

$$v = k_1 v_1 + \dots + k_n v_n$$

LEHRBUCH

Albrecht Beutelspacher

Lineare Algebra

Eine Einführung in die Wissenschaft
der Vektoren, Abbildungen und Matrizen

8. Auflage



Springer Spektrum

Lineare Algebra

Albrecht Beutelspacher

Lineare Algebra

Eine Einführung in die Wissenschaft der
Vektoren, Abbildungen und Matrizen

8., aktualisierte Auflage



Springer Spektrum

Prof. Dr. Albrecht Beutelspacher
Justus-Liebig-Universität Gießen
Giessen, Deutschland

ISBN 978-3-658-02412-3
DOI 10.1007/978-3-658-02413-0

ISBN 978-3-658-02413-0 (eBook)

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Springer Spektrum

© Springer Fachmedien Wiesbaden 1994, 1995, 1998, 2000, 2001, 2003, 2010, 2014

Dieses Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung, die nicht ausdrücklich vom Urheberrechtsgesetz zugelassen ist, bedarf der vorherigen Zustimmung des Verlags. Das gilt insbesondere für Vervielfältigungen, Bearbeitungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

Planung und Lektorat: Ulrike Schmickler-Hirzebruch | Barbara Gerlach

Gedruckt auf säurefreiem und chlorfrei gebleichtem Papier.

Springer Spektrum ist eine Marke von Springer DE. Springer DE ist Teil der Fachverlagsgruppe Springer Science+Business Media
www.springer-spektrum.de

Mathematik – eine Mutprobe?

Mein stolzes Beginnen lief darauf hinaus: Allerkleinstes – auch Prosaisches nicht ausgeschlossen – exakt und minutiös zu schildern und durch scheinbar einfachste, aber gerade deshalb schwierigste Mittel: Simplizität, Durchsichtigkeit im einzelnen und Übersichtlichkeit im ganzen, auf eine gewisse künstlerische Höhe zu heben, ja, es dadurch sogar interessant oder wenigstens lesensmöglich zu machen.

Theodor Fontane

Dies ist ein Buch für Anfänger der Mathematik. Es will sich von all seinen Vorgängern und Konkurrenten vor allem dadurch unterscheiden, dass es bewusst und direkt auf die Studierenden zugeht. Ja, unter den vielen Büchern über lineare Algebra, die Sie in der Bibliothek oder einer Buchhandlung finden, eignet sich dieses besonders dafür, Ihr *erstes* Mathematikbuch zu sein. Der Titel hätte auch lauten können „Meine erste Lineare Algebra“.

Dieses Buch soll Ihnen Mut machen, die Mathematik zu meistern, und Sie nicht durch Unverständlichkeit einschüchtern. Beim Schreiben habe ich mich daher von folgenden Ideen leiten lassen:

Keine abgehobene Sprache! Anfänger haben es schwer mit der Mathematik. Sie tun sich besonders schwer mit der mathematischen Sprache. Dieser kalte Formalismus! Diese unbarmherzige Präzision! Diese unendliche Distanz! Diese Schwindel erregende Abstraktheit!

Ja, die Mathematik ist eine Wissenschaft, die auf formalen Schlüssen basiert – das ist ihre Stärke. Die mathematische Sprache ist präzise und dadurch gegen Irrtümer gefeit. Durch Abstraktheit (was nichts anderes als „Vereinfachung“ bedeutet) wird Erkenntnisfortschritt oft erst möglich. Aber die Tatsache bleibt: Die mathematische Sprache lädt Anfänger in der Regel nicht zum Lesen oder zum Mitmachen ein.

Mit diesem Buch versuche ich eine Quadratur des Kreises, nämlich einerseits, wo es nur geht, diese Sprachbarriere abzubauen, andererseits Sie, liebe Leserin, lieber Leser, vom Nutzen der präzisen Sprache der Mathematik zu überzeugen. Insbesondere werden Sie erfahren, dass Präzision nicht unbedingt etwas mit Formalismus – und schon gar nicht mit trockenem Stil zu tun hat.

Der Stil ist für ein Mathematikbuch ganz unüblich: locker, lustig, leicht und unterhaltsam. Und vor allem habe ich versucht, die üblichen k.o.-Schläge wie etwa „wie man leicht sieht“, „trivialerweise folgt“, „man sieht unmittelbar“ zu vermeiden.

Keine unnötig abstrakte Theorie! Was ist das Ziel einer Vorlesung oder eines Buches über lineare Algebra? Ihnen sollen die wichtigsten Grundkonzepte algebraischen Denkens, algebraische Kenntnisse und Fertigkeiten sowie Anwendungen vermittelt werden. (In diesem Buch finden Sie Anwendungen in Geometrie, beim Lösen von Gleichungssystemen und in der Codierungstheorie.)

Wir werden Themen wie Äquivalenzrelationen, Faktorräume, Polynomringe und natürlich die Hauptthemen der linearen Algebra, nämlich Vektorräume, lineare Abbildungen und Diagonalisierbarkeit ausführlich behandeln.

Mir geht es nicht darum, die lineare Algebra mit allen Feinheiten und in voller Allgemeinheit zu präsentieren – in der Hoffnung, dass Kenner anerkennend mit dem Kopf nicken, aber mit dem Effekt, dass die Studierenden den Wälzer wütend an die Wand werfen.

Nicht Rechnen. Denken! Dies ist keine „Lineare Algebra light“, keine Ausgabe „für kleine Hände“. Es kommt mir mindestens so sehr auf begriffliche Klarheit wie auf technische Fertigkeiten an:

- Vektorräume werden „allgemein“ behandelt und nicht von vornherein auf K^n , \mathbb{R}^n (oder gar \mathbb{R}^3) beschränkt. Dadurch wird die Sache einfacher! Denn ein allgemeiner Vektorraum ist ein einfacheres Objekt als ein Vektorraum, bei dem man sich immer noch mit einer festen Basis herumschlagen (oder -ärgern) muss.
- Die berühmten Faktorräume werden ausführlich behandelt – obwohl man Faktorräume in der Linearen Algebra ja zur Not vermeiden könnte. Ich halte aber das Konzept des Faktorraums bzw. der Quotientenstrukturen für so wichtig, dass man das schon im ersten Semester kennen lernen sollte. (Außerdem habe ich das so gut erklärt, dass es jeder verstehen kann!)
- Auch wird in diesem Buch die Theorie der linearen Abbildungen nicht auf Matrizenbolzerei reduziert. Schwierigkeiten werden weder ausgespart noch wird über sie hinweggemogelt.

Viele Übungsaufgaben! Sie finden drei Sorten von Übungsaufgaben. Zunächst ganz einfache Kästchenaufgaben, die in der Regel aus einer „ganz dummen“ Frage bestehen. Diese dienen zur unmittelbaren Selbstkontrolle, ob Sie den Stoff verstanden haben. Lösungen zu diesen Aufgaben finden Sie am Ende des Buches.

Die eigentlichen Übungsaufgaben gehen etwas tiefer – aber auch diese sind (fast) alle leicht zu lösen. Ich habe mich bemüht, keine unnötigen Tricks einzubauen, sondern Ihnen Erfolgserlebnisse zu ermöglichen! Hinweise zur Lösung dieser Aufgaben finden Sie in einem Extra-Kapitel.

Schließlich gibt es „Projekte“; das ist eine Menge zusammengehöriger Übungsaufgaben, mit denen Sie eingeladen werden, ein neues, aber mit dem Stoff des jeweiligen Kapitels eng zusammenhängendes Thema selbständig zu erarbeiten.

