 nachrechnen und verzichten hier darauf.

3.1.27 **Beispiel**. Wir betrachten nochmal die Bilinearform $\beta\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ mit

$$\beta\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = 2x_1y_1 - x_1y_2 + x_2y_2$$

aus Beispiel 3.1.11. Setzt man den Vektor $w = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$ in die erste Komponente ein, d.h., man setzt in dieser Formel $x_1 = 3$ und $x_2 = -1$ ein, dann erhält man die Linearform

$$\iota_w^1(\beta)\left(\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = 6y_1 - 4y_2.$$

3.1.28 **Beispiel**. Wir betrachten den Vektorraum $U = K^n$ mit einer Bilinearform β. Es sei $\mathcal{B} = (e_1, \dots, e_n)$ die Standardbasis. Wir wissen aus Beispiel 3.1.15, dass β die Gestalt $\beta(u, v) = u^T A v$ für eine Matrix $A \in \mathrm{M}_{n,n}(K)$ hat.

Sei $w \in K^n$. Was ist $\iota_w^1(\beta)$? Wir erinnern uns, dass wir im Beispiel 2.5.5 einen Isomorphismus $E\colon U^* \to \mathrm{M}_{1,n}(K)$ zwischen dem Dualraum U^* und dem Raum der Zeilenvektoren $\mathrm{M}_{1,n}(K)$ angegeben haben; genauer ist $E(\alpha) = (\alpha(e_1), \dots, \alpha(e_n))$. Um zu verstehen was $\iota_w^1(\beta)$ ist, berechnen wir den zugehörigen Zeilenvektor $E(\iota_w^1(\beta))$. Dazu beobachten wir, dass $\beta(w, e_i) = w^T A e_i$ genau der i-te Eintrag des Zeilenvektors $w^T A$ ist. Damit erhalten wir

$$E(\iota_w^1(\beta)) = (\beta(w, e_1), \beta(w, e_2), \dots, \beta(w, e_n)) = w^T A.$$

Wenig überraschend entspricht das Ergebnis dem Einsetzen von w in die ersten Komponente in der Formel $\beta(u, v) = u^T A v$.

3.1.29 **Aufgabe**. Wir betrachten den Vektorraum $U = K^n$ mit einer Bilinearform $\beta(u, v) = u^T A v$. Bestimmen Sie $E(\iota_w^2(\beta))$. **L**

3.1.30 **Beispiel**. Es seien $\alpha_1, \alpha_2 \in U^*$. Wir betrachten die Bilinearform $\alpha_1 \otimes \alpha_2$ aus Beispiel 3.1.6. Sei $w \in U$. Für alle Vektoren $u \in U$ gilt dann

$$\iota_w^1(\alpha_1 \otimes \alpha_2)(u) = (\alpha_1 \otimes \alpha_2)(w, u) = \alpha_1(w)\alpha_2(u).$$

In anderen Worten ist $\iota_w^1(\alpha_1 \otimes \alpha_2) = \alpha_1(w)\alpha_2$ in U^*. Genauso kann man zeigen, dass $\iota_w^2(\alpha_1 \otimes \alpha_2) = \alpha_2(w)\alpha_1$ ist.

Man kann den Spieß jetzt umdrehen und im Ausdruck $\iota_w^k(\beta)$ den Vektor w variieren. Was erhält man dann? Für jeden Vektor $w \in U$ erhalten wir eine Linearform $\iota_w^k(\beta)$. Damit definieren wir eine Abbildung $\beta^{(k)}\colon U \to U^*$ durch die Vorschrift $w \mapsto \iota_w^k(\beta)$. Es gilt also $\beta^{(k)}(w) = \iota_w^k(\beta)$. Wir zeigen nun, dass die beiden Abbildungen $\beta^{(1)}$ und $\beta^{(2)}$ wieder K-linear sind.

3.1.31 **Satz.** *Sei U ein K-Vektorraum und sei $\beta \in \mathrm{Bil}_K(U)$. Für $k \in \{1,2\}$ definiert*

$$\beta^{(k)}\colon U \to U^* \text{ mit } w \mapsto \iota_w^k(\beta).$$

eine lineare Abbildung.

Beweis. Wir verwenden das Kriterium aus Aufgabe 2.2.18. Es seien $w, w' \in U$ und $a \in K$ gegeben. Dann gilt für alle $u \in U$

$$\begin{aligned} \beta^{(1)}(aw+w')(u) &= \iota^1_{aw+w'}(\beta)(u) = \beta(aw+w', u) \\ &= a\beta(w,u) + \beta(w',u) && (\beta \text{ bilinear}) \\ &= a\iota^1_w(\beta)(u) + \iota^1_{w'}(\beta)(u) && (\text{Def. } \iota^1_w) \\ &= a\beta^{(1)}(w)(u) + \beta^{(1)}(w')(u). \end{aligned}$$

Da u beliebig war, folgt daraus $\beta^{(1)}(aw+w') = a\beta^{(1)}(w)+\beta^{(1)}(w')$ für alle $w, w' \in U$ und $a \in K$. Das heißt, $\beta^{(1)}$ ist linear. Analog zeigt man wieder, dass auch $\beta^{(2)}$ linear ist (Übungsaufgabe!). □ L

Betrachten wir diese formalen Überlegungen nochmal aus einem anderen Blickwinkel. Was weiß man über eine Bilinearform β, wenn man $\beta^{(1)}$ oder $\beta^{(2)}$ kennt? Dazu beobachten wir zuerst, dass für alle $u, v \in U$ die Gleichheit $\beta(u,v) = \beta^{(1)}(u)(v) = \beta^{(2)}(v)(u)$ gilt. Das heißt aber, dass β sowohl durch $\beta^{(1)}$ als auch durch $\beta^{(2)}$ vollständig bestimmt ist. Der nächste Satz drückt diesen Zusammenhang aus.

3.1.32 **Satz.** *Sei U ein K-Vektorraum und sei $k \in \{1,2\}$. Die Abbildung $\cdot^{(k)}\colon \mathrm{Bil}_K(U) \to \mathrm{Hom}_K(U,U^*)$ mit $\beta \mapsto \beta^{(k)}$ ist ein Isomorphismus von Vektorräumen.*

Beweis. Zuerst zeigen wir, dass $\cdot^{(k)}$ linear ist. Es seien $\beta_1, \beta_2 \in \mathrm{Bil}_K(U)$ und $a \in K$. Dann gilt für alle $w \in U$

$$\begin{aligned} (a\beta_1+\beta_2)^{(k)}(w) &= \iota^k_w(a\beta_1+\beta_2) \\ &= a\iota^k_w(\beta_1) + \iota^k_w(\beta_2) && (\iota^k_w \text{ linear } 4.5.6) \\ &= a\beta_1^{(k)}(w) + \beta_2^{(k)}(w) \end{aligned}$$

und da w beliebig war, folgt $(a\beta_1+\beta_2)^{(k)} = a\beta_1^{(k)} + \beta_2^{(k)}$. Aus Aufgabe 2.2.18 folgt nun, dass $\cdot^{(k)}$ eine lineare Abbildung ist.

Als nächstes prüfen wir, dass $\cdot^{(k)}$ ein Isomorphismus ist. Dazu zeigen wir, dass die Abbildung bijektiv ist. Wir führen den Beweis nur für $k=1$ und lassen es wieder als kleine Übung, die passenden Änderungen im Fall $k=2$ zu finden.

Surjektivität: Sei $\gamma \in \mathrm{Hom}_K(U, U^*)$. Wir definieren $\beta \in \mathrm{Bil}_K(U)$ durch $\beta(u,v) = \gamma(u)(v)$ für alle $u, v \in U$. Warum ist β eine Bilinearform? Die Linearität in der ersten Komponente erhält man, weil γ linear ist. Die Linearität in der zweiten Komponente folgt daraus, dass $\gamma(u)$ für alle $u \in U$ eine Linearform ist. Dann gilt $\beta^{(1)}(u)(v) = \beta(u,v) = \gamma(u)(v)$ für alle $u, v \in U$ und damit $\beta^{(1)} = \gamma$.

Injektivität: Angenommen, es ist $\beta^{(1)} = 0$. Dann gilt für alle $u, v \in U$ auch

$$\beta(u,v) = \beta^{(1)}(u)(v) = 0$$

und damit $\beta = 0$. Die Abbildung $\cdot^{(1)}$ hat also einen trivialen Kern und ist somit injektiv; siehe [MG, 8.3.10]. $\square$

Wir werden bald sehen, dass es in manchen Anwendungen praktischer ist, statt der Bilinearform β mit der zugehörigen Abbildung $\beta^{(1)}$ zu arbeiten. Da $\beta^{(1)}$ und $\beta^{(2)}$ so direkt aus β entstehen, überlegen wir uns noch welchen Zusammenhang es zwischen den Matrixdarstellungen gibt. Um diese Verbindung zu sehen, darf man aber keine beliebigen Basen auf U und U^* nehmen, sondern muss zu einer Basis $\mathcal{B}$ auf U die zugehörige Dualbasis auf U^* wählen.

3.1.33 **Lemma.** *Es sei U ein endlich-dimensionaler K-Vektorraum mit einer geordneten Basis $\mathcal{B}$. Es sei $\mathcal{B}^*$ die zugehörige Dualbasis. Für jede Bilinearform $\beta \in \mathrm{Bil}_K(U)$ gelten dann*

(a) ${}_{\mathcal{B}^*}\mathrm{M}_{\mathcal{B}}(\beta^{(1)}) = \mathrm{M}_{\mathcal{B}}(\beta)^T$.

(b) ${}_{\mathcal{B}^*}\mathrm{M}_{\mathcal{B}}(\beta^{(2)}) = \mathrm{M}_{\mathcal{B}}(\beta)$.

Beweis. Es seien $\mathcal{B} = (u_1, \dots, u_n)$ und $\mathcal{B}^* = (u_1^*, \dots, u_n^*)$. Es sei ${}_{\mathcal{B}^*}\mathrm{M}_{\mathcal{B}}(\beta^{(1)}) = (a_{ij})_{i,j}$ die Matrixdarstellung von $\beta^{(1)}$. Für alle $j \in \{1, \dots, n\}$ gilt dann

$$\beta^{(1)}(u_j) = \sum_{k=1}^{n} a_{kj} u_k^*.$$

Sei $i \in \{1, \dots, n\}$. Wir setzen den Vektor u_i in $\beta^{(1)}(u_j)$ ein und erhalten

$$\beta(u_j, u_i) = \beta^{(1)}(u_j)(u_i) = \sum_{k=1}^{n} a_{kj} u_k^*(u_i) = a_{ij}.$$

Da dies für alle i, j gilt, folgt ${}_{\mathcal{B}^*}\mathrm{M}_{\mathcal{B}}(\beta^{(1)}) = \mathrm{M}_{\mathcal{B}}(\beta)^T$.

Sei nun ${}_{\mathcal{B}^*}\mathrm{M}_{\mathcal{B}}(\beta^{(2)}) = (b_{ij})_{i,j}$ die Matrixdarstellung von $\beta^{(2)}$. Für alle $j \in \{1, \ldots, n\}$ gilt dann

$$\beta^{(2)}(u_j) = \sum_{k=1}^{n} b_{kj} u_k^*.$$

Wieder setzen wir einen Basisvektor u_i in $\beta^{(2)}(u_j)$ ein:

$$\beta(u_i, u_j) = \beta^{(2)}(u_j)(u_i) = \sum_{k=1}^{n} b_{kj} u_k^*(u_i) = b_{ij}.$$

Da dies für alle i, j gilt, folgt ${}_{\mathcal{B}^*}\mathrm{M}_{\mathcal{B}}(\beta^{(2)}) = \mathrm{M}_{\mathcal{B}}(\beta)$. □

Zum Abschluss dies Abschnittes wollen wir – in Analogie zur Definition der dualen Abbildung – für eine lineare Abbildung $\varphi\colon U \to V$ eine Abbildung $\varphi^*\colon \mathrm{Bil}_K(V) \to \mathrm{Bil}_K(U)$ untersuchen. Die Abbildungsvorschrift entsteht dabei durch denselben „Trick“, aus dem wir auch die duale Abbildung konstruiert haben.

3.1.34 **Lineare Abbildungen und Bilinearformen**. Es seien U, V zwei K-Vektorräume. Ist $\varphi\colon U \to V$ eine K-lineare Abbildung, dann kann man diese verwenden, um aus einer Bilinearform auf V eine Bilinearform auf U zu konstruieren. Genauer: Für $\beta \in \mathrm{Bil}_K(V)$ definieren wir $\varphi^*(\beta) \in \mathrm{Bil}_K(U)$ durch

$$\varphi^*(\beta)(u_1, u_2) := \beta(\varphi(u_1), \varphi(u_2))$$

für alle $u_1, u_2 \in U$. Da φ linear ist, ist $\varphi^*(\beta)$ wieder in beiden Komponenten linear und damit eine Bilinearform. Man sagt auch, dass $\varphi^*(\beta)$ durch *zurückziehen* (engl. pull back) von β nach U entsteht.

Man kann zeigen, dass auch $\varphi^*\colon \mathrm{Bil}_K(V) \to \mathrm{Bil}_K(U)$ eine K-lineare Abbildung ist. Wir werden eine allgemeinere Aussage in Satz 4.5.8 kennenlernen und verzichten deshalb an dieser Stelle auf einen Beweis.

3.1.35 **Aufgabe**. Zeigen Sie, dass $\varphi^*(\beta)$ eine Bilinearform ist. **L**

3.1.36 **Aufgabe**. Sei U ein endlich-dimensionaler K-Vektorraum und seien β_1, β_2 zwei Bilinearformen auf U. Zeigen Sie, dass β_1 und β_2 genau dann kongruent sind, wenn es einen Isomorphismus $\varphi\colon U \to U$ gibt, sodass $\varphi^*(\beta_1) = \beta_2$. **L**

V. Reguläre und ausgeartete Formen

In diesem Abschnitt lernen wir eine wichtige Einteilung der Bilinearformen in zwei Klassen kennen: ausgeartete und reguläre Formen.

3.1.37 **Definition** Sei V ein K-Vektorraum. Eine Bilinearform $\beta \in \mathrm{Bil}_K(V)$ nennt man *regulär*, wenn sie folgende Eigenschaft hat: Für alle $u \neq 0$ in V existiert ein $v \in V$ mit $\beta(u,v) \neq 0$. Eine Bilinearform, die diese Bedingungen nicht erfüllt, nennen wir *ausgeartet*.

In vielen Texten werden reguläre Bilinearformen auch *nicht ausgeartet* genannt. Der nächste Satz liefert mehrere nützliche äquivalente Bedingungen für die Regularität von Bilinearformen auf endlich-dimensionalen Vektorräumen.

3.1.38 **Satz.** *Sei U ein endlich-dimensionaler K-Vektorraum und sei β eine Bilinearform auf U. Die folgenden Aussagen sind äquivalent:*

(i) β ist regulär.

(ii) $\beta^{(1)} : U \to U^*$ ist ein Isomorphismus.

(iii) Für alle $v \neq 0$ in U existiert ein $u \in U$ mit $\beta(u,v) \neq 0$.

(iv) $\beta^{(2)} : U \to U^*$ ist ein Isomorphismus.

(v) Für jede geordnete Basis $\mathcal{B}$ von U ist $\mathrm{M}_{\mathcal{B}}(\beta)$ invertierbar.

(vi) Es gibt eine geordnete Basis $\mathcal{B}$ von U, sodass $\mathrm{M}_{\mathcal{B}}(\beta)$ invertierbar ist.

Beweis. „(i) $\Longrightarrow$ (ii)“: Wir nehmen an, dass β regulär ist und zeigen, dass $\beta^{(1)}$ ein Isomorphismus ist. Da U und U^* Vektorräume derselben endlichen Dimension sind, ist es ausreichend zu zeigen, dass $\beta^{(1)}$ injektiv ist; siehe [MG, 8.3.17]. Wir weisen nach, dass $\beta^{(1)}$ einen trivialen Kern hat. Sei $u \in U$ mit $u \neq 0$ gegeben. Dann gibt es einen Vektor $v \in U$ mit $\beta^{(1)}(u)(v) = \beta(u,v) \neq 0$. Also ist $\beta^{(1)}(u) \neq 0$ und u liegt nicht im Kern. Das heißt, $\mathrm{Ker}(\beta^{(1)}) = \{0\}$ und $\beta^{(1)}$ ist injektiv.

„(iii) $\Longrightarrow$ (iv)“: Diese Implikation beweist man genau wie die Implikation „(i) $\Longrightarrow$ (ii)“ indem man die Rollen von erster und zweiter Komponente vertauscht.

„(ii) $\Longrightarrow$ (iii)“: Wir nehmen an, dass $\beta^{(1)}$ ein Isomorphismus ist. Sei $v \neq 0$ in U beliebig. Es gibt eine Linearform $\alpha \in U^*$ mit $\alpha(v) \neq 0$; siehe Lemma 2.5.6. Da $\beta^{(1)}$ surjektiv ist, gibt es einen Vektor $u \in U$ mit $\beta^{(1)}(u) = \alpha$ und damit gilt $\beta(u,v) = \beta^{(1)}(u)(v) = \alpha(v) \neq 0$. Damit haben wir (iii) nachgewiesen.

„(iv) $\Longrightarrow$ (i)“: Wir nehmen jetzt an, dass $\beta^{(2)}$ ein Isomorphismus ist. Sei $u \neq 0$ in U. Dann gibt es eine Linearform $\alpha \in U^*$ mit $\alpha(u) \neq 0$ (Lemma 2.5.6). Da $\beta^{(2)}$ surjektiv ist, finden wir einen Vektor $v \in U$ mit $\beta^{(2)}(v) = \alpha$ und damit gilt $\beta(u,v) = \beta^{(2)}(v)(u) = \alpha(u) \neq 0$; also haben wir (i) gezeigt.

„(iv) $\implies$ (v)“: Sei $\mathcal{B}$ eine beliebige geordnete Basis von U und sei $\mathcal{B}^*$ die zugehörige Dualbasis von U^*. Nach Lemma 3.1.33 ist $\mathrm{M}_{\mathcal{B}}(\beta) = {}_{\mathcal{B}^*}\mathrm{M}_{\mathcal{B}}(\beta^{(2)})$. Da $\beta^{(2)}$ ein Isomorphismus ist, ist die Matrixdarstellung ${}_{\mathcal{B}^*}\mathrm{M}_{\mathcal{B}}(\beta^{(2)})$ invertierbar; siehe Lemma 2.2.25.

Die Implikation „(v) $\implies$ (vi)“ ist offensichtlich erfüllt, weil jeder Vektorraum eine Basis besitzt.

„(vi) $\implies$ (iv)“: Sei $\mathcal{B}$ eine geordnete Basis, sodass $\mathrm{M}_{\mathcal{B}}(\beta)$ invertierbar ist. Es sei $\mathcal{B}^*$ die zugehörige Dualbasis von U^*. Nach Lemma 3.1.33 ist $\mathrm{M}_{\mathcal{B}}(\beta) = {}_{\mathcal{B}^*}\mathrm{M}_{\mathcal{B}}(\beta^{(2)})$ und aus Lemma 2.2.25 folgt, dass $\beta^{(2)}$ ein Isomorphismus ist. □

3.1.39 **Bemerkung**. In unendlich-dimensionalen Vektorräumen gilt Satz 3.1.38 nicht. Insbesondere folgt dann aus Regularität wie in Definition 3.1.37 *nicht* die duale Bedingung (iii) aus 3.1.38.

3.1.40 **Beispiel**. Es sei $U = K^n$. Die Bilinearform $\beta(u, v) = u^T Av$ für eine Matrix $A \in \mathrm{M}_{n,n}(K)$ ist genau dann regulär, wenn die Matrix A invertierbar ist. Das folgt aus Bedingung (vi) in Satz 3.1.38, denn A ist die Matrixdarstellung von β zur Standardbasis; siehe 3.1.12.

3.1.41 **Aufgabe**. Es sei U ein endlich-dimensionaler K-Vektorraum mit $\dim_K(U) \geq 2$ und es seien $\alpha_1, \alpha_2 \in U^*$. Zeigen Sie, dass die Bilinearform $\alpha_1 \otimes \alpha_2$ aus Beispiel 3.1.6 ausgeartet ist. **L**

3.1.42 **Definition** Sei V ein K-Vektorraum und $\beta \in \mathrm{Bil}_K(V)$. Ein Unterraum $U \subseteq V$ ist *β-regulär*, wenn die Einschränkung $\beta|_U$ von β auf U regulär ist.

3.1.43 **Zerlegungssatz.** *Es sei V ein endlich-dimensionaler K-Vektorraum mit einer Bilinearform $\beta \in \mathrm{Bil}_K(V)$. Sei $U \subseteq V$ ein β-regulärer Unterraum. Dann gilt $V = U \oplus W$ mit*

$$W = \{w \in V \mid \beta(u, w) = 0 \textit{ für alle } u \in U\}.$$

Beweis. Zuerst bemerken wir, dass W ein Unterraum von V ist (Übungsaufgabe!). **L**
Wir zeigen, dass $U \cap W = \{0\}$ ist. Sei dazu $v \in U \cap W$. Da v in W liegt, gilt für alle $u \in U$ schon $\beta(u, v) = 0$. Da $\beta|_U$ regulär ist und $v \in U$ liegt, folgt aus Bedingung (iii) in Satz 3.1.38, dass $v = 0$ sein muss.

Nun zeigen wir $U + W = V$. Sei dazu $v \in V$ beliebig. Wir betrachten $\alpha = \beta^{(2)}(v) \in V^*$. Diese Linearform können wir auf U einschränken (d.h., wir setzen nur noch

Vektoren aus U ein) und erhalten dadurch $\alpha|_U \in U^*$. Da $\beta|_U$ regulär ist, ist $(\beta|_U)^{(2)}$ ein Isomorphismus. Es gibt also einen Vektor $u_0 \in U$ mit $(\beta|_U)^{(2)}(u_0) = \alpha|_U$. Das heißt, für alle Vektoren $u \in U$ gilt

$$\begin{aligned} \beta(u, u_0) = \beta|_U(u, u_0) &= (\beta|_U)^{(2)}(u_0)(u) && \text{(Def. } \cdot^{(2)}) \\ &= \alpha|_U(u) && \text{(Wahl von } u_0) \\ &= \alpha(u) = \beta^{(2)}(v)(u) = \beta(u, v). \end{aligned}$$

Aus der Linearität von β in der zweiten Komponente folgt

$$\beta(u, v - u_0) = 0$$

für alle $u \in U$. Das heißt, $v - u_0 \in W$ und damit gilt $v = u_0 + (v - u_0) \in U + W$. Da $v \in V$ beliebig war, schließen wir $V = U + W$. □

Eine Bilinearform β ist regulär, wenn eine Matrixdarstellung $\mathrm{M}_{\mathcal{B}}(\beta)$ invertierbar ist, d.h., wenn $\mathrm{M}_{\mathcal{B}}(\beta)$ vollen Rang hat; siehe [MG, 4.5.4]. Andernfalls ist die Bilinearform ausgeartet. Mit dem Rang der Matrixdarstellung kann man eine noch feinere Einteilung der Bilinearformen vornehmen.

3.1.44 **Definition und Satz.** *Sei V ein endlich-dimensionaler K-Vektorraum und sei $\beta \in \mathrm{Bil}_K(V)$. Der Rang einer Matrixdarstellung*

$$\mathrm{Rg}(\mathrm{M}_{\mathcal{B}}(\beta))$$

hängt nicht von der gewählten geordneten Basis $\mathcal{B}$ von V ab.

Diese Zahl nennen wir den Rang von β *und bezeichnen sie mit* $\mathrm{Rg}(\beta)$.

Beweis. Da je zwei Matrixdarstellungen von β kongruent sind, ist es ausreichend zu zeigen, dass kongruente Matrizen denselben Rang haben. Es seien dazu $A, B \in \mathrm{M}_{n,n}(K)$ kongruent. Dann gilt $A = S^T BS$ für eine invertierbare Matrix $S \in \mathrm{GL}_n(K)$. Aus Aufgabe 1.2.34 wissen wir, dass dann auch S^T invertierbar ist. Mit [MG, 4.5.5] folgt daraus

$$\mathrm{Rg}(A) = \mathrm{Rg}(S^T BS) = \mathrm{Rg}(S^T B) = \mathrm{Rg}(B).$$ □

3.1.45 **Definition** Es sei β eine Bilinearform auf V. Dann ist

$$N_\beta = \{u \in V \mid \beta(u, v) = 0 \text{ für alle } v \in V\}$$

ein Unterraum von V, den man den (ersten) *Nullraum* zu β nennt.

Dass es sich bei N_β um einen Unterraum handelt, kann man mit dem Unterraumkriterium nachrechnen. Es folgt aber auch direkt aus der folgenden Aufgabe.

3.1.46 **Aufgabe**. Es sei β eine Bilinearform auf V. Zeigen Sie: $N_\beta = \mathrm{Ker}(\beta^{(1)})$. **L**

Analog könnte man den zweiten Nullraum als Kern von $\beta^{(2)}$ definieren. Das ist für uns aber nicht wichtig, denn wir werden uns im Folgenden nur für Situationen interessieren, in denen der erste und zweite Nullraum übereinstimmen. Mithilfe des Nullraumes kann man eine Dimensionsformel für Bilinearformen herleiten.

3.1.47 **Satz.** *Es sei V ein endlich-dimensionaler K-Vektorraum und es sei $\beta \in \mathrm{Bil}_K(V)$. Dann gilt*

$$\dim_K(V) = \dim_K(N_\beta) + \mathrm{Rg}(\beta).$$

Beweis. Der Rangsatz für lineare Abbildungen (siehe [MG, 8.3.14]) liefert uns die Gleichung

$$\dim_K(V) = \dim_K(\mathrm{Ker}(\beta^{(1)})) + \mathrm{Rg}(\beta^{(1)}).$$

Aus Aufgabe 3.1.46 folgt nun $\dim_K(\mathrm{Ker}(\beta^{(1)})) = \dim_K(N_\beta)$. Es sei $A = \mathrm{M}_{\mathcal{B}}(\beta)$ eine Matrixdarstellung von β. Nach Definition des Ranges gilt $\mathrm{Rg}(\beta) = \mathrm{Rg}(A)$. Aus Lemma 3.1.33 wissen wir, dass $A^T = {}_{\mathcal{B}^*}\mathrm{M}_{\mathcal{B}}(\beta^{(1)})$ eine Matrixdarstellung von $\beta^{(1)}$ ist. Mit Korollar 2.6.16 schließen wir

$$\mathrm{Rg}(\beta^{(1)}) = \mathrm{Rg}(A^T) = \mathrm{Rg}(A) = \mathrm{Rg}(\beta)$$

und haben damit die Dimensionsformel beweisen. □

3.2. Symmetrische Bilinearformen

In diesem Abschnitt befassen wir uns mit den *symmetrischen* Bilinearformen und wir werden versuchen in diesem Fall die in 3.1.21 und 3.1.25 aufgeworfenen Fragen zu beantworten.

I. Definition und Beispiele

3.2.1 **Definition** Sei V ein Vektorraum. Eine Bilinearform β auf V heißt *symmetrisch*, wenn

$$\beta(u, v) = \beta(v, u)$$

für alle Vektoren $u, v \in V$ gilt.

3.2.2 **Beispiel**. (a) Die Bilinearform $\beta\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ mit

$$\beta\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = 2x_1y_1 - x_1y_2 + x_2y_2$$

aus Beispiel 3.1.4 ist nicht symmetrisch, denn es gilt

$$\beta\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = -1 \neq 0 = \beta\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right).$$

(b) Die Bilinearform β aus Beispiel 3.1.5 auf dem Vektorraum der Polynome $\mathbb{R}[X]$ ist symmetrisch, weil die Multiplikation reeller Zahlen kommutativ ist:

$$\beta(f, g) = \int_0^1 f(x)g(x)\,\mathrm{d}x = \int_0^1 g(x)f(x)\,\mathrm{d}x = \beta(g, f)$$

für alle $f, g \in \mathbb{R}[X]$.

3.2.3 **Aufgabe**. Zeigen Sie: Eine Bilinearform β ist genau dann symmetrisch, wenn $\beta^{(1)} = \beta^{(2)}$ gilt. **L**

Man kann auch anhand einer Matrixdarstellung sehen, ob die Bilinearform β symmetrisch ist. Dazu benötigen wir den folgenden Begriff.

3.2.4 **Definition** Eine quadratische Matrix $A = (a_{ij})_{i,j} \in \mathrm{M}_{n,n}(K)$ nennt man *symmetrisch*, wenn

$$A = A^T$$

gilt. Das heißt, es gilt $a_{ij} = a_{ji}$ für alle $i, j \in \{1, \ldots, n\}$.

3.2.5 **Beispiel**. (a) Die Matrix

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 7 \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{Q})$$

ist symmetrisch.

(b) Die Matrix

$$\begin{pmatrix} 1 & 0 & 6 \\ 0 & 3 & 0 \\ 8 & 0 & 1 \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{Q})$$

ist nicht symmetrisch, denn die Einträge an den Stellen $(1,3)$ und $(3,1)$ sind verschieden.

3.2.6 **Lemma.** *Sei V ein endlich-dimensionaler K-Vektorraum mit einer geordneten Basis $\mathcal{B}$ und sei $\beta \in \mathrm{Bil}_K(V)$. Die Bilinearform β ist genau dann symmetrisch, wenn die Matrixdarstellung $\mathrm{M}_{\mathcal{B}}(\beta)$ symmetrisch ist.*

Beweis. Es sei $\mathcal{B} = (v_1, \dots, v_n)$.

Nehmen wir zunächst an, dass β symmetrisch ist. Dann gilt $\beta(v_i, v_j) = \beta(v_j, v_i)$ für alle $i, j \in \{1, \dots, n\}$ und somit ist $\mathrm{M}_{\mathcal{B}}(\beta)$ symmetrisch.

Umgekehrt nehmen wir nun an, dass $A = \mathrm{M}_{\mathcal{B}}(\beta)$ symmetrisch ist, d.h., $A^T = A$. Mit Satz 3.1.14 erhalten wir für alle $u, v \in V$

$$\begin{aligned}
\beta(u,v) &= \kappa_{\mathcal{B}}(u)^T A \kappa_{\mathcal{B}}(v) \\
&= (\kappa_{\mathcal{B}}(u)^T A \kappa_{\mathcal{B}}(v))^T && (c = c^T \text{ für } (1 \times 1)\text{-Matrizen}) \\
&= \kappa_{\mathcal{B}}(v)^T A^T \kappa_{\mathcal{B}}(u) && (1.2.32 \text{ (ii)}) \\
&= \kappa_{\mathcal{B}}(v)^T A \kappa_{\mathcal{B}}(u) && (A \text{ symmetrisch}) \\
&= \beta(v,u).
\end{aligned}$$

Damit ist β symmetrisch. □

Insbesondere folgt mit Lemma 3.1.19, dass entweder alle Matrizen in einer Kongruenzklasse symmetrisch sind oder keine. In der folgenden Aufgabe sollen Sie diese Aussage nochmal direkt nachweisen.

3.2.7 **Aufgabe**. Es seien $A, S \in \mathrm{M}_{n,n}(K)$. Zeigen Sie: Ist A symmetrisch, dann ist auch $S^T AS$ symmetrisch. **L**

II. Diagonalisierungssatz

Wir werden uns nun mit der in 3.1.21 gestellten Frage nach einer „einfachen" Matrixdarstellung für symmetrische Bilinearformen befassen. Wir werden zeigen, dass es für symmetrische Bilinearformen über einem geeigneten Körper immer eine diagonale Matrixdarstellung gibt. Es wird sich herausstellen, dass wir mit dem Zerlegungssatz schon das wesentliche Hilfsmittel kennen. Aber welche Körper sind „geeignet"?

3.2.8 **Was ist eigentlich** 2? Sei K ein Körper. Nach Definition gibt es in K ein Einselement 1. Wir setzen $2 := 1 + 1$. Dann gibt es zwei Fälle, die auftreten können.

$2 = 0$**:** In diesem Fall sagt man, dass K ein Körper der *Charakteristik* 2 ist. Das einzige Beispiel, das wir in diesem Text kennenlernen, ist der Körper $\mathbb{F}_2$ mit 2 Elementen.

$2 \neq 0$**:** In diesem Fall ist 2 im Körper K eine Einheit. Es gibt also ein multiplikatives Inverses $2^{-1} \in K$. Anstelle von 2^{-1} schreiben wir oft $\frac{1}{2}$. Die meisten Körper, die wir kennen, haben diese Eigenschaft: $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ und die endlichen Körper $\mathbb{F}_p$ für jede ungerade Primzahl p.

Wir werden in diesem Abschnitt immer annehmen, dass $2 \neq 0$ in K gilt. Der Grund dafür ist das nächste Ergebnis.

3.2.9 **Polarisationsformel.** *Sei K ein Körper mit $2 \neq 0$. Ist V ein K-Vektorraum mit einer symmetrischen Bilinearform β, dann gilt*

$$\beta(u, v) = \frac{1}{2}\big(\beta(u + v, u + v) - \beta(u, u) - \beta(v, v)\big)$$

für alle $u, v \in V$.

Beweis. Es seien $u, v \in V$. Mit der Linearität in beiden Komponenten (vgl. 3.1.3) erhalten wir

$$\beta(u + v, u + v) = \beta(u, u) + \beta(u, v) + \beta(v, u) + \beta(v, v).$$

Da β symmetrisch ist, gilt $\beta(u, v) = \beta(v, u)$. Durch Umstellen der Gleichung erhalten wir

$$2\beta(u, v) = \beta(u + v, u + v) - \beta(u, u) - \beta(v, v).$$

Ist nun $2 \neq 0$ in K, dann können wir mit $\frac{1}{2}$ multiplizieren und erhalten die Polarisationsformel. □

Was ist das Besondere an der Polarisationsformel? Aus ihr folgt, dass die symmetrische Bilinearform β vollständig durch die Werte $\beta(u,u)$ mit $u \in V$ bestimmt ist; also die Werte, die man erhält, wenn man zweimal denselben Vektor in die Bilinearform einsetzt. Die Abbildung $u \mapsto \beta(u,u)$ nennt man auch die *quadratische Form* zu β.

3.2.10 **Diagonalisierungssatz.** *Sei K ein Körper mit $2 \neq 0$. Es seien $V \neq \{0\}$ ein endlich-dimensionaler K-Vektorraum und β eine symmetrische Bilinearform auf V. Dann gibt es eine geordnete Basis $\mathcal{B}$ von V, sodass $\mathrm{M}_{\mathcal{B}}(\beta)$ eine Diagonalmatrix ist, d.h.,*

$$\mathrm{M}_{\mathcal{B}}(\beta) = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & a_n \end{pmatrix} \in \mathrm{M}_{n,n}(K)$$

mit Skalaren $a_1, \ldots, a_n \in K$.

3.2.11 **Beweisstrategie**. Der Beweis des Diagonalisierungssatzes ist eigentlich recht einfach, wenn man die Idee dahinter verstanden hat. Was für eine Basis suchen wir? Wir suchen eine geordnete Basis $\mathcal{B} = (v_1, \ldots, v_n)$, sodass $\mathrm{M}_{\mathcal{B}}(\beta)$ eine Diagonalmatrix ist. Das heißt, es gilt $\beta(v_i, v_j) = 0$ für alle $i \neq j$ und die Werte $a_i = \beta(v_i, v_i)$ sind die Einträge auf der Diagonalen. Ist $a_i \neq 0$, dann ist $\langle v_i \rangle$ ein ein-dimensionaler β-regulärer Unterraum von V. In der Tat, für $cv_i \in \langle v_i \rangle$ mit $c \neq 0$ gilt

$$\beta(cv_i, v_i) = c\beta(v_i, v_i) = ca_i \neq 0.$$

Das ist unsere Strategie: Wir suchen ein-dimensionale β-reguläre Unterräume und spalten diese mit dem Zerlegungssatz ab.

Beweis. Wir führen den Beweis mittels vollständiger Induktion nach $n = \dim_K(V)$. Ist $n = 1$, dann ist nichts zu tun, denn jede (1×1)-Matrix ist diagonal.

Sei nun $n > 1$. Wir nehmen an, dass der Satz für Vektorräume der Dimension $n-1$ gilt. Ist $\beta = 0$ die Nullabbildung (d.h., $\beta(u,v) = 0$ für alle $u, v \in V$), dann nehmen wir irgendeine geordnete Basis $\mathcal{B}$ von V. Dann ist $\mathrm{M}_{\mathcal{B}}(\beta) = 0$ die Nullmatrix und damit insbesondere diagonal.

Wir nehmen nun an, dass $\beta \neq 0$ ist. Nach der Polarisationsformel ist β durch die quadratische Form $u \mapsto \beta(u,u)$ vollständig bestimmt. Es muss also einen Vektor $v_1 \in V$ geben, der $\beta(v_1, v_1) = a_1 \neq 0$ erfüllt. Insbesondere ist $v_1 \neq 0$ und spannt einen ein-dimensionalen Unterraum $\langle v_1 \rangle$ auf. Wie wir in 3.2.11 besprochen haben, ist

$\langle v_1 \rangle$ ein β-regulärer Unterraum. Der Zerlegungssatz 3.1.43 liefert uns eine Zerlegeung $V = \langle v_1 \rangle \oplus W$ wobei

$$W = \{w \in V \mid \beta(v_1, w) = 0\}$$

ist. (Bemerkung: Ist $\beta(v_1, w) = 0$, dann ist auch $\beta(cv_1, w) = 0$ für alle $c \in K$.) Es gilt insbesondere $\dim_K(W) = \dim_K(V) - \dim_K(\langle v_1 \rangle) = n - 1$. Wir betrachten nun die Einschränkung $\beta|_W$. Nach Induktionsvoraussetzung gibt es eine geordnete Basis $(v_2, \ldots, v_n)$ von W, sodass die Matrixdarstellung von $\beta|_W$ diagonal ist, d.h., $\beta(v_i, v_j) = 0$ für alle $i \neq j$ mit $i, j \geq 2$. Wir definieren $\mathcal{B} = (v_1, \ldots, v_n)$. Dies ist eine geordnete Basis von V; siehe 2.3.12. Da $v_2, \ldots, v_n$ in W liegen, gilt

$$0 = \beta(v_1, v_j) = \beta(v_j, v_1)$$

für alle $j \geq 2$. Also ist $\mathrm{M}_{\mathcal{B}}(\beta)$ eine Diagonalmatrix. □

Der Rang einer Diagonalmatrix ist genau die Anzahl der Diagonaleinträge ungleich Null. Mit diesem Wissen kann man aus dem Beweis des Diagonalisierungssatzes ein Verfahren ablesen, mit dem man eine Basis $\mathcal{B}$ bestimmen kann.

3.2.12 Diagonalisieren symmetrischer Bilinearformen.

`Gegeben:` symmetrische Bilinearform $\beta \in \mathrm{Bil}_K(V)$.

`Gesucht:` geordnete Basis $\mathcal{B}$ von V, sodass $\mathrm{M}_{\mathcal{B}}(\beta)$ diagonal ist.

`Verfahren:`

Berechne den Rang $r = \mathrm{Rg}(\beta)$. Setze $W_1 = V$.

- Für $i = 1, \ldots, r$ wiederhole folgenden Schritt: Wähle einen Vektor $v_i \in W_i$ mit $\beta(v_i, v_i) \neq 0$ und berechne

$$W_{i+1} = \{w \in W_i \mid \beta(v_i, w) = 0\}.$$

  Setze $a_i = \beta(v_i, v_i)$.

- Falls $r < n$ ist, gilt nun $\beta|_{W_{r+1}} = 0$. Wähle dann eine beliebige Basis $(v_{r+1}, \ldots, v_n)$ von W_{r+1}, setze $a_j = 0$ für alle $j > r$.

Nach Abschluss der Berechnungen ist $\mathcal{B} = (v_1, \ldots, v_n)$ eine Basis, sodass $\mathrm{M}_{\mathcal{B}}(\beta)$ eine Matrixdarstellung mit Diagonaleinträgen $a_1, \ldots, a_n$ ist.

Hinweis zur Vereinfachung: Ist $r = n$, dann ist $W_{n+1} = \{0\}$. Diese Berechnung kann man sich im Schritt $i = n$ also sparen.

3.2.13 Beispiel. Wir betrachten den Vektorraum $\mathbb{R}^4$ mit der symmetrischen Bilinearform $\beta(u,v) = u^T A v$ mit der Matrix

$$A = \begin{pmatrix} 3 & 3 & -2 & -1 \\ 3 & 0 & -1 & 1 \\ -2 & -1 & 1 & 0 \\ -1 & 1 & 0 & -1 \end{pmatrix} \in \mathrm{M}_{4,4}(\mathbb{R}).$$

Wir suchen eine diagonale Matrixdarstellung für β.

Zuerst berechnen wir den Rang von A. Dazu verwenden wir das Gauß-Verfahren. Wir tauschen die letzte Zeile nach oben und addieren geeignete Vielfache dieser Zeile zu jeder anderen. Dann sehen wir

$$A = \begin{pmatrix} 3 & 3 & -2 & -1 \\ 3 & 0 & -1 & 1 \\ -2 & -1 & 1 & 0 \\ -1 & 1 & 0 & -1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} -1 & 1 & 0 & -1 \\ 0 & 6 & -2 & -4 \\ 0 & 3 & -1 & -2 \\ 0 & -3 & 1 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} -1 & 1 & 0 & -1 \\ 0 & 3 & -1 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix};$$

d.h., A hat Rang 2 und nach 3.1.44 gilt $\mathrm{Rg}(\beta) = 2$.

Wir setzen $W_1 = \mathbb{R}^4$ und suchen einen Vektor $v_1 \in W_1$ mit $\beta(v_1, v_1) \neq 0$. Der vierte Standardbasisvektor e_4 hat diese Eigenschaft, denn es gilt $\beta(e_4, e_4) = -1$. Dann berechnen wir W_2 als

$$\begin{aligned} W_2 &= \{w \in \mathbb{R}^4 \mid \beta(e_4, w) = 0\} \\ &= \{w \in \mathbb{R}^4 \mid (-1, 1, 0, -1)w = 0\} \\ &= \{w \in \mathbb{R}^4 \mid -w_1 + w_2 - w_4 = 0\} \\ &= \left\langle \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle. \end{aligned}$$

Nun suchen wir einen Vektor $v_2 \in W_2$ mit $\beta(v_2, v_2) \neq 0$. Der dritte Standardbasisvektor e_3 hat diese Eigenschaft, denn $\beta(e_3, e_3) = 1$. Nun bestimmen wir W_3:

$$\begin{aligned} W_3 &= \{w \in W_2 \mid \beta(e_3, w) = 0\} \\ &= \{w \in \mathbb{R}^4 \mid -2w_1 - w_2 + w_3 = 0\} \cap W_2 \\ &= \{w \in \mathbb{R}^4 \mid -2w_1 - w_2 + w_3 = 0 \text{ und } -w_1 + w_2 - w_4 = 0\}; \end{aligned}$$

dabei verwenden wir, dass W_2 durch die Gleichung $-w_1 + w_2 - w_4 = 0$ beschrieben wird. In anderen Worten ist W_3 der Kern der Matrix $B = \begin{pmatrix} -1 & 1 & 0 & -1 \\ -2 & -1 & 1 & 0 \end{pmatrix} \in$ $\mathrm{M}_{2,4}(\mathbb{R})$. Mit dem Gauß-Verfahren finden wir eine Basis

$$W_3 = \mathrm{Ker}(B) = \left\langle \begin{pmatrix} -1 \\ 2 \\ 0 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 3 \\ 0 \end{pmatrix} \right\rangle.$$

Da $3 = \mathrm{Rg}(\beta) + 1$ ist, nennen wir die gefundenen Basisvektoren von W_3 nun v_3 und v_4. Dann ist $\mathcal{B} = (v_1, v_2, v_3, v_4)$ die gesuchte Basis. Wenn wir diese Basisvektoren als Spalten in eine Matrix S schreiben, d.h.,

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

dann ist S die Transformationsmatrix von $\mathcal{B}$ in die Standardbasis. Aus der Transformationsformel folgt dann

$$\mathrm{M}_{\mathcal{B}}(\beta) = S^T A S = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

3.2.14 **Aufgabe**. Gegeben ist die symmetrische Matrix $A = \begin{pmatrix} \overline{1} & \overline{2} & \overline{0} \\ \overline{2} & \overline{2} & \overline{1} \\ \overline{0} & \overline{1} & \overline{0} \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{F}_3)$. **L**

Bestimmen Sie eine invertierbare Matrix $S \in \mathrm{GL}_3(\mathbb{F}_3)$, sodass $S^T A S$ eine Diagonalmatrix ist.

3.2.15 **Bemerkung**. Es sei β ein Bilinearform und $\mathcal{B} = (v_1, \ldots, v_n)$ eine geordnete Basis, sodass

$$\mathrm{M}_{\mathcal{B}}(\beta) = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & a_n \end{pmatrix}$$

eine Diagonalmatrix ist. Weder die Basis $\mathcal{B}$ noch die Diagonalmatrix sind eindeutig! Zum einen kann man die Basisvektoren in einer anderen Reihenfolge anordnen, dann erhält man eine diagonale Matrixdarstellung in der die Diagonaleinträge entsprechend umsortiert wurden. Zum anderen kann man die Basisvektoren „strecken“, d.h., man setzt $v_i' = \lambda_i v_i$ mit $\lambda_i \in K^\times$. Dann ist auch $\mathcal{B}' = (v_1', \ldots, v_n')$ eine geodnete Basis. Es gilt $\beta(v_i', v_i') = \lambda_i^2 \beta(v_i, v_i) = \lambda_i^2 a_i$ und damit erhält man die Matrixdarstellung

$$\mathrm{M}_{\mathcal{B}'}(\beta) = \begin{pmatrix} \lambda_1^2 a_1 & 0 & \cdots & 0 \\ 0 & \lambda_2^2 a_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n^2 a_n \end{pmatrix}.$$

Sogar modulo Umsortieren und Strecken ist die Diagonalmatrix im Allgemeinen *nicht* eindeutig. Deshalb ist der Diagonalisierungssatz über vielen Körpern noch ziemlich weit von einer Klassifikation der symmetrischen Bilinearformen entfernt. Zum Beispiel ist die Klassifikation symmetrischen Bilinearformen über dem Körper $\mathbb{Q}$ ein berühmtes Ergebnis der klassischen Zahlentheorie: der Satz von Hasse[2]-Minkowski[3].

In einem algebraisch abgeschlossenen Körper K ist jedes Körperelement a ein Quadrat. Dadurch hat man bei Strecken der Basisvektoren so viel Flexibilität, dass man aus dem Diagonalisierungssatz tatsächlich eine Klassifikation der symmetrischen Bilinearformen ableiten kann. Wir erklären das für den Körper $\mathbb{C}$ der komplexen Zahlen.

3.2.16 **Satz.** *Es sei V ein $\mathbb{C}$-Vektorraum mit $\dim_{\mathbb{C}}(V) = n$. Ist β eine symmetrische Bilinearform auf V, dann gibt es eine geordnete Basis $\mathcal{B}$, sodass $\mathrm{M}_{\mathcal{B}}(\beta) = \mathrm{diag}(I_r, 0_{n-r})$ eine Diagonalmatrix mit Einträgen 1 und 0 ist. Dabei ist $r = \mathrm{Rg}(\beta)$.*

Eine Matrixdarstellung $\mathrm{M}_{\mathcal{B}}(\beta) = \mathrm{diag}(I_r, 0_{n-r})$ wie im Satz nennt man eine Darstellung in *Normalform*.

Beweis. Nach dem Diagonalisierungssatz finden wir eine geordnete Basis $\mathcal{B} = (v_1, \ldots, v_n)$ von V, sodass

$$\mathrm{M}_{\mathcal{B}}(\beta) = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & a_n \end{pmatrix}$$

[2]Helmut Hasse: deutscher Mathematiker, 1898–1979.
[3]Hermann Minkowski: deutscher Mathematiker, 1864–1909.

eine Diagonalmatrix ist. Durch Umordnen der Basis können wir annehmen, dass $a_1, \dots, a_r \neq 0$ und $a_{r+1} = a_{r+2} = \dots = a_n = 0$ gilt. Dabei ist r der Rang von $\mathrm{M}_{\mathcal{B}}(\beta)$ und nach Definition 3.1.44 ist $r = \mathrm{Rg}(\beta)$.

Aus Satz 1.5.32 wissen wir, dass die komplexe Zahl a_i^{-1} (für $i \leq r$) eine Quadratwurzel λ_i hat, d.h., $\lambda_i^2 = a_i^{-1}$. Wir definieren $v_i' = \lambda_i v_i$ für $i \leq r$ und $v_i' = v_i$ für $i > r$. Dann ist $\mathcal{B}' = (v_1', v_2', \dots, v_n')$ eine geordnete Basis von V und mit den Überlegungen aus 3.2.15 folgt $\mathrm{M}_{\mathcal{B}'}(\beta) = \mathrm{diag}(I_r, 0_{n-r})$. □

3.2.17 **Beispiel**. Wir betrachten die symmetrische Matrix

$$A = \begin{pmatrix} i & 1 \\ 1 & -i \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{C})$$

und die zugehörige Bilinearform $\beta(u, v) = u^T A v$ auf $\mathbb{C}^2$. Die Matrix A hat Rang 1, denn wenn man die erste Zeile mit $-i$ multipliziert, erhält man die zweite Zeile, d.h. die Zeilen sind linear abhängig.

Wir suchen zuerst eine Matrixdarstellung in Diagonalform. Dazu wählen wir $v_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, denn es gilt $\beta(v_1, v_1) = -i$. Dann bestimmen wir das Komplement

$$W_2 = \{w \in \mathbb{C}^2 \mid \beta(v_1, w) = 0\} = \{w \in \mathbb{C}^2 \mid w_1 - i w_2 = 0\} = \langle \begin{pmatrix} i \\ 1 \end{pmatrix} \rangle.$$

Wir setzen $v_2 = \begin{pmatrix} i \\ 1 \end{pmatrix}$. Mit der geordneten Basis $\mathcal{B} = (v_1, v_2)$ erhält man die Matrixdarstellung

$$\mathrm{M}_{\mathcal{B}}(\beta) = \begin{pmatrix} -i & 0 \\ 0 & 0 \end{pmatrix}.$$

Um eine Darstellung in Normalform zu erhalten, benötigen wir eine Quadratwurzel aus $\frac{1}{-i} = i$. In Beispiel 1.5.33 haben wir gesehen, dass $z = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} i$ eine Quadratwurzel von i ist. Wir definieren also $v_1' = z v_1$. Dann ist die Matrixdarstellung zur Basis $\mathcal{B}' = (v_1', v_2)$ in Normalform, d.h.,

$$\mathrm{M}_{\mathcal{B}'}(\beta) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Mit der Matrixdarstellung in Normalform können wir nun eine Klassifikation der symmetrischen Bilinearformen über $\mathbb{C}$ angeben.

3.2.18 **Klassifikationssatz für komplexe symmetrische Bilinearformen.** *Es sei V ein endlich-dimensionaler $\mathbb{C}$-Vektorraum. Je zwei symmetrische Bilinearformen sind genau dann kongruent, wenn sie denselben Rang haben.*

Beweis. Sind zwei symmetrische Bilinearformen kongruent, dann haben sie dieselben Matrixdarstellungen und damit denselben Rang (der Rang wurde ja als Rang einer Matrixdarstellung definiert).

Es seien nun β_1 und β_2 symmetrische Bilinearformen auf V mit demselben Rang r. Dann gibt es nach Satz 3.2.16 geordnete Basen $\mathcal{B}_1$ und $\mathcal{B}_2$ von V dergestalt, dass

$$\mathrm{M}_{\mathcal{B}_1}(\beta_1) = \mathrm{diag}(I_r, 0_{n-r}) = \mathrm{M}_{\mathcal{B}_2}(\beta_2)$$

ist mit $n = \dim_{\mathbb{C}}(V)$. Also sind β_1 und β_2 kongruent. □

Aus diesem Klassifikationssatz kann man direkt den folgenden Satz über die Kongruenzklassen symmetrischer komplexer Matrizen ableiten (Übungsaufgabe!). **L**

3.2.19 **Klassifikationssatz für symmetrische komplexe Matrizen.** *Sei $n \in \mathbb{N}$. Je zwei symmetrische Matrizen in $\mathrm{M}_{n,n}(\mathbb{C})$ sind genau dann kongruent, wenn sie denselben Rang haben.*

In anderen Worten ist die Kongruenzrelation auf der Menge der komplexen symmetrischen Matrizen genau die Faserrelation der Rangabbildung.

III. Reelle symmetrische Bilinearformen

In diesem Abschnitt wollen wir einen Klassifikationssatz für reelle symmetrische Bilinearformen herleiten. Wir beginnen mit einer wichtigen Definition.

3.2.20 **Definition** Sei V ein reeller Vektorraum. Eine symmetrische Bilinearform β auf V heißt

- *positiv definit*, wenn $\beta(v, v) > 0$ für alle $v \neq 0$ in V gilt und
- *negativ definit*, wenn $\beta(v, v) < 0$ für alle $v \neq 0$ in V gilt.

Ist β weder positiv noch negativ definit, dann nennen wir β *indefinit*. Eine positiv definite symmetrische Bilinearform auf einem reellen Vektorraum nennt man auch *Skalarprodukt*.

3.2.21 **Beispiel**. (a) Auf dem $\mathbb{R}^n$ ist die reelle symmetrische Bilinearform

$$\beta\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}\right) := \sum_{i=1}^{n} x_i y_i$$

positiv definit. Um das zu sehen berechnen wir $\beta(x,x)$ und erhalten

$$\beta\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) := \sum_{i=1}^{n} x_i^2.$$

Jeder Summand x_i^2 ist dabei nicht negativ. Ist $x \neq 0$, dann ist $x_i \neq 0$ für mindestens ein i und damit gilt $\beta(x,x) > 0$. Diese Bilinearform nennt man auch das *Standardskalarprodukt* auf $\mathbb{R}^n$.

(b) Ist β eine positiv definite symmetrische Bilinearform auf V, dann ist $-\beta$ negativ definit.

(c) Die symmetrische Bilinearform β auf $\mathbb{R}^2$ mit

$$\beta\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) := x_1 y_1 - x_2 y_2$$

ist indefinit, denn $\beta(e_1, e_1) > 0$ und $\beta(e_2, e_2) < 0$.

3.2.22 **Lemma.** *Jede positiv (oder negativ) definite symmetrische Bilinearform ist regulär.*

Beweis. Es sei β eine positiv (oder negativ) definite symmetrische Bilinearform auf V. Sei $u \neq 0$ in V. Dann ist $\beta(u,u)$ positiv (bzw. negativ) und damit ungleich 0; also ist β regulär; siehe 3.1.37. □

3.2.23 **Beispiel**. Auch indefinite Bilinearformen können regulär sein. Betrachten wir noch mal die indefinite symmetrische Bilinearform β auf $\mathbb{R}^2$ mit

$$\beta\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) := x_1 y_1 - x_2 y_2.$$

Die Matrixdarstellung von β zu Standardbasis ist $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Diese Matrix ist invertierbar, also ist β regulär.

Der folgende *Trägheitssatz* geht auf J. J. Sylvester[4] zurück. Es ist der wesentliche Schritt hin zur Klassifikation der reellen symmetrischen Bilinearformen. Der Satz besagt, dass die Vorzeichen der Diagonaleinträge einer diagonalen Matrixdarstellung nicht von der gewählten Basis abhängen. Die Zahl der positiven bzw. negativen Einträge ist also „träge" und verändert sich nicht beim Basiswechsel. Es handelt sich also um Eigenschaften der Bilinearform und nicht nur um Eigenschaften einer Matrixdarstellung.

3.2.24 **Sylvester'scher Trägheitssatz.** *Sei V ein $\mathbb{R}$-Vektorraum der Dimension n und sei β eine symmetrische Bilinearform mit $r = \mathrm{Rg}(\beta)$. Dann gibt es Zahlen $r_+, r_- \in \mathbb{N}_0$ mit $r = r_+ + r_-$, sodass in jeder diagonalen Matrixdarstellung $\mathrm{M}_{\mathcal{B}}(\beta)$ genau r_+ positive, r_- negative und $n-r$ verschwindende Diagonaleinträge vorkommen.*

3.2.25 **Definition** Das Tripel $(r_+, r_-, n-r)$ nennt man auch den *Typ*, die Differenz $r_+ - r_-$ nennt man die *Signatur* der Bilinearform β.

Beweis des Trägheitssatzes. Es sei $\mathcal{B} = (v_1, \dots, v_n)$ eine geordnete Basis, sodass $\mathrm{M}_{\mathcal{B}}(\beta)$ eine Diagonalmatrix ist. Es sei $J_+ \subseteq \{1, \dots, n\}$ (bzw. J_-) die Menge aller i, sodass $\beta(v_i, v_i) > 0$ (bzw. $\beta(v_i, v_i) < 0$). Wir setzen $J_0 = \{1, \dots, n\} \setminus (J_+ \cup J_-)$. Für $i \in J_0$ gilt also $\beta(v_i, v_i) = 0$.

Für $* \in \{+, -, 0\}$ definieren wir

$$V_* = \langle \{v_i \mid i \in J_*\} \rangle.$$

Dann zerlegt sich V als direkte Summe

$$V = V_+ \oplus V_- \oplus V_0;$$

siehe Aufgabe 2.3.11. Wir setzen $r_+ = \dim_{\mathbb{R}} V_+$, $r_- = \dim_{\mathbb{R}} V_-$. Nach Definition von $V_\pm$ ist $r_\pm$ genau die Anzahl der positiven bzw. negativen Diagonaleinträge von $\mathrm{M}_{\mathcal{B}}(\beta)$.

Die Einschränkung von β auf V_+ ist positiv definit. Das sieht man so: Für $v = \sum_{i \in J_+} c_i v_i$ gilt

$$\begin{aligned} \beta(v, v) &= \sum_{i \in J_+} \sum_{j \in J_+} c_i c_j \beta(v_i, v_j) \\ &= \sum_{i \in J_+} c_i^2 \underbrace{\beta(v_i, v_i)}_{>0}. \qquad (\beta(v_i, v_j) = 0 \text{ für } i \neq j) \end{aligned}$$

[4] James Joseph SYLVESTER: britischer Mathematiker, 1814–1897.

Ist nun $v \neq 0$, dann ist $c_i \neq 0$ für einen Index i und damit ist der letzte Ausdruck positiv. Genauso sieht man, dass die Einschränkung von β auf V_- negativ definit ist.

Wir behaupten, dass $V_0 = N_\beta$ der Nullraum von β ist. Sei $j \in J_0$. Sei $w \in V$ beliebig. Wir schreiben $w = \sum_{i=1}^n c_i v_i$ als Linearkombination in der Basis $\mathcal{B}$, dann gilt

$$\beta(v_j, w) = \sum_{i=1}^n c_i \underbrace{\beta(v_j, v_i)}_{=0} = 0.$$

Da w beliebig war, schließen wir $v_j \in N_\beta$; also liegt auch die lineare Hülle V_0 der Vektoren $\{v_j \mid j \in J_0\}$ in N_β.

Sei umgekehrt $v \in N_\beta$. Wir schreiben $v = v_+ + v_- + v_0$ mit $v_+ \in V_+$, $v_- \in V_-$ und $v_0 \in V_0$. Da v im Nullraum liegt, ist

$$0 = \beta(v, v_+) = \beta(v_+, v_+)$$

und da $\beta|_{V_+}$ positiv definit ist, muss $v_+ = 0$ sein. Genauso folgt $v_- = 0$. Das heißt, $v = v_0$ liegt in V_0. Damit haben wir $V_0 = N_\beta$ gezeigt.

Aus Satz 2.3.12 und der Dimensionsformel 3.1.47 erhalten wir

$$r_+ + r_- + \dim_{\mathbb{R}} N_\beta = \dim_{\mathbb{R}}(V) = r + \dim_{\mathbb{R}} N_\beta$$

und damit $r_+ + r_- = r$.

Ist nun $\mathcal{B}'$ eine andere Basis, sodass $\mathrm{M}_{\mathcal{B}'}(\beta)$ diagonal ist, dann erhalten wir eine weitere Zerlegung

$$V = V'_+ \oplus V'_- \oplus N_\beta.$$

Dabei ist $\beta|_{V'_+}$ positiv und $\beta|_{V'_-}$ negativ definit. Für alle Vektoren $v = v_- + n \in V_- \oplus N_\beta$ gilt

$$\beta(v, v) = \underbrace{\beta(v_-, v_-)}_{\leq 0} + 2\underbrace{\beta(v_-, n)}_{=0} + \underbrace{\beta(n, n)}_{=0} \leq 0.$$

Daraus folgt $V'_+ \cap (V_- \oplus N_\beta) = \{0\}$ und damit $\dim_{\mathbb{R}} V'_+ + r_- + \dim_{\mathbb{R}} N_\beta \leq \dim_{\mathbb{R}}(V)$; siehe [MG, 7.4.6]. Wir erhalten also $\dim_{\mathbb{R}} V'_+ \leq r_+$. Dasselbe Argument mit vertauschten Rollen liefert aber auch $r_+ \leq \dim_{\mathbb{R}} V'_+$ und damit die Gleichheit $r_+ = \dim_{\mathbb{R}} V'_+$.

Ganz analog verfährt man für r_-, indem man nachweist, dass alle Vektoren $v \in V_+ \oplus N_\beta$ die Eigenschaft $\beta(v, v) \geq 0$ haben. □

3.2.26 **Beispiel**. Die symmetrische Bilinearform $\beta(u,v) = u^T A v$ auf $\mathbb{R}^5$ mit

$$A = \begin{pmatrix} -6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{3} & 0 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix} \in \mathrm{M}_{5,5}(\mathbb{R})$$

hat den Typ $(1,3,1)$, denn A ist eine diagonale Matrixdarstellung von β mit einem positiven, drei negativen und einem verschwindenden Diagonaleintrag. Die Signatur von β ist also $1-3=-2$.

3.2.27 **Aufgabe**. Für jedes $t \in \mathbb{R}$ sei **L**

$$A_t = \begin{pmatrix} t & 2 \\ 2 & -1 \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{R})$$

und es sei $\beta_t(u,v) = u^T A_t v$ die zugehörige symmetrische Bilinearform. Bestimmen Sie den Typ von β_t in Abhängigkeit von $t \in \mathbb{R}$.

3.2.28 **Bemerkung**. Aus dem Beweis des Trägheitssatzes sehen wir, dass wir wichtige Eigenschaften einer reellen symmetrischen Bilinearform direkt am Typ ablesen können. Sei β eine reelle symmetrische Bilinerform vom Typ (r_+, r_-, k) auf einem Vektorraum der Dimension n. Dann gilt

- β ist regulär $\Leftrightarrow k = 0$
- β ist ausgeartet $\Leftrightarrow k \neq 0$
- β ist positiv definit $\Leftrightarrow r_+ = n, r_- = 0, k = 0$
- β ist negativ definit $\Leftrightarrow r_+ = 0, r_- = n, k = 0$

3.2.29 **Korollar.** *Sei V ein endlich-dimensionaler $\mathbb{R}$-Vektorraum. Ist $\beta \in \mathrm{Bil}_K(V)$ eine symmetrische Bilinearform vom Typ (r_+, r_-, k), dann gibt es eine geordnete Basis $\mathcal{B}$ von V, sodass $\mathrm{M}_\mathcal{B}(\beta)$ eine Blockdiagonalmatrix der Form*

$$\mathrm{M}_\mathcal{B}(\beta) = \mathrm{diag}(I_{r_+}, -I_{r_-}, 0_k) = \left(\begin{array}{c|c|c} I_{r_+} & 0 & 0 \\ \hline 0 & -I_{r_-} & 0 \\ \hline 0 & 0 & 0 \end{array}\right)$$

ist.

Eine Matrixdarstellung einer symmetrischen reellen Bilinearform β in der Form $\operatorname{diag}(I_{r_+}, -I_{r_-}, 0_k)$ nennen wir eine Darstellung in *Normalform.*

Beweis. Es sei $r = \operatorname{Rg}(\beta) = r_+ + r_-$. Mit dem Diagonalisierungssatz 3.2.10 finden wir eine geordnete Basis $\mathcal{B}' = (v_1', \ldots, v_n')$, sodass $\mathrm{M}_{\mathcal{B}'}(\beta)$ eine Diagonalmatrix ist. Dabei gibt es in $\mathrm{M}_{\mathcal{B}'}(\beta)$ genau r_+ positive, r_- negative und $k = n - r$ verschwindende Diagonaleinträge. Nach Umordnung der Basisvektoren können wir annehmen, dass $a_i = \beta(v_i', v_i')$ positiv ist für $i \leq r_+$, negativ ist für $r_+ < i \leq r$ und Null für $i > r$. Für $i \leq r$ definieren wir

$$v_i = \frac{1}{\sqrt{|a_i|}} v_i'.$$

Dann gilt $\beta(v_i, v_i) = \frac{1}{\sqrt{|a_i|}^2} \beta(v_i', v_i') = \frac{a_i}{|a_i|} = \pm 1$ (das Vorzeichen hängt davon ab, ob a_i positiv oder negativ ist) für alle $i \leq r$. Zusätzlich setzen wir $v_i = v_i'$ für alle $i > r$. Dann ist $\mathcal{B} = (v_1, \ldots, v_n)$ eine geordnete Basis von V und nach Konstruktion gilt

$$\mathrm{M}_{\mathcal{B}}(\beta) = \operatorname{diag}(I_{r_+}, -I_{r_-}, 0_k).$$ □

3.2.30 **Bemerkung**. Aus dem Beweis 3.2.29 ergibt sich unmittelbar ein Verfahren, wie man eine geeignete Basis $\mathcal{B}$ mit $\mathrm{M}_{\mathcal{B}}(\beta) = \operatorname{diag}(I_{r_+}, -I_{r_-}, 0_k)$ berechnen kann. Zuerst bestimmen wir wie in 3.2.12 eine Basis $\mathcal{B}' = (v_1', \ldots, v_n')$, sodass $\mathrm{M}_{\mathcal{B}'}(\beta)$ diagonal ist. Dann sortieren wir die Basisvektoren nach den Vorzeichen der Werte $a_i = \beta(v_i', v_i')$ – also zuerst positiv, dann negativ und dann Null – und multiplizieren die Basisvektoren v_i' mit dem Faktor $\frac{1}{\sqrt{|a_i|}}$, falls $a_i \neq 0$ ist.

3.2.31 **Beispiel**. Gegeben ist die reelle symmetrische Matrix

$$A = \begin{pmatrix} -3 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{R}).$$

Es sei $\beta(u, v) = u^T A v$ die zugehörige Bilinearform auf dem Vektorraum $\mathbb{R}^3$. Wir wollen eine geordnete Basis finden, sodass $\mathrm{M}_{\mathcal{B}}(\beta)$ in Normalform ist.

Mit einer kurzen Rechnung sieht man, dass A den Rang $r = 2$ hat (die zweite und dritte Zeile sind linear abhängig). Wir suchen einen Vektor $v_1' \in \mathbb{R}^3$ mit $\beta(v_1', v_1') \neq 0$. Wir wählen den zweiten Standardbasisvektor $v_1' = e_2$, denn es gilt

$$\beta(e_2, e_2) = 4 \neq 0.$$

Dann berechnen wir das Komplement:

$$W_2 = \{w \in \mathbb{R}^3 \mid \beta(v_1', w) = 0\}$$

$$
\begin{aligned}
&= \operatorname{Ker}((v_1')^T A) = \operatorname{Ker}\begin{pmatrix} 2 & 4 & 2 \end{pmatrix} \\
&= \left\langle \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix} \right\rangle .
\end{aligned}
$$

Dann suchen wir $v_2' \in W_2$ mit $\beta(v_2', v_2') \neq 0$. Wir wählen $v_2' = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}$ und finden

$$
\beta(v_2', v_2') = (v_2')^T A v_2' = \begin{pmatrix} -8 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} = -16.
$$

Wir bestimmen wieder das Komplement:

$$
\begin{aligned}
W_3 &= \{ w \in W_2 \mid \beta(v_2', w) = 0 \} \\
&= W_2 \cap \operatorname{Ker}((v_2')^T A) = W_2 \cap \operatorname{Ker}\begin{pmatrix} -8 & 0 & 0 \end{pmatrix} \\
&= \left\langle \begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix} \right\rangle .
\end{aligned}
$$

Wir setzen $v_3' = \begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix}$. Da A den Rang 2 hat, ist $\beta(v_3', v_3') = 0$. Die Matrixdarstellung von β zur Basis $\mathcal{B}' = (v_1', v_2', v_3')$ ist diagonal mit den Diagonaleinträgen $4, -16, 0$. Die Reihenfolge der Vorzeichen ist schon passend, wir müssen die Basisvektoren nur entsprechend strecken. Wir definieren

$$
\begin{aligned}
v_1 &= \frac{1}{\sqrt{4}} v_1' = \begin{pmatrix} 0 \\ \frac{1}{2} \\ 0 \end{pmatrix}, \\
v_2 &= \frac{1}{\sqrt{16}} v_2' = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{4} \\ 0 \end{pmatrix}, \\
v_3 &= v_3'.
\end{aligned}
$$

Mit der geordneten Basis $\mathcal{B} = (v_1, v_2, v_3)$ gilt also $\mathrm{M}_{\mathcal{B}}(\beta) = \operatorname{diag}(1, -1, 0)$ und dies ist eine Darstellung in Normalform. Damit ist β eine Bilinearform vom Typ $(1, 1, 1)$ und mit der Signatur $1 - 1 = 0$.

3.2.32 **Aufgabe**. Wir betrachten wieder die Familie β_t von Bilinearformen auf $\mathbb{R}^2$ aus Aufgabe 3.2.27 mit der Matrixdarstellung **L**

$$A_t = \begin{pmatrix} t & 2 \\ 2 & -1 \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{R})$$

abhängig von $t \in \mathbb{R}$. Bestimmen Sie für jedes $t \in \mathbb{R}$ eine geordnete Basis $\mathcal{B}$, sodass $\mathrm{M}_{\mathcal{B}}(\beta)$ in Normalform ist.

Zum Abschluss können wir die symmetrischen reellen Bilinearformen klassifizieren.

3.2.33 **Klassifikationssatz für reelle symmetrische Bilinearformen.** *Sei V ein endlich-dimensionaler reeller Vektorraum. Je zwei symmetrische Bilinearformen auf V sind genau dann kongruent, wenn sie denselben Typ haben.*

Beweis. Es seien β_1, β_2 zwei symmetrische reelle Bilinearformen auf V. Sind β_1 und β_2 kongruent, dann haben sie dieselben Matrixdarstellungen (das haben wir im Beweis von Lemma 3.1.23 gesehen). Da man den Typ an jeder diagonalen Matrixdarstellung ablesen kann, haben β_1 und β_2 denselben Typ.

Nehmen wir umgekehrt an, dass β_1 und β_2 denselben Typ (r_+, r_-, k) haben. Dann gibt es geordnete Basen $\mathcal{B}_1$ und $\mathcal{B}_2$, sodass die Matrixdarstellungen in Normalform sind. Das heißt, es gilt

$$\mathrm{M}_{\mathcal{B}_1}(\beta_1) = \mathrm{diag}(I_{r_+}, -I_{r_-}, 0_k) = \mathrm{M}_{\mathcal{B}_2}(\beta_2)$$

und damit sind β_1 und β_2 kongruent. □

Ist $A \in \mathrm{M}_{n,n}(\mathbb{R})$ eine reelle symmetrische Matrix, dann definieren wir den Typ von A als den Typ der zugehörigen Bilinearform $\beta(u,v) = u^T A v$. Damit können wir aus dem Klassifikationssatz für reelle symmetrische Bilinearformen auch einen Klassifikationssatz für reelle symmetrische Matrizen bis auf Kongruenz ableiten.

3.2.34 **Klassifikationssatz für reelle symmetrische Matrizen.** *Sei $n \in \mathbb{N}$. Je zwei symmetrische Matrizen in $\mathrm{M}_{n,n}(\mathbb{R})$ sind genau dann kongruent, wenn sie denselben Typ haben.*

3.2.35 **Aufgabe**. Wieviele Kongruenzklassen reeller symmetrischer (4×4)-Matrizen gibt es? **L**

3.3. Alternierende Bilinearformen

In diesem Abschnitt lernen wir die *alternierenden Bilinearformen* kennen. Wir werden sehen, dass man auch in diesem Fall relativ einfach eine Klassifikation herleiten kann.

3.3.1 **Definition** Sei V ein K-Vektorraum und sei $\beta \in \mathrm{Bil}_K(V)$. Wir nennen β...

- *schiefsymmetrisch*, wenn $\beta(u,v) = -\beta(v,u)$ für alle $u, v \in V$ gilt.
- *alternierend*, wenn $\beta(u,u) = 0$ für alle $u \in V$ gilt.

3.3.2 **Anmerkung zur Definition**. (a) *Jede alternierende Bilinearform ist auch schiefsymmetrisch:*
Angenommen β ist alternierend. Für alle $u, v \in V$ gilt dann

$$0 = \beta(u+v, u+v) = \underbrace{\beta(u,u)}_{=0} + \beta(u,v) + \beta(v,u) + \underbrace{\beta(v,v)}_{=0} = \beta(u,v) + \beta(v,u)$$

und damit $\beta(u,v) = -\beta(v,u)$.

(b) *Ist $2 \neq 0$ im Körper K, dann ist auch jede schiefsymmetrische Bilinearform alternierend:*
Sei β schiefsymmetrisch. Dann gilt für alle $u \in V$

$$\beta(u,u) = -\beta(u,u)$$

und damit $2\beta(u,u) = 0$. Ist nun $2 \neq 0$, dann folgt daraus $\beta(u,u) = 0$.

(c) In einem Körper der Charakteristik 2 ist $-1 = 1$ und damit ist in diesem Fall jede symmetrische Bilinearform auch schiefsymmetrisch.

Über den meisten Körpern sind schiefsymmetrische und alternierende Formen also dasselbe. Wir werden uns hauptsächlich die alternierenden Formen anschauen, weil sich diese auch über Körpern der Charakteristik 2 gut verhalten.

3.3.3 **Beispiel**. Sei K ein Körper. Die Bilinearform $\beta \colon K^2 \times K^2 \to K$ mit

$$\beta\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = x_1 y_2 - x_2 y_1$$

ist alternierend, denn offensichtlich ist $\beta(x,x) = 0$ für alle $x \in K^2$. Die Matrixdarstellung von β zur Standardbasis $\mathcal{B}$ ist

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

3.3.4 **Aufgabe**. Zeigen Sie: Eine Bilinearform β ist genau dann schiefsymmetrisch, wenn $\beta^{(1)} = -\beta^{(2)}$ gilt. **L**

Man kann natürlich anhand einer Matrixdarstellung entscheiden, ob die Bilinearform β alternierend bzw. schiefsymmetrisch ist. Dazu benötigen wir den folgenden Begriff.

3.3.5 **Definition** Eine quadratische Matrix $A = (a_{ij})_{i,j} \in \mathrm{M}_{n,n}(K)$ nennt man *schiefsymmetrisch*, wenn

$$A^T = -A$$

gilt. Das heißt, es gilt $a_{ij} = -a_{ji}$ für alle $i, j \in \{1, \ldots, n\}$. Man nennt A *alternierend*, wenn A schiefsymmetrisch ist und $a_{ii} = 0$ für alle $i \in \{1, \ldots, n\}$ gilt.

3.3.6 **Beispiel**. (a) Die Matrix

$$\begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & 3 \\ -2 & -3 & 0 \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{Q})$$

ist alternierend.

(b) Die Matrix

$$\begin{pmatrix} \overline{1} & \overline{1} & \overline{1} \\ \overline{1} & \overline{1} & \overline{1} \\ \overline{1} & \overline{1} & \overline{1} \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{F}_2)$$

ist symmetrisch und schiefsymmetrisch, aber nicht alternierend.

3.3.7 **Lemma.** *Sei V ein endlich-dimensionaler K-Vektorraum mit einer geordneten Basis $\mathcal{B}$. Eine Bilinearform $\beta \in \mathrm{Bil}_K(V)$ ist genau dann alternierend, wenn die Matrixdarstellung $\mathrm{M}_{\mathcal{B}}(\beta)$ alternierend ist.*

Beweis. Es sei $\mathcal{B} = (v_1, \ldots, v_n)$.

Nehmen wir zunächst an, dass β alternierend ist. Dann gilt $\beta(v_i, v_i) = 0$ für alle i, denn β ist alternierend. Da β auch schiefsymmetrisch ist, gilt außerdem $\beta(v_i, v_j) = -\beta(v_j, v_i)$ für alle $i, j \in \{1, \ldots, n\}$. Also ist $\mathrm{M}_{\mathcal{B}}(\beta)$ eine alternierende Matrix.

Umgekehrt sei $A = \mathrm{M}_{\mathcal{B}}(\beta)$ alternierend. Sei $u \in V$. Wir schreiben $u = \sum_{j=1}^n c_j v_j$ in der Basis $\mathcal{B}$ mit $c_j \in K$. Da A alternierend ist, erhalten wir

$$\beta(u, u) = \sum_{i,j=1}^{n} c_i c_j \, \beta(v_i, v_j)$$

$$= \sum_{\substack{i,j=1 \\ i\neq j}}^{n} c_i c_j \, \beta(v_i, v_j) \qquad\qquad (\beta(v_i, v_i) = 0)$$

Nach Voraussetzung gilt $\beta(v_i, v_j) = -\beta(v_j, v_i)$ für alle $i \neq j$. Damit kommt jeder Term genau zweimal mit wechselnden Vorzeichen in der Summe vor; d.h., $\beta(u, u) = 0$ und β ist alternierend. □

Wir möchten nun die alternierenden Formen klassifizieren. Bei den symmetrischen Bilinearformen haben wir den Diagonalisierungssatz bewiesen indem wir Vektoren $v \in V$ mit $\beta(v, v) \neq 0$ gesucht haben. Der Unterraum $\langle v\rangle$ war dann regulär und wir konnten ihn mit dem Zerlegungssatz abspalten. Bei alternierenden Bilinearformen gilt $\beta(v, v) = 0$ für alle Vektoren v. Der Ansatz für symmetrische Bilinearformen kann also nicht funktionieren.

Es kann also keine ein-dimensionalen regulären Unterräume geben. Aber was ist mit den zwei-dimensionalen Unterräumen? Das folgende Lemma sagt uns, wie die zwei-dimensionalen Unterräume aussehen.

3.3.8 **Lemma.** *Sei V ein zwei-dimensionaler K-Vektorraum und sei β eine alternierende Form auf V. Ist $\beta \neq 0$, dann gibt es eine geordnete Basis $\mathcal{B}$ von V mit*

$$\mathrm{M}_{\mathcal{B}}(\beta) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Insbesondere ist β regulär.

Beweis. Ist $\beta \neq 0$, dann finden wir zwei Vektoren $v_1, v_2 \in V$ mit $\beta(v_1, v_2) = c \neq 0$. Wir können annehmen, dass $c = 1$ gilt, indem wir v_2 durch $\frac{1}{c}v_2$ ersetzen. Da $\beta(v_1, v_2) \neq 0$ ist, sind v_1, v_2 linear unabhängig. In der Tat, wäre $v_2 = \lambda v_1$ für ein $\lambda \in K$, dann wäre $\beta(v_1, v_2) = \lambda\beta(v_1, v_1) = 0$, weil β alternierend ist.

Wir setzen $\mathcal{B} = (v_1, v_2)$. Die Matrixdarstellung von β zur Basis $\mathcal{B}$ ist dann

$$\mathrm{M}_{\mathcal{B}}(\beta) = \begin{pmatrix} \beta(v_1, v_1) & \beta(v_1, v_2) \\ -\beta(v_1, v_2) & \beta(v_2, v_2) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Die Matrix ist invertierbar, also ist β regulär; siehe 3.1.38. □

In anderen Worten gibt es auf einem zwei-dimensionalen Raum genau zwei Kongruenzklassen von alternierenden Bilinearformen: die Nullform und die Form zu Matrixdarstellung $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Mit dieser Vorüberlegung können wir beweisen, dass sich jede alternierende Bilinearform aus diesen zwei-dimensionalen Formen „zusammensetzt“.

3.3.9 **Satz.** *Sei V ein K-Vektorraum der Dimension $n \geq 1$ mit einer alternierenden Bilinearform β. Dann gibt es eine geordnete Basis $\mathcal{B} = (u_1, v_1, u_2, v_2, \ldots, u_k, v_k, v_{2k+1}, \ldots, v_n)$ von V, sodass $\mathrm{M}_{\mathcal{B}}(\beta) \in \mathrm{M}_{n,n}(K)$ eine Blockdiagonalmatrix der Form*

$$\mathrm{M}_{\mathcal{B}}(\beta) = \mathrm{diag}(\underbrace{J, \ldots, J}_{k}, 0_{n-2k}) = \left(\begin{array}{c|c|c|c|c} J & 0 & 0 & 0 & 0 \\ \hline 0 & J & 0 & 0 & 0 \\ \hline 0 & 0 & \ddots & 0 & 0 \\ \hline 0 & 0 & 0 & J & 0 \\ \hline 0 & 0 & 0 & 0 & 0_{n-2k} \end{array}\right)$$

ist.

Eine Matrixdarstellung wie im Satz nennen wir eine Darstellung in *Normalform*.

Beweis. Wir argumentieren mit vollständiger Induktion über die Dimension n von V. Für $n = 1$ ist $\beta = 0$ und wir nehmen irgendeine Basis. Die Matrixdarstellung ist in diesem Fall die Nullmatrix.

Sei nun $n \geq 2$. Ist $\beta = 0$, dann nehmen wir wieder irgendeine geordnete Basis. Wir können also annehmen, dass $\beta \neq 0$ ist; es gibt also Vektoren $u_1, v_1 \in V$ mit $\beta(u_1, v_1) \neq 0$. Nach Lemma 3.3.8 ist der Unterraum $U = \langle u_1, v_1 \rangle$ β-regulär und wir können annehmen, dass $\beta|_U$ die Matrixdarstellung J zur Basis u_1, v_1 von U hat.

Da U ein β-regulärer Unterraum ist, folgt aus dem Zerlegungssatz $V = U \oplus W$ mit $W = \{w \in V \mid \beta(u, w) = 0 \text{ für alle } u \in U\}$. Die Einschränkung von β auf W ist eine alternierende Form. Nach Induktionsvoraussetzung finden wir eine geordnete Basis $\mathcal{B}' = (u_2, v_2, \ldots, u_k, v_k, v_{2k+1}, \ldots, v_n)$ von W wie im Satz. Wir definieren $\mathcal{B} = (u_1, v_1, u_2, v_2, \ldots, u_k, v_k, v_{2k+1}, \ldots, v_n)$. Das ist eine geordnete Basis von V und da $\beta(u, w) = 0 = \beta(w, u)$ für alle $u \in U$ und $w \in W$ gilt, hat $\mathrm{M}_{\mathcal{B}}(\beta)$ die geforderte Gestalt. □

3.3.10 **Rang alternierender Bilinearformen**. Aus der Matrixdarstellung in Normalform kann man ablesen, dass der Rang von β genau $\mathrm{Rg}(\beta) = r = 2k$ ist. Der Rang $r = 2k$ einer alternierenden Bilinearform ist also immer gerade.

3.3.11 **Normalform alternierender Bilinearformen berechnen**.

`Gegeben`: alternierende Bilinearform $\beta \in \mathrm{Bil}_K(V)$.

`Gesucht`: geordnete Basis $\mathcal{B}$ von V, sodass $\mathrm{M}_{\mathcal{B}}(\beta)$ in Normalform ist.

`Verfahren`:

Berechne den Rang $r = 2k = \mathrm{Rg}(\beta)$. Setze $W_1 = V$.

- Für $i = 1, \dots, k$ wiederhole folgenden Schritt: Finde Vektoren $u_i, v_i' \in W_i$ mit $a_i = \beta(u_i, v_i') \neq 0$ und berechne

$$W_{i+1} = \{w \in W_i \mid \beta(u_i, w) = 0 \text{ und } \beta(v_i', w) = 0\}.$$

Setze $v_i = \frac{1}{a_i} v_i'$.

- Falls $2k < n$ ist, gilt nun $\beta|_{W_{k+1}} = 0$. Wähle dann eine beliebige Basis $(v_{r+1}, \dots, v_n)$ von W_{k+1}.

Nach Abschluss der Berechnungen ist $\mathcal{B} = (u_1, v_1, u_2, v_2, \dots, u_k, v_k, v_{r+1}, \dots, v_n)$ eine Basis, sodass $\mathrm{M}_{\mathcal{B}}(\beta)$ eine Matrixdarstellung in Normalform ist.

3.3.12 **Klassifikationssatz für alternierende Bilinearformen.** *Sei V ein endlich-dimensionaler K-Vektorraum. Je zwei alternierende Bilinearformen auf V sind genau dann kongruent, wenn sie denselben Rang haben.*

Beweis. Es seien β_1, β_2 zwei alternierende Bilinearformen auf V. Sind β_1 und β_2 kongruent, dann haben sie dieselben Matrixdarstellungen und damit denselben Rang.

Nehmen wir umgekehrt an, dass β_1 und β_2 denselben Rang $r = 2k$ haben. Dann gibt es geordnete Basen $\mathcal{B}_1$ und $\mathcal{B}_2$, sodass die Matrixdarstellungen in Normalform sind. Das heißt, es gilt

$$\mathrm{M}_{\mathcal{B}_1}(\beta_1) = \mathrm{diag}(\underbrace{J, \dots, J}_{k}, 0_{n-2k}) = \mathrm{M}_{\mathcal{B}_2}(\beta_2)$$

und damit sind β_1 und β_2 kongruent. □

Aus dem Klassifikationssatz für alternierende Bilinearformen kann man auch einen Klassifikationssatz für alternierende Matrizen bis auf Kongruenz ableiten.

3.3.13 **Klassifikationssatz für alternierende Matrizen.** *Sei $n \in \mathbb{N}$. Je zwei alternierende Matrizen in $\mathrm{M}_{n,n}(K)$ sind genau dann kongruent, wenn sie denselben Rang haben.*

3.3.14 **Aufgabe**. Wieviele Kongruenzklassen alternierender (6×6)-Matrizen gibt es? L

3.4. Hermite'sche Formen

In diesem Abschnitt betrachten wir nur den Körper $\mathbb{C}$ der komplexen Zahlen. Auf komplexen Vektorräumen werden wir hier sogenannte *Hermite'sche Formen* definieren und untersuchen. Der Name wurde zu Ehren des französischen Mathematikers Charles HERMITE (1822-1901) gewählt. Bei diesen Formen handelt es sich wieder um Abbildungen $\sigma\colon V \times V \to \mathbb{C}$ mit zwei Komponenten. Diese Abbildungen sind *keine* Bilinearformen, denn sie sind nur in der zweiten Komponente linear; in der ersten Komponente sind sie stattdessen „konjugiert-linear". Wir werden sehen, dass Hermite'sche Formen sich sehr ähnlich verhalten wie symmetrische reelle Bilinearformen und, dass sich viele Überlegungen von symmetrischen Bilinearformen auf Hermite'sche Formen übertragen lassen. Das Ziel ist es wieder eine Klassifikation dieser Formen vorzunehmen; in der Tat werden wir diese aus der Klassifikation der reellen symmetrischen Bilinearform ableiten.

I. Definition und Beispiele

3.4.1 **Definition** Sei V ein $\mathbb{C}$-Vektorraum. Eine Abbildung $\sigma\colon V \times V \to \mathbb{C}$ nennen wir *Hermite'sche Form*, wenn folgende zwei Bedingungen erfüllt sind:

(a) Für alle $u, v, v' \in V$ und $a \in \mathbb{C}$ gilt

$$\sigma(u, av + v') = a\sigma(u, v) + \sigma(u, v').$$

(b) Für alle $u, v \in V$ gilt

$$\sigma(u, v) = \overline{\sigma(v, u)}.$$

Die Menge aller Hermite'schen Formen auf V bezeichnen wir mit $\operatorname{Herm}(V)$.

3.4.2 **Anmerkung zur Definition**. (a) Eine Hermite'sche Form ist per Definition linear in der zweiten Komponente. Sie ist aber nicht linear in der ersten Komponente! Aus der Bedingung $\sigma(u, v) = \overline{\sigma(v, u)}$ kann man aber folgende Formel ableiten:

$$\begin{aligned}\sigma(au + u', v) &= \overline{\sigma(v, au + u')} = \overline{a\sigma(v, u) + \sigma(v, u')} \\ &= \overline{a}\ \overline{\sigma(v, u)} + \overline{\sigma(v, u')} \qquad (1.5.19) \\ &= \overline{a}\ \sigma(u, v) + \sigma(u', v)\end{aligned}$$

für alle $u, u', v \in V$ und $a \in \mathbb{C}$. Das heißt, die Abbildung ist additiv, aber Skalare aus $\mathbb{C}$ kann man nur konjugiert aus der Abbildung rausziehen. Man sagt daher, dass σ *konjugiert-linear* in der ersten Komponente ist.

Daraus ergibt sich folgende Rechenregel:

$$\sigma(\sum_{i=1}^{n} a_i u_i, \sum_{j=1}^{m} b_j v_j) = \sum_{i=1}^{n} \sum_{j=1}^{m} \overline{a_i} b_j \ \sigma(u_i, v_j).$$

(b) Setzt man einen Vektor $v \in V$ in beide Komponenten der Hermite'schen Form σ ein, dann gilt

$$\sigma(v, v) = \overline{\sigma(v, v)}.$$

Das heißt, $\sigma(v, v)$ liegt in $\mathbb{R}$; siehe 1.5.19.

3.4.3 **Vorsicht beim Blick in andere Bücher!** Bei der Definition von Hermite'schen Formen gibt es in der Mathematik keine einheitliche Festlegung, welche der beiden Komponenten linear und welche konjugiert-linear sein soll. Falls Sie in andere Quellen nachschlagen, müssen Sie unbedingt prüfen, welche Definition dort verwendet wird.

3.4.4 **Beispiel.** Die Abbildung $\sigma\colon \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ mit der Abbildungsvorschrift

$$\sigma\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}\right) = \sum_{i=1}^{n} \overline{x_i} y_i$$

ist eine Hermite'sche Form. Die Bedingung $\sigma(x, y) = \overline{\sigma(y, x)}$ ist offensichtlich erfüllt. Die Linearität in der zweiten Komponente erhält man aus folgender Rechnung:

$$\sigma(x, ay + y') = \sum_{i=1}^{n} \overline{x_i}(ay_i + y_i') = a \sum_{i=1}^{n} \overline{x_i} y_i + \sum_{i=1}^{n} \overline{x_i} y_i' = a\sigma(x, y) + \sigma(x, y').$$

Im nächsten Abschnitt werden wir weitere Beispiele kennenlernen.

Wie für Bilinearformen möchten wir auch für Hermite'sche Formen „reguläre“ und „ausgeartete“ Formen unterscheiden. Da $\sigma(u, u)$ eine reelle Zahl ist, kann man auch bei Hermite'schen Formen wieder von positiv und negativ definiten Formen sprechen.

3.4.5 **Definition** Sei V ein $\mathbb{C}$-Vektorraum. Eine Hermite'sche Form $\sigma \in \mathrm{Herm}(V)$ heißt

- *regulär*, wenn es für alle $u \neq 0$ in V einen Vektor $v \in V$ gibt, sodass $\sigma(u, v) \neq 0$ ist.
- *ausgeartet*, wenn σ nicht regulär ist.

- *positiv definit*, wenn $\sigma(u,u) > 0$ für alle $u \neq 0$ in V gilt.
- *negativ definit*, wenn $\sigma(u,u) < 0$ für alle $u \neq 0$ in V gilt.

3.4.6 **Aufgabe**. Zeigen Sie: Jede positiv definite Hermite'sche Form ist regulär. **L**

3.4.7 **Aufgabe**. Sei $\sigma \in \mathrm{Herm}(V)$ eine Hermite'sche Form. Zeigen Sie, dass σ genau dann regulär ist, wenn gilt: Für alle $v \neq 0$ in V gibt es einen Vektor $u \in V$ mit $\sigma(u,v) \neq 0$. **L**

3.4.8 **Beispiel**. Die Hermite'sche Form σ aus Beispiel 3.4.4 mit der Abbildungsvorschrift

$$\sigma\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}\right) = \sum_{i=1}^{n} \overline{x_i} y_i$$

ist positiv definit, denn für alle $x \in \mathbb{C}^n$ gilt

$$\sigma(x,x) = \sum_{i=1}^{n} \overline{x_i} x_i = \sum_{i=1}^{n} \underbrace{|x_i|^2}_{\geq 0} \geq 0.$$

Ist $x \neq 0$, dann ist auch ein Eintrag $x_i \neq 0$. Also ist $|x_i|^2 > 0$ und damit auch $\sigma(x,x) > 0$.

3.4.9 **Beispiel**. Sei V ein komplexer Vektorraum und sei $U \subseteq V$ ein Unterraum. Ist σ eine Hermite'sche Form, dann ist auch die Einschränkung $\sigma|_U \colon U \times U \to \mathbb{C}$ eine Hermite'sche Form. Ist σ positiv (bzw. negativ) definit, dann ist auch $\sigma|_U$ positiv (bzw. negativ) definit.

3.4.10 **Definition** Sei V ein $\mathbb{C}$-Vektorraum mit einer Hermite'schen Form $\sigma \in \mathrm{Herm}(V)$. Einen Unterraum $U \subseteq V$ nennen wir σ-regulär, wenn $\sigma|_U$ regulär ist.

II. Matrixdarstellung

Wir möchten Hermite'sche Formen wieder mithilfe von Matrizen beschreiben. Vieles was wir hier machen, wird Ihnen aus dem Abschnitt über Matrixdarstellungen von Bilinearformen bekannt vorkommen. Zuerst benötigen wir aber ein paar Vorüberlegungen.

3.4.11 **Definition** Sei $A = (a_{ij})_{i,j} \in \mathrm{M}_{m,n}(\mathbb{C})$. Die Matrix, die man erhält, wenn man alle Einträge von A konjugiert, bezeichnen wir mit $\overline{A}$, d.h., $\overline{A} = (\overline{a_{ij}})_{i,j}$.

3.4.12 **Beispiel**. Wir betrachten

$$A = \begin{pmatrix} 1 & i \\ 3+i & -2-i \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{C}).$$

Dann gilt

$$\overline{A} = \begin{pmatrix} 1 & -i \\ 3-i & -2+i \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{C}).$$

3.4.13 **Rechenregeln beim Konjugieren.** *Für alle $A, B \in \mathrm{M}_{m,n}(\mathbb{C})$, $C \in \mathrm{M}_{\ell,m}(\mathbb{C})$, $z \in \mathbb{C}$ gelten folgende Rechenregeln:*

(i) $\overline{A+B} = \overline{A} + \overline{B}$

(ii) $\overline{CA} = \overline{C}\,\overline{A}$

(iii) $\overline{zA} = \overline{z}\,\overline{A}$

(iv) $\overline{\overline{A}} = A$

(v) $\overline{A}^T = \overline{(A^T)}$

Beweis. Diese Rechenregeln folgen direkt aus den Rechenregeln für die komplexe Konjugation, die wir in Lemma 1.5.19 gesehen haben. Exemplarisch beweisen wir Aussage (ii) und lassen die übrigen Rechenregeln als Übungsaufgabe.

Es sei $A = (a_{ij})_{i,j} \in \mathrm{M}_{m,n}(\mathbb{C})$ und $C = (c_{ij})_{i,j} \in \mathrm{M}_{\ell,m}(\mathbb{C})$. Wir setzen $D = CA = (d_{ij})_{i,j} \in \mathrm{M}_{\ell,n}(\mathbb{C})$ und $D' = \overline{C}\,\overline{A} = (d'_{ij})_{i,j} \in \mathrm{M}_{\ell,n}(\mathbb{C})$. Nach Definition der Matrizenmultiplikation gilt

$$\begin{aligned} \overline{d_{ij}} &= \overline{\sum_{k=1}^{m} c_{ik} a_{kj}} \\ &= \sum_{k=1}^{m} \overline{c_{ik}}\,\overline{a_{kj}} && (1.5.19) \\ &= d'_{ij} \end{aligned}$$

für alle i, j. Das heißt, $\overline{CA} = \overline{D} = D' = \overline{C}\,\overline{A}$. □

3.4.14 **Notation**. Weil wir die Operation „konjugieren und transponieren" jetzt häufig verwenden werden, führen wir für alle komplexen Matrizen die Kurzschreibweise

$$A^* := \overline{A}^T$$

ein.

3.4.15 **Aufgabe**. Bestimmen Sie A^* für die Matrix $A = \begin{pmatrix} 3i & 0 \\ 3+i & 2-i \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{C})$. **L**

3.4.16 **Aufgabe**. Es seien $A \in \mathrm{M}_{\ell,m}(\mathbb{C})$, $B \in \mathrm{M}_{m,n}(\mathbb{C})$, $C \in \mathrm{M}_{n,n}(\mathbb{C})$ für $\ell, m, n \in \mathbb{N}$. **L** Zeigen Sie,

(a) $(AB)^* = B^*A^*$

(b) $(A^*)^* = A$

(c) Ist C invertierbar, dann ist auch C^* invertierbar und es gilt $(C^*)^{-1} = (C^{-1})^*$.

3.4.17 $\boxed{\textbf{Definition}}$ Eine quadratische Matrix $A \in \mathrm{M}_{n,n}(\mathbb{C})$ mit der Eigenschaft

$$A^* = A$$

nennen wir *Hermite'sch.*

3.4.18 **Beispiel**. (a) Die Matrix

$$A = \begin{pmatrix} 1 & i & 2-i \\ -i & 0 & 1 \\ 2+i & 1 & 5 \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{C})$$

ist Hermite'sch.

(b) Die Matrix

$$B = \begin{pmatrix} 1 & 0 & 2 \\ 0 & i & 1-3i \\ 2 & 1+3i & 0 \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{C})$$

ist nicht Hermite'sch, denn

$$B^* = \begin{pmatrix} 1 & 0 & 2 \\ 0 & -i & 1-3i \\ 2 & 1+3i & 0 \end{pmatrix}.$$

Die Matrizen B und B^* unterscheiden sich an der Stelle $(2, 2)$.

Allgemein gilt: Eine Hermite'sche Matrix hat auf der Diagonalen nur reelle Einträge. In der Tat, aus $A^* = A$ folgt $\overline{a_{ii}} = a_{ii}$ und damit $a_{ii} \in \mathbb{R}$.

3.4.19 Beispiel. Jede Hermite'sche Matrix $A \in \mathrm{M}_{n,n}(\mathbb{C})$ liefert eine Hermite'sche Form $\sigma\colon \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ über die Abbildungsvorschrift

$$\sigma(u, v) = u^* A v.$$

Wir prüfen, dass es sich wirklich um eine Hermite'sche Form handelt. Es seien $u, v, v' \in \mathbb{C}^n$ und $a \in \mathbb{C}$. Dann gilt

$$\sigma(u, av + v') = u^* A(av + v') = a\, u^* A v + u^* A v' = a\sigma(u, v) + \sigma(u, v'),$$

d.h., σ ist linear in der zweiten Komponente. Für alle $u, v \in \mathbb{C}^n$ gilt außerdem

$$\begin{aligned} \overline{\sigma(v,u)} &= \overline{(v^* A u)} \\ &= \overline{(v^* A u)}^T && (v^* A u \in \mathbb{C}) \\ &= (v^* A u)^* = u^* A^* (v^*)^* && (3.4.16\ \text{(a)}) \\ &= u^* A v && (A \text{ Hermite'sch und } 3.4.16\ \text{(b)}). \end{aligned}$$

Damit haben wir gezeigt, dass σ eine Hermite'sche Form ist. Wir werden bald sehen, dass jede Hermite'sche Form aus $\mathbb{C}^n$ von dieser Art ist.

3.4.20 Definition Es sei V ein endlich-dimensionaler $\mathbb{C}$-Vektorraum mit einer Hermite'schen Form σ. Ist $\mathcal{B} = (v_1, \ldots, v_n)$ eine geordnete Basis, dann nennen wir

$$\mathrm{M}_{\mathcal{B}}(\sigma) = (\sigma(v_i, v_j))_{i,j} \in \mathrm{M}_{n,n}(\mathbb{C})$$

die *Matrixdarstellung von σ zur Basis $\mathcal{B}$*. Um die Matrixdarstellung zu erhalten setzt man also alle Kombinationen von Basisvektoren in σ ein.

3.4.21 Anmerkung zur Definition. Die Matrixdarstellung $\mathrm{M}_{\mathcal{B}}(\sigma)$ einer Hermite'schen Form ist immer eine Hermite'sche Matrix. In der Tat, für alle $i, j \in \{1, \ldots, n\}$ gilt

$$\sigma(v_i, v_j) = \overline{\sigma(v_j, v_i)};$$

d.h., der Eintrag an der Position (j, i) ist das Konjugierte des Eintrages an der Position (i, j). Also gilt $\mathrm{M}_{\mathcal{B}}(\sigma)^* = \mathrm{M}_{\mathcal{B}}(\sigma)$.

3.4.22 Beispiel. Es sei $A \in \mathrm{M}_{n,n}(\mathbb{C})$ eine Hermite'sche Matrix und $\sigma\colon \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ die zugehörige Hermite'sche Form mit $\sigma(u, v) = u^* A v$ aus Beispiel 3.4.19. Sei $\mathcal{B} = (e_1, \ldots, e_n)$ die Standardbasis. Dann gilt

$$\sigma(e_i, e_j) = e_i^* A e_j = e_i^T A e_j = a_{ij},$$

d.h., A ist die Matrixdarstellung von σ zur Standardbasis.

3.4.23 **Satz.** *Sei V ein endlich-dimensionaler $\mathbb{C}$-Vektorraum mit einer geordneten Basis $\mathcal{B}$ und sei $\sigma \in \mathrm{Herm}(V)$. Für alle $u, v \in V$ gilt dann*

$$\sigma(u,v) = \kappa_{\mathcal{B}}(u)^* \, \mathrm{M}_{\mathcal{B}}(\sigma) \, \kappa_{\mathcal{B}}(v).$$

Insbesondere ist σ vollständig durch die Matrixdarstellung $\mathrm{M}_{\mathcal{B}}(\sigma)$ bestimmt.

Beweis. Sei $n = \dim_{\mathbb{C}} V$ und sei $\mathcal{B} = (v_1, \ldots, v_n)$. Es seien $u, v \in V$ gegeben mit $\kappa_{\mathcal{B}}(u) = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ und $\kappa_{\mathcal{B}}(v) = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ mit Koordinaten $a_j, b_j \in \mathbb{C}$. Das heißt, es gelten

$$u = \sum_{j=1}^{n} a_j v_j \quad \text{und} \quad v = \sum_{k=1}^{n} b_k v_k.$$

Die Rechenregel aus 3.4.2 und die Definition der Matrizenmultiplikation liefern uns nun

$$\sigma(u,v) = \sum_{j=1}^{n} \sum_{k=1}^{n} \overline{a_j} \, \sigma(v_j, v_k) \, b_k = \kappa_{\mathcal{B}}(u)^* \, \mathrm{M}_{\mathcal{B}}(\sigma) \, \kappa_{\mathcal{B}}(v).$$ □

3.4.24 **Aufgabe**. Es sei σ eine Hermite'sche Form auf $\mathbb{C}^n$. Zeigen Sie, dass es eine Hermite'sche Matrix $A \in \mathrm{M}_{n,n}(\mathbb{C})$ gibt, sodass $\sigma(u,v) = u^* A v$ für alle $u, v \in \mathbb{C}^n$ gilt. **L**

Als nächstes müssen wir uns überlegen, wie sich die Matrixdarstellung einer Hermite'schen Form bei einem Wechsel der Basis verändert.

3.4.25 **Transformationsformel für Hermite'sche Formen.** *Sei V ein komplexer endlich-dimensionaler Vektorraum mit zwei geordneten Basen $\mathcal{B}$ und $\mathcal{C}$. Für jede Hermite'sche Form $\sigma \in \mathrm{Herm}(V)$ gilt die Transformationsformel*

$$\mathrm{M}_{\mathcal{C}}(\sigma) = S^* \mathrm{M}_{\mathcal{B}}(\sigma) S,$$

wobei $S = {}_{\mathcal{B}}\mathrm{M}_{\mathcal{C}}(\mathrm{id}_V)$ die Transformationsmatrix von $\mathcal{C}$ nach $\mathcal{B}$ ist.

Beweis. Es seien $\mathcal{B} = (u_1, \ldots, u_n)$ und $\mathcal{C} = (v_1, \ldots, v_n)$. Es sei außerdem $S = (s_{ij})_{i,j} \in \mathrm{M}_{n,n}(\mathbb{C})$ die Transformationsmatrix. Das heißt, es gilt

$$v_k = \sum_{i=1}^{n} s_{ik} u_i$$

für alle $1 \leq k \leq n$. Seien $\ell, k \in \{1, \dots, n\}$ beliebig. Wir berechnen $\sigma(v_k, v_\ell)$ mit der Rechenregel aus 3.4.2 und erhalten

$$\sigma(v_k, v_\ell) = \sigma\Big(\sum_{i=1}^{n} s_{ik} u_i, \sum_{j=1}^{n} s_{j\ell} u_j\Big) = \sum_{i=1}^{n}\sum_{j=1}^{n} \overline{s_{ik}}\ \sigma(u_i, u_j)\ s_{j\ell}$$
$$= \sum_{i=1}^{n}\sum_{j=1}^{n} s'_{ki}\ \sigma(u_i, u_j)\ s_{j\ell}$$

wobei $s'_{ki} = \overline{s_{ik}}$ der Eintrag von S^* an der Position (k, i) ist. Nach Definition der Matrizenmultiplikation ist $\sigma(v_k, v_\ell)$ damit genau der Eintrag von $S^* \mathrm{M}_{\mathcal{B}}(\beta) S$ an der Stelle (k, ℓ). □

3.4.26 **Definition** Zwei Hermite'sche Matrizen $A, B \in \mathrm{M}_{n,n}(\mathbb{C})$ heißen *Hermite'sch kongruent*, wenn es eine invertierbare Matrix $S \in \mathrm{GL}_n(\mathbb{C})$ gibt, sodass $A = S^* B S$ gilt.

3.4.27 **Aufgabe**. Für Hermite'sche Matrizen $A, B \in \mathrm{M}_{n,n}(\mathbb{C})$ schreiben wir $A \sim B$, wenn A Hermite'sch kongruent zu B ist. Zeigen Sie, dass $\sim$ eine Äquivalenzrelation ist. **L**

Die Äquivalenzklassen dieser Relation nennen wir *Hermite'sche Kongruenzklassen.*

Das folgende Lemma ist das Hermite'sche Analogon zum Lemma 3.1.19.

3.4.28 **Lemma.** *Sei V ein $\mathbb{C}$-Vektorraum der Dimension $\dim_{\mathbb{C}} V = n$. Für jede Hermite'sche Form $\sigma \in \mathrm{Herm}(V)$ ist die Menge*

$$\big\{\mathrm{M}_{\mathcal{B}}(\sigma) \mid \mathcal{B} \text{ geordnete Basis von } V\big\}$$

aller Matrixdarstellungen eine Hermite'sche Kongruenzklasse.

Beweis. Aus der Transformationsformel 3.4.25 folgt direkt, dass je zwei Matrixdarstellungen einer Hermite'schen Form Hermite'sch kongruent sind. Die Menge aller Matrixdarstellungen ist also in einer Hermite'schen Kongruenzklasse enthalten.

Umgekehrt sei nun $A \in \mathrm{M}_{n,n}(\mathbb{C})$ eine Hermite'sche Matrix, die Hermite'sch kongruent zu $\mathrm{M}_{\mathcal{B}}(\sigma)$ für eine geordnete Basis $\mathcal{B} = (u_1, \dots, u_n)$ von V ist. Es gibt also eine invertierbare Matrix $S = (s_{ij})_{i,j} \in \mathrm{GL}_n(\mathbb{C})$ mit $A = S^* \mathrm{M}_{\mathcal{B}}(\sigma) S$. Wir definieren nun eine neue Basis von V. Für $j \in \{1, \dots, n\}$ definieren wir

$$v_j = \sum_{i=1}^{n} s_{ij} u_i.$$

Da S invertierbar ist, ist $\mathcal{C} = (v_1, \dots, v_n)$ eine geordnete Basis von V. Nach Konstruktion ist S die Transformationsmatrix ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{C}}(\mathrm{id}_V)$. Aus der Transformationsformel 3.4.25 erhalten wir

$$A = S^* \mathrm{M}_{\mathcal{B}}(\sigma) S = {}_{\mathcal{B}}\mathrm{M}_{\mathcal{C}}(\mathrm{id}_V)^* \mathrm{M}_{\mathcal{B}}(\sigma) {}_{\mathcal{B}}\mathrm{M}_{\mathcal{C}}(\mathrm{id}_V) = \mathrm{M}_{\mathcal{C}}(\sigma),$$

d.h., A ist eine Matrixdarstellung von σ. Da A beliebig war, ist die Menge der Matrixdarstellungen von σ genau eine Hermite'sche Kongruenzklasse. □

3.4.29 **Aufgabe**. Es seien $A, B \in \mathrm{M}_{n,n}(\mathbb{C})$ zwei Hermite'sche Matrizen. Zeigen Sie: Sind A und B Hermite'sch kongruent, dann gilt $\mathrm{Rg}(A) = \mathrm{Rg}(B)$. **L**

3.4.30 **Definition** Sei V ein endlich-dimensionaler $\mathbb{C}$-Vektorraum und sei $\sigma \in \mathrm{Herm}(V)$. Wir definieren den *Rang* von σ als

$$\mathrm{Rg}(\sigma) = \mathrm{Rg}\big(\mathrm{M}_{\mathcal{B}}(\sigma)\big)$$

für eine beliebige geordnete Basis $\mathcal{B}$ von V. Aus Aufgabe 3.4.29 folgt, dass der Rang nicht von der Wahl der Basis $\mathcal{B}$ abhängt.

3.4.31 **Definition** Zwei Hermite'sche Formen σ_1, σ_2 auf einem endlich-dimensionalen $\mathbb{C}$-Vektorraum nennen wir *kongruent*, wenn es geordnete Basen $\mathcal{B}_1$ und $\mathcal{B}_2$ von V gibt, sodass

$$\mathrm{M}_{\mathcal{B}_1}(\sigma_1) = \mathrm{M}_{\mathcal{B}_2}(\sigma_2).$$

Da zwei Hermite'sche Kongruenzklassen entweder disjunkt oder gleich sind, folgt aus Lemma 3.4.28: σ_1, σ_2 sind genau dann kongruent, wenn jede Matrixdarstellung von σ_1 auch eine Matrixdarstellung von σ_2 ist.

3.4.32 **Aufgabe**. Zeigen Sie, dass Kongruenz von Hermite'schen Formen eine Äquivalenzrelation auf $\mathrm{Herm}(V)$ definiert. **L**

3.4.33 **Einfache Matrixdarstellungen, Klassifikation und Normalform**. Nach diesen Vorüberlegungen sind wir wieder in der Lage die wesentlichen Fragen zu Hermite'schen Formen zu formulieren:

- Besitzt jede Hermite'sche Form eine „einfache" Matrixdarstellung?
- Kann man Hermite'sche Formen bis auf Kongruenz klassifizieren?
- Können wir für Hermite'sche Formen eindeutige Matrixdarstellungen in „Normalform" festlegen? In anderen Worten: Können wir aus jeder Hermite'schen Kongruenzklasse einen eindeutigen Vertreter bestimmen?
- Und falls ja: Wie berechnet man eine Matrixdarstellung in Normalform?

Wir werden diese Fragen in den folgenden Abschnitten beantworten. Die ganze Theorie ist sehr eng verwandt mit der Theorie der reellen symmetrischen Bilinearformen. Die Klassifikation wird wieder auf einem *Sylvester'schen Trägheitssatz* basieren. In der Tat, wenn man möchte, kann man die Klassifikation der Hermite'schen Formen ganz parallel herleiten und im wesentlichen Beweis für Beweis aus dem Abschnitt über reelle symmetrische Bilinearformen auf Hermite'sche Formen übertragen. Wir werden hier allerdings anders vorgehen und die Klassifikation der Hermite'schen Formen tatsächlich auf die Klassifikation der reellen symmetrischen Bilinearformen zurückführen. Dazu müssen wir uns nur überlegen, wie man aus einer Hermite'schen Form eine reelle symmetrische Bilinearform bekommt.

III. Komplex wird reell

3.4.34 **Komplex nach reell**. Sei V ein $\mathbb{C}$-Vektorraum. Dann ist V auch ein $\mathbb{R}$-Vektorraum! Warum? Die reellen Zahlen $\mathbb{R}$ sind in den komplexen Zahlen $\mathbb{C}$ enthalten. Wir können also V mit derselben Addition betrachten, aber für die Skalarmultiplikation nur reelle Zahlen zulassen. Mit einem Blick auf die Vektorraum-Axiome aus [MG, 6.1.1] kann man sich überzeugen, dass alle Axiome auch mit der eingeschränkten Skalarmultiplikation erfüllt sind. Um keine Verwirrung zu stiften, schreiben wir $V_{\mathbb{R}}$, wenn wir V als $\mathbb{R}$-Vektorraum auffassen. Man sagt, dass $V_{\mathbb{R}}$ aus V durch *Restriktion der Skalare* entsteht.

3.4.35 **Lemma.** *Es sei V ein $\mathbb{C}$-Vektorraum mit einer Basis $\mathcal{B} = \{v_1, \dots, v_n\}$. Dann ist $\mathcal{B}_{\mathbb{R}} = \{v_1, iv_1, v_2, iv_2, \dots, v_n, iv_n\}$ eine Basis von $V_{\mathbb{R}}$.*

Insbesondere gilt $\dim_{\mathbb{R}} V_{\mathbb{R}} = 2 \dim_{\mathbb{C}} V$.

Beweis. Wir zeigen zuerst, dass $\mathcal{B}_{\mathbb{R}}$ ein Erzeugendensystem von $V_{\mathbb{R}}$ ist. Sei $v \in V$. Dann gibt es komplexe Zahlen $z_1, \dots, z_n$, sodass

$$v = \sum_{j=1}^{n} z_j v_j$$

gilt, weil $\mathcal{B}$ eine Basis von V ist. Wir schreiben $z_j = x_j + iy_j$ mit reellen Zahlen $x_j, y_j \in \mathbb{R}$. Dann erhalten wir

$$v = \sum_{j=1}^{n} z_j v_j = \sum_{j=1}^{n} x_j \underbrace{v_j}_{\in \mathcal{B}_{\mathbb{R}}} + y_j \underbrace{(iv_j)}_{\in \mathcal{B}_{\mathbb{R}}}.$$

Also erzeugt $\mathcal{B}_{\mathbb{R}}$ den reellen Vektorraum $V_{\mathbb{R}}$.