Alles in allem über 300 Übungsaufgaben!

Vor kurzem ist übrigens *Lineare Algebra interaktiv* erschienen, eine Übungs-CD zur Linearen Algebra, die noch viel, viel mehr Übungsaufgaben mit Lösungen enthält.

Wenn in einer Vorlesung an einer Universität eine Studentin oder ein Student den Stoff nicht beherrscht und deswegen keinen Schein erhält, so liegt dies – so glauben Lehrende und Lernende übereinstimmend – unzweifelhaft an der Unfähigkeit der Studentin bzw. des Studenten. Ganz anders bei professionellen Kursen im Bereich der Wirtschaft und Industrie. Dort herrschen andere Verhältnisse: Wenn ein Teilnehmer eines Kurses etwas nicht versteht, ist dies eindeutig die Schuld des Dozenten!

Mit diesem Buch stelle ich mich bewusst auf die „professionelle“ Seite: Wenn Sie etwas nicht verstehen, trage ich die Schuld daran. Falls Sie Kritik oder sogar Verbesserungsvorschläge haben, bitte ich Sie, mir ohne Hemmungen zu schreiben.

Einige Hinweise zum Aufbau des Buches: Ich habe mit vielerlei Mitteln versucht, einen lesbaren Text zu verfassen. Einige dieser Mittel sind äußerlich zu erkennen:

- Die Aussagen der Sätze sind farbig unterlegt. Die Sätze sind nicht durchnummeriert, dafür hat (fast) jeder Satz einen Namen; so können Sie ihn über das Stichwortverzeichnis finden.
- Eine Definition erkennt man nicht daran, dass davor „Definition“ steht, sondern daran, dass der zu definierende Begriff **fett** gedruckt ist.
- Das Ende eines jeden Beweises wird durch das Beweisabschlusszeichen angezeigt. Aber auch das Ende eines Satzes, der (meiner Ansicht nach) keines Beweises bedarf, wird so gekennzeichnet:

Obwohl dies ein Buch für Anfänger ist, setze ich gewisse Dinge voraus. So werden etwa Mengenlehre und Beweisprinzipien zwar behandelt – aber nicht sehr ausführlich, damit wir bald zum „eentlichen“ Stoff kommen.

Mein Dank geht an viele, die mich beim Entstehen dieses Buches unterstützt, ermutigt und beraten haben. Zuallererst danke ich den Hörerinnen und Hörern meiner Vorlesung über Lineare Algebra; für sie hatte ich ein Skriptum geschrieben, das die Grundlage für dieses Buch wurde. Und wenn das Skriptum bei den Studierenden nicht so gut angekommen wäre, wäre ich nie auf den Gedanken gekommen, dieses Buch zu schreiben.

Jörg Eisfeld, Udo Heim, Alexander Pott, Ute Rosenbaum und Johannes Ueberberg haben nicht nur das Manuskript mit Akribie und Einfühlung gelesen, sondern mir immer wieder Mut gemacht, das Buch doch so zu schreiben, wie es mir vorschwebte. Frau Susanne Hunsdorfer hat das einfühlsame Schlussbild gemalt. Allen gilt mein herzlicher Dank.

Als ich mich schon in der Hoffnung wiegte, das Buch sei fertig, habe ich es auf Anregung des Verlags nochmals einer Gruppe junger Studierender zum Lesen gegeben. Und so wurde eine bislang unentdeckte Schicht von Fehlern und Verbesserungsmöglichkeiten ans Licht befördert. Schande über mein Haupt und Tausend Dank an die Studierenden! (Und das heißt immerhin mehr als ein Dank pro entdecktem Fehler.)

Studierende, die es mit der Mathematik wagen wollen, brauchen Mut. Auch ein Autor braucht Mut – jedes neue Buch ist ein neues Wagnis. Aber auch ein Verlag braucht Mut für ein solches Buch. Daher danke ich dem Verlag Vieweg, und ganz besonders Frau Döbert und Frau Schmickler-Hirzebruch sehr, dass sie dieses Buch wagten.

Vorwort zur 8. Auflage

Schon bald nach Erscheinen dieses Buches erlebte ich eine Überraschung: Ich erhielt Fanpost. Nicht gerade wäschekörbeweise, aber immerhin ein paar Briefe pro Semester. Darin schilderten Studierende, dass ihnen dieses Buch geholfen habe, in der ersten Krise des Mathematikstudiums nicht zu verzweifeln, sondern durchzuhalten. Etwas Schöneres kann einem Autor nicht passieren. Denn genau für sie hatte ich das Buch geschrieben. Für Studierende am Anfang des Studiums; für solche, die an der rigorosen mathematischen Sprache zu scheitern drohen; für solche, die in Gefahr sind, unter der Unbarmherzigkeit mathematischer Beweisführung zusammenzubrechen. Ich freue mich sehr darüber, dass dieses Konzept eine große Zahl von Studierenden darin bestärkt hat, bei der Mathematik zu bleiben und ihr Studium erfolgreich zu beenden.

Als die erste Auflage dieser Linearen Algebra erschien, sprach noch niemand von Bachelor und Master. Aber im Grunde war dieses Buch mit seiner Konzentration auf das Wesentliche, mit den zahlreichen Übungsaufgaben und Lernhilfen schon ein Vorgriff auf die Bachelor-Ausbildung. Für die Neuauflage habe ich zahlreiche „kleine“ Änderungen realisiert. So wird dieses Buch auch in Zukunft den Bachelor- und Lehramtsstudierenden in der ersten Phase des Studiums nützlich, hilfreich und anregend sein. Dies hoffe ich jedenfalls.

Kurz vor Fertigstellung dieser Neuausgabe hatte meine Lektorin, Frau Ulrike Schmickler-Hirzebruch, eine wunderbare Idee. Sie hatte die Ausstellung „Mathe macht lustig!“ mit Mathe-Karikaturen im Mathematikum in Gießen gesehen und schlug vor, jedes Kapitel mit einer Karikatur aus dieser Ausstellung abzuschließen. In der Tat eröffnen diese Bilder ganz neue Blicke auf die Mathematik. Die Karikaturisten haben freundlicherweise zugesagt, und so geht mein Dank an Martin Zak, Til Mette, Erich Rauschenbach, Miriam Wurster, Phil Hube, NEL, Leonard Riegel, F.W. Bernstein, Rudi Hurzlmeier und, last but not least, an Ulrike Schmickler-Hirzebruch für ihre mittlerweile jahrzehntelange hervorragende Unterstützung.