Die Vektoren in $\mathcal{B}_{\mathbb{R}}$ sind $\mathbb{R}$-linear unabhängig. Denn angenommen, es sind $x_j, y_j \in \mathbb{R}$ mit

$$0 = \sum_{j=1}^{n} x_j v_j + \sum_{j=1}^{n} y_j (i v_j) = \sum_{j=1}^{n} \underbrace{(x_j + i y_j)}_{\in \mathbb{C}} v_j$$

gegeben, dann folgt $x_j + i y_j = 0$ für alle j, weil $\mathcal{B}$ $\mathbb{C}$-linear unabhängig ist. Durch Trennung von Real- und Imaginärteil schließen wir $x_j = 0$ und $y_j = 0$ für alle $j \leq n$. Damit sind die Vektoren in $\mathcal{B}_{\mathbb{R}}$ linear unabhängig (über $\mathbb{R}$). □

3.4.36 **Beispiel**. Die komplexen Zahlen $\mathbb{C}$ bilden einen $\mathbb{C}$-Vektorraum der Dimension 1. Aufgefasst als reeller Vektorraum hat $\mathbb{C}$ die Dimension 2. Das ist nicht verwunderlich, denn $\mathbb{C}$ hatten wir ursprünglich als $\mathbb{R} \times \mathbb{R}$ definiert. Eine Basis ist gegeben durch 1 und i, denn jede komplexe Zahl lässt sich eindeutig in der Form $x + yi$ mit $x, y \in \mathbb{R}$ darstellen.

3.4.37 **Unterräume**. Es sei V ein $\mathbb{C}$-Vektorraum. Ist $U \subseteq V$ ein Unterraum, dann ist U auch ein Unterraum des reellen Vektorraumes $V_{\mathbb{R}}$. In der Tat, wenn U unter der Multiplikation von Skalaren aus $\mathbb{C}$ abgeschlossen ist, dann ist U natürlich auch unter der Multiplikation mit Skalaren aus $\mathbb{R}$ abgeschlossen.

Umgekehrt klappt das übrigens nicht: $V_{\mathbb{R}}$ hat viele Unterräume, die keine Unterräume von V sind. Zum Beispiel ist $\mathbb{R}$ ein $\mathbb{R}$-Unterraum von $\mathbb{C}$, ist aber nicht unter der Multiplikation mit komplexen Zahlen abgeschlossen.

3.4.38 **Lineare Abbildungen**. Es seien V, W zwei $\mathbb{C}$-Vektorräume. Ist $\varphi\colon V \to W$ eine $\mathbb{C}$-lineare Abbildung, dann ist auch $\varphi\colon V_{\mathbb{R}} \to W_{\mathbb{R}}$ eine $\mathbb{R}$-lineare Abbildung. In der Tat, wenn die Gleichung aus 2.2.18 für alle Skalare in $\mathbb{C}$ gilt, dann insbesondere auch für alle Skalare in $\mathbb{R}$.

Damit sind wir nun in der Lage Hermite'sche Formen in reelle symmetrische Bilinearformen zu verwandeln.

3.4.39 **Satz.** *Sei V ein $\mathbb{C}$-Vektorraum und sei $\sigma \in \mathrm{Herm}(V)$. Dann ist*

$$\mathrm{Re}(\sigma)\colon V_{\mathbb{R}} \times V_{\mathbb{R}} \to \mathbb{R} \quad \textit{mit} \quad (u, v) \mapsto \mathrm{Re}(\sigma(u, v))$$

eine reelle symmetrische Bilinearform. Die Form σ ist genau dann regulär, wenn $\mathrm{Re}(\sigma)$ regulär ist.

Wir nennen $\mathrm{Re}(\sigma)$ den *Realteil* von σ.

Beweis. Zuerst machen wir uns klar, dass $\mathrm{Re}(\sigma)$ symmetrisch ist. Ist $z \in \mathbb{C}$, dann gilt $\mathrm{Re}(z) = \mathrm{Re}(\overline{z})$, denn der Realteil ändert sich bei der Konjugation nicht. Für alle $u, v \in V$ gilt also

$$\mathrm{Re}(\sigma(u,v)) = \mathrm{Re}(\overline{\sigma(v,u)}) = \mathrm{Re}(\sigma(v,u)).$$

Wir zeigen nun, dass $\mathrm{Re}(\sigma)$ eine Bilinearform ist. Dazu seien $u, v, v' \in V_{\mathbb{R}}$ und $a \in \mathbb{R}$ gegeben. Dann gilt (siehe Aufgabe 1.5.9)

$$\mathrm{Re}(\sigma(u, av + v')) = \mathrm{Re}(a\sigma(u,v) + \sigma(u,v')) = a\mathrm{Re}(\sigma(u,v)) + \mathrm{Re}(\sigma(u,v')),$$

d.h., $\mathrm{Re}(\sigma)$ ist linear in der zweiten Komponente. Daraus folgt auch direkt die Linearität in der ersten Komponente, denn $\mathrm{Re}(\sigma(u,v)) = \mathrm{Re}(\sigma(v,u))$ haben wir bereits für alle $u, v \in V$ gezeigt.

Wir nehmen nun an, dass σ regulär ist. Es sei $u \neq 0$ in V gegeben. Wir suchen einen Vektor $v \in V_{\mathbb{R}}$ mit $\mathrm{Re}(\sigma(u,v)) \neq 0$. Da σ regulär ist, gibt es einen Vektor $v' \in V$ mit $z = \sigma(u,v') \neq 0$. Ist $\mathrm{Re}(z) \neq 0$, dann setzen wir $v = v'$ und erhalten $\mathrm{Re}(\sigma)(u,v) \neq 0$. Ist $\mathrm{Re}(z) = 0$, dann ist z rein imaginär, d.h., $z = bi$ für eine reelle Zahl $b \neq 0$. In diesem Fall setzen wir $v = iv'$, denn dann gilt

$$\mathrm{Re}(\sigma(u,v)) = \mathrm{Re}(i\sigma(u,v')) = \mathrm{Re}(iz) = -b \neq 0.$$

Wir haben damit gezeigt, dass $\mathrm{Re}(\sigma)$ regulär ist.

Umgekehrt nehmen wir nun an, dass $\mathrm{Re}(\sigma)$ regulär ist. Sei $u \neq 0$ in V gegeben. Dann gibt es einen Vektor $v \in V$ mit $\mathrm{Re}(\sigma(u,v)) \neq 0$. Dann ist aber auch $\sigma(u,v) \neq 0$; folglich ist σ regulär. □

3.4.40 **Aufgabe**. Sei $\sigma \in \mathrm{Herm}(V)$. Zeigen Sie, dass σ genau dann positiv definit ist, wenn $\mathrm{Re}(\sigma)$ positiv definit ist. **L**

3.4.41 **Bemerkung**. Auf den ersten Blick könnte man meinen, dass beim Übergang von σ zu $\mathrm{Re}(\sigma)$ Information verloren geht. Das ist aber nicht der Fall, denn σ kann man vollständig durch $\mathrm{Re}(\sigma)$ beschreiben. Es gilt nämlich für alle $u, v \in V$

$$\begin{aligned}
\mathrm{Re}(\sigma(u,v)) + i\mathrm{Re}(\sigma(iu,v)) &= \mathrm{Re}(\sigma(u,v)) + i\mathrm{Re}((-i)\sigma(u,v)) && \text{(konjugiert-linear)} \\
&= \mathrm{Re}(\sigma(u,v)) + i\mathrm{Im}(\sigma(u,v)) && (1.5.8) \\
&= \sigma(u,v).
\end{aligned}$$

Mit dieser Verwandlung kann man nun relativ einfach Ergebnisse von reellen symmetrischen Bilinearformen auf Hermite'sche Formen übertragen. Wir illustrieren das mit dem Zerlegungssatz, den wir für Bilinearformen in 3.1.43 gesehen haben.

3.4.42 **Zerlegungssatz.** *Es sei V ein endlich-dimensionaler $\mathbb{C}$-Vektorraum mit einer Hermite'schen Form σ. Sei $U \subseteq V$ ein σ-regulärer Unterraum. Dann gilt $V = U \oplus W$ mit*

$$W = \{w \in V \mid \sigma(u, w) = 0 \text{ für alle } u \in U\}.$$

Beweis. Mit dem Unterraumkriterium kann man leicht zeigen, dass W ein Unterraum von V ist. Nach Voraussetzung ist $\sigma|_U$ regulär. Aus 3.4.39 folgt damit, dass auch $\mathrm{Re}(\sigma|_U) = \mathrm{Re}(\sigma)|_U$ regulär ist. Der Zerlegungssatz 3.1.43 liefert uns nun

$$V_{\mathbb{R}} = U \oplus W'$$

mit $W' = \{w \in V_{\mathbb{R}} \mid \mathrm{Re}(\sigma(u, w)) = 0 \text{ für alle } u \in U\}$. Wir behaupten nun, dass $W = W'$ ist. Da die Definition der direkten Summe die Skalarmultiplikation des Vektorraumes gar nicht verwendet, folgt daraus $V = U \oplus W$.

Offensichtlich ist $W \subseteq W'$. Für die umgekehrte Inklusion sei $w \in W'$. Für alle $u \in U$ ist auch $iu \in U$, denn U ist ein Unterraum von V. Mit 3.4.41 erhalten wir

$$\sigma(u, w) = \mathrm{Re}(\sigma(u, w)) + i\mathrm{Re}(\sigma(iu, w)) = 0$$

und, da u beliebig war, schließen wir daraus $w \in W$. □

IV. Klassifikation

In diesem Abschnitt besprechen wir die Klassifikation der Hermite'schen Formen bis auf Kongruenz. Wir werden bald sehen, dass jede Hermite'sche Form eine diagonale Matrixdarstellung besitzt. Vorab wollen wir aber wieder einen Trägheitssatz herleiten, der besagt, dass die Anzahl der positiven und negativen Einträge nicht von der gewählten Darstellung abhängt.

3.4.43 **Sylvester'scher Trägheitssatz.** *Sei V ein $\mathbb{C}$-Vektorraum der Dimension n und sei $\sigma \in \mathrm{Herm}(V)$ vom Rang $r = \mathrm{Rg}(\sigma)$. Dann gibt es Zahlen $r_+, r_- \in \mathbb{N}$ mit $r = r_+ + r_-$, sodass jede diagonale Matrixdarstellung genau r_+ positive, r_- negative und $n - r$ verschwindende Diagonaleinträge hat.*

Beweis. Sei $\mathcal{B} = (v_1, \dots, v_n)$ eine geordnete Basis, sodass $\mathrm{M}_{\mathcal{B}}(\sigma)$ diagonal ist. Wir betrachten die geordnete Basis $\mathcal{B}_{\mathbb{R}} = (v_1, iv_1, v_2, iv_2, \dots, v_n, iv_n)$ von $V_{\mathbb{R}}$, die wir aus Lemma 3.4.35 kennen. Was ist die Matrixdarstellung von $\mathrm{Re}(\sigma)$ bezüglich $\mathcal{B}_{\mathbb{R}}$? Es gilt

$$0 = \sigma(v_j, v_k) = \sigma(iv_j, v_k) = \sigma(v_j, iv_k) = \sigma(iv_j, iv_k)$$

für alle $j \neq k$. Dazu sind $\sigma(iv_j, v_j) = -i\sigma(v_j, v_j)$ und $\sigma(v_j, iv_j) = i\sigma(v_j, v_j)$ rein imaginär, d.h., $\mathrm{Re}(\sigma(v_j, iv_j)) = 0 = \mathrm{Re}(iv_j, v_j)$ für alle $j \in \{1, \dots, n\}$. Weiter gilt $\sigma(iv_j, iv_j) = (-i)i\sigma(v_j, v_j) = \sigma(v_j, v_j)$. Damit ist $\mathrm{M}_{\mathcal{B}_\mathbb{R}}(\mathrm{Re}(\sigma))$ eine diagonale Matrix in $\mathrm{M}_{2n,2n}(\mathbb{R})$, wobei jeder Diagonaleintrag von $\mathrm{M}_\mathcal{B}(\sigma)$ immer doppelt auf der Diagonalen steht. Da nach dem Trägheitssatz 3.2.24 die Anzahl der positiven, negativen und Null Einträge in einer diagonalen Matrixdarstellung von $\mathrm{Re}(\sigma)$ nicht von der gewählten geordneten Basis abhängt, hängt auch die Anzahl dieser Einträge in $\mathrm{M}_\mathcal{B}(\sigma)$ nicht von der gewählten Basis ab.

Wir formulieren dieses Ergebnis nochmal leicht um: Hat $\mathrm{Re}(\sigma)$ den Typ (s_+, s_-, k), dann sind s_+, s_- gerade und wir setzen $r_+ = \frac{s_+}{2}$ und $r_- = \frac{s_-}{2}$. Da die Diagonalmatrix $\mathrm{M}_\mathcal{B}(\sigma)$ genau $r_+ + r_-$ Einträge ungleich Null hat, ist $r_+ + r_- = r = \mathrm{Rg}(\sigma)$. □

3.4.44 **Definition** Sei V ein $\mathbb{C}$-Vektorraum der Dimension n und sei $\sigma \in \mathrm{Herm}(V)$ vom Rang $r = \mathrm{Rg}(\sigma)$. Das Tripel $(r_+, r_-, n-r)$ nennen wir den *Typ* von σ. Die Differenz $r_+ - r_-$ nennt man die *Signatur* von σ.

3.4.45 **Satz.** *Sei $V \neq \{0\}$ ein endlich-dimensionaler $\mathbb{C}$-Vektorraum mit einer Hermite'schen Form σ. Dann gibt es eine geordnete Basis $\mathcal{B}$ von V, sodass $\mathrm{M}_\mathcal{B}(\sigma)$ eine Diagonalmatrix mit Einträgen aus $\{0, 1, -1\}$ ist.*

Beweis. Ist $\sigma = 0$, dann nehmen wir irgendeine geordnete Basis. Wir können also $\sigma \neq 0$ annehmen.

Es sei $n = \dim_\mathbb{C} V$. Wir beweisen den Satz mit vollständiger Induktion nach n. Ist $n = 1$, dann nehmen wir einen Vektor $v' \neq 0$ in V. Da V ein-dimensional ist, gilt $V = \langle v' \rangle$ und damit $c = \sigma(v', v') \neq 0$. Wir erinnern uns, dass c eine reelle Zahl ist. Wir setzen $v = \frac{1}{\sqrt{|c|}} v'$, denn dann gilt

$$\sigma(v, v) = \frac{1}{|c|}\sigma(v', v') = \frac{c}{|c|} = \pm 1$$

und $\mathcal{B} = (v)$ ist die gesuchte Basis.

Es sei nun $n \geq 2$. Da $\sigma \neq 0$ ist, ist auch $\mathrm{Re}(\sigma) \neq 0$ (siehe Bemerkung 3.4.41). Aus der Polarisationsformel 3.2.9 folgt nun, dass es einen Vektor $v_1' \in V$ mit

$$\sigma(v_1', v_1') = \mathrm{Re}(\sigma(v_1', v_1')) = c \neq 0$$

gibt. Wir definieren $v_1 = \frac{1}{\sqrt{|c|}} v_1'$, denn dann gilt $\sigma(v_1, v_1) = \pm 1$ (das Vorzeichen hängt davon ab, ob c positiv oder negativ ist).

Der $\mathbb{C}$-Unterraum $\langle v_1' \rangle \subseteq V$ ist also σ-regulär. Der Zerlegungssatz 3.4.42 liefert uns

$$V = \langle v_1 \rangle \oplus W$$

mit $W = \{w \in V \mid \sigma(v_1, w) = 0\}$. Es gilt also $\dim_{\mathbb{C}} W = \dim_{\mathbb{C}} V - 1 = n - 1$. Nach Induktionsvoraussetzung gibt es eine geordnete Basis $\mathcal{B}' = (v_2, \ldots, v_n)$ von W, sodass die Matrixdarstellung von $\sigma|_W$ bzgl. $\mathcal{B}'$ eine Diagonalmatrix mit Einträgen aus $\{0, 1, -1\}$ ist., d.h., $\sigma(v_j, v_k) = 0$ und $\sigma(v_j, v_j) \in \{0, 1, -1\}$ für alle $j \neq k$ aus $\{2, \ldots, n\}$.

Wir setzen nun $\mathcal{B} = (v_1, \ldots, v_n)$. Da $v_2, \ldots, v_n$ in W liegen, haben wir $\sigma(v_1, v_j) = 0$ für alle $j \geq 2$. Damit haben wir gezeigt, dass $\mathrm{M}_{\mathcal{B}}(\sigma)$ eine Diagonalmatrix mit Einträgen aus $\{0, 1, -1\}$ ist. □

Ordnet man die Basisvektoren entsprechend der Werte $\sigma(v_j, v_j)$ in der Reihenfolge $1, -1, 0$ an, dann sprechen wir von einer Matrixdarstellung in Normalform.

3.4.46 **Definition** Sei σ eine Hermite'sche Form auf dem endlich-dimensionalen Vektorraum V. Eine Matrixdarstellung in der Form

$$\mathrm{M}_{\mathcal{B}}(\sigma) = \mathrm{diag}(I_{r_+}, -I_{r_-}, 0_k),$$

wobei (r_+, r_-, k) der Typ von σ ist, nennen wir eine Matrixdarstellung in *Normalform.*

Eine geordnete Basis, die eine Darstellung in Normalform liefert, kann man wie im Verfahren für symmetrische Bilinearformen berechnen.

3.4.47 **Normalform Hermite'scher Formen**.

`Gegeben:` Hermite'sche Form $\sigma \in \mathrm{Herm}(V)$.

`Gesucht:` geordnete Basis $\mathcal{B}$ von V, sodass $\mathrm{M}_{\mathcal{B}}(\sigma)$ in Normalform ist.

`Verfahren:`

Berechne den Rang $r = \mathrm{Rg}(\sigma)$. Setze $W_1 = V$.

- Für $j = 1, \ldots, r$ wiederhole folgenden Schritt: Wähle einen Vektor $v_j' \in W_j$ mit $\sigma(v_j', v_j') \neq 0$ und berechne

  $$W_{j+1} = \{w \in W_j \mid \sigma(v_j', w) = 0\}.$$

  Berechne $c_j = \sigma(v_j', v_j')$ und setze $v_j := \frac{1}{\sqrt{|c_j|}} v_j'$.

- Falls $r < n$ ist, gilt nun $\sigma|_{W_{r+1}} = 0$. Wähle dann eine beliebige Basis $(v_{r+1}, \ldots, v_n)$ von W_{r+1}, setze $c_j = 0$ für alle $j > r$.

Nach Abschluss der Berechnungen ist $\mathcal{B}' = (v_1, \ldots, v_n)$ eine geordnete Basis, sodass $\mathrm{M}_{\mathcal{B}}(\sigma)$ nur die Diagonaleinträge $0, 1, -1$ besitzt. Durch Umordnung der Basisvektoren in $\mathcal{B}'$ gelangt man zur gesuchten geordneten Basis $\mathcal{B}$.

3.4.48 **Beispiel**. Wir betrachten die Hermite'sche Matrix

$$A = \begin{pmatrix} 0 & 1+i & 0 \\ 1-i & 0 & i \\ 0 & -i & 4 \end{pmatrix} \in \mathrm{M}_{n,n}(\mathbb{C})$$

und die zugehörige Hermite'sche Form $\sigma(u, v) = u^* A v$ auf $\mathbb{C}^3$. Wir suchen eine Matrixdarstellung in Normalform.

Zuerst berechnen wir den Rang von A und finden $\mathrm{Rg}(A) = 3$.

Dann setzen wir $W_1 = \mathbb{C}^3$ und suchen einen Vektor $v_1' \in W_1$ mit $\sigma(v_1', v_1') \neq 0$. Wir nehmen den dritten Standardbasisvektor $e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. Dann gilt

$$\sigma(v_1', v_1') = 4.$$

Wir skalieren mit $\frac{1}{\sqrt{4}} = \frac{1}{2}$ und setzen $v_1 = \frac{1}{2} v_1'$, denn dann gilt $\sigma(v_1, v_1) = 1$.

Nun berechnen wir das Komplement

$$\begin{aligned} W_2 &= \{w \in W_1 \mid \sigma(v_1', w) = 0\} = \{w \in W_1 \mid (v_1')^* A w = 0\} \\ &= \mathrm{Ker}((v_1')^* A) = \mathrm{Ker}(\begin{pmatrix} 0 & -i & 4 \end{pmatrix}) \\ &= \langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 4 \\ i \end{pmatrix} \rangle. \end{aligned}$$

Wir suchen einen Vektor $v_2' \in W_2$ mit $\sigma(v_2', v_2') \neq 0$. Der erste Standardbasisvektor ist also nicht geeignet. Wir wählen also $v_2' = \begin{pmatrix} 0 \\ 4 \\ i \end{pmatrix}$. Es ist

$$\begin{aligned} \sigma(v_2', v_2') = (v_2')^* A v_2' &= \begin{pmatrix} 0 & 4 & -i \end{pmatrix} \begin{pmatrix} 0 & 1+i & 0 \\ 1-i & 0 & i \\ 0 & -i & 4 \end{pmatrix} \begin{pmatrix} 0 \\ 4 \\ i \end{pmatrix} \\ &= \begin{pmatrix} 4(1-i) & -1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 4 \\ i \end{pmatrix} = -4. \end{aligned}$$

Wir setzen $v_2 = \frac{1}{2}v_2'$, denn dann gilt $\sigma(v_2, v_2) = -1$.

Wir bestimmen wieder das Komplement

$$\begin{aligned} W_3 &= \{w \in W_2 \mid \sigma(v_2', w) = 0\} = W_2 \cap \{w \in \mathbb{C}^3 \mid (v_2')^* A w = 0\} \\ &= W_2 \cap \operatorname{Ker}((v_2')^* A) = \operatorname{Ker}(\begin{pmatrix} 0 & -i & 4 \\ 4(1-i) & -1 & 0 \end{pmatrix}) \\ &= \langle \begin{pmatrix} 1+i \\ 8 \\ 2i \end{pmatrix} \rangle. \end{aligned}$$

Wir setzen nun $v_3' = \begin{pmatrix} 1+i \\ 8 \\ 2i \end{pmatrix}$ und finden

$$\sigma(v_3', v_3') = (v_3')^* A v_3' = 16;$$

also definieren wir $v_3 = \frac{1}{\sqrt{16}} v_3' = \begin{pmatrix} \frac{1+i}{4} \\ 2 \\ \frac{i}{2} \end{pmatrix}$.

Um eine Matrixdarstellung in Normalform zu erhalten, müssen wir die Basisvektoren noch richtig anordnen. Wir setzen $\mathcal{B} = (v_1, v_3, v_2)$. Dann gilt

$$\mathrm{M}_{\mathcal{B}}(\sigma) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Es handelt sich also um eine Hermite'sche Form vom Typ $(2, 1, 0)$ und mit der Signatur $2 - 1 = 1$. Ist S die Transformationsmatrix von $\mathcal{B}$ in die Standardbasis, dann gilt $S^* A S = \mathrm{M}_{\mathcal{B}}(\sigma)$.

3.4.49 Aufgabe. Wir betrachten die Hermite'sche Matrix $A = \begin{pmatrix} -4 & 1+i \\ 1-i & 9 \end{pmatrix}$ und die Hermite'sche Form $\sigma(u, v) = u^* A v$ auf $\mathbb{C}^2$. Geben Sie eine geordnete Basis $\mathcal{B}$ von $\mathbb{C}^2$ an, sodass $\mathrm{M}_{\mathcal{B}}(\sigma)$ eine Darstellung in Normalform ist. **L**

Mit diesen Ergebnissen können wir nun eine Klassifikation der Hermite'schen Formen abgeben.

3.4.50 Klassifikationssatz für Hermite'sche Formen. *Sei V ein endlich-dimensionaler $\mathbb{C}$-Vektorraum. Je zwei Hermite'sche Formen sind genau dann kongruent, wenn sie denselben Typ haben.*

Beweis. Nehmen wir zuerst an, dass $\sigma_1, \sigma_2 \in \mathrm{Herm}(V)$ denselben Typ (r_+, r_-, k) haben. Dann gibt es geordnete Basen $\mathcal{B}_1, \mathcal{B}_2$ von V, sodass die eine Darstellung in Normalform liefern, d.h.

$$\mathrm{M}_{\mathcal{B}_1}(\sigma_1) = \mathrm{diag}(I_{r_+}, -I_{r_-}, 0_k) = \mathrm{M}_{\mathcal{B}_2}(\sigma_2).$$

Damit sind σ_1 und σ_2 kongruent.

Umgekehrt nehmen wir an, dass σ_1 und σ_2 kongruent sind. Dann ist jede Matrixdarstellung von σ_1 auch eine Matrixdarstellung von σ_2 (siehe 3.4.31). Da man den Typ an jeder diagonalen Matrixdarstellung ablesen kann, haben σ_1, σ_2 denselben Typ. $\square$

Für eine Hermite'sche Matrix $A \in \mathrm{M}_{n,n}(\mathbb{C})$ definieren wir den Typ von A als den Typ der zugehörigen Hermite'schen Form $\sigma(u, v) = u^*Av$. Mit dieser Definition erhalten wir nun auch eine Klassifikation der Hermite'schen Matrizen bis auf Hermite'sche Kongruenz.

3.4.51 **Klassifikationssatz für Hermite'sche Matrizen.** *Je zwei Hermite'sche Matrizen sind genau dann Hermite'sch kongruent, wenn sie denselben Typ haben.*

Beweis. Es seien $A, B \in \mathrm{M}_{n,n}(\mathbb{C})$ zwei Hermite'sche Matrizen. Es sei $\mathcal{B}$ die Standardbasis von $\mathbb{C}^n$.

Haben A und B denselben Typ, dann sind die Hermite'schen Formen $\sigma_1(u, v) = u^*Av$ und $\sigma_2(u, v) = u^*Bv$ auf $\mathbb{C}^n$ nach Satz 3.4.50 kongruent, haben also dieselben Matrixdarstellungen (siehe 3.4.31). Es gibt also eine geordnete Basis $\mathcal{C}$ von $\mathbb{C}^n$ dergestalt, dass $\mathrm{M}_{\mathcal{C}}(\sigma_2) = A$. Ist $S = {}_{\mathcal{B}}\mathrm{M}_{\mathcal{C}}(\mathrm{id}_{\mathbb{C}^n})$, dann gilt mit der Transformationsformel 3.4.25

$$S^*BS = A;$$

d.h., A und B sind Hermite'sch kongruent.

Nehmen wir umgekehrt an, dass $S^*BS = A$ für eine invertierbare Matrix $S \in \mathrm{GL}_n(\mathbb{C})$ gilt. Dann sind die Formen σ_1 und σ_2 kongruent und haben damit nach Satz 3.4.50 denselben Typ. $\square$

3.4.52 **Aufgabe**. Es sei **L**

$$A = \begin{pmatrix} a & z \\ \bar{z} & b \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{C})$$

eine Hermite'sche Matrix mit $a, b \in \mathbb{R}$ und $z \in \mathbb{C}$. Zeigen Sie: Ist $ab < 0$, dann hat A den Typ $(1, 1, 0)$.

3.4.53 **Bemerkung**. Genau wie für symmetrische reelle Bilinearformen können wir anhand des Typs einer Hermite’schen Form wichtige Eigenschaften ablesen. Sei σ eine Hermite’sche Form vom Typ (r_+, r_-, k) auf dem komplexen Vektorraum V der Dimension n. Dann gelten folgende Aussagen:

- σ ist regulär $\Leftrightarrow k = 0$,
- σ ist ausgeartet $\Leftrightarrow k > 0$,
- σ ist positiv definit $\Leftrightarrow r_+ = n$, $r_- = 0$, $k = 0$,
- σ ist negativ definit $\Leftrightarrow r_+ = 0$, $r_- = n$, $k = 0$.

Für den Beweis kann man auf die Beobachtung im Beweis des Trägheitssatzes 3.4.43 zurückgreifen, dass der Typ von $\mathrm{Re}(\sigma)$ genau $(2r_+, 2r_-, 2k)$ ist. Dann folgen die Aussagen aus 3.2.28 mithilfe von Satz 3.4.39 und Aufgabe 3.4.40.

3.5. Lösungen der Aufgaben in Lektion 3

L3.1.2 **Lösung.** Sei $\beta\colon U \times U \to K$ eine Bilinearform und sei $u \in U$. Es gilt

$$\beta(u,0) = \beta(u,0+0) = \beta(u,0) + \beta(u,0)$$

und damit $\beta(u,0) = 0$. Genauso zeigt man auch $\beta(0,u) = \beta(0,u) + \beta(0,u)$ und schließt $\beta(0,u) = 0$.

L3.1.6 **Lösung.** Für zwei Linearformen $\alpha_1, \alpha_2 \in U^*$ haben wir

$$(\alpha_1 \otimes \alpha_2)(u,v) \coloneqq \alpha_1(u)\alpha_2(v)$$

für alle $u, v \in V$ definiert. Für alle $u, v, v' \in U$ und $a \in K$ gilt dann

$$\begin{aligned}(\alpha_1 \otimes \alpha_2)(u, av + v') &= \alpha_1(u)\alpha_2(av + v') = \alpha_1(u)(a\alpha_2(v) + \alpha_2(v')) \\ &= a\alpha_1(u)\alpha_2(v) + \alpha_1(u)\alpha_2(v') \\ &= a(\alpha_1 \otimes \alpha_2)(u,v) + (\alpha_1 \otimes \alpha_2)(u,v').\end{aligned}$$

L3.1.12 **Lösung.** Wir betrachten die Bilinearform $\beta(u,v) = u^T A v$ aus Beispiel 3.1.7 für eine Matrix $A \in \mathrm{M}_{n,n}(K)$ und berechnen die Matrixdarstellung zur Standardbasis $\mathcal{B} = (e_1, \dots, e_n)$. Es seien dazu $i, j \in \{1, 2, \dots, n\}$. Der Eintrag an der Stelle k von e_i ist genau $\delta_{i,k}$; dabei ist $\delta_{s,t}$ das Kronecker-Symbol. Es gilt also

$$\beta(e_i, e_j) = e_i^T A e_j = \sum_{k=1}^{n} \sum_{\ell=1}^{n} \delta_{i,k} a_{k,\ell} \delta_{\ell,j} = a_{i,j}.$$

Das heißt, die Matrixdarstellung $\mathrm{M}_{\mathcal{B}}(\beta)$ und die Matrix A haben dieselben Einträge. Die beiden Matrizen sind also gleich.

L3.1.13 **Lösung.** Wir betrachten den Vektorraum $\mathbb{R}[X]_{\leq 2}$ der reellen Polynome vom Grad höchstens 2 mit der geordneten Basis $\mathcal{B} = (1, X, X^2)$ und der Bilinearform

$$\beta(f,g) = \int_0^1 fg \,\mathrm{d}x.$$

Aus [MG, 20.2.5] erhält man nun für alle $i, j \in \{0, 1, 2\}$ die Gleichung

$$\beta(X^i, X^j) = \int_0^1 x^{i+j} \,\mathrm{d}x = \frac{x^{i+j+1}}{i+j+1}\Big|_0^1 = \frac{1}{i+j+1}.$$

Damit ist

$$\mathrm{M}_{\mathcal{B}}(\beta) = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{pmatrix}$$

die Matrixdarstellung von β zur Basis $\mathcal{B}$.

L3.1.18 Lösung. Auf dem $\mathbb{R}^2$ bezeichne $\mathcal{B} = (e_1, e_2)$ die Standardbasis und $\mathcal{C} = (v_1, e_2)$ die Basis mit $v_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$. Die Transformationsmatrix von $\mathcal{C}$ nach $\mathcal{B}$ ist damit

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

Betrachten wir die Bilinearform β aus Beispiel 3.1.11, dann gilt mit der Transformationsformel

$$\mathrm{M}_{\mathcal{C}}(\beta) = S^T \mathrm{M}_{\mathcal{B}}(\beta) S = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 7 & -1 \\ 1 & 1 \end{pmatrix}.$$

L3.1.29 Lösung. Wir betrachten den Vektorraum $U = K^n$ mit einer Bilinearform $\beta(u, v) = u^T A v$. Sei $w \in U$. Für jedes $i \in \{1, \dots, n\}$ ist $\beta(e_i, w) = e_i^T A w$ genau der i-te Eintrag des Vektors Aw. Mit dieser Beobachtung erhalten wir

$$E(\iota_w^2(\beta)) = (\beta(e_1, w), \dots, \beta(e_n, w)) = (Aw)^T = w^T A^T.$$

L3.1.31 Lösung. Wir zeigen, dass $\beta^{(2)}$ linear ist mit dem Kriterium aus Aufgabe 2.2.18. Es seien $w, w' \in U$ und $a \in K$. Dann gilt für alle $u \in U$

$$\begin{aligned} \beta^{(2)}(aw + w')(u) &= \iota^2_{aw+w'}(\beta)(u) = \beta(u, aw + w') \\ &= a\beta(u, w) + \beta(u, w') && (\beta \text{ bilinear}) \\ &= a\iota_w^2(\beta)(u) + \iota_{w'}^2(\beta)(u) && (\text{Def. } \iota_w^2) \\ &= a\beta^{(2)}(w)(u) + \beta^{(2)}(w')(u). \end{aligned}$$

Der Vektor u war beliebig, also folgt daraus $\beta^{(2)}(aw + w') = a\beta^{(2)}(w) + \beta^{(2)}(w')$ für alle $w, w' \in U$ und $a \in K$ und $\beta^{(2)}$ ist damit linear.

L3.1.35 Lösung. Wir prüfen, dass $\varphi^*(\beta)$ eine Bilinearform ist. Es seien dazu $u, u', w \in U$ und $a \in K$. Dann gilt

$$\begin{aligned} \varphi^*(\beta)(au + u', w) &= \beta(\varphi(au + u'), \varphi(w)) \\ &= \beta(a\varphi(u) + \varphi(u'), \varphi(w)) && (\varphi \text{ linear}) \\ &= a\beta(\varphi(u), \varphi(w)) + \beta(\varphi(u'), \varphi(w)) && (\beta \text{ linear in erster Komp.}) \\ &= a\varphi^*(\beta)(u, w) + \varphi^*(\beta)(u', w). \end{aligned}$$

Also ist $\varphi^*(\beta)$ linear in der ersten Komponente.

Nun prüfen wir die Linearität in der zweiten Komponente. Sind $u, w, w' \in U$ und $a \in K$, dann hat man

$$\begin{aligned}
\varphi^*(\beta)(u, aw + w') &= \beta(\varphi(u), \varphi(aw + w')) \\
&= \beta(\varphi(u), a\varphi(w) + \varphi(w')) && (\varphi \text{ linear}) \\
&= a\beta(\varphi(u), \varphi(w)) + \beta(\varphi(u), \varphi(w')) && (\beta \text{ linear in zweiter Komp.}) \\
&= a\varphi^*(\beta)(u, w) + \varphi^*(\beta)(u, w').
\end{aligned}$$

L3.1.36 Lösung. Sei U ein endlich-dimensionaler K-Vektorraum und seien β_1, β_2 zwei Bilinearformen auf U.

Wir nehmen an, dass $\varphi^*(\beta_1) = \beta_2$ für einen Isomorphismus $\varphi\colon U \to U$ gilt. Sei $\mathcal{B} = (u_1, \ldots, u_n)$ eine geordnete Basis von U. Wir definieren $\mathcal{C} = (\varphi(u_1), \ldots, \varphi(u_n))$. Da φ ein Isomorphismus ist, ist $\mathcal{C}$ ebenfalls eine Basis von U; siehe [MG, 8.3.18]. Nach Voraussetzung gilt

$$\beta_2(u_i, u_j) = \varphi^*(\beta_1)(u_i, u_j) = \beta_1(\varphi(u_i), \varphi(u_j))$$

für alle $i, j \in \{1, \ldots, n\}$; d.h., $\mathrm{M}_{\mathcal{B}}(\beta_2) = \mathrm{M}_{\mathcal{C}}(\beta_1)$. Also sind β_1 und β_2 kongruent.

Nehmen wir umgekehrt an, dass β_1 und β_2 kongruent sind. Es gilt also $\mathrm{M}_{\mathcal{B}}(\beta_2) = \mathrm{M}_{\mathcal{C}}(\beta_1)$ für geordnete Basen $\mathcal{B} = (u_1, \ldots, u_n)$ und $\mathcal{C} = (v_1, \ldots, v_n)$. Es sei $\varphi\colon U \to U$ der eindeutige Isomorphismus mit $\varphi(u_i) = v_i$ für alle i. Dann gilt $\beta_1(\varphi(u_i), \varphi(u_j)) = \beta_2(u_i, u_j)$ für alle $i, j \in \{1, \ldots, n\}$. Damit haben β_2 und $\varphi^*(\beta_1)$ dieselbe Matrixdarstellung zur Basis $\mathcal{B}$. Aus Korollar 3.1.16 folgt damit $\varphi^*(\beta_1) = \beta_2$.

L3.1.41 Lösung. Wir betrachten die Linearform $\alpha_1\colon U \to K$. Nach dem Rangsatz gilt $\dim(U) = \mathrm{Rg}(\alpha_1) + \dim_K \mathrm{Ker}(\alpha_1)$. Da K ein ein-dimensionaler K-Vektorraum ist, gilt $\mathrm{Rg}(\alpha_1) \leq 1$. Aus der Voraussetzung $\dim_K(U) \geq 2$ folgt also $\dim_K \mathrm{Ker}(\alpha_1) > 0$, d.h., es gibt einen Vektor $u \neq 0$ mit $\alpha_1(u) = 0$. Für alle Vektoren $v \in V$ gilt dann aber

$$(\alpha_1 \otimes \alpha_2)(u, v) = \alpha_1(u)\alpha_2(v) = 0.$$

Somit ist Bedingung (i) aus Satz 3.1.38 nicht erfüllt und die Bilinearform $\alpha_1 \otimes \alpha_2$ ist ausgeartet.

L3.1.43 Lösung. Es sei V ein endlich-dimensionaler K-Vektorraum mit einer Bilinearform $\beta \in \mathrm{Bil}_K(V)$. Wir zeigen, dass

$$W = \{w \in V \mid \beta(u, w) = 0 \text{ für alle } u \in U\}$$

ein Unterraum von V ist. Es ist $\beta(u,v) = \beta^{(1)}(u)(v)$. Ein Vektor v ist damit genau dann ein Element von W, wenn er im Kern von $\beta^{(1)}(u)$ für jedes Element $u \in U$ liegt. Also gilt

$$W = \bigcap_{u \in U} \operatorname{Ker}(\beta^{(1)}(u)).$$

Der Kern einer linearen Abbildung ist ein Unterraum [MG, 8.3.8] und der Schnitt von Unterräumen ist wieder ein Unterraum [MG, 6.2.7], also ist W ein Unterraum von V.

(Alternativ kann man hier auch das Unterraumkriterium aus Aufgabe 2.2.6 anwenden.)

L3.1.46 **Lösung.** Sei β eine Bilinearform auf V. Ein Vektor $u \in V$ liegt genau dann im Kern von $\beta^{(1)}$, wenn $\beta^{(1)}(u)$ die Nullabbildung ist. Also genau dann, wenn $\beta^{(1)}(u)(v) = 0$ für alle $v \in V$ gilt. Die Gleichung $\beta^{(1)}(u)(v) = \beta(u,v)$ zeigt dann

$$\begin{aligned}\operatorname{Ker}(\beta^{(1)}) &= \{u \in V \mid \beta^{(1)}(u)(v) = 0 \text{ für alle } v \in V\ \} \\ &= \{u \in V \mid \beta(u,v) = 0 \text{ für alle } v \in V\ \} = N_\beta.\end{aligned}$$

L3.2.3 **Lösung.** Es sei $\beta \in \operatorname{Bil}_K(V)$. Ist β symmetrisch, dann gilt für alle $u, v \in V$ auch

$$\beta^{(1)}(u)(v) = \beta(u,v) = \beta(v,u) = \beta^{(2)}(u)(v).$$

Da $v \in V$ beliebig war, folgt $\beta^{(1)}(u) = \beta^{(2)}(u)$ für alle $u \in V$ und daraus folgt $\beta^{(1)} = \beta^{(2)}$.

Umgekehrt nehmen wir $\beta^{(1)} = \beta^{(2)}$ an. Für alle $u, v \in V$ gilt dann

$$\beta(u,v) = \beta^{(1)}(u)(v) = \beta^{(2)}(u)(v) = \beta(v,u),$$

d.h., β ist symmetrisch.

L3.2.7 **Lösung.** Es seien $A, S \in \mathrm{M}_{n,n}(K)$. Wir nehmen an, dass A symmetrisch ist. Mit den Rechenregeln aus 1.2.32 erhalten wir

$$\begin{aligned}(S^TAS)^T &= S^TA^T(S^T)^T && (1.2.32\ (ii)) \\ &= S^TAS. && (A = A^T \text{ und } 1.2.32\ (iv))\end{aligned}$$

Also ist S^TAS symmetrisch.

L3.2.14 Lösung. Gegeben ist die symmetrische Matrix $A = \begin{pmatrix} \overline{1} & \overline{2} & \overline{0} \\ \overline{2} & \overline{2} & \overline{1} \\ \overline{0} & \overline{1} & \overline{0} \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{F}_3)$. Diese Matrix liefert uns über die Formel

$$\beta(u, v) = u^T A v$$

eine symmetrische Bilinearform $\beta \in \mathrm{Bil}_{\mathbb{F}_3}(\mathbb{F}_3^3)$.

Gesucht ist eine invertierbare Matrix $S \in \mathrm{GL}_3(\mathbb{F}_3)$, sodass $S^T A S$ eine Diagonalmatrix ist. Dazu bestimmen wir eine geordnete Basis $\mathcal{B}$, sodass $\mathrm{M}_{\mathcal{B}}(\beta)$ eine Diagonalmatrix ist. Die gesuchte Matrix S ist die Transformationsmatrix von $\mathcal{B}$ in die Standardbasis, d.h., die Vektoren aus $\mathcal{B}$ sind die Spalten der Matrix S.

Wir berechnen zuerst den Rang von A. Mit dem Gauß-Verfahren erhalten wir

$$A = \begin{pmatrix} \overline{1} & \overline{2} & \overline{0} \\ \overline{2} & \overline{2} & \overline{1} \\ \overline{0} & \overline{1} & \overline{0} \end{pmatrix} \rightsquigarrow \begin{pmatrix} \overline{1} & \overline{2} & \overline{0} \\ \overline{0} & \overline{1} & \overline{1} \\ \overline{0} & \overline{1} & \overline{0} \end{pmatrix} \rightsquigarrow \begin{pmatrix} \overline{1} & \overline{2} & \overline{0} \\ \overline{0} & \overline{1} & \overline{0} \\ \overline{0} & \overline{0} & \overline{1} \end{pmatrix},$$

d.h., A hat den vollen Rang 3.

Wir setzen $W_1 = \mathbb{F}_3^3$ und suchen einen Vektor $v_1 \in W_1$ mit $\beta(v_1, v_1) \neq 0$. Wir wählen den ersten Standardbasisvektor $v_1 = e_1$, denn $\beta(v_1, v_1) = \overline{1}$. Mit einer kurzen Rechnung bestimmt man W_2 und erhält

$$W_2 = \{w \in W_1 \mid \beta(v_1, w) = 0\} = \left\{ \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} \in \mathbb{F}_3^3 \mid w_1 + \overline{2} w_2 = 0 \right\}$$
$$= \langle \begin{pmatrix} \overline{0} \\ \overline{0} \\ \overline{1} \end{pmatrix}, \begin{pmatrix} \overline{1} \\ \overline{1} \\ \overline{0} \end{pmatrix} \rangle.$$

Wir suchen einen Vektor $v_2 \in W_2$ mit $\beta(v_2, v_2) \neq 0$. Der dritte Standardbasisvektor e_3 ist *nicht* geeignet, denn es gilt $\beta(e_3, e_3) = 0$. Wir wählen daher $v_2 = \begin{pmatrix} \overline{1} \\ \overline{1} \\ \overline{0} \end{pmatrix}$. Dann gilt

$$\beta(v_2, v_2) = \begin{pmatrix} \overline{1} & \overline{1} & \overline{0} \end{pmatrix} \begin{pmatrix} \overline{1} & \overline{2} & \overline{0} \\ \overline{2} & \overline{2} & \overline{1} \\ \overline{0} & \overline{1} & \overline{0} \end{pmatrix} \begin{pmatrix} \overline{1} \\ \overline{1} \\ \overline{0} \end{pmatrix} = \begin{pmatrix} \overline{0} & \overline{1} & \overline{1} \end{pmatrix} \begin{pmatrix} \overline{1} \\ \overline{1} \\ \overline{0} \end{pmatrix} = \overline{1}.$$

Das Komplement W_3 ist definiert als

$$W_3 = \{w \in W_2 \mid \beta(v_2, w) = 0\} = \{w \in \mathbb{F}_3^3 \mid w_1 + \overline{2} w_2 = 0 \text{ und } w_2 + w_3 = 0\}.$$

Damit ist W_3 der Kern der Matrix $\begin{pmatrix}\overline{1} & \overline{2} & \overline{0}\\ \overline{0} & \overline{1} & \overline{1}\end{pmatrix}$. Mit dem Gauß-Verfahren findet man die Lösung

$$W_3 = \langle \begin{pmatrix}\overline{2}\\ \overline{2}\\ \overline{1}\end{pmatrix} \rangle.$$

Wir setzen $v_3 = \begin{pmatrix}\overline{2}\\ \overline{2}\\ \overline{1}\end{pmatrix}$. Mit einer kurzen Rechnung stellt man fest, dass $\beta(v_3, v_3) = \overline{2}$ ist. Damit definieren wir

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

denn dann gilt $S^T AS = \operatorname{diag}(\overline{1}, \overline{1}, \overline{2})$.

L3.2.19 **Lösung.** Sind die symmetrischen Matrizen $A, B \in \mathrm{M}_{n,n}(\mathbb{C})$ kongruent, dann gilt $B = S^T AS$ für eine invertierbare Matrix $S \in \mathrm{GL}_n(\mathbb{C})$. Aus [MG, 4.5.5] folgt, dass A und B denselben Rang haben.

Angenommen A und B haben denselben Rang, dann haben die symmetrischen Bilinearformen $\beta_1 \colon (x, y) \mapsto x^T Ay$ und $\beta_2 \colon (x, y) \mapsto x^T By$ auf $\mathbb{C}^n$ denselben Rang, denn A und B sind die zugehörigen Matrixdarstellungen zur Standardbasis; siehe 3.1.12 und 3.1.44. Durch Satz 3.2.18 wissen wir, dass β_1 und β_2 kongruent sind. Somit sind auch ihre Matrixdarstellungen A und B kongruent (siehe 3.1.23).

L3.2.27 **Lösung.** Wir betrachten

$$A_t = \begin{pmatrix}t & 2\\ 2 & -1\end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{R})$$

mit einem Parameter $t \in \mathbb{R}$. Wir führen das Verfahren aus 3.2.12 durch.

Zuerst wählen wir den Vektor $v_1 = \begin{pmatrix}0\\ 1\end{pmatrix}$. Es gilt $\beta_t(v_1, v_1) = -1 \neq 0$ unabhängig von t. Dann berechnen wir das Komplement:

$$\begin{aligned} W_2 &= \{w \in \mathbb{R}^2 \mid \beta_t(v_1, w) = 0\}\\ &= \operatorname{Ker}(v_1^T A_t) = \operatorname{Ker}\begin{pmatrix}2 & -1\end{pmatrix}\\ &= \langle \begin{pmatrix}1\\ 2\end{pmatrix} \rangle. \end{aligned}$$

Wir setzen nun $v_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$. Dann gilt

$$\beta_t(v_2, v_2) = v_2^T A v_2 = \begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} t & 2 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = t + 4.$$

Für die Basis $\mathcal{B} = (v_1, v_2)$ erhalten wir also

$$\mathrm{M}_{\mathcal{B}}(\beta_t) = \begin{pmatrix} -1 & 0 \\ 0 & t+4 \end{pmatrix}$$

für alle $t \in \mathbb{R}$.

Der Typ von β_t hängt also nur vom Vorzeichen von $t + 4$ ab.

- Ist $t > -4$, dann hat β_t den Typ $(1, 1, 0)$.
- Ist $t = -4$, dann hat β_t den Typ $(0, 1, 1)$.
- Ist $t < -4$, dann hat β_t den Typ $(0, 2, 0)$.

L3.2.32 Lösung. Für die Basis $\mathcal{B} = (v_1, v_2)$ aus der Lösung von Aufgabe 3.2.27 haben wir die Matrixdarstellung

$$\mathrm{M}_{\mathcal{B}}(\beta_t) = \begin{pmatrix} -1 & 0 \\ 0 & t+4 \end{pmatrix}$$

(in Abhängigkeit von $t \in \mathbb{R}$) gefunden.

Je nach Vorzeichen von $t+4$ müssen wir die Basisvektoren umordnen und schließlich entsprechend strecken.

Falls $t > -4$ ist, dann ist $t + 4$ positiv und wir definieren

$$w_1 = \frac{1}{\sqrt{t+4}} v_2 \quad \text{und} \quad w_2 = v_1.$$

Falls $t = -4$ ist, dann ist $t + 4 = 0$ und wir definieren

$$w_1 = v_1 \quad \text{und} \quad w_2 = v_2.$$

Falls $t < -4$ ist, dann ist $t + 4 < 0$ und wir definieren

$$w_1 = v_1 \quad w_2 = \frac{1}{\sqrt{-t-4}} v_2.$$

In jedem der drei Fälle ist (w_1, w_2) eine geordnete Basis, sodass die Matrixdarstellung von β_t in Normalform ist.

L3.2.35 Lösung. Da es zu jedem möglichen Typ (r_+, r_-, k) eine Matrix von diesem Typ gibt (nämlich $\mathrm{diag}(I_{r_+}, -I_{r_-}, 0_k)$) müssen wir nur die Anzahl der möglichen Typen bestimmen. Ist (r_+, r_-, k) der Typ einer symmetrischen (4×4)-Matrix, dann gilt $4 = r_+ + r_- + k$. In Frage kommen also die Typen

$$\begin{aligned}
&(4, 0, 0),\\
&(3, 1, 0), (3, 0, 1),\\
&(2, 2, 0), (2, 1, 1), (2, 0, 2),\\
&(1, 3, 0), (1, 2, 1), (1, 1, 2), (1, 0, 3),\\
&(0, 4, 0), (0, 3, 1), (0, 2, 2), (0, 1, 3), (0, 0, 4)
\end{aligned}$$

und damit gibt es insgesamt 15 verschiedene Kongruenzklassen.

L3.3.4 Lösung. Es sei $\beta \in \mathrm{Bil}_K(V)$. Ist β schiefsymmetrisch, dann gilt für alle $u, v \in V$ auch

$$\beta^{(1)}(u)(v) = \beta(u, v) = -\beta(v, u) = -\beta^{(2)}(u)(v).$$

Da $v \in V$ beliebig war, folgt $\beta^{(1)}(u) = -\beta^{(2)}(u)$ für alle $u \in V$ und daraus folgt $\beta^{(1)} = -\beta^{(2)}$.

Umgekehrt nehmen wir $\beta^{(1)} = -\beta^{(2)}$ an. Für alle $u, v \in V$ gilt dann

$$\beta(u, v) = \beta^{(1)}(u)(v) = -\beta^{(2)}(u)(v) = -\beta(v, u),$$

d.h., β ist schiefsymmetrisch.

L3.3.14 Lösung. Wieviele Kongruenzklassen alternierender (6×6)-Matrizen gibt es? Dazu müssen wir die möglichen Ränge zählen. Dabei müssen wir nur beachten, dass der Rang einer alternierenden Matrix gerade ist. In Frage kommen also nur die Ränge $0, 2, 4, 6$. Zu jedem dieser Ränge gibt es eine alternierende Matrix, z.B., die Matrix in Normalform. Es gibt also insgesamt 4 Kongruenzklassen.

L3.4.6 Lösung. Es sei σ eine positiv definite Hermite'sche Form auf V. Sei $u \neq 0$ in V gegeben. Dann gilt

$$\sigma(u, u) > 0$$

und damit $\sigma(u, u) \neq 0$. Also ist σ regulär.

L3.4.7 Lösung. Wir nehmen an, dass σ regulär ist. Sei $v \neq 0$ in V gegeben. Die Regularität liefert uns einen Vektor $u \in V$ mit $\sigma(v, u) \neq 0$. Damit folgt aber $\sigma(u, v) = \overline{\sigma(v, u)} \neq 0$.

Um umgekehrt die Regularität aus der genannten Bedingung abzuleiten, argumentiert man analog.

L3.4.15 Lösung. Ist $A = \begin{pmatrix} 3i & 0 \\ 3+i & 2-i \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{C})$, dann ist

$$A^* = \begin{pmatrix} -3i & 3-i \\ 0 & 2+i \end{pmatrix}.$$

L3.4.16 Lösung. Es seien $A \in \mathrm{M}_{\ell,m}(\mathbb{C})$, $B \in \mathrm{M}_{m,n}(\mathbb{C})$ für $\ell, m, n \in \mathbb{N}$. Mit den Rechenregeln aus 3.4.13 und 1.2.32 erhalten wir

$$(AB)^* = (\overline{AB})^T = (\overline{A}\,\overline{B})^T = \overline{B}^T\overline{A}^T = B^*A^*$$

und auch

$$(A^*)^* = \overline{(\overline{A}^T)}^T = \overline{(\overline{A}^T)^T} = \overline{\overline{A}} = A.$$

Nehmen wir nun an, dass $C \in \mathrm{M}_{n,n}(\mathbb{C})$ invertierbar ist. Aus der Formel $(AB)^* = B^*A^*$ schließen wir

$$(C^{-1})^*C^* = (CC^{-1})^* = I_n^* = I_n$$

und genauso

$$C^*(C^{-1})^* = (C^{-1}C)^* = I_n^* = I_n.$$

Also ist C^* invertierbar und $(C^{-1})^*$ ist das Inverse, d.h., $(C^{-1})^* = (C^*)^{-1}$.

L3.4.24 Lösung. Sei $\mathcal{B} = (e_1, \ldots, e_n)$ die Standardbasis des $\mathbb{C}^n$. Sei σ eine Hermite'sche Form auf $\mathbb{C}^n$ und es sei $A = \mathrm{M}_{\mathcal{B}}(\sigma)$ die Matrixdarstellung von σ zur Standardbasis. Wir wissen aus 3.4.21, dass A Hermite'sch ist. Für alle $u \in \mathbb{C}^n$ gilt $\kappa_{\mathcal{B}}(u) = u$ und mit Satz 3.4.23 folgt daraus

$$\sigma(u, v) = u^*Av$$

für alle $u, v \in \mathbb{C}^n$.

L3.4.27 Lösung. Wir prüfen die drei Eigenschaften aus 2.6.1 nach. Es seien $A, B, C \in \mathrm{M}_{n,n}(\mathbb{C})$ Hermite'sche Matrizen.

reflexiv: Es gilt $A \sim A$, denn mit $S = I_n$ haben wir $S^*AS = I_nAI_n = A$.

symmetrisch: Gilt $A \sim B$, dann gibt es $S \in \mathrm{GL}_n(\mathbb{C})$ mit $A = S^*BS$. Da S invertierbar ist, gilt dann auch $B = (S^*)^{-1}AS^{-1}$. Die Matrix $T = S^{-1}$ ist invertierbar und aus 3.4.16 wissen wir, dass $(S^{-1})^* = (S^*)^{-1}$ gilt. Damit haben wir $B = T^*AT$ und es gilt $B \sim A$.

transitiv: Angenommen es ist $A \sim B$ und $B \sim C$. Es gibt also $S_1, S_2 \in \mathrm{GL}_n(\mathbb{C})$ mit $A = S_1^* B S_1$ und $B = S_2^* C S_2$. Mit 3.4.16 (a) folgt

$$A = S_1^* B S_1 = S_1^* S_2^* C S_2 S_1 = S^* C S$$

mit $S = S_2 S_1$. Da $\mathrm{GL}_n(\mathbb{C})$ eine Gruppe ist (siehe 1.1.12), ist auch $S = S_2 S_1$ invertierbar.

L3.4.29 Lösung. Es seien $A, B \in \mathrm{M}_{n,n}(\mathbb{C})$ zwei Hermite'sche Matrizen. Wir nehmen an, dass A und B Hermite'sch kongruent sind, d.h., es gilt $A = S^* B S$ für eine invertierbare Matrix $S \in \mathrm{GL}_n(\mathbb{C})$. Wir erinnern uns, dass dann auch $S^* = \overline{S}^T$ invertierbar ist; siehe Aufgabe 3.4.16. Die Multiplikation mit invertierbaren Matrizen ändert den Rang nicht (siehe [MG, 4.5.5]), deshalb folgt

$$\mathrm{Rg}(A) = \mathrm{Rg}(S^* B S) = \mathrm{Rg}(B).$$

L3.4.32 Lösung. Wir zeigen, dass Kongruenz von Hermite'schen Formen eine Äquivalenzrelation auf $\mathrm{Herm}(V)$ definiert.

Die Relation ist reflexiv, denn mit $\mathcal{B} = \mathcal{B}_1 = \mathcal{B}_2$ gilt $\mathrm{M}_{\mathcal{B}_1}(\sigma) = \mathrm{M}_{\mathcal{B}_2}(\sigma)$ für alle $\sigma \in \mathrm{Herm}(V)$.

Die Relation ist symmetrisch. Seien σ_1, σ_2 kongruent. Dann folgt aus $\mathrm{M}_{\mathcal{B}_1}(\sigma_1) = \mathrm{M}_{\mathcal{B}_2}(\sigma_2)$ auch $\mathrm{M}_{\mathcal{B}_2}(\sigma_2) = \mathrm{M}_{\mathcal{B}_1}(\sigma_1)$ und damit sind σ_2 und σ_1 kongruent.

Wir zeigen, dass die Relation auch transitiv ist. Es seien dazu $\sigma_1, \sigma_2, \sigma_3 \in \mathrm{Herm}(V)$. Wir nehmen an, dass σ_1 und σ_2, sowie σ_2 und σ_3 kongruent sind. Sei $\mathcal{B}$ eine geordnete Basis. Nach Lemma 3.4.28 sind $\mathrm{M}_{\mathcal{B}}(\sigma_1)$ und $\mathrm{M}_{\mathcal{B}}(\sigma_2)$, sowie $\mathrm{M}_{\mathcal{B}}(\sigma_2)$ und $\mathrm{M}_{\mathcal{B}}(\sigma_3)$ Hermite'sch kongruent. Da Hermite'sche Kongruenz von Matrizen eine Äquivalenzrelation ist (siehe 3.4.27), sind damit $\mathrm{M}_{\mathcal{B}}(\sigma_1)$ und $\mathrm{M}_{\mathcal{B}}(\sigma_3)$ Hermite'sch kongruent. Aus Lemma 3.4.28 schließen wir, dass σ_1 und σ_3 kongruent sind. Die Relation ist also transitiv.

L3.4.40 Lösung. Sei $\sigma \in \mathrm{Herm}(V)$. Für alle $v \in V$ ist $\sigma(v,v)$ reell und damit gilt $\mathrm{Re}(\sigma)(v,v) = \sigma(v,v)$. Es gilt also $\sigma(v,v) > 0$ genau dann, wenn $\mathrm{Re}(\sigma)(v,v) > 0$ ist.

L3.4.49 Lösung. Wir betrachten die Hermite'sche Matrix $A = \begin{pmatrix} -4 & 1+i \\ 1-i & 9 \end{pmatrix}$ und die Hermite'sche Form $\sigma(u,v) = u^* A v$ auf $\mathbb{C}^2$.

Wir berechnen die Normalform. Wir stellen fest, dass A den Rang 2 hat, denn die Zeilen sind nicht linear abhängig. Wir suchen einen Vektor $v_1' \in \mathbb{C}^2$ mit $\sigma(v_1', v_1') \neq 0$.

Wir nehmen den ersten Standardbasisvektor $v_1' = e_1$, denn es ist $\sigma(v_1', v_1') = -4$. Wir setzen dann $v_1 = \frac{1}{2}v_1'$ um $\sigma(v_1, v_1) = -1$ zu erreichen. Als nächstes bestimmen wir das Komplement:

$$\begin{aligned} W_2 &= \{w \in \mathbb{C}^2 \mid \sigma(v_1', w) = 0\} = \{w \in \mathbb{C}^2 \mid (v_1')^* Aw = 0\} \\ &= \operatorname{Ker}((v_1')^* A) = \operatorname{Ker}(\begin{pmatrix}-4 & 1+i\end{pmatrix}) \\ &= \langle \begin{pmatrix} 1 \\ 2-2i \end{pmatrix} \rangle \end{aligned}$$

Wir setzen $v_2' = \begin{pmatrix} 1 \\ 2-2i \end{pmatrix}$. Dann gilt

$$(v_2')^* A v_2' = \begin{pmatrix}1 & 2+2i\end{pmatrix} \begin{pmatrix} -4 & 1+i \\ 1-i & 9 \end{pmatrix} \begin{pmatrix} 1 \\ 2-2i \end{pmatrix} = 76.$$

Wir setzen nun $v_2 = \frac{1}{\sqrt{76}} v_2'$. Die Vektoren sind noch in der falschen Reihenfolge, denn $\sigma(v_1, v_1) = -1$ aber $\sigma(v_2, v_2) = 1$. Also ist $\mathcal{B} = (v_2, v_1)$ eine geordnete Basis, sodass $\mathrm{M}_{\mathcal{B}}(\sigma) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ eine Darstellung in Normalform ist. Der Typ der Hermite'schen Form ist also $(1, 1, 0)$.

L3.4.52 **Lösung.** Es sei

$$A = \begin{pmatrix} a & z \\ \overline{z} & b \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{C})$$

eine Hermite'sche Matrix mit $a, b \in \mathbb{R}$ und $z \in \mathbb{C}$ und $ab < 0$. Das heißt, a und b haben entgegengesetzte Vorzeichen. Sagen wir a sei positiv und b sei negativ. Der andere Fall funktioniert genauso. Es sei $\sigma(u, v) = u^* A v$ die zugehörige Hermite'sche Form.

Berechnen wir eine diagonale Matrixdarstellung von σ und wählen als ersten Vektor e_1, dann erhalten wir eine Matrixdarstellung der Form

$$\begin{pmatrix} a & 0 \\ 0 & x \end{pmatrix}$$

mit einem unbekannten Eintrag $x \in \mathbb{R}$. Da $a > 0$ ist, schließen wir $r_+ \geq 1$.

Berechnen wir eine diagonale Matrixdarstellung von σ und wählen als ersten Vektor e_2, dann erhalten wir eine Matrixdarstellung der Form

$$\begin{pmatrix} b & 0 \\ 0 & y \end{pmatrix}$$

mit einem unbekannten Eintrag $y \in \mathbb{R}$. Da $b < 0$ ist, schließen wir $r_- \geq 1$.

Da $r_+ + r_- = r \leq 2$ ist, ist damit $(1, 1, 0)$ der einzige mögliche Typ.

Steffen Kionke

Lineare Algebra

Lektion 4:
Determinanten

Fakultät für
Mathematik und Informatik

FernUniversität in Hagen

Studierhinweise zur vierten Lektion

In dieser Lektion befassen wir uns ausführlich mit einem zentralen Hilfsmittel der Linearen Algebra: der Determinante. Die Determinante einer quadratischen Matrix $A \in \mathrm{M}_{n,n}(R)$ über einem kommutativen Ring R ist ein Ringelement $\det(A) \in R$.

Die Determinante hat viele Anwendungen. Beispielsweise kann man mit $\det(A)$ entscheiden, ob die Matrix A invertierbar ist. Wir werden sie auch in der nächsten Lektion benötigen, um das charakteristische Polynom einer Matrix zu definieren. Sie werden die Determinante aber auch in anderen Modulen wiedertreffen. Unter anderem ist die Determinante eng mit dem Volumen in der Maßtheorie oder Differentialgeometrie verknüpft. Diese Zusammenhänge können wir hier aber nur andeuten.

Wir werden die Determinante einer Matrix zunächst mit der komplizierten *Leibniz-Formel* definieren und uns dann auf den Weg machen, die Determinante und ihre Eigenschaften zu verstehen. Insbesondere werden wir verschiedene Methoden zur Berechnung und einige Anwendungen kennenlernen. Wir werden sehen, dass man die Determinante von Matrizen auf lineare Abbildungen (genauer: Endomorphismen endlich-dimensionaler Vektorräume) übertragen kann.

Zum Abschluss der Lektion werden wir sehen, dass die Determinante nicht nur Matrizenrechnung ist, sondern, dass man die Determinante koordinatenfrei definieren und verstehen kann. Dazu wagen wir einen Blick auf die *multilinearen Abbildungen* und die *alternierenden Multilinearformen*. Dabei handelt es sich um eine Verallgemeinerung von Bilinearformen auf Abbildungen mit vielen Komponenten. Dieser Abschnitt ist relativ abstrakt und wird später im Studienbrief auch keine Rolle mehr spielen. Sie sollten sich hier nicht festbeißen, denn Abschnitt 4.5 ist nicht klausurrelevant. Die anderen Teile dieser Lektion sind viel wichtiger.

Dieses Kapitel ist wichtig. Ich glaube aber, dass diese Lektion insgesamt etwas einfacher ist. Wenn Sie sich die Hände schmutzig machen und viele Determinanten berechnen, werden Sie schnell ein Gefühl dafür bekommen.

Lernziele

Nach dem Durcharbeiten der Lektion sollten Sie

→ die *symmetrischen Gruppen* kennen und mit *Permutationen* rechnen können,

→ mit dem *Satz von der Signatur* vertraut sein,

- → die *Leibniz-Formel* für die Determinante kennen und anwenden können,
- → mit dem *Charakterisierungssatz* und dem *Multiplikationssatz* für die Determinante umgehen können,
- → die *Determinante*, *Minoren* und die *Adjunkte* von Matrizen berechnen können,
- → den *Adjunktensatz* und seine Anwendungen kennen,
- → wissen, was die Determinante eines *Endomorphismus* ist.

Literaturhinweise

Die Determinante wird in den Lehrbüchern zur Linearen Algebra meist ausführlich besprochen, sodass Sie das meiste problemlos nachschlagen können. Der koordinatenfreie Zugang über alternierende Multilinearformen wird allerdings nur bei [Bo, 4.5] thematisiert.

- → [Bär]: 4.3 (Determinante).
- → [Beu]: Kapitel 7.
- → [Bo]: Kapitel 4.
- → [Fi]: Kapitel 4.
- → [Gö]: 2.6 (Determinante), 2.7 (Adjunkte).
- → [KaSt]: Kapitel 7.

Fahrplan durch die Lektion

4.1 → Im vorbereitenden Abschnitt 4.1 befassen wir uns mit den *symmetrischen Gruppen* und *Permutation*, die wir später zur Definition der Determinante benötigen. Im ersten Teil definieren wir die symmetrischen Gruppen S_n (4.1.1, 4.1.2) und zeigen, dass S_n genau $n!$ viele Elemente – genannt *Permutationen* – besitzt (4.1.3). Im zweiten Teil üben wir das Rechnen mit Permutationen(4.1.5–4.1.10). Wir definieren den *Träger* einer Permutation (4.1.11) und lernen *Transpositionen* kennen (4.1.14). Jede Permutation ist eine Verkettung von Transpositionen (4.1.16).

Am wichtigsten ist der dritte Teil. Hier definieren wir die Signatur (4.1.20) über die Anzahl der Fehlstände (4.1.18). Im Satz von der Signatur (4.1.24) lernen wir eine andere Formel kennen und sehen, dass die Signatur mit der Komposition verträglich ist.

In Abschnitt 4.2 definieren wir die Determinante einer Matrix und lernen wichtige Eigenschaften kennen. ← 4.2

Im ersten Teil beginnen wir mit der unhandlichen *Leibniz-Formel* (4.2.1) und versuchen diese an kleinen Matrizen zu verstehen (4.2.3, 4.2.6). Für 2×2-Matrizen, besprechen wir kurz den Zusammenhang von Determinante und Flächeninhalt (4.2.5). Wir leiten aus der Leibniz-Formel eine Formel für die Determinante von oberen Dreiecksmatrizen ab (4.2.13).

Im zweiten Teil beweisen wir den wichtigen *Charakterisierungssatz* (4.2.17), der die Determinante eindeutig durch drei Eigenschaften beschreibt. Der Beweis besteht im wesentlichen aus geschickten Umformungen der Leibniz-Formel. Aus dem Charakterisierungssatz leiten wir weitere Eigenschaften der Determinante ab (4.2.20).

Den fundamentalen *Determinantenmultiplikationssatz* $\det(AB) = \det(A)\det(B)$ (4.2.21) beweisen wir im dritten Teil. Der Beweis ist etwas anspruchsvoller, aber lassen Sie sich nicht abschrecken: Es handelt sich wieder nur um geschickte Umformungen der Leibniz-Formel. Als Anwendungen zeigen wir $\det(A^{-1}) = \det(A)^{-1}$ (4.2.22) und weisen nach, dass *ähnliche Matrizen* dieselbe Determinante haben (4.2.26).

Im letzten Teil lernen wir, wie man mit dem Gauß-Verfahren die Determinante großer Matrizen mit Einträgen aus einem Körper berechnen kann (4.2.28). Als Anwendung zeigen wir, dass eine Matrix über einem Körper genau dann invertierbar ist, wenn $\det(A) \neq 0$ ist (4.2.32).

Nicht nur die Determinante einer Matrix ist hilfreich, sondern auch die *Determinanten von Untermatrizen*: die *Minoren*. Diese studieren wir in Abschnitt 4.3. Im ersten ← 4.3
Teil besprechen wir kurz die Definition (4.3.1) und geben Beispiele (4.3.2–4.3.6). Im zweiten Teil beweisen wir den sehr nützlichen *Entwicklungssatz von Laplace* (4.3.8). Für Matrizen mit vielen verschwindenden Einträgen liefert der Entwicklungssatz ein gutes Verfahren, um die Determinante zu berechnen. Es lohnt sich den Beweis auszuarbeiten, denn er basiert auf dem Charakterisierungssatz. So bekommen Sie hoffentlich einen Eindruck, warum es nützlich sein kann, ein mathematisches Objekt eindeutig durch seine Eigenschaften zu beschreiben. Als Anwendung geben wir eine Formel für die Determinante von Blockdreiecksmatrizen an (4.3.15).

Im dritten Teil definieren wir die *Adjunkte* einer Matrix (4.3.17) und wir beweisen den *Adjunktensatz*. Als Anwendungen werden wir die Invertierbarkeit von Matrizen über kommutativen Ringen mithilfe der Determinante charakterisieren (4.3.24) und die *Cramer'sche Regel* zum Lösen linearer Gleichungssysteme herleiten (4.3.28).

4.4 → In Abschnitt 4.4 bewegen wir uns von der Matrizenrechnung weg und hin zur abstrakten Linearen Algebra. Eine lineare Abbildung $\varphi\colon V \to V$ nennt man auch *Endomorphismus* (4.4.1). In diesem Abschnitt definieren wir für Endomorphismen auf endlich-dimensionalen Vektorräumen eine Determinante (4.4.6). Dabei verwenden wir die Matrixdarstellung und zeigen mit der *Transformationsformel* für Endomorphismen (4.4.5), dass die Determinante nicht von der Wahl einer Basis abhängt.

4.5 → Im letzen fakultativen Abschnitt 4.5 gehen wir der Frage nach, ob man die Determinante auch komplett ohne Basen (also „koordinatenfrei“) definieren kann. Wir werfen einen kurzen, oberflächlichen Blick auf *multilineare Abbildungen* (4.5.2). Die Menge der n-linearen Abbildungen bildet einen Vektorraum (4.5.4). Das Einsetzen von Vektoren (4.5.5) und das „Zurückziehen“ mit linearen Abbildungen (4.5.7) liefern lineare Abbildungen zwischen den Räumen multilinearer Abbildungen (4.5.6, 4.5.8).

Im zweiten Teil diskutieren wir die *alternierenden Multilinearformen* (4.5.10). Wir sehen uns Beispiele an (4.5.11–4.5.13) und beweisen, dass die Menge $\bigwedge^n V^*$ der alternierenden n-linearen Formen ein Vektorraum ist (4.5.14). Die entscheidende Beobachtung ist, dass $\bigwedge^n V^*$ ein-dimensional ist, falls $\dim V = n$ ist (4.5.17). Im letzten Teil beweisen wir dann den entscheidenen Satz 4.5.18, aus dem man eine alternative, koordinatenfreie Definition der Determinante ableiten kann.

4.1. Die symmetrischen Gruppen

In diesem vorbereitenden Abschnitt werden wir uns mit den *symmetrischen Gruppen* vertraut machen. Dabei handelt es sich um eine wichtige Familie von Gruppen. Wir werden in den nächsten Abschnitten die Determinante kennenlernen und dabei sehen, dass die symmetrischen Gruppen auch in der Linearen Algebra von Bedeutung sind.

I. Definition

4.1.1 **Satz.** *Es seien X eine Menge und*

$$\mathrm{Sym}(X) = \{\sigma\colon X \to X \mid \sigma \text{ bijektiv } \}$$

die Menge der bijektiven Selbstabbildungen von X. Dann bildet $\mathrm{Sym}(X)$ mit der Komposition $\circ$ von Abbildungen eine Gruppe. Das neutrale Element ist die identische Abbildung id_X. Das Inverse von σ ist die Umkehrabbildung σ^{-1}.

Beweis. Wir erinnern uns, dass die Komposition bijektiver Abbildungen stets eine bijektive Abbildung ergibt (siehe [MG, 1.4.16]). Wir prüfen die vereinfachten Gruppenaxiome aus 1.1.17 nach.

(G1): Aus dem Modul „Mathematische Grundlagen" ist bekannt, dass die Komposition von Abbildungen das Assoziativgesetz erfüllt; siehe [MG, (1.4.23)].

(G2'): Die identische Abbildung id_X ist das neutrale Element, denn für alle $\sigma \in \mathrm{Sym}(X)$ und $x \in X$ gilt

$$\mathrm{id}_X \circ \sigma(x) = \mathrm{id}_X(\sigma(x)) = \sigma(x)$$

und somit $\mathrm{id}_X \circ \sigma = \sigma$.

(G3'): Es sei $\sigma \in \mathrm{Sym}(X)$. Da σ bijektiv ist, gibt es eine Umkehrabbildung $\sigma^{-1}\colon X \to X$. Die Umkehrabbildung ist wieder bijektiv, d.h. $\sigma^{-1} \in \mathrm{Sym}(X)$, und es gilt $\sigma^{-1} \circ \sigma = \mathrm{id}_X$ nach Definition der Umkehrabbildung. □

4.1.2 **Definition** Die Gruppe $(\mathrm{Sym}(X), \circ)$ heißt die *symmetrische Gruppe* von X. Für $n \in \mathbb{N}$ nennt man

$$S_n := \mathrm{Sym}(\{1, 2, \dots, n\})$$

die *symmetrische Gruppe auf n Buchstaben.*

4.1.3 **Satz.** *Es sei $n \in \mathbb{N}$. Dann gilt $|S_n| = n!$.*

Beweis. Jede injektive Abbildung $\sigma\colon \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$ ist bereits bijektiv, da das Bild aus n verschiedenen Elementen bestehen muss. Es genügt also die injektiven Selbstabbildungen von $\{1, 2, \ldots, n\}$ zu zählen.

Für $\sigma(1)$ gibt es n verschiedene Möglichkeiten. Ist $\sigma(1)$ festgelegt, dann muss $\sigma(2) \neq \sigma(1)$ sein und für $\sigma(2)$ bleiben noch $n-1$ verschiedene mögliche Werte. Für $\sigma(3)$ gibt es dann noch $n-2$ verschiedene Möglichkeiten, und so fährt man fort. Man erhält

$$|S_n| = n \cdot (n-1) \cdot (n-2) \cdots 2 \cdot 1 = n!.$$

□

4.1.4 **Beispiel**. Die symmetrische Gruppe S_2 besteht aus zwei Elementen: der identischen Abbildung id und einer Permutation σ mit $\sigma(1) = 2$ und $\sigma(2) = 1$. Man nennt σ eine Transposition.

II. Rechnen mit Permutationen

4.1.5 **Permutationen**. Sei $n \in \mathbb{N}$. Die Elemente von S_n heißen *Permutationen* und werden meistens mit den kleinen griechischen Buchstaben σ, τ, π notiert. Das Wort Permutation geht auf das lateinische „permutare" („vertauschen") zurück und so kann man sich Permutationen am besten als Vertauschungen der Zahlen $1, 2, \ldots, n$ denken.

Die Permutation $\sigma \in S_n$ bildet $i \in \{1, 2, \ldots, n\}$ auf $\sigma(i) \in \{1, 2, \ldots, n\}$ ab und wir schreiben diese Permutation als kleine „Zuordnungstabelle"

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ \sigma(1) & \sigma(2) & \sigma(3) & \ldots & \sigma(n) \end{pmatrix}$$

Man kann also aus der ersten Spalte $\sigma(1)$ ablesen, aus der zweiten Spalte $\sigma(2)$, und so weiter.

Zum Beispiel ist

$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 2 & 1 \end{pmatrix}$$

eine Permutation in der Gruppe S_4, die die Zahlen in $\{1, 2, 3, 4\}$ wie folgt vertauscht:

$$\tau(1) = 3, \quad \tau(2) = 4, \quad \tau(3) = 2, \quad \tau(4) = 1.$$

4.1.6 **Aufgabe**. Es sei $\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 5 & 3 & 1 \end{pmatrix} \in S_5$. Bestimmen Sie $\tau(3)$. **L**

4.1.7 **Komposition von Permutationen**. Aus Satz 4.1.1 wissen wir, dass S_n mit der Komposition $\circ$ eine Gruppe bildet. Um in dieser Gruppe rechnen zu können, müssen wir uns überlegen, wie die Komposition $\circ$ eigentlich aus der Tabellenschreibweise berechnet werden kann.

Seien $\sigma, \tau \in S_n$ und $i \in \{1, 2, \ldots, n\}$. Um $\sigma \circ \tau(i)$ zu bestimmen, liest man erst aus der i-ten Spalte von τ den Wert $j = \tau(i)$ ab und liest dann in der j-ten Spalte von σ den Wert $\sigma(j)$ ab. Dies ist das gesuchte Ergebnis, denn es gilt

$$\sigma \circ \tau(i) = \sigma(\tau(i)) = \sigma(j).$$

Sehen wir uns dazu ein Beispiel an. Es seien

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix} \quad \text{und} \quad \tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix}$$

in der Gruppe S_4 gegeben. Wir wollen $\sigma \circ \tau$ bestimmen. Dazu beginnen wir mit der 1 und lesen ab, dass $\tau(1) = 2$ gilt. Jetzt schauen wir in der zweiten Spalte von σ und finden $\sigma(2) = 3$. Wir erhalten also $\sigma \circ \tau(1) = 3$, d.h.,

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & & & \end{pmatrix}.$$

Nun bestimmt man auf dieselbe Weise $\sigma \circ \tau(2)$, $\sigma \circ \tau(3)$ und $\sigma \circ \tau(4)$ und erhält

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}$$

In der folgenden Aufgabe können Sie das Rechnen mit Permutationen üben.

4.1.8 **Aufgabe**. Bestimmen Sie die folgenden Kompositionen in der Gruppe S_5: **L**

(a) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 5 & 3 & 4 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 1 & 2 & 5 \end{pmatrix}$,

(b) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 3 & 5 & 4 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 2 & 1 \end{pmatrix}$.

4.1.9 **Invertieren von Permutationen**. Aus Satz 4.1.1 ist bekannt, dass S_n eine Gruppe ist. Es gibt also zu jeder Permutation $\sigma \in S_n$ eine inverse Permutation σ^{-1}. Wie findet man σ^{-1}, wenn σ in Tabellenschreibweise gegeben ist?

Sei $i \in \{1, 2, \ldots, n\}$. Um $\sigma^{-1}(i)$ zu bestimmen suchen wir den Eintrag i in der zweiten Zeile von σ. Steht dieser Eintrag in Spalte j, dann gilt

$$\sigma^{-1}(i) = \sigma^{-1}(\sigma(j)) = j.$$

Man liest hier also σ von unten nach oben ab. Das liefert uns ein einfaches Verfahren um σ^{-1} zu bestimmen:

(1) Vertausche die Zeilen von σ.

(2) Ordne die Spalten so an, dass die erste Zeile wieder 1, 2, …, n ist.

Wir betrachten wieder ein Beispiel. Es sei $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix} \in S_4$. Das Verfahren liefert uns

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix} \overset{\text{Zeilen tauschen}}{\rightsquigarrow} \begin{pmatrix} 2 & 3 & 1 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix} \overset{\text{Spalten sortieren}}{\rightsquigarrow} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 2 & 4 \end{pmatrix}$$

und damit $\sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 2 & 4 \end{pmatrix}$.

4.1.10 **Aufgabe**. Bestimmen Sie die inverse Permutation von $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 5 & 3 & 1 \end{pmatrix}$ in der Gruppe S_5. **L**

4.1.11 $\boxed{\textbf{Definition}}$ Es sei $\sigma \in S_n$. Eine Zahl $i \in \{1, 2, \ldots, n\}$ heißt *Fixpunkt* von σ, wenn $\sigma(i) = i$ gilt.

Der *Träger* von σ ist die Menge $\mathrm{Trg}(\sigma)$ aller Zahlen in $\{1, 2, \ldots, n\}$ die *keine* Fixpunkte sind, d.h.,

$$\mathrm{Trg}(\sigma) = \big\{ i \in \{1, 2, \ldots, n\} \mid \sigma(i) \neq i \big\}.$$

4.1.12 **Aufgabe**. Bestimmen Sie den Träger von $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 5 & 4 & 2 \end{pmatrix} \in S_5$. **L**

4.1.13 **Lemma.** *Seien $n \in \mathbb{N}$ und $\sigma \in S_n$. Ist $i \in \mathrm{Trg}(\sigma)$, dann ist auch $\sigma(i) \in \mathrm{Trg}(\sigma)$.*

Beweis. Sei $i \in \{1, 2, \ldots, n\}$. Wir führen den Beweis durch Kontraposition und nehmen an, dass $\sigma(i)$ nicht im Träger liegt. Dann ist $\sigma(i)$ ein Fixpunkt und es gilt

$$\sigma(\sigma(i)) = \sigma(i).$$

Da σ injektiv ist, folgt $\sigma(i) = i$, d.h., i ist ein Fixpunkt und liegt nicht im Träger. □

4.1.14 **Definition** Eine Permutation $\tau \in S_n$ deren Träger aus zwei Elementen besteht, nennt man *Transposition*.

4.1.15 **Anmerkung zur Definition**. Es sei τ in S_n eine Transposition. Dann gibt es zwei Zahlen $i \neq j$ mit $\operatorname{Trg}(\tau) = \{i, j\}$. In diesem Fall vertauscht τ die Zahlen i und j, d.h.,

$$\tau(i) = j, \quad \tau(j) = i \quad \text{und} \quad \tau(k) = k$$

für alle $k \neq i, j$ in $\{1, 2, \ldots, n\}$. Wir verwenden die Notation $\tau = (i \;\; j)$. Weil die Reihenfolge von i und j hier unwichtig ist, gilt $(i \;\; j) = (j \;\; i)$. Wegen

$$(i \;\; j) \circ (i \;\; j) = \mathrm{id},$$

gilt außerdem $(i \;\; j)^{-1} = (i \;\; j)$.

Der nächste Satz besagt, dass jede symmetrische Gruppe von Transpositionen „erzeugt“ wird.

4.1.16 **Satz.** *Sei $n \geq 2$. Jede Permutation $\sigma \in S_n$ kann als Komposition von Transpositionen geschrieben werden.*

Beweis. Es sei $\sigma \in S_n$. Wir führen den Beweis durch Induktion nach der Größe des Trägers $|\operatorname{Trg}(\sigma)|$. Ist $|\operatorname{Trg}(\sigma)| = 0$, so ist $\sigma = \mathrm{id}$ und es gilt $\mathrm{id} = (1 \;\; 2) \circ (1 \;\; 2)$.

Es sei nun $|\operatorname{Trg}(\sigma)| = m \geq 1$. Wir nehmen an, dass jede Permutation, deren Träger aus weniger als m Elementen besteht, als Komposition von Transpositionen geschrieben werden kann. Wir wählen ein $i \in \operatorname{Trg}(\sigma)$ und definieren $\pi := (i \;\; \sigma(i)) \circ \sigma$. Es gilt dann $\pi(i) = i$, d.h., i liegt nicht im Träger von π. Nach Lemma 4.1.13 ist auch $\sigma(i) \in \operatorname{Trg}(\sigma)$ und wir stellen fest, dass

$$\operatorname{Trg}(\pi) \subseteq \operatorname{Trg}(\sigma) \setminus \{i\}$$

gilt. Nach Voraussetzung kann π als Komposition $\pi = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_k$ von Transpositionen $\tau_1, \ldots, \tau_k$ geschrieben werden. Damit ist auch

$$\sigma = (i \;\; \sigma(i)) \circ (i \;\; \sigma(i)) \circ \sigma = (i \;\; \sigma(i)) \circ \pi = (i \;\; \sigma(i)) \circ \tau_1 \circ \cdots \circ \tau_k$$

als Komposition von Transpositionen darstellbar. □

4.1.17 **Beispiel**. Die Permutation $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix} \in S_4$ kann auf folgende Weise durch Transpositionen dargestellt werden:

$$\sigma = (1 \;\; 2) \circ (2 \;\; 3) \circ (3 \;\; 4).$$

III. Die Signaturabbildung

4.1.18 **Definition** Sei $n \in \mathbb{N}$ und sei $\sigma \in S_n$. Ein *Fehlstand* von σ ist ein Paar $(i, j) \in \{1, 2, \dots, n\}^2$, das

$$i < j \quad \text{und} \quad \sigma(i) > \sigma(j)$$

erfüllt. Die Menge aller Fehlstände von σ bezeichnen wir mit $\mathrm{FS}(\sigma)$.

4.1.19 **Beispiel**. Wir betrachten die Permutation $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 1 & 2 & 5 & 4 \end{pmatrix} \in S_5$. Das Paar $(1, 2)$ ist ein Fehlstand, denn $\sigma(1) = 3 > 1 = \sigma(2)$. Das Paar $(2, 3)$ ist hingegen kein Fehlstand, denn $\sigma(2) = 1$ und $\sigma(3) = 2$, es gilt also $\sigma(2) < \sigma(3)$. Insgesamt findet man die Fehlstände

$$\mathrm{FS}(\sigma) = \{(1, 2), (1, 3), (4, 5)\}.$$

4.1.20 **Definition** Seien $n \in \mathbb{N}$ und $\sigma \in S_n$. Die Zahl

$$\mathrm{sgn}(\sigma) = (-1)^{|\mathrm{FS}(\sigma)|}$$

heißt *Signatur* der Permutation σ.

Eine Permutation hat also die Signatur 1 oder -1 je nachdem, ob sie eine gerade oder ungerade Anzahl von Fehlständen hat.

4.1.21 **Beispiel**. Die identische Permutation id hat keine Fehlstände, also ist $\mathrm{sgn}(\mathrm{id}) = 1$.

4.1.22 **Aufgabe**. Bestimmen Sie die Fehlstände und die Signatur der Transposition $(1 \;\; 3) \in S_5$. **L**

Im nächsten Satz werden wir die Signatur von Transpositionen allgemein bestimmen.

4.1.23 **Satz.** *Jede Transposition τ hat Signatur* $\mathrm{sgn}(\tau) = -1$.

Beweis. Es seien $s < t \leq n$ gegeben. Wir bestimmen die Fehlstände der Transposition $\tau = (s \;\; t)$.

Es seien $i < j \leq n$. Wir unterscheiden vier Fälle:

Fall 1: $i \neq s$ und $j \neq t$. Die Zahl s ist die einzige, die auf eine echt größere Zahl abgebildet wird. Genauso ist die Zahl t die einzige, die auf eine echt kleinere abgebildet. Es gilt also $\tau(i) \leq i < j \leq \tau(j)$ und damit ist (i, j) kein Fehlstand von τ.

Fall 2: $i = s$ und $j = t$. Das Paar (s, t) ist ein Fehlstand, denn $\tau(s) = t > s = \tau(t)$.

Fall 3: $i = s$ und $j \neq t$. Das Paar (s, j) ist ein Fehlstand von τ genau dann, wenn $j > s$ und $j < \tau(s) = t$ gilt, d.h., es gilt $s + 1 \leq j \leq t - 1$.

Fall 4: $j = t$ und $i \neq s$. Das Paar (i, t) ist ein Fehlstand von τ genau dann, wenn $s + 1 \leq i \leq t - 1$ gilt.

Wir haben damit die Fehlstände

$$(s, s+1), (s, s+2), \ldots, (s, t-1), (s, t), (s+1, t), (s+2, t), \ldots, (t-1, t)$$

gefunden. Also hat τ genau $2(t - s) - 1$ verschiedene Fehlstände. Dies ist eine ungerade Zahl, also gilt $\operatorname{sgn}(\tau) = -1$. □

4.1.24 **Satz von der Signatur.** *Sei $n \in \mathbb{N}$. Für alle $\sigma \in S_n$ gilt die Formel*

$$\operatorname{sgn}(\sigma) = \prod_{\{i,j\}} \frac{\sigma(j) - \sigma(i)}{j - i},$$

dabei läuft das Produkt über alle zwei-elementigen Teilmengen von $\{1, 2, \ldots, n\}$. Für alle $\sigma, \tau \in S_n$ gilt

$$\operatorname{sgn}(\sigma \circ \tau) = \operatorname{sgn}(\sigma) \cdot \operatorname{sgn}(\tau).$$

Beweis. Es sei P_2 die Menge aller zwei-elementigen Teilmengen von $\{1, 2, \ldots, n\}$. Wir beginnen mit zwei Vorüberlegungen.

1.) Jede Permutation $\sigma \in S_n$ liefert uns auch eine Abbildung

$$\overline{\sigma} \colon P_2 \to P_2 \quad \text{mit} \quad \{i, j\} \mapsto \{\sigma(i), \sigma(j)\}.$$

Die Abbildung $\overline{\sigma}$ ist bijektiv, denn $\overline{\sigma^{-1}}$ ist offensichtlich eine Umkehrabbildung.

2.) Für $i \neq j$ gilt $\frac{\sigma(j)-\sigma(i)}{j-i} = \frac{\sigma(i)-\sigma(j)}{i-j}$. Insbesondere hängt dieser Ausdruck nur von der zwei-elementigen Menge $\{i, j\}$ ab und nicht von der Anordnung von i, j in diesem Ausdruck.

Es sei $\sigma \in S_n$ und es seien $i < j$ gegeben. Dann ist $\sigma(j) - \sigma(i)$ genau dann negativ, wenn (i, j) ein Fehlstand ist, d.h.,

$$\frac{\sigma(j) - \sigma(i)}{j - i} = \pm \frac{|\sigma(j) - \sigma(i)|}{|j - i|}$$

mit negativem Vorzeichen genau dann, wenn $(i, j) \in \operatorname{FS}(\sigma)$ ist.

Daraus erhalten wir

$$\prod_{\{i,j\}} \frac{\sigma(j)-\sigma(i)}{j-i} = (-1)^{|\operatorname{FS}(\sigma)|} \cdot \prod_{\{i,j\}} \frac{|\sigma(j)-\sigma(i)|}{|j-i|}$$
$$= \operatorname{sgn}(\sigma) \cdot \frac{\prod_{\{i,j\}} |\sigma(j)-\sigma(i)|}{\prod_{\{i,j\}} |j-i|}. \tag{4.1.a}$$

Der Betrag $b(\{i,j\}) := |i-j|$ hängt nur von der Menge $\{i,j\}$ ab. Aus Überlegung 1.) folgt nun

$$\prod_{\{i,j\}} |\sigma(j)-\sigma(i)| = \prod_{M\in P_2} b(\overline{\sigma}(M)) = \prod_{M'\in P_2} b(M') = \prod_{\{i',j'\}} |j'-i'|.$$

Damit hebt sich der Bruch in (4.1.a) weg und wir haben die Formel bewiesen.

Es seien nun $\sigma, \tau \in S_n$. Dann gilt

$$\operatorname{sgn}(\sigma\circ\tau) = \prod_{\{i,j\}} \frac{\sigma(\tau(j))-\sigma(\tau(i))}{j-i} = \prod_{\{i,j\}} \left(\frac{\sigma(\tau(j))-\sigma(\tau(i))}{j-i} \cdot \frac{\tau(j)-\tau(i)}{\tau(j)-\tau(i)} \right)$$
$$= \prod_{\{i,j\}} \frac{\sigma(\tau(j))-\sigma(\tau(i))}{\tau(j)-\tau(i)} \cdot \underbrace{\prod_{\{i,j\}} \frac{\tau(j)-\tau(i)}{j-i}}_{=\operatorname{sgn}(\tau)}.$$

Der zweite Faktor in diesem Produkt ist $\operatorname{sgn}(\tau)$. Um einzusehen, dass der erste Faktor die Signatur von σ ist, kann man wieder mit Überlegung 1.) argumentieren, denn jede zwei-elementige Teilmenge $\{i',j'\}$ kann auf eindeutige Weise in der Form $\{i',j'\} = \overline{\tau}(\{i,j\})$ geschrieben werden. Es gilt also

$$\prod_{\{i,j\}} \frac{\sigma(\tau(j))-\sigma(\tau(i))}{\tau(j)-\tau(i)} = \prod_{\{i',j'\}} \frac{\sigma(j')-\sigma(i')}{j'-i'} = \operatorname{sgn}(\sigma). \qquad \square$$

4.1.25 **Korollar.** *Sei $\sigma \in S_n$ mit $n \geq 2$. Es gilt $\operatorname{sgn}(\sigma) = 1$ genau dann, wenn σ Komposition einer geraden Anzahl von Transpositionen ist.*

Beweis. Es sei $\sigma \in S_n$. Nehmen wir zunächst an, dass $\sigma = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_{2m}$ Komposition einer geraden Anzahl von Transpositionen $\tau_1, \ldots, \tau_{2m}$ ist. Mit dem Satz von der Signatur und Satz 4.1.23 erhalten wir

$$\operatorname{sgn}(\sigma) = \operatorname{sgn}(\tau_1)\cdot\operatorname{sgn}(\tau_2)\cdots\operatorname{sgn}(\tau_{2m}) = (-1)^{2m} = 1.$$

Sei umgekehrt σ eine Permutation mit $\operatorname{sgn}(\sigma) = 1$. Nach Satz 4.1.16 können wir σ als Komposition $\sigma = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_k$ von Transpositionen $\tau_1, \ldots, \tau_k$ schreiben. Mit dem Satz von der Signatur und Satz 4.1.23 erhalten wir

$$1 = \operatorname{sgn}(\sigma) = \operatorname{sgn}(\tau_1) \cdot \operatorname{sgn}(\tau_2) \cdots \operatorname{sgn}(\tau_k) = (-1)^k;$$

damit muss k gerade sein. □

4.1.26 **Korollar.** *Sei $\sigma \in S_n$. Es gilt* $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$.

Beweis. Es gilt

$$\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma \circ \mathrm{id}) = \operatorname{sgn}(\sigma \circ \sigma \circ \sigma^{-1}) = \operatorname{sgn}(\sigma)^2 \operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma^{-1}),$$

denn $\operatorname{sgn}(\sigma)^2 = 1$. □

4.1.27 **Definition und Satz.** *Sei $n \in \mathbb{N}$. Die Menge*

$$\mathrm{Alt}_n = \{\sigma \in S_n \mid \operatorname{sgn}(\sigma) = 1\} \subseteq S_n$$

ist eine Untergruppe von S_n. Die Gruppe Alt_n *heißt* alternierende Gruppe *auf n Buchstaben.*

Beweis. Wir prüfen nach, dass die Bedingungen aus Definition 1.1.18 erfüllt sind.

(UG1): Alt_n enthält die Identität, denn es gilt $\operatorname{sgn}(\mathrm{id}) = 1$; siehe Beispiel 4.1.21.

(UG2): Es seien $\sigma, \tau \in \mathrm{Alt}_n$. Aus dem Satz von der Signatur 4.1.24 folgt

$$\operatorname{sgn}(\sigma \circ \tau) = \operatorname{sgn}(\sigma) \cdot \operatorname{sgn}(\tau) = 1 \cdot 1 = 1.$$

Also ist $\sigma \circ \tau \in \mathrm{Alt}_n$.

(UG3): Es sei $\sigma \in \mathrm{Alt}_n$. Aus Korollar 4.1.26 folgt $\operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma) = 1$ und damit $\sigma^{-1} \in \mathrm{Alt}_n$. □

4.1.28 **Aufgabe.** Bestimmen Sie alle Elemente der Gruppe Alt_3. **L**

4.1.29 **Aufgabe.** Sei $n \geq 2$ und sei $\tau \in S_n$ eine Transposition. Es sei $B_n \subseteq S_n$ die Menge aller Permutationen mit Signatur -1. Zeigen Sie, dass die Zuordnung $\sigma \mapsto \sigma \circ \tau$ eine bijektive Abbildung **L**

$$f_\tau \colon \mathrm{Alt}_n \to B_n$$

definiert.

4.2. Die Determinante von Matrizen

In diesem Abschnitt werden wir ein sehr wichtiges Konzept kennenlernen: die Determinante einer quadratischen Matrix.

I. Definition

Wir beginnen mit der sogenannten Leibniz[1]-Formel für die Determinante, die zunächst nicht sehr erhellend ist. Wir werden uns dann auf den Weg machen, die Determinante besser zu verstehen.

4.2.1 **Definition** Es sei $n \geq 1$ und sei R ein kommutativer Ring. Für eine $(n \times n)$-Matrix $A = (a_{ij})_{i,j} \in \mathrm{M}_{n,n}(R)$ definieren wir die *Determinante* $\det(A) \in R$ durch

$$\begin{aligned} \det(A) &:= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot a_{\sigma(1),1} \cdot a_{\sigma(2),2} \cdots a_{\sigma(n),n} \qquad (4.2.a) \\ &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{j=1}^{n} a_{\sigma(j),j} \end{aligned}$$

4.2.2 **Anmerkungen zur Definition.** (a) Die Formel (4.2.a) nennt man die *Leibniz-Formel* für die Determinante.

(b) In (4.2.a) bezeichnet $\operatorname{sgn}(\sigma)$ die Signatur der Permutation σ aufgefasst als Element ± 1 im Ring R.

Im folgenden Beispiel werden wir uns die Leibniz-Formel für kleine Matrizen genauer anschauen.

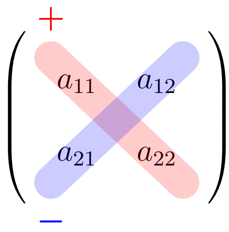
4.2.3 **Beispiel.** (a) Für $n = 1$ ist $A = a_{11} \in R$ ein Ringelement und die Leibniz-Formel liefert $\det(A) = a_{11}$.

(b) Für $n = 2$ besteht die symmetrische Gruppe aus zwei Elementen: $S_2 = \{\mathrm{id}, (1\ 2)\}$; siehe 4.1.4. Die Transposition hat Signatur -1. Die Determinante von $A = (a_{ij})_{i,j} \in \mathrm{M}_{2,2}(R)$ ist demnach

$$\det(A) = a_{11}a_{22} + (-1)a_{21}a_{12} = a_{11}a_{22} - a_{21}a_{12}.$$

Es gilt also die (2×2)-Merkregel: Produkt der Diagonaleinträge minus Produkt der anderen Einträge.

[1] Gottfried Wilhelm Leibniz: deutscher Universalgelehrter, 1646–1716.



4.2.4 **Aufgabe**. Berechnen Sie die Determinante von **L**

$$A = \begin{pmatrix} 3 & -5 \\ 1 & -3 \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{R}).$$

4.2.5 **Flächeninhalt und Determinante**. Die Determinante einer reellen 2×2-Matrix hat eine geometrische Bedeutung. Sind $u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \in \mathbb{R}^2$ und $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \in \mathbb{R}^2$ linear unabhängig, dann spannen diese Vektoren ein *Parallelogramm* $P(u, v)$ auf. In diesem

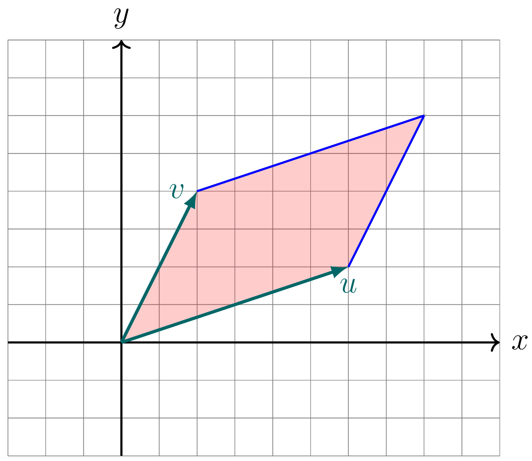


Abbildung 4.1.: Von zwei Vektoren aufgespanntes Parallelogramm im $\mathbb{R}^2$.

Fall ist der Flächeninhalt F von $P(u, v)$ genau

$$F = \left| \det \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix} \right|.$$

Wir überprüfen diese Aussage in einem speziellen Fall. Wir nehmen an, dass alle Einträge von u und v positiv sind und, dass $u_2/u_1 < v_2/v_1$ ist. Das heißt, wir sind in der Situation aus Abbildung 4.1. Wir berechnen den Flächeninhalt F als

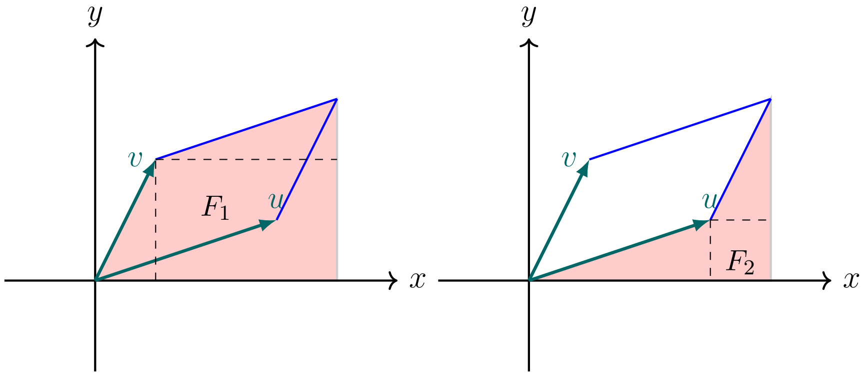


Abbildung 4.2.: Fläche des Parallelogramms als Differenz zweier Flächen.

Differenz der Flächeninhalte F_1 und F_2 aus Abbildung 4.2. Zerlegt man die Flächen in Rechtecke und rechtwinklige Dreiecke, erhält man

$$\begin{aligned} F_1 &= u_1v_2 + \frac{1}{2}v_1v_2 + \frac{1}{2}u_1u_2 \\ F_2 &= u_2v_1 + \frac{1}{2}v_1v_2 + \frac{1}{2}u_1u_2. \end{aligned}$$

Die Flächeninhalte der Dreiecke heben sich weg und man erhält

$$F = F_1 - F_2 = u_1v_2 - u_2v_1 = \det\begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix}.$$

Das Vorzeichen der Determinante hat ebenfalls eine geometrische Bedeutung. Das Vorzeichen beschreibt die *Orientierung der Vektoren* in $\mathbb{R}^2$. Man nennt u, v positiv orientiert, wenn $\det\begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix} > 0$ ist. Das ist genau dann der Fall, wenn man u gegen den Uhrzeigersinn drehen muss, um v nach einer Drehung von weniger als $180°$ zu erreichen.

4.2.6 3×3**: Die Regel von Sarrus**. Sei $n = 3$. Die symmetrische Gruppe S_3 besteht aus 6 Elementen:

$$S_3 = \{\mathrm{id}, (1\ 2), (1\ 3), (2\ 3), \sigma, \tau\}$$

mit

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \quad \text{und} \quad \tau = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$

Die drei Transpositionen haben Signatur -1 und $\mathrm{id}, \sigma, \tau$ haben Signatur 1; siehe Aufgabe 4.1.28. Die Determinante von $A = (a_{ij})_{i,j} \in \mathrm{M}_{3,3}(R)$ ergibt sich aus der

Leibniz-Formel als

$$\det(A) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32}.$$

Diese Formel kann man sich mit der Regel von Sarrus gut merken. Dazu schreibt man sich die ersten beiden Spalten der Matrix A nochmal (mit Bleistift) rechts neben die Matrix A, d.h.,

$$\begin{array}{c} \begin{array}{ccccc} + & + & + & & \end{array} \\ \left(\begin{array}{ccc|cc} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{array}\right) \\ \begin{array}{ccccc} - & - & - & & \end{array} \end{array}$$

Die Terme, die mit der Signatur $+1$ in der Determinante auftauchen, sind die Produkte über die „Diagonallinien" von links oben nach rechts unten (angefangen bei den Einträgen der ersten Zeile der Matrix A):

$$a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}.$$

Die Terme, die mit der Signatur -1 in der Determinante auftreten, sind die Produkte über die Einträge auf den „Diagonallinien" von links unten nach rechts oben (angefangen bei den Einträgen der letzten Zeile der Matrix A):

$$-a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32}.$$

4.2.7 **Beispiel**. Wir betrachten die (3×3)-Matrix

$$A = \begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{Z}).$$

Wir wenden die Regel von Sarrus an und schreiben die ersten zwei Spalten nochmal rechts daneben

$$\begin{array}{c} \begin{array}{ccccc} + & + & + & & \end{array} \\ \left(\begin{array}{ccc|cc} 0 & 1 & 2 & 0 & 1 \\ 3 & 4 & 5 & 3 & 4 \\ 6 & 7 & 8 & 6 & 7 \end{array}\right) \\ \begin{array}{ccccc} - & - & - & & \end{array} \end{array}$$

Wir erhalten dann

$$\begin{aligned}\det(A) &= 0\cdot 4\cdot 8+1\cdot 5\cdot 6+2\cdot 3\cdot 7-6\cdot 4\cdot 2-7\cdot 5\cdot 0-8\cdot 3\cdot 1\\ &= 0+30+42-48-0-24=0.\end{aligned}$$

4.2.8 **Aufgabe**. Berechnen Sie die Determinante der Matrix **L**

$$A=\begin{pmatrix}\overline{3} & \overline{1} & \overline{1}\\ \overline{2} & \overline{0} & \overline{5}\\ \overline{1} & \overline{4} & \overline{3}\end{pmatrix}\in \mathrm{M}_{3,3}(\mathbb{F}_7)$$

mit der Regel von Sarrus.

4.2.9 **Determinante und Volumen**. Die Determinante reeller (3×3)-Matrizen hat eine geometrische Bedeutung. Der Betrag der Determinante ist das Volumen des von den Spalten aufgespannten *Parallelepipeds*. Das Vorzeichen der Determinante beschreibt wieder die *Orientierung* der Spaltenvektoren im $\mathbb{R}^3$.

Diese Zusammenhänge bleiben auch allgemein im $\mathbb{R}^n$ gültig. Dazu muss man aber zuerst das Volumen im $\mathbb{R}^n$ definieren. Man könnte hier auch umgekehrt vorgehen und die Determinante verwenden, um „Volumen“ im $\mathbb{R}^n$ zu definieren.

4.2.10 **Achtung bei großen Matrizen**. Für $(n\times n)$-Matrizen mit $n\geq 4$ gibt es *keine* einfache Merkregel für die Terme der Leibniz-Formel. Überhaupt ist die Leibniz-Formel zur Berechnung der Determinante großer Matrizen im Allgemeinen nicht geeignet. Der Grund ist, dass die symmetrische Gruppe S_n sehr viele Elemente hat: $|S_n|=n!$ (siehe Satz 4.1.3). Die Leibniz-Formel hat für (4×4)-Matrizen also 24 Terme, für (5×5)-Matrizen bereits 120 Terme und für (6×6)-Matrizen sogar schon 720 Terme. Wir werden bald andere Methoden kennenlernen um die Determinante zu berechnen.

Falls eine Matrix nur wenige Einträge ungleich 0 hat, dann kann es passieren, dass viele Terme der Leibniz-Formel verschwinden und man einen sehr einfachen Ausdruck für die Determinante bekommt.

4.2.11 **Beispiel** (Diagonalmatrizen). Sei R ein kommutativer Ring und sei $A=(a_{ij})_{i,j}\in \mathrm{M}_{n,n}(R)$ eine Matrix. Man nennt A eine *Diagonalmatrix*, wenn alle Einträge, die keine Diagonalelemente sind, verschwinden, d.h., $a_{ij}=0$ falls $i\neq j$ ist. In anderen Worten hat A die Gestalt

$$A=\begin{pmatrix}a_{1,1} & 0 & \cdots & 0\\ 0 & a_{2,2} & & \vdots\\ \vdots & & \ddots & 0\\ 0 & \cdots & 0 & a_{n,n}\end{pmatrix}.$$

Sei $\sigma \in S_n$ eine Permutation. Ist $\sigma \neq \mathrm{id}$, dann gibt es ein $i \in \{1, 2, \ldots, n\}$ mit $\sigma(i) \neq i$. Ist A eine Diagonalmatrix, dann gilt also $a_{\sigma(i),i} = 0$ und der Ausdruck

$$a_{\sigma(1),1} \cdot a_{\sigma(2),2} \cdots a_{\sigma(n),n}$$

in der Leibniz-Formel (4.2.a) verschwindet. Um die Determinante einer Diagonalmatrix zu bestimmen, muss man also nur den Term zur Permutation $\sigma = \mathrm{id}$ der Leibniz-Formel berechnen. Wir erhalten also folgende Aussage:

Ist $A \in \mathrm{M}_{n,n}(R)$ eine Diagonalmatrix, dann ist die Determinante das Produkt aller Diagonaleinträge, d.h.,

$$\det(A) = a_{1,1} \cdot a_{2,2} \cdots a_{n,n} = \prod_{i=1}^{n} a_{i,i}.$$

Dieses Ergebnis kann man auch etwas allgemeiner für Dreiecksmatrizen beweisen.

4.2.12 **Definition** Eine Matrix $A \in \mathrm{M}_{n,n}(R)$ heißt *obere Dreiecksmatrix*, wenn $a_{i,j} = 0$ für alle i, j mit $i > j$ gilt. Analog heißt A *untere Dreiecksmatrix*, wenn $a_{i,j} = 0$ für alle i, j mit $i < j$ gilt.

In anderen Worten: A ist genau dann eine obere Dreiecksmatrix, wenn A^T eine untere Dreiecksmatrix ist.

4.2.13 **Satz.** *Sei R ein kommutativer Ring und sei $A \in \mathrm{M}_{n,n}(R)$ eine obere Dreiecksmatrix. Dann gilt*

$$\det(A) = \prod_{i=1}^{n} a_{i,i}.$$

Beweis. Sei $\sigma \in S_n$ mit $\sigma \neq \mathrm{id}$. Wir zeigen, dass der Term

$$a_{\sigma(1),1} \cdot a_{\sigma(2),2} \cdots a_{\sigma(n),n}$$

in der Leibniz-Formel (4.2.a) verschwindet.

Sei $i \in \{1, 2, \ldots, n\}$ die kleinste Zahl im Träger von σ, d.h., i ist minimal mit $\sigma(i) \neq i$. Da nach Lemma 4.1.13 auch $\sigma(i)$ im Träger liegt, gilt $\sigma(i) > i$. Ist A eine obere Dreiecksmatrix, dann folgt $a_{\sigma(i),i} = 0$ und damit verschwindet der zu σ gehörende Summand in der Leibniz-Formel. Es bleibt in (4.2.a) nur der Summand zu $\sigma = \mathrm{id}$ übrig und wir erhalten

$$\det(A) = a_{1,1} \cdot a_{2,2} \cdots a_{n,n} = \prod_{i=1}^{n} a_{i,i}. \qquad \square$$

Satz 4.2.13 gilt natürlich auch für untere Dreiecksmatrizen. Dazu könnte man den Beweis einfach leicht umschreiben. Man kann das aber auch auf eine andere Weise einsehen. Ist A eine untere Dreiecksmatrix, dann ist A^T eine obere Dreiecksmatrix und damit ist die Aussage eine Konsequenz aus folgendem Lemma.

4.2.14 **Lemma.** *Sei R ein kommutativer Ring und sei $A \in \mathrm{M}_{n,n}(R)$. Dann gilt*

$$\det(A) = \det(A^T).$$

Beweis. Für $\sigma \in S_n$ kommt jede der Zahlen $1, 2, \ldots, n$ genau einmal in der Liste $\sigma(1), \sigma(2), \ldots, \sigma(n)$ vor. Ist $\sigma(i) = j$, dann ist $i = \sigma^{-1}(j)$. Wir können also jeden Summanden der Leibniz-Formel (4.2.a) umsortieren und erhalten:

$$\operatorname{sgn}(\sigma) \cdot a_{\sigma(1),1} \cdot a_{\sigma(2),2} \cdots a_{\sigma(n),n} = \operatorname{sgn}(\sigma) \cdot a_{1,\sigma^{-1}(1)} \cdot a_{2,\sigma^{-1}(2)} \cdots a_{n,\sigma^{-1}(n)}.$$

Aus Korollar 4.1.26 ist bekannt, dass $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$ gilt. Da die Abbildung $\sigma \mapsto \sigma^{-1}$ bijektiv ist (siehe 1.1.4), können wir die Leibniz-Formel wie folgt umschreiben:

$$\begin{aligned} \det(A) &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{j=1}^{n} a_{\sigma(j),j} \\ &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma^{-1}) \prod_{i=1}^{n} a_{i,\sigma^{-1}(i)} \\ &= \sum_{\tau \in S_n} \operatorname{sgn}(\tau) \prod_{i=1}^{n} a_{i,\tau(i)}. \end{aligned}$$

Der letzte Ausdruck ist genau die Determinante von A^T, denn der Eintrag von A^T an der Stelle $(\tau(i), i)$ ist $a_{i,\tau(i)}$. □

4.2.15 **Aufgabe**. Es sei $A \in \mathrm{M}_{n,n}(\mathbb{C})$. Zeigen Sie, dass $\det(A^*) = \overline{\det(A)}$ gilt. **L**

II. Charakterisierungssatz

Wir sind nun in der Lage die Determinante eindeutig anhand Ihrer Eigenschaften zu beschreiben. Dazu benötigen wir noch einen weiteren Begriff. Dabei schreiben wir $A = (a_1|a_2|\ldots|a_n)$, wenn $A \in \mathrm{M}_{n,n}(R)$ die Spalten $a_1, a_2, \ldots, a_n \in R^n$ hat.

4.2.16 **Definition** Wir nennen eine Abbildung $f\colon \mathrm{M}_{n,n}(R) \to R$ *linear in jeder Spalte*, wenn für alle $k \in \{1, 2, \ldots, n\}$ gilt

(a) für alle Spalten $a_1, \dots, a_n, b_k \in R^n$ ist

$$f(a_1|\dots|a_{k-1}|a_k + b_k|a_{k+1}|\dots|a_n) \\ = f(a_1|\dots|a_{k-1}|a_k|a_{k+1}|\dots|a_n) + f(a_1|\dots|a_{k-1}|b_k|a_{k+1}|\dots|a_n),$$

(b) für alle Spalten $a_1, \dots, a_n \in R^n$ und $r \in R$ gilt

$$f(a_1|\dots|a_{k-1}|ra_k|a_{k+1}|\dots|a_n) = rf(a_1|\dots|a_{k-1}|a_k|a_{k+1}|\dots|a_n).$$

Wir werden in Abschnitt 4.5 allgemeiner über *multilineare Abbildungen* reden. Ist $R = K$ ein Körper und fasst man $\mathrm{M}_{n,n}(K)$ als $(K^n)^{\times n}$ auf (jede Spalte ist ein Vektor im K^n), dann ist eine Abbildung genau dann linear in jeder Spalte, wenn sie multilinear im Sinne von Definition 4.5.2 ist.

4.2.17 **Charakterisierungssatz.** *Sei R ein kommutativer Ring und sei $n \geq 1$. Die Determinantenabbildung* $\det\colon \mathrm{M}_{n,n}(R) \to R$ *hat folgende Eigenschaften*

(i) Die Determinante ist linear in jeder Spalte.

(ii) $\det(A) = 0$ *für jede Matrix A mit zwei gleichen Spalten.*

(iii) $\det(I_n) = 1$

Die Determinante ist die einzige Abbildung mit diesen drei Eigenschaften.

4.2.18 **Anmerkungen zum Satz**. (a) Zu Eigenschaft (ii) sagt man auch: Die Determinante ist *alternierend.*

(b) Bedingung (iii) ist eine Normierung der Determinantenabbildung. Ohne diese Bedingung wäre die Determinante nicht eindeutig, denn die Abbildung $A \mapsto \lambda \det(A)$ mit $\lambda \in R$ hat ebenfalls die beiden Eigenschaften (i) und (ii).

(c) Aus Lemma 4.2.14 folgt direkt, dass man im Charakterisierungssatz auch „Spalte“ durch „Zeile“ ersetzen kann.

Erster Teil des Beweises von Satz 4.2.17. Wir zeigen zuerst, dass die Determinante diese Eigenschaften besitzt. Eigenschaft (iii) folgt direkt aus Beispiel 4.2.11.

Zu (ii): Nehmen wir an, dass A eine Matrix mit zwei gleichen Spalten ist. Sagen wir die k-te und die ℓ-te Spalte stimmen überein mit $\ell \neq k$, d.h., $a_{ik} = a_{i\ell}$ für alle $i \in \{1, 2, \dots, n\}$. Sei $\tau = (k\ \ \ell) \in S_n$ die Transposition, die k und ℓ vertauscht. Dann gelten

$$a_{\sigma(\ell),\ell} = a_{\sigma(\ell),k} = a_{\sigma(\tau(k)),k} \quad \text{und} \quad a_{\sigma(k),k} = a_{\sigma(k),\ell} = a_{\sigma(\tau(\ell)),\ell}$$

für alle $\sigma \in S_n$. Da für alle $j \notin \{k, \ell\}$ auch $a_{\sigma(j),j} = a_{\sigma(\tau(j)),j}$ gilt, erhalten wir

$$\prod_{j=1}^{n} a_{\sigma(j),j} = \prod_{j=1}^{n} a_{\sigma(\tau(j)),j}.$$

Wir wissen aus Aufgabe 4.1.29, dass $\sigma \mapsto \sigma \circ \tau$ eine Bijektion zwischen Alt_n und der Menge B_n aller Permutationen mit Signatur -1 definiert. Damit hebt sich in der Leibniz-Formel (4.2.a) jeweils der Summand zu $\sigma \in \mathrm{Alt}_n$ und der Summand zu $\sigma \circ \tau \in B_n$ weg und die Aussage ist bewiesen.

Zu (i): Es sei $A \in \mathrm{M}_{n,n}(R)$ eine Matrix mit Spalten $a_1, \ldots, a_n$, es sei B (bzw. C) die Matrix, die man aus A erhält indem man die k-te Spalte durch b_k (bzw. $a_k + b_k$) ersetzt. Dann gilt

$$\begin{aligned}
\det(C) &= \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma)(a_{\sigma(k),k} + b_{\sigma(k),k}) \prod_{\substack{j=1 \\ j \neq k}}^{n} a_{\sigma(j),j} \\
&= \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) a_{\sigma(k),k} \prod_{\substack{j=1 \\ j \neq k}}^{n} a_{\sigma(j),j} + \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) b_{\sigma(k),k} \prod_{\substack{j=1 \\ j \neq k}}^{n} a_{\sigma(j),j} \\
&= \det(A) + \det(B).
\end{aligned}$$

Sei nun $r \in R$ und sei D die Matrix, die man erhält indem man die k-te Spalte in A durch ra_k ersetzt. Dann erhalten wir

$$\begin{aligned}
\det(D) &= \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma)(r a_{\sigma(k),k}) \prod_{\substack{j=1 \\ j \neq k}}^{n} a_{\sigma(j),j} \\
&= r \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \prod_{j=1}^{n} a_{\sigma(j),j} = r \det(A).
\end{aligned}$$

Also ist die Determinante linear in jeder Spalte. □

Um zu zeigen, dass die Determinante durch die Eigenschaften des Charakterisierungssatzes eindeutig bestimmt ist, benötigen wir zuerst eine kleine Vorüberlegung.

4.2.19 **Lemma.** *Sei $\Delta \colon \mathrm{M}_{n,n}(R) \to R$ alternierend und linear in jeder Spalte. Entsteht A' aus A durch Vertauschen zweier Spalten, dann ist $\Delta(A') = -\Delta(A)$.*

Beweis. Sei $A \in \mathrm{M}_{n,n}(R)$ mit Spalten $a_1, \dots, a_n$. Sagen wir $A' = (a_2|a_1|a_3|\dots|a_n)$ entsteht aus A durch Vertauschen der ersten beiden Spalten (andere Vertauschungen kann man analog behandeln). Dann bilden wir die Matrix $B = (a_1 + a_2|a_1 + a_2|a_3|\dots|a_n)$. Da B zwei gleiche Spalten hat, gilt $\Delta(B) = 0$. Aus der Linearität in jeder Spalte folgt nun

$$0 = \Delta(B) = \underbrace{\Delta(a_1|a_1 \mid a_3 \mid \dots)}_{=0} + \underbrace{\Delta(a_2 \mid a_2 \mid a_3| \ \dots)}_{=0} + \Delta(A) + \Delta(A')$$

und damit $\Delta(A') = -\Delta(A)$. □

Beweis der Eindeutigkeitsaussage in 4.2.17. Es sei $\Delta\colon \mathrm{M}_{n,n}(R) \to R$ eine Abbildung mit den Eigenschaften (i), (ii) und (iii). Wir zeigen, dass Δ die Determinante ist. Sei $A \in \mathrm{M}_{n,n}(R)$ eine Matrix. Sei $e_i \in R^n$ die Spalte mit dem Eintrag 1 an der Stelle i und 0 an allen anderen Einträgen. Die j-te Spalte von A ist

$$a_j = \sum_{i=1}^{n} a_{ij} e_i$$

und da Δ linear in jeder Spalte ist, erhalten wir durch „Ausmultiplizieren“

$$\Delta(A) = \sum_{1 \le i_1, \dots, i_n \le n} \prod_{j=1}^{n} a_{i_j, j} \ \Delta(e_{i_1}|e_{i_2}|\dots|e_{i_n}).$$

Dabei ist jeweils $(e_{i_1}|e_{i_2}|\dots|e_{i_n})$ eine Matrix, deren Einträge fast alle gleich Null sind abgesehen von genau einer 1 in jeder Spalte.

Was sind nun die Terme $\Delta(e_{i_1}|e_{i_2}|\dots|e_{i_n})$? Ist $i_j = i_k$ für $j \neq k$, dann hat die Matrix zwei gleiche Spalten und dieser Term verschwindet. Wir müssen uns also nur Tupel $i_1, \dots, i_n$ ansehen, deren Elemente paarweise verschieden sind. Dies ist genau denn der Fall, wenn

$$\sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{pmatrix}$$

eine Permutation ist. Wir können die Summe also umschreiben und erhalten

$$\Delta(A) = \sum_{\sigma \in S_n} \prod_{j=1}^{n} a_{\sigma(j), j} \ \Delta(e_{\sigma(1)}|e_{\sigma(2)}|\dots|e_{\sigma(n)}).$$

Kann man σ als Komposition von ℓ Transpositionen schreiben, dann kann man die Matrix $(e_{\sigma(1)}|e_{\sigma(2)}|\dots|e_{\sigma(n)})$ durch ℓ-maliges Vertauschen zweier Spalten in die

Einheitsmatrix überführen. Da $\Delta(I_n) = 1$ ist, erhalten wir aus Lemma 4.2.19 und Korollar 4.1.25

$$\Delta(e_{\sigma(1)}|e_{\sigma(2)}|\dots|e_{\sigma(n)}) = (-1)^{\ell}\Delta(I_n) = \operatorname{sgn}(\sigma).$$

Damit ist

$$\Delta(A) = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)\prod_{j=1}^{n} a_{\sigma(j),j} = \det(A). \qquad \square$$

Aus dem Charakterisierungssatz können wir nun weitere Eigenschaften der Determinante ableiten.

4.2.20 Korollar.

(a) Ist $A \in \mathrm{M}_{n,n}(R)$ eine Matrix mit einer Nullspalte oder Nullzeile, dann gilt $\det(A) = 0$.

(b) Entsteht A' aus A durch Vertauschen zweier Spalten (oder zweier Zeilen), dann ist $\det(A') = -\det(A)$.

(c) Ist $A \in \mathrm{M}_{n,n}(R)$ und $r \in R$, dann gilt $\det(rA) = r^n \det(A)$.

Beweis. Wir beweisen jeweils nur die Aussage für die Spalten. Die analoge Aussage über Zeilen folgt dann aus Lemma 4.2.14.

Sei $A \in \mathrm{M}_{n,n}(R)$ mit Spalten $a_1, \dots, a_n$.

Zu (a): Angenommen A hat eine Nullspalte. Ist $a_k = 0$, dann ist $a_k = 0 \cdot a_k$. Da die Determinante linear in jeder Spalte ist, folgt

$$\det(A) = \det(a_1|\dots|a_{k-1}|0\cdot a_k|a_{k+1}|\dots|a_n) = 0\cdot\det(A) = 0.$$

Zu (b): Dies folgt aus Lemma 4.2.19, weil die Determinante alternierend und linear in jeder Spalte ist.

Zu (c): Die Matrix rA entsteht aus A, indem man *jede* Spalte mit r multipliziert. Die Linearität in jeder Spalte erlaubt es uns r aus jeder der insgesamt n Spalten einmal herauszuziehen; d.h., $\det(rA) = r^n \det(A)$. $\square$

III. Multiplikationssatz

4.2.21 **Determinantenmultiplikationssatz.** *Sei R ein kommutativer Ring. Für alle $A, B \in \mathrm{M}_{n,n}(R)$ gilt*

$$\det(AB) = \det(A)\det(B).$$

Beweis. Sei $A = (a_{i,j})_{i,j}$ und $B = (b_{i,j})_{i,j}$. Dann hat $C := AB$ die Einträge

$$c_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j}. \tag{4.2.b}$$

Wir beginnen den Beweis des Multiplikationssatzes indem wir zuerst $\det(C)$ mithilfe der Leibniz-Formel berechnen.

$$\begin{aligned}
\det(C) &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{j=1}^{n} c_{\sigma(j),j} \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{j=1}^{n} \sum_{k=1}^{n} a_{\sigma(j),k} b_{k,j} && \text{(mit (4.2.b))} \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \sum_{1 \leq k_1,\dots,k_n \leq n} \prod_{j=1}^{n} a_{\sigma(j),k_j} b_{k_j,j} && \text{(ausmultiplizieren)} \\
&= \sum_{1 \leq k_1,\dots,k_n \leq n} \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{j=1}^{n} a_{\sigma(j),k_j} b_{k_j,j} && \text{(Summen vertauschen)} \\
&= \sum_{1 \leq k_1,\dots,k_n \leq n} \det(C(k_1,\dots,k_n))
\end{aligned}$$

Dabei definieren wir für jedes Tupel $(k_1, k_2, \dots, k_n) \in \{1, \dots, n\}^n$ die Hilfsmatrix $C(k_1, \dots, k_n) \in \mathrm{M}_{n,n}(R)$, deren Eintrag an der Stelle (i, j) gerade $a_{i,k_j} b_{k_j,j}$ ist, d.h.,

$$C(k_1, \dots, k_n) = (a_{i,k_j} b_{k_j,j})_{i,j}. \tag{4.2.c}$$

Jetzt benötigen wir eine Hilfsüberlegung.

Behauptung: Gilt $k_s = k_t$ für $s \neq t$, dann ist $\det(C(k_1, \dots, k_n)) = 0$.

Gilt $k_s = k_t$ für $s \neq t$, dann ist

$$a_{i,k_s} = a_{i,k_t}.$$

Wir setzen

$$v = \begin{pmatrix} a_{1,k_s} \\ \vdots \\ a_{n,k_s} \end{pmatrix} \in R^n.$$

Die s-te Spalte von $C(k_1, \dots, k_n)$ ist $b_{k_s,s}v$, die t-te Spalte von $C(k_1, \dots, k_n)$ ist $b_{k_t,t}v$. Da die Determinante linear in jeder Spalte ist, können wir die Faktoren $b_{k_s,s}$ und $b_{k_t,t}$ bei der Berechnung herausziehen und erhalten so eine Matrix, deren s-te und t-te Spalte jeweils v ist. Damit ist

$$\det(C(k_1, \dots, k_n)) = b_{k_s,s} b_{k_t,t} \cdot 0 = 0.$$

In der oben hergeleiteten Formel für $\det(C)$ verschwinden also viele Summanden und wir müssen nur solche Tupel $(k_1, \dots, k_n)$ betrachten, die aus paarweise verschiedenen Elementen bestehen. Dies ist genau dann der Fall, wenn

$$\begin{pmatrix} 1 & 2 & \dots & n \\ k_1 & k_2 & \dots & k_n \end{pmatrix}$$

eine Permutation ist. Wir erhalten damit

$$\begin{aligned} \det(C) &= \sum_{\tau \in S_n} \det(C(\tau(1), \dots, \tau(n))) \\ &= \sum_{\tau,\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{j=1}^{n} a_{\sigma(j),\tau(j)} b_{\tau(j),j}. \end{aligned}$$

Für jede Permutation $\tau \in S_n$ ist die Abbildung $\sigma \mapsto \sigma \circ \tau$ auf S_n bijektiv (siehe Aufgabe 1.1.6). Wenn also σ alle Elemente von S_n durchläuft, dann auch $\sigma \circ \tau$ und wir können σ durch $\sigma \circ \tau$ in der Summe ersetzen. Somit ergibt sich

$$\begin{aligned} \det(C) &= \sum_{\tau,\sigma \in S_n} \operatorname{sgn}(\sigma \circ \tau) \prod_{j=1}^{n} a_{\sigma(\tau(j)),\tau(j)} b_{\tau(j),j} \\ &= \sum_{\tau,\sigma \in S_n} \operatorname{sgn}(\sigma) \operatorname{sgn}(\tau) \prod_{j=1}^{n} a_{\sigma(\tau(j)),\tau(j)} b_{\tau(j),j} && \text{(mit 4.1.24)} \\ &= \sum_{\tau,\sigma \in S_n} \operatorname{sgn}(\sigma) \operatorname{sgn}(\tau) \prod_{k=1}^{n} a_{\sigma(k),k} \prod_{j=1}^{n} b_{\tau(j),j} && \text{(setze } k = \tau(j)\text{)} \\ &= \det(A)\det(B). \end{aligned}$$

□

4.2.22 **Korollar.** *Sei R ein kommutativer Ring. Ist $A \in \mathrm{M}_{n,n}(R)$ invertierbar, dann ist $\det(A)$ eine Einheit in R und es gilt*

$$\det(A)^{-1} = \det(A^{-1}).$$

<u>*Beweis.*</u> Ist A^{-1} die inverse Matrix zu A, dann folgt aus dem Determinantenmultiplikationssatz

$$\det(A)\det(A^{-1}) = \det(AA^{-1}) = \det(I_n) = 1.$$

Also ist $\det(A)$ eine Einheit in R und $\det(A^{-1}) = \det(A)^{-1}$. □

4.2.23 **Aufgabe.** Sei K ein Körper. Zeigen Sie, dass **L**

$$\mathrm{SL}_n(K) = \{A \in \mathrm{GL}_n(K) \mid \det(A) = 1\}$$

eine Untergruppe von $\mathrm{GL}_n(K)$ ist (siehe 1.2.15).

Die Gruppe $\mathrm{SL}_n(K)$ nennt man die *spezielle lineare Gruppe* über K.

4.2.24 **Definition** Sei R ein kommutativer Ring. Zwei Matrizen $A, B \in \mathrm{M}_{n,n}(R)$ nennt man *ähnlich*, wenn es eine invertierbare Matrix $S \in \mathrm{M}_{n,n}(R)$ gibt, sodass

$$A = S^{-1}BS$$

gilt.

4.2.25 **Ähnlichkeit ist eine Äquivalenzrelation.** Sei R ein kommutativer Ring. Für Matrizen $A, B \in \mathrm{M}_{n,n}(R)$ schreiben wir $A \sim B$, wenn diese ähnlich im Sinne von 4.2.24 sind. Dann ist $\sim$ eine Äquivalenzrelation.

reflexiv: Für alle $A \in \mathrm{M}_{n,n}(R)$ gilt $A \sim A$, denn die Einheitsmatrix I_n ist invertierbar und es gilt $A = I_n^{-1}AI_n$.

symmetrisch: Seien $A, B \in \mathrm{M}_{n,n}(R)$ mit $A \sim B$ gegeben. Es gibt also $S \in \mathrm{GL}_n(R)$ mit $S^{-1}BS = A$. Multipliziert man diese Gleichung von links mit S und von rechts mit S^{-1}, dann erhält man

$$B = (SS^{-1})B(SS^{-1}) = SAS^{-1} = (S^{-1})^{-1}AS^{-1}.$$

Da die Matrix S^{-1} ebenfalls invertierbar ist, schließen wir daraus $B \sim A$.

transitiv: Seien $A, B, C \in \mathrm{M}_{n,n}(R)$ mit $A \sim B$ und $B \sim C$ gegeben. Es gibt also invertierbare Matrizen $S, T \in \mathrm{GL}_n(R)$ mit

$$A = S^{-1}BS \quad \text{und} \quad B = T^{-1}CT.$$

Daraus folgt

$$A = S^{-1}BS = S^{-1}T^{-1}CTS = (TS)^{-1}C(TS).$$

Da $\mathrm{GL}_n(R)$ mit der Matrizenmultiplikation eine Gruppe ist (siehe 1.2.28), liegt auch TS in $\mathrm{GL}_n(R)$ (ist also invertierbar). Nach Definition von Ähnlichkeit gilt damit $A \sim C$.

Sei $R = K$ ein Körper. Die Äquivalenzklassen der Ähnlichkeitsrelation auf der Menge $\mathrm{M}_{n,n}(K)$ nennt man *Ähnlichkeitsklassen*. Das wesentliche Ziel der nächsten beiden Lektionen ist es, die Ähnlichkeitsklassen von Matrizen $\mathrm{M}_{n,n}(K)$ zu verstehen. Wir werden versuchen, ein Vertretersystem der Ähnlichkeitsklassen zu bestimmen und feststellen, dass dies ziemlich schwierig ist. Insbesondere werden wir ein konkretes Vertretersystem nur dann angeben können, wenn der Körper K algebraisch abgeschlossen ist.

4.2.26 **Korollar.** *Sei R ein kommutativer Ring. Sind $A, B \in \mathrm{M}_{n,n}(R)$ ähnlich, dann gilt*

$$\det(A) = \det(B).$$

Beweis. Es sei S eine invertierbare Matrix mit $A = S^{-1}BS$. Dann gilt

$$\begin{aligned}
\det(A) &= \det\big(S^{-1}BS\big) \\
&= \det(S^{-1})\det(B)\det(S) && (4.2.21) \\
&= \det(S)^{-1}\det(B)\det(S) && (4.2.22) \\
&= \det(B). && \square
\end{aligned}$$

IV. Matrizen über Körpern

Bisher haben wir sehr allgemeine Beobachtungen zur Determinante angestellt, die im Wesentlichen direkt aus der Leibniz-Formel hergeleitet werden konnten und über allgemeinen kommutativen Ringen gültig sind. Wir kennen aber noch keine gute Methode um die Determinante einer großen Matrix auszurechnen. Dazu wollen wir uns auf Matrizen über Körpern einschränken, denn in diesem Fall kann man die Determinante gut mithilfe von elementaren Zeilenumformungen berechnen. Wir werden uns zunächst überlegen, wie sich elementare Zeilenumformungen auf die Determinante auswirken.

4.2.27 **Satz.** *Sei R ein kommutativer Ring und sei $A \in \mathrm{M}_{n,n}(R)$. Entsteht A' aus A …*

(1) durch Vertauschen zweier Zeilen, dann ist $\det(A') = -\det(A)$,

(2) indem man eine Zeile mit $r \in R$ multipliziert, dann gilt $\det(A') = r\det(A)$.

(3) indem man ein Vielfaches einer Zeile zu einer anderen addiert, dann ist $\det(A') = \det(A)$.

Beweis. Aussage (1) kennen wir bereits aus Korollar 4.2.20. Die zweite Aussage gilt, weil die Determinante linear in jeder Zeile ist; vgl. 4.2.18 (c).

Zu (3): Sei $s \in R$ und seien $i \neq j$. Angenommen A' entsteht aus A indem man das s-fache der j-ten zur i-ten Zeile addiert. Dann ist

$$A' = T_{ij}(s)A$$

wobei $T_{ij}(s) = I_n + sE_{ij}$ die zugehörige Elementarmatrix ist; siehe Abschnitt 3.2.2 in [MG]. Wir bemerken, dass $T_{ij}(s)$ eine Dreiecksmatrix ist (genauer: eine obere für $i < j$ und eine untere für $i > j$) wobei alle Diagonaleinträge 1 sind. Daher ist $\det(T_{ij}(s)) = 1$ wegen Satz 4.2.13 und der Determinantenmultiplikationssatz liefert

$$\det(A') = \det(T_{ij}(s)) \det(A) = \det(A). \qquad \square$$

Über einem Körper kann jede Matrix mithilfe von elementaren Zeilenumformungen in eine Matrix in Treppennormalform überführt werden; siehe Abschnitt 4.2 in [MG]. Zur Berechnung der Determinante müssen wir aber gar nicht bis zur Treppennormalform umformen. Zur Berechnung der Determinante ist es ausreichend mit elementaren Zeilenumformungen zu einer oberen Dreiecksmatrix zu gelangen, weil man dann mit Satz 4.2.13 die Determinante bestimmen kann. Damit erhält man das folgende einfache Verfahren.

4.2.28 **Berechnung der Determinante mit elementaren Zeilenumformungen**. Sei K ein Körper.

`Gegeben:` $A \in \mathrm{M}_{n,n}(K)$

`Gesucht:` $\det(A)$

`Verfahren:`

Überführe A mithilfe der elementaren Zeilenumformungen (mit $i \neq j$)

$V_{i,j}$: Vertausche die i-te und j-te Zeile

$Z_i(r)$: Multipliziere die i-te Zeile mit $r \in K^\times$

$Z_{i,j}(s)$: Addiere das s-fache der j-ten zur i-ten Zeile

in eine obere Dreiecksmatrix $B = (b_{ij})_{i,j}$. Ist t die Anzahl der insgesamt vorgenommenen Vertauschungen $V_{i,j}$ und wurde Z_i insgesamt u-mal mit $r_1, \ldots, r_u \in K^\times$ eingesetzt, dann gilt

$$\det(A) = (-1)^t r_1^{-1} \cdots r_u^{-1} \prod_{j=1}^{n} b_{jj}.$$

4.2.29 Bemerkung. Im Verfahren aus 4.2.28 kann man auf die elementare Zeilenumformung $Z_i(r)$ verzichten, denn auch ohne diesen Umformungsschritt kann man zu einer oberen Dreiecksmatrix gelangen.

Da $Z_i(r)$ zur Berechnung der Determinante nicht benötigt wird und immer eine gewisse Gefahr besteht, dass man in der Formel am Ende den Korrekturterm r^{-1} vergißt, empfehle ich diese Umformung möglichst zu vermeiden.

4.2.30 Beispiel. Wir berechnen die Determinante von

$$A = \begin{pmatrix} 0 & 0 & 3 & 1 & -1 \\ 2 & 0 & 1 & 3 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 6 & -5 & 1 \\ 0 & 0 & 0 & 2 & -3 \end{pmatrix} \in \mathrm{M}_{n,n}(\mathbb{Q}).$$

Zuerst vertauschen wir die erste und die dritte Zeile und erhalten

$$A_1 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 2 & 0 & 1 & 3 & 0 \\ 0 & 0 & 3 & 1 & -1 \\ 0 & 0 & 6 & -5 & 1 \\ 0 & 0 & 0 & 2 & -3 \end{pmatrix}.$$

Dann ziehen wir das 2-fache der ersten von der zweiten Zeile ab; weiter ziehen wir das 2-fache der dritten von der vierten Zeile ab.

$$A_3 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & -2 & -1 & 1 & -2 \\ 0 & 0 & 3 & 1 & -1 \\ 0 & 0 & 0 & -7 & 3 \\ 0 & 0 & 0 & 2 & -3 \end{pmatrix}$$

Als letzten Schritt addieren wir das $\frac{2}{7}$-fache der vierten zur letzten Zeile und erhalten damit die obere Dreiecksmatrix

$$B = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & -2 & -1 & 1 & -2 \\ 0 & 0 & 3 & 1 & -1 \\ 0 & 0 & 0 & -7 & 3 \\ 0 & 0 & 0 & 0 & -\frac{15}{7} \end{pmatrix}.$$

Wir haben genau eine Vertauschung vorgenommen, also ist

$$\det(A) = (-1)^1 \cdot \left(1 \cdot (-2) \cdot 3 \cdot (-7) \cdot (-\frac{15}{7})\right) = 90.$$

(Da die Einträge der Matrix A ganze Zahlen sind, wissen wir aus der Leibniz-Formel, dass die Determinante auch eine ganze Zahl sein muss. Dadurch besteht die Möglichkeit Rechenfehler zu erkennen: Die im Verfahren auftretenden Brüche müssen sich wieder wegheben!)

4.2.31 **Aufgabe**. Berechnen Sie mit dem Verfahren aus 4.2.28 die Determinante der Matrix **L**

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 & 5 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 2 & 3 & 0 & 2 \\ 1 & 1 & 0 & 3 & 12 \\ -1 & -1 & 0 & 0 & -3 \end{pmatrix} \in \mathrm{M}_{n,n}(\mathbb{Q}).$$

4.2.32 **Korollar.** *Sei K ein Körper und sei $A \in \mathrm{M}_{n,n}(K)$. Folgende Aussagen sind äquivalent:*

(i) $\mathrm{Ker}(A) = \{0\}$.

(ii) A hat Rang n.

(iii) A ist invertierbar.

(iv) $\det(A) \neq 0$

Beweis. Der Rangsatz [MG, 8.3.14] besagt, dass

$$n = \mathrm{Rg}(A) + \dim_K(\mathrm{Ker}(A))$$

gilt. Daraus folgt die Äquivalenz von (i) und (ii), denn $\dim_K(\mathrm{Ker}(A)) = 0$ genau dann, wenn $\mathrm{Ker}(A) = \{0\}$ ist.

Die Äquivalenz von (ii) und (iii) ist aus Proposition 4.5.4 in [MG] bekannt. Aus Korollar 4.2.22 folgt „(iii) $\implies$ (iv)“.

Wir nehmen nun $\det(A) \neq 0$ an und zeigen, dass A vollen Rang hat. Durch elementare Zeilenumformungen bringen wir A auf ihre Treppennormalform T. Aus Satz 4.2.27 sehen wir, dass die Determinante durch eine elementare Zeilenumformung nicht Null werden kann (zur Erinnerung: es ist nur die Multiplikation einer Zeile mit $r \neq 0$ erlaubt), d.h., es gilt

$$\det(T) \neq 0.$$

Insbesondere hat T keine Nullzeile (siehe 4.2.20) und es muss somit $T = I_n$ sein. Mit Proposition 4.5.4 aus [MG] schließen wir, dass A Rang n hat. □

4.3. Minoren und die Adjunkte

In diesem Abschnitt werden wir uns mit den sogenannten Minoren einer Matrix beschäftigen. Dabei handelt es sich um Determinanten von Untermatrizen. Mithilfe von Minoren werden wir dann die *Adjunkte* einer Matrix definieren.

I. Minoren

4.3.1 $\boxed{\textbf{Definition}}$ Sei R ein kommutativer Ring und sei $A \in \mathrm{M}_{m,n}(R)$. Sei $1 \leq k \leq \min(m,n)$. Ein *k-reihiger Minor* von A ist die Determinante einer $(k \times k)$-Matrix, die aus A durch das Löschen von $(m-k)$ Zeilen und $(n-k)$ Spalten entsteht.

Minore werden gelegentlich auch *Unterdeterminanten* genannt.

4.3.2 **Beispiel**. Sei

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 0 & 0 \end{pmatrix} \in \mathrm{M}_{3,4}(\mathbb{R}).$$

Streichen wir beispielsweise die erste Zeile sowie die zweite und vierte Spalte von A,

$$\begin{pmatrix} \cancel{1} & \cancel{2} & \cancel{3} & \cancel{4} \\ 5 & \cancel{6} & 7 & \cancel{8} \\ 9 & \cancel{10} & 0 & \cancel{0} \end{pmatrix}$$

dann bleibt die Untermatrix

$$A' = \begin{pmatrix} 5 & 7 \\ 9 & 0 \end{pmatrix}$$

übrig. Die Determinante von A' ist $\det(A') = 0 - 9 \cdot 7 = -63$. Also ist -63 ein 2-reihiger Minor von A.

4.3.3 **Beispiel**. Sei $A \in \mathrm{M}_{m,n}(R)$. Die 1-reihigen Minoren sind genau die Einträge von A. Ist A quadratisch, d.h. $n = m$, dann ist die Determinante $\det(A)$ der einzige n-reihige Minor von A.

Uns interessieren im Folgenden speziell $(n-1)$-reihige Minore von $(n \times n)$-Matrizen.

4.3.4 **Notation**. Sei $A \in \mathrm{M}_{n,n}(R)$. Für $i, j \in \{1, 2, \ldots, n\}$ bezeichnen wir mit A_{ij} die Matrix, die aus A durch Entfernen der i-ten Zeile und j-ten Spalte entsteht.

4.3.5 **Beispiel**. Wir bestimmen die 2-reihigen Minoren der Matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 0 & -1 \\ 4 & 1 & 4 \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{Q}).$$

Dazu bestimmen wir die neun Matrizen A_{ij} mit $1 \le i, j \le 3$.

i/j	1	2	3
1	$A_{11} = \begin{pmatrix} 0 & -1 \\ 1 & 4 \end{pmatrix}$	$A_{12} = \begin{pmatrix} -1 & -1 \\ 4 & 4 \end{pmatrix}$	$A_{13} = \begin{pmatrix} -1 & 0 \\ 4 & 1 \end{pmatrix}$
2	$A_{21} = \begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix}$	$A_{22} = \begin{pmatrix} 1 & 3 \\ 4 & 4 \end{pmatrix}$	$A_{23} = \begin{pmatrix} 1 & 2 \\ 4 & 1 \end{pmatrix}$
3	$A_{31} = \begin{pmatrix} 2 & 3 \\ 0 & -1 \end{pmatrix}$	$A_{32} = \begin{pmatrix} 1 & 3 \\ -1 & -1 \end{pmatrix}$	$A_{33} = \begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix}$

Mit 4.2.3 finden wir damit die folgenden Minoren.

i/j	1	2	3
1	$\det(A_{11}) = 1$	$\det(A_{12}) = 0$	$\det(A_{13}) = -1$
2	$\det(A_{21}) = 5$	$\det(A_{22}) = -8$	$\det(A_{23}) = -7$
3	$\det(A_{31}) = -2$	$\det(A_{32}) = 2$	$\det(A_{33}) = 2$

4.3.6 **Aufgabe**. Bestimmen Sie alle 2-reihigen Minoren der Matrix **L**

$$A = \begin{pmatrix} \overline{1} & \overline{4} & \overline{3} \\ \overline{0} & \overline{2} & \overline{3} \\ \overline{1} & \overline{0} & \overline{2} \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{F}_5).$$

4.3.7 **Aufgabe**. Sei A eine $(n \times n)$-Matrix. Erläutern Sie, warum folgende Beziehung gilt: **L**

$$(A^T)_{ij} = (A_{ji})^T.$$

II. Entwicklungssatz

Mithilfe der Minoren können wir nun eine neue Methode zur Berechnung von Determinanten herleiten: den Entwicklungssatz von Laplace[2]. Dabei wird die Determinante $\det(A)$ mithilfe ihrer „Kofaktoren“ (Minoren mit Vorzeichen) ausgedrückt. Man kann somit die Determinante einer $(n \times n)$-Matrix berechnen, indem man die Determinanten mehrerer Matrizen der Größe $(n-1) \times (n-1)$ bestimmt. Wiederholt man diesen Prozess, gelangt man irgendwann zu Matrizen der Größe 3×3 oder 2×2 und in diesen Fällen können wir die Determinanten leicht berechnen.

4.3.8 **Laplace'scher Entwicklungssatz.** *Sei R ein kommutativer Ring und sei $n \geq 2$. Sei $A \in \mathrm{M}_{n,n}(R)$.*

(a) *Für jedes $j \in \{1, 2, \ldots, n\}$ gilt*

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}).$$

(b) *Für jedes $i \in \{1, 2, \ldots, n\}$ gilt*

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}).$$

Die Gleichung in (a) nennt man die Formel für die *Entwicklung an der j-ten Spalte*, die Gleichung aus (b) nennt man die Formel für die *Entwicklung an der i-ten Zeile*. Bevor wir einen Beweis des Entwicklungssatzes geben, sehen wir uns die Entwicklung an einer Zeile im Beispiel an.

4.3.9 **Beispiel**. Wir betrachten nochmal die (3×3)-Matrix

$$A = \begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{Z})$$

aus Beispiel 4.2.7. Wir bestimmen die Determinante nun durch Entwicklung an der ersten Zeile. Das heißt, wir gehen die Einträge der ersten Zeile der Reihe nach durch. Beim j-ten Eintrag bestimmen wir die Untermatrix $A_{1,j}$, die durch streichen der ersten Zeile und der j-ten Spalte entsteht und bestimmen jeweils den zugehörigen

[2]Pierre-Simon Laplace: französischer Gelehrter, 1749–1827.

Minor $\det(A_{1,j})$. Das Ergebnis multiplizieren wir mit $(-1)^{1+j}a_{1,j}$ und addieren auf. Damit erhalten wir:

$$\begin{aligned}\det(A) &= (-1)^2 \cdot 0 \cdot \det\begin{pmatrix}4 & 5\\ 7 & 8\end{pmatrix} + (-1)^3 \cdot 1 \cdot \det\begin{pmatrix}3 & 5\\ 6 & 8\end{pmatrix} + (-1)^4 \cdot 2 \cdot \det\begin{pmatrix}3 & 4\\ 6 & 7\end{pmatrix}\\ &= -1 \cdot (24-30) + 2 \cdot (21-24) = 6 - 6 = 0.\end{aligned}$$

4.3.10 **Aufgabe**. Berechnen Sie die Determinante der Matrix **L**

$$A = \begin{pmatrix}2 & 3+i & 0\\ 2 & 4 & 2\\ 3 & 5i & 1+i\end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{C})$$

durch Entwicklung an der dritten Spalte.

4.3.11 **Bemerkungen zum Entwicklungssatz**. (a) Ist $a_{ij} = 0$, dann kann man sich die Berechnung des Minors $\det(A_{ij})$ sparen. Zur Entwicklung sollte man also möglichst eine Zeile oder Spalte wählen, die nur wenige Einträge ungleich Null enthält. Findet man keine solche Zeile oder Spalte, dann ist (über einem Körper) die Berechnung mithilfe von elementaren Zeilenumformungen meist einfacher.

(b) Die Vorzeichen $(-1)^{i+j}$ im Entwicklungssatz kann man sich gut merken, indem man sich folgendes Schachbrettmuster einprägt:

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

4.3.12 **Aufgabe**. Berechnen Sie mit dem Entwicklungssatz die Determinante der Matrix **L**

$$A = \begin{pmatrix}2 & -1 & 0 & 4\\ 3 & 2 & 4 & 28\\ -2 & 0 & 0 & -5\\ 1 & 1 & 0 & -2\end{pmatrix} \in \mathrm{M}_{n,n}(\mathbb{R}).$$

4.3.13 Beweis des Entwicklungssatzes. Als ersten Schritt leiten wir die Spaltenentwicklungsformel 4.3.8 (a) aus der Formel für die Entwicklung an einer Zeile her. Dazu verwenden wir wieder die Transponierte A^T. Es seien $\tilde{a}_{ij}$ die Einträge von A^T, d.h., $\tilde{a}_{ij} = a_{ji}$. Wir erinnern außerdem an die Beziehung $(A^T)_{ij} = (A_{ji})^T$ aus Aufgabe 4.3.7. Sei nun $j \in \{1, 2 \dots, n\}$ gegeben. Wir setzen 4.3.8 (b) als bekannt voraus und entwickeln dazu A^T an der j-ten Zeile. Dann erhalten wir mit Lemma 4.2.14

$$\begin{aligned} \det(A) = \det(A^T) &= \sum_{k=1}^{n} (-1)^{j+k} \tilde{a}_{jk} \det((A^T)_{jk}) \\ &= \sum_{k=1}^{n} (-1)^{k+j} a_{kj} \det((A_{kj})^T) \\ &= \sum_{k=1}^{n} (-1)^{k+j} a_{kj} \det(A_{kj}). && \text{(Lemma 4.2.14)} \end{aligned}$$

Das ist genau die Formel für die Entwicklung an der j-ten Spalte von A.

Es ist also ausreichend 4.3.8 (b) zu beweisen. Sei dazu $i \in \{1, 2, \dots, n\}$ fest gewählt. Den Beweis führen wir mithilfe des Charakterisierungssatzes. Dazu definieren wir eine Funktion $\Delta\colon \mathrm{M}_{n,n}(R) \to R$ durch

$$\Delta(B) = \sum_{j=1}^{n} (-1)^{i+j} b_{ij} \det(B_{ij}).$$

Es gilt

$$\Delta(I_n) = (-1)^{i+i} \det(I_{n-1}) = 1.$$

Wir zeigen nun, dass Δ alternierend und linear in jeder Spalte ist. Aus dem Charakterisierungssatz 4.2.17 folgt dann $\Delta(B) = \det(B)$ für alle $B \in \mathrm{M}_{n,n}(R)$ und damit gilt der Entwicklungssatz.

Δ ist linear in jeder Spalte:
Sei dazu $k \in \{1, 2, \dots, n\}$. Es seien

$$B = (b_1|b_2|\dots|b_k|\dots|b_n) \quad \text{und} \quad B' = (b_1|b_2|\dots|b'_k|\dots|b_n)$$

zwei Matrizen, die sich höchstens in der k-ten Spalte unterscheiden. Wir setzen $C = (b_1|b_2|\dots|b_k + b'_k|\dots|b_n)$. Für $k < j$, ist die k-te Spalte von C_{ij} die Summe der k-ten Spalte von B_{ij} und B'_{ij}. Für $k > j$ ist die $(k-1)$-te Spalte von C_{ij} die Summe der $(k-1)$-ten Spalte von B_{ij} und B'_{ij}. Für $k = j$ ist $C_{ij} = B_{ij} = B'_{ij}$. Daraus ergibt sich

$$\Delta(C) = \sum_{j \neq k} (-1)^{i+j} b_{ij} \det(C_{ij}) + (-1)^{i+k} (b_{ik} + b'_{ik}) \det(C_{ik})$$

$$\begin{aligned}
&= \sum_{j\neq k} (-1)^{i+j} b_{ij} \big(\det(B_{ij}) + \det(B'_{ij})\big) + (-1)^{i+k}\big(b_{ik}\det(B_{ik}) + b'_{ik}\det(B'_{ik})\big) \\
&= \sum_{j=1}^{n} (-1)^{i+j} b_{ij} \det(B_{ij}) + \sum_{j=1}^{n} (-1)^{i+j} b'_{ij} \det(B'_{ij}) \\
&= \Delta(B) + \Delta(B').
\end{aligned}$$

Sei $r \in R$ und sei nun $B' = (b_1|b_2|\cdots|rb_k|\ldots|b_n)$. Wir überlassen es der Leserin/dem Leser als Übungsaufgabe auch die Gleichung **L**

$$\Delta(B') = r\Delta(B)$$

nachzurechnen. Insgesamt ist Δ damit linear in jeder Spalte.

Δ *ist alternierend:*
Es sei $B = (b_1|b_2|\ldots|b_n)$ eine Matrix mit zwei gleichen Spalten; sagen wir $b_s = b_t$ mit $s < t$. Ist $j \neq s$ und $j \neq t$, so haben die Matrizen B_{ij} jeweils zwei gleiche Spalten und es gilt $\det(B_{ij}) = 0$. Wir müssen also nur noch die Terme B_{is} und B_{it} betrachten. Die Matrizen B_{is} und B_{it} unterscheiden sich nur leicht: die $(t-1)$-Spalte von B_{is} taucht in B_{it} schon als s-te Spalte auf. Man kann also B_{is} in B_{it} überführen, indem man die $(t-1)$-Spalte Schritt für Schritt mit den Spalten $(t-2), (t-3), \ldots, s$ vertauscht. Es gilt also $\det(B_{it}) = (-1)^{t-s-1}\det(B_{is})$. Insgesamt erhalten wir mit $b_{is} = b_{it}$ nun

$$\begin{aligned}
\Delta(B) &= (-1)^{i+s} b_{is} \det(B_{is}) + (-1)^{i+t} b_{it} \det(B_{it}) \\
&= (-1)^{i+s} b_{is} \det(B_{is}) + (-1)^{i+t+t-s-1} b_{is} \det(B_{is}) \\
&= (-1)^{i+s} b_{is} \det(B_{is}) - (-1)^{i+s} b_{is} \det(B_{is}) = 0.
\end{aligned}$$

□

Als Anwendung des Entwicklungssatzes wollen wir eine Formel für die Determinante von „Blockdreiecksmatrizen“ herleiten.

4.3.14 $\boxed{\textbf{Definition}}$ Sei $A = (a_{ij})_{i,j} \in \mathrm{M}_{n,n}(R)$. Wir sagen, dass A eine *Blockdreiecksmatrix* ist, falls es $k < n$ gibt, sodass $a_{ij} = 0$ für alle $j \leq k < i$ gilt.

Das heißt, A hat folgende Gestalt

$$\begin{array}{cc}
\begin{array}{cc} \overbrace{\quad}^{k} & \overbrace{\quad}^{n-k} \end{array} & \\
\left(\begin{array}{c|c} B & D \\ \hline 0 & C \end{array}\right) & \begin{array}{l} \}\,k \\ \}\,n-k \end{array}
\end{array}$$

mit einer $(n-k \times k)$-Nullmatrix unten links und Matrizen $B \in \mathrm{M}_{k,k}(R), C \in \mathrm{M}_{n-k,n-k}(R), D \in \mathrm{M}_{k,n-k}(R)$.

4.3.15 **Satz.** *Sei $A \in \mathrm{M}_{n,n}(R)$ eine Blockdreiecksmatrix der Form*

$$A = \left(\begin{array}{c|c} B & D \\ \hline 0 & C \end{array}\right)$$

mit $B \in \mathrm{M}_{k,k}(R), C \in \mathrm{M}_{n-k,n-k}(R), D \in \mathrm{M}_{k,n-k}(R)$. Dann gilt

$$\det(A) = \det(B)\det(C).$$

Beweis. Wir schreiben A als Produkt $A = C'D'B'$ der drei Blockdreiecksmatrizen

$$B' = \left(\begin{array}{c|c} B & 0 \\ \hline 0 & I_{n-k} \end{array}\right), \quad C' = \left(\begin{array}{c|c} I_k & 0 \\ \hline 0 & C \end{array}\right), D' = \left(\begin{array}{c|c} I_k & D \\ \hline 0 & I_{n-k} \end{array}\right).$$

Um nachzuprüfen, dass $A = C'D'B'$ ist, muss man sich einmal davon überzeugen, dass man mit den Blöcken nach denselben Regeln wie bei der Multiplikation von (2×2)-Matrizen rechnen darf. Nach dem Determinantenmultiplikationssatz 4.2.21 gilt nun

$$\det(A) = \det(C')\det(D')\det(B').$$

Da D' eine obere Dreiecksmatrix mit Einsern auf der Diagonale ist, gilt $\det(D') = 1$ nach Satz 4.2.13. Als nächstes bestimmen wir die Determinante von

$$C' = \left(\begin{array}{cccc|ccc} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & & \vdots & \vdots & & \vdots \\ \vdots & & \ddots & 0 & \vdots & & \vdots \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ \hline 0 & \dots & \dots & 0 & & & \\ \vdots & & & \vdots & & C & \\ 0 & \dots & \dots & 0 & & & \end{array}\right).$$

Dazu entwickeln wir an der ersten Spalte. Da in der ersten Spalte fast alle Einträge verschwinden und der Eintrag an der Stelle $(1,1)$ gleich 1 ist, erhalten wir $\det(C') = \det(C'_{11})$ aus dem Entwicklungssatz 4.3.8. Wir können also die erste Zeile und Spalte streichen ohne die Determinante zu verändern. Diesen Trick können wir nacheinander auf die ersten k Spalten anwenden und erhalten $\det(C') = \det(C)$. Ähnlich verfahren wir mit der Matrix B'. Wir entwickeln an der letzten Spalte und erhalten $\det(B') = \det(B'_{nn})$. Wir wiederholen diesen Schritt $(n-k)$-mal um am Ende $\det(B') = \det(B)$ zu folgern. □

4.3.16 **Beispiel**. Gegeben ist die Matrix

$$A = \begin{pmatrix} 3 & 4 & 231 & -72 \\ 2 & 3 & 111 & -32 \\ 0 & 0 & 7 & -1 \\ 0 & 0 & 6 & 2 \end{pmatrix} \in \mathrm{M}_{4,4}(\mathbb{Z}).$$

Da A eine Blockdreiecksmatrix ist, können wir die Determinante mit Satz 4.3.15 und 4.2.3 berechnen. Wir erhalten

$$\det(A) = \det \begin{pmatrix} 3 & 4 \\ 2 & 3 \end{pmatrix} \cdot \det \begin{pmatrix} 7 & -1 \\ 6 & 2 \end{pmatrix} = (9-8)(14+6) = 20.$$

III. Adjunktensatz

In diesem Abschnitt befassen wir uns mit der „Adjunkten“.

4.3.17 $\boxed{\textbf{Definition}}$ Sei R ein kommutativer Ring und sei $A \in \mathrm{M}_{n,n}(R)$. Die *Adjunkte* von A ist die Matrix $\mathrm{Adj}(A) \in \mathrm{M}_{n,n}(R)$ deren Eintrag an der Stelle (i,j) genau $(-1)^{i+j}\det(A_{ji})$ ist.

4.3.18 **Reihenfolge beachten!**. Bei der Definition der Adjunkten muss man genau lesen: der Eintrag an der Stelle (i,j) ist wirklich $(-1)^{i+j}\det(A_{ji})$. Im Index stehen dabei j und i in der umgekehrten Reihenfolge!

4.3.19 **Beispiel**. Wir betrachten eine allgemeine (2×2)-Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \mathrm{M}_{2,2}(R).$$

Die 1-reihigen Minoren kann man direkt ablesen:

$$\det(A_{11}) = a_{22}, \quad \det(A_{12}) = a_{21}, \quad \det(A_{21}) = a_{12}, \quad \det(A_{22}) = a_{11}.$$

Damit erhalten wir folgende Beschreibung für die Adjunkte:

$$\mathrm{Adj}(A) = \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}.$$

4.3.20 Beispiel. Wir betrachten nochmal die Matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 0 & -1 \\ 4 & 1 & 4 \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{Q})$$

aus Beispiel 4.3.5. Die Minoren haben wir im Beispiel 4.3.5 bestimmt. Wir müssen diese noch in der richtigen Anordnung und mit dem passenden Vorzeichen in die Matrix $\mathrm{Adj}(A)$ übertragen. Das Ergebnis ist

$$\mathrm{Adj}(A) = \begin{pmatrix} 1 & -5 & -2 \\ 0 & -8 & -2 \\ -1 & 7 & 2 \end{pmatrix}.$$

Nun machen wir ein Experiment und multiplizieren A mit der Adjunkten $\mathrm{Adj}(A)$. Wir erhalten ein überraschendes Ergebnis:

$$A\,\mathrm{Adj}(A) = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 0 & -1 \\ 4 & 1 & 4 \end{pmatrix} \begin{pmatrix} 1 & -5 & -2 \\ 0 & -8 & -2 \\ -1 & 7 & 2 \end{pmatrix} = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

Wenn man jetzt auch noch die Determinante von A ausrechnet, stellt man fest, dass $\det(A) = -2$ ist. Ist das ein Zufall?

4.3.21 Aufgabe. Bestimmen Sie die Adjunkte der Matrix A aus Aufgabe 4.3.6 und Berechnen Sie damit $A\,\mathrm{Adj}(A)$. **L**

4.3.22 Aufgabe. Zeigen Sie, dass $\mathrm{Adj}(A^T) = \mathrm{Adj}(A)^T$ für jede quadratische Matrix A gilt. **L**

4.3.23 Adjunktensatz. *Sei R ein kommutativer Ring. Für jede Matrix $A \in \mathrm{M}_{n,n}(R)$ gilt*

$$A\,\mathrm{Adj}(A) = \mathrm{Adj}(A)A = \det(A)I_n.$$

Beweis. Seien $i, k \in \{1, 2, \ldots, n\}$. Nach Definition der Matrizenmultiplikation ist der Eintrag c_{ik} an der Stelle (i,k) von $A\,\mathrm{Adj}(A)$ gegeben durch

$$c_{ik} = \sum_{j=1}^{n} a_{i,j}(-1)^{k+j}\det(A_{kj}).$$

Für $k = i$ ergibt sich daraus die Formel für die Entwicklung an der i-ten Zeile von A und der Entwicklungssatz 4.3.8 liefert $c_{ii} = \det(A)$.

Sei nun $i \neq k$. Es bleibt zu zeigen, dass dann $c_{ik} = 0$ gilt. Es sei $\widetilde{A} = (\tilde{a}_{st})_{s,t}$ die Matrix, die sich aus A ergibt indem man die k-te Zeile durch die i-te Zeile von A ersetzt, d.h., $\tilde{a}_{st} = a_{st}$ für $s \neq k$ und $\tilde{a}_{kt} = a_{it}$ für alle $1 \leq t \leq n$. Die Matrix $\tilde{A}$ hat insbesondere zwei gleiche Zeilen und es ist $\det(\widetilde{A}) = 0$; siehe 4.2.17 und 4.2.18 (c). Damit erhalten wir

$$\begin{aligned} c_{ik} &= \sum_{j=1}^{n} a_{i,j}(-1)^{k+j}\det(A_{kj}) \\ &= \sum_{j=1}^{n} \tilde{a}_{k,j}(-1)^{k+j}\det(\tilde{A}_{kj}) && (A_{kj} = \widetilde{A}_{kj}) \\ &= \det(\widetilde{A}) = 0 \end{aligned}$$

und die Gleichheit $A\operatorname{Adj}(A) = \det(A)I_n$ ist gezeigt.

Die Gleichung $\operatorname{Adj}(A)A = \det(A)I_n$ können wir wieder mithilfe der transponierten Matrix daraus ableiten. Es gilt

$$\begin{aligned} (\operatorname{Adj}(A)A)^T &= A^T \operatorname{Adj}(A)^T && (1.2.32(ii)) \\ &= A^T \operatorname{Adj}(A^T) && (\text{Aufgabe } 4.3.22) \\ &= \det(A^T)I_n = \det(A)I_n && (\text{Lemma } 4.2.14). \end{aligned}$$

und damit auch $\operatorname{Adj}(A)A = ((\operatorname{Adj}(A)A)^T)^T = \det(A)I_n^T = \det(A)I_n$. □

Als Anwendung des Adjunktensatzes können wir nun auch die Invertierbarkeit von Matrizen über Ringen mithilfe der Determinante beschreiben.

4.3.24 **Satz.** *Sei R ein kommutativer Ring. Eine Matrix $A \in \mathrm{M}_{n,n}(R)$ ist genau dann invertierbar, wenn $\det(A)$ eine Einheit in R ist. Es gilt dann*

$$A^{-1} = \det(A)^{-1}\operatorname{Adj}(A).$$

Beweis. Aus Korollar 4.2.22 wissen wir bereits, dass $\det(A)$ eine Einheit ist, falls A invertierbar ist. Nehmen wir umgekehrt an, dass $\det(A)$ eine Einheit in R ist. Dann existiert das multiplikative Inverse $\det(A)^{-1} \in R$. Wir definieren

$$B = \det(A)^{-1}\operatorname{Adj}(A).$$

Mit dem Adjunktensatz erhalten wir

$$AB = A(\det(A)^{-1}\operatorname{Adj}(A)) = \det(A)^{-1}A\operatorname{Adj}(A) = \det(A)^{-1}\det(A)I_n = I_n$$

und auch

$$BA = \det(A)^{-1}\operatorname{Adj}(A)A = \det(A)^{-1}\det(A)I_n = I_n.$$

Also ist A invertierbar und B ist die inverse Matrix zu A; siehe 1.2.13 für die Definition von Invertierbarkeit. □

4.3.25 **Beispiel**. Wir betrachten nochmal die rationale Matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 0 & -1 \\ 4 & 1 & 4 \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{Q})$$

aus Beispiel 4.3.5. Die Adjunkte haben wir bereits in 4.3.20 bestimmt. Es gilt außerdem $\det(A) = -2$. Daher ist die Matrix A invertierbar und es gilt

$$A^{-1} = -\frac{1}{2}\operatorname{Adj}(A) = \begin{pmatrix} -\frac{1}{2} & \frac{5}{2} & 1 \\ 0 & 4 & 1 \\ \frac{1}{2} & -\frac{7}{2} & -1 \end{pmatrix}.$$

4.3.26 **Beispiel**. Sei R ein kommutativer Ring und sei

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \mathrm{M}_{2,2}(R).$$

Falls $\det(A) = a_{11}a_{22} - a_{12}a_{21}$ eine Einheit in R ist, dann ist A invertierbar und Satz 4.3.24 liefert mit Beispiel 4.3.19 auch direkt eine Formel für die inverse Matrix:

$$A^{-1} = \det(A)^{-1}\operatorname{Adj}(A) = (a_{11}a_{22} - a_{12}a_{21})^{-1}\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}.$$

4.3.27 **Beispiel**. Ist $R = K$ ein Körper, dann sind die Einheiten von K alle von 0 verschiedenen Elemente. Die Bedingung $\det(A) \neq 0$ für Invertierbarkeit kennen wir in diesem Fall schon aus 4.2.32. In allgemeinen Ringen folgt aus $\det(A) \neq 0$ *nicht*, dass die Matrix A invertierbar ist. Betrachten wir den Ring $R = \mathbb{Z}/6\mathbb{Z}$. Da 6 keine Primzahl ist, ist R kein Körper; das folgt aus Satz 1.3.14.

Die Matrix

$$A = \begin{pmatrix} \overline{2} & \overline{2} \\ \overline{0} & \overline{2} \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{Z}/6\mathbb{Z})$$

hat die Determinante

$$\det(A) = \overline{2}\cdot\overline{2} = \overline{4}.$$

Es gilt $\det(A) \neq \overline{0}$, aber trotzdem ist A nicht invertierbar: $\overline{4}$ ist *keine* Einheit in $\mathbb{Z}/6\mathbb{Z}$ (Übungsaufgabe!). **L**

Als weitere schöne Anwendung des Adjunktensatzes wollen wir die sogenannte *Cramer'sche Regel* zum Lösen von linearen Gleichungssystemen herleiten.

4.3.28 **Cramer'sche Regel.** *Sei K ein Körper und sei $A = (a_1|a_2|\cdots|a_n) \in \mathrm{M}_{n,n}(K)$. Sei $b \in K^n$. Wir schreiben*

$$A(b,j) = (a_1|a_2|\ldots|a_{j-1}|b|a_{j+1}|\ldots|a_n)$$

für die Matrix, die man erhält, indem man die j-te Spalte von A durch b ersetzt.

Ist A invertierbar, dann ist die eindeutige Lösung $x \in K^n$ des linearen Gleichungssystems

$$Ax = b$$

gegeben durch

$$x_j = \det(A)^{-1} \det\big(A(b,j)\big)$$

für alle $j \in \{1, 2 \ldots, n\}$.

Man kann die Lösung des linearen Gleichungssystems $Ax = b$ also bestimmen, indem man insgesamt $n+1$ Determinanten berechnet: die Determinante von A sowie die Determinanten der Matrizen $A(b,j)$ mit $1 \leq j \leq n$. Die Cramer'sche Regel ist in der Praxis nicht nützlich, denn die Berechnung der Determinanten ist im Allgemeinen deutlich aufwändiger als das Gauß-Verfahren.

Beweis. Wir berechnen die Determinante von $A(b,j)$. Es sei $e_i \in K^n$ der i-te Standardbasisvektor. Es gilt $b = \sum_{i=1}^n b_i e_i$. Da die Determinante linear in jeder Spalte ist (siehe 4.2.17), gilt

$$\det(A(b,j)) = \sum_{i=1}^{n} b_i \det(A(e_i,j)). \tag{4.3.a}$$

Die Matrix $A(e_i,j)$ hat in der j-ten Spalte genau einen Eintrag: eine 1 in der i-ten Zeile. Durch Entwicklung an der j-ten Spalte erhalten wir

$$\det(A(e_i,j)) = (-1)^{i+j} \det(A(e_i,j)_{ij}) = (-1)^{i+j} \det(A_{ij}),$$

denn $A(e_i,j)_{ij} = A_{ij}$. Wir setzen dieses Ergebnis in Gleichung (4.3.a) ein und finden

$$\det(A(b,j)) = \sum_{i=1}^{n} b_i \underbrace{(-1)^{i+j} \det(A_{ij})}_{\text{Eintrag } (j,i) \text{ von } \mathrm{Adj}(A)}.$$

Damit ist $\det(A(b,j))$ der j-te Eintrag des Vektors $\operatorname{Adj}(A)b$. Ist x die Lösung des Gleichungssystems $Ax = b$, dann gilt

$$x = A^{-1}b = \det(A)^{-1}\operatorname{Adj}(A)b$$

und daraus folgt die Behauptung. □

4.3.29 Beispiel. Wir betrachten das lineare Gleichungssystem

$$\begin{aligned} 2x_1 + 2x_2 &= -1 \\ 3x_1 + 4x_2 &= 1 \end{aligned}$$

über den reellen Zahlen. Die Koeffizientenmatrix

$$A = \begin{pmatrix} 2 & 2 \\ 3 & 4 \end{pmatrix}$$

hat die Determinante $\det(A) = 8 - 6 = 2$. Damit ist A invertierbar (siehe 4.2.32) und das Gleichungssystem hat eine eindeutige Lösung. Wir bestimmen die Lösung mit der Cramer'schen Regel. Wir berechnen

$$\begin{aligned} \det(A(b,1)) &= \det\begin{pmatrix} -1 & 2 \\ 1 & 4 \end{pmatrix} = -6 \\ \det(A(b,2)) &= \det\begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix} = 5 \end{aligned}$$

und schließen, dass

$$x_1 = \frac{-6}{2} = -3 \quad \text{und} \quad x_2 = \frac{5}{2}$$

die Lösung des Gleichungssystems ist.

4.3.30 Beispiel. Wir betrachten nochmal die komplexe Matrix

$$A = \begin{pmatrix} 2 & 3+i & 0 \\ 2 & 4 & 2 \\ 3 & 5i & 1+i \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{C})$$

aus Aufgabe 4.3.10; aus dieser Aufgabe wissen wir, dass $\det(A) = 22 - 14i$ gilt und A somit invertierbar ist (siehe 4.3.24). Wir wollen das lineare Gleichungssystem $Ax = b$ mit

$$b = \begin{pmatrix} 0 \\ 10 \\ 0 \end{pmatrix}$$

lösen. Dazu berechnen wir die Determinanten der Matrizen $A(b,j)$ mit $j \in \{1,2,3\}$ jeweils durch Entwicklung an der j-ten Spalte:

$$\det(A(b,1)) = \det\begin{pmatrix} 0 & 3+i & 0 \\ 10 & 4 & 2 \\ 0 & 5i & 1+i \end{pmatrix} = -10 \cdot \det\begin{pmatrix} 3+i & 0 \\ 5i & 1+i \end{pmatrix} = -20-40i$$

$$\det(A(b,2)) = \det\begin{pmatrix} 2 & 0 & 0 \\ 2 & 10 & 2 \\ 3 & 0 & 1+i \end{pmatrix} = 10 \cdot \det\begin{pmatrix} 2 & 0 \\ 3 & 1+i \end{pmatrix} = 20+20i$$

$$\det(A(b,3)) = \det\begin{pmatrix} 2 & 3+i & 0 \\ 2 & 4 & 10 \\ 3 & 5i & 0 \end{pmatrix} = -10 \cdot \det\begin{pmatrix} 2 & 3+i \\ 3 & 5i \end{pmatrix} = 90-70i.$$

Aus der Cramer'schen Regel erhalten wir die Lösung des Gleichungssystems als

$$x_1 = \frac{-20-40i}{22-14i} = \frac{-10-20i}{11-7i} = \frac{(-10-20i)(11+7i)}{170} = \frac{3-29i}{17}$$

$$x_2 = \frac{20+20i}{22-14i} = \frac{10+10i}{11-7i} = \frac{(10+10i)(11+7i)}{170} = \frac{4+18i}{17}$$

$$x_3 = \frac{90-70i}{22-14i} = \frac{45-35i}{11-7i} = \frac{(45-35i)(11+7i)}{170} = \frac{74-7i}{17}.$$

4.3.31 **Aufgabe**. Sei **L**

$$A = \begin{pmatrix} \overline{3} & \overline{1} & \overline{1} \\ \overline{2} & \overline{0} & \overline{5} \\ \overline{1} & \overline{4} & \overline{3} \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{F}_7)$$

die Matrix aus Aufgabe 4.2.8 und sei $b = \begin{pmatrix} \overline{1} \\ \overline{0} \\ \overline{1} \end{pmatrix}$. Berechnen Sie die Lösung des Gleichungssystems $Ax = b$ mithilfe der Cramer'schen Regel.

4.4. Determinanten von Endomorphismen

Ein Endomorphismus eines Vektorraumes V ist eine lineare Abbildung $\varphi\colon V \to V$. In diesem Abschnitt werden wir eine Determinante für Endomorphismen definieren.

4.4.1 **Definition** Sei K ein Körper und sei V ein K-Vektorraum. Ein *Endomorphismus* von V ist eine K-lineare Abbildung $\varphi\colon V \to V$. Wir schreiben $\mathrm{End}(V)$ für die Menge aller Endomorphismen von V.

4.4.2 **Bemerkung**. Es gilt $\mathrm{End}(V) = \mathrm{Hom}_K(V, V)$. Also wissen wir aus 2.5.1, dass $\mathrm{End}(V)$ mit der punktweisen Addition und Skalarmultiplikation ein K-Vektorraum ist.

Auf der Menge der Endomorphismen gibt es eine weitere Rechenoperation: die Komposition $\circ$. Sind $\varphi, \psi \in \mathrm{End}(V)$, dann ist auch $\varphi \circ \psi$ ein Endomorphismus. Das kann man direkt nachrechnen. Seien dazu $v, v_1, v_2 \in V$ und $a \in K$, dann gilt

$$\begin{aligned}
\varphi \circ \psi(v_1 + v_2) &= \varphi(\psi(v_1 + v_2)) = \varphi(\psi(v_1) + \psi(v_2)) && (\psi \text{ linear})\\
&= \varphi(\psi(v_1)) + \varphi(\psi(v_2)) && (\varphi \text{ linear})\\
&= \varphi \circ \psi(v_1) + \varphi \circ \psi(v_2)
\end{aligned}$$

und genauso auch

$$\varphi \circ \psi(av) = \varphi(\psi(av)) = \varphi(a\psi(v)) = a\varphi(\psi(v)) = a(\varphi \circ \psi)(v).$$

4.4.3 **Aufgabe**. Sei V ein K-Vektorraum. Zeigen Sie, dass $(\mathrm{End}(V), +, \circ)$ ein Ring ist. L

4.4.4 **Beispiel**. Sei V ein Vektorraum and sei $a \in K$. Dann ist $f_a\colon V \to V$ mit $f_a(v) = av$ ein Endomorphismus. Falls $\dim_K(V) = 1$ ist, dann ist jeder Endomorphismus von dieser Form. In der Tat, sei $\varphi\colon V \to V$ ein Endomorphismus. Ist $u \neq 0$ in V, dann bildet u eine Basis von V. Weil $\varphi(u)$ in V liegt, gibt es ein eindeutiges Element $a \in K$ mit $\varphi(u) = au$. Da eine lineare Abbildung eindeutig durch die Bilder einer Basis festgelegt ist (siehe [MG, 8.4.1]), muss $\varphi = f_a$ gelten.

Sei V ein n-dimensionaler K-Vektorraum und sei φ ein Endomorphismus von V. Ist $\mathcal{B}$ eine Basis von V, dann erhalten wir eine Matrixdarstellung

$${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi) \in \mathrm{M}_{n,n}(K).$$

Im Vergleich zu allgemeinen linearen Abbildungen, wo man für den Definitionsbereich und den Wertebereich verschiedene Basen betrachten kann, nimmt man im Fall von Endomorphismen – wenn also Definitionsbereich und der Wertebereich übereinstimmen – nur eine Basis von V. Für die Matrixdarstellungen von Endomorphismen gilt folgende Transformationsformel:

4.4.5 **Transformationsformel für Endomorphismen.** *Sei V ein K-Vektorraum von endlicher Dimension und sei $\varphi\colon V \to V$ ein Endomorphismus. Sind $\mathcal{B}$ und $\mathcal{C}$ geordnete Basen von V, dann gilt*

$$ {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi) = ({}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\mathrm{id}_V))^{-1}\, {}_{\mathcal{C}}\mathrm{M}_{\mathcal{C}}(\varphi)\, {}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\mathrm{id}_V) $$

Beweis. Die Transformationsmatrix ${}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\mathrm{id}_V)$ ist invertierbar und es gilt

$$ ({}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\mathrm{id}_V))^{-1} = {}_{\mathcal{B}}\mathrm{M}_{\mathcal{C}}(\mathrm{id}_V); $$

siehe [MG, 9.3.2]. Damit folgt das Ergebnis aus der Transformationsformel 2.2.28. □

Aus der Transformationsformel erhalten wir eine wichtige Beobachtung: *Je zwei Matrixdarstellungen eines Endomorphismus sind ähnlich*; mit Ähnlichkeit im Sinne von Definition 4.2.24. Da die Determinanten ähnlicher Matrizen gleich sind (siehe Korollar 4.2.26) erhalten wir:

4.4.6 **Definition und Satz.** *Sei V ein endlich-dimensionaler Vektorraum und sei $\varphi\colon V \to V$ ein Endomorphismus. Wir definieren die* Determinante *von φ als*

$$ \det(\varphi) = \det({}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)) $$

für eine geordnete Basis $\mathcal{B}$ von V. Der Wert von $\det(\varphi)$ hängt nicht von der Wahl der Basis $\mathcal{B}$ ab.

4.4.7 **Determinante und Flächeninhalt.** Eine Vorstellung von der Bedeutung der Determinante kann man sich wieder mithilfe des Flächeninhaltes machen. Sei $\varphi\colon \mathbb{R}^2 \to \mathbb{R}^2$ ein Endomorphismus. Ist $X \subseteq \mathbb{R}^2$ eine Teilmenge mit Flächeninhalt F (z.B. ein Rechteck, ein Parallelogramm oder ein allgemeineres Polygon), dann ist der Flächeninhalt von $\varphi(X)$ genau $|\det(\varphi)|F$. Die Determinante gibt also an, wie stark sich der Flächeninhalt durch Anwenden von φ verändert.

Die viele Eigenschaften der Determinante von Matrizen kann man nun auch für die Determinante von Endomorphismen formulieren. Als Beispiel betrachten wir den Determinantenmultiplikationssatz.

4.4.8 **Determinantenmultiplikationssatz.** *Es sei V ein endlich-dimensionaler K-Vektorraum und es seien $\varphi, \psi \in \mathrm{End}(V)$. Es gilt*

$$ \det(\psi \circ \varphi) = \det(\varphi)\det(\psi). $$

4.4.9 **Aufgabe.** Geben Sie mit dem Determinantenmultiplikationssatz 4.2.21 für Matrizen einen Beweis von Satz 4.4.8. **L**

4.5. Alternierende Multilinearformen*

Auf den ersten Blick wirkt es sonderbar, dass die Determinante, die wir mit der komplizierten Leibniz-Formel (4.2.a) definiert haben, so viele nützliche Eigenschaften hat. Handelt es sich um einen Zufall, dass die Determinante ähnlicher Matrizen übereinstimmt und man über die Determinante von Endomorphismen sprechen kann? In diesem Abschnitt werden einen koordinatenfreien Zugang zur Determinante entwickeln. Das heißt, wir wollen $\det(\varphi)$ für einen Endomorphismus $\varphi \in \operatorname{End}(V)$ beschreiben, ohne dazu eine Basis des Vektorraumes V zu wählen. Dazu besprechen wir alternierende Multilinearformen. Dieser Abschnitt ist anspruchsvoll, weil wir viele neue Begriffe benötigen, die wir nur hier verwenden. Dieser Abschnitt ist *nicht klausurrelevant*.

I. Multilineare Abbildungen

Bilineare Abbildungen sind Abbildungen mit zwei Komponenten, die beide linear sind. Was ist mit Abbildungen, die drei, vier oder mehr lineare Komponenten haben? Die Theorie dieser Abbildungen wird oft *multilineare Algebra* genannt. In diesem Abschnitt geben wir einen sehr kurzen Einblick. Wir werden dieses Thema hier allerdings nur anschneiden um Begriffe einzuführen, die im weiteren Verlauf gelegentlich vorkommen.

Zu jeder natürlichen Zahl n und zu K-Vektorräumen U, V werden wir hier jeweils einen neuen Vektorraum definieren: den Raum $L_K^{(n)}(U, V)$ der n-linearen Abbildungen zwischen U und V. Das Schöne daran ist, dass es zwischen diesen vielen verschiedenen Vektorräumen diverse *natürliche* lineare Abbildungen gibt, die aus einfachen Operationen (z.B. Einsetzen eines Vektors) hergeleitet werden. Wir werden uns dazu zwei Beispiele genauer anschauen. Eigentlich ist das Ganze nicht sehr tiefgründig und dieser Abschnitt besteht im Wesentlichen aus einigen formalen Rechnungen. Beim ersten Lesen ist der abstrakte Blickwinkel aber wahrscheinlich schwer bekömmlich und es genügt, wenn Sie sich zunächst einen groben Eindruck verschaffen. Wir werden die multilinearen Abbildungen nur einmal benötigen, wenn wir in Lektion 4 die Determinante besprechen.

4.5.1 **Notation**. Sei $n \geq 1$ und sei U ein Vektorraum. Mit $U^{\times n}$ bezeichnen wir das n-fache kartesische Produkt

$$U^{\times n} := \underbrace{U \times \cdots \times U}_{n \text{ mal}}.$$

4.5.2 **Definition** Sei K ein Körper und seien U, V zwei K-Vektorräume. Sei $n \in \mathbb{N}$. Eine Abbildung

$$\beta\colon U^{\times n} \to V$$

nennt man *n-linear* über K, wenn die Abbildung linear in jeder der n Komponenten ist. Das heißt, die folgenden Bedingungen sind erfüllt:

(a) Für alle $u_1, \ldots, u_n \in U$, $j \in \{1, 2, \ldots, n\}$ und $u'_j \in U$ gilt

$$\begin{aligned}&\beta(u_1, \ldots, u_{j-1}, u_j + u'_j, u_{j+1}, \ldots, u_n)\\ &\quad = \beta(u_1, \ldots, u_j, \ldots, u_n) + \beta(u_1, \ldots, u_{j-1}, u'_j, u_{j+1}, \ldots, u_n).\end{aligned}$$

(b) Für alle $u_1, \ldots, u_n \in U$, $j \in \{1, 2, \ldots, n\}$ und $a \in K$ gilt

$$\beta(u_1, \ldots, u_{j-1}, au_j, u_{j+1}, \ldots, u_n) = a\beta(u_1, \ldots, u_j, \ldots, u_n).$$

Die Menge aller n-linearen Abbildungen $U^{\times n} \to V$ bezeichnen wir mit $L_K^{(n)}(U, V)$.

4.5.3 **Anmerkungen zur Definition**. (a) Die 1-linearen Abbildungen sind also genau die bereits bekannten linearen Abbildungen von U nach V. Es ist also $L_K^{(1)}(U, V) = \mathrm{Hom}_K(U, V)$.

(b) Eine n-lineare Abbildung $U^{\times n} \to K$ nennt man auch eine n-lineare Form. Eine 2-lineare Form ist also eine *Bilinearform*. Es gilt $\mathrm{Bil}_K(U) = L_K^{(2)}(U, K)$.

(c) Die konstante Nullabbildung $(u_1, \ldots, u_n) \mapsto 0$ von $U^{\times n}$ nach V ist eine n-lineare Abbildung.

(d) Manchmal ist es praktisch sich die Elemente von V als 0-lineare Abbildungen vorzustellen. Deshalb definieren wir $L_K^{(0)}(U, V) := V$

4.5.4 **Satz.** *Sei $n \in \mathbb{N}$. Sei K ein Körper und seien U, V zwei K-Vektorräume. Dann ist $L_K^{(n)}(U, V)$ ein Unterraum des Vektorraumes $\mathrm{Abb}(U^{\times n}, V)$ aller Abbildungen von $U^{\times n}$ nach V.*

Beweis. Wir erinnern uns, dass die Addition und Skalarmultiplikation auf $\mathrm{Abb}(U^{\times n}, V)$ punktweise definiert wurden; siehe 2.2.3. Weil die Nullabbildung in $L_K^{(n)}(U, V)$ liegt, müssen wir nur zeigen, dass $L_K^{(n)}(U, V)$ unter der Addition und Skalarmultiplikation abgeschlossen ist.

Es seien $\beta, \gamma \in L_K^{(n)}(U, V)$. Für alle $j \in \{1, 2, \ldots, n\}$, $u_1, \ldots, u_n, u'_j \in U$ gilt nun

$$(\beta + \gamma)(u_1, \ldots, u_{j-1}, u_j + u'_j, u_{j+1}, \ldots, u_n)$$

$$\begin{aligned}
&= \beta(u_1,\dots,u_{j-1},u_j+u_j',u_{j+1},\dots,u_n) + \gamma(u_1,\dots,u_{j-1},u_j+u_j',u_{j+1},\dots,u_n)\\
&= \beta(u_1,\dots,u_j,\dots,u_n) + \beta(u_1,\dots,u_j',\dots,u_n)\\
&\qquad + \gamma(u_1,\dots,u_j,\dots,u_n) + \gamma(u_1,\dots,u_j',\dots,u_n)\\
&= (\beta+\gamma)(u_1,\dots,u_j,\dots,u_n) + (\beta+\gamma)(u_1,\dots,u_j',\dots,u_n)
\end{aligned}$$

und für alle $a \in K$ gilt auch

$$\begin{aligned}
(\beta+\gamma)(u_1,\dots,au_j,\dots,u_n) &= \beta(u_1,\dots,au_j,\dots,u_n) + \gamma(u_1,\dots,au_j,\dots,u_n)\\
&= a\beta(u_1,\dots,u_n) + a\gamma(u_1,\dots,u_n)\\
&= a(\beta+\gamma)(u_1,\dots,u_n).
\end{aligned}$$

Also ist $(\beta+\gamma)$ wieder n-linear.

Bei der Skalarmultiplikation verfährt man genauso und zeigt, dass für $\beta \in L_K^{(n)}(U,V)$ und $c \in K$ auch $c\beta$ eine n-lineare Abbildung ist (Übungsaufgabe!). □ **L**

4.5.5 **Einsetzen**. Hat man eine n-lineare Abbildung mit $n \geq 1$, dann kann man durch Einsetzen eines fest gewählten Vektors in eine der Komponenten eine $(n-1)$-lineare Abbildung erzeugen.

Es sei $\beta \in L_K^{(n)}(U,V)$ eine n-lineare Abbildung mit $n \geq 1$. Für $k \in \{1,2,\dots,n\}$ und $w \in U$ definieren wir

$$\iota_w^k(\beta)\colon U^{\times n-1} \to V$$

durch die Vorschrift

$$\iota_w^k(\beta)(u_1,\dots,u_{n-1}) = \beta(u_1,\dots,u_{k-1},\underbrace{w}_{k},u_k,\dots,u_{n-1}).$$

Es folgt direkt aus der Definition, dass $\iota_w^k(\beta)$ eine $(n-1)$-lineare Abbildung ist.

4.5.6 **Lemma.** *Seien U,V zwei K-Vektorräume und sei $n \geq 1$. Für alle $w \in U$ und $k \in \{1,2,\dots,n\}$ ist*

$$\iota_w^k\colon L_K^{(n)}(U,V) \to L_K^{(n-1)}(U,V)$$

eine K-lineare Abbildung.

Beweis. Es seien $\beta,\gamma \in L_K^{(n)}(U,V)$. Dann gilt für alle $u_1,\dots,u_{n-1}$

$$\begin{aligned}
\iota_w^k(\beta+\gamma)(u_1,\dots,u_{n-1}) &= (\beta+\gamma)(u_1,\dots,u_{k-1},w,u_k,\dots,u_{n-1})\\
&= \beta(u_1,\dots,u_{k-1},w,u_k,\dots,u_{n-1}) + \gamma(u_1,\dots,u_{k-1},w,u_k,\dots,u_{n-1})\\
&= \iota_w^k(\beta)(u_1,\dots,u_{n-1}) + \iota_w^k(\gamma)(u_1,\dots,u_{n-1})
\end{aligned}$$

$$= (\iota_w^k(\beta) + \iota_w^k(\gamma))(u_1, \ldots, u_{n-1}).$$

Wir schließen, dass $\iota_w^k(\beta+\gamma) = \iota_w^k(\beta) + \iota_w^k(\gamma)$ ist.

Sei $c \in K$ und $\beta \in L_K^{(n)}(U,V)$. Es gilt

$$\begin{aligned}\iota_w^k(c\beta)(u_1, \ldots, u_{n-1}) &= (c\beta)(u_1, \ldots, u_{k-1}, w, u_k, \ldots, u_{n-1})\\ &= c\beta(u_1, \ldots, u_{k-1}, w, u_k, \ldots, u_{n-1})\\ &= c\iota_w^k(\beta)(u_1, \ldots, u_{n-1})\end{aligned}$$

und damit ist ι_w^k eine K-lineare Abbildung. □

4.5.7 **Lineare Abbildungen und multilineare Abbildungen**. Es seien U, U', V drei K-Vektorräume. Ist $\varphi\colon U \to U'$ eine K-lineare Abbildung, dann kann man diese verwenden, um aus einer n-linearen Abbildung von $(U')^{\times n}$ nach V eine n-lineare Abbildung von $U^{\times n}$ nach V zu konstruieren. Genauer: Für $\beta \in L_K^{(n)}(U',V)$ definieren wir $\varphi^*(\beta) \in L_K^{(n)}(U,V)$ durch

$$\varphi^*(\beta)(u_1, \ldots, u_n) \coloneqq \beta(\varphi(u_1), \ldots, \varphi(u_n))$$

für alle $u_1, \ldots, u_n \in U$. Da φ linear ist, ist $\varphi^*(\beta)$ wieder n-linear.

Im nächsten Satz stellen wir fest, dass φ^* wieder K-linear ist.

4.5.8 **Satz.** *Es seien U, U', V drei K-Vektorräume und sei $\varphi\colon U \to U'$ eine K-lineare Abbildung. Dann ist*

$$\varphi^*\colon L_K^{(n)}(U',V) \to L_K^{(n)}(U,V)$$

K-linear.

Ist $\psi\colon U' \to U''$ eine weitere K-lineare Abbildung, dann gilt

$$(\psi \circ \varphi)^* = \varphi^* \circ \psi^*.$$

<u>*Beweis*</u>. Wir zeigen, dass φ^* K-linear ist. Für alle $\beta, \gamma \in L_K^{(n)}(U',V)$ und alle $u_1, \ldots, u_n \in U$ gilt

$$\begin{aligned}\varphi^*(\beta+\gamma)(u_1, \ldots, u_n) &= (\beta+\gamma)(\varphi(u_1), \ldots, \varphi(u_n))\\ &= \beta(\varphi(u_1), \ldots, \varphi(u_n)) + \gamma((\varphi(u_1), \ldots, \varphi(u_n))\\ &= \varphi^*(\beta)(u_1, \ldots, u_n) + \varphi^*(\gamma)(u_1, \ldots, u_n)\\ &= (\varphi^*(\beta) + \varphi^*(\gamma))(u_1, \ldots, u_n).\end{aligned}$$

Da $u_1, \dots, u_n$ beliebig waren, schließen wir $\varphi^*(\beta + \gamma) = \varphi^*(\beta) + \varphi^*(\gamma)$.

Seien nun $c \in K$ und $\beta \in L_K^{(n)}(U', V)$. Dann gilt

$$\begin{aligned}\varphi^*(c\beta)(u_1, \dots, u_n) &= (c\beta)(\varphi(u_1), \dots, \varphi(u_n)) \\ &= c\beta(\varphi(u_1), \dots, \varphi(u_n)) \\ &= c\varphi^*(\beta)(u_1, \dots, u_n)\end{aligned}$$

für alle $u_1, \dots, u_n \in U$ und damit $\varphi^*(c\beta) = c\varphi^*(\beta)$.

Ist $\psi\colon U' \to U''$ eine weitere K-lineare Abbildung, dann erhalten wir für alle $\beta \in L_K^{(n)}(U'', V)$ und alle $u_1, \dots, u_n \in U$

$$\begin{aligned}(\psi \circ \varphi)^*(\beta)(u_1, \dots, u_n) &= \beta(\psi(\varphi(u_1)), \dots, \psi(\varphi(u_n))) \\ &= \psi^*(\beta)(\varphi(u_1), \dots, \varphi(u_n)) \\ &= \varphi^*(\psi^*(\beta))(u_1, \dots, u_n) \\ &= (\varphi^* \circ \psi^*)(\beta)(u_1, \dots, u_n);\end{aligned}$$

d.h., $(\psi \circ \varphi)^* = \varphi^* \circ \psi^*$. □

4.5.9 **Aufgabe**. Es seien U, U', V drei K-Vektorräume und sei $\varphi\colon U \to U'$ eine K-lineare Abbildung. Es seien $n \in \mathbb{N}$ und $k \in \{1, 2, \dots, n\}$ und $w \in U$ gegeben. Zeigen Sie die Gleichheit: **L**

$$\iota_w^k \circ \varphi^* = \varphi^* \circ \iota_{\varphi(w)}^k.$$

Das beendet unseren kurzen Ausflug in die multilineare Algebra. Wenn Ihnen dieser Abschnitt Kopfzerbrechen bereitet, lesen Sie am besten einfach weiter und kommen zurück, wenn Sie in Abschnitt 4.4 wieder auf die multilinearen Abbildungen treffen.

II. Alternierende Multilinearformen

4.5.10 **Definition** Sei V ein K-Vektorraum und sei $n \geq 1$. Eine n-lineare Form $\omega \in L_K^{(n)}(V, K)$ nennt man *alternierend*, wenn

$$\omega(v_1, \dots, v_n) = 0$$

gilt, sobald $v_i = v_j$ für irgendwelche $i \neq j$ gilt. In anderen Worten: Setzt man zwei gleiche Vektoren in ω ein, dann erhält man immer das Ergebnis 0.

Die Menge aller alternierenden n-linearen Formen bezeichnen wir mit $\bigwedge^n V^*$.

Man nennt $\bigwedge^n V^*$ auch das *n-fache äußere Produkt des Dualraums*. Wir wollen uns aber hier nicht die Mühe machen das äußere Produkt allgemein zu besprechen. Es ist auf den ersten Blick etwas verwunderlich, dass der Dualraum V^* in der Notation auftaucht. In den nächsten Beispielen werden wir sehen, wie man alternierende n-lineare Formen mithilfe des Dualraumes angeben kann.

4.5.11 **Beispiel**. Jede Linearform $\alpha \in V^*$ ist alternierend, denn man hat gar nicht die Möglichkeit zwei gleiche Vektoren einzusetzen. Das heißt, $V^* = \bigwedge^1 V^*$.

4.5.12 **Beispiel**. Es sei V ein K-Vektorraum. Für Elemente $\xi_1, \xi_2 \in V^*$ definieren wir eine alternierende Bilinearform $\xi_1 \wedge \xi_2$[3]. Zur Erinnerung: 2-lineare Formen sind genau die Bilinearformen. Für v_1, v_2 setzen wir

$$\xi_1 \wedge \xi_2(v_1, v_2) \coloneqq \xi_1(v_1)\xi_2(v_2) - \xi_1(v_2)\xi_2(v_1).$$

Als Übungsaufgabe kann man prüfen, dass $\xi_1 \wedge \xi_2$ wirklich bilinear ist. Warum ist diese Bilinearform alternierend? Ist $v_1 = v = v_2$, dann erhält man **L**

$$\xi_1 \wedge \xi_2(v, v) = \xi_1(v)\xi_2(v) - \xi_1(v)\xi_2(v) = 0.$$

4.5.13 **Beispiel**. Es sei V ein K-Vektorraum und sei $n \geq 1$. Für Elemente $\xi_1, \dots, \xi_n \in V^*$ definieren wir eine alternierende n-lineare Form $\xi_1 \wedge \xi_2 \wedge \cdots \wedge \xi_n$ durch

$$\xi_1 \wedge \cdots \wedge \xi_n(v_1, \dots, v_n) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{j=1}^{n} \xi_{\sigma(j)}(v_j).$$

Diese Formel erinnert stark an die Leibnizformel. Um zu zeigen, dass diese n-lineare Form alternierend ist, kann man den Beweis, dass die Determinante alternierend ist (siehe 4.2.17) leicht anpassen (Übungsaufgabe). **L**

4.5.14 **Satz.** *Seien V und W zwei K-Vektorräume und sei $n \geq 1$.*

(a) $\bigwedge^n V^*$ ist ein Unterraum von $L_K^{(n)}(V, K)$.

(b) Sei $\psi \colon V \to W$ eine lineare Abbildung. Ist $\omega \in \bigwedge^n W^*$, dann ist $\psi^*(\omega) \in \bigwedge^n V^*$.

Hier ist ψ^* die Abbildung aus 4.5.7.

[3] Gesprochen: ξ_1 Dach ξ_2

Beweis. Wir erinnern uns, dass die Addition und Skalarmultiplikation in $L_K^{(n)}(V,K)$ punktweise definiert wurde; vgl. 4.5.4. Da die Nullform alternierend ist, ist $\bigwedge^n V^*$ nicht leer. Seien $\omega_1, \omega_2 \in \bigwedge^n V^*$ und $a \in K$. Sind $v_1, \dots, v_n \in V$ und $v_i = v_j$ für irgendwelche $i \neq j$, dann gilt

$$(a\omega_1 + \omega_2)(v_1, \dots, v_n) = a \underbrace{\omega_1(v_1, \dots, v_n)}_{=0} + \underbrace{\omega_2(v_1, \dots, v_n)}_{=0} = 0.$$

Also ist $a\omega_1 + \omega_2$ alternierend und nach 2.2.6 ist $\bigwedge^n V^*$ ein Unterraum von $L_K^{(n)}(V,K)$.

Sei nun $\psi \colon V \to W$ eine K-lineare Abbildung und sei ω eine alternierende n-lineare Form auf W. Wir zeigen, dass auch $\psi^*(\omega)$ alternierend ist. Seien dazu $v_1, \dots, v_n \in V$ und sei $v_i = v_j$ für $i \neq j$. Es gilt $\psi(v_i) = \psi(v_j)$ und damit

$$\psi^*(\omega)(v_1, \dots, v_n) = \omega(\psi(v_1), \dots, \psi(v_n)) = 0.$$

Also ist $\psi^*(\omega)$ alternierend. □

4.5.15 **Beispiel**. Seien V, W zwei K-Vektorräume und sei $\psi \colon V \to W$ eine K-lineare Abbildung. Wir schauen uns nochmal die Formen aus Beispiel 4.5.13 an. Seien $\xi_1, \dots, \xi_n \in W^*$. Dann gilt

$$\psi^*(\xi_1 \wedge \cdots \wedge \xi_n) = \psi^*(\xi_1) \wedge \cdots \wedge \psi^*(\xi_n);$$

dabei bezeichnet ψ^* auf der linken Seite die Abbildung aus 4.5.7 und auf der rechten Seite die duale Abbildung $\psi^* \colon W^* \to V^*$ aus 2.5.16. Um das zu sehen, nehmen wir Vektoren $v_1, \dots, v_n \in V$ und setzen ein:

$$\begin{aligned}
\psi^*(\xi_1 \wedge \cdots \wedge \xi_n)(v_1, \dots, v_n) &= \xi_1 \wedge \cdots \wedge \xi_n(\psi(v_1), \dots, \psi(v_n)) \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{j=1}^n \xi_{\sigma(j)}(\psi(v_j)) \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{j=1}^n \psi^*(\xi_{\sigma(j)})(v_j) && \text{(Def. 2.5.16)} \\
&= \psi^*(\xi_1) \wedge \cdots \wedge \psi^*(\xi_n)(v_1, \dots, v_n).
\end{aligned}$$

Da dies für alle $v_1, \dots, v_n \in V$ gilt, folgt die Behauptung.

4.5.16 **Lemma.** *Sei V ein K-Vektorraum und sei $\omega \in \bigwedge^n V^*$. Für alle $v_1, \dots, v_n \in V$ und alle $\sigma \in S_n$ gilt*

$$\omega(v_{\sigma(1)}, \dots, v_{\sigma(n)}) = \operatorname{sgn}(\sigma)\omega(v_1, \dots, v_n).$$

Beweis. Wir zeigen die Aussage für den Fall, dass σ eine Transposition ist. Der allgemeine Fall folgt daraus, indem man σ als Produkt von m Transpositionen schreibt und verwendet, dass $\operatorname{sgn}(\sigma) = (-1)^m$ gilt; siehe Korollar 4.1.25.

Sei $\sigma = (k\ \ \ell)$ die Transposition, die k und ℓ (mit $k < \ell$) vertauscht. Wir setzen nun $v_1, \ldots, v_n$ in ω ein, wobei wir in die k-te und ℓ-te Komponente jeweils $v_k + v_\ell$ einsetzen. Da ω alternierend ist, erhalten wir

$$\begin{aligned} 0 &= \omega(v_1, \ldots, v_k + v_\ell, \ldots, v_k + v_\ell, \ldots, v_n) \\ &= \underbrace{\omega(v_1, \ldots, v_k, \ldots, v_k, \ldots, v_n)}_{=0} + \omega(v_1, \ldots, v_\ell, \ldots, v_k, \ldots, v_n) \\ &\quad + \omega(v_1, \ldots, v_k, \ldots, v_\ell, \ldots, v_n) + \underbrace{\omega(v_1, \ldots, v_\ell, \ldots, v_\ell, \ldots, v_n)}_{=0} \end{aligned}$$

und damit

$$\omega(v_1, \ldots, v_\ell, \ldots, v_k, \ldots, v_n) = -\omega(v_1, \ldots, v_k, \ldots, v_\ell, \ldots, v_n). \qquad \square$$

4.5.17 **Satz.** *Sei V ein K-Vektorraum der Dimension $\dim_K(V) = n$. Dann hat $\bigwedge^n V^*$ die Dimension 1.*

Sei $\mathcal{B} = (e_1, \ldots, e_n)$ eine geordnete Basis von V. Jede Form $\omega \in \bigwedge^n V^*$ ist eindeutig bestimmt durch den Wert $\omega(e_1, \ldots, e_n)$.

Beweis. Wir definieren eine K-lineare Abbildung $f\colon \bigwedge^n V^* \to K$ durch $f(\omega) = \omega(e_1, \ldots, e_n)$. Es ist nicht schwer zu sehen, dass f eine K-lineare Abbildung ist. Dazu kann man verwenden, dass f durch Einsetzen von Vektoren definiert wurde – es gilt $f = \iota^1_{e_1} \circ \iota^2_{e_2} \circ \cdots \circ \iota^n_{e_n}$ – und als Verknüpfung linearer Abbildungen ist f damit linear; siehe Lemma 4.5.6. Wir zeigen, dass f ein Isomorphismus ist. Da isomorphe Vektorräume dieselbe Dimension haben, ist damit der Satz bewiesen.

Wir zeigen zuerst, dass f injektiv ist. Dazu zeigen wir, dass $\operatorname{Ker}(f) = \{0\}$ gilt (der Nullvektor ist in diesem Fall die n-lineare Form, die konstant den Wert 0 ergibt). Sei $\omega \in \operatorname{Ker}(f)$, d.h. es gilt $\omega(e_1, \ldots, e_n) = 0$. Sind $v_1, \ldots, v_n \in V$ gegeben, dann schreiben wir $v_1, \ldots, v_n$ als Linearkombination der Basisvektoren

$$v_j = \sum_{i=1}^{n} a_{ij} e_i$$

mit eindeutig bestimmten Skalaren $a_{ij} \in K$. Durch Einsatz der Multilinearität (grob gesagt: ausmultiplizieren!) erhalten wir

$$\omega(v_1, \ldots, v_n) = \sum_{i_1, \ldots, i_n} \left(\prod_{j=1}^{n} a_{i_j j} \right) \omega(e_{i_1}, \ldots, e_{i_n}).$$

Aus Lemma 4.5.16 und $\omega(e_1, \dots, e_n) = 0$ schließen wir, dass alle Terme verschwinden. Es gilt also $\omega = 0$ und f ist injektiv.

Wir zeigen nun, dass f auch surjektiv ist. Sei $e_1^*, \dots, e_n^*$ die Dualbasis zu $e_1, \dots, e_n$. Wir erinnern uns daran, dass $e_i^*(e_j) = \delta_{ij}$ ist. Für alle $a \in K$ gilt dann

$$\begin{aligned} f(ae_1^* \wedge \dots \wedge e_n^*) &= ae_1^* \wedge \dots \wedge e_n^*(e_1, \dots, e_n) \\ &= a \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{j=1}^n e_{\sigma(j)}^*(e_j) = a, \end{aligned}$$

denn $\prod_{j=1}^n e_{\sigma(j)}^*(e_j)$ verschwindet, wenn σ nicht die identische Permutation ist. Damit ist f surjektiv und insgesamt ein Isomorphismus. Da f bijektiv ist, ist jede Form $\omega \in \bigwedge^n V^*$ eindeutig durch den Wert $f(\omega) = \omega(e_1, \dots, e_n)$ bestimmt. □

III. Determinante ohne Matrizen

Der unscheinbare Satz 4.5.17 ist der Schlüssel zum Verständnis der Determinante eines Endomorphismus $\varphi\colon V \to V$. Wir wissen aus Satz 4.5.14, dass ein Endomorphismus φ auch einen Endomorphismus $\varphi^*\colon \bigwedge^n V^* \to \bigwedge^n V^*$ auf dem Raum der alternierenden n-linearen Formen liefert. Ist nun $n = \dim_K(V)$, dann ist $\bigwedge^n V^*$ ein ein-dimensionaler Vektorraum. Was ist ein Endomorphismus eines ein-dimensionalen Raumes? Die Multiplikation mit einem Skalar $a \in K$ (siehe Beispiel 4.4.4)! Wir können also durch den Umweg über $\bigwedge^n V^*$ jedem Endomorphismus φ eindeutig ein $a \in K$ zuordnen. Welches Element könnte das sein? Die Determinante!

4.5.18 **Satz.** *Sei V ein K-Vektorraum der Dimension n und sei $\varphi \in \operatorname{End}(V)$. Für alle $\omega \in \bigwedge^n V^*$ gilt*

$$\varphi^*(\omega) = \det(\varphi)\omega. \tag{4.5.a}$$

Beweis. Sei $\mathcal{B} = (e_1, \dots, e_n)$ eine geordnete Basis von V und sei $\mathcal{B}^* = (e_1^*, \dots, e_n^*)$ die zugehörige Dualbasis. Wir wissen aus dem Beweis von Satz 4.5.17, dass das Element $\omega = e_1^* \wedge \dots \wedge e_n^*$ nicht Null ist und damit (aus Dimensionsgründen) $\bigwedge^n V^*$ aufspannt. Es genügt also Gleichung (4.5.a) für dieses eine Element ω zu beweisen.

Dazu nehmen wir die Matrixdarstellung ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi) = (a_{ij})_{i,j}$ von φ. Es gilt also

$$\varphi(e_j) = \sum_{i=1}^n a_{ij} e_i$$

und damit $e_k^*(\varphi(e_j)) = a_{kj}$. Damit erhalten wir

$$\begin{aligned}
\varphi^*(\omega)(e_1, \ldots, e_n) &= \omega(\varphi(e_1), \ldots, \varphi(e_n)) \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{j=1}^{n} e_{\sigma(j)}^*(\varphi(e_j)) \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{j=1}^{n} a_{\sigma(j),j} \\
&= \det({}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)) && \text{(Leibniz-Formel)} \\
&= \det(\varphi).
\end{aligned}$$

Aus dem Beweis von Satz 4.5.17 wissen wir, dass $\omega(e_1, \ldots, e_n) = 1$ gilt. Also ist $\varphi^*(\omega)(e_1, \ldots, e_n) = \det(\varphi)\omega(e_1, \ldots, e_n)$ und aus Satz 4.5.17 folgt nun die Behauptung. □

Ist $\omega \neq 0$, dann ist $\det(\varphi)$ durch die Gleichung $\varphi^*(\omega) = \det(\varphi)\omega$ eindeutig bestimmt. Wir könnten also mit Satz 4.5.18 eine koordinatenfreie Definition der Determinante angeben und die rechenlastigen Vorüberlegungen zu Matrizen (teilweise) vergessen. In der Tat kann man nun manche Ergebnisse sehr einfach ohne Matrixrechnung herleiten. Ein schönes Beispiel dafür ist der Determinantenmultiplikationssatz 4.4.8.

Koordinatenfreier Beweis von Satz 4.4.8. Sei $n = \dim_K(V)$. Sei $\omega \in \bigwedge^n V^*$ mit $\omega \neq 0$. Mit Satz 4.5.8 und Satz 4.5.18 erhalten wir

$$\begin{aligned}
\det(\psi \circ \varphi)\omega = (\psi \circ \varphi)^*(\omega) &= \varphi^*(\psi^*(\omega)) && (4.5.8) \\
&= \varphi^*(\det(\psi)\omega) && (4.5.18) \\
&= \det(\psi)\varphi^*(\omega) && (\varphi^* \text{ linear}) \\
&= \det(\varphi)\det(\psi)\omega. && (4.5.18)
\end{aligned}$$

Da $\omega \neq 0$ ist, folgt $\det(\psi \circ \varphi) = \det(\psi)\det(\varphi)$. □

Ich hoffe, dass Sie meine Begeisterung für diesen koordinatenfreien Blick auf die Determinante nachvollziehen können. Traut man sich an abstrakte Konzepte wie alternierende Multilinearformen heran, dann kann man anspruchsvolle Rechnungen mit der Leibniz-Formel durch kurze, elegante Umformungen ersetzen. Lineare Algebra ist eben mehr als nur Matrizenrechnung.

4.6. Lösungen der Aufgaben in Lektion 4

L4.1.6 Lösung. Es sei $\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 5 & 3 & 1 \end{pmatrix} \in S_5$. Um $\tau(3)$ zu bestimmen sehen wir in der dritten Spalte nach und finden $\tau(3) = 5$.

L4.1.8 Lösung.

(a) $$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 5 & 3 & 4 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 1 & 2 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 3 & 1 & 2 & 4 \end{pmatrix}$$

(b) $$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 3 & 5 & 4 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 5 & 3 & 1 & 2 \end{pmatrix}$$

L4.1.10 Lösung. Es sei $\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 5 & 3 & 1 \end{pmatrix} \in S_5$. Um τ^{-1} zu bestimmen wenden wir das Verfahren aus 4.1.9 an. Wir erhalten

$$\begin{aligned} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 5 & 3 & 1 \end{pmatrix} &\underset{\rightsquigarrow}{\text{Zeilen tauschen}} \begin{pmatrix} 4 & 2 & 5 & 3 & 1 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix} \\ &\underset{\rightsquigarrow}{\text{Spalten sortieren}} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 2 & 4 & 1 & 3 \end{pmatrix} \end{aligned}$$

und finden damit $\tau^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 2 & 4 & 1 & 3 \end{pmatrix}$.

L4.1.12 Lösung. Die Permutation $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 5 & 4 & 2 \end{pmatrix}$ hat genau die zwei Fixpunkte 1 und 4. Der Träger ist daher

$$\operatorname{Trg}(\sigma) = \{2, 3, 5\}.$$

L4.1.22 Lösung. Wir bestimmen zunächst die Fehlstände der Transposition $\sigma = (1\ \ 3) \in S_5$. Dazu überprüft man alle zehn Paare $(i, j) \in \{1, 2, 3, 4, 5\}^2$ mit $i < j$ und findet folgende Fehlstände

$$\operatorname{FS}(\sigma) = \{(1, 2), (1, 3), (2, 3)\}.$$

Daraus folgt $\operatorname{sgn}(\sigma) = (-1)^3 = -1$.

L4.1.28 **Lösung.** Die Gruppe S_3 hat 6 Elemente; siehe 4.1.3. Davon sind drei Elemente Transpositionen – nämlich $(1\ \ 2)$, $(1\ \ 3)$ und $(2\ \ 3)$ – und liegen wegen Satz 4.1.23 nicht in Alt_3. Um alle Elemente der Gruppe Alt_3 zu bestimmen, müssen wir also nur die Signaturen der übrigen drei Elemente

$$\mathrm{id}, \quad \sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \quad \text{und} \quad \tau = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$

bestimmen. Wir wissen, dass $\mathrm{id} \in \mathrm{Alt}_3$ liegt. Es gilt außerdem

$$\sigma = (1\ \ 2) \circ (2\ \ 3) \quad \text{und} \quad \tau = (2\ \ 3) \circ (1\ \ 2)$$

und aus Korollar 4.1.25 folgt $\operatorname{sgn}(\sigma) = 1 = \operatorname{sgn}(\tau)$. Damit ist $\mathrm{Alt}_3 = \{\mathrm{id}, \sigma, \tau\}$.

L4.1.29 **Lösung.** Sei $n \geq 2$ und sei $\tau \in S_n$ eine Transposition. Aus dem Satz von der Signatur 4.1.24 und Satz 4.1.23 folgt $\operatorname{sgn}(\sigma \circ \tau) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau) = -\operatorname{sgn}(\sigma)$. Also bildet die Zuordnung $\sigma \mapsto \sigma \circ \tau$ die Menge Alt_n nach B_n ab. Die Abbildung

$$f_\tau \colon \mathrm{Alt}_n \to B_n$$

ist durch $\sigma \mapsto \sigma \circ \tau$ also sinnvoll definiert. Da τ eine Transposition ist, gilt $\tau \circ \tau = \mathrm{id}$ und damit auch $\sigma \circ \tau \circ \tau = \sigma$. Also beschreibt dieselbe Zuordnungsvorschrift $\sigma \mapsto \sigma \circ \tau$ auch die Umkehrfunktion von B_n nach Alt_n und f_τ ist bijektiv.

L4.2.4 **Lösung.** Die Determinante von

$$A = \begin{pmatrix} 3 & -5 \\ 1 & -3 \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{R}).$$

ergibt sich aus der (2×2)-Merkregel als

$$\det(A) = 3 \cdot (-3) - (-5) \cdot 1 = -9 + 5 = -4.$$

L4.2.8 **Lösung.** Wir bestimmen die Determinante der Matrix

$$A = \begin{pmatrix} \overline{3} & \overline{1} & \overline{1} \\ \overline{2} & \overline{0} & \overline{5} \\ \overline{1} & \overline{4} & \overline{3} \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{F}_7)$$

mit der Regel von Sarrus. Wir erhalten

$$\begin{aligned} \det(A) &= \overline{0} + \overline{5} + \overline{8} - \overline{0} - \overline{6} - \overline{60} \\ &= -\overline{60} = -\overline{4} = \overline{3}. \end{aligned}$$

L4.2.15 Lösung. Es sei $A \in \mathrm{M}_{n,n}(\mathbb{C})$. Nach Definition gilt $A^* = \overline{A}^T$. Aus Lemma 4.2.14 wissen wir, dass $\det(A^*) = \det(\overline{A}^T) = \det(\overline{A})$ gilt. Wir behaupten, dass $\det(\overline{A}) = \overline{\det(A)}$ ist. Dazu verwenden wir die Leibniz-Formel und die Rechenregeln aus 1.5.19:

$$\begin{aligned}
\det(\overline{A}) &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{j=1}^{n} \overline{a_{\sigma(j),j}} \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \overline{\prod_{j=1}^{n} a_{\sigma(j),j}} && (1.5.19\ (b)) \\
&= \overline{\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{j=1}^{n} a_{\sigma(j),j}} && (1.5.19\ (a)) \\
&= \overline{\det(A)}.
\end{aligned}$$

L4.2.23 Lösung. Sei K ein Körper und sei $\mathrm{SL}_n(K)$ die Menge der invertierbaren $(n \times n)$-Matrizen mit Determinante 1. Um zu zeigen, dass $\mathrm{SL}_n(K)$ eine Untergruppe von $\mathrm{GL}_n(K)$ ist, prüfen wir die Untergruppenaxiome aus 1.1.18. Die Einheitsmatrix liegt in $\mathrm{SL}_n(K)$, also ist $\mathrm{SL}_n(K)$ nicht leer.

(UG2): Sind $A, B \in \mathrm{SL}_n(K)$, dann gilt nach dem Determinantenmultiplikationssatz

$$\det(AB) = \det(A)\det(B) = 1 \cdot 1 = 1$$

und somit auch $AB \in \mathrm{SL}_n(K)$.

(UG3): Ist $A \in \mathrm{SL}_n(K)$, dann folgt aus Korollar 4.2.22

$$\det(A^{-1}) = \det(A)^{-1} = 1.$$

Also ist $A^{-1} \in \mathrm{SL}_n(K)$.

L4.2.31 Lösung. Wir bestimmen die Determinante der Matrix

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 & 5 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 2 & 3 & 0 & 2 \\ 1 & 1 & 0 & 3 & 12 \\ -1 & -1 & 0 & 0 & -3 \end{pmatrix} \in \mathrm{M}_{n,n}(\mathbb{Q})$$

mit elementaren Zeilenumformungen. Zuerst ziehen wir die erste von der vierten Zeile ab. Dann addieren wir die erste Zeile zur letzten. Wir erhalten

$$A_3 = \begin{pmatrix} 1 & 1 & 0 & 0 & 5 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 2 & 3 & 0 & 2 \\ 0 & 0 & 0 & 3 & 7 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}.$$

Schließlich vertauschen wir die zweite und die dritte Zeile und erhalten die obere Dreiecksmatrix

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

Wir haben insgesamt eine Vertauschung eingesetzt, somit gilt

$$\det(A) = (-1)^1 \left(1 \cdot 2 \cdot (-1) \cdot 3 \cdot 2\right) = 12.$$

L4.3.6 **Lösung.** Dazu bestimmen wir die neun Matrizen A_{ij} mit $1 \leq i, j \leq 3$.

i/j	1	2	3
1	$A_{11} = \begin{pmatrix} \overline{2} & \overline{3} \\ \overline{0} & \overline{2} \end{pmatrix}$	$A_{12} = \begin{pmatrix} \overline{0} & \overline{3} \\ \overline{1} & \overline{2} \end{pmatrix}$	$A_{13} = \begin{pmatrix} \overline{0} & \overline{2} \\ \overline{1} & \overline{0} \end{pmatrix}$
2	$A_{21} = \begin{pmatrix} \overline{4} & \overline{3} \\ \overline{0} & \overline{2} \end{pmatrix}$	$A_{22} = \begin{pmatrix} \overline{1} & \overline{3} \\ \overline{1} & \overline{2} \end{pmatrix}$	$A_{23} = \begin{pmatrix} \overline{1} & \overline{4} \\ \overline{1} & \overline{0} \end{pmatrix}$
2	$A_{31} = \begin{pmatrix} \overline{4} & \overline{3} \\ \overline{2} & \overline{3} \end{pmatrix}$	$A_{32} = \begin{pmatrix} \overline{1} & \overline{3} \\ \overline{0} & \overline{3} \end{pmatrix}$	$A_{33} = \begin{pmatrix} \overline{1} & \overline{4} \\ \overline{0} & \overline{2} \end{pmatrix}$

Mit 4.2.3 finden wir damit die folgenden Minoren.

i/j	1	2	3
1	$\det(A_{11}) = \overline{4}$	$\det(A_{12}) = \overline{2}$	$\det(A_{13}) = \overline{3}$
2	$\det(A_{21}) = \overline{3}$	$\det(A_{22}) = \overline{4}$	$\det(A_{23}) = \overline{1}$
3	$\det(A_{31}) = \overline{1}$	$\det(A_{32}) = \overline{3}$	$\det(A_{33}) = \overline{2}$

L4.3.7 Lösung. Da Formeln diesen Sachverhalt unnötig kompliziert erscheinen lassen, begnügen wir uns mit einer anschaulichen Begründung: Das Transponieren verwandelt die Zeilen von A in Spalten und die Spalten in Zeilen. Löscht man also aus A^T die i-te Zeile und j-te Spalte, dann kann man genauso gut die i-te Spalte und j-te Zeile aus der ursprünglichen Matrix A löschen und danach Zeilen in Spalten und Spalten in Zeilen verwandeln. Das heißt, $(A^T)_{ij} = (A_{ji})^T$.

L4.3.10 Lösung. Wir berechnen die Determinante von

$$A = \begin{pmatrix} 2 & 3+i & 0 \\ 2 & 4 & 2 \\ 3 & 5i & 1+i \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{C})$$

durch Entwicklung an der dritten Spalte. Wir gehen die Einträge der dritten Spalte der Reihe nach durch. Beim i-ten Eintrag bestimmen wir die Untermatrix $A_{i,3}$ und den Minor $\det(A_{i,3})$. Das Ergebnis multiplizieren wir $(-1)^{i+3}a_{i,3}$ und addieren auf. Damit erhalten wir

$$\begin{aligned} \det(A) &= (-1)^4 \cdot 0 + (-1)^5 \cdot 2 \cdot \det \begin{pmatrix} 2 & 3+i \\ 3 & 5i \end{pmatrix} + (-1)^6 \cdot (1+i) \cdot \det \begin{pmatrix} 2 & 3+i \\ 2 & 4 \end{pmatrix} \\ &= -2 \cdot (-9+7i) + (1+i)(2-2i) = 18 - 14i + 4 = 22 - 14i. \end{aligned}$$

L4.3.12 Lösung. Wir entwickeln an der dritten Spalte, da in dieser Spalte nur der Eintrag $a_{2,3} = 4$ nicht Null ist. Wir erhalten somit

$$\det(A) = (-1) \cdot 4 \cdot \det \begin{pmatrix} 2 & -1 & 4 \\ -2 & 0 & -5 \\ 1 & 1 & -2 \end{pmatrix}.$$

Die Determinante der (3×3)-Matrix könnte man nun mit der Regel von Sarrus berechnen. Wir wenden hier den Entwicklungssatz nochmal an und entwickeln an der zweiten Spalte. Wir erhalten

$$\begin{aligned} \det(A) &= (-4) \cdot \left((-1) \cdot (-1) \cdot \det \begin{pmatrix} -2 & -5 \\ 1 & -2 \end{pmatrix} + (-1) \cdot \det \begin{pmatrix} 2 & 4 \\ -2 & -5 \end{pmatrix} \right) \\ &= (-4)\big(9 + 2\big) = -44. \end{aligned}$$

L4.3.13 Lösung. Da B und B' sich nur in der k-ten Spalte unterscheiden, ist $B_{ik} = B'_{ik}$. Ist $j < k$, dann ist die $(k-1)$-te Spalte von B'_{ij} genau das r-fache der $(k-1)$-ten Spalte von B_{ij}. Ist $j > k$, dann ist die k-te Spalte von B'_{ij} genau das r-fache der

k-ten Spalte von B_{ij}. Da die Determinante linear in jeder Spalte ist, gilt in diesen Fällen $\det(B'_{ij}) = r\det(B_{ij})$ und wir erhalten damit

$$\begin{aligned}\Delta(B') &= \sum_{j\neq k}(-1)^{i+j}b_{ij}\det(B'_{ij}) + (-1)^{i+k}rb_{ik}\det(B'_{ik})\\ &= \sum_{j\neq k}(-1)^{i+j}b_{ij}r\det(B_{ij}) + (-1)^{i+k}rb_{ik}\det(B_{ik}) = r\Delta(B).\end{aligned}$$

L4.3.21 Lösung. Die Minoren der Matrix A haben wir in der Lösung zu Aufgabe 4.3.6 bestimmt. Damit ergibt sich die Adjunkte als

$$\mathrm{Adj}(A) = \begin{pmatrix} \overline{4} & -\overline{3} & \overline{1} \\ -\overline{2} & \overline{4} & -\overline{3} \\ \overline{3} & -\overline{1} & \overline{2} \end{pmatrix} = \begin{pmatrix} \overline{4} & \overline{2} & \overline{1} \\ \overline{3} & \overline{4} & \overline{2} \\ \overline{3} & \overline{4} & \overline{2} \end{pmatrix}.$$

Wir bilden das Matrizenprodukt und erhalten

$$A\,\mathrm{Adj}(A) = \begin{pmatrix} \overline{1} & \overline{4} & \overline{3} \\ \overline{0} & \overline{2} & \overline{3} \\ \overline{1} & \overline{0} & \overline{2} \end{pmatrix}\begin{pmatrix} \overline{4} & \overline{2} & \overline{1} \\ \overline{3} & \overline{4} & \overline{2} \\ \overline{3} & \overline{4} & \overline{2} \end{pmatrix} = \begin{pmatrix} \overline{0} & \overline{0} & \overline{0} \\ \overline{0} & \overline{0} & \overline{0} \\ \overline{0} & \overline{0} & \overline{0} \end{pmatrix}.$$

L4.3.22 Lösung. Sei $A \in \mathrm{M}_{n,n}(R)$. Wir behaupten, dass $\mathrm{Adj}(A^T) = \mathrm{Adj}(A)^T$ gilt. Der Eintrag von $\mathrm{Adj}(A^T)$ an der Stelle (i,j) ist

$$\begin{aligned}(-1)^{i+j}\det((A^T)_{ji}) &= (-1)^{i+j}\det((A_{ij})^T) && \text{(Aufgabe 4.3.7)}\\ &= (-1)^{i+j}\det(A_{ij}) && \text{(Lemma 4.2.14)}\end{aligned}$$

und dies ist genau der Eintrag an der Stelle (j,i) von $\mathrm{Adj}(A)$; in anderen Worten, der Eintrag an der Stelle (i,j) von $\mathrm{Adj}(A)^T$. Da alle Einträge übereinstimmen schließen wir $\mathrm{Adj}(A^T) = \mathrm{Adj}(A)^T$.