Gießen, im November 2013

Albrecht Beutelspacher

Inhaltsverzeichnis

1	Was wir wissen müssen, bevor wir anfangen können	1
1.1	Mengen	1
1.2	Äquivalenzrelationen	4
1.3	Abbildungen	7
1.4	Wann haben zwei Mengen gleich viele Elemente?	13
1.5	Die Σ -Notation	18
1.6	Beweisprinzipien	20
1.7	Verständnisfragen, Übungen und Tipps	22
2	Körper	29
2.1	Die Definition	29
2.1.1	Gesetze der Addition	29
2.1.2	Gesetze der Multiplikation	30
2.1.3	Distributivgesetz	30
2.2	Beispiele von Körpern	32
2.2.1	Der Körper der komplexen Zahlen	33
2.2.2	Der Quaternionenschiefkörper	36
2.2.3	Einige endliche Körper	40
2.2.4	Konstruktion eines Körpers mit vier Elementen	44
2.3	Automorphismen von Körpern	46
2.3.1	Die Definitionen	47
2.3.2	Der Körper der rationalen Zahlen	47
2.3.3	Der Körper der reellen Zahlen	50
2.3.4	Konjugiert-komplexe Zahlen	51
2.4	Verständnisfragen, Übungen und Tipps	52
3	Vektorräume	59
3.1	Die Definition	59
3.2	Beispiele von Vektorräumen	61
3.2.1	Vektorräume mit Hilfe von Geometrie	61
3.2.2	Der Vektorraum K^n	62

3.2.3	Der Vektorraum aller $m \times n$ -Matrizen	63
3.2.4	Der Vektorraum aller unendlichen Folgen	64
3.2.5	Ein Vektorraum unendlicher Folgen	64
3.2.6	Vektorräume von Funktionen	64
3.2.7	Lösungen eines Gleichungssystems	65
3.2.8	Teilmengen einer Menge	65
3.2.9	Körper als Vektorräume	65
3.3	Elementare Theorie der Vektorräume	66
3.3.1	Der Begriff der Basis	67
3.3.2	Der Steinitzsche Austauschsatz	75
3.3.3	Der Dimensionssatz	83
3.3.4	Faktorräume	85
3.4	Zur Geschichte der linearen Algebra	92
3.5	Verständnisfragen, Übungen und Tipps	94
4	Anwendungen von Vektorräumen	105
4.1	Lineare Gleichungssysteme	105
4.1.1	Begriffe und Fragen	105
4.1.2	Exkurs über Matrizen	106
4.1.3	Lösbarkeit von linearen Gleichungssystemen	111
4.1.4	Der Gaußsche Algorithmus	116
4.2	Affine Geometrie	122
4.2.1	Affine Räume	123
4.2.2	Unterräume	126
4.3	Codierungstheorie	129
4.3.1	Grundlegende Begriffe	129
4.3.2	Lineare Codes	133
4.4	Verständnisfragen, Übungen und Tipps	139
5	Lineare Abbildungen	147
5.1	Definitionen und grundlegende Eigenschaften	147
5.2	Darstellung von linearen Abbildungen durch Matrizen	154
5.3	Der Homomorphiesatz	162
5.4	Der Dualraum	166
5.5	Verständnisfragen, Übungen und Tipps	171
6	Polynomringe	177
6.1	Ringe	177
6.1.1	Gesetze der Addition	177
6.1.2	Gesetz der Multiplikation	178
6.1.3	Distributivgesetze	178

6.2	Was ist eigentlich x ?	179
6.3	Polynomdivision	187
6.4	Ideale von $K[x]$	192
6.5	Verständnisfragen, Übungen und Tipps	195
7	Determinanten	203
7.1	Die Determinantenfunktion	203
7.2	Permutationen	207
7.3	Gerade und ungerade Permutationen	211
7.4	Die Leibnizsche Determinantenformel	218
7.5	Wie berechnet man eine Determinante?	222
7.6	Der Multiplikationssatz	233
7.7	Verständnisfragen, Übungen und Tipps	236
8	Diagonalisierbarkeit	241
8.1	Einführung	241
8.2	Eigenvektoren und Eigenwerte	243
8.3	Das charakteristische Polynom	249
8.4	Das Minimalpolynom	256
8.5	Verständnisfragen, Übungen und Tipps	265
9	Elementarste Gruppentheorie	271
9.1	Beispiele von Gruppen	271
9.1.1	Gruppen in bekannten Strukturen	273
9.1.2	Gruppen aus bekannten Objekten	274
9.1.3	Gruppen aus Permutationen	276
9.2	Einfache Strukturaussagen für Gruppen	278
9.2.1	Untergruppen	278
9.2.2	Zyklische Gruppen	282
9.2.3	Der Homomorphiesatz	285
9.3	Verständnisfragen, Übungen und Tipps	289
10	Skalarprodukte	295
10.1	Ein Beispiel	295
10.2	Bilinearformen	297
10.3	Skalarprodukte	307
10.4	Orthogonale Abbildungen	315
10.5	... und eine zweite symmetrische Bilinearform?	324
10.6	Verständnisfragen, Übungen und Tipps	329

11	Lösungen	337
11.1	Lösungsvektoren der \square -Aufgaben	337
11.2	Tipps zur Lösung der Übungsaufgaben	339
Literatur		359
Sachverzeichnis		361

In diesem einleitenden Kapitel führen wir die Bezeichnungen und Begriffe ein, mit denen wir im Folgenden routinemäßig umgehen werden.

1.1 Mengen

Seit Georg Cantor (1845–1918) ist die gesamte Mathematik auf dem Mengenbegriff aufgebaut. Was eine Menge „ist“ (das heißt, wie man axiomatisch mit Mengen umgehen kann), werden wir hier nicht diskutieren. Als einführende Lektüre empfehle ich Ihnen den „Klassiker“ *Naive Mengenlehre* von P. R. Halmos [Hal] oder das neuere Buch *Mengenlehre für den Mathematiker* [FrPr].

Mengen bestehen aus **Elementen**. Ist ein Objekt x Element der Menge X , so schreiben wir dafür $x \in X$; ist x kein Element von X , so wird das durch $x \notin X$ bezeichnet.

Eine Menge Y heißt **Teilmenge** (oder **Untermenge**) der Menge X , falls jedes Element von Y auch ein Element von X ist; man schreibt dafür $Y \subseteq X$. Eine Teilmenge Y von X heißt eine **echte** Teilmenge, falls $Y \neq X$ ist.

Mengen können prinzipiell auf zwei Weisen beschrieben werden, durch Aufzählung ihrer Elemente oder durch ihre Eigenschaften. Ein Beispiel für die *Definition einer Menge durch Auflistung ihrer Elemente* ist

$$M := \{1, 2, 3, 4, 5, 6\}.$$

Damit kann man auch unendliche Mengen definieren:

$$\mathbb{N} := \{0, 1, 2, 3, \dots\}$$

ist die Menge der **natürlichen Zahlen**, und

$$\mathbf{Z} := \{0, 1, -1, 2, -2, 3, -3, \dots\}$$

ist die Menge der **ganzen Zahlen**.

Mit **Q** bezeichnen wir die Menge der **rationalen** und mit **R** die Menge der **reellen Zahlen**.

Eine Menge spielt eine ausgezeichnete Rolle: die **leere Menge**, die mit \emptyset bezeichnet wird. Dies ist die Menge, die kein Element enthält.

Die *Definition einer Menge durch eine Eigenschaft* geht von einer (bereits definierten) Menge X (oder mehreren Mengen) aus und definiert damit eine neue Menge gemäß folgendem Muster:

$$Y := \{x \in X | E(x)\},$$

wobei E eine Eigenschaft des Elements x ist. Zum Beispiel wird durch

$$2\mathbf{Z} := \{z \in \mathbf{Z} | z \text{ ist gerade}\}$$

die Menge der geraden ganzen Zahlen definiert. Hier ist die Eigenschaft $E(z)$ erklärt als „ z ist gerade“.

Die **Differenz** $X \setminus Y$ der Mengen X und Y (manchmal auch als $X - Y$ bezeichnet) ist definiert als die Menge der Elemente von X , die nicht in Y liegen:

$$X \setminus Y := \{x \in X | x \notin Y\}.$$

Wir setzen bei der Bildung der Differenz $X \setminus Y$ nicht voraus, dass Y eine Teilmenge von X ist; wenn das aber so ist, so heißt $X \setminus Y$ das **Komplement von Y in X** .