L4.3.27 Lösung. Wir prüfen durch Ausprobieren, dass es kein Inverses zu $\overline{4}$ in $\mathbb{Z}/6\mathbb{Z}$ gibt.

$$\begin{aligned}\overline{1}\cdot\overline{4} &= \overline{4} \neq \overline{1}\\ \overline{2}\cdot\overline{4} &= \overline{2} \neq \overline{1}\\ \overline{3}\cdot\overline{4} &= \overline{0} \neq \overline{1}\\ \overline{4}\cdot\overline{4} &= \overline{4} \neq \overline{1}\\ \overline{5}\cdot\overline{4} &= \overline{2} \neq \overline{1}\end{aligned}$$

Damit ist $\overline{4}$ keine Einheit in $\mathbb{Z}/6\mathbb{Z}$.

L4.3.31 Lösung. Gegeben sind

$$A = \begin{pmatrix} \overline{3} & \overline{1} & \overline{1} \\ \overline{2} & \overline{0} & \overline{5} \\ \overline{1} & \overline{4} & \overline{3} \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{F}_7) \quad \text{und} \quad b = \begin{pmatrix} \overline{1} \\ \overline{0} \\ \overline{1} \end{pmatrix} \in \mathbb{F}_7^3.$$

Aus Aufgabe 4.2.8 wissen wir, dass $\det(A) = \overline{3}$ gilt. Da $\mathbb{F}_7$ ein Körper ist, ist $\overline{3}$ eine Einheit und A ist invertierbar. Wir können also die Cramer'sche Regel anwenden. Dabei bestimmen wir die Determinanten jeweils mit der Regel von Sarrus.

$$\det(A(b,1)) = \det \begin{pmatrix} \overline{1} & \overline{1} & \overline{1} \\ \overline{0} & \overline{0} & \overline{5} \\ \overline{1} & \overline{4} & \overline{3} \end{pmatrix} = \overline{5} - \overline{20} = \overline{6}$$

$$\det(A(b,2)) = \det \begin{pmatrix} \overline{3} & \overline{1} & \overline{1} \\ \overline{2} & \overline{0} & \overline{5} \\ \overline{1} & \overline{1} & \overline{3} \end{pmatrix} = \overline{5} + \overline{2} - \overline{6} - \overline{15} = \overline{0}$$

$$\det(A(b,3)) = \det \begin{pmatrix} \overline{3} & \overline{1} & \overline{1} \\ \overline{2} & \overline{0} & \overline{0} \\ \overline{1} & \overline{4} & \overline{1} \end{pmatrix} = \overline{8} - \overline{2} = \overline{6}$$

In $\mathbb{F}_7$ gilt $\overline{3}^{-1} = \overline{5}$ und es ist $\overline{6} \cdot \overline{5} = \overline{2}$. Wir erhalten damit die Lösung des linearen Gleichungssystems $Ax = b$ als

$$x = \begin{pmatrix} \overline{2} \\ \overline{0} \\ \overline{2} \end{pmatrix}.$$

L4.4.3 Lösung. Wir prüfen nach, dass $(\mathrm{End}(V), +, \circ)$ die Ringaxiome aus 1.2.1 erfüllt.

Ad (R1): Wir wissen bereits, dass $\mathrm{End}(V)$ mit der Addition $+$ und der punktweisen Skalarmultiplikation ein Vektorraum ist. Insbesondere ist $(\mathrm{End}(V), +)$ eine abelsche Gruppe.

Ad (R2): Dass die Komposition $\circ$ von Abbildungen assoziativ ist, ist aus den Mathematischen Grundlagen bekannt; siehe [MG, 1.4.23].

Ad (R3): Die identische Abbildung $\mathrm{id}_V \colon V \to V$ ist ein Endomorphismus. Da $\mathrm{id}_V \circ \varphi = \varphi$ und $\varphi \circ \mathrm{id}_V = \varphi$ gelten, ist id_V das Einselement.

Ad (R4): Wir prüfen die Distributivgesetze. Seien $\varphi, \psi, \psi' \in \mathrm{End}(V)$. Für alle $v \in V$ gilt

$$\varphi \circ (\psi + \psi')(v) = \varphi(\psi(v) + \psi'(v)) = \varphi(\psi(v)) + \varphi(\psi'(v)) = (\varphi \circ \psi + \varphi \circ \psi')(v).$$

Da v beliebig war, folgt daraus das erste Distributivgesetz

$$\varphi \circ (\psi + \psi') = \varphi \circ \psi + \varphi \circ \psi'.$$

Das zweite Distributiv zeigt man analog (interessanterweise wird dabei die Linearität der Abbildungen gar nicht benötigt).

L4.4.9 **Lösung.** Wir wählen eine geordnete Basis $\mathcal{B}$ von V. Dann gilt mit [MG, 9.3.1]

$${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\psi \circ \varphi) = {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\psi){}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi).$$

Mit dem Determinantenmultiplikationssatz 4.2.21 für Matrizen schließen wir

$$\det(\psi \circ \varphi) = \det {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\psi \circ \varphi) = \det\left({}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\psi)\right)\det\left({}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)\right) = \det(\psi)\det(\varphi).$$

L4.5.4 **Lösung.** Sei $c \in K$ und $\beta \in L_K^{(n)}(U,V)$. Wir rechnen nach, dass $c\beta$ wieder n-linear ist, indem wir direkt die Definition 4.5.2 verwenden. Wir erhalten

$$\begin{aligned}
&(c\beta)(u_1,\dots,u_{j-1},u_j+u_j',u_{j+1},\dots,u_n)\\
&\quad= c\,\beta(u_1,\dots,u_{j-1},u_j+u_j',u_{j+1},\dots,u_n)\\
&\quad= c\,\beta(u_1,\dots,u_j,\dots,u_n) + c\,\beta(u_1,\dots,u_j',\dots,u_n)\\
&\quad= (c\beta)(u_1,\dots,u_j,\dots,u_n) + (c\beta)(u_1,\dots,u_j',\dots,u_n)
\end{aligned}$$

und

$$\begin{aligned}
(c\beta)(u_1,\dots,au_j,\dots,u_n) &= c\,\beta(u_1,\dots,au_j,\dots,u_n)\\
&= ca\beta(u_1,\dots,u_n)\\
&= a(c\beta)(u_1,\dots,u_n).
\end{aligned}$$

für alle $j \in \{1,2,\dots,n\}$, $u_1,u_2,\dots,u_n \in U$, $u_j' \in U$ und $a \in K$.

L4.5.9 **Lösung.** Es sei $\beta \in L_K^n(U',V)$ und es seien $u_1,\dots,u_{n-1} \in U$ beliebig. Durch Anwenden der Definitionen 4.5.5 und 4.5.7 erhalten wir

$$\begin{aligned}
\iota_w^k(\varphi^*(\beta))(u_1,\dots,u_{n-1}) &= \varphi^*(\beta)(u_1,\dots,u_{k-1},w,u_k,\dots,u_{n-1})\\
&= \beta(\varphi(u_1),\dots,\varphi(u_{k-1}),\varphi(w),\varphi(u_k),\dots,\varphi(u_{n-1}))\\
&= \iota_{\varphi(w)}^k(\beta)(\varphi(u_1),\dots,\varphi(u_{n-1}))\\
&= \varphi^*(\iota_w^k(\beta))(u_1,\dots,u_{n-1}).
\end{aligned}$$

Da diese Gleichung für alle β und alle $u_1,\dots,u_{n-1}$ gilt, folgt daraus die gesuchte Gleichung

$$\iota_w^k \circ \varphi^* = \varphi^* \circ \iota_{\varphi(w)}^k.$$

L4.5.12 Lösung. Für v_1, v_2 ist

$$\xi_1 \wedge \xi_2(v_1, v_2) \coloneqq \xi_1(v_1)\xi_2(v_2) - \xi_1(v_2)\xi_2(v_1).$$

Sei $v_1' \in V$. Dann gilt

$$\begin{aligned}\xi_1 \wedge \xi_2(v_1 + v_1', v_2) &= \xi_1(v_1 + v_1')\xi_2(v_2) - \xi_1(v_2)\xi_2(v_1 + v_1') \\ &= \xi_1(v_1)\xi_2(v_2) + \xi_1(v_1')\xi_2(v_2) - \xi_1(v_2)\xi_2(v_1) - \xi_1(v_2)\xi_2(v_1') \\ &= \xi_1(v_1)\xi_2(v_2) - \xi_1(v_2)\xi_2(v_1) + \xi_1(v_1')\xi_2(v_2) - \xi_1(v_2)\xi_2(v_1') \\ &= \xi_1 \wedge \xi_2(v_1, v_2) + \xi_1 \wedge \xi_2(v_1', v_2).\end{aligned}$$

Für alle $a \in K$ gilt auch

$$\begin{aligned}\xi_1 \wedge \xi_2(av_1, v_2) &= \xi_1(av_1)\xi_2(v_2) - \xi_1(v_2)\xi_2(av_1) \\ &= a\xi_1(v_1)\xi_2(v_2) - a\xi_1(v_2)\xi_2(v_1) = a\xi_1 \wedge \xi_2(v_1, v_2).\end{aligned}$$

Auf dieselbe Weise zeigt man die Linearität in der zweiten Komponente. Also ist $\xi_1 \wedge \xi_2$ eine Bilinearform.

L4.5.13 Lösung. Mit einer ähnlichen Rechnung wie in 4.5.12 prüft man, dass $\xi_1 \wedge \cdots \wedge \xi_n$ multilinear ist. Nehmen wir nun an, dass $v_1, \ldots, v_n \in V$ zwei gleiche Vektoren enthält; sagen wir $v_k = v_\ell$. Sei $\tau = (k\ \ell) \in S_n$ die Transposition, die k und ℓ vertauscht. Dann gelten

$$\xi_{\sigma(k)}(v_k) = \xi_{\sigma(k)}(v_\ell) = \xi_{\sigma(\tau(\ell))}(v_\ell) \quad \text{und} \quad \xi_{\sigma(\ell)}(v_\ell) = \xi_{\sigma(\ell)}(v_k) = \xi_{\sigma(\tau(k))}(v_k)$$

für alle $\sigma \in S_n$. Da für alle $j \notin \{k, \ell\}$ auch $\xi_{\sigma(j)}(v_j) = \xi_{\sigma(\tau(j))}(v_j)$ gilt, erhalten wir

$$\prod_{j=1}^{n} \xi_{\sigma(j)}(v_j) = \prod_{j=1}^{n} \xi_{\sigma(\tau(j))}(v_j).$$

Da $\sigma \mapsto \sigma \circ \tau$ eine Bijektion zwischen Alt_n und der Menge B_n aller Permutationen mit Signatur -1 definiert (siehe Aufgabe 4.1.29), heben sich jeweils der Summand zu $\sigma \in \mathrm{Alt}_n$ und der Summand zu $\sigma \circ \tau \in B_n$ weg und wir erhalten

$$\xi_1 \wedge \cdots \wedge \xi_n(v_1, \ldots, v_n) = 0.$$

Steffen Kionke

Lineare Algebra

Lektion 5:
Eigenwerte und Eigenvektoren

Fakultät für
Mathematik und Informatik

FernUniversität in Hagen

Studierhinweise zur fünften Lektion

In dieser und der nächsten Lektion wenden wir uns dem wichtigsten Thema dieses Moduls zu. Grob gesagt möchten wir untersuchen, welche Endomorphismen von endlich-dimensionalen Vektorräumen es gibt. Was soll das heißen? Haben zwei Endomorphismen dieselben Matrixdarstellungen zu verschiedenen Basen, dann handelt es sich im Wesentlichen – bis auf einen Wechsel des Koordinatensystems – um „dieselbe“ Abbildung. Das heißt, diese Abbildungen haben dieselben Eigenschaften. Gesucht ist also eine Klassifikation der Endomorphismen bis auf Basiswechsel.

Dem zugrunde liegt das sogenannte *Normalformproblem*. Hat man einen Endomorphismus gegeben, dann gibt es viele verschiedene Matrixdarstellungen. Man möchte daher eine möglichst „einfache“ Matrixdarstellung wählen, aus der sich die wesentlichen Eigenschaften des Endomorphismus ablesen lassen. Eine solche Matrixdarstellung werden wir mit der *Jordan'schen Normalform* in der Lektion 6 herleiten. In diesem Kapitel lernen wir dafür die wichtigsten Werkzeuge kennen.

Der zentrale Begriff dieser Lektion ist der des *Eigenvektors*. Dabei handelt es sich um Vektoren, die durch den Endomorphismus mit einem Faktor λ (dem *Eigenwert*) gestreckt (bzw. gestaucht) werden. Daraus entwickeln wir die Eigenwerttheorie, mit deren Hilfe wir unter anderem klären können, welche Endomorphismen eine diagonale Matrixdarstellung besitzen.

Um die Eigenwerte eines Endomorphismus bestimmen zu können, definieren wir das *charakteristische Polynom*. Es stellt sich heraus, dass die Eigenwerte genau die Nullstellen dieses Polynoms sind. Das charakteristische Polynom ist die Determinante einer Matrix und man kann es mit den Methoden aus Lektion 4 gut berechnen. Wir lernen außerdem das Minimalpolynom von Endomorphismen kennen. Auch dieses Polynom hat die Eigenschaft, dass die Nullstellen genau die Eigenwerte sind, aber es ist im Allgemeinen schwierig zu berechnen. Hier werden wir zeigen, dass das Minimalpolynom ein Teiler des charakteristischen Polynoms ist. Die Bedeutung des Minimalpolynoms werden wir aber erst in der nächsten Lektion verstehen.

Die Eigenwerttheorie hat auch zahlreiche Anwendungen außerhalb der Mathematik: unter anderem in der Physik, der Biologie und der Informatik. Um einen kleinen Ausblick zu geben, besprechen wir den Einsatz der Eigenwerttheorie in Googles PageRank-Verfahren. Dabei lernen wir auch *stochastische Matrizen* und den *Satz von Perron* kennen, die in vielen Anwendungen von Bedeutung sind.

Diese Lektion ist etwas einfacher, enthält aber viele zentrale Begriffe. Bitte schauen Sie sich diese gut an, sonst werden Sie in der nächsten Lektion große Probleme bekommen.

Lernziele

Nach dem Durcharbeiten der Lektion sollten Sie

→ mit dem *Normalformproblem* vertraut sein,

→ wissen, was ein *invarianter Unterraum* ist,

→ die Begriffe *Eigenvektor* und *Eigenwert* kennen,

→ wissen, wie man *Eigenwerte* und *Eigenräume* bestimmen kann,

→ verstehen, welche Endomorphismen *diagonalisierbar* sind und eine diagonale Matrixdarstellung berechnen können,

→ das *charakteristische Polynom* einer Matrix berechnen können,

→ die Definition des *Minimalpolynoms* verstehen,

→ den *Satz von Cayley-Hamilton* kennen und anwenden können,

→ wissen, was eine *stochastische Matrix* ist und mit dem *Satz von Perron* vertraut sein.

Literaturhinweise

Die Eigenwerttheorie wird in allen Lehrbüchern zur Linearen Algebra sehr gut besprochen. Ergebnisse zu stochastischen Matrizen findet man allerdings nur in [Bär] und [Gö].

→ [Bär]: 6.2 – 6.5 und A.4.

→ [Beu]: Kapitel 8.

→ [Bo]: 6.1, 6.2.

→ [Fi]: 5.1 – 5.4 und 5.6.

→ [Gö]: 6.1 und 8.3.

→ [KaSt]: 8.1 – 8.4 und 8.9.

Fahrplan durch die Lektion

Im einführenden Abschnitt 5.1 erklären wir das *Normalformproblem* (5.1.2, 5.1.4). Unser Weg zur Lösung basiert auf *invarianten Unterräumen*, die wir im zweiten Teil definieren (5.1.5) und dazu zwei Beispiele betrachten (5.1.8, 5.1.11). Wir erklären, warum eine Zerlegung in invariante Unterräume bei der Lösung des Normalformproblems hilfreich ist (5.1.13). Dieser Abschnitt dient hauptsächlich dazu, den roten Faden aufzuzeigen. ← 5.1

Abschnitt 5.2 enthält viele zentrale Begriffe, die Sie unbedingt verstehen müssen. Nehmen Sie sich für diesen Abschnitt etwas mehr Zeit. ← 5.2

Im ersten Teil definieren wir *Eigenvektoren und Eigenwerte* (5.2.1) von Endomorphismen und Matrizen. Zu jedem Eigenwert λ definieren wir den *Eigenraum* (5.2.8) und die *geometrische Vielfachheit* (5.2.10). Das unscheinbare Lemma 5.2.13 sollte man nicht übersehen (die Summe der Eigenräume ist immer direkt); wir zeigen damit, dass eine $(n \times n)$-Matrix höchstens n verschiedene Eigenwerte besitzt.

Im zweiten Teil untersuchen wir welche Endomorphismen *diagonalisierbar* sind, d.h., welche Endomorphismen eine diagonale Matrixdarstellung zulassen (5.2.16). Die Antwort gibt der *Satz zur Diagonalisierbarkeit* (5.2.19). Daraus erhalten wir auch ein Verfahren um Matrizen zu diagonalisieren (5.2.23).

Im Abschnitt 5.3 lernen wir, wie man einem Endomorphismus zwei Polynome zuordnen kann: das charakteristische Polynom und das Minimalpolynom. Diese Polynome enthalten sehr wertvolle Informationen über den Endomorphismus. ← 5.3

Im ersten Abschnitt definieren wir das charakteristische Polynom χ_A einer Matrix (5.3.2). Wir beweisen einen Satz über die Koeffizienten des charakteristischen Polynoms (5.3.5) und zeigen dann, dass die Nullstellen von χ_A genau die Eigenwerte von A sind (5.3.7). Wir definieren damit die *algebrische Vielfachheit* eines Eigenwertes (5.3.10). Wir zeigen, dass ähnliche Matrizen dasselbe charakteristische Polynom haben; das macht es möglich das charakteristische Polynom eines Endomorphismus zu definieren (5.3.17).

Im zweiten Teil befassen wir uns mit dem Minimalpolynom. Wir besprechen kurz, dass $\mathrm{M}_{n,n}(K)$ und $\mathrm{End}(V)$ *Algebren* sind (5.3.21) und erklären, wie man Elemente einer Algebra in ein Polynom einsetzen kann (5.3.22, 5.3.27). Damit sind wir in der Lage zu erläutern, was es heißt, dass ein Polynom g einen Endomorphismus φ *annulliert* – kurz: $g(\varphi) = 0$ – und wir definieren das Verschwindungsideal (5.3.29).

Den eindeutigen normierten Erzeuger μ_φ des Verschwindungsideals nennt man *Minimalpolynom* von φ (5.3.32). Wir zeigen, dass die Nullstellen des Minimalpolynoms von φ genau die Eigenwerte sind (5.3.38).

Im dritten Teil beweisen wir den wichtigen Satz von Cayley-Hamilton, der die Verbindung zwischen beiden Polynomen herstellt: χ_φ teilt μ_φ. Der Beweis ist anspruchsvoll und verwendet den Adjunktensatz.

5.4 $\rightarrow$ Im Abschnitt 5.4 werfen wir einen kurzen Blick auf Anwendungen der Eigenwerttheorie. Dazu definieren wir stochastische Matrizen (5.4.1) und beweisen den nützlichen Satz von Perron: Jede positive stochastische Matrix hat den Eigenwert 1 mit geometrischer Vielfachheit 1 (5.4.6). Wir skizzieren kurz wie diese Matrizen als Übergangsmatrizen von Markov-Ketten auftreten (5.4.9) und geben ein Beispiel (5.4.10). Im zweiten Teil erklären wir, wieso das PageRank-Verfahren der Suchmaschine Google wesentlich auf dem Satz von Perron beruht.

5.1. Das Normalformproblem

I. Was ist das Normalformproblem?

In dieser und der nächsten Lektion werden wir Endomorphismen auf Vektorräumen untersuchen. Ist V ein endlich-dimensionaler Vektorraum und $\varphi \in \mathrm{End}(V)$ ein Endomorphismus, dann möchten wir verstehen, was φ mit den Elementen von V genau macht. Dazu können wir eine geordnete Basis $\mathcal{B}$ von V wählen und die zugehörige Matrixdarstellung ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)$ betrachten. Das Ergebnis hängt aber von der gewählten Basis ab und es kann passieren, dass die Matrixdarstellung zu einer Basis sehr „einfach" ist, während die Matrixdarstellung zu einer anderen Basis schrecklich kompliziert aussieht.

5.1.1 **Beispiel**. Wir betrachten den reellen Vektorraum $\mathbb{R}^2$ mit dem Endomorphismus $\varphi \in \mathrm{End}(\mathbb{R}^2)$:

$$\varphi(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}) = \begin{pmatrix} 7x_1 - 30x_2 \\ x_1 - 4x_2 \end{pmatrix}.$$

Die Matrixdarstellung von φ bzgl. der Standardbasis $\mathcal{B} = (e_1, e_2)$ von $\mathbb{R}^2$ ist

$${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi) = \begin{pmatrix} 7 & -30 \\ 1 & -4 \end{pmatrix}.$$

Anhand dieser Beschreibung ist es schwierig sich den Endomorphismus φ vorzustellen.

Jetzt nehmen wir eine andere geordnete Basis $\mathcal{C} = (u_1, u_2)$, die aus den Vektoren

$$u_1 = \begin{pmatrix} 6 \\ 1 \end{pmatrix} \quad \text{und} \quad u_2 = \begin{pmatrix} 5 \\ 1 \end{pmatrix}$$

besteht. Wir bestimmen die Matrixdarstellung von φ bezüglich $\mathcal{C}$. Es gilt

$$\varphi(u_1) = \begin{pmatrix} 42 - 30 \\ 6 - 4 \end{pmatrix} = \begin{pmatrix} 12 \\ 2 \end{pmatrix} = 2u_1$$

und

$$\varphi(u_2) = \begin{pmatrix} 35 - 30 \\ 5 - 4 \end{pmatrix} = \begin{pmatrix} 5 \\ 1 \end{pmatrix} = u_2.$$

Die Matrixdarstellung ist also

$${}_{\mathcal{C}}\mathrm{M}_{\mathcal{C}}(\varphi) = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}.$$

Hier versteht man nun ganz gut, was φ macht: Der erste Basisvektor wird mit dem Faktor 2 gestreckt, der zweite bleibt unverändert.

5.1.2 **Das Normalformproblem für Endomorphismen**. Gegeben sei ein Endomorphismus φ eines endlich-dimensionalen K-Vektorraumes V. Wir werden nun das grundlegende Problem der Lektionen 5 und 6 formulieren.

> Man gebe eine geordnete Basis $\mathcal{B}$ von V an dergestalt, dass die Matrixdarstellung ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)$ besonders „einfach" ist.

Natürlich wollen wir dieses Problem nicht nur für einen speziellen Endomorphismus lösen. Vielmehr interessiert uns, ob es ein Verfahren gibt mit dem man eine solche Basis zu einem jeden gegebenen Endomorphismus bestimmen kann.

Darüber, welche Matrizen besonders „einfach" sind, kann man natürlich streiten. Die Lösung, die wir kennenlernen werden, sind Matrixdarstellungen in der sogenannten *Jordan'schen Normalform*. Aus diesen Matrizen kann man alle wesentlichen Eigenschaften des Endomorphismus φ direkt ablesen.

Das Normalformproblem kann man auch vollständig in der Sprache der Matrizen formulieren. Dazu halten wir zunächst die folgende Beobachtung fest.

5.1.3 **Lemma.** *Sei V ein Vektorraum mit $\dim_K(V) = n$ und sei $\mathcal{B}$ eine geordnete Basis von V. Sei $\varphi\colon V \to V$ ein Endomorphismus. Eine Matrix $A \in \mathrm{M}_{n,n}(K)$ ist genau dann eine Matrixdarstellung von φ, wenn A und ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)$ ähnlich sind.*

Beweis. Wir wissen aus der Transformationsformel 4.4.5, dass je zwei Matrixdarstellungen eines Endomorphismus ähnlich sind. Sei nun umgekehrt A ähnlich zu ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)$. Es gibt also eine invertierbare Matrix $S = (s_{ij})_{i,j} \in \mathrm{M}_{n,n}(K)$ mit $A = S^{-1}{}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)S$.

Es sei $\mathcal{B} = (v_1, \dots, v_n)$. Dann definieren wir für jedes $j \in \{1, 2, \dots, n\}$

$$w_j = \sum_{i=1}^{n} s_{ij} v_i \in V.$$

Da S invertierbar ist, ist $\mathcal{C} = (w_1, \dots, w_n)$ eine geordnete Basis von V. Das kann man beispielsweise mit folgender Überlegung sehen: Mithilfe der Einträge von S^{-1} kann man jeden Vektor aus $\mathcal{B}$ als Linearkombination der Vektoren aus $\mathcal{C}$ schreiben. Damit ist $\mathcal{C}$ ein Erzeugendensystem und da $\mathcal{C}$ genau $n = \dim_K(V)$ Elemente besitzt, handelt es sich um eine Basis; siehe [MG, 7.4.4].

Es gilt $S = {}_{\mathcal{B}}\mathrm{M}_{\mathcal{C}}(\mathrm{id})$ nach Definition der Matrixdarstellung (siehe 2.2.23). Aus der Transformationsformel 4.4.5 erhalten wir dann

$$A = S^{-1}{}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)S = {}_{\mathcal{C}}\mathrm{M}_{\mathcal{C}}(\varphi),$$

d.h., A ist eine Matrixdarstellung von φ. $\square$

Ist φ ein Endomorphismus von V, dann können wir eine beliebige Basis $\mathcal{B}$ von V wählen. Die Menge aller Basisdarstellungen von φ ist nach Lemma 5.1.3 genau die Ähnlichkeitsklasse von ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)$. Wir können das Normalformproblem also auch vollständig für Matrizen formulieren.

5.1.4 **Das Normalformproblem für Matrizen**. Sei K ein Körper. Man gebe ein Vertretersystem der Ähnlichkeitsklassen in $\mathrm{M}_{n,n}(K)$ an, das aus möglichst „einfachen“ Matrizen besteht.

Man gebe außerdem ein Verfahren an, mit dem man zu jeder Matrix $A \in \mathrm{M}_{n,n}(K)$ den Vertreter der Ähnlichkeitsklasse von A berechnen kann.

Wir werden uns der Lösung des Normalformproblems in dieser und der nächsten Lektion Schritt für Schritt nähern und dabei einige neue und wichtige Konzepte kennenlernen. Dabei verfolgen wir einen induktiven Ansatz und versuchen das Problem auf kleinere Teilprobleme zurückzuführen. Dazu ist das Normalformproblem von Endomorphismen deutlich besser geeignet, als das Normalformproblem für quadratische Matrizen. Unser wichtigstes Hilfsmittel sind die invarianten Unterräume, die wir im nächsten Abschnitt definieren.

II. Invariante Unterräume

Beginnen wir mit einer Überlegung. Nehmen wir an, wir verstehen Endomorphismen auf Vektorräumen kleiner Dimension bereits gut. Es seien nun V ein K-Vektorraum einer größeren Dimension und ein Endomorphismus $\varphi\colon V \to V$ gegeben. Wie können wir das Wissen über Endomorphismen auf Räumen niedriger Dimension einsetzen? Es ist naheliegend Unterräume von V zu untersuchen. Allerding muss man dabei vorsichtig sein. Ist U ein Unterraum von V, dann gilt im Allgemeinen $\varphi(U) \not\subseteq U$. In anderen Worten: φ bewegt den Unterraum U weg und wir haben auf U keinen Endomorphismus, der aus φ hervorgeht. Relevant sind also nur Unterräume, die nicht „wegbewegt“ werden. Damit gelangen wir zu folgender Definition.

5.1.5 $\boxed{\textbf{Definition}}$ Sei V ein K-Vektorraum und sei $\varphi \in \mathrm{End}(V)$. Ein Unterraum $U \subseteq V$ heißt *φ-invariant*, wenn

$$\varphi(u) \in U$$

für alle $u \in U$ gilt.

5.1.6 **Beobachtung**. Sei $\varphi \in \mathrm{End}(V)$. Ist $U \subseteq V$ ein φ-invarianter Unterraum, dann ist die Einschränkung $\varphi|_U\colon U \to U$ von φ auf U ein Endomorphismus von U.

5.1.7 **Beispiel**. Es sei $\varphi \in \mathrm{End}(V)$ ein Endomorphismus. Der ganze Vektorraum V und der Nullraum $\{0\} \subseteq V$ sind immer φ-invariant.

5.1.8 **Beispiel**. Wir betrachten den reellen Vektorraum $\mathbb{R}^3$ mit dem Endomorphismus φ, definiert durch

$$\varphi\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right) = \begin{pmatrix} 2x_1 - x_2 - x_3 \\ -4x_1 + x_2 \\ 5x_3 \end{pmatrix}.$$

Der Unterraum $U = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 \mid x_3 = 0 \right\}$ ist φ-invariant. Das sieht man so: Sei $\begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} \in U$. Dann gilt

$$\varphi\left(\begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 2x_1 - x_2 \\ -4x_1 + x_2 \\ 0 \end{pmatrix} \in U.$$

Um zu prüfen, ob ein Unterraum φ-invariant ist, ist es ausreichend das Verhalten des Endomorphismus auf einer Basis des Unterraumes zu untersuchen.

5.1.9 **Lemma.** *Sei V ein K-Vektorraum und sei $\varphi \in \operatorname{End}(V)$. Es sei $U \subseteq V$ ein Unterraum und es sei $\mathcal{B}$ eine Basis von U.*

Der Unterraum U ist φ-invariant genau dann, wenn $\varphi(w) \in U$ für alle $w \in \mathcal{B}$ gilt.

Beweis. Ist U invariant unter φ, dann gilt nach Definition $\varphi(u) \in U$ für alle $u \in U$ und insbesondere für alle u in der Basis $\mathcal{B}$.

Nehmen wir nun an, dass $\varphi(w) \in U$ für alle $w \in \mathcal{B}$ gilt. Sei $u \in U$ beliebig. Da $\mathcal{B}$ eine Basis ist, können wir u als Linearkombination

$$u = \sum_{i=1}^{k} a_i w_i$$

mit Basisvektoren $w_i \in \mathcal{B}$ und Skalaren $a_i \in K$ schreiben. Da φ linear ist, erhalten wir

$$\varphi(u) = \sum_{i=1}^{k} a_i \underbrace{\varphi(w_i)}_{\in U}.$$

Da U ein Unterraum ist, liegt die rechte Seite wieder in U. Das heißt, U ist φ-invariant. □

5.1.10 **Aufgabe**. Es sei $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{R})$. Wir betrachten den Endomorphismus $f_A \colon \mathbb{R}^2 \to \mathbb{R}^2$ mit $x \mapsto Ax$. Beschreiben Sie alle f_A-invarianten Unterräume im $\mathbb{R}^2$. **L**

5.1.11 **Beispiel**. Sei K ein Körper. Es sei $A = \operatorname{diag}(A_1, \ldots, A_s) \in \mathrm{M}_{n,n}(K)$ eine Blockdiagonalmatrix vom Blocktyp $p_1, \ldots, p_s$ (siehe 1.2.35). Wir schreiben $(a_{ij})_{i,j}$ für die Einträge von A. Für $t \in \{0, 1, \ldots, s\}$ definieren wir $P_t = \sum_{r=1}^{t} p_r$.

Wir betrachten den Endomorphismus $f_A \colon K^n \to K^n$ mit $x \mapsto Ax$. Es bezeichne $(e_1, e_2, \ldots, e_n)$ die Standardbasis von K^n. Für jedes $t \in \{1, \ldots, s\}$ ist

$$U_t = \langle \{ e_j \mid P_{t-1} < j \leq P_t \} \rangle$$

ein f_A-invarianter Unterraum. Wegen Lemma 5.1.9 genügt es, die gegebene Basis $\mathcal{B}_t = (e_{P_{t-1}+1}, \ldots, e_{P_t})$ von U_t zu betrachten. Für $P_{t-1} < j \leq P_t$ gilt

$$\begin{aligned} f_A(e_j) = Ae_j &= \sum_{i=1}^{n} a_{i,j} e_i \\ &= \sum_{i=P_{t-1}+1}^{P_t} a_{i,j} e_i \in U_t \qquad\qquad (A \text{ vom Blocktyp } p_1, \ldots, p_s). \end{aligned}$$

Damit ist U_t ein f_A-invarianter Unterraum und die Matrixdarstellung von $f_A|_{U_t}$ zur geordneten Basis $\mathcal{B}_t$ ist genau A_t.

Es gilt also $f_A(U_t) \subseteq U_t$ für alle t. In dieser Situation können wir insbesondere das Ergebnis aus Aufgabe 2.3.13 anwenden. Hier gilt also $\operatorname{Ker}(f_A) = \bigoplus_{t=1}^{s} \operatorname{Ker}(f_A|_{U_t})$. Da A_t die Matrixdarstellung von $f_A|_{U_t}$ zur geordneten Basis $\mathcal{B}_t$ ist, erhalten wir insbesondere

$$\dim_K \operatorname{Ker}(A) = \sum_{t=1}^{s} \dim_K \operatorname{Ker}(A_t).$$

Wir werden in dieser Lektion noch einige Beispiele invarianter Unterräume kennenlernen.

5.1.12 **Aufgabe**. Sei V ein K-Vektorraum und sei $\varphi \in \operatorname{End}(V)$. Es seien φ-invariante Unterräume $U_1, \ldots, U_k$ von V gegeben. **L**

(a) Zeigen Sie, dass $\bigcap_{i=1}^{k} U_i$ wieder ein φ-invarianter Unterraum ist.

(b) Zeigen Sie, dass $\sum_{i=1}^{k} U_i$ wieder ein φ-invarianter Unterraum ist.

Zum Abschluss können wir nun erklären, was es heißt, einen Vektorraum in invariante Unterräume zu *zerlegen*. Dazu sei V ein Vektorraum mit einem Endomorphismus φ. Wenn sich $V = U_1 \oplus U_2$ als direkte Summe zweier φ-invarianter Unterräume schreiben lässt, sprechen wir von einer *Zerlegung*.

Wieso ist das hilfreich, wenn man eine einfache Matrixdarstellung finden will? Der folgende Satz gibt die Antwort. Zerlegt sich $V = U_1 \oplus U_2$ in φ-invariante Unterräume, dann gibt es eine Matrixdarstellung des Endomorphismus φ in Blockdiagonalgestalt.

5.1.13 **Satz.** *Sei V ein endlich-dimensionaler K-Vektorraum und sei $\varphi \in \mathrm{End}(V)$. Wir nehmen an, dass $V = U_1 \oplus U_2$ direkte Summe zweier φ-invarianter Unterräume ist. Ist $\mathcal{B}_1 = (v_1, \ldots, v_k)$ eine geordnete Basis von U_1 und $\mathcal{B}_2 = (v_{k+1}, \ldots, v_n)$ eine geordnete Basis von U_2, dann ist $\mathcal{B} = (v_1, v_2, \ldots, v_n)$ eine geordnete Basis von V und ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)$ ist eine Blockdiagonalmatrix der Form*

$$
{}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi) = \left(\begin{array}{c|c} {}_{\mathcal{B}_1}\mathrm{M}_{\mathcal{B}_1}(\varphi|_{U_1}) & 0 \\ \hline 0 & {}_{\mathcal{B}_2}\mathrm{M}_{\mathcal{B}_2}(\varphi|_{U_2}) \end{array}\right).
$$

Beweis. Wir wissen aus Satz 2.3.12, dass $\mathcal{B}$ eine geordnete Basis von V ist. Sei $j \leq k$. Da U_1 ein φ-invarianter Unterraum ist, gilt $\varphi(v_j) \in U_1$ und damit

$$
\varphi(v_j) = \sum_{i=1}^{k} a_{ij} v_i;
$$

d.h., die Vektoren $v_{k+1}, \ldots, v_n$ werden nicht benötigt um $\varphi(v_j)$ in der Basis $\mathcal{B}$ auszudrücken. Die entsprechenden Matrixeinträge a_{ij} mit $i > k$ verschwinden also. Wegen $\varphi|_{U_1}(v_j) = \varphi(v_j)$, sind die Einträge a_{ij} mit $i, j \leq k$ genau die Matrixeinträge von ${}_{\mathcal{B}_1}\mathrm{M}_{\mathcal{B}_1}(\varphi|_{U_1})$.

Analog verfährt man mit U_2. Da U_2 ebenfalls φ-invariant ist, gilt für $j > k$

$$
\varphi(v_j) = \sum_{i=k+1}^{n} a_{ij} v_i;
$$

wir benötigen also die Vektoren aus $\mathcal{B}_1$ nicht, um $\varphi(v_j)$ als Linearkombination auszudrücken, d.h., $a_{ij} = 0$ für $i \leq k$. Nach Definition sind die Skalare $c_{ij} := a_{i+k,j+k}$ genau die Matrixeinträge von ${}_{\mathcal{B}_2}\mathrm{M}_{\mathcal{B}_2}(\varphi|_{U_2})$, denn v_j ist der $(j-k)$-te Vektor in der geordneten Basis $\mathcal{B}_2$. □

Damit beenden wir diesen Abschnitt. Weitere Ergebnisse zu invarianten Unterräumen besprechen wir in der nächsten Lektion.

5.2. Eigenwerte, Eigenvektoren und Diagonalisierbarkeit

In diesem Abschnitt legen wir den Grundstein für die Lösung des Normalformproblems. Wir studieren hier die ein-dimensionalen invarianten Unterräume von Endomorphismen. Diese kann man sehr gut verstehen und mit diesem Wissen werden wir für bestimmte Endomorphismen bereits eine besonders einfache Matrixdarstellung angeben können.

I. Eigenwerte und Eigenvektoren

Es sei V ein K-Vektorraum und es sei $\varphi \in \operatorname{End}(V)$ ein Endomorphismus von V. Ist $U \subseteq V$ ein Unterraum der Dimension 1, dann wird U von einem Vektor $v \neq 0$ aufgespannt, d.h., $U = \{av \mid a \in K\}$. Nehmen wir nun an, dass U ein φ-invarianter Unterraum ist, dann gilt $\varphi(v) \in U$. Es gibt also ein $\lambda \in K$ dergestalt, dass $\varphi(v) = \lambda v$ ist. Damit gelangen wir zur folgenden wichtigen Definition.

5.2.1 **Definition** Sei V ein K-Vektorraum und sei $\varphi \in \operatorname{End}(V)$ ein Endomorphismus von V. Ein Vektor $v \in V$ heißt *Eigenvektor* von φ, wenn $v \neq 0$ ist und

$$\varphi(v) = \lambda v$$

für ein $\lambda \in K$ gilt. Der Skalar λ heißt dann *Eigenwert* zum Eigenvektor v.

Wir nennen $\lambda \in K$ einen Eigenwert von φ, wenn es einen Eigenvektor mit Eigenwert λ gibt.

5.2.2 **Anmerkungen zur Definition**. (a) Ein Vektor $v \in V$ spannt genau dann einen ein-dimensionalen φ-invarianten Unterraum auf, wenn v ein Eigenvektor von φ ist.

(b) Sei $A \in \mathrm{M}_{n,n}(K)$ eine quadratische Matrix. Die Eigenvektoren und Eigenwerte der linearen Abbildung $f_A\colon K^n \to K^n$ mit $f_A(x) = Ax$ (vgl. 2.2.19) nennen wir Eigenwerte und Eigenvektoren von A. Das heißt, ein Vektor $v \neq 0$ in K^n ist ein Eigenvektor von A, wenn $Av = \lambda v$ für ein $\lambda \in K$ gilt.

5.2.3 **Eigenvektoren sind nicht** 0! Der Nullvektor ist per Definition kein Eigenvektor (obwohl $\varphi(0) = \lambda 0$ für alle λ gilt). Der Eigenwert $\lambda = 0$ ist aber erlaubt.

5.2.4 **Beispiel**. Der Vektor $v = \begin{pmatrix} 2 \\ -3 \end{pmatrix} \in \mathbb{R}^2$ ist ein Eigenvektor von

$$A = \begin{pmatrix} 2 & 2 \\ 0 & -1 \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{R}),$$

denn es ist

$$Av = \begin{pmatrix} 2 & 2 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 2 \\ -3 \end{pmatrix} = \begin{pmatrix} -2 \\ 3 \end{pmatrix} = -v.$$

Der zugehörige Eigenwert ist $\lambda = -1$.

5.2.5 **Aufgabe**. Gegeben ist die Matrix **L**

$$A = \begin{pmatrix} \frac{3}{4} & -\frac{3}{4} \\ -\frac{1}{4} & \frac{1}{4} \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{Q}).$$

Welche der folgenden Vektoren sind Eigenvektoren von A? Bestimmen Sie ggf. den zugehörigen Eigenwert.

$$v_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 3 \\ -1 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad v_4 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

5.2.6 **Aufgabe**. Sei φ ein Endomorphismus von V. Wir nehmen an, dass $v \in V$ ein Eigenvektor von φ zum Eigenwert $\lambda \in K$ ist. Zeigen Sie, dass für jedes $j \in \mathbb{N}$ der Vektor v ein Eigenvektor von **L**

$$\varphi^j := \underbrace{\varphi \circ \cdots \circ \varphi}_{j \text{ mal}}$$

zum Eigenwert λ^j ist.

5.2.7 **Lemma.** *Sei V ein endlich-dimensionaler K-Vektorraum und sei $\varphi \in \mathrm{End}(V)$. Sei $\mathcal{B} = (v_1, \dots, v_n)$ eine geordnete Basis von V. Ist $A = {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)$ die Matrixdarstellung von φ, so haben φ und A dieselben Eigenwerte.*

Beweis. Um diese Aussage einzusehen, verwenden wir die Koordinatenabbildung $\kappa_{\mathcal{B}} \colon V \to K^n$ zur Basis $\mathcal{B}$ (siehe 2.2.22). Ist $\lambda \in K$ ein Eigenwert von φ und ist $v \in V$ ein zugehöriger Eigenvektor, dann gilt mit (2.2.b) aus 2.2.23

$$A\kappa_{\mathcal{B}}(v) = \kappa_{\mathcal{B}}(\varphi(v)) = \kappa_{\mathcal{B}}(\lambda v) = \lambda \kappa_{\mathcal{B}}(v);$$

da $\kappa_{\mathcal{B}}$ ein Isomorphismus ist, ist $\kappa_{\mathcal{B}}(v) \neq 0$ und damit auch ein Eigenvektor von A zum Eigenwert λ.

Sei umgekehrt λ ein Eigenwert von A und sei $x \in K^n$ ein zugehöriger Eigenvektor. Wir argumentieren wie zuvor, aber mithilfe der Umkehrabbildung $\kappa_{\mathcal{B}}^{-1} \colon K^n \to V$ – die nach [MG, Prop. 8.1.8] ebenfalls ein Isomorphismus ist. Wir definieren $v = \kappa_{\mathcal{B}}^{-1}(x)$. Mit (2.2.b) aus 2.2.23 erhalten wir

$$\varphi(v) = \kappa_{\mathcal{B}}^{-1}(A\kappa_{\mathcal{B}}(v)) = \kappa_{\mathcal{B}}^{-1}(Ax) = \kappa_{\mathcal{B}}^{-1}(\lambda x) = \lambda v;$$

wir schließen, dass v ein Eigenvektor von φ zum Eigenwert λ ist. $\square$

5.2.8 $\boxed{\textbf{Definition}}$ Sei V ein K-Vektorraum und sei φ ein Endomorphismus. Für $\lambda \in K$ definieren wir

$$\mathrm{Eig}_\varphi(\lambda) = \{v \in V \mid \varphi(v) = \lambda v\}.$$

5.2.9 **Anmerkungen zur Definition.** (a) $\mathrm{Eig}_\varphi(\lambda)$ ist ein φ-invarianter Unterraum von V. Dass es sich um einen Unterraum handelt, kann man mit dem Kriterium aus Aufgabe 2.2.6 zeigen (Übungsaufgabe!). Man kann das auch auf anderem Wege **L**
einsehen. Es gilt

$$\mathrm{Eig}_\varphi(\lambda) = \{v \in V \mid \varphi(v) - \lambda v = 0\} = \mathrm{Ker}(\varphi - \lambda\,\mathrm{id}_V)$$

und der Kern einer linearen Abbildung ist bekanntlich ein Unterraum; siehe [MG, Bemerkung 8.3.8].

Um zu zeigen, dass $\mathrm{Eig}_\varphi(\lambda)$ auch φ-invariant ist, nehmen wir einen Vektor $v \in \mathrm{Eig}_\varphi(\lambda)$. Dann gilt

$$\varphi(v) = \lambda v \in \mathrm{Eig}_\varphi(\lambda),$$

denn $\mathrm{Eig}_\varphi(\lambda)$ ist ein Unterraum von V.

(b) Ein Vektor $v \in \mathrm{Eig}_\varphi(\lambda)$ mit $v \neq 0$ erfüllt $\varphi(v) = \lambda v$ und ist damit ein Eigenvektor von φ zum Eigenwert λ. In anderen Worten: $\mathrm{Eig}_\varphi(\lambda) \neq \{0\}$ genau dann, wenn λ ein Eigenwert von φ ist. In diesem Fall nennen wir $\mathrm{Eig}_\varphi(\lambda)$ den *Eigenraum* von φ zum Eigenwert λ.

(c) Sei $a \in K$. Ist $a + \lambda$ ein Eigenwert von φ, dann ist λ ein Eigenwert von $\varphi - a\,\mathrm{id}$. Es gilt

$$\mathrm{Eig}_\varphi(a+\lambda) = \{v \in V \mid \varphi(v) = av+\lambda v\} = \{v \in V \mid \varphi(v)-av = \lambda v\} = \mathrm{Eig}_{\varphi - a\,\mathrm{id}}(\lambda).$$

(d) Ist $A \in \mathrm{M}_{n,n}(K)$ eine quadratische Matrix, dann schreiben wir $\mathrm{Eig}_A(\lambda)$ anstelle von $\mathrm{Eig}_{f_A}(\lambda)$, d.h.,

$$\mathrm{Eig}_A(\lambda) = \{v \in K^n \mid Av = \lambda v\}.$$

Ist λ ein Eigenwert von A, dann nennen wir $\mathrm{Eig}_A(\lambda)$ den Eigenraum von A zum Eigenwert λ.

5.2.10 $\boxed{\textbf{Definition}}$ Ist λ ein Eigenwert von $\varphi \in \mathrm{End}(V)$, dann nennt man

$$\gamma(\lambda) = \dim_K \mathrm{Eig}_\varphi(\lambda)$$

die *geometrische Vielfachheit* von λ. Analog definieren wir die geometrische Vielfachheit des Eigenwertes λ einer Matrix A durch $\gamma(\lambda) = \dim_K \mathrm{Eig}_A(\lambda)$.

5.2.11 **Beispiel**. Wir betrachten die Matrix

$$A = \begin{pmatrix} 5 & 1 & 1 \\ 1 & 5 & 1 \\ 1 & 1 & 5 \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{R}).$$

Es sei $\lambda = 4$. Der Raum $\mathrm{Eig}_A(4)$ ist der Kern der Matrix $A - 4I_3$. Diesen Kern können wir mit dem Gaußverfahren berechnen:

$$\mathrm{Eig}_A(4) = \mathrm{Ker}(A - 4I_3) = \mathrm{Ker}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

Wir ziehen jeweils die erste Zeile von der zweiten und dritten Zeile ab und erhalten

$$\mathrm{Eig}_A(4) = \mathrm{Ker}\begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \langle \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \rangle.$$

Damit ist 4 ein Eigenwert von A und der Eigenraum $\mathrm{Eig}_A(4)$ hat die Dimension 2, d.h., $\gamma(4) = 2$ ist die geometrische Vielfachheit des Eigenwertes 4.

5.2.12 **Aufgabe**. Bestimmen Sie die Eigenwerte und die zugehörigen geometrischen Vielfachheiten der Matrix **L**

$$A = \begin{pmatrix} \overline{1} & \overline{2} \\ \overline{2} & \overline{1} \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{F}_3).$$

5.2.13 **Lemma.** *Sei V ein endlich-dimensionaler K Vektorraum und sei $\varphi \in \mathrm{End}(V)$. Sind $\lambda_1, \ldots, \lambda_k$ paarweise verschiedene Eigenwerte von φ, dann ist die Summe $\sum_{i=1}^k \mathrm{Eig}_\varphi(\lambda_i)$ direkt. Insbesondere gilt $\sum_{i=1}^k \gamma(\lambda_i) \leq \dim_K V$.*

Beweis. Wir beweisen die Aussage per Induktion über die Anzahl k der Eigenwerte. Für $k = 1$ ist nichts zu zeigen. Sei nun $k > 1$. Wir nehmen an, dass die Aussage für $k - 1$ Eigenwerte bekannt ist. Um zu zeigen, dass die Summe direkt ist verwenden wir Satz 2.3.6 (iii).

Für $1 \leq i \leq k$ sei $u_i \in \operatorname{Eig}_\varphi(\lambda_i)$ gegeben. Wir nehmen an, dass $\sum_{i=1}^k u_i = 0$ ist und müssen zeigen, dass $u_i = 0$ für alle i gilt. Wir beginnen mit einer Umformung und stellen fest

$$\begin{aligned} 0 = \lambda_k \cdot 0 - \varphi(0) &= \lambda_k \sum_{i=1}^k u_i - \sum_{i=1}^k \varphi(u_i) \\ &= \sum_{i=1}^k (\lambda_k - \lambda_i) u_i && (\text{denn } \varphi(u_i) = \lambda_i u_i) \\ &= \sum_{i=1}^{k-1} (\lambda_k - \lambda_i) u_i. \end{aligned}$$

Da die Eigenräume Unterräume von V sind, gilt $(\lambda_k - \lambda_i) u_i \in \operatorname{Eig}_\varphi(\lambda_i)$ für alle $i \leq k-1$. Aus der Induktionsvoraussetzung wissen wir, dass die Summe von $(k-1)$ verschiedenen Eigenräumen direkt ist. Mit 2.3.6 folgt $(\lambda_k - \lambda_i) u_i = 0$ für alle $i \leq k-1$. Die Eigenwerte sind paarweise verschieden, also gilt $\lambda_k - \lambda_i \neq 0$ und durch Kürzen folgt $u_i = 0$ für alle $i \leq k-1$. Aus $\sum_{i=1}^k u_i = 0$ folgt dann aber auch $u_k = 0$. Die Summe der Eigenräume ist also direkt.

Wir wenden uns nun der Ungleichung zu. Es sei $W = \bigoplus_{i=1}^k \operatorname{Eig}_\varphi(\lambda_i)$. Da W ein Unterraum von V ist, gilt $\dim_K(W) \leq \dim_K(V)$; siehe [MG, 7.4.5]. Insbesondere sind auch alle Eigenräume endlich-dimensional. Aus Satz 2.3.12 schließen wir

$$\dim_K(V) \geq \dim_K(W) = \sum_{i=1}^k \dim_K \operatorname{Eig}_\varphi(\lambda_i) = \sum_{i=1}^k \gamma(\lambda_i). \qquad \square$$

5.2.14 **Korollar.** *Sei V ein endlich-dimensionaler K-Vektorraum der Dimension n. Jeder Endomorphismus von V hat höchstens n verschiedene Eigenwerte.*

Beweis. Übungsaufgabe! $\square$ **L**

Insbesondere hat auch jede $(n \times n)$-Matrix höchstens n verschiedene Eigenwerte.

5.2.15 **Aufgabe**. Sei $0 < k < n$. Geben Sie eine Matrix $A \in \mathrm{M}_{n,n}(\mathbb{R})$ an, die genau k verschiedene Eigenwerte besitzt. **L**

II. Diagonalisierbarkeit

In diesem Abschnitt werden wir das Wissen über Eigenvektoren und Eigenwerte anwenden und uns mit besonders einfachen Endomorphismen auseinandersetzen. Zunächst wollen wir diesen Endomorphismen einen Namen geben.

5.2.16 **Definition** Sei V ein endlich-dimensionaler K-Vektorraum. Ein Endomorphismus $\varphi \in \operatorname{End}(V)$ heißt *diagonalisierbar*, wenn es eine geordnete Basis $\mathcal{B}$ von V gibt, sodass die Matrixdarstellung ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)$ eine Diagonalmatrix ist.

Ist φ diagonalisierbar, dann gibt es also eine geordnete Basis $\mathcal{B} = (v_1, \ldots, v_n)$ und Skalare $\lambda_1, \ldots, \lambda_n \in K$ mit

$$
{}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}.
$$

In anderen Worten, es gilt $\varphi(v_i) = \lambda_i v_i$, d.h., der i-te Basisvektor ist ein Eigenvektor von φ zum Eigenwert λ_i. Der Endomorphismus aus Beispiel 5.1.1 ist diagonalisierbar.

Man kann den Begriff der Diagonalisierbarkeit auch auf Matrizen übertragen.

5.2.17 **Definition** Eine quadratische Matrix $A \in \mathrm{M}_{n,n}(K)$ heißt *diagonalisierbar*, wenn A ähnlich zu einer Diagonalmatrix ist.

Im folgenden Lemma formulieren wir den Zusammenhang zwischen diesen beiden Konzepten aus.

5.2.18 **Lemma.** *Sei V ein endlich-dimensionaler K-Vektorraum und sei $\varphi \in \operatorname{End}(V)$. Sei $\mathcal{C}$ eine geordnete Basis von V. Der Endomorphismus φ ist genau dann diagonalisierbar, wenn die Matrixdarstellung ${}_{\mathcal{C}}\mathrm{M}_{\mathcal{C}}(\varphi)$ diagonalisierbar ist.*

Beweis. Ist φ diagonalisierbar, dann gibt es eine geordnete Basis $\mathcal{B}$, sodass ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)$ eine Diagonalmatrix ist. Da je zwei Matrixdarstellungen ähnlich sind (siehe Lemma 5.1.3), schließen wir, dass ${}_{\mathcal{C}}\mathrm{M}_{\mathcal{C}}(\varphi)$ diagonalisierbar ist.

Ist umgekehrt ${}_{\mathcal{C}}\mathrm{M}_{\mathcal{C}}(\varphi)$ diagonalisierbar, so ist ${}_{\mathcal{C}}\mathrm{M}_{\mathcal{C}}(\varphi)$ ähnlich zu einer Diagonalmatrix und diese ist nach Lemma 5.1.3 ebenfalls eine Matrixdarstellung von φ. □

Welche Endomorphismen und Matrizen diagonalisierbar sind, können wir mit den Begriffen aus dem vorangegangenen Abschnitt gut beschreiben.

5.2.19 **Satz zur Diagonalisierbarkeit.** *Es sei V ein K-Vektorraum der Dimension $n = \dim_K V$. Sei $\varphi \in \operatorname{End}(V)$ und es seien $\lambda_1, \ldots, \lambda_r$ die paarweise verschiedenen Eigenwerte von φ. Die folgenden Aussagen sind äquivalent*

(i) φ ist diagonalisierbar.

(ii) Es gibt eine Basis von V, die aus Eigenvektoren von φ besteht.

(iii) Es gibt ein Erzeugendensystem von V, das aus Eigenvektoren von φ besteht.

(iv) Die Summe der geometrischen Vielfachheiten ist n, d.h., $\sum_{i=1}^r \gamma(\lambda_i) = n$.

Beweis. „(i) $\implies$ (ii)“: Ist φ diagonalisierbar, so gibt es eine geordnete Basis $\mathcal{B} = (v_1, \ldots, v_n)$ von V dergestalt, dass ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)$ eine Diagonalmatrix mit Diagonaleinträgen $a_1, \ldots, a_n \in K$ ist. Das heißt aber $\varphi(v_i) = a_i v_i$ und damit sind alle Vektoren der Basis $\mathcal{B}$ Eigenvektoren von φ.

„(ii) $\implies$ (iii)“: Da jede Basis ein Erzeugendensystem ist, ist hier nichts zu zeigen.

„(iii) $\implies$ (iv)“: Jeder Eigenvektor liegt in einem Eigenraum $\mathrm{Eig}_\varphi(\lambda_i)$ zu einem der Eigenwerte λ_i. Wenn es also ein Erzeugendensystem E aus Eigenvektoren von φ gibt, dann gilt

$$V = \langle E \rangle \subseteq \left\langle \bigcup_{i=1}^r \mathrm{Eig}_\varphi(\lambda_i) \right\rangle = \sum_{i=1}^r \mathrm{Eig}_\varphi(\lambda_i)$$

und damit $V = \sum_{i=1}^r \mathrm{Eig}_\varphi(\lambda_i)$. Da die Summe der Eigenräume direkt ist, erhalten wir aus Satz 2.3.12

$$n = \dim_K(V) = \sum_{i=1}^r \dim_K \mathrm{Eig}_\varphi(\lambda_i) = \sum_{i=1}^r \gamma(\lambda_i).$$

„(iv) $\implies$ (ii)“: Die Summe der Eigenräume ist direkt, also hat der Unterraum $W = \bigoplus_{i=1}^r \mathrm{Eig}_\varphi(\lambda_i)$ von V nach Satz 2.3.12 die Dimension $\dim_K(W) = \sum_{i=1}^r \gamma(\lambda_i) = n$ und daraus folgt mit [MG, 7.4.5]

$$V = \bigoplus_{i=1}^r \mathrm{Eig}_\varphi(\lambda_i).$$

Für jedes $i \in \{1, \ldots, r\}$ wählen wir eine Basis $\mathcal{B}_i$ von $\mathrm{Eig}_\varphi(\lambda_i)$. Da die Vektoren einer Basis immer $\neq 0$ sind, sind die Elemente von $\mathcal{B}_i$ Eigenvektoren von φ.

Aus Satz 2.3.12 wissen wir, dass dann $\mathcal{B} = \bigcup_{i=1}^r \mathcal{B}_i$ eine Basis von V ist, die nach Konstruktion aus Eigenvektoren von φ besteht.

„(ii) $\implies$ (i)“: Es sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine geordnete Basis, die aus Eigenvektoren von φ zu den Eigenwerten $a_1, \ldots, a_n \in K$ besteht. Dann gilt

$$\varphi(v_i) = a_i v_i$$

und damit ist ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)$ die Diagonalmatrix mit den Diagonaleinträgen $a_1, \ldots, a_n$. Also ist φ diagonalisierbar. □

Diese Ergebnis kann man nun auf Matrizen übertragen. Das folgende Korollar lässt sich mit Lemma 5.2.18 herleiten (Übungsaufgabe). **L**

5.2.20 **Korollar.** *Es sei $A \in \mathrm{M}_{n,n}(K)$ eine Matrix und es seien $\lambda_1, \dots, \lambda_r \in K$ die paarweise verschiedenen Eigenwerte von A. Folgende Aussagen sind äquivalent:*

(i) A ist diagonalisierbar.

(ii) Es gibt eine Basis des K^n, die aus Eigenvektoren von A besteht.

(iii) Die Summe der geometrischen Vielfachheiten ist n, d.h., $\sum_{i=1}^{r} \gamma(\lambda_i) = n$.

5.2.21 **Aufgabe**. Ist die Matrix $A = \begin{pmatrix} \overline{1} & \overline{1} \\ \overline{0} & \overline{0} \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{F}_2)$ diagonalisierbar? **L**

Wie wir im folgenden Beispiel sehen werden, sind nicht alle quadratischen Matrizen diagonalisierbar.

5.2.22 **Beispiel**. Wir betrachten die Matrix $A = \begin{pmatrix} \overline{0} & \overline{1} \\ \overline{0} & \overline{0} \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{F}_2)$. Im Körper $\mathbb{F}_2$ gibt es nur zwei mögliche Eigenwerte von A.

Es gilt

$$\mathrm{Eig}_A(\overline{0}) = \mathrm{Ker}(A) = \langle \begin{pmatrix} \overline{1} \\ \overline{0} \end{pmatrix} \rangle;$$

damit ist $\overline{0}$ ein Eigenwert von A mit geometrischer Vielfachheit $\gamma(\overline{0}) = 1$.

Der Skalar $\overline{1} \in \mathbb{F}_2$ ist hingegen kein Eigenwert, denn $\mathrm{Eig}_A(\overline{1}) = \mathrm{Ker}(A - I_2) = \{0\}$. Also ist die Summe aller geometrischen Vielfachheiten gleich 1 und damit kleiner als 2. Aus Satz 5.2.19 folgt, dass die Matrix A nicht diagonalisierbar ist.

Im Beweis des Satzes haben wir gesehen, dass ein diagonalisierbarer Endomorphismus φ eine diagonale Matrixdarstellung bezüglich einer beliebigen Basis aus Eigenvektoren besitzt. Damit erhält man ein einfaches Verfahren, wie man eine quadratische Matrix diagonalisieren kann. Man bestimmt Basen aller Eigenräume und fügt diese (wenn möglich) zu einer Basis von K^n zusammen. Dazu muss man aber alle Eigenwerte von A kennen. Wie man die Eigenwerte einer Matrix berechnen kann, untersuchen wir im nächsten Abschnitt.

5.2.23 **Diagonalisieren von Matrizen**.

`Gegeben:` $A \in \mathrm{M}_{n,n}(K)$ und alle Eigenwerte $\lambda_1, \dots, \lambda_r$ von A.

`Gesucht:` $S \in \mathrm{GL}_n(K)$, sodass $S^{-1}AS$ eine Diagonalmatrix ist.

`Verfahren:`

Für $i = 1, \dots, r$ berechne eine Basis $\mathcal{B}_i$ von $\mathrm{Eig}_A(\lambda_i) = \mathrm{Ker}(A - \lambda_i I_n)$ mit dem Gaußverfahren. Dann ist $\gamma(\lambda_i) = |\mathcal{B}_i|$ die geometrische Vielfachheit von λ_i.

- Ist $\sum_{i=1}^r \gamma(\lambda_i) < n$, dann ist die Matrix A nicht diagonalisierbar und S existiert nicht.

- Ist $\sum_{i=1}^r \gamma(\lambda_i) = n$, dann erhält man die gesuchte Matrix S, indem man die Vektoren aus den Basen $\mathcal{B}_1, \dots, \mathcal{B}_r$ als Spalten in die Matrix S schreibt.

Die Reihenfolge in der die Vektoren in die Matrix S geschrieben werden, bestimmt die Reihenfolge der Diagonaleinträge. Ist die k-te Spalte ein Vektor aus $\mathcal{B}_i$, dann ist λ_i der k-te Eintrag auf der Diagonalen.

5.2.24 **Beispiel**. Gegeben ist die Matrix

$$A = \begin{pmatrix} 0 & -1 & -3 \\ 0 & 1 & 0 \\ 1 & 1 & 4 \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{R}).$$

Die Eigenwerte dieser Matrix sind $\lambda_1 = 1$ und $\lambda_2 = 3$. Wir werden im nächsten Abschnitt erfahren, wie man die Eigenwerte finden kann. Wir berechnen nun eine Matrix $S \in \mathrm{GL}_n(\mathbb{R})$, sodass $S^{-1}AS$ eine Diagonalmatrix ist.

Wir berechnen den Eigenraum zum Eigenwert $\lambda_1 = 1$. Es gilt

$$\mathrm{Eig}_A(1) = \mathrm{Ker}(A - I_3) = \mathrm{Ker} \begin{pmatrix} -1 & -1 & -3 \\ 0 & 0 & 0 \\ 1 & 1 & 3 \end{pmatrix}.$$

Wir addieren die letzte Zeile zur ersten und tauschen diese dann noch oben. Damit erhalten wir

$$\mathrm{Eig}_A(1) = \mathrm{Ker} \begin{pmatrix} 1 & 1 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \\ -1 \end{pmatrix} \right\rangle.$$

Die zwei angegebenen Vektoren bilden eine Basis des Eigenraumes; es gilt also $\gamma(1) = 2$.

Nun berechnen wir den Eigenraum zum Eigenwert $\lambda_2 = 3$.

$$\mathrm{Eig}_A(3) = \mathrm{Ker}(A - 3I_3) = \mathrm{Ker} \begin{pmatrix} -3 & -1 & -3 \\ 0 & -2 & 0 \\ 1 & 1 & 1 \end{pmatrix} = \mathrm{Ker} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right\rangle.$$

Die geometrische Vielfachheit des Eigenwertes $\lambda_2 = 3$ ist also $\gamma(3) = 1$.

Wir schreiben die gefundenen Vektoren als Spalten in die Matrix S, d.h. $S = \begin{pmatrix} 1 & 3 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & -1 \end{pmatrix}$ und erhalten damit

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

(Bestimmen Sie die Inverse der Matrix S und machen Sie die Probe!) **L**

5.2.25 **Aufgabe**. Wir betrachten die Matrix **L**

$$A = \begin{pmatrix} \overline{1} & \overline{2} \\ \overline{2} & \overline{1} \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{F}_3).$$

aus Aufgabe 5.2.12. Bestimmen Sie eine invertierbare Matrix $S \in \mathrm{GL}_2(\mathbb{F}_3)$, sodass $S^{-1}AS$ eine Diagonalmatrix ist.

5.2.26 **Aufgabe**. Es sei **L**

$$A = \begin{pmatrix} \overline{0} & \overline{0} & \overline{1} & \overline{1} \\ \overline{0} & \overline{1} & \overline{0} & \overline{0} \\ \overline{0} & \overline{1} & \overline{1} & \overline{1} \\ \overline{0} & \overline{1} & \overline{0} & \overline{0} \end{pmatrix} \in \mathrm{M}_{4,4}(\mathbb{F}_2).$$

Geben Sie, wenn möglich, eine invertierbare Matrix $S \in \mathrm{GL}_4(\mathbb{F}_2)$ an, sodass $S^{-1}AS$ eine Diagonalmatrix ist.

5.3. Charakteristisches Polynom und Minimalpolynom

In diesem Abschnitt werden wir jeder quadratischen Matrix und jedem Endomorphismus zwei Polynome zuordnen: das charakteristische Polynom und das Minimalpolynom. Wir beginnen mit dem charakteristischen Polynom $\det(XI_n - A)$, dass man mit den Verfahren aus der letzten Lektion leicht berechnen kann. Wir werden sehen, dass die Nullstellen genau die Eigenwerte von A sind. Im Anschluss definieren wir das Minimalpolynom. Auch das Minimalpolynom hat genau die Eigenwerte als Nullstellen, allerdings ist es im Allgemeinen deutlich schwerer zu bestimmen. Es wird aber in der nächsten Lektion ein entscheidendes Werkzeug für uns sein. Die Verbindung zwischen beiden Polynomen stellt der Satz von Cayley-Hamilton her.

I. Charakteristisches Polynom

In diesem Abschnitt werden wir eine Methode kennenlernen um die Eigenwerte eines Endomorphismus oder einer quadratischen Matrix zu bestimmen. Zuerst konzentrieren wir uns auf Matrizen.

Sei K ein Körper und sei $A \in \mathrm{M}_{n,n}(K)$ eine quadratische Matrix. Wie erkennt man ob $\lambda \in K$ ein Eigenwert von A ist? Dazu machen wir eine einfache aber wichtige Beobachtung.

5.3.1 **Satz.** *Sei K ein Körper und sei $A \in \mathrm{M}_{n,n}(K)$ eine quadratische Matrix. Dann ist $\lambda \in K$ genau dann ein Eigenwert von A, wenn*

$$\det(\lambda I_n - A) = 0$$

ist.

Beweis. Wir wissen, dass λ genau dann ein Eigenwert ist, wenn $\mathrm{Eig}_A(\lambda) \neq \{0\}$ ist. Da

$$\mathrm{Eig}_A(\lambda) = \mathrm{Ker}(A - \lambda I_n) = \mathrm{Ker}(\lambda I_n - A)$$

gilt, ist λ genau dann ein Eigenwert, wenn die Matrix $\lambda I_n - A$ einen nicht-trivialen Kern besitzt. Nach Korollar 4.2.32 hat $\lambda I_n - A$ genau dann einen nicht-trivialen Kern, wenn $\det(\lambda I_n - A) = 0$ gilt. □

Da wir nicht die Determinante der Matrizen $\lambda I_n - A$ für alle Skalare $\lambda \in K$ einzeln bestimmen können, ist es naheliegend anstelle von λ eine Unbestimmte X zu verwenden. Wir betrachten also den Polynomring $K[X]$ und fassen K als den Unterring der konstanten Polynome in $K[X]$ auf; siehe 1.4.10. Die Matrix $A \in \mathrm{M}_{n,n}(K)$ können wir also als Matrix in $\mathrm{M}_{n,n}(K[X])$ auffassen. Die Matrix

$XI_n - A$ ist dann eine quadratische Matrix mit Einträgen im Polynomring $K[X]$.

5.3.2 **Definition** Sei K ein Körper und sei $A \in \mathrm{M}_{n,n}(K)$. Das Polynom

$$\chi_A := \det(XI_n - A) \in K[X]$$

nennt man das *charakteristische Polynom* von A.

5.3.3 **Beispiel**. (a) Wir betrachten die Matrix $A = \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{Q})$. Dann erhalten wir mit 4.2.3 das charakteristische Polynom als

$$\chi_A = \det(XI_2 - A) = \det \begin{pmatrix} X & -1 \\ -2 & X-3 \end{pmatrix} = X(X-3) - 2 = X^2 - 3X - 2.$$

(b) Sei $a \in K$. Dann folgt aus 4.2.20 (c), dass das charakteristische Polynom von $A = aI_n$ durch

$$\chi_A = \det(XI_n - aI_n) = (X-a)^n$$

gegeben ist.

5.3.4 **Aufgabe**. Berechnen Sie das charakteristische Polynom der Matrix A aus Beispiel 5.2.24. **L**

Bevor wir besprechen, wie uns das charakteristische Polynom bei der Bestimmung der Eigenwerte helfen kann, überlegen wir uns zunächst, was wir über die Koeffizienten dieses Polynoms sagen können. Der folgende Satz ist auch hilfreich, um nach der Berechnung eines charakteristischen Polynoms das Ergebnis auf Plausibilität zu überprüfen.

5.3.5 **Satz über die Koeffizienten des charakteristischen Polynoms.** *Sei $n \in \mathbb{N}$ und sei $A \in \mathrm{M}_{n,n}(K)$. Das charakteristische Polynom χ_A ist ein normiertes Polynom vom Grad n. Schreibt man*

$$\chi_A = X^n + c_{n-1}X^{n-1} + \cdots + c_1 X + c_0,$$

dann gilt

$$c_{n-1} = -\operatorname{Spur}(A) \quad \text{und} \quad c_0 = (-1)^n \det(A).$$

Beweis. Sei $XI_n - A = (b_{ij})_{i,j} \in \mathrm{M}_{n,n}(K[X])$. Mit der Leibnizformel erhalten wir

$$\chi_A = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{j=1}^{n} b_{\sigma(j),j}$$

dabei ist $b_{ij} = -a_{ij}$ für $i \neq j$ und $b_{ii} = X - a_{ii}$. Der Grad von b_{ij} ist also ≤ 0 für $i \neq j$ und der Grad von b_{ii} ist 1. Aus Lemma 1.4.7 folgt, dass der Grad des Produktes $\prod_{j=1}^{n} b_{\sigma(j),j}$ höchstens die Anzahl der Fixpunkte von σ ist. Er ist genau dann gleich n, wenn σ die identische Permutation ist. Da jede andere Permutation höchstens $n-2$ Fixpunkte besitzt[1], haben alle anderen Terme einen Grad $\leq n-2$. Die beiden höchsten Koeffizienten von χ_A stimmen also mit denen des Polynoms

$$\prod_{j=1}^{n} b_{j,j} = \prod_{j=1}^{n} (X - a_{jj}) = X^n - \sum_{j=1}^{n} a_{jj} X^{n-1} + \dots$$

überein. Folglich ist χ_A normiert vom Grad n und es gilt $c_{n-1} = -\sum_j a_{jj} = -\operatorname{Spur}(A)$.

Es gilt $c_0 = \chi_A(0)$ und im folgenden Satz 5.3.7 werden wir zeigen, dass $\chi_A(0) = \det(-A)$ gilt. Mit 4.2.20 (c) erhalten wir $c_0 = \det(-A) = (-1)^n \det(A)$. □

5.3.6 **Beispiel**. Mit Satz 5.3.5 und Beispiel 4.2.3 kann man das charakteristische Polynom von (2×2)-Matrizen allgemein angeben. Sei $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$. Es gilt

$$\chi_A = X^2 - \operatorname{Spur}(A)X + \det(A) = X^2 - (a_{11} + a_{22})X + a_{11}a_{22} - a_{12}a_{21}.$$

Wie hilft uns das charakteristische Polynom nun beim Finden der Eigenwerte?

5.3.7 **Satz von den Eigenwerten als Nullstellen.** *Sei $A \in \mathrm{M}_{n,n}(K)$ eine quadratische Matrix über einem Körper K. Für alle $\lambda \in K$ gilt*

$$\chi_A(\lambda) = \det(\lambda I_n - A).$$

Die Eigenwerte von A sind genau die Nullstellen des charakteristischen Polynoms χ_A in K.

[1] Wenn $\sigma(i) \neq i$ gilt, dann muss auch $\sigma(\sigma(i)) \neq \sigma(i)$ sein.

Beweis. Sei $\lambda \in K$ und sei $XI_n - A = (b_{ij})_{i,j} \in \mathrm{M}_{n,n}(K[X])$. Dabei ist $b_{ij} = -a_{ij}$ für $i \neq j$ und $b_{ii} = X - a_{ii}$. Aus Lemma 1.4.19 ist bekannt, dass das Einsetzen von Körperelementen in Polynome mit den Rechenoperationen verträglich ist. Aus der Leibniz-Formel 4.2.1 erhalten wir damit

$$\begin{aligned}\chi_A(\lambda) &= \Big(\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{j=1}^{n} b_{\sigma(j),j}\Big)(\lambda) \\ &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{j=1}^{n} b_{\sigma(j),j}(\lambda) = \det(\lambda I_n - A).\end{aligned}$$

Aus Satz 5.3.1 folgt nun, dass die Eigenwerte von A genau die Nullstellen von χ_A in K sind. □

Der Satz von den Eigenwerten als Nullstellen gibt uns nun die Möglichkeit die Eigenwerte einer quadratischen Matrix A zu berechnen. Dazu bestimmen wir zuerst das charakteristische Polynom χ_A und suchen dann seine Nullstellen. Wie bereits erwähnt wurde, kann es sehr schwierig sein die Nullstellen eines Polynoms zu bestimmen (vgl. 1.4.34). Hier wollen wir dieses Problem ignorieren. Bei Übungsaufgaben dürfen Sie aber davon ausgehen, dass man die Nullstellen durch Ausprobieren oder Anwenden der Lösungsformel für quadratische Gleichungen finden kann.

5.3.8 **Beispiel**. Gegeben ist die Matrix

$$A = \begin{pmatrix} 0 & -1 & -3 \\ 0 & 1 & 0 \\ 1 & 1 & 4 \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{R}).$$

In Aufgabe 5.3.4 haben wir das charakteristische Polynom berechnet

$$\chi_A = X^3 - 5X^2 + 7X - 3.$$

Durch Ausprobieren finden wir die Nullstelle 1 und spalten diese mittels Polynomdivision ab:

$$\chi_A = (X-1)(X^2 - 4X + 3).$$

Das quadratische Polynom $X^2 - 4X + 3$ hat die beiden Lösungen

$$\frac{4 - \sqrt{16 - 12}}{2} = 1 \quad \text{und} \quad \frac{4 + \sqrt{16 - 12}}{2} = 3.$$

Also zerfällt das charakteristische Polynom in Linearfaktoren

$$\chi_A = (X-1)^2(X-3).$$

Die Eigenwerte von A sind also die beiden Nullstellen $\lambda_1 = 1$ und $\lambda_2 = 3$.

5.3.9 **Bemerkung**. Da uns die Nullstellen des charakteristischen Polynoms interessieren, möchten wir das charakteristische Polynom wie im Satz über die Nullstellen 1.4.30 schreiben. Dabei ist es hilfreich beim Berechnen von χ_A umsichtig vorzugehen und *nicht* alle Terme direkt auszumultiplizieren. Oft findet man einen Linearfaktor, der sich abspalten lässt.

Sehen Sie sich nochmal die Lösung zu Aufgabe 5.3.4 an: In der Berechnung von χ_A hatten wir die Gleichung $\chi_A = (X-1)(X^2-4X+3)$ und haben diese dann ausmultipliziert.

5.3.10 **Definition** Sei $A \in \mathrm{M}_{n,n}(K)$ und sei $\lambda \in K$ ein Eigenwert von A. Die Vielfachheit der Nullstelle λ in χ_A nennen wir die *algebraische Vielfachheit* von λ und notieren diese mit $\alpha(\lambda)$.

Es gilt also
$$\chi_A(X) = (X-\lambda)^{\alpha(\lambda)} h$$
für ein Polynom $h \in K[X]$ mit $h(\lambda) \neq 0$; siehe Lemma 1.4.26.

5.3.11 **Beispiel**. In Beispiel 5.3.8 haben wir gesehen, dass die Matrix A das charakteristische Polynom
$$\chi_A = (X-1)^2(X-3)$$
besitzt. Damit ist die algebraische Vielfachheit des Eigenwertes 1 genau $\alpha(1) = 2$ und die algebraische Vielfachheit des Eigenwertes 3 ist $\alpha(3) = 1$.

5.3.12 **Aufgabe**. Bestimmen Sie die Eigenwerte der Matrix **L**

$$A = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix} \in \mathrm{M}_{n,n}(\mathbb{R})$$

sowie die algebraischen und geometrischen Vielfachheiten.

5.3.13 **Satz.** *Sei K ein Körper und sei $A \in \mathrm{M}_{n,n}(K)$. Die transponierte Matrix A^T hat dasselbe charakteristische Polynom wie A, d.h.,*
$$\chi_A = \chi_{A^T}.$$
Dazu gilt: A und A^T haben dieselben Eigenwerte und die geometrischen und algebraischen Vielfachheiten jedes Eigenwertes stimmen überein.

Beweis. Aus Lemma 4.2.14 folgt mithilfe der Rechenregeln 1.2.32 direkt

$$\chi_A = \det(XI_n - A) = \det((XI_n - A)^T) = \det(XI_n - A^T) = \chi_{A^T}.$$

Insbesondere haben A und A^T wegen Satz 5.3.7 dieselben Eigenwerte. Ist λ ein Eigenwert von A und A^T, dann stimmen die algebraischen Vielfachheiten überein, weil die charakteristischen Polynome übereinstimmen. Was ist mit den geometrischen Vielfachheiten? Die geometrische Vielfachheit von λ bzgl. A ist die Dimension des Eigenraumes. Damit erhalten wir

$$\begin{aligned}
\gamma_A(\lambda) &= \dim_K \mathrm{Eig}_A(\lambda) = \dim_K \mathrm{Ker}(A - \lambda I_n) \\
&= n - \mathrm{Rg}(A - \lambda I_n) && \text{([MG, 8.3.14])} \\
&= n - \mathrm{Rg}((A - \lambda I_n)^T) && \text{(Satz 2.6.16)} \\
&= n - \mathrm{Rg}(A^T - \lambda I_n) && \text{(Rechenregeln 1.2.32)} \\
&= \dim_K \mathrm{Ker}(A^T - \lambda I_n) = \dim_K \mathrm{Eig}_{A^T}(\lambda) = \gamma_{A^T}(\lambda). && \square
\end{aligned}$$

Wir kommen nun zu einer wichtigen Beobachtung: Ähnliche Matrizen haben dasselbe charakteristische Polynom.

5.3.14 **Satz.** *Sei K ein Körper. Sind $A, B \in \mathrm{M}_{n,n}(K)$ ähnlich, dann gilt $\chi_A = \chi_B$.*

Beweis. Übungsaufgabe! Zeigen Sie zuerst, dass $XI_n - A$ und $XI_n - B$ ähnlich sind. $\square$ **L**

Aus Satz 5.3.14 und Satz 5.3.5 folgt insbesondere:

5.3.15 **Korollar.** *Ähnliche Matrizen haben dieselbe Spur.*

Im Allgemeinen ist weder die Gleichheit der Spuren noch die Gleichheit der charakteristischen Polynome hinreichend für die Ähnlichkeit von Matrizen.

5.3.16 **Aufgabe**. Geben Sie zwei Matrizen an, die dasselbe charakteristische Polynom haben, aber nicht ähnlich sind. **L**

Sei V ein endlich-dimensionaler K-Vektorraum und sei $\varphi \in \mathrm{End}(V)$. Aus der Transformationsformel für Endomorphismen 4.4.5 wissen wir, dass je zwei Matrixdarstellungen von φ ähnlich sind. Mit Satz 5.3.14 haben also alle Matrixdarstellungen von φ dasselbe charakteristische Polynom. Dadurch ist die folgende Definition sinnvoll.

5.3.17 **Definition** Sei V ein endlich-dimensionaler K-Vektorraum und sei $\varphi \in \mathrm{End}(V)$. Wir definieren das *charakteristische Polynom* von φ als

$$\chi_\varphi = \chi_A$$

für eine Matrixdarstellung $A = {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)$ von φ. Das charakteristische Polynom hängt nicht von der gewählten Matrixdarstellung A ab.

Aus dem Satz über die Eigenwerte als Nullstellen 5.3.7 und Lemma 5.2.7 erhalten wir das folgende Ergebnis.

5.3.18 **Korollar.** *Sei V ein endlich-dimensionaler K-Vektorraum und sei $\varphi \in \mathrm{End}(V)$. Die Eigenwerte von φ sind genau die Nullstellen von χ_φ in K.*

Die Vielfachheit der Nullstelle λ von χ_φ, nennen wir die *algebraische Vielfachheit* des Eigenwertes λ.

Was kann man über das charakteristische Polynom sagen, wenn sich V in eine direkte Summe invarianter Unterräume zerlegt?

5.3.19 **Satz.** *Sei V ein endlich-dimensionaler K-Vektorraum und sei $\varphi \in \mathrm{End}(V)$. Ist $V = U_1 \oplus U_2$ direkte Summe zweier φ-invarianter Unterräume, dann gilt*

$$\chi_\varphi = \chi_{\varphi|_{U_1}} \chi_{\varphi|_{U_2}}.$$

Beweis. Das charakteristische Polynom eines Endomorphismus ist das charakteristische Polynom einer beliebigen Matrixdarstellung; siehe 5.3.17. Wir wählen Basen von U_1 und U_2 wie in Satz 5.1.13, sodass die Matrixdarstellung ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi) = \mathrm{diag}({}_{\mathcal{B}_1}\mathrm{M}_{\mathcal{B}_1}(\varphi|_{U_1}), {}_{\mathcal{B}_2}\mathrm{M}_{\mathcal{B}_2}(\varphi|_{U_2}))$ blockdiagonal ist. Damit ist auch $XI_n - {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)$ eine Blockdiagonalmatrix. Sagen wir $\dim_K(V) = n$ und $\dim_K(U_1) = k$, dann folgt aus Satz 4.3.15

$$\chi_\varphi = \det(XI_k - {}_{\mathcal{B}_1}\mathrm{M}_{\mathcal{B}_1}(\varphi|_{U_1}))\det(XI_{n-k} - {}_{\mathcal{B}_2}\mathrm{M}_{\mathcal{B}_2}(\varphi|_{U_2})) = \chi_{\varphi|_{U_1}}\chi_{\varphi|_{U_2}}. \qquad \square$$

Zum Abschluss wollen wir noch eine nützliche Überlegung anstellen: Was haben die geometrischen und algebraischen Vielfachheiten miteinander zu tun? Damit lässt sich auch die Diagonalisierbarkeit besser verstehen.

5.3.20 **Satz.** *Sei V ein endlich-dimensionaler K-Vektorraum und sei $\varphi \in \mathrm{End}(V)$. Für jeden Eigenwert $\lambda \in K$ von φ gilt*

$$\gamma(\lambda) \leq \alpha(\lambda).$$

Der Endomorphismus φ ist genau dann diagonalisierbar, wenn χ_φ in Linearfaktoren zerfällt und $\gamma(\lambda) = \alpha(\lambda)$ für jeden Eigenwert λ gilt.

Beweis. Es sei λ ein Eigenwert mit geometrischer Vielfachheit $k = \gamma(\lambda)$. Wir wählen eine geordnete Basis $(v_1, \dots, v_k)$ des Eigenraumes $\mathrm{Eig}_\varphi(\lambda)$ und ergänzen diese zu einer geordneten Basis $\mathcal{B} = (v_1, \dots, v_k, v_{k+1}, \dots, v_n)$ von V. Die Matrixdarstellung ${}_\mathcal{B}\mathrm{M}_\mathcal{B}(\varphi)$ ist dann eine Blockdreiecksmatrix, wobei der obere linke $(k \times k)$-Block λI_k ist. Verwendet man diese Matrixdarstellung um das charakteristische Polynom χ_φ zu berechnen, erhält man mit Satz 4.3.15, dass $(X - \lambda)^k$ ein Teiler von χ_φ ist. Das heißt, $\gamma(\lambda) = k \leq \alpha(\lambda)$.

Wir nehmen nun an, dass χ_φ in Linearfaktoren zerfällt. Die Summe der algebraischen Vielfachheiten ist also genau $n = \mathrm{Grad}(\chi_\varphi) = \dim_K V$. Bezeichnen wir mit $\lambda_1, \dots, \lambda_r \in K$ die paarweise verschiedenen Eigenwerte von φ, dann gilt $n = \sum_{i=1}^r \alpha(\lambda_i)$. Gilt zusätzlich $\gamma(\lambda_i) = \alpha(\lambda_i)$ für alle $i \in \{1, 2, \dots, r\}$, dann folgt die Diagonalisierbarkeit von φ aus Satz 5.2.19 (iv).

Umgekehrt nehmen wir nun an, dass φ diagonalisierbar ist. Sei ${}_\mathcal{C}\mathrm{M}_\mathcal{C}(\varphi)$ eine diagonale Matrixdarstellung; dabei erscheint jeder Eigenwert λ_i gemäß seiner geometrischen Vielfachheit $\gamma(\lambda_i)$ auf der Diagonalen. Mit 4.2.11 folgt $\chi_\varphi = \prod_{i=1}^r (X - \lambda_i)^{\gamma(\lambda_i)}$ und damit $\gamma(\lambda_i) = \alpha(\lambda_i)$ für alle $i \in \{1, 2, \dots, r\}$. □

II. Minimalpolynom

5.3.21 $\mathrm{M}_{n,n}(K)$ **und** $\mathrm{End}(V)$ **als Algebren**. Sei K ein Körper. Aus den Mathematischen Grundlagen [MG, 6.1.3] ist bekannt, dass $\mathrm{M}_{n,n}(K)$ ein K-Vektorraum ist. Wir wissen außerdem aus Satz 1.2.27, dass $\mathrm{M}_{n,n}(K)$ mit der Addition und Multiplikation von Matrizen ein Ring ist. Auf $\mathrm{M}_{n,n}(K)$ gibt es also eine Addition, eine Multiplikation und eine Skalarmultiplikation mit Elementen aus K. Dazu vertauschen die Skalarmultiplikation und die Ringmultiplikation, d.h.,

$$(ab)(AB) = (aA)(bB) \tag{5.3.a}$$

für alle $a, b \in K$ und $A, B \in \mathrm{M}_{n,n}(K)$; siehe [MG, 2.4.1 (c)].

Ist V ein K-Vektorraum, dann hat $\mathrm{End}(V)$ ebenfalls diese Eigenschaften: (1) $\mathrm{End}(V)$ ist ein K-Vektorraum (siehe 2.5.1) und (2) $\mathrm{End}(V)$ ist mit der Addition und Komposition von Endomorphismen ein Ring (siehe 4.4.3). Auf $\mathrm{End}(V)$ gibt es also ebenfalls eine Addition, eine Multiplikation (die Komposition) und eine Skalarmultiplikation mit Elementen aus K. Dazu gilt wieder (Übungsaufgabe!) **L**

$$(ab)(\varphi \circ \psi) = (a\varphi) \circ (b\psi)$$

für alle $a, b \in K$ und alle $\varphi, \psi \in \mathrm{End}(V)$.

Man nennt $\mathrm{M}_{n,n}(K)$ und $\mathrm{End}(V)$ aufgrund dieser Eigenschaften *Algebren*[2] über K.

[2]Einzahl: Algebra; Mehrzahl: Algebren

Wir sind in diesem Lehrtext auch schon einigen weiteren Algebren begegnet. Zum Beispiel ist der Körper $\mathbb{C}$ der komplexen Zahlen eine Algebra über dem Körper $\mathbb{R}$ der reellen Zahlen. Weil uns hier nur $\mathrm{M}_{n,n}(K)$ und $\mathrm{End}(V)$ interessieren, werden wir darauf verzichten, eine allgemeine Definition des Begriffs „Algebra" zu geben.

In 1.4.16 haben wir gesehen, dass jedes Polynom $f \in K[X]$ eine Polynomfunktion definiert, indem man $x \in K$ für die Unbestimmte X einsetzt. Wie wir nun sehen werden, kann man in ein Polynom auch Elemente einer Algebra über K einsetzen.

5.3.22 **Einsetzen in Polynome**. Sei $\mathcal{A}$ eine Algebra über K (also für uns $\mathcal{A} = \mathrm{M}_{n,n}(K)$ oder $\mathcal{A} = \mathrm{End}(V)$). Sei K ein Körper und sei $g \in K[X]$ ein Polynom. Wir schreiben $g = \sum_i a_i X^i$ mit Koeffizienten $a_i \in K$.

Ist $x \in \mathcal{A}$ gegeben, dann definieren wir

$$g(x) \coloneqq \sum_i a_i x^i \in \mathcal{A}.$$

Was soll das bedeuten? Hier ist x^i das Element, das man erhält, wenn man x mit der Ringmultiplikation von $\mathcal{A}$ genau i-mal mit sich multipliziert. Dabei trifft man die Festlegung, dass x^0 das Einselement von $\mathcal{A}$ ist. Also $x^0 = I_n$ falls $\mathcal{A} = \mathrm{M}_{n,n}(K)$ ist bzw. $x^0 = \mathrm{id}_V$ falls $\mathcal{A} = \mathrm{End}(V)$ ist. Dann ist $a_i x^i$ durch die Skalarmultiplikation von K auf $\mathcal{A}$ definiert und die Summe bildet man mithilfe der Addition in $\mathcal{A}$. Da fast alle Koeffizienten a_i gleich Null sind, handelt es sich um eine Summe mit endlich vielen Summanden.

Wir sehen also, dass wir für diese Definition genau die drei Rechenoperationen Addition, Multiplikation und Skalarmultiplikation benötigen.

5.3.23 **Beispiel**. Wir betrachten den Körper $\mathbb{R}$ der reellen Zahlen und das Polynom $g = X^2 - 4X - 1 \in \mathbb{R}[X]$. Wir möchten die quadratische Matrix

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{R})$$

in das Polynom g einsetzen. Nach Definition gilt $A^0 = I_2$ und $A^1 = A$. Dazu berechnen wir

$$A^2 = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 8 \\ 0 & 9 \end{pmatrix}.$$

Damit erhalten wir

$$g(A) = A^2 - 4A - I_2 = \begin{pmatrix} 1 & 8 \\ 0 & 9 \end{pmatrix} + \begin{pmatrix} -4 & -8 \\ 0 & -12 \end{pmatrix} + \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} -4 & 0 \\ 0 & -4 \end{pmatrix}.$$

5.3.24 **Aufgabe**. Es seien $f = X^3 + \overline{2}X + \overline{2} \in \mathbb{F}_5[X]$ und **L**

$$A = \begin{pmatrix} \overline{2} & \overline{4} \\ \overline{2} & \overline{3} \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{F}_5).$$

Bestimmen Sie $f(A)$.

5.3.25 **Beispiel**. Sei V ein K-Vektorraum und sei $\lambda \in K$. Wir betrachten den Endomorphismus $\varphi = \lambda \operatorname{id}_V$. Sei $g = \sum_{i=0}^n a_i X^i \in K[X]$ ein Polynom. Um zu verstehen, was der Endomorphismus $g(\varphi)$ macht, setzen wir einen Vektor $v \in V$ ein:

$$g(\varphi)(v) = \sum_{i=0}^n a_i(\lambda \operatorname{id}_V)^i(v) = \sum_{i=0}^n a_i \lambda^i \operatorname{id}_V(v) = \sum_{i=0}^n a_i \lambda^i v = g(\lambda)v.$$

Das heißt, es gilt $g(\lambda \operatorname{id}_V) = g(\lambda) \operatorname{id}_V$.

5.3.26 **Beispiel**. Es sei $A \in \mathrm{M}_{n,n}(K)$ gegeben und $f_A \colon K^n \to K^n$ sei der zugehörige Endomorphismus mit $f_A(v) = Av$. Für jedes Polynom $g = \sum_i a_i X^i \in K[X]$ gilt

$$g(f_A) = f_{g(A)}.$$

Um das einzusehen, setzen wir einen Vektor $v \in K^n$ ein:

$$\begin{aligned} g(f_A)(v) &= \sum_i a_i (f_A)^i(v) = \sum_i a_i(\underbrace{f_A \circ \cdots \circ f_A}_{i \text{ mal}})(v) \\ &= \sum_i a_i A^i v = g(A)v = f_{g(A)}(v). \end{aligned}$$

5.3.27 **Lemma.** (Rechenregeln beim Einsetzen) *Seien K ein Körper, V ein K-Vektorraum und $f, g \in K[X]$ zwei Polynome.*

(i) Für jede Matrix $A \in \mathrm{M}_{n,n}(K)$ gelten

$$(f+g)(A) = f(A) + g(A) \quad \textit{und} \quad (f \cdot g)(A) = f(A)g(A).$$

(ii) Für jeden Endomorphismus $\varphi \in \operatorname{End}(V)$ gelten

$$(f+g)(\varphi) = f(\varphi) + g(\varphi) \quad \textit{und} \quad (f \cdot g)(\varphi) = f(\varphi) \circ g(\varphi).$$

<u>*Beweis*</u>. Der Beweis ist im wesentlichen derselbe wie für Lemma 1.4.19. Wir beweisen die Aussage hier für Matrizen und überlassen es der Leserin/dem Leser sich zu überlegen, weshalb derselbe Beweis auch für Endomorphismen funktioniert.

Es seien $f = \sum_i a_i X^i$ und $g = \sum_i b_i X^i$. Nach Definition der Addition im Polynomring (siehe 1.4.5) gilt $f + g = \sum_i (a_i + b_i) X^i$. Damit erhalten wir

$$\begin{aligned}(f+g)(A) &= \sum_i (a_i + b_i) A^i \\ &= \sum_i (a_i A^i + b_i A^i) && \text{([MG, 2.2.3 (b)])} \\ &= \sum_i a_i A^i + \sum_i b_i A^i = f(A) + g(A).\end{aligned}$$

Nach Definition der Multiplikation im Polynomring (siehe 1.4.5) gilt $f \cdot g = \sum_i c_i X^i$ mit

$$c_i = \sum_{j=0}^{i} a_j b_{i-j}.$$

Damit erhalten wir

$$\begin{aligned}(f \cdot g)(A) &= \sum_i c_i A^i = \sum_i \sum_{j=0}^{i} a_j b_{i-j} A^i \\ &= \sum_i \sum_{j=0}^{i} (a_j A^j)(b_{i-j} A^{i-j}) && (5.3.a) \\ &= \sum_{j,j'} (a_j A^j)(b_{j'} A^{j'}) && (\text{setze } j' = i - j) \\ &= \Big(\sum_j a_j A^j\Big)\Big(\sum_{j'} b_{j'} A^{j'}\Big) = f(A) g(A). && \square\end{aligned}$$

5.3.28 **Aufgabe**. Sei V ein K-Vektorraum. Es sei $\varphi \in \operatorname{End}(V)$ und es sei $U \subseteq V$ ein φ-invarianter Unterraum. Zeigen Sie: Ist $g \in K[X]$, dann ist U auch ein $g(\varphi)$-invarianter Unterraum. **L**

5.3.29 **Definition** Sei V ein K-Vektorraum und sei $\varphi \in \operatorname{End}(V)$. Wir sagen das Polynom $g \in K[X]$ *annulliert* φ, wenn

$$g(\varphi) = 0$$

gilt. Die Menge $\mathcal{I}_\varphi$ aller Polynome, die φ annullieren, nennen wir das *Verschwindungsideal* von φ, d.h.,

$$\mathcal{I}_\varphi = \{g \in K[X] \mid g(\varphi) = 0\}.$$

5.3.30 **Anmerkungen zur Definition**. (a) Ist $A \in \mathrm{M}_{n,n}(K)$, dann schreiben wir $\mathcal{I}_A$ anstelle von $\mathcal{I}_{f_A}$ und nennen $\mathcal{I}_A$ das Verschwindungsideal von A. Aus Beispiel 5.3.26 folgt

$$\mathcal{I}_A = \{g \in K[X] \mid g(A) = 0\}.$$

(b) Das Verschwindungsideal ist nie leer, denn es enthält immer das Nullpolynom.

(c) Das Verschwindungsideal $\mathcal{I}_\varphi$ ist ein Ideal im Sinne von Definition 1.4.45. Es ist eine gute Übungsaufgabe, die Eigenschaften (I1) und (I2) nachzuweisen. **L**

5.3.31 **Beispiel**. Sei V ein K-Vektorraum. Es sei $0 \in \mathrm{End}(V)$ die Nullabbildung, d.h., $0(v) = 0$ für alle $v \in V$. Sei $g = \sum_i a_i X^i \in K[X]$. Dann gilt

$$g(0) = \sum_i a_i 0^i = a_0 \operatorname{id}_V + a_1 0 + a_2 0 + \cdots = a_0 \operatorname{id}_V .$$

Das Polynom g annulliert 0 also genau dann, wenn der konstante Koeffizient a_0 gleich 0 ist. Anders ausgedrückt heißt das: g annulliert 0 genau dann, wenn X ein Teiler von g ist. Das Verschwindungsideal $\mathcal{I}_0$ ist in diesem Fall das von X erzeugte Hauptideal $XK[X]$.

Gibt es im Allgemeinen immer ein nicht-triviales Polynom, dass einen Endomorphismus φ annulliert? Im nächsten Satz zeigen wir, dass dies in endlich-dimensionalen Räumen immer der Fall ist. Das führt uns zum wichtigen Begriff des Minimalpolynoms.

5.3.32 **Definition und Satz.** *Sei V ein endlich-dimensionaler K-Vektorraum und sei $\varphi \in \mathrm{End}(V)$. Dann existiert ein Polynom $g \neq 0$ in $K[X]$, das φ annulliert. Es gilt also $\mathcal{I}_\varphi \neq \{0\}$.*

Nach Satz 1.4.49 besitzt $\mathcal{I}_\varphi$ einen eindeutigen normierten Erzeuger, den wir mit μ_φ bezeichnen. Man nennt μ_φ das Minimalpolynom *von φ.*

Beweis. Es sei $n = \dim_K(V)$ die Dimension von V. Die Menge aller Endomorphismen $\mathrm{End}(V) = \mathrm{Hom}_K(V, V)$ ist ebenfalls ein Vektorraum und es gilt $\dim_K(\mathrm{End}(V)) = n^2$; siehe 2.5.1. Betrachten wir also die Elemente

$$\operatorname{id}_V = \varphi^0, \varphi, \varphi^2, \varphi^3, \ldots, \varphi^{n^2}$$

von $\mathrm{End}(V)$, so können diese nicht linear unabhängig sein, denn $n^2+1 > \dim_K(\mathrm{End}(V))$. Sie sind also linear abhängig und wir finden Skalare $a_0, \ldots, a_{n^2} \in K$ (nicht alle gleich 0), sodass

$$0 = \sum_{i=0}^{n^2} a_i \varphi^i$$

ist. Mit diesen gefundenen Skalaren definieren wir das Polynom

$$g = \sum_{i=0}^{n^2} a_i X^i$$

und stellen fest, dass nun $g(\varphi) = 0$ gilt, d.h., $g \in \mathcal{I}_\varphi$. Da die Koeffizienten a_i nicht alle gleich 0 sind, ist g nicht das Nullpolynom. $\square$

5.3.33 Anmerkungen zum Minimalpolynom. (a) Aus dem Beweis von Satz 1.4.49 ist bekannt, dass der normierte Erzeuger eines Ideals das eindeutige normierte Polynom von kleinstem Grad im Ideal ist. Das heißt, das Minimalpolynom von φ ist das normierte Polynom vom kleinsten Grad, das φ annulliert.

(b) Ist $A \in \mathrm{M}_{n,n}(K)$, dann schreiben wir μ_A anstelle von μ_{f_A} und nennen dieses Polynom das *Minimalpolynom von A*.

(c) Ist $\dim_K(V) > 0$ (also $V \neq \{0\}$), dann hat das Minimalpolynom μ_φ immer mindestens Grad 1. In der Tat, falls $g = aX^0$ Polynom vom Grad 0 ist (d.h. $a \neq 0$), dann gilt $g(\varphi) = a\,\mathrm{id}_V$. Da $V \neq \{0\}$ ist, ist die identische Abbildung nicht die Nullabbildung. Also wird φ nicht von einem Polynom vom Grad 0 annulliert.

5.3.34 Beispiel. Wir gehen zurück zu Beispiel 5.3.31. Dort haben wir festgestellt, dass das Verschwindungsideal der Nullabbildung $0 \in \mathrm{End}(V)$ das von X erzeugte Hauptideal ist. Also ist $\mu_0 = X$ das Minimalpolynom der Nullabbildung.

5.3.35 Beispiel. Wir betrachten die Matrix

$$A = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{Q}).$$

Wir stellen zunächst fest, dass das Polynom $p = (X-3)^2 = X^2 - 6X + 9$ die Matrix A annulliert:

$$p(A) = A^2 - 6A + 9I_2 = \begin{pmatrix} 9 & 6 \\ 0 & 9 \end{pmatrix} + \begin{pmatrix} -18 & -6 \\ 0 & -18 \end{pmatrix} + \begin{pmatrix} 9 & 0 \\ 0 & 9 \end{pmatrix} = 0.$$

Handelt es sich bei p um das Minimalpolynom von A? Aus 5.3.33 (c) wissen wir, dass Polynome vom Grad 0 nicht als Minimalpolynom in Frage kommen. Also müssen wir prüfen, ob ein normiertes Polynom vom Grad 1 die Matrix A annulliert. Ist $g \in \mathbb{Q}[X]$ ein normiertes Polynom vom Grad 1, dann hat g die Gestalt $g = X - a$ für ein $a \in \mathbb{Q}$. Es gilt dann

$$g(A) = A - aI_2 = \begin{pmatrix} 3-a & 1 \\ 0 & 3-a \end{pmatrix} \neq 0.$$

Da das Verschwindungsideal keine Polynome von Grad 1 oder 0 enthält, muss p das Minimalpolynom von A sein.

5.3.36 **Aufgabe**. Bestimmen Sie das Minimalpolynom der Matrix **L**

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{R}).$$

Die wichtigste Eigenschaft des Minimalpolynoms halten wir in folgendem Satz fest.

5.3.37 **Satz.** *Seien V ein endlich-dimensionaler K-Vektorraum und $\varphi \in \mathrm{End}(V)$. Sei $g \in K[X]$ ein Polynom. Dann sind folgende Aussagen äquivalent*

(a) g annulliert φ

(b) μ_φ teilt g.

Beweis. „(a) $\Longrightarrow$ (b)“: Gilt $g(\varphi) = 0$, dann liegt g im Verschwindungsideal $\mathcal{I}_\varphi$. Da μ_φ dieses Ideal erzeugt, gibt es ein Polynom $h \in K[X]$ mit $\mu_\varphi h = g$, d.h., μ_φ teilt g.

„(b) $\Longrightarrow$ (a)“: Teilt μ_φ das Polynom g, dann gibt es ein Polynom $h \in K[X]$ mit $g = \mu_\varphi h$. Mithilfe von Lemma 5.3.27 erhalten wir

$$g(\varphi) = \mu_\varphi(\varphi) \circ h(\varphi) = 0;$$

d.h. g annulliert φ. □

5.3.38 **Satz über die Nullstellen des Minimalpolynoms.** *Sei V ein K-Vektorraum von endlicher Dimension und sei $\varphi \in \mathrm{End}(V)$. Die Nullstellen des Minimalpolynoms μ_φ sind genau die Eigenwerte von φ.*

Beweis. Sei $\mu_\varphi = \sum_i a_i X^i$ das Minimalpolynom von φ. Sei λ ein Eigenwert von φ und sei $v \neq 0$ ein zugehöriger Eigenvektor. Wir bemerken, dass $\varphi^i(v) = \lambda^i v$ für alle $i \in \mathbb{N}_0$ gilt. Damit erhalten wir

$$0 = \mu_\varphi(\varphi)(v) = \sum_i a_i \varphi^i(v) = \sum_i a_i \lambda^i v = \mu_\varphi(\lambda) v.$$

Da $v \neq 0$ ist, folgt $\mu_\varphi(\lambda) = 0$ und λ ist eine Nullstelle von μ_φ.

Sei nun umgekehrt $a \in K$ eine Nullstelle von μ_φ. Wir müssen zeigen, dass a ein Eigenwert von φ ist. Mit dem Abspaltungssatz 1.4.21 schreiben wir

$$\mu_\varphi = (X - a)h \tag{5.3.b}$$

für ein Polynom $h \in K[X]$ von kleinerem Grad als μ_φ. Aus Lemma 1.4.7 sehen wir, dass μ_φ das Polynom h nicht teilt. Mit Satz 5.3.37 folgt, dass h den Endomorphismus φ nicht annulliert, d.h., $h(\varphi) \neq 0$. Also gibt es einen Vektor $v \in V$ mit der Eigenschaft

$$h(\varphi)(v) \neq 0.$$

Wir setzen $w = h(\varphi)(v)$ und behaupten, dass dies ein Eigenvektor von φ ist. Aus Gleichung (5.3.b) und $\mu_\varphi(\varphi) = 0$ folgt $\varphi \circ h(\varphi) = ah(\varphi)$ und damit

$$\varphi(w) = \varphi(h(\varphi)(v)) = ah(\varphi)(v) = aw.$$

Also ist w ein Eigenvektor von φ zum Eigenwert a. □

III. Der Satz von Cayley-Hamilton

Sei V ein endlich-dimensionaler K-Vektorraum und sei $\varphi \in \mathrm{End}(V)$. Wir kennen nun zwei Polynome – das charakteristische Polynom χ_φ und das Minimalpolynom μ_φ – die beide die Eigenschaft haben, dass Ihre Nullstellen genau die Eigenwerte von φ sind. Gibt es einen Zusammenhang zwischen diesen Polynomen? Die Antwort darauf gibt der wichtige Satz von Cayley[3] und Hamilton[4].

5.3.39 **Satz von Cayley-Hamilton.** *Sei V ein endlich-dimensionaler K-Vektorraum und sei $\varphi \in \mathrm{End}(V)$. Das charakteristische Polynom χ_φ annulliert φ, d.h.,*

$$\chi_\varphi(\varphi) = 0.$$

Insbesondere folgt aus dem Satz von Cayley-Hamilton und Satz 5.3.37 direkt

5.3.40 **Korollar.** *Das Minimalpolynom μ_φ teilt das charakteristische Polynom χ_φ.*

Man kann den Satz von Cayley-Hamilton auch für quadratische Matrizen formulieren.

[3] Arthur Cayley: britischer Mathematiker, 1821–1895.
[4] William Rowan Hamilton: irischer Mathematiker, 1805–1865.

5.3.41 **Satz von Cayley-Hamilton für Matrizen.** *Sei K ein Körper und sei $A \in \mathrm{M}_{n,n}(K)$. Das charakteristische Polynom χ_A annulliert A, d.h.,*

$$\chi_A(A) = 0.$$

Das Minimalpolynom μ_A teilt χ_A.

Aus dem Satz von Cayley-Hamilton erhalten wir eine weitere nützliche Folgerung.

5.3.42 **Korollar.** *Das Minimalpolynom einer $(n \times n)$-Matrix hat höchstens Grad n.*

Beweis. Da μ_A das charakteristische Polynom χ_A teilt, gilt mit Satz 5.3.5

$$\mathrm{Grad}(\mu_A) \leq \mathrm{Grad}(\chi_A) = n. \qquad \square$$

5.3.43 **Aufgabe.** Verifizieren Sie die Aussage des Satzes von Cayley-Hamilton für die Matrix $A = \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{Q})$ aus Beispiel 5.3.3. **L**

5.3.44 **Aufgabe.** Verwenden Sie die Formel aus Beispiel 5.3.6, um den Satz von Cayley-Hamilton für 2×2-Matrizen zu beweisen. **L**

Wir werden den Satz für Matrizen beweisen und dann daraus den Satz für Endomorphismen ableiten.

Beweis von Satz 5.3.41. Es sei $B = XI_n - A \in \mathrm{M}_{n,n}(K[X])$. Wir betrachten die Adjunkte $\mathrm{Adj}(B) \in \mathrm{M}_{n,n}(K[X])$. Die Einträge der Adjunkten sind die Kofaktoren $(-1)^{i+j}\det(B_{ji})$ und damit (bis auf das Vorzeichen) Determinanten von $(n-1) \times (n-1)$-Matrizen. Mit der Leibniz-Formel (4.2.a) und Lemma 1.4.7 machen wir eine hilfreiche Beobachtung: Da die Einträge von B Polynome vom Grad ≤ 1 sind, sind die Einträge von $\mathrm{Adj}(B)$ Polynome vom Grad höchstens $n-1$. In der Tat, in der Leibniz-Formel für die Determinante einer $(n-1) \times (n-1)$-Matrix werden jeweils $n-1$ Matrixeinträge multipliziert, die dann aufsummiert werden. Wir können also $\mathrm{Adj}(B)$ in der Form

$$\mathrm{Adj}(B) = \sum_{j=0}^{n-1} D_j X^j$$

für Matrizen $D_j \in \mathrm{M}_{n,n}(K)$ schreiben.

Wir verwenden den Adjunktensatz 4.3.23 und erhalten

$$\mathrm{Adj}(B)B = \det(B)I_n = \chi_A I_n. \tag{5.3.c}$$

Es seien $c_0, \dots, c_n$ die Koeffizienten des charakteristischen Polynoms (es gilt $c_n = 1$), d.h.,

$$\chi_A = \sum_{j=0}^{n} c_j X^j.$$

Wir multiplizieren nun $\mathrm{Adj}(B)B$ aus und sortieren nach den Potenzen von X; damit erhalten wir

$$\begin{aligned}
\mathrm{Adj}(B)B &= \Big(\sum_{j=0}^{n-1} D_j X^j\Big)(XI_n - A) \\
&= -D_0A + \Big(\sum_{j=1}^{n-1}(D_{j-1} - D_jA)X^j\Big) + D_{n-1}X^n.
\end{aligned}$$

Aus Gleichung (5.3.c) sehen wir durch Vergleich der Koeffizienten von X^j

$$\begin{aligned}
c_0 I_n &= -D_0 A \\
c_1 I_n &= D_0 - D_1 A \\
c_2 I_n &= D_1 - D_2 A \\
&\ \vdots \\
c_{n-1} I_n &= D_{n-2} - D_{n-1} A \\
c_n I_n &= D_{n-1}
\end{aligned}$$

Multiplizieren wir $c_j I_n$ von rechts mit A^j erhalten wir $c_j A^j$. Wenn wir diese Terme aufsummieren, erhalten wir $\chi_A(A) = \sum_{j=0}^{n} c_j A^j$. Damit erhalten wir einen Ausdruck in dem sich alle Terme wegheben:

$$\begin{aligned}
\chi_A(A) = \sum_{j=0}^{n} c_j A^j &= \sum_{j=0}^{n} (c_j I_n) A^j \\
&= -D_0A + \Big(\sum_{j=1}^{n-1}(D_{j-1} - D_jA)A^j\Big) + D_{n-1}A^n \\
&= -D_0A + \Big(\sum_{j=1}^{n-1} D_{j-1}A^j - D_jA^{j+1}\Big) + D_{n-1}A^n \\
&= -D_0A + D_0A - D_1A^2 + D_1A^2 - \cdots - D_{n-1}A^n + D_{n-1}A^n = 0.
\end{aligned}$$

Also annulliert χ_A die Matrix A und damit ist das Minimalpolynom μ_A ein Teiler von χ_A. $\square$

Um den Satz für Endomorphismen aus der Aussage für Matrizen herzuleiten, machen wir eine nützliche Beobachtung: Die Matrixdarstellung von Endomorphismen verträgt sich mit dem Einsetzen in Polynome.

5.3.45 **Lemma.** *Sei V ein endlich-dimensionaler K-Vektorraum und sei $\varphi \in \mathrm{End}(V)$. Sei $\mathcal{B}$ eine geordnete Basis von V. Für jedes Polynom $g \in K[X]$ gilt*

$$ {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(g(\varphi)) = g({}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)). $$

Beweis. Aus den Mathematischen Grundlagen [MG, 9.1.9] ist bekannt, dass die Abbildung ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}} \colon \mathrm{End}(V) \to \mathrm{M}_{n,n}(K)$ ein Isomorphismus ist. Insbesondere ist die Matrixdarstellung mit Addition und Skalarmultiplikation verträglich. Aus [MG, 9.3.1] wissen wir außerdem, dass ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}$ Komposition in Matrizenmultiplikation verwandelt. Insbesondere gilt ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi^j) = {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)^j$. Es sei nun $g = \sum_j a_j X^j$. Mit diesen Eigenschaften erhalten wir

$$ \begin{aligned} {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(g(\varphi)) &= {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\sum_j a_j \varphi^j) \\ &= \sum_j a_j \, {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi^j) && \text{([MG, 9.1.9])} \\ &= \sum_j a_j \, {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)^j && \text{([MG, 9..3.1])} \\ &= g({}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)). && \square \end{aligned} $$

5.3.46 **Bemerkung**. Im Beweis von Lemma 5.3.45 haben wir verwendet, dass die bijektive Abbildung ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}} \colon \mathrm{End}(V) \to \mathrm{M}_{n,n}(K)$ die Addition, Skalarmultiplikation und Ringmultiplikaktion erhält. Aus Sicht der Algebra würde man sagen, dass ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}} \colon \mathrm{End}(V) \to \mathrm{M}_{n,n}(K)$ ein *Isomorphismus von Algebren über* K ist.

Nun können wir den Satz von Cayley-Hamilton für Endomorphismen herleiten.

Beweis von Satz 5.3.39. Es sei $\varphi \in \mathrm{End}(V)$. Wir nehmen eine geordnete Basis $\mathcal{B}$ von V. Es sei $A = {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)$ die zugehörige Matrixdarstellung von φ. Nach Definition 5.3.17 gilt $\chi_A = \chi_\varphi$. Aus dem Satz von Cayley-Hamilton 5.3.41 und Lemma 5.3.45 schließen wir

$$ \begin{aligned} 0 &= \chi_A(A) && \text{(Cayley-Hamilton)} \\ &= \chi_\varphi({}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)) && (\chi_A = \chi_\varphi) \\ &= {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\chi_\varphi(\varphi)) && \text{(Lemma 5.3.45).} \end{aligned} $$

Da die Abbildung ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}\colon \mathrm{End}(V) \to \mathrm{M}_{n,n}(K)$ ein Isomorphismus ist (siehe [MG, 9.1.9]), wird nur die Nullabbildung auf die Nullmatrix abgebildet, d.h., wir erhalten $\chi_\varphi(\varphi) = 0$. □

5.3.47 **Bemerkung**. Man kann auch umgekehrt den Satz von Cayley-Hamilton für Matrizen aus dem Satz für Endomorphismen ableiten. Dazu wendet man den Satz von Cayley-Hamilton auf den Endomorphismus $f_A\colon K^n \to K^n$ an. Aus Beispiel 5.3.26 schließen wir

$$0 = \chi_A(f_A) = f_{\chi_A(A)}$$

und damit auch $\chi_A(A) = 0$.