Wenn zum *Beispiel* X die Menge aller Studierenden der Mathematik und Y die Menge derjenigen Mathematikstudenten ist, die regelmäßig die Vorlesungen besuchen, dann ist $X \setminus Y = \dots$?

Die folgenden Mengenbildungen sind äußerst wichtig:

$$X \cap Y := \{x \in X | x \in Y\}$$

ist der **Durchschnitt** von X und Y ; die Mengen X und Y heißen **disjunkt**, falls $X \cap Y = \emptyset$ ist. Die **Vereinigung** von X und Y ist definiert als die Menge $X \cup Y$, die aus all den Elementen besteht, die in X oder in Y liegen:

$$X \cup Y := \{z | z \in X \text{ oder } z \in Y\}.$$

(Wenn wir „oder“ sagen, lassen wir stets zu, dass beide Möglichkeiten zutreffen.)

Schließlich besteht das **kartesische Produkt** der nichtleeren Mengen X_1, \dots, X_n aus den **Folgen der Länge n (oder n -Tupeln)** (x_1, \dots, x_n) mit $x_1 \in X_1, \dots, x_n \in X_n$:

$$X_1 \times \dots \times X_n := \{(x_1, \dots, x_n) | x_i \in X_i\}.$$

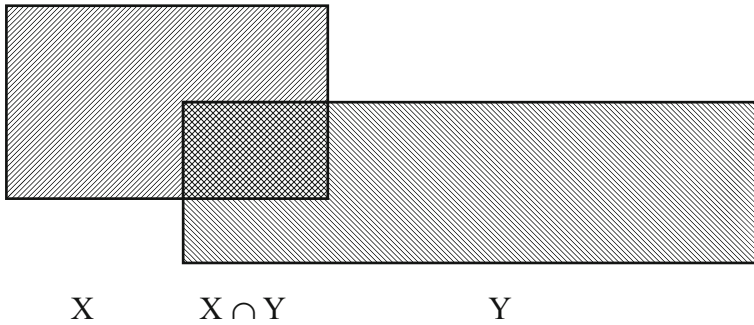


Abb. 1.1 Die Summenformel

(Dies wird nach René Descartes (1596–1650) genannt, der erkannt hat, dass man die Punkte der Ebene durch Zahlenpaare darstellen kann.)

Wenn wir zum *Beispiel* mit XX die Menge aller Frauen und mit XY die Menge aller Männer bezeichnen, so ist $XX \times XY$ die Menge aller getrenntgeschlechtlichen Paare.

Bemerkung Bei der Definition neuer Mengen verwendet man häufig nicht das Gleichheitszeichen, sondern das Zeichen $:=$. Der Doppelpunkt steht dabei auf derjenigen Seite des Gleichheitszeichens, auf der das neu definierte Objekt steht. So kann man zum Beispiel schreiben: Sei

$$n\mathbf{Z} := \{nz \mid z \in \mathbf{Z}\}$$

die Menge der durch n teilbaren ganzen Zahlen.

Um uns an die Konzepte zu gewöhnen, betrachten wir endliche Mengen und rechnen die Anzahl der neu gebildeten Mengen aus.

Eine Menge heißt **endlich**, falls sie nur endlich viele Elemente hat (also etwa 5 Elemente oder 1000 oder $2^{64} - 1$). Die Anzahl der Elemente einer Menge X bezeichnen wir mit $|X|$.

Seien für den Rest dieses Abschnitts X , Y , sowie X_1, \dots, X_n endliche Mengen.

Wenn Y eine Teilmenge von X ist, so gilt

$$|X \setminus Y| = |X| - |Y|;$$

denn $X \setminus Y$ umfasst genau die Elemente von X , die nicht in Y liegen.

Die Anzahlen der Elemente von $X \cap Y$ und $X \cup Y$ sind durch folgende **Summenformel** miteinander gekoppelt:

$$|X \cup Y| = |X| + |Y| - |X \cap Y|.$$

In Worten: Die Anzahl der Elemente in der Vereinigung zweier Mengen ist gleich der Summe der Anzahlen der einzelnen Mengen minus der Anzahl der Elemente im Durchschnitt (siehe Abb. 1.1).

Der *Beweis* der Summenformel ist einfach:

Da in der Summe $|X| + |Y|$ die Elemente von $X \setminus Y$ und die Elemente $Y \setminus X$ genau einmal, die Elemente von $X \cap Y$ aber genau zweimal gezählt werden, wird in der Summe $|X| + |Y| - |X \cap Y|$ jedes Element von $X \cup Y$ genau einmal erfasst. Also ist

$$|X| + |Y| - |X \cap Y| = |X \cup Y|.$$

Schließlich berechnen wir die Anzahl der Elemente des kartesischen Produkts mit Hilfe der folgenden **Produktformel**:

$$|X_1 \times X_2 \times \dots \times X_n| = |X_1| \cdot |X_2| \cdot \dots \cdot |X_n|.$$

Für den *Beweis* müssen wir die Anzahl der Elemente (x_1, x_2, \dots, x_n) mit $x_i \in X_i$ berechnen: Für x_1 gibt es genau $|X_1|$ Möglichkeiten. Für jede dieser $|X_1|$ Möglichkeiten gibt es genau $|X_2|$ Möglichkeiten, x_2 zu wählen. Für jede der $|X_1| \cdot |X_2|$ Möglichkeiten für x_1 und x_2 gibt es $|X_3|$ Möglichkeiten, x_3 zu wählen und so weiter. Insgesamt gibt es daher genau $|X_1| \cdot |X_2| \cdot \dots \cdot |X_n|$ Möglichkeiten, eine Folge (x_1, x_2, \dots, x_n) zu wählen. Es gilt also die oben angegebene Produktformel. \square

Es gibt viele mehr oder weniger eindruckliche Veranschaulichungen dieses Sachverhalts: Professor MacMath ist im Besitz von drei Jacketts, sechs Hemden und zwei Hosen; er kleidet sich an, indem er jeweils ein Kleidungsstück von einer Sorte zufällig wählt. Auf wie viele Weisen kann er sich kleiden? Die Antwort lautet: Auf $3 \cdot 6 \cdot 2 = 36$ Weisen. (In Übungsaufgabe 2 finden Sie ein realistischeres Beispiel.)

1.2 Äquivalenzrelationen

In der Umgangssprache bezeichnen wir zwei Vorgänge als äquivalent, wenn sie gleichwertig, ziemlich ähnlich, aber im Allgemeinen nicht gleich sind. Die Mathematik lehnt sich an diesen Sprachgebrauch an und präzisiert ihn in charakteristischer Weise.

Durch Relationen werden Beziehungen, die Elemente einer Menge untereinander haben können, beschrieben. Formal definieren wir eine **Relation** auf der Menge X als eine Teilmenge ρ von $X \times X$. Gilt $(x, y) \in \rho$, so sagen wir „ x steht in der Relation ρ mit y “ und schreiben dafür meist $x \rho y$. Da wir in der Regel jeweils nur eine Relation betrachten, brauchen wir nicht verschiedene Relationen durch verschiedene griechische Buchstaben zu unterscheiden; wir bezeichnen eine Relation daher meist mit \sim und schreiben $x \sim y$ für $x \rho y$ bzw. $(x, y) \in \rho$.

Relationen, die aus alltäglichen Situationen bekannt sind, sind zum Beispiel

$$x \sim y \Leftrightarrow x \text{ ist befreundet mit/verwandt mit/per Du mit/verliebt in } y$$

Die wichtigsten Eigenschaften mathematischer Relationen sind Reflexivität, Symmetrie und Transitivität. Für eine Relation \sim auf der Menge X definieren wir

- Die Relation \sim heißt **reflexiv**, falls gilt: $x \sim x$ für alle $x \in X$,
- die Relation \sim heißt **symmetrisch**, falls gilt: $(x \sim y \Leftrightarrow y \sim x)$ für alle $x, y \in X$,
- die Relation \sim heißt **transitiv**, falls gilt: $(x \sim y, y \sim z \Rightarrow x \sim z)$ für alle $x, y, z \in X$.

Die wichtigsten Relationen sind Äquivalenzrelationen. Eine Relation heißt eine **Äquivalenzrelation**, falls sie reflexiv, symmetrisch und transitiv ist.

Zunächst einige *Beispiele* von Äquivalenzrelationen:

- Die Relation \sim definiert auf der Menge der Bürger von Deutschland durch

$$x \sim y \Leftrightarrow x \text{ und } y \text{ haben ihren ersten Wohnsitz in derselben Stadt}$$

ist eine Äquivalenzrelation.

- Eine andere Äquivalenzrelation \sim auf der Menge der Einwohner der Bundesrepublik Deutschland ist definiert durch

$$x \sim y \Leftrightarrow x \text{ und } y \text{ werden durch denselben direkt gewählten Bundestagsabgeordneten vertreten.}$$

- Die **Gleichheitsrelation** = ist eine Äquivalenzrelation.
- Die Relation \parallel definiert auf den Geraden der Ebene durch

$$g \parallel h \Leftrightarrow g \text{ und } h \text{ sind parallel}$$

ist eine Äquivalenzrelation.

- Für jede natürliche Zahl n ist die Relation \sim_n auf der Menge \mathbf{Z} der ganzen Zahlen, die wie folgt definiert ist:

$$x \sim_n y \Leftrightarrow y - x \text{ ist ein ganzzahliges Vielfaches von } n$$

eine Äquivalenzrelation.

Warum sind Äquivalenzrelationen wichtig? In vielen Fällen bilden Äquivalenzrelationen gute Beschreibungen von Situationen. Sie erlauben es nämlich, ohne Informationsverlust eine „größere“ Struktur, jedenfalls eine Struktur mit weniger Elementen zu betrachten. Im ersten Beispiel braucht man nicht die unübersehbare Fülle aller 80 Millionen Deutschen zu betrachten, sondern es genügt, den Blick auf die nur etwa 16.000 Gemeinden zu werfen.

Der mathematische Begriff, der hier die zweite Hauptrolle spielt, ist der der „Äquivalenzklasse“:

Sei \sim eine Äquivalenzrelation auf der Menge X . Für ein $x \in X$ sei $A(x)$ die Menge aller Elemente von X , die mit x in der Relation \sim stehen; man nennt $A(x)$ die **Äquivalenzklasse** von x . Wir drücken diese Definition nochmals in der Formelsprache aus:

$$A(x) := \{y \in X \mid y \sim x\}.$$

Die zentrale und wichtigste Tatsache über Äquivalenzrelationen ist die folgende:

Satz über Äquivalenzklassen

Äquivalente Elemente haben dieselbe Äquivalenzklasse; das heißt:

$$x \sim y \Rightarrow A(x) = A(y) .$$

Diesen Satz *beweisen* wir wie folgt: Sei $x \sim y$. Dann ist zu zeigen $A(x) = A(y)$. Dies beweisen wir dadurch, dass wir sowohl $A(x) \subseteq A(y)$ als auch $A(y) \subseteq A(x)$ zeigen.

„ $A(x) \subseteq A(y)$ “: Sei $z \in A(x)$. Das bedeutet $z \sim x$. Wegen $x \sim y$ und der Transitivität von \sim folgt daraus auch $z \sim y$. Dies bedeutet aber $z \in A(y)$. Somit ist jedes Element von $A(x)$ in $A(y)$, also gilt $A(x) \subseteq A(y)$.

„ $A(y) \subseteq A(x)$ “: Wir haben zwei Möglichkeiten, dies zu beweisen.

Sture Möglichkeit: Man geht analog wie im ersten Fall vor. Dies sollen Sie in Übungsaufgabe 5 tun.

Intelligente Möglichkeit: Wegen der Symmetrie folgt aus $x \sim y$ auch $y \sim x$. Also folgt aus dem ersten Fall (wenn man die Buchstaben x und y vertauscht) $A(y) \subseteq A(x)$. \square

Man kann diese fundamentale Tatsache auch wie folgt ausdrücken:

Satz über Gleichheit von Äquivalenzklassen

Zwei Äquivalenzklassen sind gleich oder disjunkt.

Beweis Haben die Äquivalenzklassen $A(x)$ und $A(y)$ auch nur ein Element z gemeinsam, so ergibt sich aus der gerade bewiesenen Tatsache sowohl $A(x) = A(z)$ (da $x \sim z$) als auch $A(y) = A(z)$ (da $y \sim z$). Das bedeutet: Sind die Äquivalenzklassen $A(x)$ und $A(y)$ nicht disjunkt, so sind sie gleich. \square

Den obigen Satz kann man auch wie folgt formulieren:

Jedes Element von X liegt in genau einer Äquivalenzklasse. \square

Dies drückt man oft mit Hilfe des Begriffs einer Partition aus. Eine **Partition** einer Menge X ist eine Menge $\pi = \{Y_1, Y_2, \dots\}$ von nichtleeren Teilmengen von X , die folgende Eigenschaften haben:

- $Y_i \cap Y_j = \emptyset$ für $i \neq j$,
- $\cup Y_i = X$, wobei $\cup Y_i$ die Vereinigung der Mengen Y_i bezeichnet.

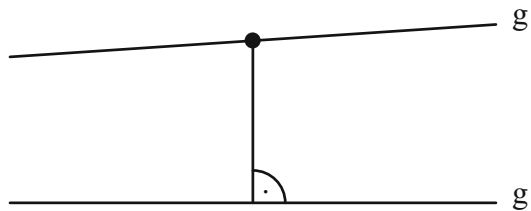
Damit kann man sagen:

Die Menge der Äquivalenzklassen bildet eine Partition von X .

Beweis Wegen der Reflexivität liegt jedes Element x von X in mindestens einer Äquivalenzklasse, nämlich in $A(x)$; die Disjunktheit der Äquivalenzklassen folgt aus der ersten Umformulierung des Satzes über die Gleichheit von Äquivalenzklassen. \square

In der Mathematik definiert man oft Eigenschaften von Äquivalenzklassen über Eigenschaften einzelner Elemente (**Repräsentanten**) der Äquivalenzklassen. Dann muss man sich überzeugen, dass diese Eigenschaft unabhängig von den einzelnen Elementen definiert ist. Dies bezeichnet man dann als **Wohldefiniertheit**. Das wird später noch eine wichtige Rolle spielen. An dieser Stelle erwähnen wir nur ein Beispiel, um uns die Problematik klarzumachen.

Was ist der *Abstand zweier Geraden* g, g' der gewöhnlichen euklidischen Ebene? Ganz einfach: Man wählt einen Punkt auf g und bestimmt den Abstand dieses Punktes zu g' .