5.3.48 **Beispiel**. Das charakteristische Polynom und das Minimalpolynom sind im Allgemeinen nicht gleich. Betrachten wir zum Beispiel die Matrix

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{R})$$

aus Aufgabe 5.3.36. Dort haben wird das Minimalpolynom bestimmt: $\mu_A = X^2 - X$. Das charakteristische Polynom von A ist aber

$$\begin{aligned} \chi_A = \det(XI_3 - A) &= \det \begin{pmatrix} X & 0 & 0 \\ 0 & X-1 & 0 \\ 0 & 0 & X-1 \end{pmatrix} \\ &= X(X-1)^2 = (X^2 - X)(X-1). \end{aligned}$$

5.4. Stochastische Matrizen und Anwendungen

Wahrscheinlich hat kein anderes Konzept der Linearen Algebra so viele verschiedene Anwendungen wie Eigenwerte und Eigenvektoren. Die Anwendungsgebiete sind vielfältig und man trifft die Eigenwerttheorie unter anderem in der Biologie, der Physik und der Informatik.

In Anwendungen benötigt man oft zusätzliche Resultate aus der Eigenwerttheorie. Das liegt daran, dass man es in der Praxis meistens mit sehr speziellen, reellen Matrizen zu tun hat und man durch zusätzliche Eigenschaften auch mehr über die Eigenwerte dieser Matrizen sagen kann. Vor solchen speziellen Ergebnissen der Eigenwerttheorie müssen Sie keine Angst haben. Meist handelt es sich um Ergebnisse, die Sie mit den Methoden aus diesem Studienbrief gut verstehen können.

In diesem Abschnitt werden wir die stochastischen Matrizen untersuchen, die in vielen Anwendungen eine Rolle spielen. Wir definieren zunächst stochastische Matrizen und besprechen den Satz von Perron[5]. Dann lernen wir eine wichtige Anwendung kennen: das PageRank-Verfahren der Suchmaschine Google.

I. Stochastische Matrizen

5.4.1 **Definition** Sei $A = (a_{ij})_{i,j} \in \mathrm{M}_{n,n}(\mathbb{R})$. Wir nennen A *nicht-negativ* (bzw. *positiv*), wenn $a_{ij} \geq 0$ (bzw. $a_{ij} > 0$) für alle i, j gilt.

Eine nicht-negative Matrix $A \in \mathrm{M}_{n,n}(\mathbb{R})$, sodass die Summe über jede Spalte den Wert 1 ergibt, nennt man *stochastisch*. Das heißt, es gilt

$$\sum_{i=1}^{n} a_{ij} = 1$$

für jeden Spaltenindex $j \in \{1, 2, \ldots, n\}$.

5.4.2 **Beispiel**. Die Matrix

$$A = \begin{pmatrix} 0.2 & 0.1 & 0 \\ 0.8 & 0.7 & 1 \\ 0 & 0.2 & 0 \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{R})$$

ist eine stochastische Matrix. Sie ist nicht-negativ, aber nicht positiv, denn drei der Einträge sind 0.

[5] Oskar Perron: deutscher Mathematiker, 1880–1975

5.4.3 **Aufgabe**. Es seien $A, B \in \mathrm{M}_{n,n}(\mathbb{R})$ zwei stochastische Matrizen. Sei $p \in \mathbb{R}$ mit $0 \leq p \leq 1$. Zeigen Sie, dass auch $pA + (1-p)B$ eine stochastische Matrix ist. **L**

5.4.4 **Aufgabe**. Seien $A, B \in \mathrm{M}_{n,n}(\mathbb{R})$ zwei stochastische Matrizen. Zeigen Sie, dass auch AB eine stochastische Matrix ist. **L**

5.4.5 **Lemma.** *Jede stochastische Matrix hat den Eigenwert* 1. *Es gibt einen Eigenvektor* $w \in \mathrm{Eig}_A(1)$ *der nur nicht-negative Einträge hat, d.h.,* $w_i \geq 0$ *für alle* i.

Beweis. Sei A eine stochastische Matrix. Da A und A^T dieselben Eigenwerte haben (siehe Satz 5.3.13), genügt es zu zeigen, dass $B = A^T$ den Eigenwert 1 besitzt. Sei $\mathbb{1} \in \mathbb{R}^n$ der Vektor, dessen Einträge alle gleich 1 sind, d.h. $\mathbb{1}_i = 1$ für alle i. Dann gilt

$$(B\mathbb{1})_i = \sum_{j=1}^n b_{ij}\mathbb{1}_i = \sum_{j=1}^n b_{ij} = \underbrace{\sum_{j=1}^n a_{ji}}_{\text{Summe über Spalte } i} = 1,$$

weil die Matrix A stochastisch ist. Es gilt also $A^T\mathbb{1} = \mathbb{1}$ und $\mathbb{1}$ ist damit ein Eigenvektor zum Eigenwert 1 von A^T.

Sei nun $v \in \mathbb{R}^n$ ein beliebiger Eigenvektor von A zum Eigenwert 1 - insbesondere ist $v \neq 0$. Wir betrachten den Vektor $w \in \mathbb{R}^n$ mit den Einträgen $w_i = |v_i|$ und zeigen, dass auch w ein Eigenvektor zum Eigenwert 1 ist. Für alle $i \in \{1, 2, \ldots, n\}$ gilt

$$\begin{aligned} (Aw)_i &= \sum_{j=1}^n a_{ij}|v_j| \\ &\geq \left|\sum_{j=1}^n a_{ij}v_j\right| && \text{(Dreiecksungl. [MG, 12.2.23])} \\ &= |(Av)_i| = |v_i| = w_i. \end{aligned}$$

Wir zeigen, dass alle diese Ungleichungen sogar Gleichungen sind. Um das zu sehen, summieren wir alle Einträge von Aw und erhalten

$$\sum_{i=1}^n (Aw)_i = \sum_{i=1}^n \sum_{j=1}^n a_{ij}w_j = \sum_{j=1}^n \underbrace{\left(\sum_{i=1}^n a_{ij}\right)}_{=1} w_j = \sum_{j=1}^n w_j,$$

da A stochastisch ist. Damit ist $(Aw)_i > w_i$ unmöglich und wir schließen $Aw = w$, d.h., w ist ein Eigenvektor zum Eigenwert 1. $\square$

5.4.6 **Satz von Perron.** *Sei $A \in \mathrm{M}_{n,n}(\mathbb{R})$ eine stochastische Matrix. Wir nehmen an, dass A^m für ein $m \in \mathbb{N}$ positiv ist. Dann hat A den Eigenwert 1 mit geometrischer Vielfachheit $\gamma(1) = 1$. Es gibt einen Eigenvektor $w \in \mathrm{Eig}_A(1)$ der nur positive Einträge hat.*

Beweis. Ist v ein Eigenvektor zum Eigenwert 1 von A, dann gilt

$$A^m v = A^{m-1} v = \cdots = Av = v,$$

d.h., $\mathrm{Eig}_A(1) \subseteq \mathrm{Eig}_{A^m}(1)$. Also gilt $\gamma_A(1) \leq \gamma_{A^m}(1)$. Aus Lemma 5.4.5 ist $\gamma_A(1) \geq 1$ bekannt, daher genügt es, die Aussage für A^m zu beweisen. Da A^m nach Aufgabe 5.4.4 ebenfalls stochastisch ist, können wir nun annehmen, dass A positiv ist.

Da A und A^T dieselben Eigenwerte mit jeweils gleichen geometrischen Vielfachheiten besitzen (siehe Satz 5.3.13), zeigen wir, dass A^T den Eigenwert 1 mit der geometrischen Vielfachheit 1 besitzt. Wir wissen aus dem Beweis von Lemma 5.4.5, dass der Vektor $\mathbb{1}$ – der Vektor, dessen Einträge alle gleich 1 sind – ein Eigenvektor von A^T zum Eigenwert 1 ist. Sei $v \in \mathbb{R}^n$ ein Eigenvektor von A^T zum Eigenwert 1. Sei $i \in \{1, 2, \ldots, n\}$, sodass $v_i \in \mathbb{R}$ ein Eintrag von v mit dem größten Betrag $|v_i|$ ist. Wir können annehmen, dass $v_i > 0$ ist (sonst ersetzen wir v durch $-v$). Dann gilt

$$\begin{aligned}
v_i = (A^T v)_i &= \sum_{j=1}^{n} a_{ji} v_j \\
&\overset{\star}{\leq} \Big(\sum_{j=1}^{n} a_{ji}\Big) v_i && (v_i \text{ maximal, } A \text{ positiv}) \\
&= v_i && (A \text{ stochastisch}).
\end{aligned}$$

Weil für alle Einträge von A die Ungleichung $a_{ji} > 0$ gilt, ist die mit $\star$ markierte Ungleichung eine echte Ungleichung, falls $v_j < v_i$ für ein $j \neq i$ gilt. Es muss also $v_i = v_j$ für alle j sein, d.h., v ist ein Vielfaches von $\mathbb{1}$. Damit gilt $\mathrm{Eig}_{A^T}(1) = \langle \mathbb{1} \rangle$ und die geometrische Vielfachheit ist 1.

Aus Lemma 5.4.5 wissen wir, dass A einen Eigenvektor w mit ausschließlich nicht-negativen Einträgen hat. Da $w \neq 0$ gilt, gibt es mindestens einen positiven Eintrag. Sagen wir $w_k > 0$. Es gilt $w = Aw$ und damit $w_i = \sum_{j=1}^{n} a_{ij} w_j \geq a_{ik} w_k > 0$. Die Einträge von w sind also alle positiv. $\square$

5.4.7 **Aufgabe**. Zeigen Sie, dass die stochastische Matrix **L**

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

die Voraussetzung des Satzes von Perron erfüllt und berechnen Sie $\mathrm{Eig}_A(1)$.

5.4.8 **Näherungsweise Berechnung von Eigenvektoren**. Ist $A \in \mathrm{M}_{n,n}(\mathbb{R})$ eine stochastische Matrix wie im Satz von Perron, dann kann man einen Eigenvektor zum Eigenwert 1 relativ einfach näherungsweise berechnen. Ist $v_0 \in \mathbb{R}^n$ ein beliebiger Vektor mit ausschließlich positiven Einträgen, dann *konvergieren* die Vektoren

$$v_k = A^k v_0$$

im $\mathbb{R}^n$ gegen einen positiven Eigenvektor zum Eigenwert 1. Das heißt, für große Werte von k unterscheiden sich die Einträge von v_k nur durch sehr kleine reelle Zahlen von den Einträgen eines Eigenvektors. Diese Aussage werden wir hier nicht beweisen, weil wir dazu einige Begriffe der Analysis bräuchten.

5.4.9 **Stochastische Matrizen und Markov-Ketten**. Stochastische Matrizen treten in vielen Anwendungen als Übergangsmatrizen von sogenannten *Markov-Ketten* auf. Dabei hat man ein dynamisches System mit endlich vielen Zuständen; sagen wir $\{1, 2, \ldots, n\}$. Befindet sich das System im Zustand j, dann springt es im nächsten Schritt zufällig in einen anderen Zustand. Die Wahrscheinlichkeit, dass das System von Zustand j nach Zustand i springt sei p_{ij}. Die Matrix $p = (p_{ij})$ ist dann eine stochastische Matrix: die *Übergangsmatrix*.

5.4.10 **Beispiel: CarSharing mit festen Stationen**. Ein CarSharing Anbieter möchte in einer Stadt ein Leihsystem aufbauen, bei dem die Leihautos an festgelegten Orten in der Stadt abgeholt und wieder abgestellt werden können. Wieviele Autos soll der CarSharing Anbieter an den Stationen platzieren? Wieviel Platz wird benötigt? Kann der Anbieter es vermeiden, dass sich die Autos an einer Station häufen und somit Angestellte die Autos zwischen den Stationen transferieren müssen?

Nehmen wir an, der CarSharing Anbieter plant n Stationen in der Stadt. Er ermittelt (oder schätzt) den durchschnittlichen Anteil q_{ij} der Autos die morgens bei Station j stehen aber am Abend bei Station i abgestellt sind. Wir nehmen an, dass dieser Anteil unabhängig von der Anzahl der Autos ist, die morgens bei Station j stehen. Beispielsweise heißt das: Die Fahrt vom Marktplatz zum Bahnhof ist doppelt so beliebt, wie die Fahrt vom Marktplatz zur Universität unabhängig von der Anzahl der Fahrzeuge, die am Marktplatz zur Verfügung stehen. Die Anzahl der Fahrzeuge ist also klein im Vergleich zum Bedarf.

Sei Q die Matrix mit den Einträgen $(q_{ij})_{i,j}$. Betrachten wir nun einen durchschnittlichen Tag. Dazu sei w_j die Anzahl der Autos, die am Morgen an der Station j stehen. Wieviele Autos w_i' stehen Abends an der Station i? Es gilt

$$w_i' = \sum_{j=1}^{n} q_{ij} w_j.$$

Die Verteilung am Ende des Tages wird also durch den Vektor $w' = Qw$ beschrieben.

Wie könnte die CarSharing Firma die Autos nun auf die Stationen verteilen, um zu vermeiden, dass sich die Verteilung der Autos nach einem durchschnittlichen Tag ändert? Dazu muss die Gleichung

$$w = Qw$$

erfüllt sein. Der CarSharing Anbieter sucht also einen Eigenvektor zum Eigenwert 1 der Matrix Q. Gibt es einen solchen Eigenvektor? Ja, das folgt aus Lemma 5.4.5, denn die Matrix Q ist stochastisch!

Wir nehmen an, es gibt in der Stadt drei CarSharing Stationen: (1) am Marktplatz, (2) am Bahnhof und (3) an der Universität. Die Matrix Q sei gegeben durch

$$Q = \begin{pmatrix} \frac{1}{4} & \frac{3}{4} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}.$$

Von den Autos, die Morgens am Bahnhof stehen, parken also 3/4 abends am Marktplatz und 1/4 an der Universität.

Berechnet man den Eigenraum zum Eigenwert 1, dann erhält man

$$\mathrm{Eig}_Q(1) = \langle \begin{pmatrix} 11 \\ 7 \\ 6 \end{pmatrix} \rangle.$$

Der CarSharing Anbieter könnte also 24 Autos einsetzen und davon 11 Autos am Marktplatz abstellen, 7 am Bahnhof und 6 an der Universität. Er hat dadurch auch eine grobe Schätzung für die benötigte Anzahl von Parkplätzen.

Die Matrix Q aus diesem Beispiel erfüllt die Voraussetzungen des Satzes von Perron, denn

$$Q^2 = \begin{pmatrix} \frac{9}{16} & \frac{5}{16} & \frac{7}{16} \\ \frac{3}{16} & \frac{7}{16} & \frac{5}{16} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$$

ist positiv. Der CarSharing Anbieter muss sich also in diesem Fall keine Sorgen über die Verteilung machen. Auch wenn er die Autos anders verteilt (z.B. 8 Autos an jeder Station), wird sich die Verteilung der Autos mit der Zeit dem Eigenvektor $\begin{pmatrix} 11 \\ 7 \\ 6 \end{pmatrix}$ annähern; vgl. 5.4.8. Das trifft auch dann zu, wenn sich der Anbieter beim Schätzen der Parameter q_{ij} leicht vertan hat, denn die Voraussetzungen des Satzes von Perron sind resistent gegen kleine Störungen.

II. Google PageRank

Der schnelle Aufstieg der Suchmaschine Google um die Jahrtausendwende basiert auf dem sogenannten PageRank-Verfahren. In den 90er Jahren haben Suchmaschinen die Relevanz einer Webseite bezüglich einer Suchanfrage anhand einer Textanalyse der einzelnen Webseiten bewertet. Vereinfacht gesagt, bestimmte die Häufigkeit des Auftretens des Suchbegriffes die Reihung der angezeigten Suchergebnisse. Dies führte zu unbefriedigenden Suchergebnissen, die Webseiten mit viel Werbung aber ohne sinnvollen Inhalt enthielten. Die Google-Gründer Larry Page und Sergey Brin hatten die clevere Idee auch die „Wichtigkeit“ (oder „Beliebtheit“) einer Webseite bei der Reihung der Suchergebnisse einzubeziehen. Aber wie erkennt man eine „wichtige“ Webseite? Der Ansatz von Page und Brin beruht auf folgender Idee: Eine Webseite ist „wichtig“, wenn es viele andere Webseiten gibt, die diese Seite verlinken! Das PageRank-Verfahren der Google-Gründer versucht die „Wichtigkeit“ einer Webseite mithilfe der Anzahl und „Wichtigkeit“ der anderen Webseiten zu bemessen, die auf diese Seite verweisen. Die „Wichtigkeit“ einer Seite wird in diesem Ansatz durch den sogenannten *PageRank* quantifiziert. Jeder Internetseite wird also ein PageRank zwischen 0 und 1 zugeordnet. Eine Webseite mit einem hohen PageRank ist wichtig, eine Webseite mit einem niedrigen PageRank ist eher unwichtig.

5.4.11 **Das Zufallssurfer-Modell**. Um den PageRank zu definieren, entwerfen Page und Brin das *Zufallssurfer-Modell* (engl.: *random surfer*). Ein *Zufallssurfer* nutzt das Internet auf recht einfache Weise: Er folgt von jeder Webseite zufällig einem Link um zur nächsten Webseite zu gelangen. Falls er auf einer Webseite ankommt, von der keine Links ausgehen, springt er zufällig zu einer anderen Webseite.

Wenn nun sehr viele Zufallssurfer für lange Zeit durch das Internet surfen, sollte sich (nach und nach) eine stabile Verteilung der Surfer auf alle Websites einstellen in der wichtige Websites viele Besucher und unwichtige Websites wenige Besucher haben. Der PageRank R_j einer Webseite j sollte also proportional zur Anzahl der

Zufallssurfer sein, die in dieser stabilen Verteilung auf der Webseite j zu finden sind.

Es sei n die Anzahl aller Internetseiten. Wir stellen uns vor, dass wir die Webseiten von 1 bis n nummeriert haben. Ist j eine Internetseite, dann bezeichnen wir mit L_j die Anzahl der Links, die auf Seite j zu finden sind und mit $L_j(i)$ die Anzahl dieser Links mit denen man auf Seite i springen kann. Ist $L_j \neq 0$, dann definieren wir

$$s_{ij} = \frac{L_j(i)}{L_j}.$$

Das heißt, s_{ij} ist der Anteil der Links von Webseite j, mit denen man zur Seite i gelangt. Falls $L_j = 0$ ist – wenn also von Seite j kein Link ausgeht – definieren wir

$$s_{ij} = \frac{1}{n}$$

für alle i. Die Matrix $S = (s_{ij})_{i,j} \in \mathrm{M}_{n,n}(\mathbb{R})$ nennt man die *Übergangsmatrix*. Der Eintrag s_{ij} ist genau die Wahrscheinlichkeit, dass ein Zufallssurfer von Seite j auf Seite i wechselt.

Der PageRank Vektor $R \in \mathbb{R}^n$ ist der Spaltenvektor, dessen Eintrag R_j der PageRank der Seite j ist. Der PageRank Vektor enthält eine „stabile Verteilung von Zufallssurfern“ genau dann, wenn

$$SR = R$$

gilt. In anderen Worten: Der PageRank Vektor ist ein Eigenvektor der Übergangsmatrix S zum Eigenwert 1.

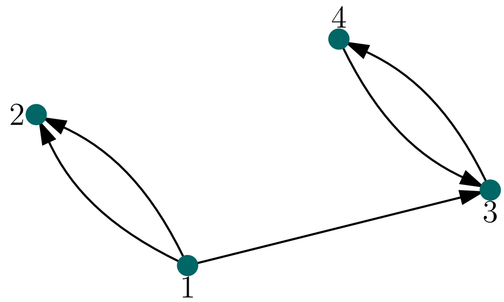
Aber funktioniert das überhaupt? Hat die Übergangsmatrix S den Eigenwert 1? Die Antwort ist „ja“, weil die Matrix S eine *stochastische Matrix* ist; siehe Lemma 5.4.5.

Wir überlegen uns kurz, warum die Übergangsmatrix S stochastisch ist. Ist j eine Webseite ohne ausgehende Links, dann ist die Spalte j von S genau $\frac{1}{n}\mathbb{1}$ und offensichtlich ist $\sum_{i=1}^{n} \frac{1}{n} = 1$. Nehmen wir nun an, dass $L_j \neq 0$ ist. Dann ergibt die Spaltensumme

$$\sum_{i=1}^{n} \frac{L_j(i)}{L_j} = \frac{1}{L_j} \sum_{i=1}^{n} L_j(i) = \frac{L_j}{L_j} = 1.$$

Die Existenz eines Eigenvektors von S zum Eigenwert 1 ist also gesichert. Allerdings können andere Probleme auftreten, wie wir im nächsten Beispiel sehen werden.

5.4.12 Beispiel. Wir betrachten ein Beispiel mit $n = 4$ Webseiten. Wir zeichnen einen Pfeil um anzudeuten, dass es einen Link von Seite j nach Seite i gibt.



Von Seite 1 gehen $L_1 = 3$ Links ab. Zwei verweisen auf Seite 2 und einer auf Seite 3. Es gilt also

$$L_1(1) = 0, \quad L_1(2) = 2, \quad L_1(3) = 1, \quad L_1(4) = 0.$$

Von Seite 2 gehen keine Links ab. Die Übergangsmatrix S ist gegeben durch

$$S = \begin{pmatrix} 0 & \frac{1}{4} & 0 & 0 \\ \frac{2}{3} & \frac{1}{4} & 0 & 0 \\ \frac{1}{3} & \frac{1}{4} & 0 & 1 \\ 0 & \frac{1}{4} & 1 & 0 \end{pmatrix}.$$

Die Übergangmatrix S besitzt wirklich den Eigenwert 1 mit der geometrischen Vielfachheit $\gamma(1) = 1$. Der Eigenraum zum Eigenwert 1 wird vom Vektor

$$R = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

aufgespannt.

Dieses Ergebnis scheint etwas zu extrem zu sein um eine nützliche Bewertung zu ergeben. Demnach wären die Webseiten 3 und 4 extrem wichtig und die Webseiten 1 und 2 absolut unbedeutend.

Noch schlechter wird das Ergebnis, wenn wir zwei weitere Webseiten hinzufügen, die sich nur gegenseitig verlinken. Die neue Übergangsmatrix

$$S' = \begin{pmatrix} 0 & \frac{1}{6} & 0 & 0 & 0 & 0 \\ \frac{2}{3} & \frac{1}{6} & 0 & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{6} & 0 & 1 & 0 & 0 \\ 0 & \frac{1}{6} & 1 & 0 & 0 & 0 \\ 0 & \frac{1}{6} & 0 & 0 & 0 & 1 \\ 0 & \frac{1}{6} & 0 & 0 & 1 & 0 \end{pmatrix}$$

hat den Eigenwert 1 mit der geometrischen Vielfachheit $\gamma(1) = 2$. Beispielsweise sind

$$R_1 = \begin{pmatrix} 0\\0\\1\\1\\0\\0 \end{pmatrix} \quad \text{und} \quad R_2 = \begin{pmatrix} 0\\0\\0\\0\\1\\1 \end{pmatrix}$$

und jede Linearkombination dieser Vektoren Eigenvektoren zum Eigenwert 1. Welcher dieser Vektoren ist nun der gesuchte PageRank?

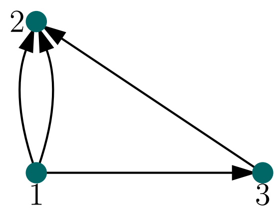
5.4.13 **Das Zufallssurfer-Modell mit Dämpfung**. Um diese Probleme zu beheben, variieren Brin und Page das Zufallssurfer-Modell mit einem Trick. Sie nehmen an, dass der Zufallssurfer nur mit großer Wahrscheinlichkeit den Links auf der Webseite folgt, aber gelegentlich einfach eine zufällige Internetadresse in seinen Browser eingibt. Sei $p \in [0, 1]$ die Wahrscheinlichkeit, dass der Zufallssurfer den Links folgt. Ist beispielsweise $p = 0.9$, dann folgt der Zufallssurfer in 90% der Fälle den Links auf der Webseite. In 10% der Fälle (Wahrscheinlichkeit $1 - p$) wechselt er auf eine ganz andere Webseite. Dadurch verändert sich die Übergangsmatrix und man erhält die sogenannte *Google-Matrix*

$$G = pS + (1-p)\frac{1}{n}E.$$

Dabei ist E die Matrix, in der jeder Eintrag gleich 1 ist. Weil S und $\frac{1}{n}E$ beide stochastische Matrizen sind, ist auch G eine stochastische Matrix; siehe Aufgabe 5.4.3. Man nennt p auch den *Dämpfungsfaktor*.

Was ist der Vorteil der Google-Matrix G gegenüber der Übergangsmatrix S? Die Matrix G ist eine *positive* stochastische Matrix, was uns erlaubt den Satz von Perron (5.4.6) anzuwenden.

5.4.14 **Aufgabe**. Wir betrachten ein Beispiel mit $n = 3$ Webseiten. **L**



Bestimmen Sie die Google-Matrix zum Dämpfungsfaktor $p = \frac{2}{3}$.

5.4.15 **Bemerkung: Normalisierung des PageRank-Vektors**. Die Google-Matrix G hat also nach dem Satz von Perron einen ein-dimensionalen Eigenraum zum Eigenwert 1 und einen Eigenvektor v mit ausschließlich positiven Einträgen. Damit ist der PageRank-Vektor bis auf Streckung mit $r > 0$ bestimmt. Um den PageRank eindeutig festzulegen, einigt man sich auf eine Normalisierung.

Hier legen wir fest: *Der PageRank R einer Google-Matrix G ist der eindeutige Eigenvektor zum Eigenwert 1, sodass die Summe aller Einträge 1 ergibt.*

5.4.16 **Beispiel**. Wir betrachten nochmal die Situation mit vier Webseiten aus Beispiel 5.4.12. Wir wählen den Dämpfungsfaktor $p = \frac{3}{4}$. Damit erhalten wir die Google-Matrix

$$G = \frac{3}{4}S + \frac{1}{4} \cdot \frac{1}{4}E = \frac{1}{16}\begin{pmatrix} 1 & 4 & 1 & 1 \\ 9 & 4 & 1 & 1 \\ 5 & 4 & 1 & 13 \\ 1 & 4 & 13 & 1 \end{pmatrix}.$$

Mit einer kurzen Rechnung bestimmt man den Eigenraum zum Eigenwert 1:

$$\mathrm{Eig}_G(1) = \langle \begin{pmatrix} 14 \\ 21 \\ 64 \\ 62 \end{pmatrix} \rangle.$$

Den PageRank-Vektor erhalten wir durch Normalisierung, d.h., wir teilen durch die Summe N aller Einträge $N = 14 + 21 + 64 + 62 = 161$ und erhalten

$$R = \frac{1}{161}\begin{pmatrix} 14 \\ 21 \\ 64 \\ 62 \end{pmatrix} \approx \begin{pmatrix} 0.087 \\ 0.130 \\ 0.398 \\ 0.385 \end{pmatrix}.$$

Mit diesem Ergebnis würden wir also die Webseite 3 als „Wichtigste" und Webseite 1 „Unwichtigste" einschätzen.

5.4.17 **PageRank Berechnung in der Praxis**. Aktuell indiziert Google wahrscheinlich mehr als 25 Milliarden Webseiten. Das heißt, die zugehörige Google-Matrix ist wirklich unvorstellbar groß. Es ist auch mit modernen Computern hoffnungslos, den Eigenraum durch Lösen eines Gleichungssystems mit 25 Milliarden Unbekannten zu bestimmen. Aber wie kann man den PageRank dann berechnen?

Die Lösung kommt wieder aus der Bemerkung 5.4.8: Man berechnet den PageRank nur näherungsweise! Dazu multipliziert man die Google-Matrix G wiederholt mit einem positiven Vektor v_0: $v_k = G^k v_0$. Ist die Summe aller Einträge von v_0 gleich 1, dann konvergiert diese Folge gegen den PageRank-Vektor R. Diese Berechnung ist iterativ möglich, denn

$$v_k = Gv_{k-1}.$$

Dafür benötigt man nur Matrix-Vektor-Multiplikationen, die relativ einfach umgesetzt werden können, weil die Google-Matrix G sehr dünn besetzt (engl.: „sparse“) ist. Durch die Wahl des Dämpfungsfaktors p kann man außerdem die Konvergenzgeschwindigkeit beeinflussen.

Von dieser Idee ist es natürlich noch ein weiter Weg bis zu einer funktionierenden Suchmaschine. Mehr Details zu Suchmaschinen und Googles PageRank-Verfahren findet man in [LM11].

5.5. Lösungen der Aufgaben in Lektion 5

L5.1.10 Lösung. Wir untersuchen den Endomorphismus f_A von $\mathbb{R}^2$ mit $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

Der ganze Raum $\mathbb{R}^2$ ist f_A-invariant und auch der Nullraum $\{0\}$ ist f_A-invariant. Alle anderen Unterräume sind ein-dimensional und werden von einem Vektor $w \neq 0$ aufgespannt. Aus Lemma 5.1.9 wissen wir, dass der Unterraum $U = \langle w \rangle$ genau dann f_A-invariant ist, wenn $f_A(w) \in \langle w \rangle$ gilt, d.h., wenn $f_A(w) = \lambda w$ für ein Skalar $\lambda \in \mathbb{R}$ gilt. Wir machen einen Ansatz $w = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$ und suchen Lösungen $w_1, w_2, \lambda \in \mathbb{R}$ der Gleichung $Aw = \lambda w$. Das heißt, es sollen

$$w_1 = \lambda w_2 \quad \text{und} \quad w_2 = \lambda w_1$$

erfüllt sein. Mindestens ein Eintrag von w, sagen wir w_1, ist nicht Null. Dann erhalten wir $w_1 = \lambda w_2 = \lambda^2 w_1$ und durch Kürzen $\lambda \in \{\pm 1\}$. Dieselbe Schlussfolgerung erhält man, wenn $w_2 \neq 0$ ist.

Mit $\lambda = 1$ finden wir den f_A-invarianten Unterraum $U = \langle \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rangle$. Mit $\lambda = -1$ finden wir $U = \langle \begin{pmatrix} 1 \\ -1 \end{pmatrix} \rangle$. Insgesamt gibt es also vier verschiedene f_A-invariante Unterräume.

L5.1.12 Lösung. Gegeben sind φ-invariante Unterräume $U_1, \ldots, U_k$ von V. (a) Aus den Mathematischen Grundlagen ist bekannt, dass der Schnitt von Unterräumen $\bigcap_{i=1}^k U_i$ wieder ein Unterraum ist; vgl. [MG, 6.2.7]. Wir müssen also nur zeigen, dass dieser Unterraum φ-invariant ist. Sei dazu $u \in \bigcap_{i=1}^k U_i$. Sei $i \in \{1, 2, \ldots, k\}$. Da U_i ein φ-invarianter Unterraum ist, gilt $\varphi(u) \in U_i$. Da i beliebig war, folgt aus der Definition des Schnittes bereits $\varphi(u) \in \bigcap_{i=1}^k U_i$.

(b) Aus 2.3.1 wissen wir, dass $\sum_{i=1}^k U_i$ ein Unterraum ist. Sei nun $u \in \sum_{i=1}^k U_i$. Nach Definition der Summe finden wir Vektoren $u_i \in U_i$ für $i \in \{1, 2, \ldots, k\}$ mit $u = \sum_{i=1}^k u_i$. Wir erhalten

$$\varphi(u) = \varphi\Big(\sum_{i=1}^k u_i\Big) = \sum_{i=1}^k \varphi(u_i).$$

Da U_i invariant unter φ ist, gilt $\varphi(u_i) \in U_i$ und damit $\varphi(u) \in \sum_{i=1}^k U_i$. Also ist die Summe wieder ein φ-invarianter Unterraum.

L5.2.5 **Lösung.** Gegeben ist die Matrix

$$A = \begin{pmatrix} \frac{3}{4} & -\frac{3}{4} \\ -\frac{1}{4} & \frac{1}{4} \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{Q}).$$

Der Vektor v_1 ist der Nullvektor und per Definition kein Eigenvektor von A.

Der Vektor v_2 ist ein Eigenvektor von A zum Eigenwert $\lambda = 1$, denn es gilt

$$Av_2 = \begin{pmatrix} \frac{3}{4} & -\frac{3}{4} \\ -\frac{1}{4} & \frac{1}{4} \end{pmatrix} \begin{pmatrix} 3 \\ -1 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \end{pmatrix} = v_2.$$

Der Vektor v_3 ist ein Eigenvektor von A zum Eigenwert $\lambda = 0$, denn es gilt

$$Av_3 = \begin{pmatrix} \frac{3}{4} & -\frac{3}{4} \\ -\frac{1}{4} & \frac{1}{4} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0 \cdot v_3.$$

Der Vektor v_4 ist kein Eigenvektor von A, denn

$$Av_4 = \begin{pmatrix} \frac{3}{4} & -\frac{3}{4} \\ -\frac{1}{4} & \frac{1}{4} \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{3}{4} \\ -\frac{1}{4} \end{pmatrix}$$

und dieser Vektor kann nicht in der Form λv_4 geschrieben werden.

L5.2.6 **Lösung.** Nach Voraussetzung gelten $v \neq 0$ und $\varphi(v) = \lambda v$. Damit ergibt sich die Aussage für $j = 1$. Wir zeigen die Aussage mit vollständiger Induktion für $j > 1$. Wir nehmen an, die Aussage sei bekannt für $j - 1$. Dann erhalten wir

$$\begin{aligned} \varphi^j(v) &= \varphi^{j-1}(\varphi(v)) \\ &= \varphi^{j-1}(\lambda v) && (v \text{ Eigenvektor von } \varphi) \\ &= \lambda \varphi^{j-1}(v) && (\varphi^{j-1} \text{ ist linear}) \\ &= \lambda \lambda^{j-1} v && (\text{Induktionsannahme}) \\ &= \lambda^j v \end{aligned}$$

Damit ist v ein Eigenvektor von φ^j zum Eigenwert λ^j.

L5.2.9 **Lösung.** Zunächst stellen wir fest, dass $\mathrm{Eig}_\varphi(\lambda)$ nicht leer ist. In der Tat, der Nullvektor erfüllt $\varphi(0) = 0 = \lambda 0$ und liegt damit in $\mathrm{Eig}_\varphi(\lambda)$. Sind $v, v' \in \mathrm{Eig}_\varphi(\lambda)$ und $a \in K$, dann gilt

$$\varphi(av + v') = a\varphi(v) + \varphi(v') = a\lambda v + \lambda v' = \lambda(av + v')$$

und damit liegt $av + v'$ in $\mathrm{Eig}_\varphi(\lambda)$.

L5.2.12 Lösung. Da wir noch keine bessere Methode kennen, müssen wir die drei Elemente $\lambda = \overline{0}, \overline{1}, \overline{2}$ von $\mathbb{F}_3$ durchgehen und $\mathrm{Eig}(A, \lambda)$ bestimmen.

$\mathrm{Eig}(A, \overline{0})$: Wir berechnen $\mathrm{Ker}(A)$. Es gilt

$$\mathrm{Ker}(A) = \mathrm{Ker}\begin{pmatrix} \overline{1} & \overline{2} \\ \overline{2} & \overline{1} \end{pmatrix} = \begin{pmatrix} \overline{1} & \overline{2} \\ \overline{0} & \overline{0} \end{pmatrix} = \langle \begin{pmatrix} \overline{1} \\ \overline{1} \end{pmatrix} \rangle.$$

Der Raum $\mathrm{Eig}(A, \overline{0})$ ist also ein-dimensional. Damit ist $\overline{0}$ ein Eigenwert von A mit der geometrischen Vielfachheit $\gamma(\overline{0}) = 1$.

$\mathrm{Eig}(A, \overline{1})$: Wir berechnen $\mathrm{Ker}(A - I_2)$ und erhalten

$$\mathrm{Ker}(A) = \mathrm{Ker}\begin{pmatrix} \overline{0} & \overline{2} \\ \overline{2} & \overline{0} \end{pmatrix} = \{0\},$$

denn $A - I_2$ ist invertierbar. Damit ist $\mathrm{Eig}(A, \overline{1}) = \{0\}$ und $\overline{1}$ ist kein Eigenwert von A.

$\mathrm{Eig}(A, \overline{2})$: Wir berechnen $\mathrm{Ker}(A - 2I_2)$. Es gilt

$$\mathrm{Ker}(A) = \mathrm{Ker}\begin{pmatrix} \overline{2} & \overline{2} \\ \overline{2} & \overline{2} \end{pmatrix} = \begin{pmatrix} \overline{1} & \overline{1} \\ \overline{0} & \overline{0} \end{pmatrix} = \langle \begin{pmatrix} \overline{1} \\ \overline{2} \end{pmatrix} \rangle.$$

Der Raum $\mathrm{Eig}(A, \overline{2})$ ist also ein-dimensional und $\overline{2}$ damit ein Eigenwert von A mit der geometrischen Vielfachheit $\gamma(\overline{2}) = 1$.

L5.2.14 Lösung. Seien $\lambda_1, \dots, \lambda_r$ paarweise verschiedene Eigenwerte von $\varphi \in \mathrm{End}(V)$. Da $\mathrm{Eig}_\varphi(\lambda_i)$ ungleich $\{0\}$ ist und damit mindestens Dimension 1 hat, erhalten wir mit Lemma 5.2.13

$$n = \dim_K(V) \geq \sum_{i=1}^{r} \gamma(\lambda_i) \geq \sum_{i=1}^{r} 1 = r.$$

Also hat φ höchstens n paarweise verschiedene Eigenwerte.

L5.2.15 Lösung. Wir nehmen uns k verschiedene reelle Zahlen, sagen wir $\lambda_1 = 1, \lambda_2 = 2, \dots, \lambda_k = k$. Wir setzen $\lambda_j = k$ für alle $k < j \leq n$ und betrachten die Diagonalmatrix

$$A = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}.$$

Der i-te Standardbasisvektor e_i ist ein Eigenvektor von A zum Eigenwert λ_i, denn $Ae_i = \lambda_i e_i$. Wir wollen zeigen, dass $\lambda_1 = 1, \ldots, \lambda_k = k$ wirklich alle Eigenwerte von A sind. Für $i < k$ gilt

$$\mathrm{Eig}_A(\lambda_i) = \mathrm{Ker}(A - \lambda_i I_n) = \langle e_i \rangle$$

und damit $\gamma(\lambda_i) = 1$. Weiter ist

$$\mathrm{Eig}_A(k) = \mathrm{Ker}(A - kI_n) = \langle e_k, e_{k+1}, \ldots, e_n \rangle.$$

und damit $\gamma(\lambda_i) = n - k + 1$. Insgesamt erhalten wir

$$\sum_{i=1}^{k} \gamma(\lambda_i) = k - 1 + n - k + 1 = n.$$

und mit der Ungleichung aus Lemma 5.2.13 folgt, dass es keine weiteren Eigenwerte geben kann.

L5.2.20 **Lösung.** Die Matrix $A \in \mathrm{M}_{n,n}(K)$ ist genau dann diagonalisierbar, wenn der Endomorphismus $f_A \colon K^n \to K^n$ mit $x \mapsto Ax$ diagonalisierbar ist; das folgt aus Lemma 5.2.18, denn A ist die Matrixdarstellung von f_A zur Standardbasis. Per Definition sind die Eigenvektoren von f_A genau die Eigenvektoren von A und auch die geometrischen Vielfachheiten stimmen überein (siehe 5.2.9 (d)), also folgt das Korollar nun direkt aus Satz 5.2.19.

L5.2.21 **Lösung.** Gegeben ist die Matrix $A = \begin{pmatrix} \overline{1} & \overline{1} \\ \overline{0} & \overline{0} \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{F}_2)$. Im Körper $\mathbb{F}_2$ gibt es nur zwei mögliche Eigenwerte: $\overline{0}$ und $\overline{1}$. Wir prüfen, ob dies Eigenwerte sind und berechnen die geometrischen Vielfachheiten. Es gilt

$$\mathrm{Eig}_A(\overline{0}) = \mathrm{Ker}(A) = \langle \begin{pmatrix} \overline{1} \\ \overline{1} \end{pmatrix} \rangle$$

und damit ist $\overline{0}$ ein Eigenwert mit geometrischer Vielfachheit $\gamma(\overline{0}) = 1$. Außerdem gilt

$$\mathrm{Eig}_A(\overline{1}) = \mathrm{Ker}(A - I_2) = \langle \begin{pmatrix} \overline{1} \\ \overline{0} \end{pmatrix} \rangle$$

Damit ist auch $\overline{1}$ ein Eigenwert mit geometrischer Vielfachheit $\gamma(\overline{1}) = 1$. Wegen $2 = \gamma(\overline{0}) + \gamma(\overline{1})$ ist A nach Korollar 5.2.20 diagonalisierbar.

L5.2.24 Lösung. Die Inverse der Matrix S ist
$$S^{-1} = \begin{pmatrix} 0 & -1 & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & -\frac{3}{2} \end{pmatrix}.$$

L5.2.25 Lösung. Wir betrachten wieder die Matrix
$$A = \begin{pmatrix} \overline{1} & \overline{2} \\ \overline{2} & \overline{1} \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{F}_3).$$

In der Lösung zu Aufgabe 5.2.12 haben wir gesehen, dass $\overline{0}$ und $\overline{2}$ die Eigenwerte von A sind; beide haben die geometrische Vielfachheit 1. Die Eigenräume sind
$$\mathrm{Eig}(A, \overline{0}) = \langle \begin{pmatrix} \overline{1} \\ \overline{1} \end{pmatrix} \rangle \quad \text{und} \quad \mathrm{Eig}(A, \overline{2}) = \langle \begin{pmatrix} \overline{1} \\ \overline{2} \end{pmatrix} \rangle.$$
Wir schreiben die gefundenen Eigenvektoren als Spalten in die Matrix S und erhalten
$$S = \begin{pmatrix} \overline{1} & \overline{1} \\ \overline{1} & \overline{2} \end{pmatrix}.$$
Es gilt
$$S^{-1}AS = \begin{pmatrix} \overline{0} & \overline{0} \\ \overline{0} & \overline{2} \end{pmatrix}.$$

L5.2.26 Lösung. Da $\mathbb{F}_2 = \{\overline{0}, \overline{1}\}$ nur zwei Elemente besitzt, hat A höchstens diese zwei Eigenwerte in $\mathbb{F}_2$. Wir berechnen $\mathrm{Eig}_A(\lambda)$ für alle $\lambda \in \mathbb{F}_2$.

$\mathrm{Eig}_A(\overline{0}) = \mathrm{Ker}(A)$: Es gilt
$$\mathrm{Ker}(A) = \mathrm{Ker} \begin{pmatrix} \overline{0} & \overline{0} & \overline{1} & \overline{1} \\ \overline{0} & \overline{1} & \overline{0} & \overline{0} \\ \overline{0} & \overline{1} & \overline{1} & \overline{1} \\ \overline{0} & \overline{1} & \overline{0} & \overline{0} \end{pmatrix} = \mathrm{Ker} \begin{pmatrix} \overline{0} & \overline{1} & \overline{0} & \overline{0} \\ \overline{0} & \overline{0} & \overline{1} & \overline{1} \\ \overline{0} & \overline{0} & \overline{0} & \overline{0} \\ \overline{0} & \overline{0} & \overline{0} & \overline{0} \end{pmatrix} = \langle \begin{pmatrix} \overline{1} \\ \overline{0} \\ \overline{0} \\ \overline{0} \end{pmatrix}, \begin{pmatrix} \overline{0} \\ \overline{0} \\ \overline{1} \\ \overline{1} \end{pmatrix} \rangle.$$

Es ist $\mathrm{Eig}_A(\overline{0}) \neq \{0\}$ und somit ist $\overline{0}$ ein Eigenwert von A: Die beiden gefundenen Eigenvektoren sind linear unabhängig, bilden also eine Basis von $\mathrm{Eig}_A(\overline{0})$. Daraus erhalten wir die geometrische Vielfachheit $\gamma(\overline{0}) = 2$.

$\mathrm{Eig}_A(\overline{1}) = \mathrm{Ker}(A - I_4)$: Es gilt
$$\mathrm{Ker}(A - I_4) = \mathrm{Ker} \begin{pmatrix} \overline{1} & \overline{0} & \overline{1} & \overline{1} \\ \overline{0} & \overline{0} & \overline{0} & \overline{0} \\ \overline{0} & \overline{1} & \overline{0} & \overline{1} \\ \overline{0} & \overline{1} & \overline{0} & \overline{1} \end{pmatrix} = \mathrm{Ker} \begin{pmatrix} \overline{1} & \overline{0} & \overline{1} & \overline{1} \\ \overline{0} & \overline{1} & \overline{0} & \overline{1} \\ \overline{0} & \overline{0} & \overline{0} & \overline{0} \\ \overline{0} & \overline{0} & \overline{0} & \overline{0} \end{pmatrix} = \langle \begin{pmatrix} \overline{1} \\ \overline{0} \\ \overline{1} \\ \overline{0} \end{pmatrix}, \begin{pmatrix} \overline{0} \\ \overline{1} \\ \overline{1} \\ \overline{1} \end{pmatrix} \rangle.$$

Also ist auch $\overline{1}$ ein Eigenwert von A mit geometrischer Vielfachheit $\gamma(\overline{1}) = 2$.

Es gilt $\gamma(\overline{0}) + \gamma(\overline{1}) = 4$ und damit ist die (4×4)-Matrix A diagonalisierbar. Wir schreiben die gefundenen Vektoren als Spalten in die Matrix S und erhalten

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

Dann gilt

$$S^{-1}AS = \begin{pmatrix} \overline{0} & \overline{0} & \overline{0} & \overline{0} \\ \overline{0} & \overline{0} & \overline{0} & \overline{0} \\ \overline{0} & \overline{0} & \overline{1} & \overline{0} \\ \overline{0} & \overline{0} & \overline{0} & \overline{1} \end{pmatrix}.$$

L5.3.4 Lösung. Gegeben ist die Matrix

$$A = \begin{pmatrix} 0 & -1 & -3 \\ 0 & 1 & 0 \\ 1 & 1 & 4 \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{R}).$$

Das charakteristische Polynom berechnen wir mit der Regel von Sarrus 4.2.6 und erhalten

$$\begin{aligned} \chi_A = \det(XI_3 - A) &= \det \begin{pmatrix} X & 1 & 3 \\ 0 & X-1 & 0 \\ -1 & -1 & X-4 \end{pmatrix} \\ &= X(X-1)(X-4) - (-1) \cdot 3 \cdot (X-1) \\ &= (X-1)(X(X-4)+3) = (X-1)(X^2-4X+3) \\ &= X^3 - 5X^2 + 7X - 3. \end{aligned}$$

L5.3.12 Lösung. Das charakteristische Polynom von

$$A = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix} \in \mathrm{M}_{n,n}(\mathbb{R})$$

können wir mithilfe von Satz 4.2.13 berechnen, denn $XI_4 - A$ ist eine obere Dreiecksmatrix. Wir erhalten:

$$\chi_A = \det \begin{pmatrix} X & -1 & 0 & -1 \\ 0 & X & -2 & -1 \\ 0 & 0 & X & -1 \\ 0 & 0 & 0 & X+1 \end{pmatrix} = X^3(X+1).$$

Die Eigenwerte sind damit $\lambda_1 = 0$ und $\lambda_2 = -1$. Die algebraischen Vielfachheiten sind $\alpha(0) = 3$ und $\alpha(-1) = 1$. Um die geometrischen Vielfachheiten zu bestimmen, müssen wir noch die Eigenräume berechnen.

$\mathrm{Eig}_A(0)$: Es gilt

$$\mathrm{Ker}(A) = \mathrm{Ker} \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \mathrm{Ker} \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & \frac{1}{2} \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \langle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \rangle$$

d.h., $\gamma(0) = 1$.

$\mathrm{Eig}_A(-1)$: Es gilt

$$\mathrm{Ker}(A + I_4) = \mathrm{Ker} \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \langle \begin{pmatrix} -2 \\ 1 \\ -1 \\ 1 \end{pmatrix} \rangle$$

d.h., $\gamma(-1) = 1$.

L5.3.14 Lösung. Sind A und B ähnlich, dann gibt es eine Matrix $S \in \mathrm{GL}_n(K)$ mit $A = S^{-1}BS$. Wegen $I_n = S^{-1}S$ gilt auch

$$S^{-1}(XI_n - B)S = XS^{-1}S - S^{-1}BS = XI_n - A,$$

d.h., auch $XI_n - B$ und $XI_n - A$ sind ähnlich. Aus Korollar 4.2.26 erhalten wir

$$\chi_A = \det(XI_n - A) = \det(XI_n - B) = \chi_B.$$

L5.3.16 Lösung. Zum Beispiel haben die Matrizen

$$A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \ B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{R})$$

beide das charakteristische Polynom X^2 (und Spur 0). Weil A die Nullmatrix ist, gilt aber für alle $S \in \mathrm{GL}_2(\mathbb{R})$ die Gleichung $S^{-1}AS = A$; d.h., A und B sind nicht ähnlich.

L5.3.21 Lösung. Es seien $a, b \in K$ und $\varphi, \psi \in \mathrm{End}(V)$. Wir prüfen die Gleichheit

$$(ab)(\varphi \circ \psi) = (a\varphi) \circ (b\psi),$$

durch Einsetzen von Vektoren. Es sei $v \in V$ beliebig. Dann gilt aufgrund der K-Linearität

$$(ab)(\varphi \circ \psi)(v) = ab\varphi(\psi(v)) = a\varphi(b\psi(v)) = (a\varphi)((b\psi)(v)) = (a\varphi) \circ (b\psi)(v).$$

Da v beliebig war, folgt die behauptete Gleichheit.

L5.3.24 Lösung. Wir berechnen

$$\begin{aligned} A^2 &= \begin{pmatrix} \overline{2} & \overline{4} \\ \overline{2} & \overline{3} \end{pmatrix} \begin{pmatrix} \overline{2} & \overline{4} \\ \overline{2} & \overline{3} \end{pmatrix} = \begin{pmatrix} \overline{2} & \overline{0} \\ \overline{0} & \overline{2} \end{pmatrix} \\ A^3 &= A^2 A = \begin{pmatrix} \overline{4} & \overline{3} \\ \overline{4} & \overline{1} \end{pmatrix} \end{aligned}$$

Damit erhalten wir

$$f(A) = A^3 + \overline{2}A + \overline{2}I_2 = \begin{pmatrix} \overline{4} & \overline{3} \\ \overline{4} & \overline{1} \end{pmatrix} + \begin{pmatrix} \overline{4} & \overline{3} \\ \overline{4} & \overline{1} \end{pmatrix} + \begin{pmatrix} \overline{2} & \overline{0} \\ \overline{0} & \overline{2} \end{pmatrix} = \begin{pmatrix} \overline{0} & \overline{1} \\ \overline{3} & \overline{4} \end{pmatrix}.$$

L5.3.28 Lösung. Es sei $U \subseteq V$ ein φ-invarianter Unterraum und es sei $g \in K[X]$. Wir argumentieren mit Induktion nach dem Grad von g. Ist $\mathrm{Grad}(g) \leq 0$, dann ist $g = aX^0$ mit $a \in K$ und $g(\varphi) = a\,\mathrm{id}_V$. Ist $u \in U$, dann gilt $g(\varphi)(u) = au \in U$, denn U ist ein Unterraum.

Für den Induktionsschritt nehmen wir $\mathrm{Grad}(g) \geq 1$ an und schreiben $g = g'X + a$ mit $a \in K$ und einem Polynom g' mit $\mathrm{Grad}(g') = \mathrm{Grad}(g) - 1$ (Wir klammern also X aus allen Monomen von Grad ≥ 1 aus). Sei $u \in U$. Mit Lemma 5.3.27 gilt dann

$$g(\varphi)(u) = (g'(\varphi) \circ \varphi + a\,\mathrm{id}_V)(u) = g'(\varphi)(\varphi(u)) + au.$$

Da U invariant unter φ ist, gilt $u' = \varphi(u) \in U$ und nach Induktionsvoraussetzung auch $g'(\varphi)(u') \in U$. Da U ein Unterraum ist, liegt auch die Linearkombination

$$g(\varphi)(u) = g'(\varphi)(u') + au$$

wieder in U. Das heißt, U ist $g(\varphi)$-invariant.

L5.3.30 Lösung. Zu (I1): Sind g, g' zwei Polynome, die φ annullieren, dann gilt wegen Lemma 5.3.27

$$(g + g')(\varphi) = g(\varphi) + g'(\varphi) = 0 + 0 = 0,$$

d.h., $g + g' \in \mathcal{I}_\varphi$.

Zu (I2): Seien $h \in K[X]$ und $g \in \mathcal{I}_\varphi$. Dann gilt mit Lemma 5.3.27

$$(h \cdot g)(\varphi) = h(\varphi) \circ g(\varphi) = h(\varphi) \circ 0 = 0;$$

d.h., $hg \in \mathcal{I}_\varphi$.

L5.3.36 Lösung. Wir wissen aus 5.3.33 (c), dass Polynome vom Grad 0 nicht als Minimalpolynom in Frage kommen. Sei nun $g = X - a \in \mathbb{R}[X]$ ein normiertes Polynom von Grad 1. Dann gilt

$$g(A) = \begin{pmatrix} -a & 0 & 0 \\ 0 & 1-a & 0 \\ 0 & 0 & 1-a \end{pmatrix} \tag{5.5.a}$$

Diese Matrix ist für keinen Wert $a \in \mathbb{R}$ die Nullmatrix, also hat das Minimalpolynom von A einen Grad größer als 1.

Können wir ein Polynom von Grad 2 finden, das A annulliert? Beim Blick auf die Gleichung (5.5.a) kommt man auf die Idee die Polynome $(X - 1)$ und $X = (X - 0)$ zu multiplizieren. Wir setzen also $g = X^2 - X$ und überprüfen, dass dieses Polynom A annulliert. Wir erhalten

$$g(A) = A^2 - A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = 0.$$

Damit ist $\mu_A = X^2 - X$ das Minimalpolynom von A.

L5.3.43 Lösung. Aus Beispiel 5.3.3 wissen wir, dass $\chi_A = X^2 - 3X - 2$ gilt. Wir setzen A ein und erhalten

$$\chi_A(A) = A^2 - 3A - 2I_2 = \begin{pmatrix} 2 & 3 \\ 6 & 11 \end{pmatrix} - \begin{pmatrix} 0 & 3 \\ 6 & 9 \end{pmatrix} - \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

L5.3.44 Lösung. Es sei $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ eine allgemeine 2×2-Matrix. Dann gilt

$$A^2 = \begin{pmatrix} a_{11}^2 + a_{12}a_{21} & a_{12}a_{11} + a_{12}a_{22} \\ a_{11}a_{21} + a_{22}a_{21} & a_{12}a_{21} + a_{22}^2 \end{pmatrix}.$$

Verwendet man die Formel aus Beispiel 5.3.6 um $\chi_A(A)$ zu berechnen, dann erhält man mit einer unangenehmen Rechnung

$$\begin{aligned}\chi_A(A) &= A^2 - (a_{11}+a_{22})A + (a_{11}a_{22} - a_{12}a_{21})I_2 \\ &= \begin{pmatrix} a_{11}^2 + a_{12}a_{21} & a_{12}a_{11} + a_{12}a_{22} \\ a_{11}a_{21} + a_{22}a_{21} & a_{12}a_{21} + a_{22}^2 \end{pmatrix} - (a_{11}+a_{22})\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22}\end{pmatrix} \\ &\quad + \begin{pmatrix} a_{11}a_{22} - a_{12}a_{21} & 0 \\ 0 & a_{11}a_{22} - a_{12}a_{21}\end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.\end{aligned}$$

L5.4.3 **Lösung.** Es seien $A = (a_{ij}), B = (b_{ij}) \in \mathrm{M}_{n,n}(\mathbb{R})$ zwei stochastische Matrizen. Sei $p \in \mathbb{R}$ mit $0 \le p \le 1$. Wir bemerken, dass $pA + (1-p)B$ nicht-negativ ist, weil A, B, p und $1-p$ nicht-negativ sind. Sei $j \in \{1, 2, \ldots, n\}$. Wir summieren die Einträge der j-ten Spalte der Matrix $pA + (1-p)B$:

$$\begin{aligned}\sum_{i=1}^n pa_{ij} + (1-p)b_{ij} &= p\sum_{i=1}^n a_{ij} + (1-p)\sum_{i=1}^n b_{ij} \\ &= p \cdot 1 + (1-p)\cdot 1 && (A, B \text{ stochastisch}) \\ &= 1.\end{aligned}$$

Folglich ist $pA + (1-p)B$ eine stochastische Matrix. (Man nennt $pA + (1-p)B$ eine *Konvexkombination* von A und B.)

L5.4.4 **Lösung.** Seien $A = (a_{ij}), B = (b_{ij}) \in \mathrm{M}_{n,n}(\mathbb{R})$ zwei stochastische Matrizen. Sei $j \in \{1, 2 \ldots, n\}$. Wir summieren die Einträge in der j-ten Spalte der Matrix $C = AB$ und erhalten

$$\begin{aligned}\sum_{i=1}^n c_{ij} &= \sum_{i=1}^n \sum_{k=1}^n a_{ik}b_{kj} \\ &= \sum_{k=1}^n b_{kj} \underbrace{\Big(\sum_{i=1}^n a_{ik}\Big)}_{=1} \\ &= \sum_{k=1}^n b_{kj} && (A \text{ stochastisch}) \\ &= 1 && (B \text{ stochastisch}).\end{aligned}$$

Aus der Formel für die Matrizenmultiplikation sieht man, dass $C = AB$ wieder eine nicht-negative Matrix ist; es handelt sich also um eine stochastische Matrix.

L5.4.7 **Lösung.** Wir quadrieren die stochastische Matrix A und erhalten

$$A^2 = \frac{1}{9}\begin{pmatrix} 1 & 6 & 4 \\ 2 & 0 & 4 \\ 6 & 3 & 1 \end{pmatrix}.$$

Diese Matrix ist noch nicht positiv, also nehmen wir die dritte Potenz

$$A^3 = \frac{1}{27}\begin{pmatrix} 13 & 12 & 6 \\ 2 & 12 & 8 \\ 12 & 3 & 13 \end{pmatrix}.$$

Da A^3 positiv ist, sind die Voraussetzung des Satzes von Perron erfüllt. Wir berechnen den Eigenraum zum Eigenwert 1:

$$\mathrm{Eig}_A(1) = \mathrm{Ker}\frac{1}{3}\begin{pmatrix} -2 & 0 & 2 \\ 2 & -3 & 0 \\ 0 & 3 & -2 \end{pmatrix} = \langle \begin{pmatrix} 3 \\ 2 \\ 3 \end{pmatrix} \rangle.$$

L5.4.14 **Lösung.** Aus dem gegebenen Graphen erhalten wir die Übergangsmatrix

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

und damit die Google-Matrix

$$G = \frac{2}{3}S + \frac{1}{9}E = \frac{1}{9}\begin{pmatrix} 1 & 3 & 1 \\ 5 & 3 & 7 \\ 3 & 3 & 1 \end{pmatrix}.$$

Steffen Kionke

Lineare Algebra

Lektion 6:
Die Jordan’sche Normalform

Fakultät für
Mathematik und Informatik

FernUniversität in Hagen

Studierhinweise zur sechsten Lektion

In dieser Lektion werden wir eine Lösung des Normalformproblems besprechen. Wir werden zeigen, dass jede komplexe quadratische Matrix ähnlich zu einer Matrix in *Jordan'scher Normalform* ist. Der Name wurde zu Ehren des französischen Mathematikers Camille JORDAN (1838–1922) gewählt[1]. Grob gesagt ist eine Matrix A in Normalform, wenn alle Einträge außerhalb der Diagonalen und der Nebendiagonalen (also an den Stellen $(i, i+1)$) verschwinden. Dabei sollen die Diagonaleinträge $a_{i,i}$ die Eigenwerte von A sein und jeder Eintrag der Nebendiagonalen ist entweder 0 oder 1. Allgemein gibt es zu einem Endomorphismus φ genau dann eine Matrixdarstellung in Jordan'scher Normalform, wenn das charakteristische Polynom in Linearfaktoren zerfällt. Die komplexen Zahlen braucht man nur, um sicherzustellen, dass diese Bedingung für jede Matrix erfüllt ist.

Die Jordan'sche Normalform ist eine gute Lösung des Normalformproblems, weil man alle wesentlichen Eigenschaften eines Endomorphismus direkt aus einer Matrixdarstellung in Normalform ablesen kann. Insbesondere kann man damit die komplexen Matrizen bis auf Ähnlichkeit klassifizieren. Die Jordan'sche Normalform hat innerhalb der Mathematik noch einige weitere Anwendungen.

Um die Jordan'sche Normalform zu verstehen, müssen wir uns zuerst genauer mit invarianten Unterräumen beschäftigen. Wir werden sehen, dass jede Faktorisierung des Minimalpolynoms in teilerfremde Faktoren uns eine Zerlegung in invariante Unterräume liefert. Damit können wir insbesondere die Zerlegung in *Haupträume* herleiten. Der schwierigste Schritt auf dem Weg zur Jordan'schen Normalform ist die Normalform für nilpotente Matrizen und Endomorphismen. Wir werden uns ausgiebig mit diesem Fall beschäftigen. Wenn wir die nilpotenten Endomorphismen einmal verstanden haben, können wir die allgemeine Jordan'sche Normalform mithilfe der Zerlegung in Haupträume auf den nilpotenten Fall zurückführen.

Diese Lektion ist wahrscheinlich die anspruchsvollste. Die Jordan'sche Normalform ist nicht leicht zu verstehen und wenn man sie berechnen möchte, kommt es wirklich auf die Details an! Bleiben Sie am Ball und fragen Sie frühzeitig nach, wenn Sie etwas nicht verstehen.

Es lohnt sich, wenn Sie sich zuerst ausführlich mit den nilpotenten Matrizen und Endomorphismen befassen. Wenn Sie die nilpotente Normalform berechnen können, dann ist es zur Jordan'schen Normalform nicht mehr weit.

[1] Man spricht den Namen „Jordan“ hier also ungefähr „Schordoan“ und nicht „Tschorden“

Lernziele

Nach dem Durcharbeiten der Lektion sollten Sie

→ wissen, wie man aus dem Minimalpolynom eines Endomorphismus eine Zerlegung in *invariante Unterräume* bekommt,

→ wissen, wie man anhand des Minimalpolynoms erkennt, ob ein Endomorphismus diagonalisierbar ist,

→ verstehen, was ein *nilpotenter* Endomorphismus ist,

→ wissen, was ein *zyklischer Unterraum* ist,

→ wissen, dass es zu jedem nilpotenten Endomorphismus eine Zerlegung in *zyklische Unterräume* gibt,

→ mit Partitionen umgehen können,

→ die *nilpotente Normalform* kennen und berechnen können,

→ *die Jordan'sche Normalform* und den Hauptsatz kennen,

→ die Haupträume und die Jordan'sche Normalform einer Matrix berechnen können,

→ Eigenschaften einer Matrix aus der Jordan'schen Normalform ablesen können,

→ Anwendungen der Jordan'schen Normalform nennen können.

Literaturhinweise

Die Jordan'sche Normalform wird in vielen Lehrbüchern zur Linearen Algebra besprochen (nur in [Beu] fehlt das Thema). Allerdings wird das Thema oft kurz abgehandelt (teilweise werden Beweise ausgelassen). Die Perspektive auf das Thema und auch die Terminologie unterscheidet sich von unserer. Unter anderem wird die Rangpartition in keiner der genannten Quellen diskutiert. Die Begriffe „Jordanblock“ und „Jordankästchen“ werden nicht einheitlich verwendet.

→ [Bär]: 6.6 und B.4.

→ [Bo]: 6.3 – 6.5 (Achtung: ganz anderer Weg!).

→ [Fi]: 5.5 – 5.7.

→ [Gö]: Kapitel 7.

→ [KaSt]: 8.7, 8.8.

Fahrplan durch die Lektion

In Abschnitt 6.1 befassen wir uns nochmal mit invarianten Unterräumen. Ist φ ein Endomorphismus und $g \in K[X]$ ein Polynom, dann definieren wir $U_\varphi(g) = \operatorname{Ker}(g(\varphi))$ und zeigen, dass dies ein φ-invarianter Unterraum ist (6.1.1). Diese Unterräume werden wir im folgenden genauer untersuchen. Wir beginnen mit einigen Beispielen (6.1.3–6.1.6). Das wesentliche Ergebnis ist Satz 6.1.8: Für teilerfremde Polynome g_1, g_2 gilt $U_\varphi(g_1 g_2) = U_\varphi(g_1) \oplus U_\varphi(g_2)$. Damit erhalten wir zu jeder *Zerlegung des Minimalpolynoms* μ_φ in teilerfremde Faktoren eine *Zerlegung von* V in φ-invariante Unterräume (6.1.10). ← 6.1

Im zweiten Teil dieses Abschnitts, setzen wir diese Methode ein, um ein neues Kriterium für *Diagonalisierbarkeit* herzuleiten. Wir zeigen, dass φ genau dann diagonalisierbar ist, wenn μ_φ in paarweise verschiedene Linearfaktoren zerfällt (6.1.11).

In Abschnitt 6.2 untersuchen wir ausführlich die *nilpotenten* Endomorphismen und Matrizen (siehe Definition 6.2.1). Dieser Abschnitt ist anspruchsvoll, aber auch wichtig zum Verständnis der Jordan'schen Normalform. ← 6.2

Im ersten Teil definieren wir die *Stufe* von Vektoren (6.2.7) und die *Filtrierung* zu einem nilpotenten Endomorphismus (6.2.13, 6.2.12). Mit diesen Begriffen kann man gut verstehen, was ein nilpotenter Endomorphismus macht. Wir beweisen grundlegende Charakterisierungssätze für nilpotente Endomorphismen und Matrizen (6.2.17, 6.2.18).

Im zweiten Teil definieren wir *zyklische Unterräume* (6.2.20). Wir zeigen, dass diese invariant sind (6.2.21) und beschreiben im nilpotenten Fall eine Basis (6.2.23). Damit beweisen wir, dass man jeden Vektorraum mit einem nilpotenten Endomorphismus in zyklische Unterräume zerlegen kann (6.2.25).

Im dritten Teil machen wir einen kleinen Exkurs und besprechen *Partitionen* (6.2.27). Eine Partition von $n \in \mathbb{N}$ ist ein Tupel $(p_1, \dots, p_s)$ natürlicher Zahlen mit $p_1 \geq \cdots \geq p_s$ und $p_1 + p_2 + \cdots + p_s = n$. Wir besprechen das *Diagramm* (6.2.30) einer Partition p und die *duale Partition* p^* (6.2.32). Wir zeigen, dass $p = (p^*)^*$ gilt (6.2.36) und leiten eine alternative Beschreibung der dualen Partition her (6.2.40).

Im vierten Teil besprechen wir die Normalform für nilpotente Endomorphismen. Zu jedem nilpotenten Endomorphismus φ definieren wir die *Rangpartition* $p(\varphi)$ (6.2.41). Wir erläutern den Zusammenhang zwischen der Rangpartition und einer Zerlegung in zyklische Unterräume (6.2.48). Wir definieren, was eine nilpotente Matrix in *Normalform* ist (6.2.49) und beweisen den wichtigen Satz 6.2.51: Jeder nilpotente Endomorphismus hat eine Matrixdarstellung in Normalform. Diese Darstellung ist *eindeutig* durch die Rangpartition bestimmt. Wir schließen, dass zwei nilpotente Matrizen genau dann ähnlich sind, wenn sie dieselbe Rangpartition besitzen (6.2.54). Zum Abschluss schauen wir uns an, wie man die Normalform einer nilpotenten Matrix berechnet (6.2.56).

6.3 → Zur *Jordan'schen Normalform* kommen wir im Abschnitt 6.3. Wir definieren *Jordanblöcke* (6.3.1) und *Jordankästchen* (6.3.2). Damit können wir klären was eine Matrix in Jordan'scher Normalform ist (6.3.5). Es ist entscheidend, dass Sie sich diese Begriffe einprägen und die Unterschiede kennen! Wir formulieren dann den Hauptsatz für Endomorphismen (6.3.7) und Matrizen (6.3.8) deren charakteristisches Polynom in Linearfaktoren zerfällt. Bevor wir den Hauptsatz beweisen, klären wir, wie man Eigenschaften aus der Jordan'schen Normalform ablesen kann (6.3.9).

Der zweite Teil dieses Abschnitts enthält den Beweis des Hauptsatzes. Wir lernen dabei die *verallgemeinerten Eigenräume* (6.3.14) und die *Haupträume* (6.3.15) kennen. Mit den Ergebnissen aus 6.1 zerlegen wir den Vektorraum in Haupträume. Auf dem Hauptraum $H_\varphi(\lambda)$ zum Eigenwert λ ist $\varphi - \lambda\,\mathrm{id}$ nilpotent, sodass wir die Ergebnisse aus Abschnitt 6.2 anwenden können.

Im dritten Teil sehen wir uns an, wie man die Normalform bestimmen kann (6.3.19). Das Verfahren basiert auf dem Verfahren für die Normalform von nilpotenten Endomorphismen. Als Vorbereitung überlegen wir uns, wie man den Hauptraum zu einem Eigenwert berechnen kann (6.3.16).

Im vierten Teil besprechen wir Anwendungen der Jordan'schen Normalform. Wir klassifizieren die komplexen Matrizen bis auf Ähnlichkeit (6.3.22) und erklären, wie man mit der Jordan'schen Normalform eine Formel für die Potenzen einer Matrix bestimmen kann (6.3.23–6.3.29).

6.1. Invariante Unterräume

In diesem Abschnitt erarbeiten wir uns ein Werkzeug, das wir zur Lösung des Normalformproblems benötigen. Wir betrachten immer einen endlich-dimensionalen K-Vektorraum V und einen Endomorphismus $\varphi \in \mathrm{End}(V)$. Wir erinnern uns daran, dass wir $U \subseteq V$ einen φ-invarianten Unterraum nennen, wenn U ein Unterraum ist und

$$\varphi(u) \in U$$

für alle $u \in U$ gilt.

Unser Ziel ist es hier, V in eine direkte Summe φ-invarianter Unterräume zu zerlegen. Wir werden sehen, dass wir dazu das Minimalpolynom verstehen müssen. Als Anwendung können wir ein weiteres Kriterium für Diagonalisierbarkeit herleiten.

I. Zerlegung durch das Minimalpolynom

In diesem Teilabschnitt werden wir zu jedem Polynom $g \in K[X]$ und jedem Endomorphismus φ einen φ-invarianten Unterraum $U_\varphi(g)$ definieren. Unser Ziel ist es, zu zeigen, dass jede Faktorisierung des Minimalpolynoms $\mu_\varphi = g_1 \cdots g_r$ in ein Produkt teilerfremder Polynome, eine Zerlegung

$$V = U_\varphi(g_1) \oplus \cdots \oplus U_\varphi(g_r)$$

liefert (das ist Korollar 6.1.10 unten). Wir beginnen mit der Definition von $U_\varphi(g)$.

6.1.1 **Lemma.** *Sei V ein K-Vektorraum. Sei $\varphi \in \mathrm{End}(V)$ und sei $g \in K[X]$. Dann ist*

$$U_\varphi(g) := \mathrm{Ker}(g(\varphi))$$

ein φ-invarianter Unterraum von V.

Beweis. Kerne von linearen Abbildungen sind Unterräume, insbesondere ist $U_\varphi(g)$ ein Unterraum von V.

Wir schreiben $g = \sum_i a_i X^i \in K[X]$. Sei $u \in U_\varphi(g)$ beliebig. Dann gilt

$$\begin{aligned} g(\varphi)(\varphi(u)) &= \sum_i a_i \varphi^i(\varphi(u)) = \sum_i a_i \varphi^{i+1}(u) \\ &= \varphi\Big(\sum_i a_i \varphi^i(u)\Big) && (\varphi \text{ linear}) \\ &= \varphi(g(\varphi)(u)) \end{aligned}$$

$$= \varphi(0) = 0 \qquad\qquad (u \in \operatorname{Ker}(g(\varphi))).$$

Das heißt, $\varphi(u)$ liegt in $U_\varphi(g)$. Damit ist $U_\varphi(g)$ ein φ-invarianter Unterraum. $\square$

6.1.2 **Aufgabe**. Sei V ein K-Vektorraum. Sei $\varphi \in \operatorname{End}(V)$ und sei $g \in K[X]$. Zeigen Sie, dass $\operatorname{Bild}(g(\varphi))$ ein φ-invarianter Unterraum von V ist. **L**

6.1.3 **Beispiel**. Sei $\varphi \in \operatorname{End}(V)$ und sei $f \in K[X]$. Falls f den Endomorphismus φ annulliert, dann gilt

$$U_\varphi(f) = \operatorname{Ker}(f(\varphi)) = \operatorname{Ker}(0) = V.$$

Das gilt insbesondere dann, wenn $f = \mu_\varphi$ das Minimalpolynom von φ ist.

6.1.4 **Beispiel**. Es sei V ein K-Vektorraum. Sei $\varphi \in \operatorname{End}(V)$ und sei $\lambda \in K$. Ist $g = X - \lambda$, dann gilt

$$U_\varphi(X - \lambda) = \operatorname{Ker}(\varphi - \lambda \operatorname{id}_V) = \{v \in V \mid \varphi(v) = \lambda v\} = \operatorname{Eig}_\varphi(\lambda).$$

In diesem Fall sind die Unterräume $U_\varphi(g)$ also die Eigenräume, die wir aus Lektion 5 kennen.

6.1.5 **Bemerkung**. Sei K ein Körper. Ist $A \in \mathrm{M}_{n,n}(K)$ eine quadratische Matrix, dann schreiben wir $U_A(g)$ anstelle von $U_{f_A}(g)$. Das heißt,

$$U_A(g) = \operatorname{Ker}(g(A)).$$

6.1.6 **Beispiel**. Sei

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \in \mathrm{M}_{n,n}(\mathbb{R}).$$

Wir betrachten das Polynom $g = X^2$. Dann gilt

$$U_A(g) = \operatorname{Ker}(A^2) = \operatorname{Ker} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Der Unterraum $U_A(g)$ besteht also aus allen Vektoren $v \in \mathbb{R}^3$, deren letzter Eintrag 0 ist.

6.1.7 **Aufgabe**. Seien $g, h \in K[X]$ und sei $\varphi \in \operatorname{End}(V)$. Zeigen Sie: Ist g ein Teiler von h, dann gilt $U_\varphi(g) \subseteq U_\varphi(h)$. **L**

6.1.8 **Satz.** *Sei V ein endlich-dimensionaler K-Vektorraum und sei $\varphi \in \mathrm{End}(V)$. Ist $g = g_1 g_2 \in K[X]$ das Produkt teilerfremder Polynome g_1 und g_2, dann gilt*

$$U_\varphi(g) = U_\varphi(g_1) \oplus U_\varphi(g_2).$$

Beweis. Wir bemerken zunächst, dass $U_\varphi(g_1)$ und $U_\varphi(g_2)$ in $U_\varphi(g)$ enthalten sind; siehe 6.1.7. Da g_1 und g_2 teilerfremd sind, gilt $\mathrm{ggT}(g_1, g_2) = 1$ und es gibt Polynome $h_1, h_2 \in K[X]$ mit

$$h_1 g_1 + h_2 g_2 = 1;$$

das folgt aus 1.4.54 (c).

Aus Korollar 2.3.7 sehen wir, dass wir folgende zwei Eigenschaften nachweisen müssen:

$$U_\varphi(g) = U_\varphi(g_1) + U_\varphi(g_2) \quad \text{und} \quad U_\varphi(g_1) \cap U_\varphi(g_2) = \{0\}.$$

Dazu machen wir eine Beobachtung: Für jeden Vektor $v \in U_\varphi(g)$ gilt $h_1 g_1(\varphi)(v) \in U_\varphi(g_2)$, denn

$$\begin{aligned} g_2(\varphi)(h_1 g_1(\varphi)(v)) &= g_2 h_1 g_1(\varphi)(v) && \text{(Lemma 5.3.27)} \\ &= h_1 \underbrace{g_1 g_2}_{=g}(\varphi)(v) \\ &= h_1(\varphi)(\underbrace{g(\varphi)(v)}_{=0}) && \text{(Lemma 5.3.27)} \\ &= 0. \end{aligned}$$

Genauso gilt $h_2 g_2(\varphi)(v) \in U_\varphi(g_1)$. Daraus folgt

$$v = \mathrm{id}_V(v) = (h_1 g_1 + h_2 g_2)(\varphi)(v) = h_1 g_1(\varphi)(v) + h_2 g_2(\varphi)(v) \in U_\varphi(g_2) + U_\varphi(g_1)$$

für jeden Vektor $v \in U_\varphi(g)$. Also ist $U_\varphi(g)$ die Summe der beiden Teilräume.

Sei nun $v \in U_\varphi(g_1) \cap U_\varphi(g_2)$. Dann ist

$$\begin{aligned} v &= h_1 g_1(\varphi)(v) + h_2 g_2(\varphi)(v) \\ &= h_1(\varphi)(g_1(\varphi)(v)) + h_2(\varphi)(g_2(\varphi)(v)) && \text{(Lemma 5.3.27)} \\ &= 0 + 0 = 0, && (v \in U_\varphi(g_1) \cap U_\varphi(g_2)) \end{aligned}$$

d.h., $U_\varphi(g_1) \cap U_\varphi(g_2) = \{0\}$. Also ist $U_\varphi(g)$ die direkte Summe der beiden Unterräume. □

6.1.9 **Aufgabe**. Wir betrachten die Matrix $A = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & -1 \\ 1 & 2 & -1 \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{R})$ und das **L**
Polynom $g = X^3 + X$. Zeigen Sie, dass $U_A(g) = \mathbb{R}^3$ ist und berechnen Sie die Zerlegung $\mathbb{R}^3 = U_A(X) \oplus U_A(X^2+1)$.

6.1.10 **Korollar.** *Sei $\varphi \in \mathrm{End}(V)$. Ist das Minimalpolynom μ_φ ein Produkt*

$$\mu_\varphi = g_1 g_2 \cdots g_r$$

paarweise teilerfremder Polynome $g_i \in K[X]$, dann gilt

$$V = \bigoplus_{i=1}^{r} U_\varphi(g_i).$$

Ist g_i normiert, dann ist g_i das Minimalpolynom von $\varphi|_{U_\varphi(g_i)}$.

Beweis. Die Aussage zur direkten Summe beweisen wir durch Induktion nach r. Der Induktionsanfang mit $r = 2$ ist durch Satz 6.1.8 und der Beobachtung $U_\varphi(\mu_\varphi) = V$ (Beispiel 6.1.3) abgedeckt. Im Induktionsschritt sei $r \geq 3$ und wir nehmen an, dass die Aussage für $r-1$ bekannt ist. Wir setzen $g' = g_{r-1}g_r$ und stellen fest, dass $\mu_\varphi = g_1 \dots g_{r-2}g'$ gilt. Aus Aufgabe 1.4.57 folgt, dass diese Faktoren wieder paarweise teilerfremd sind. Aus der Induktionsannahme folgt

$$V = \bigoplus_{i=1}^{r-2} U_\varphi(g_i) \oplus U_\varphi(g').$$

Da $g' = g_{r-1}g_r$ ein Produkt teilerfremder Polynome ist, erhalten wir aus Satz 6.1.8

$$U_\varphi(g') = U_\varphi(g_{r-1}) \oplus U_\varphi(g_r)$$

und damit die Behauptung.

Nehmen wir nun an, dass g_1 normiert ist und schreiben $g = g_1 h$. Dabei sind g_1 und $h = g_2 \cdots g_r$ teilerfremd (siehe 1.4.57). Wir zeigen, dass g_1 das Minimalpolynom von $\varphi_1 := \varphi|_{U_\varphi(g_1)}$ ist. Nach Definition von $U_\varphi(g_1)$ gilt $g_1(\varphi_1) = 0$. Also ist das Minimalpolynom $\mu_1 := \mu_{\varphi_1}$ von φ_1 ein Teiler von g_1 (siehe Satz 5.3.37). Sei $v = u_1 + u_2 \in V$ mit $u_1 \in U_\varphi(g_1)$ und $u_2 \in U_\varphi(h)$. Dann gilt

$$\begin{aligned}(\mu_1 h)(\varphi)(v) &= (\mu_1 h)(\varphi)(u_1) + (\mu_1 h)(\varphi)(u_2) \\ &= h(\varphi)(\underbrace{\mu_1(\varphi)(u_1)}_{=0}) + \mu_1(\varphi)(\underbrace{h(\varphi)(u_2)}_{=0}) = 0.\end{aligned}$$

Das heißt, $\mu_1 h$ annulliert φ und μ_φ teilt $\mu_1 h$. Also gilt

$$\operatorname{Grad}(g_1) + \operatorname{Grad}(h) = \operatorname{Grad}(\mu_\varphi) \leq \operatorname{Grad}(\mu_1 h) = \operatorname{Grad}(\mu_1) + \operatorname{Grad}(h)$$

und damit $\operatorname{Grad}(g_1) \leq \operatorname{Grad}(\mu_1)$. Da μ_1 und g_1 normiert sind und μ_1 das Polynom g_1 teilt, folgt daraus schon $\mu_1 = g_1$. Da die Reihenfolge der Polynome im Produkt keine Rolle spielt, folgt aus demselben Argument auch $\mu_i = g_i$ für alle $i \in \{1, 2, \ldots, r\}$. □

II. Diagonalisierbarkeit

Mit diesen neuen Hilfsmitteln können wir ein weiteres Kriterium für die Diagonalisierbarkeit herleiten.

6.1.11 **Satz.** *Sei V ein endlich-dimensionaler K-Vektorraum. Ein Endomorphismus $\varphi \in \operatorname{End}(V)$ ist genau dann diagonalisierbar, wenn μ_φ in paarweise verschiedene Linearfaktoren zerfällt, d.h.,*

$$\mu_\varphi = (X - \lambda_1)(X - \lambda_2) \cdots (X - \lambda_r)$$

mit paarweise verschiedenen $\lambda_1, \ldots, \lambda_r \in K$.

Beweis. Nehmen wir zunächst an, dass μ_φ in paarweise verschiedene Linearfaktoren zerfällt. Sind die $\lambda_1, \ldots, \lambda_r$ paarweise verschieden, dann sind die Linearfaktoren $(X - \lambda_i)$ paarweise teilerfremd; siehe 1.4.56. Aus Korollar 6.1.10 folgt somit

$$V = \bigoplus_{i=1}^{r} U_\varphi(X - \lambda_i).$$

Wir wissen aus Beispiel 6.1.4, dass $U_\varphi(X - \lambda_i) = \operatorname{Eig}_\varphi(\lambda_i)$ gilt. Also wird V von den Eigenvektoren von φ erzeugt und nach Satz 5.2.19 ist φ damit diagonalisierbar.

Sei umgekehrt φ diagonalisierbar. Es seien $\lambda_1, \ldots, \lambda_r$ die paarweise verschiedenen Eigenwerte von φ. Wir definieren $g = (X - \lambda_1) \cdots (X - \lambda_r)$ und behaupten, dass g das Minimalpolynom von φ ist. Da die Eigenwerte von φ die Nullstellen von μ_φ sind (siehe Satz 5.3.38), folgt aus dem Satz über die Nullstellen 1.4.30, dass g ein Teiler von μ_φ ist.

Da φ diagonalisierbar ist, existiert nach Satz 5.2.19 eine Basis $\mathcal{B}$ von V, die aus Eigenvektoren von φ besteht. Sei nun $v \in \mathcal{B}$ ein Eigenvektor zum Eigenwert λ_i. Wir schreiben $g = g' \cdot (X - \lambda_i)$ und beobachten

$$g(\varphi)(v) = g'(\varphi)(\varphi(v) - \lambda_i v) = 0.$$

Da der Endomorphismus $g(\varphi)$ durch die Bilder der Basisvektoren $v \in \mathcal{B}$ eindeutig bestimmt ist (siehe [MG, 8.4.1]), folgt daraus $g(\varphi) = 0$. Nach Satz 5.3.37 ist damit μ_φ ein Teiler von g. Also teilen sich g und μ_φ gegenseitig. Weil aber g und μ_φ normiert sind, muss $g = \mu_\varphi$ sein. □

Wenden wir diesen Satz auf den Endomorphismus $f_A \colon x \mapsto Ax$ für eine Matrix $A \in \mathrm{M}_{n,n}(K)$ an, dann erhalten wir:

6.1.12 **Korollar.** *Sei K ein Körper. Eine Matrix $A \in \mathrm{M}_{n,n}(K)$ ist genau dann diagonalisierbar, wenn das Minimalpolynom μ_A in paarweise verschiedene Linearfaktoren zerfällt.*

Dieses Kriterium ist gut geeignet um zu zeigen, dass eine gegebene Matrix nicht diagonalisierbar ist.

6.1.13 **Beispiel**. Wir betrachten die Matrix

$$A = \begin{pmatrix} 0 & 3 \\ 0 & 0 \end{pmatrix}.$$

Es gilt $A^2 = 0$, also ist das Minimalpolynom ein Teiler von X^2. Da A nicht die Nullmatrix ist, annulliert X die Matrix A nicht. Also ist $\mu_A = X^2$ das Minimalpolynom. Aus Korollar 6.1.12 folgt nun, dass A nicht diagonalisierbar ist, denn der Linearfaktor X kommt mit Vielfachheit 2 im Minimalpolynom vor.

6.1.14 **Beispiel**. Sei $A \in \mathrm{M}_{n,n}(\mathbb{C})$ mit $A^k = I_n$ für ein $k \in \mathbb{N}$ gegeben. Dann ist A diagonalisierbar. In der Tat, das Polynom $X^k - 1$ annulliert A. In den komplexen Zahlen zerfällt $X^k - 1$ aber in paarweise verschiedene Linearfaktoren:

$$X^k - 1 = \prod_{j=0}^{k-1} (X - \omega_j);$$

dabei ist $\omega_j = \exp(2\pi i \frac{j}{k})$ eine sogenannte k-te *Einheitswurzel*. Das heißt, es gilt $\omega_j^k = 1$. Das Minimalpolynom von A ist also ein Teiler von $X^k - 1$ und zerfällt folglich in paarweise verschiedene Linearfaktoren (vgl. Korollar 1.4.39).

6.1.15 **Aufgabe**. Zeigen Sie, dass eine Matrix $A \in \mathrm{M}_{n,n}(\mathbb{F}_2)$ genau dann diagonalisierbar ist, wenn $A^2 = A$ gilt. **L**

6.2. Nilpotente Endomorphismen

Die diagonalisierbaren Endomorphismen verstehen wir bereits gut. Wir wissen nun, dass dies genau jene Endomorphismen sind, deren Minimalpolynom in paarweise verschiedene Linearfaktoren zerfällt. Was passiert aber, wenn das Minimalpolynom nicht in paarweise verschiedene Linearfaktoren zerfällt? Hier wollen wir den speziellen Fall betrachten, dass das Minimalpolynom die einfache Form X^k besitzt. Das Minimalpolynom zerfällt also in Linearfaktoren, aber diese Linearfaktoren sind alle gleich X und damit nicht paarweise verschieden.

I. Definition und Charakterisierung

6.2.1 **Definition** Sei V ein K-Vektorraum. Ein Endomorphismus $\varphi \in \mathrm{End}(V)$ heißt *nilpotent*, wenn es ein $m \in \mathbb{N}_0$ gibt mit $\varphi^m = 0$.

Eine quadratische Matrix $A \in \mathrm{M}_{n,n}(K)$ heißt *nilpotent*, wenn es ein $m \in \mathbb{N}_0$ gibt mit $A^m = 0$.

In beiden Fällen nennt man die kleinste Zahl $m \in \mathbb{N}_0$ mit dieser Eigenschaft den *Nilpotenzindex* von φ bzw. A. Den Nilpotenzindex bezeichnen wir mit $\mathrm{Ni}(\varphi)$ bzw. $\mathrm{Ni}(A)$.

6.2.2 **Anmerkungen zur Definition.** (a) Sei V ein endlich-dimensionaler Vektorraum und sei $\varphi \in \mathrm{End}(V)$. Der Endomorphismus φ ist genau dann nilpotent, wenn die Matrixdarstellung ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)$ bezüglich einer geordneten Basis $\mathcal{B}$ nilpotent ist. In der Tat, wegen Proposition 9.3.1 aus den Mathematischen Grundlagen [MG] gilt für alle $m \in \mathbb{N}_0$ die Gleichheit

$${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi^m) = {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)^m.$$

Daraus folgt die Behauptung, denn es gilt $\varphi^m = 0$ genau dann, wenn für die Matrixdarstellung ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi^m) = 0$ gilt.

(b) Ist $\varphi \in \mathrm{End}(V)$ ein nilpotenter Endomorphismus mit Nilpotenzindex $\mathrm{Ni}(\varphi) = 0$, dann gilt $0 = \varphi^0 = \mathrm{id}_V$ und damit muss $V = \{0\}$ sein.

(c) Ist $V \neq \{0\}$ und $\varphi \in \mathrm{End}(V)$ nilpotent, dann ist $\mathrm{Ker}(\varphi) \neq \{0\}$. Um das zu sehen sei $k = \mathrm{Ni}(\varphi)$ der Nilpotenzindex von φ. Da k minimal mit der Eigenschaft $\varphi^k = 0$ ist, ist $\varphi^{k-1} \neq 0$, d.h., es gibt einen Vektor $v \in V$ mit $w = \varphi^{k-1}(v) \neq 0$. Dann liegt w im Kern von φ, denn $\varphi(w) = \varphi^k(v) = 0$. In anderen Worten: 0 ist ein Eigenwert von φ.

(d) Hat $\varphi \in \operatorname{End}(V)$ einen Eigenwert $\lambda \neq 0$, dann ist φ *nicht* nilpotent. In der Tat, ist $v \in V$ ein Eigenvektor zum Eigenwert λ, dann gilt $\varphi^m(v) = \lambda^m v \neq 0$, denn $\lambda^m \neq 0$ und $v \neq 0$ (Definition 5.2.1).

6.2.3 **Beispiel**. (a) Die Nullmatrix $A = 0 \in \mathrm{M}_{n,n}(K)$ ist nilpotent vom Nilpotenzindex $\mathrm{Ni}(A) = 1$.

(b) Die Matrix

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in \mathrm{M}_{2,2}(K)$$

ist nilpotent, denn es gilt $A^2 = 0$. Weil $A \neq 0$ ist, ist $\mathrm{Ni}(A) = 2$ der Nilpotenzindex von A.

(c) Sei $\dim_K(V) = 1$ und sei $\varphi \in \operatorname{End}(V)$. Dann ist $V = \langle u \rangle$ für einen Vektor $u \neq 0$ in V und u ist ein Eigenvektor von φ; das wissen wir aus Beispiel 4.4.4. Es gilt also $\varphi(u) = \lambda u$ für ein $\lambda \in K$. Ist $\lambda \neq 0$, dann ist φ nicht nilpotent wegen 6.2.2 (d). Ist $\lambda = 0$, dann ist φ die Nullabbildung und ist damit nilpotent mit Nilpotenzindex 1.

6.2.4 **Aufgabe**. Zeigen Sie, dass die Matrix **L**

$$A = \begin{pmatrix} -2 & 4 & -6 & 12 \\ -1 & 2 & -3 & 6 \\ 0 & 0 & -2 & 4 \\ 0 & 0 & -1 & 2 \end{pmatrix} \in \mathrm{M}_{4,4}(\mathbb{Q})$$

nilpotent ist und bestimmen Sie den Nilpotenzindex.

6.2.5 **Aufgabe**. Zeigen Sie, dass die Matrix **L**

$$B = \begin{pmatrix} \overline{2} & \overline{2} & \overline{0} \\ \overline{2} & \overline{0} & \overline{2} \\ \overline{0} & \overline{2} & \overline{1} \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{F}_3)$$

nilpotent ist und bestimmen Sie den Nilpotenzindex.

6.2.6 **Aufgabe**. Es seien $A, B \in \mathrm{M}_{n,n}(K)$ ähnliche Matrizen. Zeigen Sie: Ist A nilpotent, dann ist auch B nilpotent und es gilt $\mathrm{Ni}(A) = \mathrm{Ni}(B)$. **L**

6.2.7 **Definition** Sei $\varphi \in \operatorname{End}(V)$ nilpotent. Die *Stufe* von $v \in V$ ist die kleinste Zahl $j \in \mathbb{N}_0$, sodass $\varphi^j(v) = 0$ gilt.

6.2.8 **Bemerkung**. Der Nullvektor ist per Definition der einzige Vektor der Stufe 0. Die Elemente $v \neq 0$ in $\mathrm{Ker}(\varphi)$ sind die Vektoren erster Stufe. Die höchste Stufe, die ein Vektor haben kann, ist der Nilpotenzindex von φ. Auch umgekehrt gilt: Der Nilpotenzindex von φ ist das Maximum der Stufen aller Vektoren in V.

6.2.9 **Aufgabe**. Es sei $\varphi \in \mathrm{End}(V)$ nilpotent und es sei $v \in V$ ein Vektor der Stufe $j > 0$. Zeigen Sie, dass $\varphi(v)$ ein Vektor der Stufe $j-1$ ist. **L**

6.2.10 **Beispiel**. (a) Wir betrachten die Matrix $A = \begin{pmatrix} 0 & 4 \\ 0 & 0 \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{R})$. Der Vektor $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ ist ein Vektor zweiter Stufe für A, denn es ist $A^2 e_2 = 0$ aber $Ae_2 = 4e_1 \neq 0$.

(b) Wir betrachten die Matrix A aus Aufgabe 6.2.4. Der erste Basisvektor ist ein Vektor zweiter Stufe für A, denn es gilt

$$Ae_1 = \begin{pmatrix} -2 \\ -1 \\ 0 \\ 0 \end{pmatrix} \neq 0$$

aber $A^2 e_1 = 0$. Daraus folgt auch, dass der Vektor Ae_1 ein Vektor erster Stufe ist.

6.2.11 **Aufgabe**. Sei $\varphi \in \mathrm{End}(V)$. Machen Sie sich klar, dass für alle $j \in \mathbb{N}_0$ die Beziehungen $\mathrm{Ker}(\varphi^j) \subseteq \mathrm{Ker}(\varphi^{j+1})$, $\mathrm{Bild}(\varphi^j) \supseteq \mathrm{Bild}(\varphi^{j+1})$, $\varphi^{-1}(\mathrm{Ker}(\varphi^j)) = \mathrm{Ker}(\varphi^{j+1})$ und $\varphi(\mathrm{Bild}(\varphi^j)) = \mathrm{Bild}(\varphi^{j+1})$ gelten. **L**

Wie kann man sich einen nilpotenten Endomorphismus vorstellen? Jede Anwendung reduziert die Stufe eines Vektors $v \neq 0$ um 1, bis man irgendwann den Nullvektor erreicht. Das haben wir schon in Aufgabe 6.2.9 gesehen. Das folgende Lemma erklärt das nochmal in etwas anderer Form.

6.2.12 **Lemma.** *Sei $V \neq \{0\}$ ein K-Vektorraum und sei $\varphi \in \mathrm{End}(V)$ nilpotent mit $k = \mathrm{Ni}(\varphi)$.*

(a)
$$\{0\} = \mathrm{Ker}(\varphi^0) \subsetneq \mathrm{Ker}(\varphi) \subsetneq \mathrm{Ker}(\varphi^2) \subsetneq \cdots \subsetneq \mathrm{Ker}(\varphi^k) = V$$
ist eine echt aufsteigende Folge φ-invarianter Unterräume.

(b)
$$V = \mathrm{Bild}(\varphi^0) \supsetneq \mathrm{Bild}(\varphi) \supsetneq \mathrm{Bild}(\varphi^2) \supsetneq \cdots \supsetneq \mathrm{Bild}(\varphi^k) = \{0\}$$
ist eine echt absteigende Folge φ-invarianter Unterräume.

Beweis. Es gilt $\mathrm{Ker}(\varphi^j) = U_\varphi(X^j)$ und wir wissen aus Lemma 6.1.1, dass dies ein φ-invarianter Unterraum ist. Warum steigt die Kette der Kerne echt auf? Nehmen wir an, dass $\mathrm{Ker}(\varphi^j) = \mathrm{Ker}(\varphi^{j+1})$ für ein $j \leq k$ gilt. Dann gilt auch $\mathrm{Ker}(\varphi^{j+2}) = \varphi^{-1}(\mathrm{Ker}(\varphi^{j+1})) = \varphi^{-1}(\mathrm{Ker}(\varphi^j)) = \mathrm{Ker}(\varphi^{j+1}) = \mathrm{Ker}(\varphi^j)$. Dieses Argument können wir beliebig oft wiederholen und erhalten schließlich

$$V = \mathrm{Ker}(\varphi^k) = \mathrm{Ker}(\varphi^j).$$

Daraus folgt aber $\varphi^j = 0$ und da k der Nilpotenzindex von φ ist, muss $j = k$ sein.

Im Beweis von Aussage (b) nehmen wir aus Bequemlichkeit an, dass V endlich-dimensional ist. Sei $j \in \{0, 1, \ldots, k-1\}$ gegeben. Wir wissen aus Aufgabe 6.1.2, dass das $\mathrm{Bild}(\varphi^j)$ ein φ-invarianter Unterraum ist. Aus Aufgabe 6.2.11 kennen wir die Inklusion $\mathrm{Bild}(\varphi^{j+1}) \subseteq \mathrm{Bild}(\varphi^j)$. Der Rangsatz [MG, 8.3.14] liefert uns

$$\dim_K(V) = \dim_K(\mathrm{Ker}(\varphi^j)) + \dim_K(\mathrm{Bild}(\varphi^j)).$$

Wenn also die Dimension des Kerns mit jedem Schritt $j \rightsquigarrow j+1$ ansteigt, nimmt die Dimension des Bildes in jedem Schritt ab, d.h., die Inklusion $\mathrm{Bild}(\varphi^{j+1}) \subsetneq \mathrm{Bild}(\varphi^j)$ ist echt für alle $j \leq k-1$. □

Die Kerne von φ^j werden also immer größer mit j, die Bilder hingegen werden immer kleiner. Solange $\varphi^j \neq 0$ ist, muss damit das Bild von φ^j den Kern von φ nicht-trivial schneiden.

6.2.13 **Definition** Sei $\varphi \in \mathrm{End}(V)$ nilpotent mit $k = \mathrm{Ni}(\varphi)$. Wir schreiben $V_j = \mathrm{Ker}(\varphi^j)$ und nennen die aufsteigende Folge φ-invarianter Unterräume

$$\{0\} = V_0 \subsetneq V_1 \subsetneq V_2 \subsetneq \cdots \subsetneq V_k = V$$

die *Filtrierung* zu φ. Die Elemente in $V_j \setminus V_{j-1}$ sind also genau die Vektoren der Stufe j.

Ist $\varphi = f_A$ für eine quadratische Matrix A, dann sprechen wir auch von der Filtrierung zu A.

6.2.14 **Aufgabe**. Bestimmen Sie die Filtrierung zur Matrix **L**

$$B = \begin{pmatrix} \overline{2} & \overline{2} & \overline{0} \\ \overline{2} & \overline{0} & \overline{2} \\ \overline{0} & \overline{2} & \overline{1} \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{F}_3)$$

aus Aufgabe 6.2.5.

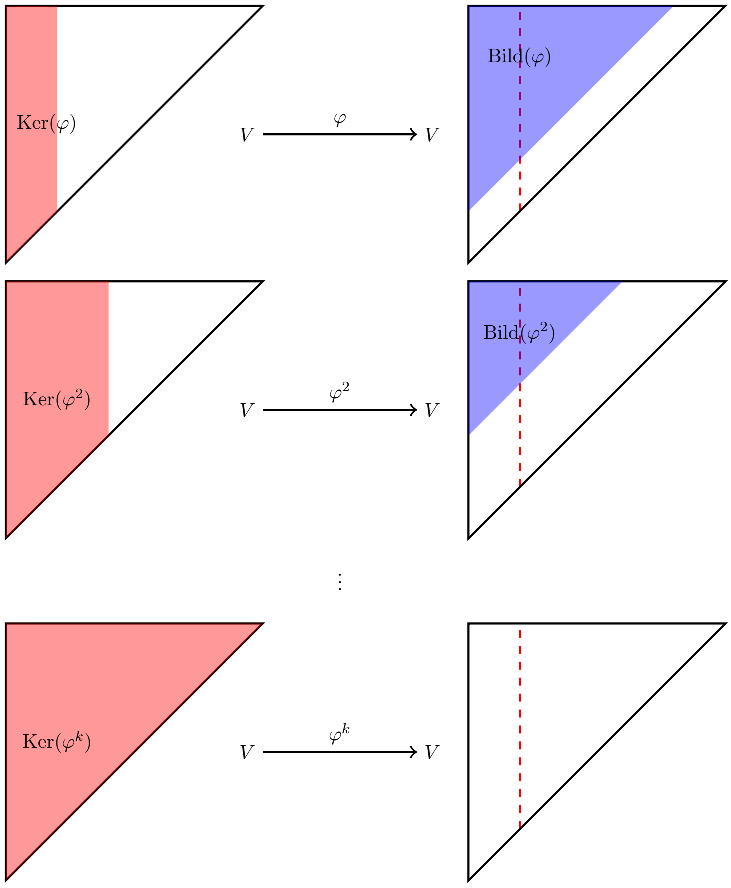


Abbildung 6.1.: Kern und Bild der Potenzen eines nilpotenten Endomorphismus

6.2.15 $\boxed{\textbf{Definition}}$ Eine obere Dreiecksmatrix Matrix $A = (a_{ij})_{i,j} \in \mathrm{M}_{n,n}(K)$ nennen wir *strikte obere Dreiecksmatrix*, wenn alle Diagonaleinträge verschwinden, d.h. $a_{ii} = 0$ für alle $i \in \{1, \dots, n\}$.

6.2.16 **Beispiel**. Wir betrachten die beiden reellen (3×3)-Matrizen

$$A = \begin{pmatrix} 0 & 5 & -1 \\ 0 & 0 & 9 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 1 & 5 & -1 \\ 0 & 0 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Beide sind obere Dreiecksmatrizen, aber nur A ist eine strikte obere Dreiecksmatrix.

Der folgende Satz liefert eine erste hilfreiche Charakterisierung der nilpotenten Endomorphismen.

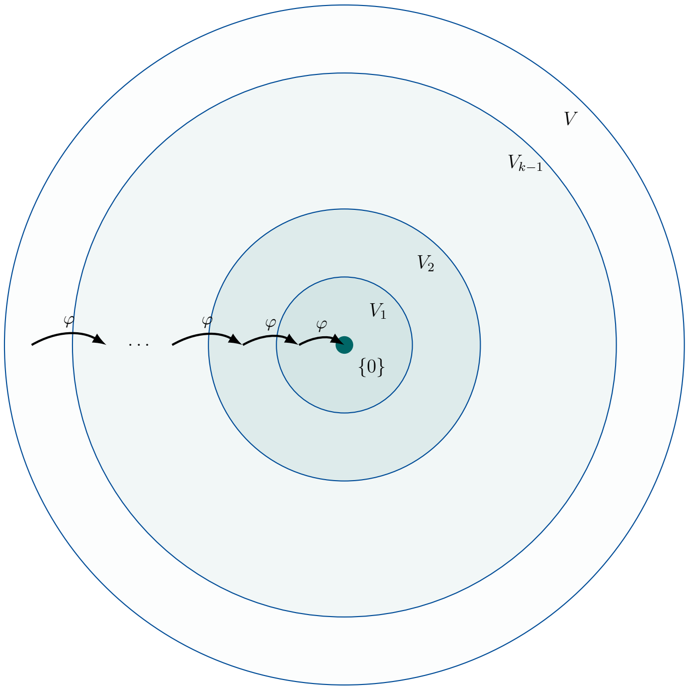


Abbildung 6.2.: Der nilpotente Endomorphismus φ verschiebt die Vektoren entlang der Filtrierung hin zum Nullvektor.

6.2.17 **Satz.** *Sei V ein K-Vektorraum der endlichen Dimension $n = \dim_K(V) \geq 1$. Sei $\varphi \in \operatorname{End}(V)$. Folgende Aussagen sind äquivalent:*

(i) φ ist nilpotent,

(ii) es gibt eine geordnete Basis $\mathcal{B}$ von V, sodass ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)$ eine strikte obere Dreiecksmatrix ist,

(iii) $\chi_\varphi = X^n$,

(iv) $\mu_\varphi = X^k$ für ein $k \in \mathbb{N}$.

Sind sie erfüllt, dann ist k aus (iv) der Nilpotenzindex von φ.

Beweis. „(i) $\Longrightarrow$ (ii)“: Wir nehmen an, dass φ nilpotent mit $\operatorname{Ni}(\varphi) = k$ ist. Wir wählen eine geordnete Basis $\mathcal{B}_1 = (v_1, \dots, v_{\ell_1})$ von V_1, ergänzen diese zu einer geordneten Basis $\mathcal{B}_2 = (v_1, \dots, v_{\ell_1}, \dots, v_{\ell_2})$ von V_2, ergänzen diese nochmal zu einer geordneten Basis $\mathcal{B}_3$ von V_3 und so weiter, bis wir eine geordnete Basis $\mathcal{B}_k = (v_1, \dots, v_n)$ von V_k haben. Ein Basisvektor v_i, der im j-ten Schritt hinzugefügt wurde, ist also ein Vektor der Stufe j. Damit liegt $\varphi(v_i)$ in V_{j-1} und wir können $\varphi(v_i)$ als Linearkombination der Vektoren aus $\mathcal{B}_{j-1}$ schreiben. Das heißt, es werden nur Basisvektoren benötigt, die vor v_i in der geordneten Basis stehen. Damit ist ${}_{\mathcal{B}_k}\mathrm{M}_{\mathcal{B}_k}(\varphi)$ eine strikte obere Dreiecksmatrix.

„(ii) $\Longrightarrow$ (iii)“: Übungsaufgabe! **L**

„(iii) $\Longrightarrow$ (iv)“: Wir nehmen an, dass $\chi_\varphi = X^n$ gilt. Aus dem Satz von Cayley-Hamilton 5.3.41 wissen wir, dass das Minimalpolynom μ_φ ein Teiler von χ_φ ist. Also gilt $\mu_\varphi = X^k$ für ein $k \leq n$; das haben wir in Lemma 1.4.37 gesehen.

„(iv) $\Longrightarrow$ (i)“: Wir nehmen an, dass $\mu_\varphi = X^k$ gilt. Das Minimalpolynom annulliert φ, also gilt $\varphi^k = 0$. Damit ist φ nilpotent.

Es gilt in diesem Fall sicherlich $\mathrm{Ni}(\varphi) \leq k$. Umgekehrt annulliert das Polynom $X^{\mathrm{Ni}(\varphi)}$ den Endomorphismus φ. Da der Grad von μ_φ minimal unter den nicht-trivialen Polynomen im Verschwindungsideal ist, gilt auch $k \leq \mathrm{Ni}(\varphi)$. Insgesamt folgt daraus $k = \mathrm{Ni}(\varphi)$. □

Man kann einen entsprechenden Satz natürlich auch für Matrizen formulieren:

6.2.18 **Satz.** *Seien K ein Körper und $A \in \mathrm{M}_{n,n}(K)$. Folgende Aussagen sind äquivalent:*

(i) *A ist nilpotent,*

(ii) *A ist ähnlich zu einer strikten oberen Dreiecksmatrix,*

(iii) $\chi_A = X^n$,

(iv) $\mu_A = X^k$ *für ein* $k \in \mathbb{N}$.

In diesem Fall ist $k = \mathrm{Ni}(A)$ der Nilpotenzindex von A.

Beweis. Wir betrachten den Endomorphismus $f_A \colon x \mapsto Ax$ von K^n. Dann ist A genau dann nilpotent, wenn f_A nilpotent ist; siehe 6.2.2 (a). Da A die Matrixdarstellung von f_A zur Standardbasis ist, wissen wir aus Lemma 5.1.3, dass die Matrixdarstellungen von f_A genau die Matrizen sind, die zu A ähnlich sind. Da per Definition außerdem $\chi_A = \chi_{f_A}$ und $\mu_A = \mu_{f_A}$ gelten, folgt die Aussage unmittelbar aus Satz 6.2.17. □

Durch den Satz von Cayley-Hamilton ist der Grad des Minimalpolynoms immer höchstens der Grad des charakteristischen Polynoms. Damit erhalten wir aus diesem Satz eine hilfreiche Information.

6.2.19 **Korollar.** *Ist $A \in \mathrm{M}_{n,n}(K)$ nilpotent, dann ist $\mathrm{Ni}(A) \leq n$.*

Das heißt, wenn wir den Nilpotenzindex einer nilpotenten Matrix $A \in \mathrm{M}_{n,n}(K)$ bestimmen wollen, müssen wir höchstens die Potenzen $A, A^2, \ldots, A^n$ berechnen.

II. Zyklische Unterräume

6.2.20 **Definition** Sei V ein K-Vektorraum und sei $\varphi \in \operatorname{End}(V)$. Sei $v \in V$. Dann heißt

$$Z_\varphi(v) = \langle \{\varphi^k(v) \mid k \in \mathbb{N}_0\} \rangle$$

der von v erzeugte *zyklische Unterraum* von V.

Der Unterraum $Z_\varphi(v)$ wird also von den Vektoren $v, \varphi(v), \varphi^2(v), \varphi^3(v), \ldots$ aufgespannt.

6.2.21 **Lemma.** *Seien V ein K-Vektorraum, $v \in V$ und $\varphi \in \operatorname{End}(V)$. Der zyklische Unterraum $Z_\varphi(v)$ ist φ-invariant.*

Beweis. Da $Z_\varphi(v)$ als lineare Hülle definiert wurde, handelt es sich um einen Unterraum von V. Wir zeigen, dass dieser Unterraum φ-invariant ist. Sei dazu $w \in Z_\varphi(v)$. Nach Definition der linearen Hülle gibt es eine Zahl $k \in \mathbb{N}_0$ und Skalare $a_0, \ldots, a_k \in K$, sodass

$$w = \sum_{j=0}^{k} a_j \varphi^j(v)$$

ist. Wir wenden φ auf w an und erhalten

$$\varphi(w) = \varphi\Big(\sum_{j=0}^{k} a_j \varphi^j(v)\Big) = \sum_{j=0}^{k} a_j \varphi^{j+1}(v) \in Z_\varphi(v);$$

damit ist $Z_\varphi(v)$ invariant unter φ. □

6.2.22 **Aufgabe**. Zeigen Sie, dass $Z_\varphi(v)$ der kleinste φ-invariante Unterraum von V ist, der v enthält. **L**

Für nilpotente Endomorphismen sind zyklische Unterräume nützlich, denn man kann die Stufe eines Vektors $v \in V$ mit dem zyklischen Unterraum $Z_\varphi(v)$ bestimmen.

6.2.23 **Lemma.** *Sei V ein endlich-dimensionaler K-Vektorraum und sei $\varphi \in \operatorname{End}(V)$ nilpotent. Ist $v \in V$ ein Vektor der Stufe j, dann gilt $\dim_K Z_\varphi(v) = j$ und*

$$\mathcal{B} = \{\varphi^\ell(v) \mid 0 \leq \ell \leq j-1\}$$

ist eine Basis von $Z_\varphi(v)$.

Beweis. Per Definition wird der zyklische Unterraum $Z_\varphi(v)$ von der Menge $\{\varphi^\ell(v) \mid \ell \in \mathbb{N}_0\}$ aufgespannt. Da v ein Vektor der Stufe j ist, gilt $\varphi^\ell(v) = 0$ für alle $\ell \geq j$. Also ist $\mathcal{B} = \{\varphi^\ell(v) \mid \ell < j\}$ ein Erzeugendensystem von $Z_\varphi(v)$.

Wir zeigen, dass die Vektoren in $\mathcal{B}$ linear unabhängig sind. Ist $j = 0$ (also $v = 0$), dann ist $\mathcal{B}$ leer und damit linear unabhängig. Wir nehmen nun $v \neq 0$ an. Wir zeigen mittels Induktion nach $r \in \{1, \dots, j\}$, dass $\{\varphi^{j-r}(v), \dots, \varphi^{j-1}(v)\}$ linear unabhängig ist. Zum Induktionsanfang betrachten wir $r = 1$. In diesem Fall gilt $\varphi^{j-1}(v) \neq 0$, weil v ein Vektor der Stufe j ist, d.h., $\{\varphi^{j-1}(v)\}$ ist linear unabhängig. Im Induktionsschritt sei $1 < r \leq j$; wir nehmen an, dass $\{\varphi^{j-(r-1)}(v), \dots, \varphi^{j-1}(v)\}$ linear unabhängig ist. Nun betrachten wir eine Linearkombination

$$\sum_{\ell=j-r}^{j-1} a_\ell \varphi^\ell(v) = 0 \tag{6.2.a}$$

mit Koeffizienten $a_\ell \in K$ und zeigen, dass alle Koeffizienten a_ℓ verschwinden. Wir wenden φ auf die Gleichung (6.2.a) an und erhalten

$$\begin{aligned} 0 = \varphi(0) = \varphi\left(\sum_{\ell=j-r}^{j-1} a_\ell \varphi^\ell(v)\right) & \\ = \sum_{\ell=j-r}^{j-1} a_\ell \varphi^{\ell+1}(v) & \qquad (\varphi \text{ linear}) \\ = \sum_{\ell=j-r}^{j-2} a_\ell \varphi^{\ell+1}(v) & \qquad (\varphi^j(v) = 0) \\ = \sum_{\ell'=j-(r-1)}^{j-1} a_{\ell'-1} \varphi^{\ell'}(v) & \qquad (\ell' = \ell + 1). \end{aligned}$$

Mit der Induktionsannahme schließen wir nun $a_\ell = 0$ für alle $j - 1 > \ell \geq j - r$. Damit verschwinden in (6.2.a) alle Terme bis auf $a_{j-1}\varphi^{j-1}(v)$. Da $\varphi^{j-1}(v) \neq 0$ ist, folgt daraus $a_{j-1} = 0$. Damit sind die Vektoren in $\mathcal{B}$ linear unabhängig und $\mathcal{B}$ ist eine Basis von $Z_\varphi(v)$. Insbesondere gilt $j = |\mathcal{B}| = \dim_K Z_\varphi(v)$. □

6.2.24 **Beispiel**. Sei V ein K-Vektorraum mit einem nilpotenten Endomorphismus φ von Nilpotenzindex k. Wir nehmen an, dass es einen Vektor $v \in V$ gibt, sodass $V = Z_\varphi(v)$ gilt. In diesem Fall muss v ein Vektor der höchsten Stufe k sein, d.h., v liegt in $V \setminus \mathrm{Ker}(\varphi^{k-1})$. Wir definieren $v_i = \varphi^{k-i}(v)$. Zum Beispiel sind $v_1 = \varphi^{k-1}(v)$

und $v_k = \varphi^0(v) = v$. Wie wir gerade gesehen haben, ist nun $\mathcal{B} = (v_1, v_2, \ldots, v_k) = (\varphi^{k-1}(v), \varphi^{k-2}(v), \ldots, \varphi(v), v)$ eine geordnete Basis von V und es gilt

$$\varphi(v_i) = \begin{cases} v_{i-1} & \text{falls } i > 1 \\ 0 & \text{falls } i = 1 \end{cases}.$$

Die Matrixdarstellung von φ ist also

$$_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi) := \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & & \vdots \\ 0 & 0 & 0 & \ddots & 0 & 0 \\ \vdots & & & \ddots & 1 & 0 \\ & & & & 0 & 1 \\ 0 & & \cdots & 0 & 0 & 0 \end{pmatrix}.$$

Das heißt, alle Einträge verschwinden, nur die Einträge an den Positionen $(i-1, i)$ für $i \in \{2, 3, \ldots, k\}$ sind 1. Diese Matrix nennen wir $N(k)$.