Gut, aber was passiert, wenn man einen anderen Punkt von g wählt? Genau das ist das Problem! Es geht darum, ob der Abstand von g und g' wohldefiniert ist. Dies ist genau dann der Fall, wenn *jeder* Punkt von g den gleichen Abstand von g' hat.

In Übungsaufgabe 6 soll folgendes gezeigt werden: Wenn man überprüft hat, dass drei Punkte von g den gleichen Abstand von g' haben, so hat jeder Punkt von g den gleichen Abstand von g' .

1.3 Abbildungen

Der Begriff der Abbildung ist zwar historisch relativ jung, hat sich aber als außerordentlich nützliches und heute nicht mehr wegzudenkendes Werkzeug der modernen Mathematik erwiesen.

Seien X und Y Mengen. Eine **Abbildung von X nach Y** (oder „von X in Y “) ist eine Vorschrift f , die jedem $x \in X$ genau ein Element aus Y zuordnet; dieses Element wird mit $f(x)$ bezeichnet. Wenn f eine Abbildung von X nach Y ist, so schreibt man dafür auch $f: X \rightarrow Y$. Die Elemente von X heißen die **Urbilder** der Abbildung f , diejenigen Elemente y aus Y , für die es ein $x \in X$ gibt mit $f(x) = y$, heißen die **Bilder** von f . Man nennt X den **Definitions-** und Y den **Bildbereich** von f (siehe Abb. 1.2).

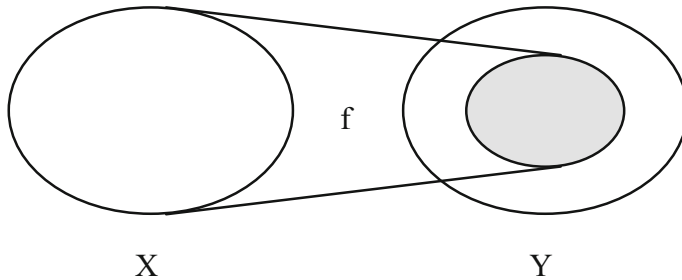


Abb. 1.2 Eine Abbildung

Bemerkung Wenn Ihnen diese Definition einer Abbildung nicht genau genug ist (in der Tat können Sie fragen, was der schwammige Begriff „Vorschrift“ bedeuten soll), so kann Ihnen vielleicht mit der folgenden alternativen Definition geholfen werden: Eine **Abbildung** von X nach Y ist eine Relation $f \subseteq X \times Y$ mit der Eigenschaft, dass es zu jedem $x \in X$ genau ein $y \in Y$ gibt mit $(x, y) \in f$. Dieses eindeutig bestimmte y bezeichnet man mit $f(x)$.

Machen Sie sich klar, dass wir nicht zwei verschiedene Dinge definiert haben, sondern dass es sich im Grunde um das gleiche Konzept handelt, das nur auf zwei verschiedenen Sprachebenen dargestellt wurde.

Wie kann man konkret eine Abbildung beschreiben? – Ganz einfach dadurch, dass man für jedes Element x von X das Bild $f(x)$ angeben muss. Dies kann auf vielfältige Weise geschehen; an den folgenden *Beispielen* wird das klar werden.

1. Sei $f: \mathbf{Z} \rightarrow \mathbf{N}$ die durch $f(z) := z^2$ definierte Abbildung.
2. Sei f die Abbildung von \mathbf{R} in sich, die jeder reellen Zahl ihren Absolutbetrag zuordnet.
3. Sei f die Abbildung der Menge $\{1, 2, 3\}$ in sich, die folgendermaßen erklärt ist

$$1 \rightarrow 1, 2 \rightarrow 3, 3 \rightarrow 2.$$

4. Die Multiplikation reeller Zahlen ist die Abbildung

$$\mathbf{R} \times \mathbf{R} \rightarrow \mathbf{R}, (r, s) \rightarrow r \cdot s.$$

Abbildungen gibt es in Hülle und Fülle. Eine verdient einen besonderen Namen, auch schon deshalb, weil sie sonst nicht auffällt: Ist X eine Menge, so nennen wir die Abbildung $\text{id} = \text{id}_X$, von X in sich, die durch

$$\text{id}(x) := x \quad \text{für alle } x \in X$$

definiert ist, die **Identität** (oder **identische Abbildung**) auf X .

Die im Zusammenhang mit Abbildungen am häufigsten gebrauchten Wörter sind die Wörter „injektiv“, „surjektiv“ und „bijektiv“. Es hilft nichts: Sie müssen diese Begriffe wie Vokabeln einer fremden Sprache stur pauken. Dies lohnt sich, denn diese Begriffe begegnen Ihnen während Ihres Studiums garantiert tausendfach.

Sei f eine Abbildung von X nach Y .

- (a) Man nennt f **injektiv**, wenn keine zwei verschiedenen Elemente von X auf dasselbe Element von Y abgebildet werden.
- (b) Man nennt f **surjektiv**, wenn es zu jedem Element $y \in Y$ ein $x \in X$ gibt mit $f(x) = y$. (Wenn f eine surjektive Abbildung von X nach Y ist, so darf man auch sagen, dass f eine Abbildung von X **auf** Y ist.)
- (c) Eine Abbildung heißt **bijektiv**, wenn sie injektiv und surjektiv ist.

Wir betrachten einige *Beispiele*.

1. Die Abbildung $f: \mathbf{Z} \rightarrow \mathbf{N}$ mit $f(z) := z^2$ ist weder injektiv (denn $f(-z) = (-z)^2 = z^2 = f(z)$ für alle $z \in \mathbf{Z}$) noch surjektiv (denn zum Beispiel hat die Zahl 3 kein Urbild).
2. Die Abbildung $f: \mathbf{R} \rightarrow \mathbf{R}^+ (:= \{r \in \mathbf{R} \mid r \geq 0\})$, die jeder reellen Zahl ihren Absolutbetrag zuordnet, ist nicht injektiv, aber surjektiv.
3. Die Abbildung f der Menge $\{1, 2, 3\}$ in sich mit

$$f(1) = 1, f(2) = 3, f(3) = 2$$

ist sowohl injektiv als auch surjektiv, also bijektiv.

Ist $f: X \rightarrow Y$ eine Abbildung, so definieren wir

$$\text{Bild}(f) := \{y \in Y \mid \text{es existiert ein } x \in X \text{ mit } f(x) = y\}.$$

In $\text{Bild}(f)$ sind also genau die Elemente von Y enthalten, die wirklich als Bild eines Elementes von X auftreten.

Offenbar kann man f auch als Abbildung von X nach $\text{Bild}(f)$ auffassen; diese Abbildung ist in jedem Fall surjektiv:

Man kann also jede Abbildung f durch Einschränkung des Bildbereichs Y auf $\text{Bild}(f)$ surjektiv „machen“; entsprechend kann man auch jede Abbildung durch Einschränkung des Definitionsbereichs injektiv machen: Man muss so viele Elemente aus X streichen, dass für jedes Element $y \in \text{Bild}(f)$ nur ein $x \in X$ mit $f(x) = y$ übrig bleibt.

Übrigens kann man auch jede surjektive Abbildung durch Vergrößerung des Bildbereichs nicht-surjektiv machen. Kann man auch jede injektive Abbildung nicht-injektiv machen?

Für die Bedeutung der bijektiven Abbildungen ist die Tatsache verantwortlich, dass die bijektiven Abbildungen genau die invertierbaren Abbildungen sind. Bevor wir dies erläutern können, brauchen wir noch eine Definition.