Wir kommen nun zu einem sehr wichtigen Ergebnis.

6.2.25 **Satz.** *Sei V ein endlich-dimensionaler K-Vektorraum und sei φ ein nilpotenter Endomorphismus. Dann ist V eine direkte Summe zyklischer Unterräume. Das heißt, es gibt Vektoren $w_1, \ldots, w_p \in V \setminus \{0\}$, sodass*

$$V = \bigoplus_{i=1}^{p} Z_\varphi(w_i)$$

gilt. Dabei ist $p = \dim_K \operatorname{Ker}(\varphi)$.

Beweis. Wir beweisen die Aussage mit Induktion über die Dimension von V. Für $V = \{0\}$ ist nichts zu tun ($p = 0$). Wir nehmen also $V \neq \{0\}$ an; dann ist $k = \operatorname{Ni}(\varphi) \geq 1$. Als Induktionsvoraussetzung nehmen wir an, dass der Satz für alle Vektorräume mit Dimension kleiner $\dim_K(V)$ gilt. Als ersten Schritt wählen wir einen Vektor $w \in V$ der Stufe k.

Behauptung: Es gibt einen φ-invarianten Unterraum $U \subseteq V$ mit $U \oplus Z_\varphi(w) = V$.

Haben wir diese Behauptung bewiesen, dann folgt der Satz aus der Induktionsannahme, denn $\varphi|_U$ ist nilpotent und es gilt

$$\dim_K V = \dim_K(U) + \dim_K(Z_\varphi(w)) = \dim_K(U) + k > \dim_K(U).$$

Zum Beweis der Behauptung nehmen wir einen *maximalen* φ-invarianten Unterraum U mit der Eigenschaft $U \cap Z_\varphi(w) = \{0\}$. Die Summe von U und $Z_\varphi(w)$ ist also direkt und wir müssen nur zeigen, dass $U + Z_\varphi(w) = V$ ist.

Für einen Widerspruch nehmen wir an, dass es einen Vektor $v \in V$ gibt, der nicht in $U + Z_\varphi(w)$ liegt. Da $\varphi^k(v) = 0 \in U + Z_\varphi(w)$ gilt, gibt es eine kleinste Zahl $j \geq 1$, sodass $\varphi^j(v)$ in $U + Z_\varphi(w)$ liegt. Es gilt also $v' := \varphi^{j-1}(v) \notin U + Z_\varphi(w)$, aber $\varphi(v') \in U + Z_\varphi(w)$. Wir schreiben $\varphi(v')$ nun als Linearkombination

$$\varphi(v') = u + \sum_{i=0}^{k-1} a_i \varphi^i(w)$$

mit $u \in U$ und $a_i \in K$. Damit erhalten wir

$$\begin{aligned} 0 = \varphi^k(v') &= \varphi^{k-1}\big(u + \sum_{i=0}^{k-1} a_i \varphi^i(w)\big) \\ &= \varphi^{k-1}(u) + \sum_{i=0}^{k-1} a_i \varphi^{i+k-1}(w) \\ &= \underbrace{\varphi^{k-1}(u)}_{\in U} + \underbrace{a_0 \varphi^{k-1}(w)}_{\in Z_\varphi(w)}. \qquad (\varphi^{i+k-1}(w) = 0 \text{ für } i \geq 1) \end{aligned}$$

Die Summe $U + Z_\varphi(w)$ ist direkt und $\varphi^{k-1}(w) \neq 0$ (w ist ein Vektor der Stufe k), also ist $a_0 = 0$. Wir setzen nun

$$x := v' - \sum_{i=1}^{k-1} a_i \varphi^{i-1}(w).$$

Da v' nicht in $U + Z_\varphi(w)$ liegt, kann auch x nicht in diesem Unterraum liegen. Außerdem gilt

$$\begin{aligned} \varphi(x) &= \varphi\Big(v' - \sum_{i=1}^{k-1} a_i \varphi^{i-1}(w)\Big) \\ &= \varphi(v') - \sum_{i=1}^{k-1} a_i \varphi^i(w) = u \end{aligned}$$

Man kann nun zeigen, dass der Unterraum $U' = \langle x \rangle + U$ wieder φ-invariant ist und $U' \cap Z_\varphi(w) = \{0\}$ erfüllt (Übungsaufgabe!). Dies widerspricht der Maximalität von U mit dieser Eigenschaft und wir sind fertig. □ **L**

6.2.26 **Bemerkung**. Im ersten Schritt im Beweis von Satz 6.2.25 wählt man einen Vektor w der Stufe $k = \mathrm{Ni}(\varphi)$. Diese Wahl ist nicht eindeutig! In der Tat sind die Vektoren $w_1, \ldots, w_p$ und die zyklischen Unterräume in Satz 6.2.25 nicht eindeutig. Wir werden aber noch feststellen, dass die Dimensionen der auftretenden zyklischen Unterräume – also nach Lemma 6.2.23 die Stufen der Vektoren w_i – bis auf die Reihenfolge eindeutig sind.

III. Einschub: Partitionen

Um die nilpotenten Endomorphismen noch etwas besser beschreiben zu können, benötigen wir den Begriff der „Partition“. In diesem kurzen Einschub tragen wir alles zu Partitionen zusammen, was wir im Anschluss benötigen.

6.2.27 **Definition** Sei $n \in \mathbb{N}$. Eine *Partition* p von n ist ein Tupel von natürlichen Zahlen $p = (p_1, \ldots, p_k)$ mit

$$n = \sum_{i=1}^{k} p_i \quad \text{und} \quad p_1 \geq p_2 \geq \cdots \geq p_k > 0.$$

6.2.28 **Beispiel**. (a) $(5, 2, 2, 1)$ ist eine Partition von 10.

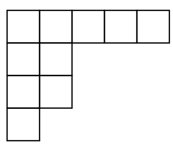
(b) Die Zahl 4 besitzt genau fünf verschiedene Partitionen:

$$(4), (3, 1), (2, 2), (2, 1, 1), (1, 1, 1, 1).$$

6.2.29 **Aufgabe**. Bestimmen Sie alle Partitionen der Zahl 5. **L**

6.2.30 **Das Diagramm einer Partition**. Gelegentlich wird es hilfreich sein, eine Partition mit einem *Diagramm* zu veranschaulichen. Ist $p = (p_1, \ldots, p_r)$ eine Partition von n, dann besteht das Diagramm zu p aus n Kästchen, die in r Zeilen angeordnet sind. In Zeile i werden dabei genau p_i Kästchen abgebildet.

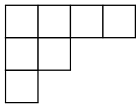
Das Diagramm zur Partition $(5, 2, 2, 1)$ ist:



Ist ein Diagramm gegeben, dann kann man die zugehörige Partition ermitteln, indem man die Kästchen in jeder Zeile zählt.

6.2.31 Aufgabe. (a) Zeichnen Sie das Diagramm zur Partition $(4, 4, 4, 1)$. **L**

(b) Bestimmen Sie die Partition zu folgendem Diagramm:



6.2.32 Definition und Satz. *Sei $p = (p_1, \dots, p_r)$ eine Partition von $n \in \mathbb{N}$. Für $j \le p_1$ sei $p_j^* \in \{1, \dots, r\}$ die Anzahl der Indizes $i \in \{1, \dots, r\}$, die $p_i \ge j$ erfüllen. Dann ist $p^* = (p_1^*, \dots, p_{p_1}^*)$ eine Partition von n.*

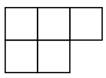
Man nennt p^* die duale Partition zu p.

<u>*Beweis*</u>. Offensichtlich ist $p_j^* \ge p_{j+1}^*$, denn wenn $p_i \ge j+1$ gilt, dann gilt auch $p_i \ge j$. Es bleibt zu zeigen, dass die Summe über die Zahlen p_j^* wieder n ergibt. Das sieht man anhand folgender Rechnung:

$$\begin{aligned} n &= \sum_{i=1}^{r} p_i = \sum_{i=1}^{r} \sum_{j=1}^{p_i} 1 \\ &= \sum_{j=1}^{p_1} |\{i \in \{1, \dots, r\} \mid p_i \ge j\}| && \text{(Vertauschen der Summen)} \\ &= \sum_{j=1}^{p_1} p_j^*. && \square \end{aligned}$$

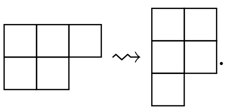
6.2.33 Beispiel. Wir betrachten die Partition $p = (3, 2)$. Die duale Partition ist $p^* = (2, 2, 1)$, denn $p_1, p_2 \ge 2 \ge 1$ aber nur p_1 erfüllt $p_1 \ge 3$.

6.2.34 Die duale Partition am Diagramm ablesen. Sei $p = (p_1, \dots, p_r)$ eine Partition von n. Die duale Partition kann man gut am Diagramm ablesen, denn p_i^* ist genau die Anzahl der Kästchen in der i-ten Spalte des Diagramms. Das Diagramm



der Partition $(3, 2)$ hat 2 Kästchen in der ersten, 2 Kästchen in der zweiten und 1 Kästchen in der dritten Spalte. Das heißt, die duale Partition zu $(3, 2)$ ist $(2, 2, 1)$.

In anderen Worten: Das Diagramm der dualen Partition p^* entsteht aus dem Diagramm der Partition p durch das „Transponieren" (Vertauschen von Zeilen und Spalten):



6.2.35 **Aufgabe**. Bestimmen Sie die duale Partition zu $p = (5, 4, 4, 1)$. **L**

6.2.36 **Lemma.** *Für jede Partition p gilt $p = (p^*)^*$.*

Beweis. Das sieht man am besten, wenn man an das Diagramm zur Partition p denkt: Wenn man das Diagramm zweimal transponiert, dann erhält man wieder das Ausgangsdiagramm. □

6.2.37 **Summe von Partitionen**. Sei $p = (p_1, \ldots, p_r)$ eine Partition von n und sei $q = (q_1, \ldots, q_s)$ eine Partition von m. Wir setzen $p_i = 0$ für alle $i > r$ und $q_i = 0$ für alle $i > s$. Es sei $t = \max(r, s)$. Dann ist

$$p + q \coloneqq (p_1 + q_1, p_2 + q_2, \ldots, p_t + q_t)$$

eine Partition von $n + m$, denn es gelten

$$\sum_{i=1}^{t} p_i + q_i = \sum_{i=1}^{r} p_i + \sum_{i=1}^{s} q_i = n + m$$

und

$$p_i + q_i \geq p_{i+1} + q_{i+1}$$

für alle $i \leq t$.

6.2.38 **Beispiel**. Zum Beispiel gilt

$$(2, 2, 1) + (5, 3, 1, 1, 1) = (7, 5, 2, 1, 1).$$

6.2.39 **Bemerkung**. Die Addition „+“ von Partitionen erfüllt das Assoziativgesetz und das Kommutativgesetz, d.h., für alle Partitionen p, q, r gelten

(1) $(p + q) + r = p + (q + r)$ und

(2) $p + q = q + p$.

Das folgt direkt aus der komponentenweisen Definition der Addition, denn die Addition von natürlichen Zahlen erfüllt das Assoziativgesetz und das Kommutativgesetz. Insbesondere müssen wir beim Rechnen mit Partitionen keine Klammern setzen; siehe 1.1.16.

Die duale Partition kann man auch mithilfe der Summe von Partitionen beschreiben. Für $n \in \mathbb{N}$ schreiben wir $\mathbb{1}_n$ für die Partition $(1, 1, \ldots, 1)$ von n.

6.2.40 **Lemma.** *Es sei $p = (p_1, p_2, \ldots, p_r)$ eine Partition von n. Dann gilt*

$$p^* = \mathbb{1}_{p_1} + \mathbb{1}_{p_2} + \mathbb{1}_{p_3} + \cdots + \mathbb{1}_{p_r}.$$

Beweis. Sei $q = 1\!\!1_{p_1} + 1\!\!1_{p_2} + \cdots + 1\!\!1_{p_r}$. Was ist q_j? Die Partition $1\!\!1_{p_i}$ trägt zu q_j eine 1 bei genau dann, wenn $p_i \geq j$ gilt. Das heißt, q_j ist genau die Anzahl der $i \in \{1, 2, \ldots, r\}$, die $p_i \geq j$ erfüllen; also $q = p^*$. □

IV. Normalform

Wir wissen bereits, dass ein nilpotenter Endomorphismus immer eine Matrixdarstellung besitzt, die eine strikte obere Dreiecksmatrix ist. In diesem Abschnitt wollen wir zeigen, dass wir für nilpotente Endomorphismen immer eine sehr einfache darstellende strikte obere Dreiecksmatrix finden können: die sogenannte nilpotente Normalform. Um eine entsprechende Basis zu konstruieren, verwenden wir die Zerlegung in zyklische Unterräume aus Satz 6.2.25. Die Eindeutigkeit der Normalform werden wir dann mithilfe der Partitionen begründen.

6.2.41 **Definition und Satz.** *Sei $V \neq \{0\}$ ein endlich-dimensionaler K-Vektorraum und sei $\varphi \in \operatorname{End}(V)$ nilpotent mit $k = \operatorname{Ni}(\varphi)$. Wir schreiben $V_j := \operatorname{Ker}(\varphi^j)$ und definieren*

$$p_j(\varphi) := \dim_K(V_j) - \dim_K(V_{j-1})$$

für alle $j \in \{1, \ldots, k\}$. Dann ist $p(\varphi) = (p_1(\varphi), p_2(\varphi), \ldots, p_k(\varphi))$ eine Partition von $n = \dim_K(V)$. Man nennt $p(\varphi)$ die Rangpartition *von φ.*

Beweis. Die Gleichheit $\sum_{j=1}^{k} p_j(\varphi) = \dim_K(V)$ lassen wir als Übungsaufgabe! **L**

Sei nun $j \in \{1, \ldots, k-1\}$. Wir müssen die Ungleichung $p_j(\varphi) \geq p_{j+1}(\varphi)$ zeigen. Wir erinnern uns daran, dass V_j für jedes j ein φ-invarianter Unterraum ist (6.1.1) und daran, dass $\varphi(V_j) \subseteq V_{j-1}$ für alle $j \geq 1$ gilt (6.2.11).

Es sei $\pi\colon V_j \to V_j/V_{j-1}$ die kanonische Projektion auf den Faktorraum. Wir betrachten die lineare Abbildung $\psi\colon V_{j+1} \to V_j/V_{j-1}$, die durch $\psi = \pi \circ \varphi|_{V_{j+1}}$ definiert ist. Wir beobachten, dass der Kern von ψ genau V_j ist. In der Tat, es gilt

$$\operatorname{Ker}(\psi) = \varphi^{-1}(\pi^{-1}(\{0\})) = \varphi^{-1}(V_{j-1}) = V_j.$$

Aus der Dimensionsformel 2.4.22 für Faktorräume folgt $p_j(\varphi) = \dim_K(V_j/V_{j-1})$. Damit erhalten wir

$$\begin{aligned}
p_j(\varphi) &= \dim_K(V_j/V_{j-1}) && (2.4.22)\\
&\geq \dim_K \operatorname{Bild}(\psi) && ([\text{MG}, 7.4.5])\\
&= \dim_K(V_{j+1}) - \dim_K \operatorname{Ker}(\psi) && ([\text{MG}, 8.3.14])\\
&= \dim_K(V_{j+1}) - \dim_K(V_j) && (V_j = \operatorname{Ker}(\psi))\\
&= p_{j+1}(\varphi).
\end{aligned}$$
□

6.2.42 Bemerkungen. (a) Sei $\dim_K(V) = n$ und sei $\varphi \in \mathrm{End}(V)$ nilpotent. Aus dem Rangsatz [MG, 8.3.14] folgt,

$$\begin{aligned} p_j(\varphi) &= \dim_K\bigl(\mathrm{Ker}(\varphi^j)\bigr) - \dim_K\bigl(\mathrm{Ker}(\varphi^{j-1})\bigr) \\ &= (n - \mathrm{Rg}(\varphi^j)) - (n - \mathrm{Rg}(\varphi^{j-1})) = \mathrm{Rg}(\varphi^{j-1}) - \mathrm{Rg}(\varphi^j). \end{aligned}$$

Dies liefert eine alternative Definition für $p(\varphi)$ und erklärt den Namen „Rangpartition".

(b) Der letzte Term der Rangpartition ist $p_k(\varphi)$, wobei $k = \mathrm{Ni}(\varphi)$ der Nilpotenzindex von φ ist. Ist j grösser als der Nilpotenzindex $k = \mathrm{Ni}(\varphi)$, dann gilt $\mathrm{Ker}(\varphi^{j-1}) = \mathrm{Ker}(\varphi^j) = V$ und damit $p_j(\varphi) = 0$. Die Zahlen $p_{k+1}(\varphi), p_{k+2}(\varphi), \ldots$ usw. gehören also nicht mehr zur Partition $p(\varphi)$. Weil es in einigen Rechnungen nützlich ist, schreiben wir manchmal $p_j(\varphi)$ anstelle von 0 für $j > \mathrm{Ni}(\varphi)$.

(c) Ist $A \in \mathrm{M}_{n,n}(K)$ eine nilpotente Matrix mit $k = \mathrm{Ni}(A)$, dann schreiben wir $p_j(A)$ anstelle von $p_j(f_A)$. Das heißt,

$$p_j(A) = \dim_K \mathrm{Ker}(A^j) - \dim_K \mathrm{Ker}(A^{j-1}).$$

Wir nennen $p(A) = (p_1(A), p_2(A), \ldots, p_k(A))$ die Rangpartition von A. Da A eine $(n \times n)$-Matrix ist, ist $p(A)$ eine Partition von n.

(d) Sei $\varphi \in \mathrm{End}(V)$ nilpotent und sei $\mathcal{B}$ eine geordnete Basis von V. Aus 6.2.2 (a) wissen wir, dass ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)^j$ eine Matrixdarstellung von φ^j ist. Aus den Mathematischen Grundlagen ist bekannt [MG, 9.2.3], dass $\dim_K \mathrm{Ker}(\varphi^j) = \dim_K \mathrm{Ker}({}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi^j))$ gilt. Das heißt, φ und die Matrixdarstellung ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)$ haben dieselbe Rangpartition.

6.2.43 Beispiel. Wir betrachten nochmal die Matrix

$$A = \begin{pmatrix} -2 & 4 & -6 & 12 \\ -1 & 2 & -3 & 6 \\ 0 & 0 & -2 & 4 \\ 0 & 0 & -1 & 2 \end{pmatrix} \in \mathrm{M}_{4,4}(\mathbb{Q})$$

aus Aufgabe 6.2.4. Dort haben wir gesehen, dass A nilpotent mit $\mathrm{Ni}(A) = 2$ ist. Wir bestimmen die Rangpartition. Wegen $A^2 = 0$ gilt $\dim_K \mathrm{Ker}(A^2) = 4$. Wir müssen nur noch die Dimension des Kerns von A bestimmen. Ziehen wir die zweite Zeile von A zweimal von der ersten ab, und ziehen wir die vierte Zeile zweimal von der dritten ab, dann erhalten wir nach Vertauschen der Zeilen

$$\begin{pmatrix} -2 & 4 & -6 & 12 \\ -1 & 2 & -3 & 6 \\ 0 & 0 & -2 & 4 \\ 0 & 0 & -1 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} -1 & 2 & -3 & 6 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Insbesondere hat A den Rang 2 und $\dim_K \operatorname{Ker}(A) = 4 - \operatorname{Rg}(A) = 2$. Die Rangpartition von A ist

$$p(A) = (2, 4-2) = (2, 2).$$

6.2.44 **Aufgabe**. Bestimmen Sie die Rangpartition der Matrix **L**

$$B = \begin{pmatrix} \overline{2} & \overline{2} & \overline{0} \\ \overline{2} & \overline{0} & \overline{2} \\ \overline{0} & \overline{2} & \overline{1} \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{F}_3)$$

aus Aufgabe 6.2.5 und 6.2.14.

6.2.45 **Beispiel**. Sei $n \in \mathbb{N}$. Wir betrachten die Matrix $N(n) = (a_{i,j})_{i,j} \in \mathrm{M}_{n,n}(K)$ mit $a_{i,j} = 1$ für $j = i+1$ und $a_{i,j} = 0$ in allen anderen Fällen. Für großes n kann man sich das vorstellen als

$$N(n) = \begin{pmatrix} 0 & 1 & 0 & 0 & & \cdots & 0 \\ 0 & 0 & 1 & 0 & & & \\ 0 & 0 & 0 & 1 & & & \vdots \\ & & & \ddots & \ddots & \ddots & \\ \vdots & & & & 0 & 1 & 0 \\ & & & & 0 & 0 & 1 \\ 0 & & \cdots & & 0 & 0 & 0 \end{pmatrix}.$$

Was passiert, wenn man $N(n)$ mit dem Standardbasisvektor e_i multipliziert? Es gilt sicherlich $N(n)e_1 = 0$. Für $i \geq 2$ wird e_i auf den Vektor e_{i-1} abgebildet, d.h. $N(n)e_i = e_{i-1}$. Wendet man $N(n)$ zweimal auf den Basisvektor e_i an, dann gilt $N(n)^2 e_i = N(n)e_{i-1} = e_{i-2}$ für $i \geq 3$ und die Vektoren e_1, e_2 liegen im Kern von $N(n)^2$. Allgemein erhalten wir

$$N(n)^k e_i = \begin{cases} 0 & \text{für } i \leq k \\ e_{i-k} & \text{für } i > k \end{cases}.$$

Insbesondere gilt $N(n)^n = 0$ und $N(n)$ ist nilpotent. Der Nilpotenzindex von $N(n)$ ist n, denn $N(n)^{n-1}e_n = e_1$, d.h., $N(n)^{n-1} \neq 0$.

Was ist die Rangpartition von $N(n)$? Der Kern von $N(n)^k$ ist

$$\operatorname{Ker}(N(n)^k) = \langle e_1, \ldots, e_k \rangle$$

und hat somit die Dimension k. Die Rangpartition von $N(n)$ ist

$$p(N(n)) = (1, 2-1, 3-2, \ldots, n-(n-1)) = \mathbb{1}_n.$$

6.2.46 **Beispiel**. Sei V ein K-Vektorraum der Dimension n und sei $\varphi \in \operatorname{End}(V)$ nilpotent. Wir nehmen an, dass $V = Z_\varphi(w)$ für einen Vektor $w \in V$ gilt. Dann ist $p(\varphi) = \mathbb{1}_n$.

Das folgt aus Beispiel 6.2.45 und 6.2.42(d), denn $N(n)$ ist eine Matrixdarstellung von φ; siehe 6.2.24.

6.2.47 **Lemma.** *Es sei V ein endlich dimensionaler K-Vektorraum und es sei $\varphi \in \operatorname{End}(V)$. Wir nehmen an, V sei die direkte Summe*

$$V = \bigoplus_{i=1}^{k} W_i$$

von φ-invarianten Unterräumen W_i. Dann ist φ genau dann nilpotent, wenn die Einschränkung $\varphi|_{W_i}$ für alle i nilpotent ist. In diesem Fall gelten

(i) $\operatorname{Ni}(\varphi) = \max\{\operatorname{Ni}(\varphi|_{W_i}) \mid 1 \leq i \leq k\}$ *und*

(ii) $p(\varphi) = p(\varphi|_{W_1}) + p(\varphi|_{W_2}) + \cdots + p(\varphi|_{W_k})$.

Beweis. Aussage (i) lassen wir als Übungsaufgabe. **L**

Sei $j \in \mathbb{N}$. Da W_i invariant unter φ ist, ist W_i auch invariant unter φ^j; siehe 5.3.28. Aus Aufgabe 2.3.13 erhalten wir $\operatorname{Ker}(\varphi^j) = \bigoplus_{i=1}^k \operatorname{Ker}(\varphi^j|_{W_i})$.

Aus Satz 2.3.12 schließen wir nun

$$\begin{aligned} p_j(\varphi) &= \dim_K \operatorname{Ker}(\varphi^j) - \dim_K \operatorname{Ker}(\varphi^{j-1}) \\ &= \sum_{i=1}^{k} \dim_K \operatorname{Ker}(\varphi^j|_{W_i}) - \dim_K \operatorname{Ker}(\varphi^{j-1}|_{W_i}) \\ &= \sum_{i=1}^{k} p_j(\varphi|_{W_i}) \end{aligned}$$

(wobei wir $p_j(\psi) = 0$ setzen, falls j größer als der Nilpotenzindex von ψ ist). Das heißt, $p(\varphi) = \sum_{i=1}^k p(\varphi|_{W_i})$. □

6.2.48 **Korollar.** *Sei V ein K-Vektorraum mit $\dim_K V = n$ und sei φ in $\operatorname{End}(V)$ nilpotent. Ist V die direkte Summe zyklischer Unterräume*

$$V = \bigoplus_{i=1}^{s} Z_\varphi(w_i)$$

der Dimensionen $q_i = \dim_K Z_\varphi(w_i)$ mit $q_1 \geq q_2 \geq \cdots \geq q_s$, dann ist $q = (q_1, \ldots, q_s)$ eine Partition von n und $q^* = p(\varphi)$.

Beweis. Es gilt
$$n = \dim_K V = \sum_{i=1}^{s} \dim_K Z_\varphi(w_i) = \sum_{i=1}^{s} q_i,$$
also ist q eine Partition von n. Aus Lemma 6.2.47 folgt nun
$$\begin{aligned} p(\varphi) &= p(\varphi|_{Z_\varphi(w_1)}) + \cdots + p(\varphi|_{Z_\varphi(w_s)}) && (6.2.47)\\ &= \mathbb{1}_{q_1} + \mathbb{1}_{q_2} + \cdots + \mathbb{1}_{q_s} && (6.2.46)\\ &= q^*. && (6.2.40) \ \square \end{aligned}$$

6.2.49 **Definition** Sei K ein Körper. Sei $n \in \mathbb{N}$ und sei $q = (q_1, \ldots, q_s)$ eine Partition von n. Die $(n \times n)$-Blockdiagonalmatrix
$$\mathcal{N}(q) = \operatorname{diag}(N(q_1), N(q_2), \ldots, N(q_s)) = \begin{pmatrix} N(q_1) & & & \\ & N(q_2) & & \\ & & \ddots & \\ & & & N(q_s) \end{pmatrix}$$
nennen wir *nilpotente Normalform* zur Partition q. Aus Korollar 6.2.48 folgt nun, dass die Rangpartition von $\mathcal{N}(q)$ genau die duale Partition q^* ist.

6.2.50 **Beispiel**. Es ist hilfreich, die Aussage von Korollars 6.2.48 einmal an einem Beispiel nachzurechnen. Es sei $q = (3, 2)$. Wir betrachten die Matrix
$$\mathcal{N}(q) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0\\ 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \in \mathrm{M}_{5,5}(K).$$
Diese Matrix ist eine obere Dreiecksmatrix, die den Rang 3 hat. Wir können den Kern von $\mathcal{N}(q)$ daran ablesen und erhalten $\langle e_1, e_4\rangle$. Wir berechnen $\mathcal{N}(q)^2$:
$$\mathcal{N}(q)^2 = \begin{pmatrix} 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$
Damit hat $\mathcal{N}(q)^2$ den Rang 1 und $\operatorname{Ker}(\mathcal{N}(q)^2) = \langle e_1, e_2, e_4, e_5\rangle$. Schließlich stellen wir fest, dass $\mathcal{N}(q)^3 = 0$ gilt. Die Rangpartition ist damit
$$p(\mathcal{N}(q)) = (2, 4-2, 5-4) = (2, 2, 1).$$

Dies ist die duale Partition von $q = (3, 2)$; siehe 6.2.34. Die Situation kann man sich gut mit dem Diagramm zur Partition q veranschaulichen:

$$\begin{array}{|c|c|c|} \hline e_1 & e_2 & e_3 \\ \hline e_4 & e_5 & \\ \cline{1-2} \end{array}$$

Multiplikation mit der Matrix $\mathcal{N}(q)$ verschiebt die Vektoren in jeder Zeile von rechts nach links. Die Vektoren in der ersten Spalte von links sind die Vektoren im Kern von $\mathcal{N}(q)$. Die Vektoren in den ersten beiden Spalten liegen im Kern von $\mathcal{N}(q)^2$. Bezeichnet man die Vektoren anhand ihrer Position im Diagramm, dann kann man die Multiplikation mit $\mathcal{N}(q)$ relativ einfach beschreiben. Wir setzen also $v_{1,1} = e_1$, $v_{1,2} = e_2$, $v_{1,3} = e_3$, $v_{2,1} = v_4$ und $v_{2,2} = e_5$. Dann gilt

$$\mathcal{N}(q)v_{i,j} = \begin{cases} v_{i,j-1} & \text{falls } j \neq 1 \\ 0 & \text{sonst} \end{cases}.$$

6.2.51 **Satz von der nilpotenten Normalform.** *Es sei V ein endlich-dimensionaler K-Vektorraum. Ist $\varphi \in \mathrm{End}(V)$ nilpotent, dann gibt es eine geordnete Basis $\mathcal{B}$ von V, sodass*

$$_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi) = \mathcal{N}(q)$$

gilt. Dabei ist $q = p(\varphi)^*$ die duale Partition zur Rangpartition von φ und q_1 ist der Nilpotenzindex von φ.

Beweis. Nach Satz 6.2.25 ist $V = \bigoplus_{i=1}^{p_1} Z_\varphi(w_i)$ die direkte Summe zyklischer Unterräume. Dabei ist $p_1 = p_1(\varphi) = \dim_K V_1$. Es sei $q_i = \dim_K Z_\varphi(w_i)$. Wir können durch Umordnen annehmen, dass $q_1 \geq q_2 \geq \cdots \geq q_s$ gilt. Dann ist $q = (q_1, \ldots, q_s)$ eine Partition von $n = \dim_K V$ und die Rangpartition ist $p(\varphi) = q^*$; siehe 6.2.48. Aus Lemma 6.2.36 schließen wir nun $q = (q^*)^* = p(\varphi)^*$. Der Nilpotenzindex k ist wegen Lemma 6.2.47 das Maximum der Zahlen $q_1, \ldots, q_s$, also $k = q_1$. Wir definieren $v_{i,j} = \varphi^{q_i - j}(w_i)$ für $i \leq p_1$ und $1 \leq j \leq q_i$. Aus Lemma 6.2.23 wissen wir, dass

$$\mathcal{B}_i = (v_{i,1}, v_{i,2}, \ldots, v_{i,q_i-1}, v_{i,q_i})$$

eine geordnete Basis von $Z_\varphi(w_i)$ ist. Also ist

$$\mathcal{B} = (v_{1,1}, \ldots, v_{1,q_1}, v_{2,1}, \ldots, v_{2,q_2}, \ldots, v_{p_1,1}, \ldots, v_{p_1,q_{p_1}})$$

eine geordnete Basis von V; siehe 2.3.12. Es gilt nun

$$\varphi(v_{i,j}) = \begin{cases} v_{i,j-1} & \text{falls } j \neq 1 \\ 0 & \text{falls } j = 1 \end{cases}$$

und (wie in Beispiel 6.2.24) folgt daraus $_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi) = \mathcal{N}(q)$. □

6.2.52 **Korollar.** *Sei K ein Körper und es sei $A \in \mathrm{M}_{n,n}(K)$ nilpotent. Dann gibt es eine invertierbare Matrix $S \in \mathrm{GL}_n(K)$, sodass*

$$S^{-1}AS = \mathcal{N}(q)$$

gilt, wobei $q = p(A)^*$ die duale Partition der Rangpartition ist.

Beweis. Wir wenden den Satz von der nilpotenten Normalform auf den Endomorphismus $f_A \colon K^n \to K^n$ an. Es gibt also eine geordnete Basis $\mathcal{B}$ von K^n mit ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(f_A) = \mathcal{N}(q)$. Es sei $\mathcal{S}$ die Standardbasis von K^n. Wir setzen $S = {}_{\mathcal{S}}\mathrm{M}_{\mathcal{B}}(\mathrm{id}_{K^n})$. Mit der Transformationsformel für Endomorphismen 4.4.5 erhalten wir

$$\mathcal{N}(q) = {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(f_A) = S^{-1}{}_{\mathcal{S}}\mathrm{M}_{\mathcal{S}}(f_A)S \overset{2.2.24}{=} S^{-1}AS. \qquad \square$$

6.2.53 **Aufgabe**. Wir betrachten die Matrix $B \in \mathrm{M}_{3,3}(\mathbb{F}_3)$ aus den Aufgaben 6.2.5, 6.2.14 und 6.2.44. Geben Sie eine Matrix in nilpotenter Normalform an, die ähnlich zu B ist. **L**

6.2.54 **Korollar.** *Zwei nilpotente Matrizen $A, B \in \mathrm{M}_{n,n}(K)$ sind genau dann ähnlich, wenn sie dieselbe Rangpartition haben.*

Beweis. Angenommen A und B sind ähnlich. Dann gibt es eine invertierbare Matrix $S \in \mathrm{GL}_n(K)$ mit $S^{-1}AS = B$. Durch Potenzieren gilt auch

$$B^j = (S^{-1}AS)^j = S^{-1}ASS^{-1}AS \cdots S^{-1}AS = S^{-1}A^jS$$

Also gilt

$$\begin{aligned}
\mathrm{Ker}(B^j) &= \{x \in K^n \mid B^jx = 0\} \\
&= \{x \in K^n \mid A^jSx = 0\} && (S \text{ invertierbar}) \\
&= \{S^{-1}y \in K^n \mid A^jy = 0\} && (\text{setze } y = Sx) \\
&= S^{-1}\mathrm{Ker}(A^j).
\end{aligned}$$

Da S invertierbar ist, ist die Abbildung $f_{S^{-1}} \colon K^n \to K^n$ ein Isomorphismus und es gilt

$$\dim_K \mathrm{Ker}(B^j) = \dim_K S^{-1}\mathrm{Ker}(A^j) = \dim_K \mathrm{Ker}(A^j).$$

Daraus folgt $p(B) = p(A)$.

Nehmen wir nun umgekehrt an, dass $p(A) = p(B)$ gilt. Sei $q = p(A)^* = p(B)^*$ die duale Partition. Nach Korollar 6.2.52 sind sowohl A als auch B ähnlich zur nilpotenten Normalform $\mathcal{N}(q)$. Da Ähnlichkeit eine Äquivalenzrelation ist (siehe 4.2.25), folgt daraus, dass A und B ähnlich sind. $\square$

Man könnte dieses Korollar auch so formulieren: Die Ähnlichkeit auf der Menge der nilpotenten Matrizen ist die Faserrelation der Rangpartitionsabbildung.

6.2.55 **Bemerkung**. Da es für jede Partition p von n eine Matrix mit Rangpartition p gibt (nämlich $\mathcal{N}(p^*)$), lässt Korollar 6.2.54 eine überraschende Schlussfolgerung zu: Die Anzahl der Ähnlichkeitsklassen nilpotenter Matrizen in $\mathrm{M}_{n,n}(K)$ ist genau die Anzahl $P(n)$ der Partitionen von n. Insbesondere hängt diese Anzahl gar nicht vom Körper K ab!

Die Funktion $P(n)$ wächst schnell mit n an. Es gilt beispielsweise $P(6) = 11, P(7) = 15, P(8) = 22, P(9) = 30, P(10) = 42, \ldots$. Es ist ziemlich schwierig eine geschlossene Formel für die Funktion $P(n)$ anzugeben. Ein bekanntes Ergebnis der Mathematiker Hardy[2] und Ramanujan[3] besagt, dass sich $P(n)$ für große Werte von n wie die Funktion

$$\frac{1}{4n\sqrt{3}} \exp\left(\pi\sqrt{\frac{2n}{3}}\right)$$

verhält.

6.2.56 **Wie berechnet man die Normalform einer nilpotenten Matrix**? Zum Abschluss wollen wir noch die Frage beantworten, wie man die Normalform einer nilpotenten Matrix oder eines nilpotenten Endomorphismus eigentlich berechnet. Der Beweis von Satz 6.2.25 verrät uns leider nur, dass wir Vektoren der höchsten möglichen Stufe wählen müssen. Wenn man das genauer aufdröselt (was wir hier nicht tun wollen), findet man folgendes Verfahren zur Berechnung der nilpotenten Normalform.

`Gegeben:` nilpotente Matrix $A \in \mathrm{M}_{n,n}(K)$

`Gesucht:` $S \in \mathrm{GL}_n(K)$, sodass $S^{-1}AS$ in Normalform ist.

`Verfahren:`

Schritt 1: Berechne $A, A^2, A^3, \ldots, A^{k-1}$ bis $A^k = 0$ gilt. Dann ist $k = \mathrm{Ni}(A)$.

Schritt 2: Berechne die Unterräume $V_j = \mathrm{Ker}(A^j) \subseteq K^n$ für $j \in \{1, 2, \ldots, k\}$. Dann ist

$$\{0\} = V_0 \subsetneq V_1 \subsetneq V_2 \subsetneq \ldots V_{k-1} \subsetneq V_k = K^n$$

die zu A gehörige Filtrierung.

Schritt 3: Berechne die Rangpartition von A

$$p = (\dim_K V_1 - \dim_K V_0, \dim_K V_2 - \dim_K V_1, \ldots, \dim_K V_k - \dim_K V_{k-1})$$

[2] Godfrey Harold Hardy: britischer Mathematiker, 1877–1947
[3] Srinivasa Ramanujan: indischer Mathematiker, 1887–1920

und die duale Partition $q = p^*$.

Zeichne das Diagramm zur dualen Partition q. (Für jedes Kästchen im Diagramm

$v_{1,1}$	$v_{1,2}$	$v_{1,3}$	$v_{1,4}$
$v_{2,1}$	$v_{2,2}$	$v_{2,3}$	$v_{2,4}$
$v_{3,1}$	$v_{3,2}$	$v_{3,3}$	
$v_{4,1}$			

$A^3v_{1,4}$	$A^2v_{1,4}$	$Av_{1,4}$	$v_{1,4}$
$A^3v_{2,4}$	$A^2v_{2,4}$	$Av_{2,4}$	$v_{2,4}$
$A^2v_{3,3}$	$Av_{3,3}$	$v_{3,3}$	
$v_{4,1}$			

Abbildung 6.3.: Beispieldiagramm mit $q = (4, 4, 3, 1)$
Links: Bezeichnung der Basisvektoren
Rechts: Gelb markierte Vektoren müssen gewählt werden

suchen wir nun einen Vektor $v_{i,j} \in K^n$. Das Diagramm wird in den nächsten Schritten von rechts nach links mit Vektoren gefüllt. Am Ende bilden die Vektoren $v_{i,j}$ eine Basis von K^n.)

Schritt 4: Wähle linear unabhängige Vektoren $v_{1,k}, \dots, v_{p_k,k} \in K^n \setminus V_{k-1}$, die ein Komplement von V_{k-1} in V_k erzeugen.

Setze $j = k - 1$. (Die nächsten beiden Schritte werden wiederholt bis $j = 0$ gilt.)

Schritt 5 (Nach links „verschieben“): Definiere $v_{i,j} = Av_{i,j+1}$ für alle $i \leq p_{j+1}$.

Schritt 6 (Zur Basis ergänzen): Wähle $v_{i,j} \in V_j \setminus V_{j-1}$ für alle $i \in \{p_{j+1} + 1, \dots, p_j\}$ dergestalt, dass $v_{1,j}, \dots, v_{p_j,j}$ linear unabhängig sind und ein Komplement zu V_{j-1} in V_j erzeugen.

Ist $j > 1$, dann setze $j \to j - 1$ und gehe zu Schritt 5.

Schritt 7: Die gesuchte Matrix S erhält man, indem man die gefundenen Vektoren $v_{i,j}$ in der Reihenfolge $v_{1,1}, v_{1,2}, v_{1,3}, \dots, v_{1,q_1}, v_{2,1}, \dots, v_{p_1,1}, \dots, v_{p_1,q_{p_1}}$ (d.h., Zeilen des Diagramms folgend) als Spalten in eine Matrix einträgt.

Dann gilt $S^{-1}AS = \mathcal{N}(q)$.

6.2.57 **Bemerkungen**. (a) Möchte man nur wissen, zu welcher Matrix $\mathcal{N}(q)$ in nilpotenter Normalform die Matrix A ähnlich ist, dann kann man das Verfahren nach dem dritten Schritt abbrechen.

(b) Die Matrix S, die man am Ende des Verfahrens erhält, ist invertierbar, d.h., die Spalten $v_{i,j}$ bilden eine Basis. Ist dies nicht der Fall, dann hat man irgendwo einen Fehler gemacht.

(c) Die Schritte 4 und 6 sind die unangenehmen Teile dieses Verfahrens, weil man die Vektoren wirklich *wählen* muss. Das sind die Vektoren zu den gelben Feldern in Abbildung 6.3. Die übrigen Vektoren muss man aus den gewählten Vektoren berechnen. In Schritt 6 ist es mühsam zu prüfen, ob die Vektoren $v_{1,j}, \dots, v_{p_j,j}$ linear unabhängig sind und ein Komplement zu V_{j-1} in V_j aufspannen. Wird diese Bedingung verletzt, dann bilden die Vektoren $v_{i,j}$ am Ende keine Basis. Man kann sich also etwas Arbeit sparen, indem man „auf Risiko spielt" und erst ganz am Ende prüft, ob die gefundenen Vektoren eine Basis bilden.

Das Verfahren hört sich komplizierter an, als es ist. Sehen wir uns dazu einige Beispiele an.

6.2.58 **Beispiel**. Wir betrachten nochmal die Matrix

$$A = \begin{pmatrix} -2 & 4 & -6 & 12 \\ -1 & 2 & -3 & 6 \\ 0 & 0 & -2 & 4 \\ 0 & 0 & -1 & 2 \end{pmatrix} \in \mathrm{M}_{4,4}(\mathbb{Q})$$

aus Aufgabe 6.2.4. Wir wissen, dass A nilpotent ist, da $A^2 = 0$ gilt. Der Nilpotenzindex ist $k = 2$. Wir berechnen den Kern von A und erhalten

$$V_1 = \mathrm{Ker}(A) = \left\langle \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 2 \\ 1 \end{pmatrix} \right\rangle.$$

Es gilt $V_0 = \{0\}$ und $V_2 = \mathbb{Q}^4$. Die Rangpartition ist demnach $p(A) = (2, 2)$. Die duale Partition dazu ist $q = p(A)^* = (2, 2)$.

Wir zeichnen das Diagramm zur dualen Partition q

$v_{1,1}$	$v_{1,2}$
$v_{2,1}$	$v_{2,2}$

und suchen die Vektoren $v_{i,j}$ von rechts nach links.

Wir wählen $v_{1,2}, v_{2,2} \in \mathbb{Q}^4 \setminus V_1$ als

$$v_{1,2} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad v_{2,2} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix},$$

denn diese Vektoren sind linear unabhängig und $\langle v_{1,2}, v_{2,2} \rangle$ ist ein Komplement zu V_1 in $\mathbb{Q}^4$.

Nun kommt Schritt 5. Wir setzen

$$v_{1,1} = Av_{1,2} = \begin{pmatrix} -2 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \quad v_{2,1} = Av_{2,2} = \begin{pmatrix} -6 \\ -3 \\ -2 \\ -1 \end{pmatrix}.$$

Schließlich müssen wir die Vektoren nur in der richtigen Reihenfolge (im Diagramm zeilenweise von links nach rechts) in die gesuchte Matrix S schreiben:

$$S = \begin{pmatrix} -2 & 1 & -6 & 0 \\ -1 & 0 & -3 & 0 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}.$$

Es gilt dann $S^{-1}AS = \mathcal{N}(2, 2)$.

6.2.59 **Beispiel**. Wir betrachten die Matrix

$$A = \begin{pmatrix} 2 & -1 & 0 & 0 & 0 \\ 6 & -3 & -1 & 0 & 0 \\ -2 & 1 & 1 & 0 & 0 \\ 9 & -3 & 3 & -1 & -1 \\ -9 & 3 & 0 & 1 & 1 \end{pmatrix} \in \mathrm{M}_{5,5}(\mathbb{R}).$$

Die Matrix ist nilpotent mit $\mathrm{Ni}(A) = 3$, denn es ist

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

und $A^3 = 0$. Wir suchen nun eine invertierbare Matrix S, sodass $S^{-1}AS$ in nilpotenter Normalform ist. Dazu bestimmen wir zuerst die Filtrierung. Es gilt sicherlich $V_0 = \{0\}$ und wegen $A^3 = 0$ auch $V_3 = \mathbb{R}^5$. Dazu findet man mit einer kurzen Rechnung

$$V_1 = \mathrm{Ker}(A) = \mathrm{Ker}\begin{pmatrix} 2 & -1 & 0 & 0 & 0 \\ 6 & -3 & -1 & 0 & 0 \\ -2 & 1 & 1 & 0 & 0 \\ 9 & -3 & 3 & -1 & -1 \\ -9 & 3 & 0 & 1 & 1 \end{pmatrix} = \left\langle \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 0 \\ 3 \\ 0 \end{pmatrix} \right\rangle,$$

$$V_2 = \mathrm{Ker}(A^2) = \mathrm{Ker}\begin{pmatrix} -2 & 1 & 1 & 0 & 0 \\ -4 & 2 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ -6 & 3 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = \left\langle \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \\ 0 \end{pmatrix} \right\rangle.$$

Die Rangpartition ist somit

$$p(A) = (2-0, 4-2, 5-4) = (2, 2, 1)$$

und die duale Partition ist $q = (3, 2)$. Wir zeichnen das Diagramm der dualen Partition

$v_{1,1}$	$v_{1,2}$	$v_{1,3}$
$v_{2,1}$	$v_{2,2}$	

und suchen die Vektoren $v_{i,j}$ von rechts nach links. Für $v_{1,3}$ suchen wir einen Vektor in $V_3 \setminus V_2$. Wir wählen $v_{1,3} = \left(\begin{smallmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{smallmatrix}\right)$. Wir „verschieben nach links" mit A und erhalten $v_{1,2} = Av_{1,3} = \left(\begin{smallmatrix} 0 \\ -1 \\ 1 \\ 3 \\ 0 \end{smallmatrix}\right)$ und $v_{1,1} = Av_{1,2} = A^2 v_{1,3} = \left(\begin{smallmatrix} 1 \\ 2 \\ 0 \\ 3 \\ 0 \end{smallmatrix}\right)$.

Jetzt kommt der knifflige 6. Schritt im Verfahren. Wir suchen einen Vektor $v_{2,2} \in V_2 \setminus V_1$, sodass $\langle v_{1,2}, v_{2,2} \rangle$ ein Komplement zu V_1 in V_2 aufspannt. Wir wählen $v_{2,2} = \left(\begin{smallmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{smallmatrix}\right)$. Hier könnten wir (etwas pedantisch) nun prüfen, dass $\langle v_{1,2}, v_{2,2} \rangle$ wirklich ein Komplement zu V_1 in V_2 ist. Wir nehmen einfach an, dies sei der Fall und spielen auf Risiko, d.h., wir rechnen erstmal weiter.

Wir verschieben mit A nach links und erhalten $v_{2,1} = Av_{2,2} = \left(\begin{smallmatrix} 0 \\ 0 \\ 0 \\ -1 \\ 1 \end{smallmatrix}\right)$. Wir bilden nun die Matrix S, indem wir die Vektoren $v_{i,j}$ aus dem Diagramm (zeilenweise von links nach rechts) als Spalten eintragen, d.h.,

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

Wenn wir keinen Fehler gemacht haben, dann ist die Matrix S invertierbar, d.h., die Vektoren $v_{i,j}$ bilden eine Basis von $\mathbb{R}^5$. Weil S eine untere Dreiecksmatrix ist, sehen wir, dass $\det(S) = 1$ gilt (siehe Satz 4.2.13) und damit ist S invertierbar (siehe Korollar 4.2.32). Also haben wir $v_{2,2}$ richtig gewählt und es gilt $S^{-1}AS = \mathcal{N}(3, 2)$.

6.2.60 **Aufgabe**. Wir betrachten nochmal die Matrix **L**

$$B = \begin{pmatrix} \overline{2} & \overline{2} & \overline{0} \\ \overline{2} & \overline{0} & \overline{2} \\ \overline{0} & \overline{2} & \overline{1} \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{F}_3)$$

aus den Aufgaben 6.2.5, 6.2.14 und 6.2.44. Bestimmen Sie eine invertierbare Matrix $S \in \mathrm{GL}_3(\mathbb{F}_3)$, sodass $S^{-1}BS$ in nilpotenter Normalform ist.

6.2.61 **Aufgabe**. Zeigen Sie, dass die Matrix **L**

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & -2 & 1 & -1 \\ 0 & -1 & 0 & 0 \\ -2 & 4 & -2 & 2 \end{pmatrix} \in \mathrm{M}_{4,4}(\mathbb{R})$$

nilpotent ist und bestimmen Sie eine invertierbare Matrix $S \in \mathrm{GL}_4(\mathbb{R})$, sodass $S^{-1}AS$ in nilpotenter Normalform ist.

6.3. Die Jordan'sche Normalform

Wir sind jetzt in der Lage, unser Wissen über diagonalisierbare und nilpotente Endomorphismen zu kombinieren und eine Normalform für allgemeine Endomorphismen anzugeben: die *Jordan'sche Normalform*. Dabei arbeiten wir unter der Voraussetzung, dass das charakteristische Polynom in Linearfaktoren zerfällt. Über einem algebraisch abgeschlossenem Körper (z.B. $\mathbb{C}$) ist diese Voraussetzung immer erfüllt. Zunächst müssen wir klären, welche Matrizen in „Jordan'scher Normalform“ sind.

I. Was ist die Jordan'sche Normalform?

6.3.1 **Definition** Sei K ein Körper und sei $\lambda \in K$. Ist q eine Partition von n, dann nennen wir die Matrix

$$J(\lambda, q) = \lambda I_n + \mathcal{N}(q) \in \mathrm{M}_{n,n}(K)$$

einen *Jordanblock* zum Eigenwert λ und zur Partition q. Die Zahl n nennen wir die Größe des Jordanblocks.

6.3.2 **Anmerkungen**. (a) Sei $q = (q_1, q_2, \dots, q_s)$ eine Partition. Da $\mathcal{N}(q)$ eine Blockdiagonalmatrix ist (siehe 6.2.49), ist auch jeder Jordanblock $J(\lambda, q)$ eine Blockdiagonalmatrix der Form

$$J(\lambda, q) = \mathrm{diag}(\lambda I_{q_1} + N(q_1), \lambda I_{q_2} + N(q_2), \dots, \lambda I_{q_s} + N(q_s)).$$

Diese „kleinen“ Blöcke auf der Diagonalen nennen wir *Jordankästchen*. Die Zahlen $q_1, \dots, q_s$ sind genau die Größen der Jordankästchen.

(b) Ein Jordanblock $J(\lambda, q)$ der Größe n ist eine obere Dreiecksmatrix. Alle Diagonaleinträge sind gleich λ, also gilt $\chi_{J(\lambda,q)} = (X - \lambda)^n$; siehe Satz 4.2.13. Damit ist λ der einzige Eigenwert von $J(\lambda, q)$.

6.3.3 **Beispiel**. (a) Die Matrix

$$J(7, (2, 1, 1)) = 7I_4 + \mathcal{N}(2, 1, 1) = \left(\begin{array}{cc|c|c} 7 & 1 & 0 & 0 \\ 0 & 7 & 0 & 0 \\ \hline 0 & 0 & 7 & 0 \\ \hline 0 & 0 & 0 & 7 \end{array}\right) \in \mathrm{M}_{4,4}(\mathbb{Q})$$

ist ein Jordanblock zum Eigenwert 7 und zur Partition $(2, 1, 1)$. Der Jordanblock $J(7, (2, 1, 1))$ hat Größe 4 und besitzt drei Jordankästchen.

(b) Die Matrix

$$J(i,(3,2)) = iI_5 + \mathcal{N}(3,2) = \left(\begin{array}{ccc|cc} i & 1 & 0 & 0 & 0 \\ 0 & i & 1 & 0 & 0 \\ 0 & 0 & i & 0 & 0 \\ \hline 0 & 0 & 0 & i & 1 \\ 0 & 0 & 0 & 0 & i \end{array}\right) \in \mathrm{M}_{5,5}(\mathbb{C})$$

ist ein Jordanblock zum Eigenwert i und zur Partition $(3,2)$ mit zwei Jordankästchen.

6.3.4 **Aufgabe**. Schreiben Sie den Jordanblock $J(4,(3,1,1))$ auf. Wie viele Jordankästchen gibt es? **L**

6.3.5 **Definition** Sei K ein Körper. Eine Matrix $A \in \mathrm{M}_{n,n}(K)$ ist in *Jordan'scher Normalform*, wenn A eine Blockdiagonalmatrix der Form

$$A = \operatorname{diag}(J(\lambda_1, q^{(1)}), J(\lambda_2, q^{(2)}), \ldots, J(\lambda_r, q^{(r)})) \in \mathrm{M}_{n,n}(K)$$

ist, wobei die Eigenwerte $\lambda_1, \ldots, \lambda_r \in K$ paarweise verschieden sind. Dabei nennen wir $q^{(i)}$ die Partition zum Eigenwert λ_i.

6.3.6 **Beispiel**. (a) Die Matrix

$$\operatorname{diag}\big(J(3,(2)), J(-1,(2,1))\big) = \left(\begin{array}{cc|ccc} 3 & 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ \hline 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{array}\right) \in \mathrm{M}_{5,5}(\mathbb{R})$$

ist eine Matrix in Jordan'scher Normalform mit zwei Jordanblöcken zu den Eigenwerten 3 und -1.

(b) Die Matrix

$$\operatorname{diag}\big(J(i,(2,2)), J(\sqrt{2},(1)), J(0,(1))\big) = \left(\begin{array}{cccc|c|c} i & 1 & 0 & 0 & 0 & 0 \\ 0 & i & 0 & 0 & 0 & 0 \\ 0 & 0 & i & 1 & 0 & 0 \\ 0 & 0 & 0 & i & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & \sqrt{2} & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \in \mathrm{M}_{6,6}(\mathbb{C})$$

ist in Jordan'scher Normalform und besitzt insgesamt drei Jordanblöcke zu den Eigenwerten i, $\sqrt{2}$ und 0.

Das Ziel dieses Abschnittes ist es, den folgenden wichtigen Satz zu beweisen, der über algebraisch abgeschlossenen Körpern das in 5.1.2 vorgestellte Normalformproblem beantwortet.

6.3.7 **Hauptsatz zur Jordan'schen Normalform.** *Sei K ein Körper und sei V ein endlich-dimensionaler K-Vektorraum. Ist $\varphi \in \operatorname{End}(V)$ ein Endomorphismus, dessen charakteristisches Polynom χ_φ in Linearfaktoren zerfällt, dann gibt es eine geordnete Basis $\mathcal{B}$ von V, sodass ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)$ in Jordan'scher Normalform ist.*

Die Matrix in Jordan'scher Normalform ist bis auf die Reihenfolge der Eigenwerte $\lambda_1, \ldots, \lambda_r$ eindeutig bestimmt.

Wendet man diesen Satz auf den Endomorphismus $f_A\colon K^n \to K^n$ für eine Matrix $A \in \mathrm{M}_{n,n}(K)$ an, erhält man auch eine Lösung des Normalformproblems für Matrizen 5.1.4.

6.3.8 **Satz.** *Sei K ein Körper. Eine Matrix $A \in \mathrm{M}_{n,n}(K)$ ist genau dann ähnlich zu einer Matrix in Jordan'scher Normalform, wenn ihr charakteristisches Polynom χ_A in Linearfaktoren zerfällt.*

Die Jordan'sche Normalform ist bis auf die Reihenfolge der Jordanblöcke eindeutig bestimmt.

Eine Matrixdarstellung in Jordan'scher Normalform ist nützlich, weil man anhand dieser Darstellung alle wesentlichen Eigenschaften eines Endomorphismus direkt ablesen kann.

6.3.9 **Eigenschaften aus der Jordan'schen Normalform ablesen.** *Es sei $A = \operatorname{diag}(J(\lambda_1, q^{(1)}), J(\lambda_2, q^{(2)}), \ldots, J(\lambda_r, q^{(r)})) \in \mathrm{M}_{n,n}(K)$ eine Matrix in Jordan'scher Normalform mit paarweise verschiedenen $\lambda_1, \ldots, \lambda_r \in K$. Dann gelten folgende Aussagen:*

(i) Die Diagonaleinträge $\lambda_1, \ldots, \lambda_r$ sind die Eigenwerte von A.

(ii) Die algebraische Vielfachheit von λ_i ist die Größe des zugehörigen Jordanblocks $J(\lambda_i, q^{(i)})$.

(iii) Die geometrische Vielfachheit von λ_i ist die Anzahl der Jordankästchen im Jordanblock $J(\lambda_i, q^{(i)})$.

(iv) Die Vielfachheit der Nullstelle λ_i im Minimalpolynom μ_A ist die Größe des größten Jordankästchens im Jordanblock $J(\lambda_i, q^{(i)})$.

Wir werden uns gleich anschauen, wie man diese Aussagen beweist (siehe 6.3.12). Um die Aussagen zu verdeutlichen betrachten wir zuerst ein Beispiel.

6.3.10 Beispiel. Gegeben ist die Matrix

$$\operatorname{diag}\big(J(3,(3)), J(-1,(2,1))\big) = \left(\begin{array}{ccc|ccc} 3 & 1 & 0 & 0 & 0 & 0 \\ 0 & 3 & 1 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{array}\right) \in \mathrm{M}_{6,6}(\mathbb{R}).$$

Die Matrix ist in Jordan'scher Normalform und hat die Eigenwerte 3 und -1. Das charakteristische Polynom ist $(X-3)^3(X+1)^3$. Der Jordanblock zum Eigenwert 3 hat nur ein Kästchen. Die geometrische Vielfachheit von 3 ist also $\gamma(3) = 1$. Der Jordanblock zum Eigenwert -1 hat zwei Kästchen, also gilt $\gamma(-1) = 2$. Das größte Kästchen im Block $J(3,(3))$ ist ein 3×3-Kästchen und das größte Kästchen im Block $J(-1,(2,1))$ ist ein 2×2-Kästchen. Also ist das Minimalpolynom von A gegeben durch $\mu_A = (X-3)^3(X+1)^2$.

6.3.11 Aufgabe. Geben Sie eine reelle Matrix A in Jordan'scher Normalform mit **L**

$$\chi_A = (X-2)^3(X-5)^4 \quad \text{und} \quad \mu_A = (X-2)^2(X-5)^2$$

an.

6.3.12 Beweis der Aussagen aus 6.3.9. Es sei

$$A = \operatorname{diag}(J(\lambda_1, q^{(1)}), J(\lambda_2, q^{(2)}), \ldots, J(\lambda_r, q^{(r)})) \in \mathrm{M}_{n,n}(K)$$

eine Matrix in Jordan'scher Normalform. Dabei sei n_i die Größe des Blocks $J(\lambda_i, q^{(i)})$. Die Matrix A ist eine obere Dreiecksmatrix, deshalb können wir das charakteristische Polynom mit Satz 4.2.13 berechnen und erhalten

$$\chi_A = (X-\lambda_1)^{n_1} \cdots (X-\lambda_r)^{n_r}.$$

Aus Satz 5.3.7 schließen wir, dass $\lambda_1, \ldots, \lambda_r$ die Eigenwerte von A sind. Weiter ist n_i genau die Vielfachheit der Nullstelle λ_i in χ_A, d.h., $n_i = \alpha(\lambda_i)$ ist die algebraische Vielfachheit von λ_i. Damit haben wir die Aussagen (i) und (ii) bewiesen.

Für die Aussage (iii) betrachten wir zuerst den speziellen Fall, dass A ein Jordanblock $J(\lambda, q) = \lambda I_n + \mathcal{N}(q)$ ist. Dabei ist $q = (q_1, \ldots, q_s)$ eine Partition von n. Die Zahl s ist genau die Anzahl der Kästchen im Jordanblock. Der Eigenraum zum Eigenwert

λ ist der Kern der Matrix $J(\lambda, q) - \lambda I_n = \mathcal{N}(q)$. Da $\mathcal{N}(q)$ nilpotent ist und $p = q^*$ die Rangpartition ist, gilt $\dim_K \operatorname{Ker}\big(\mathcal{N}(q)\big) = p_1$. Aus Lemma 6.2.40 folgt nun, dass $p_1 = s$ ist. Das heißt, die geometrische Vielfachheit ist die Anzahl der Kästchen im Jordanblock.

Ist $\lambda' \neq \lambda$ und ist q' eine Partition von n, dann ist $J(\lambda', q') - \lambda I_n$ eine obere Dreiecksmatrix mit nicht verschwindenden Diagonaleinträgen, d.h., insbesondere ist $J(\lambda', q') - \lambda I_n$ invertierbar und hat einen trivialen Kern.

Kehren wir zum allgemeinen Fall zurück. Die Matrizen A und $A - \lambda_i I_n$ sind Blockdiagonalmatrizen. Aus der Diskussion aus Beispiel 5.1.11 erhalten wir daher

$$\begin{aligned}
\gamma(\lambda_i) &= \dim_K \operatorname{Ker}(A - \lambda_i I_n) \\
&= \sum_{t=1}^{r} \dim_K \operatorname{Ker}(J(\lambda_t, q^{(t)}) - \lambda_i I_{n_t}) && (5.1.11) \\
&= \dim_K \operatorname{Ker}(J(\lambda_i, q^{(i)}) - \lambda_i I_{n_i})
\end{aligned}$$

und das ist, wie wir gerade gesehen haben, genau die Anzahl der Jordankästchen im Block zum Eigenwert λ_i.

Aussage (iv) beweisen wir weiter unten in 6.3.17.

II. Beweis des Hauptsatzes

In diesem Abschnitt beweisen wir den Hauptsatz zur Jordan'schen Normalform 6.3.7. Es sei V ein endlich-dimensionaler K-Vektorraum und es sei $\varphi \in \operatorname{End}(V)$ ein Endomorphismus. Wir betrachten zuerst den Spezialfall, dass das Minimalpolynom von φ die Gestalt $\mu_\varphi = (X - \lambda)^k$ für ein $\lambda \in K$ und ein $k \in \mathbb{N}$ hat. Diesen Fall erhalten wir einfach aus der Theorie der nilpotenten Endomorphismen.

6.3.13 **Lemma.** *Es sei V ein endlich-dimensionaler K-Vektorraum und es sei $\varphi \in \operatorname{End}(V)$. Gilt $\mu_\varphi = (X - \lambda)^k$ für ein $\lambda \in K$ und ein $k \in \mathbb{N}$, dann gibt es eine geordnete Basis $\mathcal{B}$ von V, sodass ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi) = J(\lambda, q)$ ein Jordanblock zum Eigenwert λ ist. Die Partition q ist dabei eindeutig bestimmt.*

Beweis. Es sei $n = \dim_K V$. Der Endomorphismus $\psi = \varphi - \lambda\,\mathrm{id}_V$ ist nilpotent, denn es gilt

$$0 = \mu_\varphi(\varphi) = (\varphi - \lambda\,\mathrm{id}_V)^k = \psi^k.$$

Wir wenden den Satz von der nilpotenten Normalform auf ψ an. Es gibt also eine geordnete Basis $\mathcal{B}$ von V, sodass ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\psi) = \mathcal{N}(q)$ ist. Dabei ist q die duale Partition der Rangpartition von ψ. Insbesondere ist q eindeutig durch φ bestimmt. Wir erhalten

$$\mathcal{N}(q) = {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\psi) = {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi) - \lambda{}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\mathrm{id}_V) = {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi) - \lambda I_n$$

und durch Umstellen der Gleichung folgt die Behauptung. □

Den Hauptsatz können wir nun aus diesem Lemma und Korollar 6.1.10 herleiten. Dazu führen wir einen weiteren Begriff ein.

6.3.14 **Definition** Es sei V ein K-Vektorraum. Sei $\varphi \in \mathrm{End}(V)$ und sei $\lambda \in K$ ein Eigenwert von φ. Ist $\ell \in \mathbb{N}$ gegeben, dann nennen wir

$$\mathrm{Eig}_{\varphi}^{\ell}(\lambda) := U_{\varphi}((X-\lambda)^{\ell})$$

den *verallgemeinerten Eigenraum der Stufe* ℓ zum Eigenwert λ.

Mit dieser Notation ist $\mathrm{Eig}_{\varphi}^{1}(\lambda) = U_{\varphi}(X-\lambda) = \mathrm{Eig}_{\varphi}(\lambda)$ der Eigenraum zum Eigenwert λ.

Beweis von Satz 6.3.7. Es sei V ein endlich-dimensionaler K-Vektorraum und es sei $\varphi \in \mathrm{End}(V)$ ein Endomorphismus. Wir nehmen an, dass χ_{φ} in Linearfaktoren zerfällt. Da das Minimalpolynom ein Teiler des charakteristischen Polynoms ist (Satz von Cayley-Hamilton 5.3.41), zerfällt auch μ_{φ} in Linearfaktoren; siehe 1.4.39. Es seien $\lambda_1, \ldots, \lambda_r$ die paarweise verschiedenen Eigenwerte von φ in K. Dann gilt

$$\mu_{\varphi} = \prod_{i=1}^{r}(X-\lambda_i)^{k_i},$$

wobei $k_i \in \mathbb{N}$ die Vielfachheit der Nullstelle λ_i in μ_{φ} ist.

Für $i \neq j$ sind die Polynome $X-\lambda_i$ und $X-\lambda_j$ teilerfremd; das folgt aus Beispiel 1.4.56. Mit Aufgabe 1.4.58 können wir schließen, dass auch $(X-\lambda_i)^{k_i}$ und $(X-\lambda_j)^{k_j}$ teilerfremd sind. Aus Korollar 6.1.10 erhalten wir nun eine Zerlegung von V in φ-invariante Unterräume (siehe 6.1.1)

$$V = \bigoplus_{i=1}^{r} \mathrm{Eig}_{\varphi}^{k_i}(\lambda_i).$$

Dabei ist $H_i := \mathrm{Eig}_{\varphi}^{k_i}(\lambda_i)$ der verallgemeinerte Eigenraum der Stufe k_i zum Eigenwert λ_i. Aus Korollar 6.1.10 wissen wir auch, dass $\varphi|_{H_i}$ ein Endomorphismus mit Minimalpolynom $(X-\lambda_i)^{k_i}$ ist. Aus Lemma 6.3.13 erhalten wir eine geordnete

Basis $\mathcal{B}_i = (v_1^i, \ldots, v_{k_i}^i)$ von H_i, sodass ${}_{\mathcal{B}_i}\mathrm{M}_{\mathcal{B}_i}(\varphi|_{H_i}) = J(\lambda_i, q^{(i)})$ ein Jordanblock zum Eigenwert λ_i und einer Partition $q^{(i)}$ ist.

Es sei nun $\mathcal{B} = (v_1^1, \ldots, v_{k_1}^1, \ldots, v_1^r, \ldots, v_{k_r}^r)$ die geordnete Basis von V, die wir durch Zusammenfügen der Basen $\mathcal{B}_1, \ldots, \mathcal{B}_r$ erhalten. Wie in 5.1.13 folgt nun, dass ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)$ eine Blockdiagonalmatrix mit den Blöcken ${}_{\mathcal{B}_i}\mathrm{M}_{\mathcal{B}_i}(\varphi|_{H_i}) = J(\lambda_i, q^{(i)})$ ist, d.h., ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)$ ist eine Matrix in Jordan'scher Normalform.

Warum ist die Jordan'sche Normalform bis auf die Reihenfolge der Eigenwerte (d.h., bis auf die Reihenfolge der Jordanblöcke) eindeutig? Aus 6.3.9 wissen wir, dass wir die Eigenwerte, die algebraischen Vielfachheiten, die geometrischen Vielfachheiten und die Vielfachheiten der Eigenwerte im Minimalpolynom aus der Jordan'schen Normalform ablesen können. Umgekehrt legen also diese Eigenschaften von φ die Jordan'sche Normalform fast vollständig fest. Wir müssen uns nur noch überlegen, warum die Partitionen $q^{(i)}$ eindeutig bestimmt sind. Was ist $q^{(i)}$? Die Partition $q^{(i)}$ ist dual zur Rangpartition des nilpotenten Endomorphismus $(\varphi - \lambda_i \operatorname{id}_V)|_{H_i}$. Es ist hilfreich, sich klar zu machen, dass der Raum H_i eindeutig durch φ und das Minimalpolynom festgelegt ist. Nach Lemma 6.3.13 ist die Partition $q^{(i)}$ dadurch eindeutig. Etwas konkreter: Ist ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)$ irgendeine Matrixdarstellung von φ, dann können wir die duale Partition $(q^{(i)})^*$ bestimmen, indem wir die Kerne der Potenzen $({}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi) - \lambda_i I_n)^j$ berechnen. Die Differenzen der Dimensionen liefern die Rangpartition von $(\varphi - \lambda_i \operatorname{id}_V)|_{H_i}$. Insbesondere kann eine Matrix in Jordan'scher Normalform nur dann eine Matrixdarstellung von φ sein, wenn der Jordanblock zum Eigenwert λ_i genau $J(\lambda_i, q^{(i)})$ ist. In anderen Worten, die Jordanblöcke können höchstens in einer anderen Reihenfolge angeordnet werden.

Jede Umordnung der Jordanblöcke führt zu einer ähnlichen Matrix, denn diese Umordnung entsteht durch eine Umordnung der Basisvektoren, also durch einen Basiswechsel. □

Im Beweis des Hauptsatzes taucht ein wichtiges neues Konzept auf.

6.3.15 $\boxed{\textbf{Definition}}$ Sei V ein endlich-dimensionaler K-Vektorraum. Sei $\varphi \in \operatorname{End}(V)$ und sei λ ein Eigenwert von φ. Ist k die Vielfachheit von λ im Minimalpolynom μ_φ, dann nennt man den verallgemeinerten Eigenraum der Stufe k

$$H_\varphi(\lambda) := \operatorname{Eig}_\varphi^k(\lambda)$$

den *Hauptraum* zum Eigenwert λ.

Ist $A \in \mathrm{M}_{n,n}(K)$ eine Matrix, so schreiben wir $H_A(\lambda)$ für $H_{f_A}(\lambda) \subseteq K^n$ und nennen diesen Raum den Hauptraum von A zum Eigenwert λ.

III. Berechnung der Jordan'schen Normalform

Sei V ein endlich-dimensionaler K-Vektorraum und sei $\varphi \in \mathrm{End}(V)$ ein Endomorphismus, dessen charakteristisches Polynom in Linearfaktoren zerfällt. Wir wissen aus dem Hauptsatz, dass es eine Basis $\mathcal{B}$ von V gibt dergestalt, dass ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)$ eine Matrix in Jordan'scher Normalform ist. Wie kann man eine solche Basis bestimmen?

Der Beweis des Hauptsatzes gibt uns einen guten Hinweis. Wenn wir die Eigenwerte $\lambda_1, \ldots, \lambda_r$ von φ kennen, dann zerlegt sich der Vektorraum V in die direkte Summe der Haupträume, d.h.,

$$V = \bigoplus_{i=1}^{r} H_\varphi(\lambda_i).$$

Die Haupträume sind φ-invariante Unterräume (siehe Lemma 6.1.1) und die Einschränkung von $\varphi - \lambda_i \,\mathrm{id}_V$ auf $H_\varphi(\lambda_i)$ ist nilpotent. Die Einschränkung können wir also mit den Methoden aus dem Abschnitt über nilpotente Endomorphismen behandeln.

Wie berechnet man aber die Haupträume? Darüber müssen wir noch etwas nachdenken, denn es gibt dabei ein Problem: Das Minimalpolynom μ_φ kennen wir (meistens) nicht! In der Tat ist es im Allgemeinen etwa genauso schwer das Minimalpolynom zu bestimmen wie die Haupträume zu berechnen. Wir müssen also den Hauptraum zum Eigenwert λ berechnen, ohne die Vielfachheit der Nullstelle λ im Minimalpolynom zu kennen. Dabei hilft uns der nächste Satz.

6.3.16 **Satz.** *Sei V ein endlich-dimensionaler K-Vektorraum und sei $\varphi \in \mathrm{End}(V)$ ein Endomorphismus. Es sei λ ein Eigenwert von φ. Falls $\dim_K \mathrm{Eig}_\varphi^m(\lambda)$ die algebraische Vielfachheit $\alpha(\lambda)$ für ein $m \in \mathbb{N}$ ist, dann gilt*

$$H_\varphi(\lambda) = \mathrm{Eig}_\varphi^m(\lambda).$$

Ist m die kleinste Zahl mit dieser Eigenschaft, dann ist m genau die Vielfachheit von λ in μ_φ.

Beweis. Es sei k die Vielfachheit der Nullstelle λ im Minimalpolynom μ_φ. Aus dem Beweis des Hauptsatzes sehen wir, dass $\dim_K H_\varphi(\lambda)$ genau die Größe des Jordanblocks zum Eigenwert λ ist. Mit 6.3.9 (ii) können wir schließen, dass dies genau die algebraische Vielfachheit $\alpha(\lambda)$ ist.

Wir schreiben $\mu_\varphi = (X-\lambda)^k h$, dabei ist λ keine Nullstelle von h, d.h., h ist teilerfremd zu $(X-\lambda)$. Aus dem Zerlegungssatz 6.1.8 folgt

$$V = U_\varphi(\mu_\varphi) = H_\varphi(\lambda) \oplus U_\varphi(h).$$

Sei $m \in \mathbb{N}$. Wir definieren $g = (X-\lambda)^m h$. Dann folgt aus dem Zerlegungssatz
$$U_\varphi(g) = \operatorname{Eig}_\varphi^m(\lambda) \oplus U_\varphi(h).$$
Angenommen es gilt $\dim_K \operatorname{Eig}_\varphi^m(\lambda) = \alpha(\lambda)$. Mit der Dimensionsformel aus Satz 2.3.12 erhalten wir
$$\begin{aligned}\dim_K U_\varphi(g) &= \dim_K \operatorname{Eig}_\varphi^m(\lambda) + \dim_K U_\varphi(h)\\ &= \dim_K H_\varphi(\lambda) + \dim_K U_\varphi(h) = \dim_K V\end{aligned}$$
und damit $V = U_\varphi(g)$. Das heißt, g annulliert φ und μ_φ teilt g. Da k die Vielfachheit der Nullstelle λ in μ_φ ist, muss $m \geq k$ sein. Insbesondere ist k die kleinste Zahl, die $\dim_K \operatorname{Eig}_\varphi^k(\lambda) = \alpha(\lambda)$ erfüllt.

Umgekehrt sei nun $m > k$, dann teilt μ_φ das Polynom g und es gilt $g(\varphi) = 0$. Das heißt $V = U_\varphi(g)$. Wir rechnen
$$\dim_K H_\varphi(\lambda) + \dim_K U_\varphi(h) = \dim_K V = \dim_K \operatorname{Eig}_\varphi^m(\lambda) + \dim_K U_\varphi(h)$$
und schließen daraus, dass $\dim_K H_\varphi(\lambda) = \dim_K \operatorname{Eig}_\varphi^m(\lambda)$ ist. Aus Aufgabe 6.1.7 folgt aber, dass
$$H_\varphi(\lambda) = \operatorname{Eig}_\varphi^k(\lambda) \subseteq \operatorname{Eig}_\varphi^m(\lambda)$$
gilt und aus der Gleichheit der Dimension erhalten wir $H_\varphi(\lambda) = \operatorname{Eig}_\varphi^m(\lambda)$. □

6.3.17 **Nachtrag zu 6.3.9**. Mit diesem Satz können wir nun auch Aussage (iv) aus 6.3.9 herleiten. Wir haben gerade gesehen, dass die Vielfachheit k von λ im Minimalpolynom genau die kleinste Zahl ist, sodass $(\varphi - \lambda \operatorname{id}_V)^k|_{H_\varphi(\lambda)} = 0$ ist. In anderen Worten ist k genau der Nilpotenzindex von $(\varphi - \lambda \operatorname{id}_V)|_{H_\varphi(\lambda)}$. Aus dem Satz zur nilpotenten Normalform 6.2.51 wissen wir aber, dass $k = q_1$ ist, wenn $q = p^*$ die duale Partition der Rangpartition von $(\varphi - \lambda \operatorname{id}_V)|_{H_\varphi(\lambda)}$ ist. Aus der Definition von $\mathcal{N}(q)$ sehen wir, dass q_1 die Größe des größten Kästchens in $\mathcal{N}(q)$ und damit im Jordanblock $J(\lambda, q)$ zum Eigenwert λ ist.

Mit Satz 6.3.16 haben wir ein einfaches Verfahren gefunden um den Hauptraum zu berechnen. Wir bestimmen die verallgemeinerten Eigenräume $V_m = U_\varphi((X-\lambda)^m) = \operatorname{Ker}(\varphi - \lambda \operatorname{id}_V)^m$. Dies ist eine aufsteigende Folge von Unterräumen $V_1 \subseteq V_2 \subseteq V_3 \subseteq \dots$ bis irgendwann $\dim_K V_m = \alpha(\lambda)$ gilt. Dann ist $V_m = H_\varphi(\lambda)$ der Hauptraum und es gilt $V_{m+i} = V_m$ für alle $i \geq 1$.

6.3.18 **Beispiel**. Wir betrachten die Matrix
$$A = \begin{pmatrix} \overline{3} & \overline{4} & \overline{0} & \overline{0} \\ \overline{4} & \overline{4} & \overline{4} & \overline{3} \\ \overline{0} & \overline{0} & \overline{4} & \overline{2} \\ \overline{0} & \overline{0} & \overline{1} & \overline{0} \end{pmatrix} \in \mathrm{M}_{4,4}(\mathbb{F}_5).$$

Das charakteristische Polynom der Matrix A kann man mit Satz 4.3.15 ausrechnen, da A eine Blockdreiecksmatrix ist. Wir erhalten damit

$$\begin{aligned}\chi_A &= ((X-\overline{3})(X-\overline{4})-\overline{16})((X-\overline{4})X-\overline{2})\\ &= (X^2-\overline{2}X+\overline{1})(X^2-\overline{4}X+\overline{3}) = (X-\overline{1})^3(X-\overline{3}).\end{aligned}$$

Damit ist $\overline{1}$ ein Eigenwert von A mit der algebraischen Vielfachheit $\alpha(\overline{1}) = 3$. Wir berechnen nun den Hauptraum von A zum Eigenwert $\overline{1}$. Zuerst berechnen wir den Eigenraum $\mathrm{Eig}_A(\overline{1}) = \mathrm{Ker}(A - I_4)$

$$\begin{aligned}\mathrm{Ker}(A-I_4) &= \mathrm{Ker}\begin{pmatrix}\overline{2} & \overline{4} & \overline{0} & \overline{0}\\ \overline{4} & \overline{3} & \overline{4} & \overline{3}\\ \overline{0} & \overline{0} & \overline{3} & \overline{2}\\ \overline{0} & \overline{0} & \overline{1} & \overline{4}\end{pmatrix} = \mathrm{Ker}\begin{pmatrix}\overline{2} & \overline{4} & \overline{0} & \overline{0}\\ \overline{0} & \overline{0} & \overline{4} & \overline{3}\\ \overline{0} & \overline{0} & \overline{3} & \overline{2}\\ \overline{0} & \overline{0} & \overline{1} & \overline{4}\end{pmatrix}\\ &= \mathrm{Ker}\begin{pmatrix}\overline{2} & \overline{4} & \overline{0} & \overline{0}\\ \overline{0} & \overline{0} & \overline{1} & \overline{4}\\ \overline{0} & \overline{0} & \overline{0} & \overline{2}\\ \overline{0} & \overline{0} & \overline{0} & \overline{0}\end{pmatrix} = \left\langle \begin{pmatrix}\overline{2}\\ \overline{4}\\ \overline{0}\\ \overline{0}\end{pmatrix}\right\rangle.\end{aligned}$$

Da die Dimension des Eigenraumes echt kleiner als 3 ist, berechnen wir nun den verallgemeinerten Eigenraum der zweiten Stufe $\mathrm{Eig}_A^2(\overline{1}) = \mathrm{Ker}(A - I_4)^2$. Nach einer kurzen Rechnung erhält man

$$\mathrm{Ker}(A-I_4)^2 = \mathrm{Ker}\begin{pmatrix}\overline{0} & \overline{0} & \overline{1} & \overline{2}\\ \overline{0} & \overline{0} & \overline{2} & \overline{4}\\ \overline{0} & \overline{0} & \overline{1} & \overline{4}\\ \overline{0} & \overline{0} & \overline{2} & \overline{3}\end{pmatrix} = \left\langle \begin{pmatrix}\overline{1}\\ \overline{0}\\ \overline{0}\\ \overline{0}\end{pmatrix}, \begin{pmatrix}\overline{0}\\ \overline{1}\\ \overline{0}\\ \overline{0}\end{pmatrix}\right\rangle.$$

Der verallgemeinerte Eigenraum ist zwei-dimensional und ist damit noch nicht der Hauptraum. Wir berechnen also den verallgemeinerten Eigenraum der dritten Stufe $\mathrm{Eig}_A^3(\overline{1}) = \mathrm{Ker}(A - I_4)^3$. Nach einer kurzen Rechnung finden wir

$$\mathrm{Ker}(A-I_4)^3 = \mathrm{Ker}\begin{pmatrix}\overline{0} & \overline{0} & \overline{0} & \overline{0}\\ \overline{0} & \overline{0} & \overline{0} & \overline{0}\\ \overline{0} & \overline{0} & \overline{2} & \overline{3}\\ \overline{0} & \overline{0} & \overline{4} & \overline{1}\end{pmatrix} = \left\langle \begin{pmatrix}\overline{1}\\ \overline{0}\\ \overline{0}\\ \overline{0}\end{pmatrix}, \begin{pmatrix}\overline{0}\\ \overline{1}\\ \overline{0}\\ \overline{0}\end{pmatrix}, \begin{pmatrix}\overline{0}\\ \overline{0}\\ \overline{1}\\ \overline{1}\end{pmatrix}\right\rangle.$$

Also ist $H_A(\overline{1}) = \mathrm{Ker}(A - I_4)^3$ der Hauptraum zum Eigenwert $\overline{1}$.

6.3.19 Berechnung der Jordan'schen Normalform. Damit haben wir alle Hilfsmittel zusammen, um die Jordan'sche Normalform auszurechnen. Dieses Verfahren lässt sich aus dem Beweis des Hauptsatzes ableiten. Es sei V ein endlich-dimensionaler K-Vektorraum.

`Gegeben:` $\varphi \in \mathrm{End}(V))$

`Gesucht:` geordnete Basis $\mathcal{B}$, sodass ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)$ in Jordan'scher Normalform ist.

`Verfahren:`

Schritt 1: Berechne das charakteristische Polynom von φ. Bestimme die Eigenwerte $\lambda_1, \ldots, \lambda_r$ und ihre algebraischen Vielfachheiten $\alpha(\lambda_i)$.

Falls χ_φ nicht in Linearfaktoren zerfällt, beende das Verfahren: φ hat keine Matrixdarstellung in Jordan'scher Normalform.

Für alle $i \in \{1, \ldots, r\}$ führe die folgenden Schritte 2 und 3 aus.

Schritt 2: Berechne die verallgemeinerten Eigenräume $V_j = \mathrm{Eig}^j_\varphi(\lambda_i) \subseteq V$ für $j = 1, 2, \ldots$ bis $\dim_K V_k = \alpha(\lambda_i)$ gilt. Dann ist $H_\varphi(\lambda_i) = V_k$ der Hauptraum und $(\varphi - \lambda_i \,\mathrm{id}_V)|_{H_\varphi(\lambda_i)}$ ist nilpotent. Die aufsteigende Folge der verallgemeinerten Eigenräume

$$\{0\} = V_0 \subsetneq V_1 \subsetneq V_2 \subsetneq \ldots V_{k-1} \subsetneq V_k = H_\varphi(\lambda_i)$$

ist die zu $(\varphi - \lambda_i \,\mathrm{id}_V)|_{H_\varphi(\lambda_i)}$ gehörige Filtrierung.

Schritt 3: Berechne die Rangpartition von $(\varphi - \lambda_i \,\mathrm{id}_V)|_{H_\varphi(\lambda_i)}$

$$p = (\dim_K V_1 - \dim_K V_0, \dim_K V_2 - \dim_K V_1, \ldots, \dim_K V_k - \dim_K V_{k-1})$$

und die duale Partition $q^{(i)} = p^*$. Berechne mit dem Verfahren aus 6.2.56 eine geordnete Basis $\mathcal{B}_i = (v_1^{(i)}, \ldots, v_{\alpha(\lambda_i)}^{(i)})$, sodass die Matrixdarstellung von $(\varphi - \lambda_i \,\mathrm{id}_V)|_{H_\varphi(\lambda_i)}$ in nilpotenter Normalform ist, d.h., die Matrixdarstellung ist $\mathcal{N}(q^{(i)})$.

Schritt 4: Die Basis $\mathcal{B} = (v_1^{(1)}, \ldots, v_{\alpha(\lambda_1)}^{(1)}, \ldots, v_1^{(r)}, \ldots v_{\alpha(\lambda_r)}^{(r)})$ die durch Zusammenfügen der Basen $\mathcal{B}_i$ mit $1 \leq i \leq r$ entsteht, ist die gesuchte Basis und es gilt

$${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi) = \mathrm{diag}(J(\lambda_1, q^{(1)}), J(\lambda_2, q^{(2)}), \ldots, J(\lambda_r, q^{(r)})).$$

6.3.20 **Beispiel**. Wir betrachten wieder die Matrix

$$A = \begin{pmatrix} \overline{3} & \overline{4} & \overline{0} & \overline{0} \\ \overline{4} & \overline{4} & \overline{4} & \overline{3} \\ \overline{0} & \overline{0} & \overline{4} & \overline{2} \\ \overline{0} & \overline{0} & \overline{1} & \overline{0} \end{pmatrix} \in \mathrm{M}_{4,4}(\mathbb{F}_5)$$

aus Beispiel 6.3.18. Das charakteristische Polynom der Matrix A haben wir dort berechnet. Es gilt $\chi_A = (X - \overline{1})^3(X - \overline{3})$. Das Polynom zerfällt also in Linearfaktoren.

Wir betrachten zuerst den Eigenwert $\overline{1}$. In 6.3.18 haben wir den Hauptraum und die Filtrierung bestimmt:

$$\{0\} \subsetneq V_1 = \left\langle \begin{pmatrix} \overline{2} \\ \overline{4} \\ \overline{0} \\ \overline{0} \end{pmatrix} \right\rangle \subsetneq V_2 = \left\langle \begin{pmatrix} \overline{1} \\ \overline{0} \\ \overline{0} \\ \overline{0} \end{pmatrix}, \begin{pmatrix} \overline{0} \\ \overline{1} \\ \overline{0} \\ \overline{0} \end{pmatrix} \right\rangle \subsetneq V_3 = \left\langle \begin{pmatrix} \overline{1} \\ \overline{0} \\ \overline{0} \\ \overline{0} \end{pmatrix}, \begin{pmatrix} \overline{0} \\ \overline{1} \\ \overline{0} \\ \overline{0} \end{pmatrix}, \begin{pmatrix} \overline{0} \\ \overline{0} \\ \overline{1} \\ \overline{1} \end{pmatrix} \right\rangle = H_A(\overline{1}).$$

Die Rangpartition ist demnach

$$p = (1 - 0, 2 - 1, 3 - 2) = (1, 1, 1)$$

und die duale Partition ist $q = (3)$. Das Diagramm zur Partition q ist

$v_{1,1}$	$v_{1,2}$	$v_{1,3}$

Wie im Verfahren zur Bestimmung der nilpotenten Normalform suchen wir einen Vektor $v_{1,3} \in V_3 \setminus V_2$. Wir wählen $v_{1,3} = \begin{pmatrix} \overline{0} \\ \overline{0} \\ \overline{1} \\ \overline{1} \end{pmatrix}$. Dann verschieben wir mit $A - I_4$ (!) nach links und erhalten

$$v_{1,2} = (A - I_4)v_{1,3} = \begin{pmatrix} \overline{0} \\ \overline{2} \\ \overline{0} \\ \overline{0} \end{pmatrix},$$

$$v_{1,1} = (A - I_4)v_{1,2} = (A - I_4)^2 v_{1,3} = \begin{pmatrix} \overline{3} \\ \overline{1} \\ \overline{0} \\ \overline{0} \end{pmatrix}.$$

Dann ist $\mathcal{B}_1 = (v_{1,1}, v_{1,2}, v_{1,3})$ der erste Teil der Basis.

Nun betrachten wir den Eigenwert $\overline{3}$. Dieser Eigenwert hat die algebraische Vielfachheit $\alpha(\overline{3}) = 1$. In diesem Fall stimmt also der Eigenraum mit dem Hauptraum überein. Wir berechnen

$$\operatorname{Eig}_A(\overline{3}) = \operatorname{Ker}(A - \overline{3}I_4) = \operatorname{Ker} \begin{pmatrix} \overline{0} & \overline{4} & \overline{0} & \overline{0} \\ \overline{4} & \overline{1} & \overline{4} & \overline{3} \\ \overline{0} & \overline{0} & \overline{1} & \overline{2} \\ \overline{0} & \overline{0} & \overline{1} & \overline{2} \end{pmatrix} = \left\langle \begin{pmatrix} \overline{0} \\ \overline{0} \\ \overline{1} \\ \overline{2} \end{pmatrix} \right\rangle$$

und wählen $\mathcal{B}_2 = (w)$ mit $w = \begin{pmatrix} \overline{0} \\ \overline{0} \\ \overline{1} \\ \overline{2} \end{pmatrix}$ als letzten Basisvektor. Die geordnete Basis $\mathcal{B} = (v_{1,1}, v_{1,2}, v_{1,3}, w)$ hat nun die Eigenschaft, dass

$$_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(f_A) = \mathrm{diag}(J(\overline{1}, (3)), J(\overline{3}, (1))) = \begin{pmatrix} \overline{1} & \overline{1} & \overline{0} & \overline{0} \\ \overline{0} & \overline{1} & \overline{1} & \overline{0} \\ \overline{0} & \overline{0} & \overline{1} & \overline{0} \\ \overline{0} & \overline{0} & \overline{0} & \overline{3} \end{pmatrix}$$

in Jordan’scher Normalform ist.

Schreiben wir die Basisvektoren $(v_{1,1}, v_{1,2}, v_{1,3}, w)$ als Spalten in eine Matrix S, dann gilt $S^{-1}AS = \mathrm{diag}(J(\overline{1}, (3)), J(\overline{3}, (1)))$.

6.3.21 **Aufgabe**. Gegeben ist die Matrix L

$$B = \begin{pmatrix} 2 & 1 \\ -9 & -4 \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{Q}).$$

Bestimmen Sie eine invertierbare Matrix $S \in \mathrm{GL}_2(\mathbb{Q})$, sodass $S^{-1}BS$ in Jordan’scher Normalform ist.

IV. Anwendungen

Die Jordan’sche Normalform hat vor allem Anwendungen innerhalb der Mathematik. Sie ist ein wesentliches Werkzeug um Fragen über Matrizen zu beantworten. Insbesondere kann man damit das Normalformproblem für komplexe Matrizen beantworten.

6.3.22 **Satz.** *Jede Matrix $A \in \mathrm{M}_{n,n}(\mathbb{C})$ ist ähnlich zu einer Matrix in Jordan’scher Normalform. Die Normalform ist bis auf die Reihenfolge der Jordanblöcke eindeutig.*

Zwei komplexe Matrizen sind genau dann ähnlich, wenn ihre Jordan’schen Normalformen bis auf die Reihenfolge der Jordanblöcke gleich sind.

Beweis. Über den komplexen Zahlen zerfällt jedes normierte Polynom in Linearfaktoren, also folgt die erste Aussage aus Satz 6.3.8.

Sind $A, B \in \mathrm{M}_{n,n}(\mathbb{C})$ komplexe Matrizen mit derselben Jordan’schen Normalform, dann gibt es eine Matrix J in Normalform, sodass A und B ähnlich zu J sind. Da Ähnlichkeit eine Äquivalenzrelation ist (siehe Beispiel 4.2.25), sind A und B damit ähnlich.

Es seien umgekehrt A und B ähnlich. Sei J eine Matrix in Jordan'scher Normalform, die ähnlich zu A ist. Da Ähnlichkeit eine Äquivalenzrelation ist (siehe Beispiel 4.2.25), muss auch B ähnlich zu J sein. Da die Jordan'sche Normalform (bis auf Reihenfolge der Jordanblöcke) eindeutig ist, ist J eine Jordan'sche Normalform zu B, d.h., A und B haben dieselbe Jordan'sche Normalform. □

Wir wollen noch eine andere Anwendung der Jordan'schen Normalform vorstellen. Ist $A \in \mathrm{M}_{n,n}(K)$ gegeben, so ist es im Allgemeinen schwierig, die Potenzen A^m für große Werte von m zu bestimmen. Mit der Jordan'schen Normalform kann man dieses Problem aber gut lösen. Dazu benötigen wir den binomischen Lehrsatz. Zwei $n \times n$-Matrizen B und C *kommutieren miteinander*, wenn gilt $BC = CB$. In diesem Fall kann man per Induktion die vertraute Formel $(B+C)^m = \sum_{j=0}^{m} \binom{m}{j} B^{m-j} C^j$ beweisen.

6.3.23 **Berechnung von Potenzen einer Matrix in Jordan'scher Normalform**. Die Matrix $A = \operatorname{diag}(J(\lambda_1, q^{(1)}), J(\lambda_2, q^{(2)}), \ldots, J(\lambda_r, q^{(r)}))$ sei in Jordan'scher Normalform. Wir wollen die Potenzen von A bestimmen. Aufgrund der Rechenregel 1.2.38 für Produkte von Blockdiagonalmatrizen gilt

$$A^m = \operatorname{diag}(J(\lambda_1, q^{(1)})^m, J(\lambda_2, q^{(2)})^m, \ldots, J(\lambda_r, q^{(r)})^m)$$

für alle $m \in \mathbb{N}$. Es ist also ausreichend, wenn wir uns den Fall anschauen, dass $A = J(\lambda, q)$ ein Jordanblock zum Eigenwert λ ist. Dann ist $A = \lambda I_n + \mathcal{N}(q)$. Es sei $k = q_1$ der Nilpotenzindex von $\mathcal{N}(q)$. Weil die Matrizen λI_n und $\mathcal{N}(q)$ miteinander kommutieren, erhalten wir mit dem binomischen Lehrsatz die Gleichung

$$A^m = \sum_{j=0}^{m} \binom{m}{j} \lambda^{m-j} \mathcal{N}(q)^j = \lambda^m I_n + \sum_{j=1}^{k} \binom{m}{j} \lambda^{m-j} \mathcal{N}(q)^j, \tag{6.3.a}$$

denn es gilt $\mathcal{N}(q)^j = 0$ für alle $j \geq k$. Weil man die Potenzen von $\mathcal{N}(q)$ mit den Überlegungen aus Beispiel 6.2.45 gut berechnen kann, hat man damit eine geschlossene Formel für alle Potenzen von A.

6.3.24 **Beispiel**. Sei $\lambda \in \mathbb{R}$. Wir betrachten die Matrix

$$J(\lambda, (3)) = \lambda I_3 + \mathcal{N}(3) = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} \in \mathrm{M}_{n,n}(\mathbb{R}).$$

Es gelten

$$\mathcal{N}(3) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \qquad \text{und} \qquad \mathcal{N}(3)^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

und $\mathcal{N}(3)^3 = 0$. Sei nun $m \in \mathbb{N}$. Mit 6.3.a und den Formeln $\binom{m}{1} = m$ und $\binom{m}{2} = \frac{m(m-1)}{2}$ erhalten wir insgesamt einen Ausdruck für die Potenzen von $J(\lambda, (3))$:

$$J(\lambda, (3))^m = \begin{pmatrix} \lambda^m & m\lambda^{m-1} & \frac{m(m-1)}{2}\lambda^{m-2} \\ 0 & \lambda^m & m\lambda^{m-1} \\ 0 & 0 & \lambda^m \end{pmatrix}$$

6.3.25 **Aufgabe**. Sei $\lambda \in \mathbb{R}$. Geben Sie eine Formel für die Potenzen $J(\lambda, (2, 1))$ an. **L**

6.3.26 **Berechnung von Matrixpotenzen allgemein**. Ist $B \in \mathrm{M}_{n,n}(K)$ gegeben und ähnlich zu einer Matrix in Jordan’scher Normalform, dann können wir die Berechnung der Potenzen von B auf den Fall in 6.3.23 zurückführen. Es gibt eine invertierbare Matrix $S \in \mathrm{GL}_n(K)$, sodass $B = SAS^{-1}$ ist, wobei A eine Matrix in Jordan’scher Normalform ist. Dann gilt

$$B^m = (SAS^{-1})^m = SA\underbrace{S^{-1}S}_{=I_n}A\underbrace{S^{-1}S}_{=I_n}AS^{-1}\cdots SAS^{-1} = SA^mS^{-1}.$$

Wir können also wie in 6.3.23 die Potenzen von A berechnen und dann das Ergebnis mit der Matrix S konjugieren um eine Formel für die Potenzen von B zu erhalten.

6.3.27 **Beispiel**. Wir betrachten die Matrix

$$B = \begin{pmatrix} -1 & -3 \\ 2 & 4 \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{R}).$$

Wir suchen eine Formel für B^m. Dazu bestimmen wir zuerst die Jordan’sche Normalform. Das charakteristische Polynom von B ist

$$\chi_B = (X+1)(X-4) + 6 = X^2 - 3X + 2 = (X-1)(X-2).$$

Damit ist B diagonalisierbar (siehe 6.1.11) mit den Eigenwerten 1 und 2. Wir berechnen die Eigenräume

$$\mathrm{Eig}_B(1) = \mathrm{Ker}(B - I_2) = \langle \begin{pmatrix} 3 \\ -2 \end{pmatrix} \rangle,$$

und

$$\mathrm{Eig}_B(2) = \mathrm{Ker}(B - 2I_2) = \langle \begin{pmatrix} 1 \\ -1 \end{pmatrix} \rangle.$$

Wir wählen die genannten Vektoren als Basen der Eigenräume. Die Matrix $S = \begin{pmatrix} 3 & 1 \\ -2 & -1 \end{pmatrix}$ hat dann die Eigenschaft, dass

$$A = S^{-1}BS = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

gilt. Die Potenzen der Diagonalmatrix A kann man direkt berechnen. Es gilt

$$A^m = \begin{pmatrix} 1 & 0 \\ 0 & 2^m \end{pmatrix}$$

für alle $m \in \mathbb{N}$. Damit ist $SA^mS^{-1} = B^m$. Um diesen Ausdruck zu berechnen, bestimmen wir zuerst S^{-1}. Mit der Formel aus 4.3.26 erhalten wir

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

Insgesamt gilt also

$$B^m = SA^mS^{-1} = \begin{pmatrix} 3 - 2^{m+1} & 3(1 - 2^m) \\ 2^{m+1} - 2 & 3 \cdot 2^m - 2 \end{pmatrix}.$$

6.3.28 Beispiel. Die Fibonacci-Folge (f_n) ist eine rekursiv definierte Folge ganzer Zahlen. Man setzt $f_0 = 0$, $f_1 = 1$ und dann

$$f_{n+2} = f_{n+1} + f_n$$

für alle $n \geq 0$. Diese Folge kann man gut mithilfe von Matrixpotenzen beschreiben. Wir definieren

$$F = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{R})$$

und setzen $v_n = \begin{pmatrix} f_n \\ f_{n+1} \end{pmatrix}$ für $n \geq 0$. Die Einträge des Vektors v_n sind also immer die letzten beiden Einträge der Fibonacci-Folge. Für v_n gilt dann

$$Fv_n = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} f_n \\ f_{n+1} \end{pmatrix} = \begin{pmatrix} f_{n+1} \\ f_{n+1} + f_n \end{pmatrix} = \begin{pmatrix} f_{n+1} \\ f_{n+2} \end{pmatrix} = v_{n+1}.$$

Per Induktion erhalten wir damit die Formel $v_n = F^n v_0$. Wir können also eine allgemeine Formel für die Fibonacci-Folge herleiten, wenn wir eine Formel für die Potenzen von F kennen. Dazu verwenden wir das gerade beschriebene Verfahren.

Die Matrix F hat das charakteristische Polynom

$$\chi_F = X(X-1) - 1 = X^2 - X - 1.$$

Dieses Polynom hat zwei reelle Nullstellen

$$\lambda_1 = \frac{1+\sqrt{5}}{2} \quad \text{und} \quad \lambda_2 = \frac{1-\sqrt{5}}{2};$$

diese sind die Eigenwerte von F. Dabei ist λ_1 der sogenannte „goldene Schnitt". Die Matrix F ist also diagonalisierbar. Berechnet man die Eigenräume, dann findet man

$$\mathrm{Eig}_F(\lambda_1) = \left\langle \begin{pmatrix} \frac{1-\sqrt{5}}{2} \\ -1 \end{pmatrix} \right\rangle \quad \text{und} \quad \mathrm{Eig}_F(\lambda_2) = \left\langle \begin{pmatrix} 1 \\ \frac{1-\sqrt{5}}{2} \end{pmatrix} \right\rangle.$$

Ist $S = \left(\begin{smallmatrix} \lambda_2 & 1 \\ -1 & \lambda_2 \end{smallmatrix}\right)$ die Matrix, deren Spalten die gefundenen beiden Eigenvektoren sind, dann gilt $S^{-1}FS = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$. Von dieser Diagonalmatrix können wir nun die Potenzen berechnen und erhalten

$$(S^{-1}FS)^m = \begin{pmatrix} \lambda_1^m & 0 \\ 0 & \lambda_2^m \end{pmatrix}.$$

Aus der Formel 4.3.26 erhält man $S^{-1} = \frac{1}{\lambda_2^2+1} \left(\begin{smallmatrix} \lambda_2 & -1 \\ 1 & \lambda_2 \end{smallmatrix}\right)$. Mit einer Nebenrechnung kann man zeigen, dass $\lambda_2^2 + 1 = \frac{\sqrt{5}}{\lambda_1}$ gilt. Mit der Bemerkung aus 6.3.26 leiten wir eine Formel für die Potenzen von F her:

$$\begin{aligned} F^m &= S(S^{-1}FS)^m S^{-1} = S \begin{pmatrix} \lambda_1^m & 0 \\ 0 & \lambda_2^m \end{pmatrix} S^{-1} \\ &= \frac{1}{\sqrt{5}} \begin{pmatrix} \lambda_1^{m-1} - \lambda_2^{m-1} & \lambda_1^m - \lambda_2^m \\ \lambda_1^m - \lambda_2^m & \lambda_1^{m+1} - \lambda_2^{m+1} \end{pmatrix}. \end{aligned}$$

Der m-te Term f_m der Fibonacci-Folge ist der erste Eintrag des Vektors $F^m v_0$. Damit erhalten wir eine explizite Formel für die Terme der Fibonacci-Folge

$$f_m = \frac{\lambda_1^m - \lambda_2^m}{\sqrt{5}}.$$

6.3.29 **Aufgabe**. Berechnen Sie die Potenzen B^m der Matrix **L**

$$B = \begin{pmatrix} 2 & 1 \\ -9 & -4 \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{Q})$$

aus Aufgabe 6.3.21.

6.3.30 **Matrizen in Potenzreihen einsetzen**. In der Analysis gibt es Situationen, in denen man eine Potenzreihe

$$\sum_{m=0}^{\infty} c_m A^m$$

mit Koeffizienten $c_m \in \mathbb{R}$ berechnen möchte, wobei man aber keine Zahl, sondern eine reelle Matrix A in die Potenzreihe einsetzt. Dazu muss man natürlich immer sicher stellen, dass diese Reihe konvergiert. Ein bekanntes Beispiel ist das Lösungsverfahren für Systeme linearer Differentialgleichungen. Dabei muss man die Koeffizientenmatrix in die Exponentialreihe $\exp(x) = \sum_{m=0}^{\infty} \frac{1}{m!} x^m$ einsetzen.

Auch in diesen Situationen kann man die Berechnung auf den Fall zurückführen, dass A in Jordan’scher Normalform ist. In der Tat, ist $S^{-1}AS = B$ in Jordan’scher Normalform, dann gilt

$$\sum_{m=0}^{\infty} c_m A^m = S\Big(\sum_{m=0}^{\infty} c_m B^m\Big)S^{-1}.$$

6.3.31 **Die Jordan’sche Normalform ist numerisch instabil**. Die Jordan’sche Normalform ist für computergestützte numerische Berechnungen ziemlich ungeeignet. Der Grund dafür ist, dass die Jordan’sche Normalform *numerisch instabil* ist. Das heißt, auch sehr kleine Änderungen an einer reellen oder komplexen Matrix können die zugehörige Jordan’sche Normalform stark verändern. Zum Beispiel ist die Matrix

$$A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{R})$$

eine nicht diagonalisierbare Matrix in Jordan’scher Normalform mit 2 als einzigem Eigenwert. Verändern wir einen Diagonaleintrag von A mit einer beliebig kleinen reellen Zahl $\varepsilon \neq 0$, dann ist

$$\begin{pmatrix} 2+\varepsilon & 1 \\ 0 & 2 \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{R})$$

plötzlich eine diagonalisierbare Matrix mit zwei verschiedenen Eigenwerten $2+\varepsilon$ und 2. Die zugehörige Jordan’sche Normalform ist also

$$\begin{pmatrix} 2 & 0 \\ 0 & 2+\varepsilon \end{pmatrix}$$

und unterscheidet sich im Eintrag oben rechts deutlich von A.

6.4. Lösungen der Aufgaben in Lektion 6

L6.1.2 **Lösung.** Sei V ein K-Vektorraum. Sei $\varphi \in \mathrm{End}(V)$ und sei $g \in K[X]$. Da $g(\varphi)$ wieder ein Endomorphismus von V ist, ist das Bild jedenfalls ein Unterraum von V; das wissen wir aus [MG, 8.3.2]. Es bleibt zu zeigen, dass $\mathrm{Bild}(g(\varphi))$ sogar φ-invariant ist. Es sei $w \in \mathrm{Bild}(g(\varphi))$, dann gilt $w = g(\varphi)(v)$ für ein $v \in V$. Wir schreiben $g = \sum_i a_i X^i$ mit $a_i \in K$ und erhalten

$$\begin{aligned}
\varphi(w) = \varphi(g(\varphi)(v)) &= \varphi\Big(\sum_i a_i \varphi^i(v)\Big) \\
&= \sum_i a_i \varphi^{i+1}(v) && (\varphi \text{ linear}) \\
&= \sum_i a_i \varphi^i(\varphi(v)) && (\varphi^{i+1} = \varphi^i \circ \varphi) \\
&= g(\varphi)(\varphi(v)) \in \mathrm{Bild}(g(\varphi)).
\end{aligned}$$

Damit ist $\mathrm{Bild}(g(\varphi))$ ein φ-invarianter Unterraum.

L6.1.7 **Lösung.** Ist g ein Teiler von h, dann gibt es ein Polynom $p \in K[X]$ mit $pg = h$. Sei nun $v \in U_\varphi(g) = \mathrm{Ker}(g(\varphi))$. Mit den Rechenregeln aus Lemma 5.3.27 erhalten wir

$$h(\varphi)(v) = pg(\varphi)(v) = p(\varphi)(\underbrace{g(\varphi)(v)}_{=0}) = 0$$

und damit $v \in U_\varphi(h)$. Da v beliebig war, gilt $U_\varphi(g) \subseteq U_\varphi(h)$.

L6.1.9 **Lösung.** Wir berechnen A^2 und erhalten

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

Damit finden wir

$$g(A) = A(A^2 + I_3) = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & -1 \\ 1 & 2 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} = 0;$$

d.h., $U_A(g) = \mathrm{Ker}(g(A)) = \mathbb{R}^3$. Die beiden invarianten Unterräume sind

$$U_A(X) = \mathrm{Ker}(A) = \langle \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \rangle$$

und

$$U_A(X^2+1) = \operatorname{Ker}(A^2+I_3) = \operatorname{Ker}\begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 0\\ 1 & 0 & 0\end{pmatrix} = \langle \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix} \rangle.$$

L6.1.15 Lösung. Sei $A \in \mathrm{M}_{n,n}(\mathbb{F}_2)$.

Wir nehmen an, dass A diagonalisierbar ist. Dann zerfällt μ_A in paarweise verschiedene Linearfaktoren (siehe 6.1.12). Da $\mathbb{F}_2$ nur zwei Elemente $\overline{0}$ und $\overline{1}$ enthält, folgt daraus, dass μ_A das Polynom $g = X(X-\overline{1}) = X^2 - X$ teilt. Mit Satz 5.3.37 sehen wir, dass $0 = g(A) = A^2 - A$ gilt, d.h., $A^2 = A$.

Gilt umgekehrt $A^2 = A$, dann annulliert das Polynom $X^2 - X$ die Matrix A. Mit Satz 5.3.37 folgt, dass μ_A ein Teiler von $X^2 - X = X(X-\overline{1})$ ist. Damit zerfällt μ_A in paarweise verschiedene Linearfaktoren (siehe 1.4.37) und A ist diagonalisierbar durch Korollar 6.1.12.

L6.2.4 Lösung. Um zu zeigen, dass die Matrix

$$A = \begin{pmatrix} -2 & 4 & -6 & 12\\ -1 & 2 & -3 & 6\\ 0 & 0 & -2 & 4\\ 0 & 0 & -1 & 2 \end{pmatrix} \in \mathrm{M}_{4,4}(\mathbb{Q})$$

nilpotent ist, berechnen wir solange die Potenzen von A bis wir ein k mit $A^k = 0$ finden. Es gilt

$$A^2 = A \cdot A = 0.$$

Also ist A nilpotent. Da $A \neq 0$ ist, ist $\mathrm{Ni}(A) = 2$ der Nilpotenzindex von A.

L6.2.5 Lösung. Wir berechnen die Potenzen der Matrix $B = \begin{pmatrix} \overline{2} & \overline{2} & \overline{0}\\ \overline{2} & \overline{0} & \overline{2}\\ \overline{0} & \overline{2} & \overline{1} \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{F}_3)$.

Wir erhalten

$$B^2 = \begin{pmatrix} \overline{2} & \overline{1} & \overline{1}\\ \overline{1} & \overline{2} & \overline{2}\\ \overline{1} & \overline{2} & \overline{2} \end{pmatrix} \qquad \text{und} \qquad B^3 = 0.$$

Damit ist B nilpotent und es ist $\mathrm{Ni}(B) = 3$.

L6.2.6 **Lösung.** Es seien $A, B \in \mathrm{M}_{n,n}(K)$ ähnliche Matrizen und A sei nilpotent. Weil B ähnlich zu A ist, existiert eine invertierbare Matrix $S \in \mathrm{M}_{n,n}(K)$ mit $B = S^{-1}AS$. Es gilt dann für alle $m \in \mathbb{N}_0$ die Gleichung

$$B^m = (S^{-1}AS)^m = S^{-1}A\underbrace{SS^{-1}}_{=I_n}A\underbrace{SS^{-1}}_{=I_n}AS \cdots S^{-1}AS = S^{-1}A^mS.$$

Weil S invertierbar ist, ist B^m genau dann die Nullmatrix, wenn A^m die Nullmatrix ist. Folglich ist B ebenfalls nilpotent und das minimale m mit $A^m = 0$ ist das minimale m mit $B^m = 0$, d.h., $\mathrm{Ni}(A) = \mathrm{Ni}(B)$.

L6.2.9 **Lösung.** Es sei $\varphi \in \mathrm{End}(V)$ nilpotent und es sei $v \in V$ ein Vektor der Stufe $j > 0$. Wir prüfen, dass $\varphi(v)$ ein Vektor der Stufe $j-1$ ist.

Da $j \geq 1$ ist und v die Stufe j hat, gilt

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

Also hat $\varphi(v)$ höchstens Stufe $j-1$. Die Stufe kann nicht kleiner sein, denn aus $\varphi^k(\varphi(v)) = 0$ folgt $\varphi^{k+1}(v) = 0$ und damit $k+1 \geq j$, weil j minimal mit dieser Eigenschaft ist.

L6.2.11 **Lösung.** Sei $\varphi \in \mathrm{End}(V)$. Ist $x \in \mathrm{Ker}(\varphi^j)$, dann gilt auch $\varphi^{j+1}(x) = \varphi(\varphi^j(x)) = \varphi(0) = 0$, d.h., $x \in \mathrm{Ker}(\varphi^{j+1})$. Sei nun $x \in \mathrm{Bild}(\varphi^{j+1})$. Dann ist $x = \varphi^{j+1}(y)$ für ein $y \in V$. Es gilt also $x = \varphi^j(\varphi(y)) \in \mathrm{Bild}(\varphi^j)$. Für alle $x \in V$ gilt außerdem

$$\begin{aligned} x \in \varphi^{-1}(\mathrm{Ker}(\varphi^j)) \quad &\Leftrightarrow \quad \varphi(x) \in \mathrm{Ker}(\varphi^j) \\ &\Leftrightarrow \quad \varphi^j(\varphi(x)) = \varphi^{j+1}(x) = 0 \\ &\Leftrightarrow \quad x \in \mathrm{Ker}(\varphi^{j+1}). \end{aligned}$$

Mit der Beobachtung, dass x in $\mathrm{Bild}(\varphi^{j+1})$ liegt genau dann, wenn $x = \varphi^{j+1}(y) = \varphi(\varphi^j(y))$ für ein $y \in V$ gilt, zeigt man letzte Behauptung.

L6.2.14 **Lösung.** Wir bestimmen die Filtrierung zur Matrix B aus Aufgabe 6.2.5. Zuerst berechnen wir den Nullraum von B:

$$V_1 = \mathrm{Ker}(B) = \mathrm{Ker}\begin{pmatrix} \overline{2} & \overline{2} & \overline{0} \\ \overline{2} & \overline{0} & \overline{2} \\ \overline{0} & \overline{2} & \overline{1} \end{pmatrix} = \mathrm{Ker}\begin{pmatrix} \overline{1} & \overline{1} & \overline{0} \\ \overline{0} & \overline{1} & \overline{2} \\ \overline{0} & \overline{0} & \overline{0} \end{pmatrix} = \langle \begin{pmatrix} \overline{2} \\ \overline{1} \\ \overline{1} \end{pmatrix} \rangle.$$

In Aufgabe 6.2.5 haben wir bereits B^2 berechnet. Der zweite Term der Filtrierung ist der Kern von B^2:

$$V_2 = \mathrm{Ker}(B^2) = \mathrm{Ker}\begin{pmatrix} \overline{2} & \overline{1} & \overline{1} \\ \overline{1} & \overline{2} & \overline{2} \\ \overline{1} & \overline{2} & \overline{2} \end{pmatrix} = \mathrm{Ker}\begin{pmatrix} \overline{1} & \overline{2} & \overline{2} \\ \overline{0} & \overline{0} & \overline{0} \\ \overline{0} & \overline{0} & \overline{0} \end{pmatrix} = \langle \begin{pmatrix} \overline{1} \\ \overline{0} \\ \overline{1} \end{pmatrix}, \begin{pmatrix} \overline{1} \\ \overline{1} \\ \overline{0} \end{pmatrix} \rangle.$$

Aus Aufgabe 6.2.5 wissen wir, dass $B^3 = 0$ ist, also ist $V_3 = \mathbb{F}_3^3$.

L6.2.17 Lösung. Ist $A = {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)$ eine strikte obere Dreiecksmatrix, dann ist $XI_n - A$ eine obere Dreiecksmatrix, deren Diagonaleinträge alle gleich X sind. Mit Definition 5.3.17 und Satz 4.2.13 erhalten wir

$$\chi_\varphi = \chi_A = \det(XI_n - A) = X^n.$$

L6.2.22 Lösung. Warum handelt es sich bei $Z_\varphi(v)$ um den kleinsten φ-invarianten Unterraum, der v enthält? Ist U ein φ-invarianter Unterraum, der v enthält, dann enthält U auch $\varphi(v)$ und damit auch $\varphi(\varphi(v)) = \varphi^2(v)$ und damit auch $\varphi^3(v)$ und so weiter und so fort. Per Induktion kann man schließen, dass U die Vektoren $\varphi^k(v)$ für alle $k \in \mathbb{N}_0$ enthält. Da U ein Unterraum ist, enthält U dann auch alle Linearkombinationen dieser Vektoren, d.h., $U \supseteq Z_\varphi(v)$.

L6.2.25 Lösung. *U' ist φ-invariant:* Sei $u' = cx + u_0 \in U'$ mit $c \in K$ und $u_0 \in U$. Dann gilt

$$\varphi(cx + u_0) = \underbrace{c\varphi(x)}_{\in U} + \underbrace{\varphi(u_0)}_{\in U} \in U \subseteq U'$$

und der Unterraum U' ist somit φ-invariant.

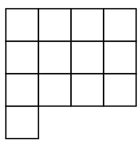
$U' \cap Z_\varphi(w) = \{0\}$: Es sei $u' = cx + u_0 \in U' \cap Z_\varphi(w)$ mit $c \in K$ und $u_0 \in U$. Durch Umformen folgt $cx \in U + Z_\varphi(w)$ und damit $c = 0$, denn x liegt nicht in $U + Z_\varphi(w)$. Wegen $U \cap Z_\varphi(w) = \{0\}$ muss dann aber auch $u_0 = 0$ sein, d.h., $u' = 0$. Da u' beliebig war, schließen wir $U' \cap Z_\varphi(w) = \{0\}$.

L6.2.29 Lösung. Die Partitionen der Zahl 5 sind

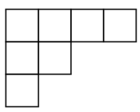
$$(5), (4,1), (3,2), (3,1,1), (2,2,1), (2,1,1,1), (1,1,1,1,1).$$

Es gibt also sieben verschiedene Partitionen der Zahl 5.

L6.2.31 Lösung. (a) Das Diagramm zur Partition $(4,4,4,1)$ ist

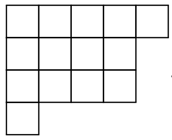


(b) Das Diagramm



hat 4 Kästchen in der ersten, 2 Kästchen in der zweiten und 1 Kästchen in der dritten Zeile, d.h, die zugehörige Partition ist $(4,2,1)$.

L6.2.35 Lösung. Das Diagramm zur Partition $p = (5, 4, 4, 1)$ ist


.

In der ersten Spalte sind 4 Kästchen, in der zweiten, dritten und vierten Spalte sind 3 Kästchen und in der fünften Spalte ist 1 Kästchen. Die duale Partition zu p ist damit $p^* = (4, 3, 3, 3, 1)$.

L6.2.41 Lösung. Die Summe über die Zahlen $p_j(\varphi)$ ist eine Teleskopsumme und wir erhalten

$$\begin{aligned}
\sum_{j=1}^{k} p_j(\varphi) &= \sum_{j=1}^{k} \big(\dim_K(V_j) - \dim_K(V_{j-1})\big) \\
&= \sum_{j=1}^{k} \dim_K(V_j) - \sum_{j'=0}^{k-1} \dim_K(V_{j'}) \qquad (j' = j-1) \\
&= \dim_K(\mathrm{Ker}(\varphi^k)) - \dim_K(\mathrm{Ker}(\varphi^0)) \\
&= \dim_K(V) - \dim_K(\{0\}) = \dim_K(V).
\end{aligned}$$

L6.2.44 Lösung. Wir wissen aus Aufgabe 6.2.5, dass $\mathrm{Ni}(B) = 3$ ist. Anhand der Filtrierung (siehe Aufgabe 6.2.14) sehen wir nun, dass $\dim_{\mathbb{F}_3}(V_1) = 1$, $\dim_{\mathbb{F}_3}(V_2) = 2$ und $\dim_{\mathbb{F}_3}(V_3) = 3$ gelten. Damit ist

$$p(A) = (1 - 0, 2 - 1, 3 - 2) = (1, 1, 1)$$

die Rangpartition von A.

L6.2.47 Lösung. Wir nehmen an, φ sei nilpotent mit $\mathrm{Ni}(\varphi) = \ell$. Dann gilt jedenfalls $\varphi^\ell(w) = 0$ für jedes $w \in W_i$, d.h., $\varphi|_{W_i}$ ist nilpotent und es gilt $\mathrm{Ni}(\varphi) \geq \mathrm{Ni}(\varphi|_{W_i})$. Insbesondere schließen wir $\mathrm{Ni}(\varphi) \geq \max\{\mathrm{Ni}(\varphi|_{W_i}) \mid 1 \leq i \leq k\}$.

Nehmen wir nun umgekehrt an, dass $\varphi|_{W_i}$ für alle i nilpotent ist. Sei $v \in V$ beliebig. Wir schreiben

$$v = \sum_{i=1}^{k} w_i$$

mit Vektoren $w_i \in W_i$. Ist $\ell = \max\{\mathrm{Ni}(\varphi|_{W_i}) \mid 1 \leq i \leq k\}$ dann erhalten wir

$$\varphi^\ell(v) = \sum_{i=1}^{k} \varphi^\ell(w_i) = \sum_{i=1}^{k} 0 = 0.$$

Damit ist φ nilpotent und $\mathrm{Ni}(\varphi) \leq \max\{\mathrm{Ni}(\varphi|_{W_i}) \mid 1 \leq i \leq k\}$. Da wir nun beide Ungleichungen haben, muss $\mathrm{Ni}(\varphi) = \max\{\mathrm{Ni}(\varphi|_{W_i}) \mid 1 \leq i \leq k\}$ sein.

L6.2.53 Lösung. In Aufgabe 6.2.44 haben wir gesehen, dass B die Rangpartition $p(B) = (1,1,1)$ besitzt. Duale Partition dazu ist $q = p(B)^* = (3)$. Aus Korollar 6.2.52 folgt nun, dass B ähnlich ist zur Matrix

$$\mathcal{N}(q) = N(3) = \begin{pmatrix} \overline{0} & \overline{1} & \overline{0} \\ \overline{0} & \overline{0} & \overline{1} \\ \overline{0} & \overline{0} & \overline{0} \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{F}_3).$$

L6.2.60 Lösung. Die Matrix B hat Nilpotenzindex 3 (siehe 6.2.5) und hat die Rangpartition $p(B) = (1,1,1)$ (siehe 6.2.44). Die duale Partition zur Rangpartition ist $q = p(B)^* = (3)$. Das Diagramm dazu ist:

$v_{1,1}$	$v_{1,2}$	$v_{1,3}$

Die Filtrierung haben wir in Aufgabe 6.2.14 bestimmt. Es ist $V_3 = \mathbb{F}_3^3$ und

$$V_2 = \mathrm{Ker}(B^2) = \langle \begin{pmatrix} \overline{1} \\ \overline{0} \\ \overline{1} \end{pmatrix}, \begin{pmatrix} \overline{1} \\ \overline{1} \\ \overline{0} \end{pmatrix} \rangle.$$

Für $v_{1,3}$ suchen wir einen Vektor in $V_3 \setminus V_2$. Wir wählen $v_{1,3} = e_3 = \begin{pmatrix} \overline{0} \\ \overline{0} \\ \overline{1} \end{pmatrix}$. Dann berechnen wir

$$v_{1,2} = Bv_{1,3} = \begin{pmatrix} \overline{0} \\ \overline{2} \\ \overline{1} \end{pmatrix}, \quad v_{1,1} = Bv_{1,2} = \begin{pmatrix} \overline{1} \\ \overline{2} \\ \overline{2} \end{pmatrix}.$$

Wir schreiben diese Vektoren in der Reihenfolge $v_{1,1}, v_{1,2}, v_{1,3}$ in die Matrix

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

Die Matrix S ist invertierbar, denn es handelt sich im eine untere Dreiecksmatrix mit $\det(S) = \overline{1} \cdot \overline{2} \cdot \overline{1} = \overline{2} \neq \overline{0}$. Es gilt dann

$$S^{-1}BS = \mathcal{N}(q) = N(3);$$

das ist die nilpotente Normalform zu B.

L6.2.61 Lösung. Wir prüfen, dass die Matrix

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & -2 & 1 & -1 \\ 0 & -1 & 0 & 0 \\ -2 & 4 & -2 & 2 \end{pmatrix} \in \mathrm{M}_{4,4}(\mathbb{R})$$

nilpotent ist, indem wir die Potenzen berechnen. Wir erhalten:

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

Also ist A nilpotent vom Nilpotenzindex 3.

Um A in nilpotente Normalform zu bringen, berechnen wir die Filtrierung.

$$\begin{aligned} V_0 &= \{0\} \\ V_1 &= \mathrm{Ker}(A) = \left\langle \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle \\ V_2 &= \mathrm{Ker}(A^2) = \left\langle \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\rangle \\ V_3 &= \mathbb{R}^4 \end{aligned}$$

Die Rangpartition zu A ist also $p(A) = (2, 1, 1)$ und die duale Partition ist $q = (3, 1)$. Wir zeichnen das Diagramm zur dualen Partition und suchen die Basisvektoren $v_{i,j}$ von rechts nach links.

$v_{1,1}$	$v_{1,2}$	$v_{1,3}$
$v_{2,1}$		

Wir wählen $v_{1,3} \in V_3 \setminus V_2$, also zum Beispiel $v_{1,3} = \left(\begin{smallmatrix} 0 \\ 0 \\ 0 \\ 1 \end{smallmatrix}\right)$. Dann finden wir die Vektoren links davon durch Anwenden von A:

$$v_{1,2} = Av_{1,3} = \begin{pmatrix} 0 \\ -1 \\ 0 \\ 2 \end{pmatrix},$$

$$v_{1,1} = Av_{1,2} = A^2 v_{1,3} = \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

Schließlich benötigen wir noch den Vektor $v_{2,1}$. Dieser muss in V_1 liegen und $v_{1,1}, v_{2,1}$ müssen eine Basis eines Komplementes zu $V_0 = \{0\}$ in V_1 sein. In diesem Fall heißt das: $v_{1,1}, v_{2,1}$ bilden eine Basis von $V_1 = \operatorname{Ker}(A)$. Wir wählen also $v_{2,1} = \left(\begin{smallmatrix} 1 \\ 0 \\ 0 \\ 1 \end{smallmatrix}\right)$. Wir tragen die gefundenen Basisvektoren (in der richtigen Reihenfolge: $v_{1,1}, v_{1,2}, v_{1,3}, v_{2,1}$) in eine Matrix S ein und erhalten

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

Dann gilt

$$S^{-1}AS = \mathcal{N}(3,1) = \left(\begin{array}{ccc|c} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right).$$

L6.3.4 **Lösung.** Es ist

$$J(4,(3,1,1)) = 4I_5 + \mathcal{N}(3,1,1) = \left(\begin{array}{ccc|c|c} 4 & 1 & 0 & 0 & 0 \\ 0 & 4 & 1 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ \hline 0 & 0 & 0 & 4 & 0 \\ \hline 0 & 0 & 0 & 0 & 4 \end{array}\right).$$

Dieser Jordanblock hat drei Jordankästchen.

L6.3.11 **Lösung.** Gesucht ist eine reelle Matrix A in Jordan'scher Normalform mit

$$\chi_A = (X-2)^3(X-5)^4 \quad \text{und} \quad \mu_A = (X-2)^2(X-5)^2.$$

Wir benötigen also zwei Jordanblöcke zu den Eigenwerten 2 und 5. Aus dem charakteristischen Polynom liest man ab, dass der Block zum Eigenwert 5 ein 4×4-Block und der Block zum Eigenwert 2 ein 3×3-Block ist. Insgesamt handelt es sich also um eine 7×7-Matrix. Aus dem Minimalpolynom kann man ablesen, dass das größte Kästchen in beiden Blöcken ein 2×2-Kästchen ist. In Frage kommen also die Matrizen

$$\begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 5 \end{pmatrix} \quad \text{oder} \quad \begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 5 \end{pmatrix}.$$

Die Jordanblöcke könnte man natürlich auch umgekehrt anordnen.

L6.3.21 **Lösung.** Gegeben ist die Matrix

$$B = \begin{pmatrix} 2 & 1 \\ -9 & -4 \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{Q}).$$

Das charakteristische Polynom ist

$$\begin{aligned} \chi_B = \det(XI_2 - B) &= \det \begin{pmatrix} X-2 & -1 \\ 9 & X+4 \end{pmatrix} \\ &= (X-2)(X+4) + 9 = X^2 + 2X + 1 = (X+1)^2. \end{aligned}$$

Da das charakteristische Polynom in Linearfaktoren zerfällt, ist B ähnlich zu einer Matrix in Jordan'scher Normalform. Die Matrix B hat nur den Eigenwert -1 mit der algebraischen Vielfachheit 2. Wir berechnen den Eigenraum zu -1:

$$V_1 = \mathrm{Eig}_B(-1) = \mathrm{Ker}(B + I_2) = \mathrm{Ker} \begin{pmatrix} 3 & 1 \\ -9 & -3 \end{pmatrix} = \langle \begin{pmatrix} -1 \\ 3 \end{pmatrix} \rangle.$$

Da der Eigenraum ein-dimensional ist, ist dies noch nicht der Hauptraum $H_B(-1)$ und $\mu_B \neq X+1$. Da das Minimalpolynom ein Teiler von χ_B ist, muss $\mu_B = (X+1)^2$. Der Hauptraum ist bekanntlich 2-dimensional (denn $\alpha(-1) = 2$), also gilt hier $H_B(-1) = \mathbb{Q}^2$. Wir suchen nun einen Vektor $v_{1,2} \in \mathbb{Q}^2$, der nicht in $\mathrm{Eig}_B(-1)$ liegt. Wir wählen

$$v_{1,2} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

und erhalten damit $v_{1,1} = (B + I_2)v_{1,2} = \begin{pmatrix} 1 \\ -3 \end{pmatrix}$. Wir schreiben die gefundenen Vektoren als Spalten in eine Matrix und erhalten

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

Dann gilt

$$S^{-1}BS = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix} = J(-1,(2))$$

und diese Matrix ist in Jordan'scher Normalform.

L6.3.25 Lösung. Sei $\lambda \in \mathbb{R}$. Wir suchen eine Formel für die Potenzen von

$$J(\lambda,(2,1)) = \lambda I_3 + \mathcal{N}(2,1) = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} \in \mathrm{M}_{n,n}(\mathbb{R}).$$

Es ist

$$\mathcal{N}(2,1) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

und $\mathcal{N}(2,1)^2 = 0$. Sei $m \in \mathbb{N}$. Aus 6.3.a und $\binom{m}{1} = m$ erhalten wir eine Formel für die Potenzen von $J(\lambda,(2,1))$:

$$J(\lambda,(2,1))^m = \begin{pmatrix} \lambda^m & m\lambda^{m-1} & 0 \\ 0 & \lambda^m & 0 \\ 0 & 0 & \lambda^m \end{pmatrix}.$$

L6.3.29 Lösung. Für die Matrix

$$B = \begin{pmatrix} 2 & 1 \\ -9 & -4 \end{pmatrix} \in \mathrm{M}_{n,n}(\mathbb{Q})$$

haben wir in Aufgabe 6.3.21 die Jordan'sche Normalform berechnet. Es gilt

$$A = S^{-1}BS = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix} = J(-1,(2))$$

mit $S = \begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix}$. Die Potenzen von A bestimmt man wie in 6.3.23; man erhält

$$A^m = \begin{pmatrix} (-1)^m & (-1)^{m-1}m \\ 0 & (-1)^m \end{pmatrix}.$$

Das Inverse von S berechnen wir mit der Formel aus 4.3.26 und finden

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

Mit der Bemerkung aus 6.3.26 gilt also

$$B^m = SA^mS^{-1} = \begin{pmatrix} (-1)^m(1-3m) & (-1)^{m-1}m \\ 9m(-1)^m & (-1)^m(1+3m) \end{pmatrix}.$$

Steffen Kionke

Lineare Algebra

Lektion 7:
Euklidische und unitäre Vektorräume

Fakultät für
Mathematik und Informatik

FernUniversität in Hagen

Studierhinweise zur siebten Lektion

In dieser Lektion befassen wir uns ausführlich mit *Euklidischen* und *unitären* Vektorräumen. Dabei handelt es sich um reelle oder komplexe Vektorräume mit einem *Skalarprodukt*, also einer positiv definiten symmetrischen Bilinearform oder einer positiv definiten Hermite'schen Form. Ein Skalarprodukt ermöglicht es *Abstände* zu messen und von *orthogonalen* Vektoren zu sprechen. Ein Euklidischer oder unitärer Vektorraum ist also geeignet um Geometrie zu betreiben.

Wenn man „Vektorräume“ durch „Vektorräume mit Skalarprodukt“ ersetzt, muss man zunächst auch einige Grundbegriffe überdenken. Insbesondere werden wir *Orthonormalbasen* kennenlernen. Nur diese Basen führen zu Koordinaten, die geeignet sind, um in Euklidischen und unitären Vektorräumen zu rechnen. Die Transformationsmatrizen zwischen Orthonormalbasen sind sogenannte *orthogonale* bzw. *unitäre* Matrizen. Allgemein hat eine lineare Abbildung eine orthogonale oder unitäre Matrixdarstellung genau dann, wenn es sich um eine abstandserhaltende Abbildung handelt. Solche Abbildungen nennen wir *Isometrien*. Beispielsweise sind Drehungen und Spiegelungen Isometrien.

Das Ziel dieser Lektion sind die *Spektralsätze*. Das „Spektrum“ eines Endomorphismus ist die Menge seiner Eigenwerte. Der Begriff kommt eigentlich aus der Funktionalanalysis, der Theorie der unendlich-dimensionalen normierten Räume (Banachräume oder Hilberträume), wo ein Operator im Allgemeinen nicht nur endlich viele, sondern ein ganzes „Spektrum“ von Eigenwerten hat. Ein Spektralsatz ist ein Satz zur Diagonalisierbarkeit bestimmter Endomorphismen, der meistens auch eine Einschränkung an die möglichen Eigenwerte beinhaltet. Beispielsweise besagt der *Spektralsatz für selbstadjungierte Endomorphismen*, dass diese „selbstadjungierten“ Endomorphismen diagonalisierbar sind und nur reelle Eigenwerte haben.

Was haben die Spektralsätze mit den Skalarprodukten zu tun? Mit dem Skalarprodukt definieren wir zu jedem Endomorphismus einen sogenannten *adjungierten Endomorphismus*. Wenn ein Endomorphismus mit seinem Adjungierten vertauscht, dann stehen die Eigenräume orthogonal aufeinander und das kann man verwenden um eine Orthonormalbasis aus Eigenvektoren zu finden.

Insgesamt ist diese Lektion einfacher als die letzte. Wir lernen hier viele schöne und nützliche Begriffe kennen und verstehen am Ende, warum der Spektralsatz für selbstadjungierte Endomorphismen einen geeigneten Schlusspunkt für diesen Studienbrief bildet.

Lernziele

Nach dem Durcharbeiten der Lektion sollten Sie

- → sicher im Rechnen mit *Skalarprodukten* sein,
- → das *Hauptminorenkriterium* anwenden können,
- → die *Norm* zu einem Skalarprodukt und die *Ungleichung von Cauchy-Schwarz* kennen,
- → mit den Begriffen *orthogonal*, *Orthonormalbasis*, *orthogonales Komplement* vertraut sein,
- → den *Basisergänzungssatz* für Orthonormalbasen kennen,
- → das *Orthonormalisierungsverfahren nach Gram-Schmidt* anwenden können,
- → wissen, was eine *adjungierte Abbildung* ist,
- → wissen, was eine *Isometrie* ist und die Verbindung zu *orthogonalen* und *unitären* Matrizen kennen,
- → die *Spektralsätze* für *normale*, *unitäre* und *selbstadjungierte* Endomorphismen kennen und anwenden können,
- → Orthonormalbasen wie in den *Spektralsätzen* berechnen können.

Literaturhinweise

Die Euklidischen und unitären Vektorräume werden Sie in allen Lehrbüchern finden. Auch die Spektralsätze werden ausführlich besprochen. Manchmal fehlt der allgemeine Spektralsatz für normale Endomorphismen.

- → [Bär]: 7.3 – 7.6.
- → [Beu]: Kapitel 10.
- → [Bo]: Kapitel 7.
- → [Fi]: Kapitel 6.
- → [Gö]: 5.1–5.3, 6.2–6.5.
- → [KaSt]: Kapitel 9.

Fahrplan durch die Lektion

In Abschnitt 7.1 klären wir die Grundbegriffe *Skalarprodukt* und *Norm*. Der erste Teil beginnt mit der Definition von Skalarprodukten (7.1.1) und einigen Beispielen (7.1.4, 7.1.5). Mit dem *Hauptminorenkriterium* (7.1.8) kann man Skalarprodukte anhand einer Matrixdarstellung erkennen. ← 7.1

Im zweiten Teil lernen wir die *Norm* (7.1.11) zu einem Skalarprodukt und die wichtigsten Eigenschaften kennen (7.1.13). Wir beweisen die nützliche Ungleichung von Cauchy-Schwarz (7.1.15) und erklären, wie man mithilfe der Norm *Abstände* messen kann (7.1.16).

Im Abschnitt 7.2 geht es um den zentralen Begriff der *Orthogonalität*. Zwei Vektoren sind orthogonal, wenn sie senkrecht aufeinander stehen (7.2.3). Damit definieren wir das *orthogonale Komplement* $U^\perp$ eines Unterraums U (7.2.7) und beweisen, dass stets $V = U \oplus U^\perp$ gilt (7.2.8). ← 7.2

Im zweiten Teil definieren wir *Orthonormalbasen* (7.2.16): Das sind Basen aus Vektoren der Länge 1, die paarweise orthogonal sind. Die Koordinaten eines Vektors bzgl. einer Orthonormalbasis kann man einfach mit dem Skalarprodukt berechnen (7.2.20). Wir zeigen, dass jeder endlich-dimensionale Euklidische oder unitäre Vektorraum eine Orthonormalbasis besitzt (7.2.21) und leiten einen *Basisergänzungssatz* her (7.2.23).

Im letzten Teil besprechen wir das *Orthonormalisierungsverfahren von Gram-Schmidt* (7.2.24). Mit diesem Verfahren kann man eine Basis in eine Orthonormalbasis überführen (7.2.25).

Im Abschnitt 7.3 erläutern wir, was *adjungierte* lineare Abbildungen sind (7.3.1). Es ist wichtig, dass Sie diesen Begriff gut verstehen. Er wird unser wichtigstes Werkzeug für die Spektralsätze sein. Im ersten Teil besprechen wir Beispiele (7.3.4–7.3.8) und zeigen dann, dass es zu jeder linearen Abbildung zwischen endlich-dimensionalen Euklidischen oder unitären Vektorräumen immer eine eindeutige adjungierte Abbildung gibt (7.3.9). Im zweiten Teil studieren wir Matrixdarstellungen adjungierter Abbildungen (7.3.11) und wir lernen, dass die adjungierte und die duale Abbildung verwandt sind (7.3.14). ← 7.3

Einen Endomorphismus, der das Skalarprodukt erhält, nennt man *Isometrie*. Diese Abbildungen nehmen wir in Abschnitt 7.4 unter die Lupe. Wir zeigen, dass Isometrien Abstände erhalten (7.4.2) und lernen einige Beispiele kennen (7.4.4). Im Satz 7.4.6 beschreiben wir Isometrien mithilfe der adjungieren Abbildung. Wir zeigen, dass die Isometrien eine Gruppe bilden (7.4.8). ← 7.4

Im zweiten Teil befassen wir uns mit den *orthogonalen* und *unitären* Matrizen (7.4.9). Wir besprechen einige äquivalente Charakterisierungen (7.4.10, 7.4.11) und Beispiele (7.4.13–7.4.16). Die orthogonalen Matrizen bilden die *orthogonale Gruppe*, die unitären Matrizen bilden die *unitäre Gruppe* (7.4.18). Wir zeigen, dass eine Matrix genau dann orthogonal bzw. unitär ist, wenn ihre Spalten eine Orthonormalbasis bilden. Damit können wir alle orthogonalen (2×2)-Matrizen beschreiben (7.4.22).

Im kurzen letzten Teil stellen wir die Verbindung zwischen Isometrien und orthogonalen/unitären Matrizen her. Ein Endomorphismus ist genau dann eine Isometrie, wenn eine Matrixdarstellung zu einer Orthonormalbasis eine orthogonale/unitäre Matrix ist (7.4.24). Damit stellt sich heraus, dass Isometrien genau jene Endomorphismen sind, die Orthonormalbasen aus Orthonormalbasen abbilden (7.4.25).

7.5 → Die Spektralsätze besprechen wir im letzten Abschnitt 7.5. Nach Vorüberlegungen zu invarianten Unterräumen und dem orthogonalen Komplement (7.5.1, 7.5.2) befassen wir uns mit *normalen* Endomorphismen (7.5.3). Wir werden feststellen, dass die Eigenräume eines normalen Endomorphismus orthogonal aufeinander stehen (7.5.6, 7.5.7). Damit können wir die Spektralsätze (komplex und reell) für normale Endomorphismen beweisen (7.5.8, 7.5.10). Mit diesen Spektralsätzen leiten wir dann relativ einfach die Spektralsätze für unitäre Abbildungen und Matrizen her (7.5.12, 7.5.13) und besprechen ein Verfahren, um unitäre Matrizen zu diagonalisieren (7.5.14).

Im dritten Teil besprechen wir selbstadjungierte Endomorphismen (7.5.16) und ihre Matrixdarstellungen (7.5.18). Zum Abschluss besprechen wir die Spektralsätze für selbstadjungierte Endomorphismen, für symmetrische reelle Matrizen und für Hermite'sche Matrizen (7.5.21, 7.5.22, 7.5.23). Es ist wichtig, dass Sie die Bedeutung dieser Sätze und ihre Verbindung zu anderen Diagonalisierungssätzen im Text verstehen (7.5.24). Wir werden sehen, wie man in konkreten Beispielen vorgeht (7.5.27, 7.5.28).

7.1. Skalarprodukt und Norm

In diesem Abschnitt lernen wir Euklidische und unitäre Vektorräume kennen. Dabei handelt es sich um Vektorräume mit einem *Skalarprodukt*. Dieses Skalarprodukt wird es uns ermöglichen Abstände zu messen und von *orthogonalen Vektoren* zu sprechen.

I. Definition und Beispiele

7.1.1 **Definition** Sei V ein reeller Vektorraum. Ein *Skalarprodukt* $\langle \cdot, \cdot \rangle$ *auf* V ist eine positiv definite symmetrische Bilinearform auf V. Ein Paar $(V, \langle \cdot, \cdot \rangle)$ bestehend aus einem reellen Vektorraum und einem Skalarprodukt nennen wir einen *Euklidischen Vektorraum*.

Sei V ein komplexer Vektorraum. Ein *Skalarprodukt* $\langle \cdot, \cdot \rangle$ *auf* V ist eine positiv definite Hermite'sche Form. Ein Paar $(V, \langle \cdot, \cdot \rangle)$ bestehend aus einem komplexen Vektorraum und einem Skalarprodukt nennen wir einen *unitären Vektorraum*.

7.1.2 **Vereinbarung**. Mit $\mathbb{K}$ bezeichnen wir ab sofort immer entweder den Körper $\mathbb{R}$ oder den Körper $\mathbb{C}$, abhängig davon, ob wir einen Euklidischen oder unitären Vektorraum betrachten.

Um beide Fälle gleichzeitig entwickeln zu können, machen wir eine kleine Beobachtung: Ist $(V, \langle \cdot, \cdot \rangle)$ ein Euklidischer Vektorraum, dann gilt wie im unitären Fall

$$\langle u, v \rangle = \overline{\langle v, u \rangle} \text{ und } \langle au, v \rangle = \overline{a} \langle u, v \rangle,$$

denn reelle Zahlen verändern sich unter der komplexen Konjugation nicht. Wir können also stets die Rechenregeln für Hermite'sche Formen anwenden, wenn wir beide Fälle gleichzeitig betrachten.

Genauso gilt für jede reelle Matrix $S \in \mathrm{M}_{n,n}(\mathbb{R})$ die Gleichung $S^* = S^T$. Wir können also beim Rechnen mit Matrizen allgemein S^* verwenden, auch wenn wir im reellen Fall eigentlich die Transponierte S^T meinen.

7.1.3 **Achtung: Verwechslungsgefahr**. Die spitzen Klammern $\langle , \rangle$ haben in dieser Lektion zwei verschiedene Bedeutungen. Einerseits verwenden wir sie für das Skalarprodukt und andererseits für die lineare Hülle. Um die Verwechslungsgefahr zu senken, schreiben wir in dieser Lektion $\langle v_1, \dots, v_r \rangle^{\mathbb{K}}$ für die $\mathbb{K}$-lineare Hülle der Vektoren $v_1, \dots, v_r$.

7.1.4 **Beispiel**. (a) Auf dem reellen Vektorraum $\mathbb{R}^n$ ist

$$\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle = \sum_{i=1}^{n} x_i y_i$$

ein Skalarprodukt; siehe Beispiel 3.2.21. Dieses Skalarprodukt nennen wir das *Standardskalarprodukt* und $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$ nennen wir den kanonischen n-dimensionalen Euklidischen Vektorraum.

(b) Auf dem reellen Vektorraum $\mathbb{R}[X]$ aller Polynome ist

$$\langle f, g \rangle = \int_0^1 f(x)g(x) \,\mathrm{d}x$$

ein Skalarprodukt. In Beispiel 3.1.5 haben wir bereits gesehen, dass es sich um eine symmetrische Bilinearform handelt. Wir müssen noch prüfen, dass diese Form positiv definit ist. Es sei dazu $f \neq 0$ ein Polynom. Dann ist

$$\langle f, f \rangle = \int_0^1 f(x)^2 \,\mathrm{d}x$$

das Integral über einer stetigen nicht-negativen Funktion. Da jedes Polynom nur endlich viele Nullstellen hat, finden wir einen Punkt $t \in (0,1)$ mit $f(t) \neq 0$ also $f(t)^2 > 0$. Da die Polynomfunktion $x \mapsto f(x)^2$ stetig ist, ist $f(x)^2 > 0$ in einer offenen Umgebung von t und aus der Definition des Riemann-Integrals folgt damit $\int_0^1 f(x)^2 \,\mathrm{d}x > 0$.

(c) Auf dem komplexen Vektorraum $\mathbb{C}^n$ ist

$$\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle = \sum_{i=1}^{n} \overline{x_i} y_i$$

ein Skalarprodukt; siehe Beispiel 3.4.8. Dieses Skalarprodukt nennen wir das *Standardskalarprodukt* und $(\mathbb{C}^n, \langle \cdot, \cdot \rangle)$ nennen wir den kanonischen n-dimensionalen unitären Vektorraum.

7.1.5 **Beispiel**. Sei (V, σ) ein unitärer Vektorraum. Dann ist $(V_{\mathbb{R}}, \mathrm{Re}(\sigma))$ ein Euklidischer Vektorraum. In der Tat, aus Aufgabe 3.4.40 folgt, dass $\mathrm{Re}(\sigma)$ ein reelles Skalarprodukt ist.

Um noch weitere Beispiele zur Verfügung zu haben, wollen wir uns überlegen, wie man anhand einer Matrixdarstellung erkennt, ob eine symmetrische Bilinearform oder eine Hermite'sche Form positiv definit ist. Rufen wir uns in Erinnerung, was positiv definite Formen so besonders macht: *Die Einschränkung auf jeden Unterraum ist wieder positiv definit und damit ist jeder Unterraum regulär.*

7.1.6 **Notation**. Sei $A \in \mathrm{M}_{n,n}(K)$ eine quadratische Matrix. Wir schreiben $A_{\leq k}$ für die Matrix, die man erhält, wenn man die letzten $n-k$ Zeilen und Spalten aus A streicht. In anderen Worten ist $A_{\leq k} = (a_{ij})_{i,j \leq k}$ die linke obere $(k \times k)$-Untermatrix von A. Die Determinante $\det(A_{\leq k})$ ist somit ein k-reihiger Minor von A. Man nennt diese Minoren auch die *Hauptminoren*.

7.1.7 **Beispiel**. Es sei

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{pmatrix} \in \mathrm{M}_{4,4}(\mathbb{R}).$$

Mit der eingeführten Notation ergibt sich damit

$$A_{\leq 1} = 1, \quad A_{\leq 2} = \begin{pmatrix} 1 & 2 \\ 5 & 6 \end{pmatrix}, \quad A_{\leq 3} = \begin{pmatrix} 1 & 2 & 3 \\ 5 & 6 & 7 \\ 9 & 10 & 11 \end{pmatrix}.$$

7.1.8 **Hauptminorenkriterium.** *Sei $\mathbb{K}$ der Körper der reellen (bzw. komplexen) Zahlen. Es sei $A \in \mathrm{M}_{n,n}(\mathbb{K})$ die Matrixdarstellung einer symmetrischen Bilinearform (bzw. einer Hermite'schen Form) β. Die Form β ist genau dann positiv definit, wenn die Hauptminoren*

$$\det(A_{\leq k}) > 0$$

für alle $k \in \{1, \ldots, n\}$ erfüllen.

Beweis. Es sei $A = \mathrm{M}_{\mathcal{B}}(\beta)$ für eine geordnete Basis $\mathcal{B} = (v_1, \ldots, v_n)$ von V.

Nehmen wir zuerst an, dass β positiv definit ist. Dann gilt $A = S^* I_n S$ für eine invertierbare Matrix $S \in \mathrm{GL}_n(\mathbb{K})$; siehe 3.2.28 (bzw. 3.4.53 im Hermite'schen Fall). Aus dem Determinantenmultiplikationssatz sehen wir

$$\begin{aligned} \det(A) &= \det(S^*) \det(I_n) \det(S) \\ &= \overline{\det(S)} \det(S) && (4.2.15) \\ &= |\det(S)|^2 \end{aligned}$$

und da S invertierbar ist, ist $\det(S) \neq 0$ und wir schließen $\det(A) > 0$. Die Matrix $A_{\leq k}$ ist eine Matrixdarstellung der Einschränkung von β auf den Unterraum $U_k = \langle v_1, \ldots, v_k \rangle^{\mathbb{K}}$. Da $\beta|_{U_k}$ auf diesem Unterraum auch positiv definit ist, folgt genauso $\det(A_{\leq k}) > 0$.

Die umgekehrte Implikation beweisen wir mit vollständiger Induktion nach n. Für $n = 1$ ist die Aussage offensichtlich. Nehmen wir also an, dass die Aussage für $(n-1) \times (n-1)$ Matrizen gilt. Weiter nehmen wir an, dass $\det(A_{\leq k}) > 0$ für alle $k \in \{1, \ldots, n\}$ ist. Wir betrachten den Unterraum $U = \langle v_1, \ldots, v_{n-1} \rangle^{\mathbb{K}}$. Dann ist $A_{\leq n-1}$ eine Matrixdarstellung von $\beta|_U$ und aus der Induktionsannahme wissen wir, dass $\beta|_U$ positiv definit ist (man muss sich dazu klar machen, dass $(A_{\leq n-1})_{\leq k} = A_{\leq k}$ ist). Also ist U ein β-regulärer Unterraum. Wir erhalten somit $V = U \oplus W$ mit $W = \{w \in V \mid \beta(u, w) = 0 \text{ für alle } u \in U\}$ aus dem Zerlegungssatz 3.1.43 bzw. 3.4.42. Es gilt $n = \dim V = \dim U + \dim W = n - 1 + \dim W$, also ist W eindimensional.

Sei $W = \langle w \rangle$. Wir behaupten, dass $\beta(w, w) > 0$ ist. Die Matrixdarstellung von β zur Basis $\mathcal{B}' = (v_1, \ldots, v_{n-1}, w)$ hat die Form

$$\mathrm{M}_{\mathcal{B}'}(\beta) = \begin{pmatrix} A_{\leq n-1} & 0 \\ 0 & \beta(w, w) \end{pmatrix}$$

und es gilt $\det \mathrm{M}_{\mathcal{B}'}(\beta) = \beta(w, w) \cdot \det(A_{\leq n-1})$; siehe Satz 4.3.15. Aus der Transformationsformel wissen wir, dass je zwei Matrixdarstellungen von β (Hermite'sch) kongruent sind, d.h., es gilt $\mathrm{M}_{\mathcal{B}'}(\beta) = S^* A S$ für eine invertierbare Matrix S. Aus dem Determinantenmultiplikationssatz folgt dann $\beta(w, w) \det(A_{\leq n-1}) = |\det(S)|^2 \det(A)$. Da $\det(A) > 0$ ist, folgt $\beta(w, w) > 0$.

Sei nun $v \in V$ beliebig. Wir schreiben $v = u + \lambda w$ mit $u \in U$ und $\lambda \in \mathbb{K}$. Es gilt

$$\begin{aligned} \beta(v, v) &= \beta(u + \lambda w, u + \lambda w) = \beta(u, u) + \lambda \underbrace{\beta(u, w)}_{=0} + \overline{\lambda} \underbrace{\beta(w, u)}_{=0} + |\lambda|^2 \beta(w, w) \\ &= \underbrace{\beta(u, u)}_{\geq 0} + |\lambda|^2 \beta(w, w) \end{aligned}$$

Ist also $v \neq 0$, dann ist

- entweder $u \neq 0$ und $\beta(u, u) > 0$
- oder $\lambda \neq 0$ und $|\lambda|^2 \beta(w, w) > 0$.

In jedem Fall ist $\beta(v, v) > 0$ und damit ist β positiv definit. □

7.1.9 **Beispiel**. Wir behaupten, dass die symmetrische Bilinearform $\beta(u,v) = u^T Av$ mit

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 5 & -1 \\ 0 & -1 & 2 \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{R})$$

ein Skalarprodukt ist. Wir zeigen, dass β positiv definit ist, indem wir die Hauptminoren berechnen.

Es ist $A_{\leq 1} = 1$ und damit $\det A_{\leq 1} = 1 > 0$.

Weiter ist $A_{\leq 2} = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$ und mit der Formel aus 4.2.3 sehen wir $\det(A_{\leq 2}) = 5 - 4 = 1 > 0$.

Schließlich ist $A_{\leq 3} = A$. Wir berechnen die Determinante mit der Regel von Sarrus 4.2.6:

$$\det(A) = 10 - 1 - 8 = 1 > 0.$$

Aus dem Hauptminorenkriterium 7.1.8 schließen wir, dass β ein Skalarprodukt ist.

7.1.10 **Aufgabe**. Für $t \in \mathbb{R}$ definieren wir die Hermite'sche Matrix **L**

$$A_t = \begin{pmatrix} 2 & i & 1 \\ -i & 2 & 0 \\ 1 & 0 & t \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{C}).$$

Für welche Werte $t \in \mathbb{R}$ ist $\sigma_t(u,v) = u^* A_t v$ ein Skalarprodukt?

II. Die Norm

Der Grund, warum wir uns für Skalarprodukte interessieren ist, dass Sie einem reellen oder komplexen Vektorraum die nötige Zusatzstruktur geben um Geometrie zu betreiben. Insbesondere wollen wir Abstände messen. Aus dem Skalarprodukt definieren wir dafür zunächst eine *Norm*.

7.1.11 **Definition** Es sei $(V, \langle\cdot,\cdot\rangle)$ ein Euklidischer oder unitärer Vektorraum. Für $v \in V$ definieren wir die *Norm* von v durch

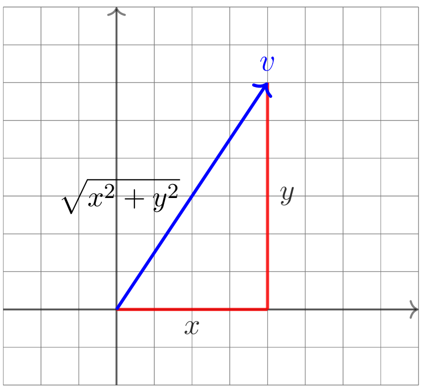
$$\|v\| = \sqrt{\langle v, v\rangle} \in \mathbb{R}.$$

Man nennt $\|v\|$ auch die *Länge* von v. Einen Vektor der Länge 1 nennt man *normierter Vektor*.
(Hinweis: Weil das Skalarprodukt positiv definit ist, ist $\langle v, v\rangle$ eine nicht-negative reelle Zahl.)

7.1.12 Beispiel. Wir betrachten den kanonischen Euklidischen Raum $(\mathbb{R}^2, \langle\cdot,\cdot\rangle)$. Für einen Vektor $v = \begin{pmatrix} x \\ y \end{pmatrix}$ gilt dann

$$\|v\| = \sqrt{\langle v, v\rangle} = \sqrt{x^2 + y^2}.$$

Nach dem Satz des Pythagoras ist die Norm $\|v\|$ genau der Abstand des Punktes $\begin{pmatrix} x \\ y \end{pmatrix}$ vom Nullpunkt in $\mathbb{R}^2$.



7.1.13 Eigenschaften der Norm. *Es sei $(V, \langle\cdot,\cdot\rangle)$ ein Euklidischer oder unitärer Vektorraum. Für alle $u, v \in V$ und $a \in \mathbb{K}$ gelten*

(N1) $\|v\| \geq 0$

(N2) $\|v\| = 0$ *genau dann, wenn* $v = 0$

(N3) Homogenität: $\|av\| = |a| \cdot \|v\|$

(N4) Dreiecksungleichung: $\|u + v\| \leq \|u\| + \|v\|$

Allgemein nennt man eine Abbildung auf einem reellen oder komplexen Vektorraum, die die Eigenschaften (N1)-(N4) erfüllt, eine *Norm*. Neben der von uns definierten Norm, gibt es auch diverse Normen, die *nicht* durch ein Skalarprodukt induziert werden.

Beweis. Zu (N1): Da ein Skalarprodukt immer positiv definit ist, folgt $\langle v, v\rangle \geq 0$ für alle $v \in V$ und damit ist auch die Wurzel $\|v\| \geq 0$.

Zu (N2): Sicherlich gilt $\|0\| = 0$. Ist $v \neq 0$, dann gilt $\langle v, v\rangle > 0$ und damit ist auch die Wurzel $\sqrt{\langle v, v\rangle}$ positiv.

Zu (N3): Für alle $a \in \mathbb{K}$ und $v \in V$ gilt

$$\sqrt{\langle av, av\rangle} = \sqrt{\overline{a}a\langle v, v\rangle} = \sqrt{|a|^2\langle v, v\rangle} = |a| \cdot \|v\|.$$

Den Beweis von (N4) liefern wir gleich nach, wenn wir die Ungleichung von Cauchy und Schwarz kennen. □

7.1.14 **Aufgabe**. Zeigen Sie, dass $\left\|\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right\| := \max(|x_1|, |x_2|)$ auf $\mathbb{R}^2$ die Eigenschaften (N1)-(N4) besitzt. **L**

Die folgende nützliche Ungleichung wird meist nach den Mathematikern Cauchy[1] und Schwarz[2] benannt. Aber auch der russische Mathematiker Wiktor Jakowlewitsch Bunjakowski (1804–1889) hat zur Entdeckung beigetragen.

7.1.15 **Cauchy-Schwarz Ungleichung.** *Es sei $(V, \langle\cdot,\cdot\rangle)$ ein Euklidischer oder unitärer Vektorraum. Für alle $u, v \in V$ gilt*

$$|\langle u, v\rangle| \leq \|u\| \cdot \|v\|.$$

Es ist $|\langle u, v\rangle| = \|u\| \cdot \|v\|$ genau dann, wenn u und v linear abhängig sind.

Beweis. Falls $u = 0$ oder $v = 0$ gilt, dann ist $|\langle u, v\rangle| = 0 = \|u\| \cdot \|v\|$ und es ist nichts weiter zu tun. Wir nehmen nun an, dass $u \neq 0$ und $v \neq 0$ ist. Wir können also durch $\|u\|$ und $\|v\|$ teilen und müssen in diesem Fall die Ungleichung

$$|\langle u', v'\rangle| \leq 1$$

herleiten mit $u' = \frac{1}{\|u\|}u$ und $v' = \frac{1}{\|v\|}v$. Die Vektoren u' und v' sind normiert, d.h, sie haben die Länge 1. Das folgt jeweils aus der Homogenität der Norm:

$$\|u'\| = \left\|\frac{1}{\|u\|}u\right\| = \frac{1}{\|u\|}\|u\| = 1.$$

Wir setzen nun $z = \overline{\langle u', v'\rangle}$. Aus der positiven Definitheit schließen wir

$$\begin{aligned} 0 \leq \langle u' - zv', u' - zv'\rangle &= \langle u', u'\rangle - \overline{z}\langle v', u'\rangle - z\langle u', v'\rangle + |z|^2\langle v', v'\rangle \\ &= \underbrace{\|u'\|^2}_{=1} - 2|z|^2 + |z|^2 \underbrace{\|v\|^2}_{=1} \\ &= 1 - |z|^2. \end{aligned}$$

[1]Augustin-Louis Cauchy: französischer Mathematiker, 1789–1857.
[2]Hermann Amandus Schwarz: deutscher Mathematiker, 1843–1921.

Es ist $|z|^2 = z\overline{z} = |\overline{z}|^2$; siehe 1.5.19. Nach Umstellen und Wurzelziehen erhalten wir somit die behauptete Ungleichung $|\langle u', v'\rangle| \leq 1$.

Für die zweite Aussage nehmen wir zunächst an, dass $|\langle u, v\rangle| = \|u\| \cdot \|v\|$ gilt. Ist $\|u\| \cdot \|v\| = 0$, dann ist $u = 0$ oder $v = 0$ und die Vektoren sind linear abhängig. Wir können also wieder $u \neq 0$ und $v \neq 0$ annehmen. Aus der gerade durchgeführten Rechnung erhalten wir eine Gleichheit genau dann, wenn $0 = \langle u' - zv', u' - zv'\rangle$ ist. Wegen der positiven Definitheit ist dann aber $0 = u' - zv' = \frac{1}{\|u\|}u - \frac{z}{\|v\|}v$ und damit sind u und v linear abhängig.

Nehmen wir umgekehrt an, die Vektoren u und v sind linear abhängig. Ohne Beschränkung der Allgemeinheit gilt dann $v = \lambda u$ für ein $\lambda \in \mathbb{K}$. Daraus folgt

$$|\langle u, v\rangle| = |\lambda| \langle u, u\rangle = |\lambda| \|u\|^2 = \|u\| \cdot \|v\|. \qquad \square$$

Mit der Cauchy-Schwarz Ungleichung können wir nun auch direkt die Dreiecksungleichung (N4) aus Satz 7.1.13 herleiten.

Beweis der Dreiecksungleichung. Gegeben sind $u, v \in V$. Wir berechnen

$$\begin{aligned}
\|u + v\|^2 = \langle u + v, u + v\rangle &= \langle u, u\rangle + \langle u, v\rangle + \langle v, u\rangle + \langle v, v\rangle \\
&= \|u\|^2 + \langle u, v\rangle + \overline{\langle u, v\rangle} + \|v\|^2 \\
&= \|u\|^2 + 2\,\mathrm{Re}(\langle u, v\rangle) + \|v\|^2.
\end{aligned}$$

Der Realteil einer komplexen Zahl ist immer höchstens so groß wie ihr Betrag (siehe Aufgabe 1.5.12). Mit der Ungleichung von Cauchy-Schwarz folgt nun

$$\begin{aligned}
\|u + v\|^2 &= \|u\|^2 + 2\,\mathrm{Re}(\langle u, v\rangle) + \|v\|^2 \\
&\leq \|u\|^2 + 2\,|\langle u, v\rangle| + \|v\|^2 \\
&\leq \|u\|^2 + 2\,\|u\| \cdot \|v\| + \|v\|^2 && (7.1.15) \\
&= (\|u\| + \|v\|)^2.
\end{aligned}$$

Wir ziehen die Wurzel und erhalten damit die Dreiecksungleichung. $\square$

Die Norm ermöglicht es uns, in Euklidischen und unitären Vektorräumen Abstände zu messen.

7.1.16 $\boxed{\textbf{Definition}}$ Sei $(V, \langle \cdot, \cdot \rangle)$ ein Euklidischer oder unitärer Vektorraum. Für $u, v \in V$ definieren wir

$$d(u, v) := \|u - v\|$$

und nennen dies den *Abstand* von u und v in V.

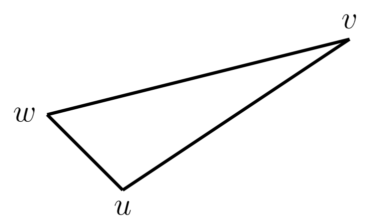
Der Abstand von u und v ist also die Länge des Vektors, der von v nach u zeigt. Der Abstand von u und v ist der Abstand von v und u, denn

$$d(u,v) = \|u - v\| = \|(-1) \cdot (v - u)\| = \|v - u\| = d(v,u);$$

die Abstandsfunktion ist also *symmetrisch*. Aus der Dreiecksungleichung (N4) erhalten wir nun die Ungleichung

$$d(u,v) = \|u - v\| = \|u - w + w - v\| \leq \|u - w\| + \|w - v\| = d(u,w) + d(w,v);$$

damit wird der Begriff *Dreiecksungleichung* verständlich: Der direkte Weg von u nach v ist immer höchstens so lang wie der Umweg von u nach w und dann von w nach v.



Eine Abstandsfunktion nennt man auch *Metrik*.

7.1.17 **Aufgabe**. Bestimmen Sie den Abstand der Vektoren **L**

$$u = \begin{pmatrix} 2 \\ 1 \\ -i \\ 0 \end{pmatrix} \quad \text{und} \quad v = \begin{pmatrix} 1-i \\ -1 \\ -1 \\ 1 \end{pmatrix}$$

im kanonischen 4-dimensionalen unitären Raum $(\mathbb{C}^4, \langle \cdot, \cdot \rangle)$.

7.2. Orthogonalität

Um Geometrie zu betreiben, möchten wir neben Abständen gerne auch Winkel messen. Wie wir gleich sehen werden, ist das in Euklidischen Räumen problemlos möglich. In unitären Räumen ist das Messen von Winkeln komplizierter, weil man sich erstmal entscheiden muss, was ein Winkel im Komplexen eigentlich ist: eine komplexe Zahl oder eine reelle Zahl? Diese Überlegungen können wir hier aber beiseite lassen, denn wir werden uns hauptsächlich für *rechte Winkel* interessieren. Hier definieren und untersuchen wir orthogonale Vektoren in Euklidischen und unitären Vektorräumen. Wir werden sehen, dass jeder Euklidische oder unitäre Vektorraum eine sogenannte „Orthonormalbasis" besitzt. Eine solche Basis besteht aus normierten und paarweise orthogonalen Vektoren.

I. Orthogonale Vektoren

7.2.1 **Definition** Sei $(V, \langle\cdot,\cdot\rangle)$ ein Euklidischer Vektorraum. Für Vektoren $u, v \in V \setminus \{0\}$ definiert man den *Winkel* $\alpha \in [0, \pi]$ zwischen u und v durch die Formel

$$\cos(\alpha) = \frac{\langle u, v\rangle}{\|u\| \cdot \|v\|}.$$

Durch die Ungleichung von Cauchy-Schwarz ist die rechte Seite immer eine Zahl im Intervall $[-1, 1]$. Der Cosinus liefert eine bijektive Abbildung $\cos\colon [0, \pi] \to [-1, 1]$, weshalb α durch diese Formel eindeutig bestimmt ist; genauer $\alpha = \arccos \frac{\langle u,v\rangle}{\|u\|\cdot\|v\|}$. Wie man aus der Formel leicht sieht, ändert sich der Winkel nicht, wenn man die Vektoren vertauscht oder die Vektoren mit positiven reellen Zahlen multipliziert.

In der Ebene liefert diese Formel tatsächlich den „Winkel", den man aus der Schulmathematik kennt. Genauer: Die Formel beschreibt immer den kleineren der beiden von den Vektoren gebildeten Winkel.

7.2.2 **Beispiel**. Wir betrachten den kanonischen Euklidischen Raum $(\mathbb{R}^2, \langle\cdot,\cdot\rangle)$. Was ist der Winkel zwischen den Vektoren $u = \begin{pmatrix}1\\1\end{pmatrix}$ und $v = \begin{pmatrix}1\\0\end{pmatrix}$? Mit der obigen Formel erhalten wir $\alpha = \arccos \frac{1}{\sqrt{2}} = \frac{\pi}{4}$. Es handelt sich also um einen Winkel von $45°$ (wie man anhand einer Skizze vermutet hätte).

Ein rechter Winkel ist ein Winkel von $\alpha = \frac{\pi}{2}$, d.h., von $90°$. Wann bilden zwei Vektoren einen rechten Winkel? Es gilt $\alpha = \frac{\pi}{2}$ genau dann, wenn $\cos(\alpha) = 0$ ist. Damit gelangt man zur folgenden wichtigen Definition (die problemlos auch für unitäre Vektorräume eingesetzt werden kann).

7.2.3 $\boxed{\textbf{Definition}}$ Es sei $(V, \langle\cdot,\cdot\rangle)$ ein Euklidischer oder unitärer Vektorraum. Zwei Vektoren $u, v \in V$ heißen *orthogonal*, wenn

$$\langle u, v\rangle = 0$$

gilt.

Wegen der Beziehung $\langle u, v\rangle = \overline{\langle v, u\rangle}$ sind u und v genau dann orthogonal, wenn v und u orthogonal sind. Es kommt also nicht auf die Reihenfolge der beiden Vektoren an.

7.2.4 **Aufgabe**. Suchen Sie einen Vektor $v \neq 0$, der orthogonal zu $u = \begin{pmatrix} 3 \\ -2 \end{pmatrix} \in \mathbb{R}^2$ im 2-dimensionalen kanonischen Euklidischen Raum ist. **L**

7.2.5 **Aufgabe**. Welche der folgenden Vektoren sind orthogonal zueinander im kanonischen unitären Raum $(\mathbb{C}^3, \langle\cdot,\cdot\rangle)$? **L**

$$u = \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}, \quad v = \begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix}, \quad w = \begin{pmatrix} 1 \\ i \\ 1+i \end{pmatrix}$$

7.2.6 **Beispiel**. Wir betrachten den reellen Vektorraum $\mathbb{R}[X]$ aller Polynome mit dem Skalarprodukt

$$\langle f, g\rangle = \int_0^1 f(x)g(x)\,\mathrm{d}x$$

aus Beispiel 7.1.4. Die Polynome $f = X$ und $g = 3X - 2$ sind orthogonal, denn mit dem Hauptsatz der Differential- und Integralrechnung erhalten wir

$$\langle f, g\rangle = \int_0^1 x(3x-2)\,\mathrm{d}x = \int_0^1 3x^2 - 2x\,\mathrm{d}x = (x^3 - x^2)\Big|_0^1 = 0.$$

7.2.7 $\boxed{\textbf{Definition}}$ Es sei $(V, \langle\cdot,\cdot\rangle)$ ein Euklidischer oder unitärer Vektorraum. Ist $S \subseteq V$ eine Teilmenge, dann definieren wir

$$S^\perp = \{v \in V \mid \langle u, v\rangle = 0 \text{ für alle } u \in S\}.$$

Ist U ein Unterraum, dann nennen wir $U^\perp$ das *orthogonale Komplement* zu U in V.

Das orthogonale Komplement ist also die Menge aller Vektoren, die zu allen Vektoren aus U orthogonal sind. Das orthogonale Komplement ist ein Spezialfall des Komplementes W, das wir aus den Zerlegungssätzen 3.1.43 (für Bilinearformen) und 3.4.42 (für Hermite'sche Formen) kennen. Aus diesen allgemeineren Überlegungen erhalten wir nun den folgenden Satz.

7.2.8 **Satz vom orthogonalen Komplement.** *Sei $(V, \langle\cdot,\cdot\rangle)$ ein endlich-dimensionaler Euklidischer oder unitärer Vektorraum. Sei $U \subseteq V$ ein Unterraum. Dann gelten*

(i) $U^\perp$ *ist ein Unterraum von* V*,*

(ii) $V = U \oplus U^\perp$*,*

(iii) $\dim_{\mathbb{K}} V = \dim_{\mathbb{K}} U + \dim_{\mathbb{K}} U^\perp$ *und*

(iv) $(U^\perp)^\perp = U$*.*

Beweis. Die ersten beiden Aussagen kennen wir aus den Zerlegungssätzen 3.1.43 (für Bilinearformen) und 3.4.42 (für Hermite'sche Formen). Die Dimensionsformel erhalten wir dann aus Satz 2.3.12. Für die letzte Aussage stellen wir fest, dass $U \subseteq (U^\perp)^\perp$ gilt, denn nach Konstruktion ist jeder Vektor aus U orthogonal zu jedem Vektor aus $U^\perp$. Die zweimalige Anwendung der Dimensionsformel liefert uns dann

$$\dim_{\mathbb{K}} U + \dim_{\mathbb{K}} U^\perp = \dim_{\mathbb{K}} V = \dim_{\mathbb{K}} U^\perp + \dim_{\mathbb{K}}(U^\perp)^\perp$$

und wir schließen $\dim_{\mathbb{K}} U = \dim_{\mathbb{K}}(U^\perp)^\perp$ und damit $U = (U^\perp)^\perp$. $\square$

7.2.9 **Aufgabe**. Sei $(V, \langle\cdot,\cdot\rangle)$ ein Euklidischer oder unitärer Vektorraum. Zeigen Sie: Ist U ein Unterraum und $S \subseteq U$ ein Erzeugendensystem von U, dann gilt $U^\perp = S^\perp$. **L**

7.2.10 **Beispiel**. Wir betrachten den 3-dimensionalen kanonischem Euklidischen Raum $(\mathbb{R}^3, \langle\cdot,\cdot\rangle)$. Es sei $u = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$ und $U = \langle u \rangle^{\mathbb{R}}$ sei der von u erzeugte Unterraum. Nach Aufgabe 7.2.9 ist $U^\perp = \{u\}^\perp$. Das orthogonale Komplement $U^\perp$ ist also der Lösungsraum einer linearen Gleichung:

$$0 = \langle u, x \rangle = x_1 + 2x_2 + x_3.$$

Mit einer kurzen Rechnung findet man eine Basis des Lösungsraumes:

$$U^\perp = \left\langle \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\rangle^{\mathbb{R}}.$$

7.2.11 **Aufgabe**. Sei $(V, \langle\cdot,\cdot\rangle)$ ein Euklidischer oder unitärer Vektorraum. Es seien $U_1, U_2 \subseteq V$ zwei Unterräume. Zeigen Sie: $(U_1 + U_2)^\perp = U_1^\perp \cap U_2^\perp$. **L**

II. Orthonormalbasen

7.2.12 **Definition** Sei $(V, \langle\cdot,\cdot\rangle)$ ein Euklidischer oder unitärer Vektorraum. Ein Tupel $S = (v_1, \ldots, v_k)$ von Vektoren aus V nennen wir *orthogonal*, wenn v_i und v_j orthogonal sind für alle $i \neq j$. Wir nennen S ein *Orthonormalsystem*, wenn S orthogonal ist und die Vektoren $v_1, \ldots, v_k$ normiert sind.

7.2.13 **Bemerkung**. Sei $S = (v_1', \ldots, v_k') \subseteq V$ ein orthogonales Tupel und es sei $v_i' \neq 0$ für alle i. Wir normieren die Vektoren aus S, d.h., wir setzen $v_i = \frac{1}{\|v_i'\|}v_i'$. Dann ist $(v_1, \ldots, v_k)$ ein Orthonormalsystem. In der Tat, die Vektoren v_i sind normiert und es gilt für $i \neq j$ auch

$$\langle v_i, v_j\rangle = \frac{1}{\|v_i'\| \cdot \|v_j'\|}\langle v_i', v_j'\rangle = 0.$$

7.2.14 **Beispiel**. Wir betrachten den kanonischen 2-dimensionalen Euklidischen Raum $(\mathbb{R}^2, \langle\cdot,\cdot\rangle)$.

(a) Das Tupel $(0, \begin{pmatrix}1\\2\end{pmatrix})$ ist orthogonal, ist aber kein Orthonormalsystem, denn keiner der beiden Vektoren ist normiert.

(b) Die Standardbasis (e_1, e_2) im kanonischen 2-dimensionalen Euklidischen Raum ist ein Orthonormalsystem.

(c) Das Tupel $(\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}\frac{3}{5}\\\frac{4}{5}\end{pmatrix})$ besteht aus normierten Vektoren. Diese sind aber nicht orthogonal, denn

$$\langle\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}\frac{3}{5}\\\frac{4}{5}\end{pmatrix}\rangle = \frac{3}{5} \neq 0.$$

Es ist also kein Orthonormalsystem.

7.2.15 **Lemma.** *Sei $(V, \langle\cdot,\cdot\rangle)$ ein Euklidischer oder unitärer Vektorraum. Ist $S = (v_1, \ldots, v_k)$ ein Orthonormalsystem, dann ist S linear unabhängig.*

Beweis. Wir nehmen an, dass
$$\sum_{i=1}^{k} a_i v_i = 0$$
mit Skalaren $a_1, \dots, a_k \in \mathbb{K}$ gilt. Dann gilt für alle $j \in \{1, \dots, k\}$

$$\begin{aligned} 0 = \langle v_j, 0 \rangle = \langle v_j, \sum_{i=1}^{k} a_i v_i \rangle & \\ = \sum_{i=1}^{k} a_i \langle v_j, v_i \rangle & \\ = a_j \cdot \langle v_j, v_j \rangle & \qquad (\langle v_j, v_i \rangle = 0 \text{ für } i \neq j). \end{aligned}$$

Wegen $\langle v_j, v_j \rangle = 1$ folgt daraus $a_j = 0$. Da j beliebig war, schließen wir, dass S linear unabhängig ist. □

7.2.16 **Definition** Sei $(V, \langle \cdot, \cdot \rangle)$ ein endlich-dimensionaler Euklidischer oder unitärer Vektorraum. Eine *Orthonormalbasis* ist ein Orthonormalsystem $\mathcal{B} = (v_1, \dots, v_n)$ von V, das auch eine Basis von V ist.

7.2.17 **Anmerkungen zur Definition**. (a) Da jedes Orthonormalsystem $S = (v_1, \dots, v_k)$ linear unabhängig ist, ist S immer eine Orthonormalbasis der linearen Hülle $\langle v_1, \dots, v_k \rangle^{\mathbb{K}} \subseteq V$. Ein Orthonormalsystem S ist also genau dann eine Orthonormalbasis von V, wenn es ein Erzeugendensystem von V ist.

(b) Eine geordnete Basis $\mathcal{B} = (v_1, \dots, v_n)$ ist genau dann eine Orthonormalbasis, wenn gilt
$$\langle v_i, v_j \rangle = \delta_{ij}$$
für alle $i, j \in \{1, \dots, n\}$. Dabei bezeichnet δ das Kronecker-Symbol (siehe 2.5.8). Das sieht man so:

- $\|v_i\|^2 = \langle v_i, v_i \rangle = 1$ gilt genau dann, wenn v_i normiert ist,
- $\langle v_i, v_j \rangle = 0$ gilt genau dann, wenn v_i und v_j orthogonal sind.

Das kann man auch so ausdrücken: $\mathcal{B}$ ist eine Orthonormalbasis genau dann, wenn die Matrixdarstellung $\mathrm{M}_{\mathcal{B}}(\langle \cdot, \cdot \rangle)$ des Skalarproduktes die Einheitsmatrix ist.

7.2.18 **Beispiel**. (a) Die Standardbasis $(e_1, \dots, e_n)$ im n-dimensionalen kanonischen Euklidischen Raum $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$ ist eine Orthonormalbasis.

(b) Die Standardbasis $(e_1, \dots, e_n)$ im n-dimensionalen kanonischen unitären Raum $(\mathbb{C}^n, \langle \cdot, \cdot \rangle)$ ist eine Orthonormalbasis.

(c) Die geordnete Basis $\mathcal{B} = \left(\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}, \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \right)$ des $\mathbb{R}^2$ ist eine Orthonormalbasis des kanonischen Euklidischen Raumes $(\mathbb{R}^2, \langle \cdot, \cdot \rangle)$.

Kennt man eine Orthonormalbasis eines Euklidischen oder unitären Vektorraumes, dann ist dadurch das Skalarprodukt bereits vollständig bestimmt. Das Ziel der folgenden Aufgabe ist es, sich diese Aussage im Euklidischen Fall klarzumachen.

7.2.19 **Aufgabe**. Es sei V ein endlich-dimensionaler reeller Vektorraum mit einer geordneten Basis $\mathcal{B}$. Zeigen Sie: Es gibt genau ein Skalarprodukt β auf V, sodass $\mathcal{B}$ eine Orthonormalbasis ist. **L**

Orthonormalbasen haben den großen Vorteil, dass es sehr einfach ist, die Koordinaten eines Vektors durch das Skalarprodukt zu bestimmen. Man muss also in diesem Fall kein Gleichungssystem lösen, um die Koordinaten zu bestimmen.

7.2.20 **Lemma.** *Sei $(V, \langle \cdot, \cdot \rangle)$ ein endlich-dimensionaler Euklidischer oder unitärer Vektorraum und sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine Orthonormalbasis. Für alle $u \in V$ gilt*

$$u = \sum_{i=1}^{n} \langle v_i, u \rangle \; v_i.$$

Beweis. Wir schreiben u in der Basis $\mathcal{B}$, d.h., wir schreiben

$$u = \sum_{i=1}^{n} a_i v_i$$

mit eindeutigen Skalaren $a_i \in \mathbb{K}$. Dann berechnen wir die Skalare a_j mithilfe des Skalarproduktes:

$$\langle v_j, u \rangle = \sum_{i=1}^{n} a_i \underbrace{\langle v_j, v_i \rangle}_{\delta_{ji}} = a_j. \qquad \square$$

Wir werden im Laufe dieser Lektion noch mehrfach sehen, dass wir in Euklidischen und unitären Vektorräumen möglichst immer nur mit Orthonormalbasen und nicht mit allgemeinen Basen rechnen sollten. Anders ausgedrückt: Ersetzen wir Vektorräume durch Vektorräume mit einem Skalarprodukt, dann sind nicht mehr alle Basen, sondern nur noch die Orthonormalbasen geeignet um sinnvoll in Koordinaten zu rechnen. Das bringt uns zu einer wichtigen Frage: Gibt es immer eine Orthonormalbasis?

7.2.21 **Satz zur Existenz von Orthonormalbasen.** *Jeder endlich-dimensionale Euklidische oder unitäre Vektorraum besitzt eine Orthonormalbasis.*

Beweis. Wir betrachten den Euklidischen Fall. Den unitären Fall behandelt man analog (Übungsaufgabe!). **L**

Sei (V, β) ein Euklidischer Vektorraum der Dimension n. Das Skalarprodukt β ist eine positiv definite reelle symmetrische Bilinearform. Es handelt sich also um eine Form vom Typ $(n, 0, 0)$ (siehe 3.2.28). Mit Korollar 3.2.29 finden wir eine geordnete Basis $\mathcal{B} = (v_1, \ldots, v_n)$ von V, sodass

$$\mathrm{M}_{\mathcal{B}}(\beta) = I_n$$

gilt. Das heißt, es gilt $\langle v_i, v_j \rangle = \delta_{ij}$ und $\mathcal{B}$ ist eine Orthonormalbasis von (V, β). $\square$

Eine Orthonormalbasis kann man mit dem Verfahren aus 3.2.30 (bzw. 3.4.47 im unitären Fall) berechnen. Wir werden im nächsten Abschnitt mit dem Verfahren von Gram-Schmidt noch einen anderen, direkteren Weg kennenlernen.

7.2.22 **Aufgabe**. Sei $(V, \langle \cdot, \cdot \rangle)$ ein endlich-dimensionaler Euklidischer oder unitärer Vektorraum und sei $U \subseteq V$ ein Unterraum. Zeigen Sie: Ist $(v_1, \ldots, v_k)$ eine Orthonormalbasis von U und $(v_{k+1}, \ldots, v_n)$ eine Orthonormalbasis von $U^\perp$, dann ist $(v_1, \ldots, v_n)$ eine Orthonormalbasis von V. **L**

Der nächste Satz ist ein Analogon des Basisergänzungssatzes.

7.2.23 **Basisergänzungssatz für Orthonormalbasen.** *Sei $(V, \langle \cdot, \cdot \rangle)$ ein endlich-dimensionaler Euklidischer oder unitärer Vektorraum. Jedes Orthonormalsystem $(v_1, \ldots, v_k)$ kann zu einer Orthonormalbasis $(v_1, \ldots, v_k, v_{k+1}, \ldots, v_n)$ von V ergänzt werden.*

Beweis. Es sei $U = \langle v_1, \ldots, v_k \rangle^{\mathbb{K}}$ der aufgespannte Unterraum. Da die Vektoren $(v_1, \ldots, v_k)$ linear unabhängig sind (siehe Lemma 7.2.15), handelt es sich um eine Orthonormalbasis von U. Mit Satz 7.2.21 wählen wir eine Orthonomalbasis $(v_{k+1}, \ldots, v_n)$ von $U^\perp$. In Aufgabe 7.2.22 haben wir gezeigt, dass dann $(v_1, \ldots, v_k, v_{k+1}, \ldots, v_n)$ eine Orthonormalbasis von V ist. $\square$

III. Das Verfahren von Gram-Schmidt

In diesem Abschnitt lernen wir ein einfaches Verfahren kennen, das es ermöglicht, aus einer Basis eines Euklidischen oder unitären Vektorraumes eine Orthonormalbasis zu berechnen. Das Verfahren geht auf Arbeiten von J. P. Gram[3] (1883) und E. Schmidt[4] (1907) zurück. Obwohl es schon zuvor in ähnlicher Form verwendet wurde (z.B. von Laplace) nennt man es heute das *Gram-Schmidt-Verfahren*.

7.2.24 **Gram-Schmidt Orthonormalisierungsverfahren**. Es sei $(V, \langle\cdot,\cdot\rangle)$ ein endlich-dimensionaler Euklidischer oder unitärer Vektorraum.

`Gegeben:` geordnete Basis $\mathcal{B} = (u_1, \ldots, u_n)$ von V.

`Gesucht:` Orthonormalbasis $\mathcal{C} = (v_1, \ldots, v_n)$ von V.

`Verfahren:`

Setze $v_1 = \frac{1}{\|u_1\|}u_1$.

Für $j = 2, \ldots, n$ wiederhole folgende zwei Schritte:

- Setze
$$v_j' := u_j - \sum_{i=1}^{j-1} \langle v_i, u_j\rangle v_i.$$
- Normiere: $v_j := \frac{1}{\|v_j'\|}v_j'$.

7.2.25 **Satz zum Orthonormalisierungsverfahren.** *Es sei $(V, \langle\cdot,\cdot\rangle)$ ein endlich-dimensionaler Euklidischer oder unitärer Vektorraum. Sei $\mathcal{B} = (u_1, \ldots, u_n)$ eine geordnete Basis von V. Sind $(v_1, \ldots, v_n)$ die durch das Gram-Schmidt-Verfahren bestimmten Vektoren, dann ist $(v_1, \ldots, v_j)$ eine Orthonormalbasis des Unterraumes $U_j = \langle u_1, \ldots, u_j\rangle^{\mathbb{K}}$ für alle $j \in \{1, \ldots, n\}$.*

Beweis. Wir beweisen die Aussage mit vollständiger Induktion nach j. Sei $j = 1$. Nach Konstruktion ist $v_1 = \frac{1}{\|u_1\|}u_1$ normiert und damit eine Orthonormalbasis des ein-dimensionalen Unterraumes $U_1 = \langle u_1\rangle^{\mathbb{K}}$.

Sei nun $j > 1$. Für den Induktionsschritt nehmen wir an, dass die Aussage für $j-1$ bekannt ist. Wir zeigen zunächst, dass der Vektor $v_j' = u_j - \sum_{i=1}^{j-1}\langle v_i, u_j\rangle v_i$ orthogonal zu $v_1, \ldots, v_{j-1}$ ist. Sei dazu $k \leq j-1$. Dann gilt
$$\langle v_k, v_j'\rangle = \langle v_k, u_j\rangle - \sum_{i=1}^{j-1} \langle v_i, u_j\rangle \cdot \langle v_k, v_i\rangle.$$

[3] Jørgen Pedersen Gram: dänischer Mathematiker, 1850–1916.
[4] Erhard Schmidt: deutsch-baltischer Mathematiker, 1879–1959.

Nach Induktionsvoraussetzung ist $\langle v_k, v_i \rangle = \delta_{ki}$ und somit ist

$$\langle v_k, v_j' \rangle = \langle v_k, u_j \rangle - \langle v_k, u_j \rangle = 0.$$

Der Vektor v_j entsteht aus v_j' durch Normieren, also ist v_j ebenfalls orthogonal zu $v_1, \ldots, v_{j-1}$. Das Tupel $(v_1, \ldots, v_j)$ ist also orthonormal. Nach Induktionsvoraussetzung gilt $U_{j-1} = \langle v_1, \ldots, v_{j-1} \rangle^{\mathbb{K}}$. Daraus schließen wir

$$\begin{aligned} U_j = \langle u_1, \ldots, u_j \rangle^{\mathbb{K}} &= U_{j-1} + \langle u_j \rangle^{\mathbb{K}} \\ &= U_{j-1} + \langle v_j' \rangle^{\mathbb{K}} && (v_j' \in u_j + U_{j-1}) \\ &= U_{j-1} + \langle v_j \rangle^{\mathbb{K}} && (\langle v_j \rangle^{\mathbb{K}} = \langle v_j' \rangle^{\mathbb{K}}) \\ &= \langle v_1, \ldots, v_{j-1}, v_j \rangle^{\mathbb{K}}. && \text{(Ind.-Ann.)} \end{aligned}$$

Das heißt, $(v_1, \ldots, v_j)$ ist eine Orthonormalbasis von U_j. □

7.2.26 **Beispiel**. Wir betrachten den kanonischen Euklidischen Raum $(\mathbb{R}^3, \langle \cdot, \cdot \rangle)$ und die geordnete Basis $\mathcal{B} = (u_1, u_2, u_3)$ mit

$$u_1 = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \quad u_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad u_3 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

Wir überführen die Basis mit dem Verfahren von Gram-Schmidt in eine Orthonormalbasis. Wir setzen

$$v_1 = \frac{1}{\|u_1\|} u_1 = \frac{1}{2} \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Dann berechnen wir v_2' als

$$v_2' = u_2 - \langle v_1, u_2 \rangle v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - 1 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

Wir normieren das Ergebnis und erhalten

$$v_2 = \frac{1}{\sqrt{2}} v_2' = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

Nun berechnen wir v_3' und finden

$$v_3' = u_3 - \langle v_1, u_3\rangle v_1 - \langle v_2, u_3\rangle v_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} - 0 - 0 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

Zum Abschluss normieren wir diesen Vektor:

$$v_3 = \frac{1}{\sqrt{2}} v_3' = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

Damit haben wir $\mathcal{B}$ in die Orthonormalbasis (v_1, v_2, v_3) überführt.

7.2.27 **Aufgabe**. Wir betrachten den Vektorraum $\mathbb{R}^3$ mit dem Skalarprodukt β aus Beispiel 7.1.9. Überführen Sie die Standardbasis mit dem Verfahren von Gram-Schmidt in eine Orthonormalbasis. **L**

7.2.28 **Aufgabe**. Verwenden Sie das Orthonormalisierungsverfahren um einen anderen Beweis des Basisergänzungssatzes für Orthonormalbasen zu geben. **L**

7.2.29 **Ausblick: Die Hadamard-Vermutung**. Mit den Begriffen aus dieser Lektion sind wir in der Lage ein bekanntes offenes Problem der Linearen Algebra zu formulieren: die Hadamard-Vermutung. Die Vermutung wurde nach dem französischen Mathematiker Jacques Hadamard (1865–1963) benannt, obwohl er die Vermutung nicht formuliert hat.

Eine Matrix $A = (a_{ij})_{i,j} \in \mathrm{M}_{n,n}(\mathbb{R})$ nennt man *Hadamard-Matrix*, wenn alle Einträge $a_{ij} \in \{1, -1\}$ erfüllen und die Spalten paarweise orthogonal sind. Orthogonal bezieht sich hier auf das Standardskalarprodukt. Zum Beispiel sind

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{pmatrix}$$

Hadamard-Matrizen. Solche Matrizen werden unter anderem zur Konstruktion fehlerkorrigierender Codes verwendet.

Ist $A \in \mathrm{M}_{n,n}(\mathbb{R})$ eine Hadamard-Matrix und $n > 2$, dann kann man zeigen, dass n durch 4 teilbar sein muss, d.h., $n = 4k$. Aber muss es für jedes k eine solche Hadamard-Matrix geben?

> **Hadamard-Vermutung:** Für jedes $k \in \mathbb{N}$ gibt es eine $(4k \times 4k)$ Hadamard-Matrix.

Es gibt inzwischen viele verschiedene Konstruktionen von Hadamard-Matrizen. Beispielsweise ist es nicht schwer zu sehen, dass es für jedes m eine $(2^m \times 2^m)$ Hadamard-Matrix gibt. Allerdings decken diese Konstruktionen nicht alle Werte von n ab. Zum Beispiel kennt man aktuell noch keine Hadamard-Matrix der Größe 668×668 .

7.3. Adjungierte Abbildungen

In diesem Abschnitt beginnen wir lineare Abbildungen zwischen Euklidischen oder unitären Vektorräumen zu untersuchen. Das Skalarprodukt liefert uns dabei eine zusätzliche Struktur. Aber wie kann man diese zusätzliche Struktur nutzen? Hier lernen wir dazu ein zentrales, neues Konzept kennen: die „Adjungierte“ einer linearen Abbildung.

I. Definition und Existenzsatz

7.3.1 **Definition** Es seien $(V, \langle\cdot,\cdot\rangle_V)$ und $(W, \langle\cdot,\cdot\rangle_W)$ zwei Euklidische oder unitäre Vektorräume. Zwei lineare Abbildungen $\varphi\colon V \to W$ und $\psi\colon W \to V$ nennt man *adjungiert* zueinander, wenn

$$\langle\varphi(v), w\rangle_W = \langle v, \psi(w)\rangle_V \tag{7.3.a}$$

für alle $v \in V$ und $w \in W$ gilt.

7.3.2 **Anmerkungen zur Definition**. Es seien $(V, \langle\cdot,\cdot\rangle_V)$ und $(W, \langle\cdot,\cdot\rangle_W)$ zwei Euklidische oder unitäre Vektorräume.

(a) Sind $\varphi\colon V \to W$ und $\psi\colon W \to V$ zueinander adjungiert, dann gilt auch

$$\langle w, \varphi(v)\rangle_W = \langle\psi(w), v\rangle_V \tag{7.3.b}$$

für alle $v \in V$ und $w \in W$ (Übungsaufgabe!). Man kann auch diese Gleichung zur Definition der Adjungiertheit verwenden. Es kommt also nicht auf die Reihenfolge der linearen Abbildungen an. **L**

(b) Es gibt *höchstens eine* zu $\varphi\colon V \to W$ adjungierte Abbildung. Nehmen wir an, dass ψ_1 und ψ_2 beide die Eigenschaft aus (7.3.a) erfüllen. Dann gilt für alle $v \in V$ und $w \in W$

$$\begin{aligned}\langle v, \psi_1(w) - \psi_2(w)\rangle_V &= \langle v, \psi_1(w)\rangle_V - \langle v, \psi_2(w)\rangle_V \\ &= \langle\varphi(v), w\rangle_W - \langle\varphi(v), w\rangle_W = 0.\end{aligned}$$

Da v beliebig war und das Skalarprodukt auf V regulär ist, folgt daraus $\psi_1(w) - \psi_2(w) = 0$ für alle $w \in W$. Das heißt, es ist $\psi_1 = \psi_2$.

Man spricht daher von *der adjungierten Abbildung* und bezeichnet diese mit φ^{ad}, wenn sie existiert.

(c) Gibt es zu φ eine adjungierte Abbildung φ^{ad}, dann gibt es auch eine zu φ^{ad} adjungierte Abbildung und es gilt

$$(\varphi^{\mathrm{ad}})^{\mathrm{ad}} = \varphi.$$

Das folgt aus Gleichung (7.3.b), denn damit erfüllt φ die definierende Gleichung (7.3.a).

7.3.3 **Adjungierte und Adjunkte**. Die Begriffe *Adjungierte* und *Adjunkte* darf man auf keinen Fall verwechseln. Es gibt *keinen* Zusammenhang zwischen der adjungierten Abbildung φ^{ad} und der Adjunkten einer Matrix $\mathrm{Adj}(A)$.

7.3.4 **Beispiel**. Es sei $\varphi\colon V \to W$ die Nullabbildung zwischen zwei Euklidischen oder unitären Vektorräumen. Dann ist φ^{ad} ebenfalls die Nullabbildung, denn es gilt

$$\langle \varphi(v), w\rangle_W = \langle 0, w\rangle_W = 0 = \langle v, 0\rangle_V$$

für alle $v \in V$ und $w \in W$.

7.3.5 **Beispiel**. Sei $(V, \langle\cdot,\cdot\rangle)$ ein Euklidischer oder unitärer Vektorraum und es sei $\varphi = \lambda\,\mathrm{id}_V$ mit $\lambda \in \mathbb{K}$. Dann ist $\varphi^{\mathrm{ad}} = \overline{\lambda}\,\mathrm{id}_V$, denn es gilt

$$\langle \varphi(v), w\rangle = \langle \lambda v, w\rangle = \overline{\lambda}\langle v, w\rangle = \langle v, \overline{\lambda} w\rangle$$

für alle $v, w \in V$.

7.3.6 **Aufgabe**. Es seien $(U, \langle\cdot,\cdot\rangle_U)$, $(V, \langle\cdot,\cdot\rangle_V)$ und $(W, \langle\cdot,\cdot\rangle_W)$ drei Euklidische oder unitäre Vektorräume mit linearen Abbildungen $\psi\colon U \to V$ und $\varphi\colon V \to W$. Zeigen Sie: $(\varphi \circ \psi)^{\mathrm{ad}} = \psi^{\mathrm{ad}} \circ \varphi^{\mathrm{ad}}$. **L**

Wir werden uns im Weiteren nur mit adjungierten linearen Abbildungen auf endlichdimensionalen Euklidischen oder unitären Vektorräumen befassen. Adjungierte Abbildungen spielen darüber hinaus aber auch in der Funktionalanalysis und ihren Anwendungen eine große Rolle. In den folgenden beiden Beispielen soll dieser wichtige Zusammenhang kurz angedeutet werden.

7.3.7 **Beispiel**. Wir betrachten den Vektorraum $W = \mathbb{R}[X]$ aller reellen Polynome mit dem Skalarprodukt aus Beispiel 7.1.4 (b) und den Vektorraum $V = \mathbb{R}$ mit dem Standardskalarprodukt. Wir betrachten die lineare Abbildung $\varphi\colon V \to W$, die $t \in \mathbb{R}$ auf das konstante Polynom $t \in \mathbb{R}[X]$ abbildet. Es gilt für alle $t \in \mathbb{R}$ und $f \in \mathbb{R}[X]$

$$\langle \varphi(t), f\rangle = \int_0^1 t f(x)\,\mathrm{d}x = t\int_0^1 f(x)\,\mathrm{d}x = \langle t, \int_0^1 f(x)\,\mathrm{d}x\rangle.$$

Damit ist $\varphi^{\mathrm{ad}}\colon f \mapsto \int_0^1 f(x)\,\mathrm{d}x$ die zu φ adjungierte lineare Abbildung.

7.3.8 **Beispiel**. Wir betrachten den Vektorraum V aller glatten Funktionen $f\colon [0,1] \to \mathbb{R}$, sodass f und alle Ableitungen von f an den Randpunkten $\{0,1\}$ verschwinden. Auch auf diesem reellen Vektorraum ist

$$\langle f, g\rangle = \int_0^1 f(x)g(x)\,\mathrm{d}x$$

ein Skalarprodukt. Die Ableitung $D\colon V \to V$ mit $D\colon f \mapsto f'$ ist eine lineare Abbildung. In diesem Fall ist $-D$ die zu D adjungierte Abbildung. Das sieht man durch partielle Integration [MG, 20.2.9]:

$$\begin{aligned}\langle D(f), g\rangle &= \int_0^1 f'(x)g(x)\,\mathrm{d}x \\ &= fg\Big|_0^1 - \int_0^1 f(x)g'(x)\,\mathrm{d}x && \text{(partielle Int.)} \\ &= \int_0^1 f(x)\cdot(-g'(x))\,\mathrm{d}x = \langle f, -D(g)\rangle.\end{aligned}$$

Nun wollen wir zeigen, dass in endlich-dimensionalen Euklidischen oder unitären Vektorräumen adjungierte Abbildungen immer existieren.

7.3.9 **Satz von der adjungierten Abbildung.** *Es seien $(V, \langle\cdot,\cdot\rangle_V)$ und $(W, \langle\cdot,\cdot\rangle_W)$ zwei endlich-dimensionale Euklidische oder unitäre Vektorräume. Ist $\varphi\colon V \to W$ eine lineare Abbildung, dann gibt es eine eindeutige adjungierte lineare Abbildung $\varphi^{ad}\colon W \to V$, d.h., es gilt*

$$\langle \varphi(v), w\rangle_W = \langle v, \varphi^{ad}(w)\rangle_V$$

für alle $v \in V$ und $w \in W$.

Beweis. Dass die adjungierte Abbildung eindeutig ist, haben wir schon in 7.3.2 gesehen. Wir müssen zeigen, dass die Abbildung φ^{ad} existiert. Dazu wählen wir eine Orthonormalbasis $\mathcal{B} = (v_1, \dots, v_n)$ von V. Sei $w \in W$. Um $\varphi^{\mathrm{ad}}(w)$ zu definieren, machen wir einen Ansatz und schreiben $\varphi^{\mathrm{ad}}(w)$ als Linearkombination der Vektoren in $\mathcal{B}$. Nach Lemma 7.2.20 gilt

$$\varphi^{\mathrm{ad}}(w) = \sum_{i=1}^n \langle v_i, \varphi^{\mathrm{ad}}(w)\rangle_V \, v_i.$$

Wenn also φ^{ad} die Eigenschaft aus (7.3.a) hat, dann muss $\langle v_i, \varphi^{\mathrm{ad}}(w)\rangle_V = \langle \varphi(v_i), w\rangle_W$ sein. Damit muss insbesondere folgende Gleichung erfüllt sein:

$$\varphi^{\mathrm{ad}}(w) = \sum_{i=1}^n \langle \varphi(v_i), w\rangle_W \, v_i.$$

Da die rechte Seite nur von φ und der Orthonormalbasis $\mathcal{B}$ abhängt, wählen wir diese Gleichung nun einfach als Definition von φ^{ad}!

Wir prüfen, dass φ^{ad} damit eine lineare Abbildung ist. Seien dazu $w_1, w_2 \in W$ und $a \in \mathbb{K}$ beliebig. Dann gilt

$$\begin{aligned}
\varphi^{\mathrm{ad}}(aw_1 + w_2) &= \sum_{i=1}^{n} \langle \varphi(v_i), aw_1 + w_2 \rangle_W \; v_i \\
&= \sum_{i=1}^{n} \big(a\langle \varphi(v_i), w_1 \rangle_W + \langle \varphi(v_i), w_2 \rangle_W \big) \; v_i && \text{(linear in 2. Komp.)} \\
&= a \sum_{i=1}^{n} \langle \varphi(v_i), w_1 \rangle_W \; v_i + \sum_{i=1}^{n} \langle \varphi(v_i), w_2 \rangle_W \; v_i \\
&= a\varphi^{\mathrm{ad}}(w_1) + \varphi^{\mathrm{ad}}(w_2).
\end{aligned}$$

Aus Aufgabe 2.2.18 schließen wir, dass φ^{ad} linear ist.

Nun zeigen wir, dass φ^{ad} Gleichung (7.3.a) erfüllt. Es seien $v \in V$ und $w \in W$ beliebig. Dann gilt

$$\begin{aligned}
\langle v, \varphi^{\mathrm{ad}}(w) \rangle_V &= \Big\langle v, \sum_{i=1}^{n} \langle \varphi(v_i), w \rangle_W \; v_i \Big\rangle_V \\
&= \sum_{i=1}^{n} \langle \varphi(v_i), w \rangle_W \cdot \langle v, v_i \rangle_V \\
&= \Big\langle \sum_{i=1}^{n} \overline{\langle v, v_i \rangle_V} \; \varphi(v_i), w \Big\rangle_W && \text{(3.4.2)} \\
&= \Big\langle \varphi\Big(\sum_{i=1}^{n} \langle v_i, v \rangle_V \; v_i \Big), w \Big\rangle_W && (\varphi \text{ linear, } \langle v_i, v \rangle = \overline{\langle v, v_i \rangle}) \\
&= \langle \varphi(v), w \rangle_W. && \text{(Lemma 7.2.20)}
\end{aligned}$$

□

II. Matrixdarstellung

Um mit konkreten Beispielen rechnen zu können, müssen wir zunächst die Matrixdarstellungen von adjungierten Abbildungen verstehen. Dabei ist es wichtig sich in Erinnerung zu rufen, dass wir in Euklidischen und unitären Vektorräumen immer an Matrixdarstellungen bezüglich *Orthonormalbasen* interessiert sind. Aber wie kann man eine Matrixdarstellung einer Abbildung bezüglich einer Orthonormalbasis berechnen? Das nächste Lemma liefert uns dazu eine einfache Formel.

7.3.10 **Lemma.** *Es seien $(V, \langle\cdot,\cdot\rangle_V)$ und $(W, \langle\cdot,\cdot\rangle_W)$ zwei endlich-dimensionale Euklidische oder unitäre Vektorräume. Sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine geordnete Basis von V und sei $\mathcal{C} = (w_1, \ldots, w_m)$ eine Orthonormalbasis von W. Für jede lineare Abbildung $\varphi\colon V \to W$ gilt*

$${}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\varphi) = \big(\langle w_i, \varphi(v_j)\rangle_W\big)_{i,j}.$$

Beweis. Um den Eintrag von ${}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\varphi)$ an der Position (i,j) zu ermitteln, müssen wir den Vektor $\varphi(v_j)$ als Linearkombination in der Basis $\mathcal{C}$ schreiben und den Koeffizienten vor w_i ablesen; siehe 2.2.23. Nach Lemma 7.2.20 gilt

$$\varphi(v_j) = \sum_{k=1}^{m} \langle w_k, \varphi(v_j)\rangle_W \, w_k.$$

Daraus erhalten wir direkt ${}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\varphi) = \big(\langle w_i, \varphi(v_j)\rangle_W\big)_{i,j}$. □

7.3.11 **Satz.** *Es seien $(V, \langle\cdot,\cdot\rangle_V)$ und $(W, \langle\cdot,\cdot\rangle_W)$ zwei endlich-dimensionale Euklidische oder unitäre Vektorräume. Sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine Orthonormalbasis von V und sei $\mathcal{C} = (w_1, \ldots, w_m)$ eine Orthonormalbasis von W. Sei $\varphi\colon V \to W$ eine lineare Abbildung.*

(i) Sind V und W Euklidisch, dann gilt

$${}_{\mathcal{B}}\mathrm{M}_{\mathcal{C}}(\varphi^{ad}) = {}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\varphi)^T.$$

(ii) Sind V und W unitär, dann gilt

$${}_{\mathcal{B}}\mathrm{M}_{\mathcal{C}}(\varphi^{ad}) = {}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\varphi)^*.$$

Beweis. Ist A eine Matrix mit reellen Einträgen, dann gilt $A^T = A^*$ – die reellen Zahlen verändern sich unter komplexer Konjugation nicht. Wir können also mit den Hinweisen aus 7.1.2 den Beweis in beiden Fällen parallel führen.

Es sei ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{C}}(\varphi^{\mathrm{ad}}) = (a_{ij})_{i,j}$. Da $\mathcal{B}$ eine Orthonormalbasis ist, können wir Lemma 7.3.10 auf die adjungierte Abbildung φ^{ad} anwenden:

$$a_{ij} = \langle v_i, \varphi^{\mathrm{ad}}(w_j)\rangle_V = \langle \varphi(v_i), w_j\rangle_W = \overline{\langle w_j, \varphi(v_i)\rangle_W}.$$

Dieser Ausdruck beschreibt (nach 7.3.10, denn $\mathcal{C}$ ist eine Orthonormalbasis) genau den Eintrag an der Position (i,j) der Matrix ${}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\varphi)^*$ □

7.3.12 **Beispiel**. Wir betrachten die lineare Abbildung

$$\varphi\colon \mathbb{C}^3 \to \mathbb{C}^2, \quad x \mapsto \begin{pmatrix} i & 1+i & 0 \\ 4 & 1 & 1-2i \end{pmatrix} x$$

zwischen den kanonischen unitären Vektorräumen $(\mathbb{C}^3, \langle\cdot,\cdot\rangle)$ und $(\mathbb{C}^2, \langle\cdot,\cdot\rangle)$. Die Standardbasen sind Orthonormalbasen und die Matrix

$$A = \begin{pmatrix} i & 1+i & 0 \\ 4 & 1 & 1-2i \end{pmatrix} \in \mathrm{M}_{2,3}(\mathbb{C})$$

ist die Matrixdarstellung von φ bezüglich der Standardbasen. Aus Satz 7.3.11 wissen wir, dass die Matrixdarstellung von $\varphi^{\mathrm{ad}}\colon \mathbb{C}^2 \to \mathbb{C}^3$ bezüglich der Standardbasen durch

$$A^* = \begin{pmatrix} -i & 4 \\ 1-i & 1 \\ 0 & 1+2i \end{pmatrix} \in \mathrm{M}_{3,2}(\mathbb{C})$$

gegeben ist.

7.3.13 **Aufgabe**. Zeigen Sie mit einem Beispiel, dass die Formeln aus Satz 7.3.11 im Allgemeinen falsch sind, wenn $\mathcal{B}$ oder $\mathcal{C}$ keine Orthonormalbasis ist. **L**

7.3.14 **Ausblick: Adjungierte und duale Abbildung**. Es besteht ein enger Zusammenhang zwischen der adjungierten linearen Abbildung φ^{ad} und der dualen linearen Abbildung φ^*, die wir aus 2.5.16 kennen. Mit diesem Zusammenhang wollen wir uns nur für Euklidische Vektorräume vertraut machen. In unitären Vektorräumen existieren ähnliche Verbindungen, es fehlen uns aber ein paar Begriffe um das konkret zu machen.

Es seien (V, β) und (W, γ) zwei endlich-dimensionale Euklidische Vektorräume. Ist $\varphi\colon V \to W$ eine lineare Abbildung, dann gilt mit (7.3.a)

$$\gamma(\varphi(v), w) = \beta(v, \varphi^{\mathrm{ad}}(w))$$

für alle $v \in V$ und $w \in W$.

Die Gleichung oben lässt sich mit der Notation aus Satz 3.1.31 so umschreiben:

$$\gamma^{(2)}(w)(\varphi(v)) = \beta^{(2)}(\varphi^{\mathrm{ad}}(w))(v).$$

Nach Definition der dualen Abbildung 2.5.16 gilt damit $\gamma^{(2)}(w)(\varphi(v)) = \varphi^*(\gamma^{(2)}(w))(v)$. Da diese Beziehung für alle $v \in V$ und $w \in W$ gilt, folgt damit die Gleichung

$$\varphi^* \circ \gamma^{(2)} = \beta^{(2)} \circ \varphi^{\mathrm{ad}}.$$

Man sagt, das folgende Diagramm *kommutiert*

$$\begin{tikzcd}
W \arrow[r, "\varphi^{\mathrm{ad}}"] \arrow[d, "\gamma^{(2)}"'] & V \arrow[d, "\beta^{(2)}"] \\
W^* \arrow[r, "\varphi^*"'] & V^*
\end{tikzcd}$$

Da β und γ regulär sind, induzieren diese Isomorphismen $\beta^{(2)} \colon V \to V^*$ und $\gamma^{(2)} \colon W \to W^*$ (siehe 3.1.38). Wir können also durch Komposition mit der jeweiligen Umkehrabbildung sowohl φ^{ad} durch φ^* beschreiben

$$\varphi^{\mathrm{ad}} = (\beta^{(2)})^{-1} \circ \varphi^* \circ \gamma^{(2)},$$

als auch φ^* durch φ^{ad} beschreiben

$$\varphi^* = \beta^{(2)} \circ \varphi^{\mathrm{ad}} \circ (\gamma^{(2)})^{-1}.$$

Die adjungierte Abbildung kann man sich also als Zwillingsschwester der dualen Abbildung vorstellen, die ohne den Dualraum auskommt. Umgekehrt bietet die duale Abbildung manchmal die Möglichkeit, Ergebnisse aus der Theorie der Euklidischen und unitären Vektorräume auf allgemeine Vektorräume (und damit auch auf andere Körper als $\mathbb{R}$ und $\mathbb{C}$) zu übertragen. Zum Beispiel ist Satz 2.5.20 eine Verallgemeinerung von Satz 7.3.11.

7.4. Orthogonale und unitäre Abbildungen

In diesem Abschnitt untersuchen wir Endomorphismen von Euklidischen oder unitären Vektorräumen, die das Skalarprodukt erhalten. Man nennt diese Abbildungen „Isometrien“.

I. Isometrien

7.4.1 **Definition** Es sei $(V, \langle\cdot,\cdot\rangle)$ ein endlich-dimensionaler Euklidischer oder unitärer Vektorraum. Einen Endomorphismus $\varphi\colon V \to V$ nennt man *Isometrie*, wenn für alle $u, v \in V$ gilt

$$\langle\varphi(u), \varphi(v)\rangle = \langle u, v\rangle.$$

Isometrien Euklidischer Vektorräume nennt man auch *orthogonale Abbildungen*. Isometrien unitärer Vektorräume nennt man auch *unitäre Abbildungen*.

Die Menge aller Isometrien von $(V, \langle\cdot,\cdot\rangle)$ bezeichnen wir mit $\operatorname{Isom}(V, \langle\cdot,\cdot\rangle)$ oder kurz $\operatorname{Isom}(V)$.

7.4.2 **Isometrien erhalten Abstände**. Sei $\varphi\colon V \to V$ eine Isometrie, dann gilt

$$\begin{aligned}
d(\varphi(u), \varphi(v)) &= \|\varphi(u) - \varphi(v)\| \\
&= \|\varphi(u - v)\| && (\varphi \text{ linear}) \\
&= \sqrt{\langle\varphi(u - v), \varphi(u - v)\rangle} && (\text{Def. der Norm}) \\
&= \sqrt{\langle u - v, u - v\rangle} && (\varphi \text{ Isometrie}) \\
&= \|u - v\| = d(u, v)
\end{aligned}$$

für alle $u, v \in V$. Das heißt, jede Isometrie erhält Abstände (7.1.16). Genauso kann man zeigen, dass Isometrien Orthogonalität und (im reellen Fall) Winkel erhalten. Isometrien erhalten also die Geometrie. In der Schulmathematik werden Isometrien auch „Kongruenzabbildungen“ genannt.

Mit dem Begriff „Isometrie“ bezeichnet man auch allgemeiner abstandserhaltende Abbildungen zwischen metrischen Räumen. Hier im Lehrtext verstehen wir unter Isometrien aber immer nur Endomorphismen auf Euklidischen oder unitären Vektorräumen im Sinne von Definition 7.4.1.

7.4.3 **Aufgabe**. Sei $\varphi \in \operatorname{Isom}(V, \langle\cdot,\cdot\rangle)$ und seien $u, v \in V$. Zeigen Sie: u und v sind genau dann orthogonal, wenn $\varphi(u)$ und $\varphi(v)$ orthogonal sind. **L**

7.4.4 **Beispiel**. Sei $(V, \langle\cdot,\cdot\rangle)$ ein endlich-dimensionaler Euklidischer oder unitärer Vektorraum.

(a) Die identische Abbildung id_V ist eine Isometrie, denn es gilt $\langle \mathrm{id}_V(u), \mathrm{id}_V(v)\rangle = \langle u, v\rangle$ für alle $u, v \in V$.

(b) Die Punktspiegelung im Ursprung $s\colon x \mapsto -x$ auf V ist eine Isometrie, denn es ist

$$\langle s(u), s(v)\rangle = \langle -u, -v\rangle = (-1)^2 \cdot \langle u, v\rangle = \langle u, v\rangle$$

für alle $u, v \in V$.

(c) Sei $(V, \langle\cdot,\cdot\rangle)$ ein unitärer Vektorraum. Ist $\omega \in \mathbb{C}$ eine komplexe Zahl mit $|\omega| = 1$, dann ist $\varphi\colon x \mapsto \omega x$ eine unitäre Abbildung. Um das zu sehen macht man für $u, v \in V$ die folgende Rechnung:

$$\begin{aligned}\langle \varphi(u), \varphi(v)\rangle &= \langle \omega u, \omega v\rangle = \omega\overline{\omega} \cdot \langle u, v\rangle \\ &= |\omega|^2 \cdot \langle u, v\rangle = \langle u, v\rangle.\end{aligned}$$

(d) Ist $w \in V$ ein normierter Vektor (d.h., $\|w\| = 1$), dann ist die Abbildung $\sigma_w\colon V \to V$ mit $\sigma_w(x) = x - 2\langle w, x\rangle w$ eine Isometrie. Wir bemerken zuerst, dass σ_w linear ist, weil das Skalarprodukt in der zweiten Komponente linear ist. Es seien nun $u, v \in V$ beliebig. Dann gilt

$$\begin{aligned}\langle \sigma_w(u), \sigma_w(v)\rangle &= \big\langle u - 2\langle w, u\rangle w, v - 2\langle w, v\rangle w\big\rangle \\ &= \langle u, v\rangle - 2\underbrace{\overline{\langle w, u\rangle}}_{=\langle u,w\rangle} \cdot \langle w, v\rangle - 2\langle w, v\rangle \cdot \langle u, w\rangle + 4\overline{\langle w, u\rangle} \cdot \langle w, v\rangle \cdot \underbrace{\langle w, w\rangle}_{=1} \\ &= \langle u, v\rangle - 4\langle w, v\rangle \cdot \langle u, w\rangle + 4\langle w, v\rangle \cdot \langle u, w\rangle \\ &= \langle u, v\rangle.\end{aligned}$$

Also ist σ_w eine Isometrie.

Wie kann man sich σ_w vorstellen? Ist x orthogonal zu w, dann ist $\langle w, x\rangle = 0$ und damit $\sigma_w(x) = x$. Dazu gilt $\sigma_w(w) = -w$. Die Abbildung σ_w ist die Spiegelung am orthogonalen Komplement $(\langle w\rangle^{\mathbb{K}})^{\perp}$ von w.

7.4.5 **Aufgabe**. Es sei (V, σ) ein unitärer Vektorraum mit einer unitären Abbildung $\varphi\colon V \to V$. Wir betrachten den Euklidischen Vektorraum $(V_{\mathbb{R}}, \mathrm{Re}(\sigma))$ aus Beispiel 7.1.5. Zeigen Sie, dass $\varphi\colon V_{\mathbb{R}} \to V_{\mathbb{R}}$ eine orthogonale Abbildung ist. **L**

7.4.6 **Satz.** *Es sei $(V, \langle\cdot,\cdot\rangle)$ ein endlich-dimensionaler Euklidischer oder unitärer Vektorraum und sei $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$. Dann sind folgende Aussagen äquivalent:*

(i) φ ist eine Isometrie,

(ii) φ ist ein Isomorphismus und $\varphi^{ad} = \varphi^{-1}$,

(iii) φ ist längentreu: Für alle $v \in V$ gilt $\|\varphi(v)\| = \|v\|$.

Beweis. „(i) $\Rightarrow$ (ii)“: Es sei φ eine Isometrie. Der Raum V ist endlich-dimensional, also ist es ausreichend die Injektivität von φ nachzuweisen (siehe [MG, 8.3.17]), um zu zeigen, dass φ ein Isomorphismus ist. Es sei $v \in \mathrm{Ker}(\varphi)$, d.h., $\varphi(v) = 0$. Dann gilt

$$\langle v, v\rangle = \langle \varphi(v), \varphi(v)\rangle = \langle 0, 0\rangle = 0.$$

Da das Skalarprodukt positiv definit ist, folgt daraus $v = 0$. Somit ist $\mathrm{Ker}(\varphi) = \{0\}$ und φ ist injektiv. Für alle $u, v \in V$ gilt

$$\langle u, v\rangle = \langle \varphi(u), \varphi(v)\rangle = \langle u, \varphi^{\mathrm{ad}}(\varphi(v))\rangle$$

und damit auch $0 = \langle u, v - \varphi^{\mathrm{ad}}(\varphi(v))\rangle$. Da u beliebig ist, folgt aus der Regularität des Skalarproduktes $v = \varphi^{\mathrm{ad}}(\varphi(v))$. Da v beliebig ist, schließen wir $\varphi^{\mathrm{ad}} \circ \varphi = \mathrm{id}_V$. Damit ist $\varphi^{\mathrm{ad}} = \varphi^{-1}$.

„(ii) $\Rightarrow$(iii) “: Angenommen φ ist ein Isomorphismus mit $\varphi^{\mathrm{ad}} = \varphi^{-1}$. Sei $v \in V$ beliebig. Dann gilt

$$\|\varphi(v)\|^2 = \langle \varphi(v), \varphi(v)\rangle = \langle v, \varphi^{-1} \circ \varphi(v)\rangle = \langle v, v\rangle = \|v\|^2$$

und damit $\|\varphi(v)\| = \|v\|$.

„(iii) $\Rightarrow$(i) “: Wir nehmen an, dass φ längentreu ist. Wir betrachten zuerst den Euklidischen Fall. Die Polarisationsformel 3.2.9 impliziert

$$\begin{aligned}
\langle \varphi(u), \varphi(v)\rangle &= \frac{1}{2}\Big(\|\varphi(u) + \varphi(v)\|^2 - \|\varphi(u)\|^2 - \|\varphi(v)\|^2\Big) && (3.2.9)\\
&= \frac{1}{2}\Big(\|(\varphi(u + v)\|^2 - \|\varphi(u)\|^2 - \|\varphi(v)\|^2\Big) && (\varphi \text{ linear})\\
&= \frac{1}{2}\Big(\|u + v\|^2 - \|u\|^2 - \|v\|^2\Big) && (\varphi \text{ längentreu})\\
&= \langle u, v\rangle && (3.2.9)
\end{aligned}$$

für alle $u, v \in V$. Also ist φ eine Isometrie.

Im unitären Fall bezeichnen wir das Skalarprodukt mit σ. Aus Aufgabe 7.4.5 und dem Euklidischen Fall folgt, dass $\varphi\colon V_{\mathbb{R}} \to V_{\mathbb{R}}$ eine Isometrie bezüglich $\mathrm{Re}(\sigma)$ ist. Da wir σ aber wie in Bemerkung 3.4.41 durch seinen Realteil ausdrücken können, folgt daraus auch, dass φ eine Isometrie bzgl. σ ist (hier geht ein, dass φ komplex-linear ist). $\square$

7.4.7 **Lemma.** *Sei* $(V, \langle\cdot,\cdot\rangle)$ *ein endlich-dimensionaler Euklidischer oder unitärer Vektorraum. Sind* $\varphi, \psi \in \operatorname{Isom}(V)$*, dann ist auch*

$$\varphi \circ \psi \in \operatorname{Isom}(V) \quad \textit{und} \quad \varphi^{-1} \in \operatorname{Isom}(V).$$

Beweis. Es seien $u, v \in V$ beliebig. Dann gilt

$$\langle \varphi(\psi(u)), \varphi(\psi(v)) \rangle = \langle \psi(u), \psi(v) \rangle = \langle u, v \rangle,$$

d.h., $\varphi \circ \psi$ ist eine Isometrie. Außerdem gilt

$$\langle \varphi^{-1}(u), \varphi^{-1}(v) \rangle = \langle \varphi(\varphi^{-1}(u)), \varphi(\varphi^{-1}(v)) \rangle = \langle u, v \rangle;$$

im zweiten Schritt wird dabei verwendet, dass φ eine Isometrie ist. $\square$

7.4.8 **Korollar.** *Sei* $(V, \langle\cdot,\cdot\rangle)$ *ein endlich-dimensionaler Euklidischer oder unitärer Vektorraum. Dann ist* $(\operatorname{Isom}(V, \langle\cdot,\cdot\rangle), \circ)$ *mit der Komposition* $\circ$ *von Abbildungen eine Gruppe.*

Man nennt diese Gruppe die *Isometriegruppe* von $(V, \langle\cdot,\cdot\rangle)$. In Euklidischen Vektorräumen spricht man auch von der *orthogonalen Gruppe*, in unitären Vektorräumen von der *unitären Gruppe*.

Beweis. Wir müssen nachweisen, dass die Gruppenaxiome aus Definition 1.1.1 erfüllt sind. Aus Lemma 7.4.7 folgt, dass die Komposition von Isometrien wieder eine Isometrie ist. Die Komposition ist also eine Verknüpfung auf der Menge $\operatorname{Isom}(V)$. Es ist bekannt, dass die Komposition von Abbildungen immer assoziativ ist; siehe [MG, 1.4.23]. Die identische Abbildung id_V ist eine Isometrie und ist das neutrale Element für die Komposition; Axiom (G2) ist also erfüllt. Aus Satz 7.4.6 wissen wir, dass jede Isometrie φ eine Umkehrabbildung φ^{-1} besitzt und nach Lemma 7.4.7 ist diese Umkehrabbildung wieder eine Isometrie. Damit ist $(\operatorname{Isom}(V), \circ)$ eine Gruppe. $\square$

II. Orthogonale und unitäre Matrizen

7.4.9 **Definition** Eine reelle quadratische Matrix $A \in \mathrm{M}_{n,n}(\mathbb{R})$ nennt man *orthogonal*, wenn A invertierbar ist und die Gleichung

$$A^T = A^{-1}$$

erfüllt. Die Menge aller orthogonalen $(n \times n)$-Matrizen bezeichnen wir mit $\mathrm{O}(n)$.

Eine komplexe quadratische Matrix $B \in \mathrm{M}_{n,n}(\mathbb{C})$, die invertierbar ist und der Gleichung

$$B^* = B^{-1}$$

genügt, nennt man *unitär*. Die Menge aller unitären $(n \times n)$-Matrizen bezeichnen wir mit $\mathrm{U}(n)$.

Bevor wir uns Beispiele anschauen, leiten wir einige nützliche äquivalente Beschreibungen orthogonaler und unitärer Matrizen her.

7.4.10 **Lemma.** *Sei $A \in \mathrm{M}_{n,n}(\mathbb{R})$. Die folgenden Aussagen sind äquivalent:*

(i) A ist orthogonal.

(ii) A^T ist orthogonal.

(iii) Es gilt $A^T A = I_n$.

(iv) Es gilt $AA^T = I_n$.

Beweis. Wir überlegen uns zuerst, warum (i) die Aussagen (iii) und (iv) impliziert. Ist A orthogonal, dann gilt $A^{-1} = A^T$ und damit auch $A^T A = A^{-1}A = I_n$ und $AA^T = AA^{-1} = I_n$.

„(i) $\Rightarrow$ (ii)“: Ist A orthogonal, dann ist A invertierbar und $A^T = A^{-1}$. Aus Aufgabe 1.2.34 wissen wir, dass A^T dann ebenfalls invertierbar ist und es gilt

$$(A^T)^{-1} = (A^{-1})^T = (A^T)^T.$$

Also ist A^T eine orthogonale Matrix.

„(ii) $\Rightarrow$ (i)“: Folgt direkt aus „(i) $\Rightarrow$ (ii)“, denn $(A^T)^T = A$; siehe 1.2.32.

„(iii) $\Rightarrow$ (i)“: Wir nehmen an es gilt $A^T A = I_n$. Aus den Mathematischen Grundlagen [MG, 4.5.6] wissen wir, dass A dann invertierbar ist und es gilt $A^{-1} = A^T$.

Mit demselben Argument folgt „(iv) $\Rightarrow$ (ii)“ und damit sind alle Aussagen äquivalent. $\square$

7.4.11 Aufgabe. Sei $B \in \mathrm{M}_{n,n}(\mathbb{C})$. Zeigen Sie, dass die folgenden Aussagen äquivalent sind: **L**

(i) B ist unitär.

(ii) B^* ist unitär.

(iii) Es gilt $B^*B = I_n$.

(iv) Es gilt $BB^* = I_n$.

7.4.12 Aufgabe. Zeigen Sie: Ist $B \in \mathrm{U}(n)$, dann ist auch $B^T \in \mathrm{U}(n)$. **L**

7.4.13 Beispiel. (a) Die Einheitsmatrix $I_n \in \mathrm{M}_{n,n}(\mathbb{R})$ ist orthogonal, denn $I_n^T I_n = I_n I_n = I_n$.

(b) Die Einheitsmatrix $I_n \in \mathrm{M}_{n,n}(\mathbb{C})$ ist unitär, denn $I_n^* I_n = I_n I_n = I_n$.

(c) Die Matrix

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{R})$$

ist orthogonal, denn es ist

$$A^T A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = I_2.$$

(d) Die Matrix

$$B = \begin{pmatrix} \frac{1+i}{2} & \frac{1+i}{2} \\ \frac{1-i}{2} & \frac{-1+i}{2} \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{C})$$

ist unitär, denn es gilt

$$B^* B = \begin{pmatrix} \frac{1-i}{2} & \frac{1+i}{2} \\ \frac{1-i}{2} & \frac{-1-i}{2} \end{pmatrix} \begin{pmatrix} \frac{1+i}{2} & \frac{1+i}{2} \\ \frac{1-i}{2} & \frac{-1+i}{2} \end{pmatrix} = I_2.$$

7.4.14 Aufgabe. Zeigen Sie: $U(1)$ ist genau der Einheitskreis in $\mathbb{C}$. **L**

7.4.15 Beispiel (Drehmatrizen). Es sei $\alpha \in [0, 2\pi)$. Die Matrix

$$D_\alpha = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{R})$$

nennt man *Drehmatrix*. Die Abbildung $x \mapsto D_\alpha x$ beschreibt eine Drehung im Ursprung um den Winkel α gegen den Uhrzeigersinn.