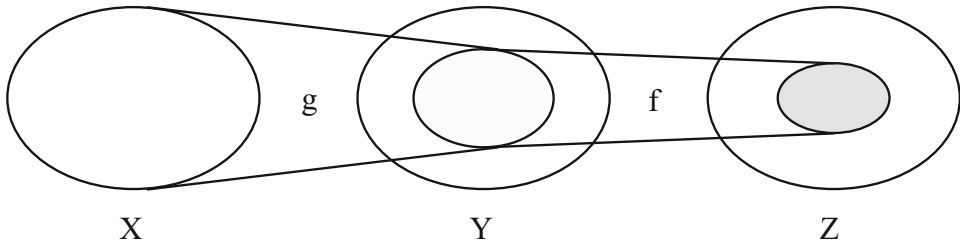


Abb. 1.3 Hintereinanderausführung von Abbildungen

Sei g eine Abbildung von X nach Y und f eine Abbildung von Y nach Z . Dann ist $f \circ g$ die Abbildung von X nach Z , die wie folgt definiert ist:

$$f \circ g(x) := f(g(x)) .$$

Um das Bild von x unter $f \circ g$ zu erhalten, wendet man also zuerst g auf x an; damit erhält man ein Element $g(x)$ von Y , und auf dieses wendet man dann f an, um schließlich $f(g(x))$ zu erhalten (siehe Abb. 1.3).

Man nennt $f \circ g$ die Abbildung, die man aus f und g durch **Hintereinanderausführung (Komposition)** von Abbildungen erhält. Man spricht kurz auch vom **Produkt** der Abbildungen f und g .

Achtung $f \circ g$ wird *von rechts nach links* ausgeführt; um dies nicht zu vergessen, benutzt man Sprechweisen wie etwa „erst g dann f “, „ f nach g “. Dies ist seltsam, aber nicht aufregend; Sie müssen sich einfach daran gewöhnen.

Sie hätten's natürlich auch anders haben können – wenn wir statt der Schreibweise $f(x)$, bei der f links von x steht, Formulierungen wie xf oder x^f benutzen würden (bei denen das Symbol für die Abbildung rechts vom Argument steht). Obwohl das unter Algebraikern als besonders fein gilt, habe ich von dieser (für Sie) ungewohnten (und in der Analysis völlig unüblichen) Schreibweise Abstand genommen. Die Komplikation bei der Hintereinanderausführung von Abbildungen ist der Preis, den wir dafür bezahlen müssen.

Es ist klar, dass man $f \circ g$ nur dann bilden kann, wenn der Bildbereich von g im Definitionsbereich von f enthalten ist. Insbesondere kann man bijektive Abbildungen einer Menge in sich stets hintereinander ausführen und erhält dabei wieder eine bijektive Abbildung.

Nun zu der angekündigten Tatsache, dass jede bijektive Abbildung invertierbar ist:

Satz über die Invertierbarkeit bijektiver Abbildungen

Sei $f: X \rightarrow Y$ eine bijektive Abbildung. Dann gibt es eine Abbildung $f': Y \rightarrow X$ derart, dass gilt

$$f \circ f' = \text{id}_Y \text{ und } f' \circ f = \text{id}_X .$$

Das bedeutet, dass für alle $x \in X$ und für alle $y \in Y$ gilt

$$f \circ f'(y) = y \text{ und } f' \circ f(x) = x .$$

Mit anderen Worten: Jede bijektive Abbildung ist **umkehrbar** (oder **invertierbar**).

Diese Tatsache müssen wir (wie alle Tatsachen in der Mathematik) *beweisen*. Was müssen wir dazu tun? In der zu beweisenden Aussage wird behauptet, dass es etwas gibt, nämlich die Abbildung f' ; also besteht unsere Aufgabe darin, ein solches f' zu finden.

Dieses f' soll eine Abbildung von Y nach X sein. Unsere Aufgabe kann also noch genauer dadurch beschrieben werden, dass wir für jedes $y \in Y$ sagen müssen, was $f'(y)$ sein soll.

Wir erinnern uns daran, dass f' eine Inverse von f werden soll, also die Wirkung von f rückgängig machen muss. Das bedeutet, dass wir ein $x \in X$ finden müssen, so dass $f(x) = y$ ist.

Gibt es ein solches x ? Ja! Denn f ist ja eine bijektive, insbesondere also eine surjektive Abbildung. Daher gibt es zu jedem $y \in Y$ ein $x \in X$ mit $f(x) = y$. Es liegt an dieser Stelle nahe, f' durch $f'(y) := x$ zu definieren.

Wir müssen aber noch einmal aufpassen: f' soll eine *Abbildung* sein; das bedeutet, dass jedem Element von Y *genau ein* Element von X zugeordnet werden muss. Die Frage ist also: Gibt es genau ein $x \in X$ mit $f(x) = y$? Auch hierauf ist die Antwort ja! Dies liegt daran, dass f injektiv ist; dies bedeutet nämlich gerade, dass es keine verschiedenen Elemente von X gibt, die auf dasselbe y abgebildet werden.

Also können wir in der Tat die Abbildung f' von Y nach X definieren durch

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

wobei x das eindeutig bestimmte Element von X ist mit $f(x) = y$. Dann gilt nach Definition

$$f \circ f'(y) = f(x) = y \text{ und } f' \circ f(x) = f'(y) = x$$

für alle $y \in Y$ und alle $x \in X$.

Damit haben wir alles gezeigt. □

Wir haben in diesem Buch schon einige Beweise geschafft – aber das ist der erste, auf den man (ich schließe mich dabei ein!) nicht so ohne weiteres gekommen wäre. Herzlichen Glückwunsch! Schauen Sie sich nach einer Erholungspause den Beweis nochmals an und überzeugen Sie sich, wie sich durch unsere immer genaueren Fragen die zu beweisenden Sachverhalte fast automatisch ergeben haben. Dies ist oft (aber nicht immer) der Fall.

In dem obigen Satz haben wir gezeigt, dass jede bijektive Abbildung eine inverse Abbildung hat. Wenn wir diese Aussage nochmals genau anschauen, so sehen wir, dass sie lautet „... hat *mindestens eine* Inverse“. Diese Formulierung löst sofort die Frage aus: Wie viele inverse Abbildungen hat eine bijektive Abbildung? Zwei? Tausend? Unendlich viele? Alles falsch: Genau eine! Dies drückt folgender Satz aus:

Eindeutigkeit der inversen Abbildung

Sei $f: X \rightarrow Y$ eine bijektive Abbildung. Dann gibt es genau eine Abbildung $f': Y \rightarrow X$ mit $f \circ f' = \text{id}_Y$ und $f' \circ f = \text{id}_X$.

Dass es mindestens eine solche Abbildung gibt, haben wir vorher gezeigt. Also müssen wir nur noch die *Eindeutigkeit* von f' zeigen.

Für solche „Eindeutigkeitsbeweise“ gibt es in der Mathematik ein Standard-Beweisschema; das lautet:

Seien f' und f'' Abbildungen der fraglichen Art; wir zeigen $f' = f''$.

Das bedeutet, dass f' und f'' dasselbe Objekt nur unter verschiedenem Namen bezeichnen. (Beachten Sie, dass wir dazu nicht annehmen müssen, dass f' und f'' verschieden sind). Natürlich kann man die Beweise auch so aufziehen:

Seien f' und f'' zwei verschiedene Abbildungen der fraglichen Art. Dann folgt ein Widerspruch. Also war $f' = f''$.