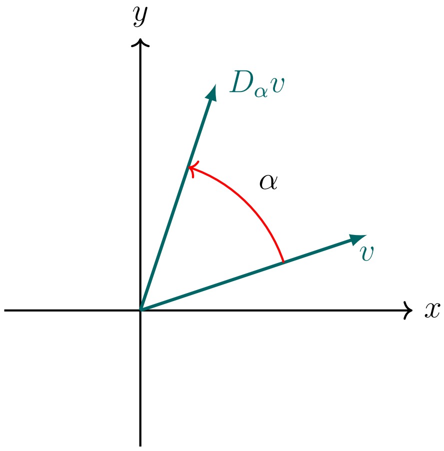


Abbildung 7.1.: Drehung um den Winkel α.

Jede Drehmatrix ist orthogonal. Um das zu sehen, muss man sich erinnern, dass $\cos^2(\alpha) + \sin^2(\alpha) = 1$ gilt. Damit erhält man

$$\begin{aligned} D_\alpha^T D_\alpha &= \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix} \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \\ &= \begin{pmatrix} \cos^2(\alpha) + \sin^2(\alpha) & 0 \\ 0 & \cos^2(\alpha) + \sin^2(\alpha) \end{pmatrix} = I_2. \end{aligned}$$

7.4.16 Beispiel (Spiegelungsmatrizen). Es sei $\alpha \in [0, 2\pi)$. Die Matrix

$$S_\alpha = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{R})$$

nennt man *Spiegelungsmatrix*. Die Abbildung $x \mapsto S_\alpha x$ beschreibt eine Spiegelung an der Geraden $\langle \begin{pmatrix} \cos(\alpha/2) \\ \sin(\alpha/2) \end{pmatrix} \rangle^{\mathbb{R}}$. Um das zu sehen, kann man mit den Formeln

$$\begin{aligned} \cos(\alpha - \beta) &= \cos(\alpha)\cos(\beta) + \sin(\alpha)\sin(\beta) \\ \sin(\alpha - \beta) &= \sin(\alpha)\cos(\beta) - \cos(\alpha)\sin(\beta) \end{aligned}$$

zeigen, dass $\begin{pmatrix} \cos(\alpha/2) \\ \sin(\alpha/2) \end{pmatrix}$ ein Eigenvektor zum Eigenwert 1 ist. Man kann nun nachrechnen, dass das orthogonale Komplement dazu der Eigenraum zum Eigenwert -1 ist.

Jede Spiegelungsmatrix ist orthogonal, denn es gilt

$$\begin{aligned} S_\alpha^T S_\alpha &= \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{pmatrix} \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{pmatrix} \\ &= \begin{pmatrix} \cos^2(\alpha) + \sin^2(\alpha) & 0 \\ 0 & \cos^2(\alpha) + \sin^2(\alpha) \end{pmatrix} = I_2. \end{aligned}$$

7.4.17 **Aufgabe**. Sei $x \neq 0$ in $\mathbb{R}^2$. Bestimmen Sie den Winkel zwischen x und $D_\alpha x$ im kanonischen Euklidischen Raum $(\mathbb{R}^2, \langle \cdot, \cdot \rangle)$ mit der Formel aus 7.2.1. **L**

7.4.18 **Satz.**

- $\mathrm{O}(n)$ *ist eine Untergruppe von* $\mathrm{GL}_n(\mathbb{R})$.
- $\mathrm{U}(n)$ *ist eine Untergruppe von* $\mathrm{GL}_n(\mathbb{C})$.

Man nennt $\mathrm{O}(n)$ die *orthogonale Gruppe* und $\mathrm{U}(n)$ die *unitäre Gruppe.*

Beweis. Wir bemerken zunächst, dass orthogonale und unitäre Matrizen per Definition invertierbar sind, also ist $\mathrm{O}(n)$ eine Teilmenge von $\mathrm{GL}_n(\mathbb{R})$ und $\mathrm{U}(n)$ eine Teilmenge von $\mathrm{GL}_n(\mathbb{C})$. Wir prüfen jetzt, dass die Eigenschaften aus Definition 1.1.18 erfüllt sind.

Die Einheitsmatrix ist orthogonal (bzw. unitär), also sind $\mathrm{O}(n)$ und $\mathrm{U}(n)$ nicht leer. Damit ist 1.1.18 (UG1) erfüllt.

Es seien $A, B \in \mathrm{O}(n)$. Mit der Regel 1.2.32 (ii) erhalten wir

$$(AB)^T(AB) = B^T \underbrace{A^T A}_{=I_n} B = B^T B = I_n.$$

Also ist $AB \in \mathrm{O}(n)$. Also ist $\mathrm{O}(n)$ abgeschlossen unter der Matrizenmultiplikation. Dass auch $\mathrm{U}(n)$ unter der Multiplikation von Matrizen abgeschlossen ist, lassen wir als kleine Übungsaufgabe. **L**

Nun zeigen wir, dass 1.1.18 (UG3) erfüllt ist. Sei dazu $A \in \mathrm{O}(n)$. Dann gilt $A^{-1} = A^T$ und aus Lemma 7.4.10 folgt damit $A^{-1} = A^T \in \mathrm{O}(n)$. Analog gilt für $A \in \mathrm{U}(n)$ auch $A^* = A^{-1} \in \mathrm{U}(n)$ (vgl. 7.4.11). $\square$

Der nächste Satz gibt uns noch eine weitere hilfreiche Beschreibung der orthogonalen und unitären Matrizen.

7.4.19 **Satz.** *Eine Matrix $A \in \mathrm{M}_{n,n}(\mathbb{R})$ ist genau dann orthogonal, wenn die Spalten von A eine Orthonormalbasis des kanonischen Euklidischen Vektorraumes $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$ bilden.*

Eine Matrix $B \in \mathrm{M}_{n,n}(\mathbb{C})$ ist genau dann unitär, wenn die Spalten von B eine Orthonormalbasis des kanonischen unitären Vektorraumes $(\mathbb{C}^n, \langle \cdot, \cdot \rangle)$ bilden.

Beweis. Wir betrachten zunächst den reellen Fall. Es sei $A = (a_{ij})_{i,j} \in \mathrm{M}_{n,n}(\mathbb{R})$. Die i-te Spalte von A bezeichnen wir mit $a^{(i)}$. Berechnet man den Eintrag an Position (i, j) von $A^T A$, so findet man

$$\sum_{k=1}^{n} a_{ki} a_{kj} = \langle a^{(i)}, a^{(j)} \rangle.$$

Der Eintrag an Position (i, j) von $A^T A$ ist also das Skalarprodukt der i-ten und j-ten Spalte. Die Matrix A ist genau dann orthogonal, wenn $A^T A = I_n$ ist, d.h., wenn der Eintrag an Position (i, j) genau δ_{ij} ist. Wie wir aus 7.2.17 wissen, gilt $\langle a^{(i)}, a^{(j)} \rangle = \delta_{ij}$ genau dann, wenn $(a^{(1)}, \ldots, a^{(n)})$ eine Orthonormalbasis ist.

Es sei $B = (b_{ij})_{i,j} \in \mathrm{M}_{n,n}(\mathbb{C})$. Die i-te Spalte von B bezeichnen wir mit $b^{(i)}$. Berechnet man den Eintrag an Position (i, j) von $B^* B$, so findet man

$$\sum_{k=1}^{n} \overline{b_{ki}} b_{kj} = \langle b^{(i)}, b^{(j)} \rangle;$$

dies ist genau das komplexe Standardskalarprodukt der i-ten und der j-ten Spalte. Damit folgt die Behauptung wie im orthogonalen Fall. □

7.4.20 **Bemerkung**. Ist A orthogonal, dann ist auch A^T orthogonal (Lemma 7.4.10) und damit gilt die Aussage von Satz 7.4.19 auch für die Spalten von A^T, d.h., für die transponierten Zeilen von A.

Ist B unitär, dann ist auch B^T unitär (Aufgabe 7.4.12) und die Aussage von Satz 7.4.19 gilt damit auch für die Spalten von B^T, d.h., für die transponierten Zeilen von B.

7.4.21 **Beispiel** (Permutationsmatrizen). Es sei $\sigma \in S_n$ eine Permutation. Wir definieren die Matrix

$$P_\sigma = (p_{ij})_{i,j} \in \mathrm{M}_{n,n}(\mathbb{R})$$

durch $p_{ij} = \delta_{i,\sigma(j)}$. Das heißt, P_σ hat in jeder Zeile und Spalte genau einen Eintrag 1 (nämlich an der Position $(\sigma(j), j)$) und alle anderen Einträge verschwinden. Die lineare Abbildung $x \mapsto P_\sigma x$ bildet also den j-ten Standardbasisvektor e_j auf den Basisvektor $e_{\sigma(j)}$ ab. Die Matrix P_σ nennt man *Permutationsmatrix* zur Permutation σ. Da die Spalten von P_σ genau die Standardbasisvektoren $e_{\sigma(1)}, e_{\sigma(2)}, \dots, e_{\sigma(n)}$ sind, ist P_σ mit Satz 7.4.19 eine orthogonale Matrix. Fasst man P_σ als komplexe Matrix in $\mathrm{M}_{n,n}(\mathbb{C})$ auf, dann ist P_σ auch unitär.

Schauen wir uns ein kleines Beispiel dazu an. Ist $\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$, dann gilt

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

Mit Satz 7.4.19 können wir nun relativ direkt alle orthogonalen 2×2-Matrizen beschreiben.

7.4.22 **Korollar.** *Jede orthogonale 2×2-Matrix ist entweder eine Drehmatrix D_α oder eine Spiegelungsmatrix S_α.*

Beweis. Es sei

$$A = \begin{pmatrix} r & t \\ s & u \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{R}).$$

Aus Satz 7.4.19 folgt, dass A genau denn orthogonal ist, wenn folgende drei Bedingungen erfüllt sind

(1) $r^2 + s^2 = 1$,

(2) $t^2 + u^2 = 1$ und

(3) $rt + su = 0$.

Die letzte Gleichung bedeutet, dass die beiden Spalten orthogonal aufeinander stehen. Wegen $\begin{pmatrix} r \\ s \end{pmatrix} \neq 0$ ist das orthogonale Komplement $U^\perp$ des Unterraumes $U = \langle \begin{pmatrix} r \\ s \end{pmatrix} \rangle^{\mathbb{R}}$ ein-dimensional (siehe 7.2.8). Man kann leicht nachprüfen, dass $U^\perp$ von $\begin{pmatrix} -s \\ r \end{pmatrix}$ aufgespannt wird. Es gilt also

$$\begin{pmatrix} t \\ u \end{pmatrix} = \lambda \begin{pmatrix} -s \\ r \end{pmatrix}$$

für ein $\lambda \in \mathbb{R}$. Aus den beiden Gleichungen $r^2+s^2 = 1 = t^2+u^2$ folgt damit $\lambda = \pm 1$. Weil die erste Spalte auf dem Kreis mit Radius 1 um den Nullpunkt liegt, gibt es einen Winkel $\alpha \in [0, 2\pi)$ mit $r = \cos(\alpha)$ und $s = \sin(\alpha)$. Die Matrix A hat also eine der beiden Formen

$$\begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \quad \text{oder} \quad \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{pmatrix}.$$

Im ersten Fall ist $A = D_\alpha$ eine Drehmatrix (siehe 7.4.15), im zweiten Fall ist $A = S_\alpha$ eine Spiegelungsmatrix (siehe 7.4.16). □

7.4.23 **Aufgabe**. Geben Sie eine orthogonale Matrix A mit erster Spalte $\frac{1}{3}\begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}$ an. **L**

III. Matrixdarstellung von Isometrien

In diesem kurzen Abschnitt werden wir die Matrixdarstellungen von Isometrien studieren. Der nächste Satz liefert uns den Zusammenhang zwischen orthogonalen Abbildungen und orthogonalen Matrizen (bzw. zwischen unitären Abbildungen und unitären Matrizen).

7.4.24 **Satz.** *Sei $(V, \langle \cdot, \cdot \rangle)$ ein endlich-dimensionaler Euklidischer (bzw. unitärer) Vektorraum und es seien $\mathcal{B}$ und $\mathcal{C}$ zwei Orthonormalbasen von V.*

Ein Endomorphismus $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$ ist genau dann eine Isometrie, wenn die Matrixdarstellung ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{C}}(\varphi)$ orthogonal (bzw. unitär) ist.

Beweis. Nehmen wir zuerst an, dass φ eine Isometrie ist. Aus Satz 7.4.6 wissen wir, dass φ invertierbar ist und $\varphi^{\mathrm{ad}} = \varphi^{-1}$ gilt. Mit Satz 7.3.11 sehen wir im reellen Fall, dass

$${}_{\mathcal{B}}\mathrm{M}_{\mathcal{C}}(\varphi)^T {}_{\mathcal{B}}\mathrm{M}_{\mathcal{C}}(\varphi) = {}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\varphi^{\mathrm{ad}}) {}_{\mathcal{B}}\mathrm{M}_{\mathcal{C}}(\varphi) = {}_{\mathcal{C}}\mathrm{M}_{\mathcal{C}}(\varphi^{\mathrm{ad}} \circ \varphi) = {}_{\mathcal{C}}\mathrm{M}_{\mathcal{C}}(\mathrm{id}_V) = I_n$$

gilt. Damit ist die Matrixdarstellung ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{C}}(\varphi)$ orthogonal.

Im unitären Fall zeigt dieselbe Rechnung mit ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{C}}(\varphi)^* = {}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\varphi^{\mathrm{ad}})$, dass ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{C}}(\varphi)$ unitär ist.

Umgekehrt nehmen wir nun an, dass ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{C}}(\varphi)$ orthogonal ist. Da ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{C}}(\varphi)$ invertierbar ist, ist dann auch φ invertierbar (siehe Lemma 2.2.25). Dazu gilt

$${}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\varphi^{\mathrm{ad}}) = {}_{\mathcal{B}}\mathrm{M}_{\mathcal{C}}(\varphi)^T = {}_{\mathcal{B}}\mathrm{M}_{\mathcal{C}}(\varphi)^{-1} = {}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\varphi^{-1}).$$

Im letzten Schritt verwenden wir dabei [MG, 9.3.2]. Da jede Matrixdarstellung die Abbildung eindeutig bestimmt (vgl. 2.2.23), folgt daraus $\varphi^{-1} = \varphi^{\mathrm{ad}}$. Aus Satz 7.4.6 schließen wir, dass φ eine Isometrie ist. Im unitären Fall funktioniert dasselbe Argument, wenn man die Beziehung ${}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\varphi^{\mathrm{ad}}) = {}_{\mathcal{B}}\mathrm{M}_{\mathcal{C}}(\varphi)^*$ verwendet. □

Aus diesem Satz können wir noch eine weitere Charakterisierung von Isometrien ableiten: Isometrien sind jene linearen Abbildungen, die Orthonormalbasen auf Orthonormalbasen abbilden.

7.4.25 **Korollar.** *Sei $(V, \langle\cdot,\cdot\rangle)$ ein endlich-dimensionaler Euklidischer oder unitärer Vektorraum und sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine Orthonormalbasis von V. Ein Endomorphismus φ ist genau dann eine Isometrie, wenn $\varphi(\mathcal{B}) = (\varphi(v_1), \ldots, \varphi(v_n))$ eine Orthonormalbasis von V ist.*

Beweis. Ist φ eine Isometrie, dann ist φ ein Isomorphismus. Also ist $\varphi(\mathcal{B})$ eine Basis von V; siehe [MG, 8.3.18]. Außerdem gilt

$$\langle\varphi(v_i), \varphi(v_j)\rangle = \langle v_i, v_j\rangle = \delta_{ij}$$

und damit handelt es sich um eine Orthonormalbasis.

Nehmen wir umgekehrt an, dass $\mathcal{C} = \varphi(\mathcal{B})$ eine Orthonormalbasis ist. Dann gilt ${}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\varphi) = I_n$. Weil dies eine Matrixdarstellung bezüglich zweier Orthonormalbasen ist und die Einheitsmatrix eine orthogonale bzw. unitäre Matrix ist, folgt aus Satz 7.4.24, dass φ eine Isometrie ist. □

7.5. Spektralsätze

In diesem Abschnitt betrachten wir Endomorphismen von Euklidischen oder unitären Vektorräumen. Das Skalarprodukt liefert uns dabei eine zusätzliche Struktur, die es uns ermöglicht, feinere Fragen als zuvor zu stellen. Insbesondere wollen wir nochmal zur Diagonalisierbarkeit zurückkehren und folgende verschärfte Frage stellen:

> Welche Endomorphismen lassen sich mithilfe einer *Orthonormalbasis* diagonalisieren?

Wir werden sehen, dass wir durch das Skalarprodukt Endomorphismen von unitären Vektorräumen besser verstehen können, als mit den allgemeinen Methoden aus den Lektionen 5 und 6. Das gilt zumindest dann, wenn der Endomorphismus gut zum Skalarprodukt „passt“. Das Resultat ist der sogenannte Spektralsatz für normale Endomorphismen. In Euklidischen Vektorräumen ist dies etwas komplizierter, aber auch hier werden wir einen Spektralsatz für selbstadjungierte Endomorphismen herleiten.

Wie kann man das Skalarprodukt dabei verwenden? Die Sätze zur Diagonalisierbarkeit und Jordan'schen Normalform basieren alle auf einer Zerlegung des Vektorraumes in invariante Unterräume. In Euklidischen und unitären Vektorräumen haben wir ein neues Hilfsmittel: Ist U ein Unterraum, dann ist $U^\perp$ ein Komplement. Wir können also mithilfe des Skalarproduktes immer ein kanonisches Komplement angeben. Diese Komplementbildung ist mit Orthonormalbasen gut verträglich (Aufgabe 7.2.22), mehr noch: Wenn wir uns für Orthonormalbasen interessieren, dann ist das orthogonale Komplement das einzige sinnvolle Komplement.

Das führt uns zur entscheidenden Frage. Angenommen U ist φ-invariant. Ist dann auch $U^\perp$ invariant unter φ? Allgemein ist die Antwort negativ. Aber das nächste Lemma liefert eine einfache, aber wichtige Alternative: $U^\perp$ ist zumindest invariant unter φ^{ad}. Die Zerlegung über orthogonale Komplemente kann also funktionieren, wenn es eine Beziehung zwischen φ und φ^{ad} gibt. Das wollen wir in diesem Abschnitt ausarbeiten.

7.5.1 **Lemma.** *Sei $(V, \langle\cdot,\cdot\rangle)$ ein endlich-dimensionaler Euklidischer oder unitärer Vektorraum und sei $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$. Ist $U \subseteq V$ ein φ-invarianter Unterraum, dann ist $U^\perp$ ein φ^{ad}-invarianter Unterraum.*

Beweis. Sei $v \in U^\perp$ und sei $u \in U$. Dann gilt

$$\langle u, \varphi^{\mathrm{ad}}(v)\rangle = \langle \varphi(u), v\rangle = 0,$$

denn nach Voraussetzung ist U invariant unter φ und damit $\varphi(u) \in U$. Da $u \in U$ beliebig war, liegt $\varphi^{\mathrm{ad}}(v)$ im orthogonalen Komplement $U^\perp$. □

7.5.2 **Korollar.** *Sei $(V, \langle\cdot,\cdot\rangle)$ ein endlich-dimensionaler Euklidischer oder unitärer Vektorraum und sei $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$. Ist $U \subseteq V$ ein Unterraum der unter φ und φ^{ad} invariant ist, dann hat auch $U^\perp$ diese Eigenschaft. Es gilt außerdem $(\varphi|_U)^{ad} = (\varphi^{ad})|_U$.*

Beweis. Dass $U^\perp$ invariant unter φ und φ^{ad} ist folgt direkt aus Lemma 7.5.1 und der Beziehung $(\varphi^{\mathrm{ad}})^{\mathrm{ad}} = \varphi$ (siehe 7.3.2 (c)).

Da die Identität $\langle\varphi(u), v\rangle = \langle u, \varphi^{\mathrm{ad}}(v)\rangle$ insbesondere für alle Vektoren $u, v \in U$ erfüllt ist, gilt $(\varphi|_U)^{\mathrm{ad}} = (\varphi^{\mathrm{ad}})|_U$. □

I. Normale Endomorphismen

7.5.3 **Definition** Sei $(V, \langle\cdot,\cdot\rangle)$ ein endlich-dimensionaler Euklidischer oder unitärer Vektorraum. Ein Endomorphismus $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$ heißt *normal*, wenn er mit seinem Adjungierten vertauscht, d.h., wenn

$$\varphi \circ \varphi^{\mathrm{ad}} = \varphi^{\mathrm{ad}} \circ \varphi$$

gilt.

7.5.4 **Beispiel**. Sei $(V, \langle\cdot,\cdot\rangle)$ ein endlich-dimensionaler Euklidischer oder unitärer Vektorraum und sei $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$.

(a) Ist $\varphi\colon V \to V$ eine Isometrie, dann ist φ normal. Um das zu sehen kann man die Beziehung $\varphi^{\mathrm{ad}} = \varphi^{-1}$ aus Satz 7.4.6 verwenden:

$$\varphi \circ \varphi^{\mathrm{ad}} = \varphi \circ \varphi^{-1} = \mathrm{id}_V = \varphi^{-1} \circ \varphi = \varphi^{\mathrm{ad}} \circ \varphi.$$

(b) Einen Endomorphismus φ mit $\varphi^{\mathrm{ad}} = \varphi$ nennt man *selbstadjungiert*. Selbstadjungierte Endomorphismen sind normal, denn es ist

$$\varphi \circ \varphi^{\mathrm{ad}} = \varphi \circ \varphi = \varphi^{\mathrm{ad}} \circ \varphi.$$

7.5.5 **Aufgabe**. Einen Endomorphismus φ mit $\varphi^{\mathrm{ad}} = -\varphi$ nennt man *schiefsymmetrisch*. **L**
Zeigen Sie: Schiefsymmetrische Endomorphismen sind normal.

Das nächste Lemma enthält die entscheidende Beobachtung um den Spektralsatz für normale Endomorphismen zu beweisen.

7.5.6 **Lemma.** *Sei $(V, \langle \cdot, \cdot \rangle)$ ein endlich-dimensionaler Euklidischer oder unitärer Vektorraum und sei φ ein normaler Endomorphismus.*

Ist $\lambda \in \mathbb{K}$ ein Eigenwert von φ, dann gilt

$$\operatorname{Eig}_{\varphi}(\lambda) = \operatorname{Eig}_{\varphi^{ad}}(\overline{\lambda}).$$

Insbesondere ist $\operatorname{Eig}_{\varphi}(\lambda)$ ein φ^{ad}-invarianter Unterraum.

Beweis. Wir zeigen zuerst, dass $U := \operatorname{Eig}_{\varphi}(\lambda)$ ein φ^{ad}-invarianter Unterraum ist. Es sei dazu $v \in \operatorname{Eig}_{\varphi}(\lambda)$. Dann gilt

$$\varphi(\varphi^{\mathrm{ad}}(v)) = \varphi^{\mathrm{ad}}(\varphi(v)) = \varphi^{\mathrm{ad}}(\lambda v) = \lambda \varphi^{\mathrm{ad}}(v).$$

Damit ist $\varphi^{\mathrm{ad}}(v) \in \operatorname{Eig}_{\varphi}(\lambda)$. Da U damit φ^{ad}-invariant und auch φ-invariant ist (siehe 5.2.9), folgt

$$\begin{aligned} \varphi^{\mathrm{ad}}|_U &= (\varphi|_U)^{\mathrm{ad}} && (7.5.2) \\ &= (\lambda \operatorname{id}_U)^{\mathrm{ad}} && (U = \operatorname{Eig}_{\varphi}(\lambda)) \\ &= \overline{\lambda} \operatorname{id}_U . && (7.3.5) \end{aligned}$$

Für alle $u \in U$ gilt also $\varphi^{\mathrm{ad}}(u) = \overline{\lambda} u$, d.h., es ist $\operatorname{Eig}_{\varphi}(\lambda) \subseteq \operatorname{Eig}_{\varphi^{\mathrm{ad}}}(\overline{\lambda})$. Dasselbe Argument mit vertauschten Rollen und die Beziehung $(\varphi^{\mathrm{ad}})^{\mathrm{ad}} = \varphi$ liefern uns die umgekehrte Inklusion. □

7.5.7 **Aufgabe**. Sei $(V, \langle \cdot, \cdot \rangle)$ ein endlich-dimensionaler Euklidischer oder unitärer Vektorraum und sei φ ein normaler Endomorphismus. **L**

Zeigen Sie: Sind $u, v \in V$ Eigenvektoren von φ zu verschiedenen Eigenwerten, dann sind u und v orthogonal.

Damit haben wir alle Hilfsmittel zusammen um den folgenden, sehr nützlichen Satz beweisen zu können. Wir betrachten zuerst nur den Fall unitärer Vektorräume.

7.5.8 **Spektralsatz für normale Endomorphismen (komplex).** *Sei $(V, \langle \cdot, \cdot \rangle)$ ein endlich-dimensionaler* unitärer *Vektorraum. Ist $\varphi \in \operatorname{End}_{\mathbb{C}}(V)$ normal, dann gibt es eine Orthonormalbasis $\mathcal{B}$ von V, die aus Eigenvektoren von φ besteht.*

Insbesondere ist φ diagonalisierbar.

Beweis. Wir zeigen den Satz durch Induktion über die Dimension $n = \dim_{\mathbb{C}} V$ des unitären Vektorraumes. Für $n = 1$ nehmen wir einen normierten Vektor $v_1 \in V$. Es gilt $\varphi(v_1) = \lambda v_1$ für geeignetes λ (vgl. 4.4.4), also ist $\mathcal{B} = (v_1)$ die gesuchte Orthonormalbasis.

Im Induktionsschritt sei $n \geq 2$ und wir nehmen an, dass die Aussage für alle unitären Räume kleinerer Dimension bekannt ist. Da wir über den komplexen Zahlen arbeiten, besitzt das charakteristische Polynom von φ mindestens eine Nullstelle $\lambda \in \mathbb{C}$ (siehe 1.5.25). Das heißt, λ ist ein Eigenwert von φ und der Eigenraum $U = \mathrm{Eig}_{\varphi}(\lambda)$ ist nicht-trivial, d.h., $\dim_{\mathbb{C}} U = k > 0$. Wir wählen eine Orthonormalbasis $(v_1, \ldots, v_k)$ von U (siehe 7.2.21).

Nach Lemma 7.5.6 ist U sowohl φ- als auch φ^{ad}-invariant. Mit Korollar 7.5.2, schließen wir, dass dann auch das orthogonale Komplement φ-invariant ist. Aus dem Satz vom orthogonalen Komplement 7.2.8 erhalten wir

$$\dim_{\mathbb{C}}(U^{\perp}) = n - \dim_{\mathbb{C}} U = n - k < n.$$

Wegen $\varphi^{\mathrm{ad}}|_{U^{\perp}} = (\varphi|_{U^{\perp}})^{\mathrm{ad}}$ ist auch $\varphi|_{U^{\perp}}$ normal. Aus der Induktionsvoraussetzung erhalten wir eine Orthonormalbasis $(v_{k+1}, \ldots, v_n)$ von $U^{\perp}$, die aus Eigenvektoren von $\varphi|_{U^{\perp}}$ (also auch von φ) besteht. Dann ist $(v_1, \ldots, v_k, v_{k+1}, \ldots, v_n)$ eine Orthonormalbasis von V, die aus Eigenvektoren von φ besteht (dass dies eine Orthonormalbasis ist wurde in Aufgabe 7.2.22 nachgeprüft). □

Der Beweis des Spektralsatzes verwendet nur an einer Stelle, dass $\mathbb{K} = \mathbb{C}$ ist: Der Fundamentalsatz der Algebra sichert uns die Existenz eines Eigenwertes von φ! Deshalb kann man den Spektralsatz nicht direkt auf Euklidische Vektorräume übertragen.

7.5.9 **Aufgabe**. Zeigen Sie mit einem Beispiel, dass der Spektralsatz in dieser Formulierung für Euklidische Vektorräume nicht stimmt. **L**

Man kann aber ganz analog einen Spektralsatz für normale Endomorphismen Euklidischer Vektorräume beweisen, wenn man die Voraussetzung hinzunimmt, dass das charakteristische Polynom in $\mathbb{R}[X]$ in Linearfaktoren zerfällt. Allerdings ist dieser Satz weniger wichtig als im Komplexen, da auf Euklidischen Vektorräumen die normalen, diagonalisierbaren Endomorphismen genau die selbstadjungierten Endomorphismen sind und diese werden wir uns gleich noch genauer ansehen.

7.5.10 **Spektralsatz für normale Endomorphismen (reell).** *Sei $(V, \langle\cdot,\cdot\rangle)$ ein Euklidischer Vektorraum von endlicher Dimension. Sei $\varphi \in \mathrm{End}_{\mathbb{R}}(V)$ normal. Falls* *das charakteristische Polynom $\chi_\varphi \in \mathbb{R}[X]$ in Linearfaktoren zerfällt, dann gibt es eine Orthonormalbasis $\mathcal{B}$ von V, die aus Eigenvektoren von φ besteht.*

Insbesondere ist φ diagonalisierbar.

Der Beweis von Satz 7.5.8 lässt sich dabei fast ohne Änderung übertragen. Da $\chi_\varphi \in \mathbb{R}[X]$ in Linearfaktoren zerfällt, finden wir wieder einen Eigenwert $\lambda \in \mathbb{R}$. Ist $U = \mathrm{Eig}_\varphi(\lambda)$, dann haben wir wieder die Zerlegung $V = U \oplus U^\perp$ als Summe invarianter Unterräume. Man muss sich dann klarmachen, dass $\varphi|_{U^\perp}$ wieder die Voraussetzungen erfüllt, d.h., dass das charakteristische Polynom in Linearfaktoren zerfällt. Da $U \oplus U^\perp$ eine Zerlegung in φ-invariante Unterräume ist, ist $\chi_{\varphi|_{U^\perp}}$ ein Teiler von χ_φ (siehe Satz 5.3.19) und zerfällt nach Korollar 1.4.39 in Linearfaktoren.

7.5.11 **Bemerkung**. Es ist nicht schwer eine Orthonormalbasis aus Eigenvektoren für normale Endomorphismen zu finden. Aus Aufgabe 7.5.7 ist bekannt, dass die Eigenräume zu verschiedenen Eigenwerten orthogonal aufeinander stehen. Das heißt, es ist ausreichend eine Orthonormalbasis jedes Eigenraumes zu wählen. Fügt man diese Orthonormalbasen zusammen, dann erhält man eine Orthonormalbasis aus Eigenvektoren; vgl. Aufgabe 7.2.22.

II. Unitäre Endomorphismen

Da unitäre Endomorphismen normal sind (siehe 7.5.4), können wir aus dem Spektralsatz für normale Endomorphismen insbesondere einen Spektralsatz für unitäre Abbildungen ableiten. Um diesen Satz zu formulieren erinnern wir uns, dass eine komplexe Zahl ω auf dem Einheitskreis liegt, wenn $|\omega| = 1$ gilt; siehe 1.5.13.

7.5.12 **Spektralsatz für unitäre Abbildungen.** *Sei $(V, \langle\cdot,\cdot\rangle)$ ein unitärer endlich-dimensionaler Vektorraum und sei $\varphi\colon V \to V$ eine unitäre Abbildung. Dann gibt es eine Orthonormalbasis $\mathcal{B}$ von V, die aus Eigenvektoren von φ besteht. Alle Eigenwerte von φ liegen auf dem Einheitskreis.*

Beweis. Da alle unitären Abbildungen normal sind (siehe 7.5.4), erhalten wir die erste Aussage direkt aus dem Spektralsatz 7.5.8. Wir müssen nur noch zeigen, dass die Eigenwerte auf dem Einheitskreis liegen. Sei λ ein Eigenwert von φ und sei $v \in V$ ein Eigenvektor zum Eigenwert λ. Dann gilt

$$\begin{aligned} \langle v, v\rangle &= \langle \varphi(v), \varphi(v)\rangle && (\varphi \text{ unitär}) \\ &= \langle \lambda v, \lambda v\rangle = \lambda\overline{\lambda}\langle v, v\rangle = |\lambda|^2\langle v, v\rangle. \end{aligned}$$

Da $v \neq 0$ ist und das Skalarprodukt positiv definit ist, gilt $\langle v, v\rangle > 0$. Durch Kürzen erhalten wir $|\lambda|^2 = 1$, d.h., λ liegt auf dem Einheitskreis. □

Aus diesem Satz kann man einen Spektralsatz für unitäre Matrizen ableiten.

7.5.13 Spektralsatz für unitäre Matrizen. *Es sei $B \in \mathrm{U}(n)$ eine unitäre Matrix. Dann gibt es eine unitäre Matrix $S \in \mathrm{U}(n)$, sodass*

$$S^{-1}BS = S^*BS = \operatorname{diag}(\omega_1, \ldots, \omega_n)$$

eine Diagonalmatrix ist. Die Einträge ω_i liegen alle auf dem Einheitskreis.

Hinweis: Es ist $S^* = S^{-1}$, weil S unitär ist.

Beweis. Wir betrachten den kanonischen unitären Raum $(\mathbb{C}^n, \langle\cdot,\cdot\rangle)$ mit dem Endomorphismus $\varphi\colon v \mapsto Bv$. Ist $\mathcal{C}$ die Standardbasis von $\mathbb{C}^n$, dann ist $B = {}_{\mathcal{C}}\mathrm{M}_{\mathcal{C}}(\varphi)$. Da B unitär ist, ist dann auch φ eine unitäre Abbildung nach Satz 7.4.24.

Nach Satz 7.5.12 gibt es also eine Orthonormalbasis $\mathcal{B}$ von $\mathbb{C}^n$, sodass $D = {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)$ eine Diagonalmatrix ist. Mit der Transformationsformel für Endomorphismen schließen wir

$$D = {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi) = {}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\mathrm{id})^{-1}\; B\; {}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\mathrm{id}).$$

Die Matrix $S = {}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\mathrm{id})$ ist dabei unitär, weil $\mathcal{B}$ und $\mathcal{C}$ Orthonormalbasen sind und die identische Abbildung $\mathrm{id}\colon \mathbb{C}^n \to \mathbb{C}^n$ eine Isometrie ist (siehe Satz 7.4.24). Es gilt also $S^* = S^{-1}$ und damit ist der Satz bewiesen. □

Aus Bemerkung 7.5.11 kann man ein Verfahren ableiten, um eine Matrix S wie in Satz 7.5.13 zu berechnen.

7.5.14 Diagonalisieren unitärer Matrizen mit Orthonormalbasen.

`Gegeben:` $B \in \mathrm{U}(n)$.

`Gesucht:` $S \in \mathrm{U}(n)$, sodass $S^{-1}BS = S^*BS$ eine Diagonalmatrix ist.

`Verfahren:`

- Berechne die Eigenwerte $\lambda_1, \ldots, \lambda_r$ von B.
- Bestimme eine Orthonormalbasis $\mathcal{B}_i$ von $\mathrm{Eig}_B(\lambda_i)$ (z.B. mit Gram-Schmidt).
- Die gesuchte Matrix S erhält man, indem man die Vektoren aus den Basen $\mathcal{B}_1, \ldots, \mathcal{B}_r$ als Spalten in die Matrix S schreibt.

Die Reihenfolge in der die Vektoren in die Matrix S geschrieben werden, bestimmt die Reihenfolge der Diagonaleinträge. Ist die k-te Spalte ein Vektor aus $\mathcal{B}_i$, dann ist λ_i der k-te Eintrag auf der Diagonalen.

7.5.15 Beispiel. Wir betrachten die Permutationsmatrix

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{C}).$$

Aus Beispiel 7.4.21 ist bekannt, dass Permutationsmatrizen unitär sind. Wir wollen P mithilfe einer unitären Matrix S diagonalisieren.

Das charakteristische Polynom von P berechnet sich mit der Regel von Sarrus als

$$\chi_P = \det \begin{pmatrix} X & -1 & 0 \\ 0 & X & -1 \\ -1 & 0 & X \end{pmatrix} = X^3 - 1.$$

Die Nullstellen von χ_P sind $\lambda_1 = 1$, $\lambda_2 = \omega$ und $\lambda_3 = \overline{\omega}$ mit

$$\omega = -\frac{1}{2} + \frac{\sqrt{3}}{2} i.$$

Die Zahl ω ist eine dritte Einheitswurzel und es gilt $\omega^2 = \overline{\omega}$. Mit einer kurzen Rechnung findet man die Eigenräume

$$\mathrm{Eig}_P(1) = \langle \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \rangle,$$

$$\mathrm{Eig}_P(\omega) = \langle \begin{pmatrix} 1 \\ \omega \\ \omega^2 \end{pmatrix} \rangle,$$

$$\mathrm{Eig}_P(\overline{\omega}) = \langle \begin{pmatrix} 1 \\ \overline{\omega} \\ \overline{\omega}^2 \end{pmatrix} \rangle.$$

Die Eigenräume sind ein-dimensional. Man kann sich leicht davon überzeugen, dass die Eigenräume orthogonal aufeinander stehen. Orthonormalbasen der Eigenräume erhalten wir hier ganz einfach, indem wir die gefundenen Eigenvektoren normieren:

$$v_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad v_2 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ \omega \\ \omega^2 \end{pmatrix}, \quad v_3 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ \overline{\omega} \\ \overline{\omega}^2 \end{pmatrix}.$$

Wir schreiben diese Vektoren als Spalten in eine Matrix und haben eine unitäre Matrix

$$S = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 1 & 1 \\ 1 & \omega & \overline{\omega} \\ 1 & \omega^2 & \overline{\omega}^2 \end{pmatrix}$$

mit $S^* P S = S^{-1} P S = \mathrm{diag}(1, \omega, \overline{\omega})$ gefunden.

III. Selbstadjungierte Endomorphismen

7.5.16 **Definition** Sei $(V, \langle\cdot,\cdot\rangle)$ ein endlich-dimensionaler Euklidischer oder unitärer Vektorraum. Ein Endomorphismus $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$ ist *selbstadjungiert*, wenn

$$\varphi^{\mathrm{ad}} = \varphi$$

gilt.

7.5.17 **Anmerkung**. Ein Endomorphismus $\varphi\colon V \to V$ ist also genau dann selbstadjungiert, wenn

$$\langle v, \varphi(w)\rangle = \langle \varphi(v), w\rangle \tag{7.5.a}$$

für alle $v, w \in V$ gilt. Diese Gleichung kann man zur Definition von Selbstadjungiertheit heranziehen.

Selbstadjungierte Endomorphismen („Operatoren") spielen auch in der Theorie der unendlich-dimensionalen Euklidischen und unitären Räume eine sehr wichtige Rolle. Hier betrachten wir aber nur endlich-dimensionale Vektorräume. Ob ein Endomorphismus selbstadjungiert ist, kann man dann anhand einer Matrixdarstellung bzgl. einer Orthonormalbasis ablesen.

7.5.18 **Lemma.** *Sei $(V, \langle\cdot,\cdot\rangle)$ ein endlich-dimensionaler Euklidischer oder unitärer Vektorraum. Es seien $\mathcal{B}$ eine Orthonormalbasis und $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$.*

(a) Ist V Euklidisch, dann gilt:

$$\varphi \text{ selbstadjungiert} \quad\Leftrightarrow\quad {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi) \text{ symmetrisch.}$$

(b) Ist V unitär, dann gilt:

$$\varphi \text{ selbstadjungiert} \quad\Leftrightarrow\quad {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi) \text{ Hermite'sch.}$$

Beweis. Angenommen V ist Euklidisch. Aus Satz 7.3.11 folgt

$${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi^{\mathrm{ad}}) = {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)^T.$$

Da die Matrixdarstellung den Endomorphismus eindeutig bestimmt (vgl. 2.2.23), gilt $\varphi = \varphi^{\mathrm{ad}}$ genau dann, wenn

$${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi) = {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi^{\mathrm{ad}}) = {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)^T$$

ist. Damit folgt die Behauptung.

Den unitären Fall kann man analog behandeln (Übungsaufgabe!). □ **L**

7.5.19 **Aufgabe**. Sei $(V, \langle\cdot,\cdot\rangle)$ ein endlich-dimensionaler Euklidischer oder unitärer Vektorraum und sei $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$. Zeigen Sie: $\varphi^{\mathrm{ad}} \circ \varphi$ ist selbstadjungiert. **L**

7.5.20 **Beispiel**. Die Matrix

$$A = \begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{R})$$

ist symmetrisch. Wir betrachten den Endomorphismus $f_A\colon x \mapsto Ax$ des kanonischen Euklidischen Raumes $(\mathbb{R}^2, \langle\cdot,\cdot\rangle)$. Die Standardbasis $\mathcal{B}$ ist eine Orthonormalbasis und $A = {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(f_A)$ (siehe 2.2.24). Also ist f_A selbstadjungiert.

Aus Beispiel 7.5.4 wissen wir, dass selbstadjungierte Endomorphismen normal sind. Daraus werden wir nun den folgenden Spektralsatz ableiten.

7.5.21 **Spektralsatz für selbstadjungierte Endomorphismen.** *Sei $(V, \langle\cdot,\cdot\rangle)$ ein endlich-dimensionaler Euklidischer oder unitärer Vektorraum und sei $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$ selbstadjungiert. Dann gibt es eine Orthonormalbasis, die aus Eigenvektoren von φ besteht. Alle Eigenwerte von φ sind reell.*

Wir verschieben den Beweis etwas nach hinten und formulieren zuerst Versionen dieses Satzes für Hermite'sche und symmetrische Matrizen.

7.5.22 **Spektralsatz für Hermite'sche Matrizen.** *Sei $B \in \mathrm{M}_{n,n}(\mathbb{C})$ eine Hermite'sche Matrix. Dann gibt es eine unitäre Matrix $S \in \mathrm{U}(n)$, sodass $S^{-1}BS = S^*BS$ eine Diagonalmatrix mit reellen Einträgen ist.*

Beweis. Wir betrachten den kanonischen unitären Raum $(\mathbb{C}^n, \langle\cdot,\cdot\rangle)$ und die lineare Abbildung $f_B\colon \mathbb{C}^n \to \mathbb{C}^n$ mit $x \mapsto Bx$. Die Standardbasis $\mathcal{C}$ ist eine Orthonormalbasis und die Matrixdarstellung $B = {}_{\mathcal{C}}\mathrm{M}_{\mathcal{C}}(f_B)$ ist nach Voraussetzung Hermite'sch. Also folgt aus Lemma 7.5.18, dass f_B selbstadjungiert ist.

Der Spektralsatz für selbstadjungierte Endomorphismen liefert uns eine Orthonormalbasis $\mathcal{B}$ von $\mathbb{C}^n$, die aus Eigenvektoren von f_B besteht. Die Transformationsformel für Endomorphismen 4.4.5 impliziert

$$D = {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(f_B) = {}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\mathrm{id})^{-1}\, B\, {}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\mathrm{id}),$$

wobei D eine reelle Diagonalmatrix ist (die Diagonaleinträge sind die Eigenwerte von B). Die Matrix $S = {}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\mathrm{id})$ ist dabei unitär, weil $\mathcal{B}$ und $\mathcal{C}$ Orthonormalbasen sind und die identische Abbildung $\mathrm{id}\colon \mathbb{C}^n \to \mathbb{C}^n$ eine Isometrie ist (siehe Satz 7.4.24). Es gilt also $S^* = S^{-1}$ und damit ist der Satz bewiesen. □

7.5.23 **Spektralsatz für reelle symmetrische Matrizen.** *Sei $A \in \mathrm{M}_{n,n}(\mathbb{R})$ eine symmetrische Matrix. Dann gibt es eine orthogonale Matrix $S \in \mathrm{O}(n)$, sodass $S^{-1}AS = S^TAS$ eine reelle Diagonalmatrix ist.*

Der Beweis funktioniert genau wie der Beweis von Satz 7.5.22. Übungsaufgabe! **L**

7.5.24 **Bedeutung der Spektralsätze: Simultane Diagonalisierung**. Die Spektralsätze für Hermite'sche und symmetrische Matrizen bilden einen bemerkenswerten Schlusspunkt für dieses Modul. Um das zu verstehen, müssen wir noch einmal zurückblicken.

Im Abschnitt über Diagonalisierbarkeit haben wir uns die Frage gestellt, welche Endomorphismen sich diagonalisieren lassen. Über die Matrixdarstellungen konnten wir das Problem in die Sprache der Matrizen übersetzen: Gegeben eine Matrix $A \in \mathrm{M}_{n,n}(K)$, suchen wir eine invertierbare Matrix $S \in \mathrm{GL}_n(K)$, sodass

$$S^{-1}AS$$

eine Diagonalmatrix ist. Auf der linken Seite steht dabei S^{-1} – eine Folge der Transformationsformel für Endomorphismen. Dieses Problem konnten wir über beliebigen Körpern K beantworten. Die Matrix A war dabei die Matrixdarstellung eines Endomorphismus.

Im Abschnitt über symmetrische Bilinearformen 3.2 haben wir den Diagonalisierungssatz 3.2.10 bewiesen. Er besagt, dass jede symmetrische Bilinearform eine diagonale Matrixdarstellung besitzt (vorausgesetzt es gilt $2 \neq 0$ in K). Mithilfe der Matrixdarstellung von Bilinearformen kann man das in der Sprache der Matrizen ausdrücken: Ist $A \in \mathrm{M}_{n,n}(K)$ eine symmetrische Matrix, dann gibt es eine invertierbare Matrix $S \in \mathrm{GL}_n(K)$, sodass

$$S^TAS$$

eine Diagonalmatrix ist. Auf der linken Seite steht dabei S^T – eine Konsequenz der Transformationsformel 3.1.17 für Bilinearformen. Die Matrix A ist dabei die Matrixdarstellung einer symmetrischen Bilinearform.

Haben wir nun eine symmetrische Matrix $A \in \mathrm{M}_{n,n}(K)$ gegeben, dann können wir diese auf zwei verschiedene Weisen auffassen.

(1) als Matrixdarstellung des Endomorphismus $f_A \colon K^n \to K^n$ oder

(2) als Matrixdarstellung der symmetrischen Bilinearform $\beta(x, y) = x^TAy$.

Können wir beide Probleme simultan lösen? Können wir also den Endomorphismus und die Bilinearform gleichzeitig diagonalisieren? Dazu benötigen wir eine invertierbare Matrix $S \in \mathrm{GL}_n(K)$ mit $S^T = S^{-1}$, sodass

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

eine Diagonalmatrix ist. Der Spektralsatz für reelle symmetrische Matrizen besagt, dass das über $K = \mathbb{R}$ möglich ist! Die Matrizen mit der Eigenschaft $S^T = S^{-1}$ sind genau die orthogonalen Matrizen.

Auch den Spektralsatz für Hermite'sche Matrizen kann man so verstehen. Eine Hermite'sche Matrix B können wir sowohl als Matrixdarstellung eines Endomorphismus verstehen, als auch als Matrixdarstellung einer Hermite'schen Form auffassen. Der Spektralsatz besagt, dass man den Endomorphismus und die Hermite'sche Form gleichzeitig in Diagonalform bringen kann:

$$S^{-1}BS = S^*BS$$

ist diagonal. Der Ausdruck S^* auf der linken Seite kommt dabei aus der Transformationsformel für Hermite'sche Formen 3.4.25.

7.5.25 Beweis des Spektralsatzes für selbstadjungierte Endomorphismen 7.5.21.

Beweis. Wir betrachten zuerst den unitären Fall. Die erste Aussage folgt hier direkt aus dem Spektralsatz für normale Endomorphismen 7.5.8. Es bleibt zu prüfen, dass sämtliche Eigenwerte reell sind. Sei dazu λ ein Eigenwert von φ und es sei v ein zugehöriger Eigenvektor. Dann ist v ein Eigenvektor von φ^{ad} zum Eigenwert $\overline{\lambda}$; das folgt aus Lemma 7.5.6. Da $\varphi = \varphi^{\mathrm{ad}}$ ist, erhalten wir

$$\lambda v = \varphi(v) = \varphi^{\mathrm{ad}}(v) = \overline{\lambda} v$$

und damit $\lambda = \overline{\lambda}$ (wegen $v \neq 0$). Das heißt, $\lambda \in \mathbb{R}$.

Nehmen wir nun an, dass $(V, \langle\cdot,\cdot\rangle)$ Euklidisch ist. Um den Spektralsatz für normale Endomorphismen 7.5.10 anzuwenden, müssen wir nachweisen, dass $\chi_\varphi \in \mathbb{R}[X]$ in Linearfaktoren zerfällt. Dazu nehmen wir eine Orthonormalbasis $\mathcal{B}$ von V. Dann ist $A = {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi) \in \mathrm{M}_{n,n}(\mathbb{R})$ eine reelle symmetrische Matrix (Lemma 7.5.18). Wir wissen, dass $\chi_\varphi = \chi_A$ ist; siehe 5.3.17. Wir wollen zeigen, dass χ_A in Linearfaktoren zerfällt.

Die reellen Zahlen liegen in den komplexen Zahlen. Wir können A also auch als Element von $\mathrm{M}_{n,n}(\mathbb{C})$ auffassen. Dann gilt $A^* = \overline{A}^T = A^T = A$, d.h., A ist eine Hermite'sche Matrix. Aus dem Spektralsatz für Hermite'sche Matrizen wissen wir aber bereits, dass A (als komplexe Matrix) diagonalisierbar ist und nur reelle Eigenwerte $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ besitzt (die Eigenwerte dürfen hier mehrfach vorkommen). Insbesondere gilt $\chi_A = (X - \lambda_1)(X - \lambda_2) \cdots (X - \lambda_n)$, d.h., χ_A zerfällt in $\mathbb{R}[X]$ in Linearfaktoren. $\square$

7.5.26 **Verfahren zur simultanen Diagonalisierung**. Um eine unitäre (bzw. orthogonale) Matrix S wie im Spektralsatz 7.5.22 (bzw. 7.5.23) zu bestimmen, geht man genau wie im Verfahren 7.5.14 vor. Man berechnet die Eigenwerte $\lambda_1, \dots, \lambda_r \in \mathbb{R}$ und die zugehörigen Eigenräume. Dann bestimmt man eine Orthonormalbasis von jedem Eigenraum. Die gesuchte Matrix S erhält man, indem man die Vektoren aus den Orthonormalbasen als Spalten in die Matrix S schreibt. Die Reihenfolge in der die Vektoren in die Matrix S geschrieben werden, bestimmt die Reihenfolge der Diagonaleinträge.

7.5.27 **Beispiel**. Wir betrachten die reelle symmetrische Matrix

$$A = \begin{pmatrix} -1 & -2 & 4 \\ -2 & 2 & 2 \\ 4 & 2 & -1 \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{R}).$$

Wir möchten eine orthogonale Matrix $S \in \mathrm{O}(3)$ bestimmen, sodass $S^T AS = S^{-1}AS$ eine Diagonalmatrix ist.

Wir berechnen zuerst das charakteristische Polynom mit der Regel von Sarrus 4.2.6:

$$\begin{aligned} \chi_A &= \det \begin{pmatrix} X+1 & 2 & -4 \\ 2 & X-2 & -2 \\ -4 & -2 & X+1 \end{pmatrix} \\ &= (X+1)^2(X-2) + 16 + 16 - 16(X-2) - 4(X+1) - 4(X+1) \\ &= X^3 - 27X + 54. \end{aligned}$$

Durch Ausprobieren finden wir die Nullstelle $\lambda_1 = 3$ von χ_A. Mit diesem Wissen kann man das charakteristische Polynom zerlegen und findet

$$\chi_A = (X-3)^2(X+6).$$

Die Eigenwerte sind also $\lambda_1 = 3$ und $\lambda_2 = -6$.

Wir bestimmen die Eigenräume und eine Orthonormalbasis von jedem Eigenraum.

$$\mathrm{Eig}_A(3) = \mathrm{Ker}(A - 3I_3) = \mathrm{Ker} \begin{pmatrix} -4 & -2 & 4 \\ -2 & -1 & 2 \\ 4 & 2 & -4 \end{pmatrix}$$

$$\begin{aligned} &= \operatorname{Ker}\begin{pmatrix} 2 & 1 & -2 \end{pmatrix} \\ &= \langle \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \rangle^{\mathbb{R}} \end{aligned}$$

Damit haben wir eine Basis von $\operatorname{Eig}_A(3)$, aber noch keine Orthogonalbasis. Wir verwenden das Verfahren nach Gram-Schmidt mit $u_1 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$ und $u_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$. Es gilt $\|u_1\| = \sqrt{1+4+4} = 3$. Wir normieren u_1 und erhalten $v_1 = \frac{1}{3}u_1 = \frac{1}{3}\begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$. Dann berechnen wir v_2' mit der Formel

$$v_2' = u_2 - \langle v_1, u_2 \rangle v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} - 1 \cdot v_1 = \frac{1}{3}\begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}.$$

Der Vektor v_2' ist bereits normiert, also setzen wir $v_2 = v_2'$. Dann ist (v_1, v_2) eine Orthonormalbasis von $\operatorname{Eig}_A(3)$.

Wir betrachten nun $\lambda_2 = -6$.

$$\begin{aligned} \operatorname{Eig}_A(-6) = \operatorname{Ker}(A + 6I_3) &= \operatorname{Ker}\begin{pmatrix} 5 & -2 & 4 \\ -2 & 8 & 2 \\ 4 & 2 & 5 \end{pmatrix} \\ &= \operatorname{Ker}\begin{pmatrix} -1 & 4 & 1 \\ 0 & 18 & 9 \\ 0 & 18 & 9 \end{pmatrix} \\ &= \operatorname{Ker}\begin{pmatrix} -1 & 4 & 1 \\ 0 & 2 & 1 \end{pmatrix} \\ &= \langle \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix} \rangle^{\mathbb{R}} \end{aligned}$$

Der Eigenraum ist ein-dimensional. Wir wählen den normierten Vektor $v_3 = \frac{1}{3}\begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}$ und haben damit eine Orthonormalbasis von $\operatorname{Eig}_A(6)$ gefunden. Die Basisvektoren v_1, v_2, v_3 bilden die Spalten der Matrix S:

$$S = \frac{1}{3}\begin{pmatrix} 1 & 2 & 2 \\ 2 & -2 & 1 \\ 2 & 1 & -2 \end{pmatrix}.$$

(Wenn man möchte, kann man hier eine Probe machen, ob S orthogonal ist.) Dann gilt

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

7.5.28 Beispiel. Wir betrachten die Hermite'sche Matrix

$$B = \begin{pmatrix} 1 & 0 & 1+i \\ 0 & -1 & 0 \\ 1-i & 0 & 0 \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{C}).$$

Unser Ziel ist es, eine unitäre Matrix $S \in \mathrm{U}(3)$ zu bestimmen, sodass $S^*BS = S^{-1}BS$ eine reelle Diagonalmatrix ist. Wir bestimmen zuerst das charakteristische Polynom mit der Regel von Sarrus und finden

$$\chi_B = \det \begin{pmatrix} X-1 & 0 & -1-i \\ 0 & X+1 & 0 \\ -1+i & 0 & X \end{pmatrix} = (X+1)^2(X-2).$$

Die Eigenwerte von B sind also $\lambda_1 = -1$ und $\lambda_2 = 2$.

Wir berechnen den Eigenraum zum Eigenwert λ_1.

$$\begin{aligned} \mathrm{Eig}_B(-1) &= \mathrm{Ker}(B + 1I_3) = \mathrm{Ker}\begin{pmatrix} 2 & 0 & 1+i \\ 0 & 0 & 0 \\ 1-i & 0 & 1 \end{pmatrix} \\ &= \mathrm{Ker}\begin{pmatrix} 2 & 0 & 1+i \end{pmatrix} && ((1-i)(1+i) = 2) \\ &= \langle \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1-i \end{pmatrix} \rangle^{\mathbb{C}} \end{aligned}$$

Die beiden Basisvektoren sind orthogonal, wir müssen sie also nur normieren um eine Orthonormalbasis zu erhalten:

$$v_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad v_2 = \frac{1}{\sqrt{3}}\begin{pmatrix} -1 \\ 0 \\ 1-i \end{pmatrix}.$$

Dann berechnen wir den Eigenraum zum Eigenwert 2.

$$\begin{aligned}\mathrm{Eig}_B(2) = \mathrm{Ker}(B - 2I_3) &= \mathrm{Ker}\begin{pmatrix} -1 & 0 & 1+i \\ 0 & -3 & 0 \\ 1-i & 0 & -2 \end{pmatrix} \\ &= \mathrm{Ker}\begin{pmatrix} -1 & 0 & 1+i \\ 0 & 1 & 0 \end{pmatrix} && ((1-i)(1+i) = 2) \\ &= \langle \begin{pmatrix} 2 \\ 0 \\ 1-i \end{pmatrix} \rangle^{\mathbb{C}}\end{aligned}$$

Wir normieren den gefundenen Basisvektor und setzen

$$v_3 = \frac{1}{\sqrt{6}} \begin{pmatrix} 2 \\ 0 \\ 1-i \end{pmatrix}.$$

Damit ist

$$S = \begin{pmatrix} 0 & -\frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \\ 1 & 0 & 0 \\ 0 & \frac{1-i}{\sqrt{3}} & \frac{1-i}{\sqrt{6}} \end{pmatrix}$$

eine unitäre Matrix und es gilt $S^*BS = S^{-1}BS = \mathrm{diag}(-1, -1, 2)$.

7.5.29 **Aufgabe**. Wir betrachten die reelle symmetrische Matrix **L**

$$A = \begin{pmatrix} 3 & 4 \\ 4 & -3 \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{R}).$$

Geben Sie eine orthogonale Matrix $S \in \mathrm{O}(2)$ an, sodass $S^T AS = S^{-1}AS$ eine Diagonalmatrix ist.

7.5.30 **Aufgabe**. Für $a \in \mathbb{R}$ setzen wir **L**

$$B_a = \begin{pmatrix} 0 & a-i \\ a+i & 0 \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{C}).$$

Bestimmen Sie eine unitäre Matrix $S_a \in \mathrm{U}(2)$, sodass $S_a^* B_a S_a = S_a^{-1} B_a S_a$ eine reelle Diagonalmatrix ist.

7.5.31 **Anwendung: Hauptachsentransformation von Quadriken**. Eine *Quadrik* $Q = Q(a, b, c, r)$ ist die Lösungsmenge im $\mathbb{R}^2$ einer quadratischen Gleichung der Form

$$ax_1^2 + 2bx_1x_2 + cx_2^2 = r$$

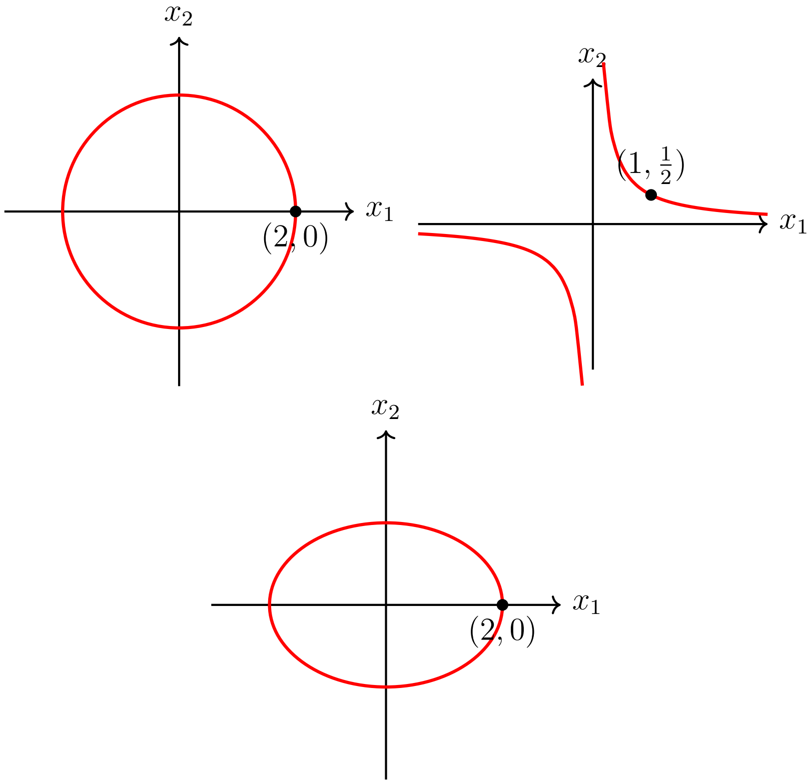


Abbildung 7.2.: $Q(1,0,1,4)$, $Q(0,1,0,1)$ und $Q(1,0,2,4)$.

mit vorgegebenen $a,b,c,r \in \mathbb{R}$. Wir nehmen dabei an, dass mindestens eine der Zahlen a,b,c nicht Null ist.

Zum Beispiel ist $Q(1,0,1,r)$ die Lösungsmenge der Gleichung $x_1^2+x_2^2=r$. Ist $r>0$, dann ist $Q(1,0,1,r)$ ein Kreis mir Radius $\sqrt{r}$. Für $r=0$ besteht $Q(1,0,1,0)=\{0\}$ nur aus einem Punkt. Ist $r<0$, dann hat diese Gleichung keine Lösung und $Q(1,0,1,r)$ ist die leere Menge. Quadriken können aber auch anders aussehen. Zum Beispiel ist $Q(0,1,0,1)$ eine Hyperbel und $Q(1,0,2,4)$ eine Ellipse.

Kann man die möglichen Formen aller Quadriken beschreiben? Ja, dazu kann man den Spektralsatz 7.5.23 verwenden. Wir definieren die symmetrische reelle Matrix

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{R})$$

und stellen fest, dass

$$ax_1^2 + 2bx_1x_2 + cx_2^2 = x^TAx$$

gilt. Ein Koordinatenwechsel, der Längen und Winkel nicht ändert, verändert auch die Gestalt der Quadrik nicht. Wir dürfen also x durch Sy mit einer orthogonalen

Matrix S ersetzen, ohne die Gestalt der Quadrik zu ändern. Wählen wir nun S wie im Spektralsatz 7.5.23, dann ist $S^T AS = \begin{pmatrix} a' & 0 \\ 0 & c' \end{pmatrix}$ eine Diagonalmatrix. Die Quadrik $Q(a, b, c, r)$ hat also dieselbe Gestalt, wie die Quadrik $Q(a', 0, c', r)$. Diese Umformung nennt man die *Hauptachsentransformation*. Um zu verstehen welche Quadriken es gibt, müssen wir uns also nur noch Gleichungen der Form

$$a'x_1^2 + c'x_2^2 = r$$

ansehen. Durch Vertauschen der beiden Koordinaten und/oder Multiplikation mit -1 können wir $a' > 0$ annehmen (dadurch ändert sich ggf. das Vorzeichen von r). Mit diesen Vorbereitungen kann man dann recht einfach alle möglichen Quadriken beschreiben. Die Details überspringen wir hier. Man kann die Rechnungen zum Beispiel im Buch von Bär [Bär, Abschnitt 7.2] finden.

c'	r'	$Q(a', 0, c', r')$
> 0	> 0	Ellipse
	$= 0$	$\{0\}$
	< 0	$\emptyset$
< 0	$\neq 0$	Hyperbel
	$= 0$	2 sich schneidende Geraden
$= 0$	> 0	2 parallele Geraden
	$= 0$	Gerade
	< 0	$\emptyset$

Abbildung 7.3.: Klassifikation der Quadriken $Q(a', 0, c', r')$ mit $a' > 0$.

7.5.32 **Aufgabe**. Es seien $a > 0$ und $c < 0$ in $\mathbb{R}$. Zeigen Sie, dass $Q(a, 0, c, 0)$ aus zwei verschiedenen, sich in der Null schneidenden Geraden besteht. **L**

7.6. Lösungen der Aufgaben in Lektion 7

L7.1.10 Lösung. Für $t \in \mathbb{R}$ definieren wir die Hermite'sche Matrix

$$A_t = \begin{pmatrix} 2 & i & 1 \\ -i & 2 & 0 \\ 1 & 0 & t \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{C}).$$

und die Hermite'sche Form $\sigma_t(u,v) = u^* A_t v$.

Wir möchten das Hauptminorenkriterium anwenden und berechnen dazu die Hauptminoren in Abhängigkeit von t.

Es ist $(A_t)_{\leq 1} = 2$ und somit $\det((A_t)_{\leq 1}) = 2 > 0$.

Es ist $(A_t)_{\leq 2} = \begin{pmatrix} 2 & i \\ -i & 2 \end{pmatrix}$ und mit 4.2.3 erhalten wir

$$\det\bigl((A_t)_{\leq 2}\bigr) = 4 - (-i)i = 4 - 1 = 3 > 0.$$

Schließlich ist $(A_t)_{\leq 3} = A_t$ und mit der Regel von Sarrus 4.2.6 erhalten wir $\det(A) = 4t - 2 - t = 3t - 2$. Dieser Minor ist nur für $t > \frac{2}{3}$ positiv. Das heißt, σ_t ist genau dann ein Skalarprodukt, wenn $t > \frac{2}{3}$ ist.

L7.1.14 Lösung. Wir betrachten die Abbildung $\left\| \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right\| := \max(|x_1|, |x_2|)$ auf $\mathbb{R}^2$.

Da der Betrag einer reellen Zahl immer nicht-negativ ist, gilt $\left\| \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right\| \geq 0$ für alle $x_1, x_2 \in \mathbb{R}$ und damit (N1). Es gilt $\left\| \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right\| = 0$ genau dann, wenn $x_1 = x_2 = 0$ ist, d.h., wenn $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ der Nullvektor ist.

Zu (N3): Sei $a \in \mathbb{R}$.

$$\begin{aligned} \left\| a \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right\| = \left\| \begin{pmatrix} ax_1 \\ ax_2 \end{pmatrix} \right\| &= \max(|ax_1|, |ax_2|) \\ &= \max(|a| \cdot |x_1|, |a| \cdot |x_2|) = |a| \cdot \max(|x_1|, |x_2|) \\ &= |a| \cdot \left\| \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right\|. \end{aligned}$$

Zu (N4): Es seien $x, y \in \mathbb{R}^2$. Dann gilt

$$\begin{aligned}\|x+y\| = \left\|\begin{pmatrix} x_1+y_1 \\ x_2+y_2 \end{pmatrix}\right\| &= \max(|x_1+y_1|, |x_2+y_2|) \\ &\le \max(|x_1|+|y_1|, |x_2|+|y_2|) \qquad ([\text{MG}, 12.2.21]) \\ &\le \max(|x_1|, |x_2|) + \max(|y_1|, |y_2|) \\ &= \|x\| + \|y\|.\end{aligned}$$

L7.1.17 Lösung. Wir berechnen den Abstand der Vektoren

$$u = \begin{pmatrix} 2 \\ 1 \\ -i \\ 0 \end{pmatrix} \quad \text{und} \quad v = \begin{pmatrix} 1-i \\ -1 \\ -1 \\ 1 \end{pmatrix}$$

im unitären Raum $(\mathbb{C}^4, \langle \cdot, \cdot \rangle)$ und erhalten:

$$\begin{aligned} d(u,v) = \|u-v\| &= \left\|\begin{pmatrix} 1+i \\ 2 \\ 1-i \\ -1 \end{pmatrix}\right\| \\ &= \sqrt{2+4+2+1} = \sqrt{9} = 3. \end{aligned}$$

L7.2.4 Lösung. Gegeben ist $u = \begin{pmatrix} 3 \\ -2 \end{pmatrix} \in \mathbb{R}^2$. Wir suchen die zu u orthogonalen Vektoren. Wir machen den Ansatz $v = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Aus der Bedingung $\langle u, v \rangle = 0$ erhalten wir nun

$$0 = \langle u, v \rangle = 3x_1 - 2x_2.$$

Diese lineare Gleichung hat den Lösungsraum

$$L = \left\{ \begin{pmatrix} 2t \\ 3t \end{pmatrix} \in \mathbb{R}^2 \mid t \in \mathbb{R} \right\}.$$

L7.2.5 Lösung. Gegeben sind folgende Vektoren im 3-dimensionalen kanonischen unitären Raum $(\mathbb{C}^3, \langle \cdot, \cdot \rangle)$:

$$u = \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}, \quad v = \begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix}, \quad w = \begin{pmatrix} 1 \\ i \\ 1+i \end{pmatrix}$$

u und v sind orthogonal, denn

$$\langle u, v\rangle = \overline{u}^T v = 1 \cdot 1 + (-i) \cdot (-i) + 0 = 1 - 1 = 0.$$

u und w sind nicht orthogonal, denn

$$\langle u, w\rangle = \overline{u}^T w = 1 \cdot 1 + (-i) \cdot i + 0 = 1 + 1 = 2.$$

v und w sind orthogonal, denn

$$\langle v, w\rangle = \overline{v}^T w = 1 \cdot 1 + i \cdot i + 0 = 1 - 1 = 0.$$

L7.2.9 Lösung. *Behauptung*: Ist U ein Unterraum von V und $S \subseteq U$ ein Erzeugendensystem von U, dann gilt $U^\perp = S^\perp$.

Da $S \subseteq U$ eine Teilmenge ist, gilt jedenfalls $U^\perp \subseteq S^\perp$. Zu zeigen ist die Inklusion $S^\perp \subseteq U^\perp$. Sei $v \in S^\perp$ beliebig. Ist $u \in U$ gegeben, dann schreiben wir u als Linearkombination von Elementen aus S

$$u = \sum_{i=1}^{n} a_i s_i$$

mit $s_i \in S$ und $a_i \in \mathbb{K}$. Das ist möglich, weil S ein Erzeugendensystem von U ist. Mit den Rechenregeln des Skalarproduktes erhalten wir

$$\langle u, v\rangle = \sum_{i=1}^{n} \overline{a_i} \underbrace{\langle s_i, v\rangle}_{=0} = 0;$$

die Terme $\langle s_i, v\rangle$ verschwinden dabei, weil $v \in S^\perp$ und $s_i \in S$ ist. Wir erhalten $v \in U^\perp$ und weil $v \in S^\perp$ beliebig war, folgt daraus $S^\perp \subseteq U^\perp$.

L7.2.11 Lösung. *Behauptung:* $(U_1 + U_2)^\perp = U_1^\perp \cap U_2^\perp$.

Der Vektorraum $U_1 + U_2$ wird erzeugt von $U_1 \cup U_2$. Mit Aufgabe 7.2.9 erhalten wir daraus

$$\begin{aligned}(U_1 + U_2)^\perp = (U_1 \cup U_2)^\perp &= \{v \in V \mid \langle u, v\rangle = 0 \text{ für alle } u \text{ in } U_1 \cup U_2\} \\ &= \{v \in V \mid \langle u, v\rangle = 0 \text{ für alle } u \in U_1 \text{ und alle } u \in U_2\} \\ &= U_1^\perp \cap U_2^\perp.\end{aligned}$$

L7.2.19 Lösung. Es sei V ein endlich-dimensionaler reeller Vektorraum mit einer geordneten Basis $\mathcal{B}$. Zeigen Sie: Es gibt genau ein Skalarprodukt β auf V, sodass $\mathcal{B}$ eine Orthonormalbasis ist.

Sei $n = \dim_{\mathbb{R}} V$. Aus Korollar 3.1.16 wissen wir, dass die Abbildung $\mathrm{M}_{\mathcal{B}} : \mathrm{Bil}_{\mathbb{R}}(V) \to \mathrm{M}_{n,n}(\mathbb{R})$ ein Isomorphismus ist. Das heißt, es gibt genau eine Bilinearform β auf V, sodass $\mathrm{M}_{\mathcal{B}}(\beta) = I_n$ ist. Aus dem Hauptminorenkriterium 7.1.8 sehen wir, dass β ein Skalarprodukt ist.

L7.2.21 Lösung. *Beweis für unitäre Vektorräume*: Sei (V, σ) ein unitärer Vektorraum der Dimension n. Das Skalarprodukt σ ist eine positiv definite Hermite'sche Form, d.h., es handelt sich um eine Hermite'sche Form vom Typ $(n, 0, 0)$ (siehe 3.4.53). Aus dem Satz über die Normalform Hermite'scher Formen 3.4.45 erhalten wir eine geordnete Basis $\mathcal{B} = (v_1, \ldots, v_n)$ von V, sodass

$$\mathrm{M}_{\mathcal{B}}(\sigma) = I_n$$

gilt. Das heißt, es gilt $\langle v_i, v_j \rangle = \delta_{ij}$ und damit ist $\mathcal{B}$ eine Orthonormalbasis von (V, σ).

L7.2.22 Lösung. Sei $(V, \langle \cdot, \cdot \rangle)$ ein endlich-dimensionaler Euklidischer oder unitärer Vektorraum und sei $U \subseteq V$ ein Unterraum. Zeigen Sie: Ist $(v_1, \ldots, v_k)$ eine Orthonormalbasis von U und $(v_{k+1}, \ldots, v_n)$ eine Orthonormalbasis von $U^\perp$, dann ist $(v_1, \ldots, v_n)$ eine Orthonormalbasis von V.

Es gilt $V = U \oplus U^\perp$ (siehe Satz 7.2.8). Aus Satz 2.3.12 schließen wir, dass $(v_1, \ldots, v_k, v_{k+1}, \ldots, v_n)$ eine geordnete Basis von V ist. Alle Vektoren dieser geordneten Basis sind normiert. Sie sind auch paarweise orthogonal. Sei $i, j \in \{1, \ldots, n\}$. Ist $i, j \leq k$, dann gilt $\langle v_i, v_j \rangle = 0$, weil v_i, v_j Elemente einer Orthonormalbasis von U sind. Ist $i, j > k$, dann gilt $\langle v_i, v_j \rangle = 0$, weil v_i, v_j Elemente einer Orthonormalbasis von $U^\perp$ sind. In allen anderen Fällen liegt einer der beiden Vektoren in U und der andere in $U^\perp$, sodass diese wieder orthogonal sind. Es handelt sich also um eine Orthonormalbasis von V.

L7.2.27 Lösung. Wir betrachten den Vektorraum $\mathbb{R}^3$ und das Skalarprodukt $\beta(u, v) = u^T A v$ mit

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 5 & -1 \\ 0 & -1 & 2 \end{pmatrix} \in \mathrm{M}_{3,3}(\mathbb{R}),$$

das wir aus Beispiel 7.1.9 kennen. Wir nehmen die Standardbasis (e_1, e_2, e_3) und überführen sie mit dem Verfahren von Gram-Schmidt in eine Orthonormalbasis.

Es gilt $\beta(e_1, e_1) = 1$. Also ist der erste Vektor normiert und wir setzen $v_1 = e_1$.

Jetzt bestimmen wir v_2':

$$v_2' = e_2 - \beta(v_1, e_2)\, v_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - 2 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}.$$

Um v_2' zu normieren, berechnen wir $\beta(v_2', v_2')$ und erhalten

$$\beta(v_2', v_2') = \begin{pmatrix} -2 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 \\ 2 & 5 & -1 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} = 1.$$

Wir setzen $v_2 = v_2'$.

Schließlich bestimmen wir v_3'. Mit der Gram-Schmidt-Formel ergibt sich

$$v_3' = e_3 - \beta(v_1, e_3)v_1 - \beta(v_2, e_3)v_2 = e_3 + v_2 = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}.$$

Nach einer kurzen Rechnung erhalten wir $\beta(v_3', v_3') = 1$ und wir setzen $v_3 = v_3'$. Dann ist (v_1, v_2, v_3) eine Orthonormalbasis bzgl. des vorgegebenen Skalarproduktes.

L7.2.28 **Lösung.** Sei $(V, \langle \cdot, \cdot \rangle)$ ein endlich-dimensionaler Euklidischer oder unitärer Vektorraum. Angenommen $(u_1, \ldots, u_k)$ ist ein Orthonormalsystem. Die Vektoren $u_1, \ldots, u_k$ sind linear unabhängig wegen 7.2.15, also können wir diese zu einer Basis $\mathcal{B} = (u_1, \ldots, u_n)$ von V ergänzen (siehe 2.2.12). Auf diese Basis wenden wir das Verfahren von Gram-Schmidt an und erhalten eine Orthonormalbasis $(v_1, \ldots, v_n)$ von V.

Behauptung: $v_j = u_j$ für alle $j \leq k$.

Wir zeigen die Behauptung mit Induktion nach j. Da u_1 normiert ist, gilt $v_1 = u_1$. Sei nun $j > 1$. Im Induktionsschritt nehmen wir an, dass $u_1 = v_1, \ldots, u_{j-1} = v_{j-1}$ gilt. Da $(u_1, \ldots, u_k)$ ein Orthonormalsystem ist, gilt

$$v_j' = u_j - \sum_{i=1}^{j-1} \langle v_i, u_j \rangle v_i = u_j - \sum_{i=1}^{j-1} \underbrace{\langle u_i, u_j \rangle}_{=0} u_i = u_j.$$

Der Vektor u_j ist normiert, also gilt $v_j = u_j$.

Damit ist $(v_1, \ldots, v_n) = (u_1, \ldots, u_k, v_{k+1}, \ldots, v_n)$ eine Ergänzung des Orthonormalsystems zu einer Orthonormalbasis.

L7.3.2 **Lösung.** Für alle $v \in V$ und $w \in W$ gilt

$$\langle w, \varphi(v) \rangle_W = \overline{\langle \varphi(v), w \rangle_W} = \overline{\langle v, \psi(w) \rangle_V} = \langle \psi(w), v \rangle_V.$$

L7.3.6 **Lösung.** Gegeben sind lineare Abbildungen $\psi \colon U \to V$ und $\varphi \colon V \to W$. Es sei $u \in U$ und $w \in W$ gegeben. Dann gilt

$$\langle \varphi(\psi(u)), w \rangle_W = \langle \psi(u), \varphi^{\mathrm{ad}}(w) \rangle_V = \langle u, \psi^{\mathrm{ad}}(\varphi^{\mathrm{ad}}(w)) \rangle_U.$$

Daraus folgt $(\varphi \circ \psi)^{\mathrm{ad}} = \psi^{\mathrm{ad}} \circ \varphi^{\mathrm{ad}}$.

L7.3.13 **Lösung.** Wir betrachten den kanonischen Euklidischen Raum $(\mathbb{R}^2, \langle \cdot, \cdot \rangle)$ mit der linearen Abbildung

$$\varphi \colon x \mapsto \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix} x.$$

Die adjungierte Abbildung hat demnach die Matrixdarstellung

$$_{\mathcal{S}}\mathrm{M}_{\mathcal{S}}(\varphi^{\mathrm{ad}}) = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}$$

zur Standardbasis $\mathcal{S}$.

Jetzt betrachten wir die geordnete Basis $\mathcal{B} = (v_1, v_2)$ mit

$$v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Diese Basis ist keine Orthonormalbasis, denn die Vektoren sind nicht orthogonal und der Vektor v_2 ist nicht normiert.

Die Basiswechselmatrix ist

$$_{\mathcal{S}}\mathrm{M}_{\mathcal{B}}(\mathrm{id}) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Wir berechnen die inverse Matrix und erhalten

$$_{\mathcal{B}}\mathrm{M}_{\mathcal{S}}(\mathrm{id}) = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

Mithilfe der Transformationsformel 2.2.28 berechnen wir die Matrixdarstellungen von φ und φ^{ad} bezüglich $\mathcal{B}$:

$$_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi) = {_{\mathcal{B}}\mathrm{M}_{\mathcal{S}}(\mathrm{id})}\,{_{\mathcal{S}}\mathrm{M}_{\mathcal{S}}(\varphi)}\,{_{\mathcal{S}}\mathrm{M}_{\mathcal{B}}(\mathrm{id})} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

$$_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi^{\mathrm{ad}}) = {_{\mathcal{B}}\mathrm{M}_{\mathcal{S}}(\mathrm{id})}\,{_{\mathcal{S}}\mathrm{M}_{\mathcal{S}}(\varphi^{\mathrm{ad}})}\,{_{\mathcal{S}}\mathrm{M}_{\mathcal{B}}(\mathrm{id})} = \begin{pmatrix} 1 & 3 \\ 0 & -1 \end{pmatrix}.$$

Die Formeln aus Satz 7.3.11 sind hier nicht erfüllt.

L7.4.3 **Lösung.** Sei $\varphi\colon V \to V$ eine Isometrie von $(V, \langle\cdot,\cdot\rangle)$. Gegeben sind Vektoren $u, v \in V$. Da φ eine Isometrie ist, erhalten wir

$$\langle\varphi(u), \varphi(v)\rangle = \langle u, v\rangle.$$

Per Definition sind u und v genau dann orthogonal, wenn $\langle u, v\rangle = 0$ gilt, also genau dann, wenn auch $\langle\varphi(u), \varphi(v)\rangle = 0$ ist.

L7.4.5 **Lösung.** Wir betrachten den Euklidischen Vektorraum $(V_{\mathbb{R}}, \operatorname{Re}(\sigma))$ aus Beispiel 7.1.5. Zunächst erinnern wir uns, dass eine $\mathbb{C}$-lineare Abbildung $\varphi\colon V \to V$ sicherlich auch $\mathbb{R}$-linear als Abbildung von $V_{\mathbb{R}}$ ist (vgl. 3.4.38). Ist φ eine unitäre Abbildung, dann gilt $\sigma(\varphi(u), \varphi(v)) = \sigma(u, v)$ und damit auch

$$\operatorname{Re}(\sigma(\varphi(u), \varphi(v))) = \operatorname{Re}(\sigma(u, v))$$

für alle $u, v \in V$. Also ist $\varphi\colon V_{\mathbb{R}} \to V_{\mathbb{R}}$ eine orthogonale Abbildung bzgl. des Skalarproduktes $\operatorname{Re}(\sigma)$.

L7.4.11 **Lösung.** Sei $B \in \mathrm{M}_{n,n}(\mathbb{C})$.

Es ist klar, dass (i) die Aussagen (iii) und (iv) impliziert: Ist B unitär, dann gilt $B^{-1} = B^*$ und damit auch $B^*B = I_n$ und $BB^* = I_n$.

„(i) $\Rightarrow$ (ii)“: Ist B unitär, dann ist B invertierbar und $B^* = B^{-1}$. In Aufgabe 3.4.16 wurde gezeigt, dass B^* dann ebenfalls invertierbar ist und, dass $(B^*)^{-1} = (B^{-1})^*$ gilt. Damit haben wir

$$(B^*)^{-1} = (B^{-1})^* = (B^*)^*$$

und B^* ist eine unitäre Matrix.

„(ii) $\Rightarrow$ (i)“: Diese Richtung folgt aus „(i) $\Rightarrow$ (ii)“, denn aus Aufgabe 3.4.16 ist bekannt, dass $(B^*)^* = B$ gilt.

„(iii) $\Rightarrow$ (i)“: Wir nehmen an, es gilt $B^*B = I_n$. Aus den Mathematischen Grundlagen [MG, 4.5.6] wissen wir, dass B dann invertierbar ist und es gilt $B^{-1} = B^*$.

Mit demselben Argument folgt „(iv) $\Rightarrow$ (ii)“ und damit die Äquivalenz aller Aussagen.

L7.4.12 **Lösung.** Sei $B \in \mathrm{U}(n)$. Dann gilt

$$\begin{aligned}
(B^T)^* &= \overline{(B^T)^T} = \overline{B} && (1.2.32\ (iv))\\
&= (\overline{B}^T)^T = (B^*)^T && (\text{Def. } *)\\
&= (B^{-1})^T && (B \text{ unitär})\\
&= (B^T)^{-1} && (1.2.34);
\end{aligned}$$

d.h., B^T ist unitär.

L7.4.14 **Lösung.** Die Menge $U(1)$ besteht genau aus den komplexen Zahlen z mit der Eigenschaft

$$1 = z^*z = \overline{z}z = |z|^2.$$

Damit ist $U(1)$ genau der Einheitskreis in $\mathbb{C}$; siehe 1.5.13.

L7.4.17 **Lösung.** Es sei $x = \big(x_1, x_2\big)$. Es ist also $\|x\| = \sqrt{x_1^2 + x_2^2}$. Dann ist

$$D_\alpha x = \begin{pmatrix} x_1 \cos(\alpha) - x_2 \sin(\alpha) \\ x_1 \sin(\alpha) + x_2 \cos(\alpha) \end{pmatrix}.$$

Dieser Vektor hat ebenfalls die Länge $\|D_\alpha x\| = \|x\|$, denn

$$\|D_\alpha x\|^2 = x_1^2 \cos^2(\alpha) + x_2^2 \sin^2(\alpha) + x_1^2 \sin^2(\alpha) + x_2^2 \cos^2(\alpha) = x_1^2 + x_2^2.$$

Hier verwenden wir wieder $\cos^2 \alpha + \sin^2 \alpha = 1$.

Mit der Formel aus 7.2.1 ist der Winkel zwischen x und $D_\alpha x$ gegeben durch

$$\arccos\Big(\frac{\langle D_\alpha x, x\rangle}{\|D_\alpha x\| \cdot \|x\|}\Big) = \arccos\Big(\frac{\cos(\alpha)(x_1^2 + x_2^2)}{\|x\|^2}\Big) = \arccos\cos(\alpha).$$

Ist $\alpha \in [0, \pi]$ (D_α beschreibt also eine Drehung von höchstens 180°), dann ist $\arccos\cos(\alpha) = \alpha$. Ist $\alpha \in (\pi, 2\pi)$, dann ist der Drehwinkel größer als 180° und $\arccos\cos(\alpha) = 2\pi - \alpha$. Es wird also der kleinere der beiden Winkel gemessen.

L7.4.18 **Lösung.** Es seien $A, B \in \mathrm{U}(n)$. Mit den Rechenregeln aus Aufgabe 3.4.16 erhält man

$$(AB)^*AB = B^* \underbrace{A^*A}_{=I_n} B = B^*B = I_n.$$

Das heißt, AB ist ebenfalls unitär.

L7.4.23 **Lösung.** Wir definieren $v_1 = \frac{1}{3}\begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}$. Da v_1 normiert ist, gibt es nach dem Basisergänzungssatz 7.2.23 eine Orthonormalbasis (v_1, v_2, v_3). Schreiben wir diese Vektoren als Spalten in eine Matrix, dann ist diese Matrix orthogonal.

Wir bestimmen eine Orthonormalbasis mit dem Verfahren von Gram-Schmidt. Dazu wählen wir zunächst

$$u_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad u_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

denn dann ist (v_1, u_2, u_3) eine Basis von $\mathbb{R}^3$. Dabei handelt es sich um eine Wahl! Hier ist natürlich auch jede andere Ergänzung zu einer Basis möglich. Wir überführen diese Basis in eine Orthonormalbasis.

Wir setzen

$$v_2' = u_2 - \underbrace{\langle v_1, u_2 \rangle}_{=0} v_1 = u_2$$

und normieren diesen Vektor zu $v_2 = \frac{1}{\sqrt{2}} u_2$.

Dann setzen wir

$$v_3' = u_3 - \langle v_1, u_3 \rangle v_1 - \underbrace{\langle v_2, u_3 \rangle}_{=0} v_2 = \begin{pmatrix} -\frac{2}{9} \\ -\frac{2}{9} \\ \frac{8}{9} \end{pmatrix}.$$

Die Norm von v_3 ist $\|v_3\| = \sqrt{\frac{72}{81}} = \frac{2\sqrt{2}}{3}$. Der Vektor

$$v_3 = \frac{3}{2\sqrt{2}} \begin{pmatrix} -\frac{2}{9} \\ -\frac{2}{9} \\ \frac{8}{9} \end{pmatrix} = \begin{pmatrix} -\frac{1}{3\sqrt{2}} \\ -\frac{1}{3\sqrt{2}} \\ \frac{2\sqrt{2}}{3} \end{pmatrix}$$

ist normiert. Insgesamt ist damit

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

eine orthogonale Matrix mit der vorgegebenen ersten Spalte.

(Wie kompliziert die Rechnung wird, hängt wesentlich davon ab, dass man eine „gute“ Basis für das Gram-Schmidt-Verfahren wählt. Was denken Sie, warum wurden die Vektoren u_2 und u_3 hier gewählt?)

L7.5.5 **Lösung.** Es sei $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$ mit $\varphi^{\mathrm{ad}} = -\varphi$. Dann gilt für alle $v \in V$

$$\varphi^{\mathrm{ad}}(\varphi(v)) = -\varphi(\varphi(v)) = \varphi(-\varphi(v)) = \varphi(\varphi^{\mathrm{ad}}(v)).$$

Da v beliebig war, folgt daraus $\varphi^{\mathrm{ad}} \circ \varphi = \varphi \circ \varphi^{\mathrm{ad}}$. Das heißt, φ ist normal.

L7.5.7 Lösung. Sei $(V, \langle \cdot, \cdot \rangle)$ ein endlich-dimensionaler Euklidischer oder unitärer Vektorraum und sei φ ein normaler Endomorphismus. Es seien $u, v \in V$ Eigenvektoren von φ zu verschiedenen Eigenwerten $\lambda_1 \neq \lambda_2$. Dann gilt

$$\begin{aligned} \lambda_2 \langle u, v \rangle &= \langle u, \lambda_2 v \rangle = \langle u, \varphi(v) \rangle \\ &= \langle \varphi^{\mathrm{ad}}(u), v \rangle = \langle \overline{\lambda_1} u, v \rangle \\ &= \lambda_1 \langle u, v \rangle. \end{aligned}$$

Wegen $\lambda_1 \neq \lambda_2$ folgt daraus $\langle u, v \rangle = 0$, d.h., u und v sind orthogonal.

L7.5.9 Lösung. Um ein Gegenbeispiel für Euklidische Räume zu finden, suchen wir eine normale Abbildung, die aber keine reellen Eigenwerte hat. Wir betrachten dazu den kanonischen Euklidischen Raum $(\mathbb{R}^2, \langle \cdot, \cdot \rangle)$ mit dem Endomorphismus

$$\varphi \colon x \mapsto \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} x.$$

Die Matrix $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ ist orthogonal (siehe 7.4.13) und nach Satz 7.4.24 ist φ damit eine orthogonale Abbildung. Allerdings hat A das charakteristische Polynom $\chi_A = X^2 + 1$ und dieses besitzt keine Nullstellen in $\mathbb{R}$. Das heißt, A hat *keine* Eigenwerte.

L7.5.18 Lösung. Es sei V unitär. Satz 7.3.11 impliziert ${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi^{\mathrm{ad}}) = {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)^*$. Die Matrixdarstellung bestimmt den Endomorphismus eindeutig (vgl. 2.2.23), also gilt $\varphi = \varphi^{\mathrm{ad}}$ dann und nur dann, wenn

$${}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi) = {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi^{\mathrm{ad}}) = {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(\varphi)^*$$

ist; d.h., genau dann, wenn die Matrixdarstellung Hermite'sch ist.

L7.5.19 Lösung. Sei $(V, \langle \cdot, \cdot \rangle)$ ein endlich-dimensionaler Euklidischer oder unitärer Vektorraum und sei $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$. Wir erinnern uns, dass $\varphi = (\varphi^{\mathrm{ad}})^{\mathrm{ad}}$ gilt; siehe 7.3.2 (c). Mit der Formel aus Aufgabe 7.3.6 ist

$$(\varphi^{\mathrm{ad}} \circ \varphi)^{\mathrm{ad}} = \varphi^{\mathrm{ad}} \circ (\varphi^{\mathrm{ad}})^{\mathrm{ad}} = \varphi^{\mathrm{ad}} \circ \varphi.$$

Das heißt, $\varphi^{\mathrm{ad}} \circ \varphi$ ist selbstadjungiert.

L7.5.23 Lösung.

<u>*Beweis*</u>. Auf dem kanonischen Euklidischen Raum $(\mathbb{R}^n, \langle\cdot,\cdot\rangle)$ betrachten wir die lineare Abbildung $f_A\colon \mathbb{R}^n \to \mathbb{R}^n$ mit $x \mapsto Ax$. Dann ist f_A selbstadjungiert. Das folgt aus Lemma 7.5.18, denn die Standardbasis $\mathcal{C}$ ist orthonormal und die Matrixdarstellung $A = {}_{\mathcal{C}}\mathrm{M}_{\mathcal{C}}(f_A)$ ist symmetrisch.

Aus dem Spektralsatz für selbstadjungierte Endomorphismen erhalten wir eine Orthonormalbasis $\mathcal{B}$ von $\mathbb{R}^n$, die aus Eigenvektoren von f_A besteht. Wir zuvor verwenden wir die Transformationsformel für Endomorphismen 4.4.5:

$$D = {}_{\mathcal{B}}\mathrm{M}_{\mathcal{B}}(f_A) = {}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\mathrm{id})^{-1}\, A\, {}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\mathrm{id});$$

dabei ist D eine reelle Diagonalmatrix.

Die Matrix $S = {}_{\mathcal{C}}\mathrm{M}_{\mathcal{B}}(\mathrm{id})$ ist orthogonal, weil $\mathcal{B}$ und $\mathcal{C}$ Orthonormalbasen sind und die identische Abbildung $\mathrm{id}\colon \mathbb{R}^n \to \mathbb{R}^n$ orthogonal ist (siehe Satz 7.4.24). Es gilt also $S^T = S^{-1}$. □

L7.5.29 Lösung. Wir betrachten die reelle symmetrische Matrix

$$A = \begin{pmatrix} 3 & 4 \\ 4 & -3 \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{R}).$$

Das charakteristische Polynom ist

$$\chi_A = (X-3)(X+3) - 16 = X^2 - 25.$$

Die Eigenwerte von A sind also $\lambda_1 = 5$ und $\lambda_2 = -5$. Wir berechnen Orthonormalbasen der Eigenräume.

$$\mathrm{Eig}_A(5) = \mathrm{Ker}(A - 5I_2) = \begin{pmatrix} -2 & 4 \\ 4 & -8 \end{pmatrix} = \langle \begin{pmatrix} 2 \\ 1 \end{pmatrix} \rangle^{\mathbb{R}}.$$

Der Eigenraum ist ein-dimensional. Es ist also ausreichend, den gefundenen Eigenvektor zu normieren. Wir setzen

$$v_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

$$\mathrm{Eig}_A(-5) = \mathrm{Ker}(A + 5I_2) = \begin{pmatrix} 8 & 4 \\ 4 & 2 \end{pmatrix} = \langle \begin{pmatrix} 1 \\ -2 \end{pmatrix} \rangle^{\mathbb{R}}.$$

Der Eigenraum ist wieder ein-dimensional und wir setzen

$$v_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ -2 \end{pmatrix}.$$

Damit ist
$$S = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ 1 & -2 \end{pmatrix}$$
eine orthogonale Matrix und $S^T AS = S^{-1}AS = \operatorname{diag}(5, -5)$.

L7.5.30 **Lösung.** Für $a \in \mathbb{R}$ setzen wir
$$B_a = \begin{pmatrix} 0 & a-i \\ a+i & 0 \end{pmatrix} \in \mathrm{M}_{2,2}(\mathbb{C}).$$

Da a reell ist, ist B_a eine Hermite'sche Matrix, d.h., $B_a^* = B_a$. Das charakteristische Polynom von B_a ist
$$\chi(B_a) = \begin{pmatrix} X & -a+i \\ -a-i & X \end{pmatrix} = X^2 - (a^2+1).$$

Damit hat B_a die beiden Eigenwerte $\lambda_1 = \sqrt{a^2+1}$ und $\lambda_2 = -\sqrt{a^2+1}$. Weil a^2+1 immer positiv ist, sind diese beiden Wurzeln reell; die Eigenwerte sind verschieden, denn $\lambda_1 > 0$ und $\lambda_2 < 0$. Der Eigenraum zum Eigenwert λ_1 ist
$$\operatorname{Eig}_{B_a}(\lambda_1) = \operatorname{Ker}(B_a - \lambda_1 I_2) = \operatorname{Ker}(\begin{pmatrix} -\sqrt{a^2+1} & a-i \\ a+i & -\sqrt{a^2+1} \end{pmatrix}).$$

Da B_a zwei verschiedene Eigenwerte hat, hat die Matrix $B_a - \lambda_1 I_2$ notwendig den Rang 1. Will man sich davon überzeugen, kann man die erste Zeile mit $-\lambda_1$ und die zweite mit $a-i$ multiplizieren um zwei gleiche Zeilen zu erhalten. Es gilt insbesondere
$$\begin{aligned} \operatorname{Eig}_{B_a}(\lambda_1) &= \operatorname{Ker}(\begin{pmatrix} -\sqrt{a^2+1} & a-i \\ a+i & -\sqrt{a^2+1} \end{pmatrix}) \\ &= \operatorname{Ker}(\begin{pmatrix} -\sqrt{a^2+1} & a-i \end{pmatrix}) = \langle \begin{pmatrix} \sqrt{a^2+1} \\ a+i \end{pmatrix} \rangle^{\mathbb{C}}. \end{aligned}$$

Der Eigenraum ist ein-dimensional. Eine Orthonormalbasis erhalten wir, indem wir den gefundenen Basisvektor normieren. Man erhält dann
$$v_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \frac{a+i}{\sqrt{a^2+1}} \end{pmatrix}.$$

Der Eigenraum zum Eigenwert λ_2 ist
$$\operatorname{Eig}_{B_a}(\lambda_2) = \operatorname{Ker}(\begin{pmatrix} \sqrt{a^2+1} & a-i \\ a+i & \sqrt{a^2+1} \end{pmatrix}) = \langle \begin{pmatrix} -\sqrt{a^2+1} \\ a+i \end{pmatrix} \rangle^{\mathbb{C}}.$$

Wir normieren den Basisvektor und haben damit

$$v_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ \frac{a+i}{\sqrt{a^2+1}} \end{pmatrix}.$$

Die unitäre Matrix

$$S_a = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ \frac{a+i}{\sqrt{a^2+1}} & \frac{a+i}{\sqrt{a^2+1}} \end{pmatrix}$$

erfüllt $S_a^* B_a S_a = S_a^{-1} B_a S_a = \operatorname{diag}(\lambda_1, \lambda_2)$.

L7.5.32 Lösung. Es seien $a > 0$ und $c < 0$ in $\mathbb{R}$. Die Quadrik $Q(a, 0, c, 0)$ ist die Lösungsmenge der Gleichung $ax_1^2 + cx_2^2 = 0$. Wir setzen $t = -c/a$. Dann ist $Q(a, 0, c, 0)$ die Lösungsmenge der Gleichung

$$x_1^2 - tx_2^2 = 0.$$

Wegen $c < 0$ und $a > 0$ ist t eine positive reelle Zahl. Wir setzen $s = \sqrt{t}$. Dann gilt

$$x_1^2 - tx_2^2 = (x_1 + sx_2)(x_1 - sx_2).$$

Dieser Ausdruck ist genau dann Null, wenn $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ auf einer der beiden Geraden $x_1 = \pm sx_2$ liegt. Da $s \neq 0$ ist sind diese Geraden verschieden.

Literaturverzeichnis

[Bär] C. Bär. *Lineare Algebra und analytische Geometrie*. Springer Spektrum, Berlin, 2018.

[Beu] A. Beutelspacher. *Lineare Algebra*. Springer Lehrbuch. Springer-Verlag, Berlin, 2014. 8. Auflage.

[Bo] S. Bosch. *Lineare Algebra*. Springer Lehrbuch. Springer-Verlag, Berlin, 2021. 6. Auflage.

[Can95] G. Cantor. Beiträge zur Begründung der transfiniten Mengenlehre. Art. I. *Math. Ann.*, 46:481–512, 1895. doi:10.1007/BF02124929.

[Fi] G. Fischer and B. Springborn. *Lineare Algebra*. Grundkurs Mathematik. Springer-Verlag, Berlin, 2020. 19. Auflage.

[Fis17] G. Fischer. *Lehrbuch der Algebra*. Springer-Lehrbuch. Springer-Verlag, Berlin, 2017. 4. Auflage.

[Gö] L. Göllmann. *Lineare Algebra*. Springer Lehrbuch. Springer-Verlag, Berlin, 2020. 2. Auflage.

[Jec03] T. Jech. *Set theory*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2003. The third millennium edition, revised and expanded.

[KaSt] C. Karpfinger and H. Stachel. *Lineare Algebra*. Springer Lehrbuch. Springer-Verlag, Berlin, 2020.

[Kne50] H. Kneser. Eine direkte Ableitung des Zornschen Lemmas aus dem Auswahlaxiom. *Math. Z.*, 53:110–113, 1950. doi:10.1007/BF01162404.

[Lew91] J. Lewin. A simple proof of Zorn's lemma. *Amer. Math. Monthly*, 98(4):353–354, 1991. doi:10.2307/2323807.

[LM11] A. Langville and C. Meyer. *Google's PageRank and Beyond: The Science of Search Engine Rankings*. Princeton University Press, 2011.

[MG] L. Unger. *Mathematische Grundlagen*. FernUniversität in Hagen, 2020.

Symbolverzeichnis

0: Nullelement, 15
$\mathbb{1}_n$: Partition $(1,\ldots,1)$ von n, 390

$\overline{A}$: komplexe Konjugation der Matrix A, 202
a^{-1}: das Inverse zu a, 18
a^{-1}: inverses Element zu a, 8
$a \mid b$: a teilt b, 26
$\mathrm{Abb}(X,Y)$: Abbildungen $X \to Y$, xiii
A_{ij}: Löschen der i-ten Zeile und j-ten Spalte von A, 264
$\alpha(\lambda)$: algebraische Vielfachheit des Eigenwertes λ, 327
Alt_n: alternierende Gruppe, 245
A^T: transponierte Matrix zu A, 22

$\mathrm{Bil}_K(U)$: Raum aller Bilinearformen auf U, 157

χ_A: charakteristisches Polynom von A, 324

$\bigwedge^n V^*$: Raum der alternierenden n-linearen Formen auf V, 284
$\delta_{i,j}$: Kronecker-Symbol, 128
$\det(A)$: Determinante von A, 246

$\mathrm{Eig}_\varphi(\lambda)$: Eigenraum von φ zum Eigenwert λ, 315
$\mathrm{End}(V)$: Endomorphismenring von V, 278

$f(x)$: x eingesetzt in f, 40
φ^*: duale Abbildung zu φ, 131
$fK[X]$: von f erzeugtes Hauptideal, 50
$\mathrm{FS}(\sigma)$: Menge aller Fehlstände von σ, 242

$\gamma(\lambda)$: geometrische Vielfachheit des Eigenwertes λ, 316
$\mathrm{GL}_n(K)$: allgemeine lineare Gruppe, 10

$H_\varphi(\lambda)$: Hauptraum zum Eigenwert λ, 410
$\mathrm{Herm}(V)$: Menge der Hermite'schen Formen auf V., 200
$\mathrm{Hom}_K(V,W)$: Menge aller K-linearen Abb. von V nach W, 126

i: imaginäre Einheit, 57
$\mathcal{I}_\varphi$: Verschwindungsideal, 333
$\mathrm{Isom}(V)$: Isometriegruppe, 466

$J(\lambda,q)$: Jordanblock zum Eigenwert λ und zur Partition q, 404

$k+n\mathbb{Z}$: Restklasse von k modulo n, 27
$\mathrm{Ker}(\varphi)$: Kern von φ, 101
$K^\times$: Einheitengruppe von K, 10
$K[X]$: Polynomring über K, 34

$L_K^{(n)}(U,V)$: Raum aller n-linearen Abbildungen, 281

$M_{\mathcal{B}}(\beta)$: Matrixdarstellung der Bilinearform β, 162
$M_{m,n}(R)$: Menge der $(m \times n)$-Matrizen über R, 20
$M_{\mathcal{B}}(\sigma)$: Matrixdarstellung der Hermite'schen Form σ, 205
μ_φ: Minimalpolynom von φ, 334

$\mathbb{N}$: natürliche Zahlen $1, 2, 3, \ldots$, xiii
$\mathbb{N}_0$: natürliche Zahlen mit 0 , xiii
$\mathrm{Ni}(A)$: Nilpotenzindex von A, 377

$\mathrm{O}(n)$: orthogonale Gruppe, 470

$p(\varphi)$: Rangpartition von φ, 391
$\mathcal{P}(X)$: Potenzmenge von X, 91
p^*: duale Partition zu p, 389

$\mathbb{Q}$: die rationalen Zahlen, xiii

$\mathbb{R}$: die reellen Zahlen, xiii
$\mathrm{Rg}(\beta)$: Rang der Bilinearform β, 176
$R^\times$: Einheiten in R, 18

S^1 : Einheitskreis in $\mathbb{C}$, 58
$\mathrm{sgn}(\sigma)$: Signatur von σ, 242
$\mathrm{SL}_n(K)$: spezielle lineare Gruppe, 259
$\mathrm{Spur}(A)$: Spur der Matrix A, 21

$\sum_{i=1}^n I_i$: Summe der Ideale $I_1, \ldots, I_n$, 52
$\mathrm{Sym}(X)$: symmetrische Gruppe, 237
S_n: symmetrische Gruppe, 237

$\mathrm{Trg}(\sigma)$: Träger einer Permutation σ, 240

$\mathrm{U}(n)$: unitäre Gruppe, 470
$U_1 + U_2$: Summe der Unterräume U_1, U_2, 105
$U_1 \oplus U_2$: direkte Summe von Unterräumen, 107

V^*: Dualraum zu V, 126
$v + U$: affiner Unterraum zu U durch v., 116
V/U: Faktorraum von V modulo U, 117

$\mathbb{Z}$: die ganzen Zahlen, xiii
$Z_\varphi(v)$: von v aufgespannter zyklischer Unterraum, 384
$\mathbb{Z}/n\mathbb{Z}$: Restklassenring modulo n, 27
$\overline{z}$: Die zu z komplex konjugierte Zahl, 59
$|z|$: Betrag von $z \in \mathbb{C}$, 57

Index

A

abgeschlossen, 12
Abspaltungssatz, 42
Abstand, 446
adjungiert, 459
Adjunkte, 271
affiner Unterraum, 116
ähnliche Matrizen, 259
Ähnlichkeitsklassen, 260
Algebra, 330
algebraisch abgeschlossen, 48
algebraische Vielfachheit, 327, 329
allgemeine lineare Gruppe, 10, 21
alternierend, 253
alternierende
  Bilinearform, 195
  Gruppe, 245
  Matrix, 196
  Multilinearform, 284
annullieren, 333
antisymmetrisch, 91
äquivalente Matrizen, 138
Äquivalenzklasse, 136
Äquivalenzrelation, 134
Assoziativgesetz, 7
ausgeartete
  Bilinearform, 174
  Hermite'sche Form, 201
Auswahlaxiom, 90

B

Basis, 98
Basisergänzungssatz, 99
  für Orthonormalbasen, 454
Basiswechsel, 104
Betrag
  einer komplexen Zahl, 57
Bilinearform, 157, 281
binomischer Lehrsatz, 417
Blockdiagonalmatrix, 23
Blockdreiecksmatrix, 269

C

Charakterisierungssatz, 253
charakteristisches Polynom
  einer Matrix, 324
  eines Endomorphismus, 329
Cramer'sche Regel, 275

D

Darstellungsmatrix, 102
Determinante, 246
  eines Endomorphismus, 279
diagonalisierbar, 318
Diagonalmatrix, 250
Diagramm
  einer Partition, 388
direkte Summe, 107
Diskriminante, 62
Distributivgesetze, 14
Division mit Rest

ganzer Zahlen, 26
Polynome, 38
Drehmatrix, 471
Dreiecksmatrix, 251
Dualbasis, 129
Duale Abbildung, 131
duale Partition, 389
Dualraum, 126

E

Eigenraum, 315
verallgemeinerter, 409
Eigenvektor, 313
Eigenwert, 313
Einheit, 18
Einheitengruppe, 19
Einheitskreis, 58
Einheitsmatrix, 20
Einheitswurzel, 376
Einschränkung
von Bilinearformen, 161
Einselement, 15
Einsetzen
in Polynome, 40, 331
Eintrag
einer Matrix, 20
endlich-dimensional, 100
Endomorphismus, 278
Entwicklungssatz von Laplace, 266
Euklidischer Algorithmus, 32
Euklidischer Vektorraum, 439
kanonisch, 440

F

Faktorabbildung, 120
Faktorraum, 118
Faserrelation, 135
Fehlstand
einer Permutation, 242
Fibonacci-Folge, 419
Filtrierung, 380
Fixpunkt, 240

G

geometrische Vielfachheit, 316
geordnete Basis, 100
goldener Schnitt, 420
Google-Matrix, 350
Grad eines Polynoms, 35
Gram-Schmidt-Verfahren, 455
größter gemeinsamer Teiler
von Polynomen, 52
Gruppe, 7
abelsche, 9
Gruppenaxiome, 7
Gruppenmultiplikation, 7
Gruppenverknüpfung, 7

H

Hauptachsentransformation, 492
Hauptideal, 50
Hauptminorenkriterium, 441
Hauptraum, 410
Hermite'sche Form, 200
Hermite'sche Kongruenz, 207
Hermite'sche Matrix, 204
Homomorphismus, 126

I

Ideal, 49
imaginäre Einheit, 57
Imaginärteil, 57
indefinite
reelle symmetrische Bilinearform, 187
induzierte Ordnung, 92
invarianter Unterraum, 309
inverses Element, 7, 8
invertierbar, 18
Isometrie, 466

Isometriegruppe, 469
Isomorphismus, 101
    kanonisch, 127

J
Jordanblock, 404
Jordankästchen, 404
Jordan'sche Normalform, 404, 405

K
kanonische Projektion, 120
Kern, 101
Kette, 92
Klassifikation
    Hermite'scher Formen, 216
    reeller symmetrischer
        Bilinearformen, 194
Kommutativgesetz, 9
Komplement, 110
komplementär, 110
komplexe Konjugation, 59
Kongruenz
    von Bilinearformen, 167
    von Hermite'schen Formen, 208
    von Matrizen, 137
Kongruenzklasse
    von Matrizen, 137
Koordinatenvektor, 101
Körper, 19
    der komplexen Zahlen, 56
Kronecker-Symbol, 128
Kürzungsregeln
    in Gruppen, 8

L
Länge eines Vektors, 443
Leitkoeffizient, 35
Linear in jeder Spalte, 252
linear unabhängig, 97
lineare Abbildung, 100
Linearfaktor, 42
Linearform, 126

M
Matrix, 20
Matrixdarstellung, 102
    einer Bilinearform, 162
    einer Hermite'schen Form, 205
Matrizenmultiplikation, 20
maximales Element, 91
Metrik, 447
Minimalpolynom, 334
Minor, 264
Mitternachtsformel, 64

N
Nebenklasse, 116
negative Definitheit
    Hermite'scher Formen, 202
    reeller symmetrischer
        Bilinearformen, 187
neutrales Element, 7
nicht-negative Matrix, 342
nilpotent, 377
Nilpotenzindex, 377
n-linear, 281
Norm, 443
normaler Endomorphismus, 479
Normalform, 169
    Hermite'scher Formen, 214
    reeller symmetrischer
        Bilinearformen, 192
    symmetrischer Bilinearformen
        über $\mathbb{C}$, 185
Normalformproblem
    für Bilinearformen, 169
normiert, 35
normierter Erzeuger
    eines Ideals, 50
Nullelement, 15

Nullmatrix, 20
Nullraum, 101
Nullring, 16
Nullstelle
  eine Polynoms, 41
Nullvektor, 95

O
obere Schranke, 91
Orientierung, 248
orthogonal, 451
orthogonale
  Abbildung, 466
  Gruppe, 469, 473
  Matrix, 470
  Vektoren, 449
orthogonales Komplement, 449
Orthonormalbasis, 452
Orthonormalsystem, 451

P
Parallelenschar, 115
Parallelepiped, 250
partielle Ordnung, 90
Partition, 388
Permutation, 238
Permutationsmatrix, 475
Polarisationsformel, 180
Polynom, 34
  konstantes, 35
Polynomdivision mit Rest, 38
Polynomfunktion, 40
Polynomring, 34
positive Definitheit
  Hermite'scher Formen, 202
  reeller symmetrische
    Bilinearformen, 187
positive Matrix, 342
Potenzmenge, 91
Primzahl, 30

Q
quadratische Form, 181
Quadratwurzel, 62
Quotient, 26, 38

R
Rang
  einer Bilinearform, 176
Rangpartition, 391
Realteil, 57
  einer Hermite'schen Form, 210
reflexiv, 91
Regel von Sarrus, 248
reguläre
  Bilinearform, 174
  Hermite'sche Form, 201
Rest, 26, 38
Restklasse, 27
Restklassenring, 29
Restriktion der Skalare, 209
Ring, 14
  kommutativer, 15
  unitärer, 15
Russell'sches Paradoxon, 89

S
Satz
  vom Faktorraum, 118
  vom Komplement, 111
  vom orthogonalen Komplement,
    450
  von Cayley-Hamilton, 337
  von den Restklassen, 27
  von der nilpotenten Normalform,
    396
  von der Signatur, 243
schiefsymmetrisch
  Bilinearform, 195
  Matrix, 196
selbstadjungiert, 479

Signatur
- einer Hermite'schen Form, 213
- einer Permutation, 242
- einer reellen symmetrischen Bilinearform, 189

Skalarprodukt, 187, 439
Spektralsatz
- für normale Endomorphismen, 480, 482
- für selbstadjungierte Endomorphismen, 486
- für unitäre Abbildungen, 482

spezielle lineare Gruppe, 259
Spiegelungsmatrix, 472
Spur
- einer Matrix, 21

Standardbasis, 102
Standardskalarprodukt, 188, 440
stochastische Matrix, 342
Stufe, 379
- verallgemeinerter Eigenraum, 409

Stützvektor, 116
Summe
- von Idealen, 51
- von Unterräumen, 105

symmetrisch
- Bilinearform, 178
- Matrix, 178

symmetrische Gruppe, 237

T

Teiler, 26
- im Polynomring, 40

teilerfremd, 53
Tensorprodukt, 160
Totalordnung, 91
Träger
- einer Permutation, 240

Trägheitssatz
- für Hermite'sche Formen, 212
- für symmetrische Bilinearformen, 189

Transformationsformel
- für Bilinearformen, 164
- für Endomorphismen, 278
- für Hermite'sche Formen, 206

Transformationsmatrix, 103
transitiv, 91
transponieren, 22
Transposition, 241
Typ
- einer Hermite'schen Form, 213
- einer reellen symmetrischen Bilinearform, 189

U

Übergangsmatrix, 348
Unbestimmte, 34
unendlich-dimensional, 100
unitär
- Abbildung, 466
- Gruppe, 469, 473
- Matrix, 470
- Vektorraum, 439

unitärer Vektoraum
- kanonischer, 440

Untergruppe, 12
Unterkörper, 19
Unterring, 16

V

Vektorraum, 95
Verschwindungsideal, 333
Vertreter
- einer Restklasse, 28

Vielfachheit, 43

W

wohldefiniert, 29, 121

Z

zerfällt in Linearfaktoren, 45
Zerlegung in invariante Unterräume, 312
Zufallssurfer-Modell, 347
zurückziehen
    einer Bilinearform, 173
zyklischer Unterraum, 384