(Versuchen Sie stets zuerst die erste Beweisvariante.)

Nun ans Werk! Seien f' und f'' Abbildungen von Y nach X mit

$$f \circ f'(y) = y \text{ und } f \circ f''(y) = y \text{ für alle } y \in Y$$

sowie

$$f' \circ f(x) = x \text{ und } f'' \circ f(x) = x \text{ für alle } x \in X .$$

Wir zeigen, dass $f' = f''$ ist. Um die Gleichheit der Abbildungen f' und f'' nachzuweisen, müssen wir zeigen, dass für jedes $y \in Y$ gilt $f'(y) = f''(y)$.

Dies ergibt sich aber einfach wie folgt: Da f bijektiv ist, gibt es zu jedem $y \in Y$ genau ein $x \in X$ mit $f(x) = y$. Jede zu f inverse Abbildung muss y auf x abbilden. Also gilt

$$f'(y) = x \text{ und } f''(y) = x ,$$

also

$$f'(y) = f''(y) .$$

Wenn $f: X \rightarrow Y$ eine bijektive Abbildung ist, so nennen wir die (nach obigem Satz eindeutig bestimmte) Abbildung $f': Y \rightarrow X$ mit $f' \circ f(x) = x$ für alle $x \in X$ und $f \circ f'(y) = y$ für alle $y \in Y$ die zu f **inverse** Abbildung und schreiben in Zukunft f^{-1} für f' .

1.4 Wann haben zwei Mengen gleich viele Elemente?

Dumme Frage! Zwei Mengen haben offensichtlich gleich viele Elemente, wenn es eine Zahl n gibt (nämlich die Anzahl ihrer Elemente), so dass sowohl die eine Menge als auch die andere genau n Elemente hat.

Wenn man dies so erklärt, stellen sich zwei Fragen. Die eine Frage ist eher praktischer Natur: Muss man wirklich die beiden Mengen abzählen, um herauszubekommen, ob sie gleich viele Elemente haben? Die andere Frage scheint theoretischer Natur zu sein, ist aber sehr wichtig: Kann man auch von unendlichen Mengen sagen, dass sie „gleich viele“ Elemente haben?

Zunächst zur ersten Frage. Wenn wir im täglichen Leben feststellen wollen, ob zwei Mengen die gleiche Anzahl von Elementen haben, machen wir das häufig *nicht* so, dass wir die eine Menge abzählen, dann die andere Menge abzählen und dann die Zahlen vergleichen. Stellen wir uns ein Beispiel vor. Wir wollen herausbekommen, ob in einer populären Vorlesung jeder Student einen Sitzplatz erhält. Man könnte natürlich vorher die Plätze und die Studenten zählen (wie soll das gehen?) und die Zahlen vergleichen – oder man könnte die Studierenden Platz nehmen lassen und dann einfach sehen.

Das ist die Idee! Diese führt unschwer zu folgender Definition.

Zwei Mengen X und Y heißen **gleichmächtig**, wenn es eine bijektive Abbildung von X nach Y gibt.

So einfach ist das! Im Studenten-Hörsaal-Beispiel wird die bijektive Abbildung einfach durch die Zuordnung eines Studenten zu seinem Platz hergestellt. Wir gewöhnen uns an diesen Begriff, indem wir einige Beispiele betrachten.

Gleichmächtigkeit endlicher Mengen

Seien X und Y endliche Mengen. Dann gilt: Genau dann sind X und Y (im soeben definierten Sinne) gleichmächtig, wenn sie gleich viele Elemente besitzen.

Wir müssen zwei Richtungen *beweisen*. Wenn X und Y gleichmächtig sind, dann gibt es eine bijektive Abbildung f von X auf Y . Da X endlich ist, können wir X schreiben als

$$X = \{x_1, x_2, \dots, x_n\}$$

für eine nichtnegative ganze Zahl n . (Das bedeutet, dass X genau n Elemente hat.) Nun betrachten wir die Bilder $y_i = f(x_i)$ der Elemente von X . Da f surjektiv ist, kommt jedes Element y von Y unter diesen Elementen vor; das heißt, dass es ein i gibt mit $y = f(x_i)$. Da f injektiv ist, gibt es nur ein solches i . Also sind die Elemente $y_1 = f(x_1), \dots, y_n = f(x_n)$ genau die Elemente von Y . Mit anderen Worten: Auch Y hat genau n Elemente.

Umgekehrt mögen die Mengen X und Y beide die gleiche Anzahl n von Elementen besitzen. Jetzt müssen wir eine bijektive Abbildung von X nach Y finden. Dazu nummerieren

wir die Elemente von X und Y von 1 bis n durch:

$$X = \{x_1, x_2, \dots, x_n\} \text{ und } Y = \{y_1, y_2, \dots, y_n\}.$$

(Das ist möglich, da sowohl X als auch Y endliche Mengen mit genau n Elementen sind.)
Dann definieren wir die Abbildung f von X nach Y durch

$$f(x_i) := y_i.$$

Daraus ergibt sich unmittelbar, dass diese Abbildung f sowohl surjektiv (denn jedes Element $y_i \in Y$ hat als Urbild x_i) als auch injektiv ist (denn y_i hat x_i als einziges Urbild). Also ist f bijektiv, und die Behauptung ist bewiesen. \square

Was ist eine unendliche Menge? Ganz einfach: Eine Menge ist **unendlich**, wenn sie nicht endlich ist. Genauer gesagt: Eine Menge ist **unendlich**, wenn es keine nichtnegative ganze Zahl n gibt, so dass sie genau n Elemente hat.

Beispiele unendlicher Mengen gibt es in Hülle und Fülle. Die Menge \mathbf{N} aller natürlichen Zahlen ist unendlich, die Menge \mathbf{Z} ist unendlich. Auch die Menge $2\mathbf{Z}$ aller geraden ganzen Zahlen ist eine unendliche Menge. Und die Menge \mathbf{R} aller reellen Zahlen ist unendlich. Manche dieser Mengen kann man durch Aufzählung so beschreiben, dass man „drei Pünktchen“ verwendet. Dies sind die „abzählbaren Mengen“; dazu gehören \mathbf{N} , \mathbf{Z} , $2\mathbf{Z}$. Bei anderen Mengen, wie zum Beispiel bei \mathbf{R} , geht dies nicht; solche Mengen nennt man „überabzählbar“. Solche Mengen werden uns aber nur am Rande beschäftigen.

Der Begriff der Gleichmächtigkeit führt bei unendlichen Mengen zu besonders bemerkenswerten (und merkwürdigen) Resultaten.

Gleichmächtigkeit von \mathbf{Z} und $2\mathbf{Z}$

Die Mengen \mathbf{Z} und $2\mathbf{Z}$ sind gleichmächtig.

Zum *Beweis* müssen wir eine bijektive Abbildung von \mathbf{Z} nach $2\mathbf{Z}$ angeben.

Dazu überlegen wir uns zunächst, aus welchen Elementen die Menge $2\mathbf{Z}$ besteht. Dies ist genau die Menge der geraden Zahlen; also

$$\begin{aligned} 2\mathbf{Z} &= \{z \in \mathbf{Z} \mid z \text{ ist gerade}\} = \{z \in \mathbf{Z} \mid \text{es gibt ein } x \in \mathbf{Z} \text{ mit } z = 2x\} \\ &= \{2x \in \mathbf{Z} \mid x \in \mathbf{Z}\}. \end{aligned}$$

Nun bietet sich als Abbildung von \mathbf{Z} nach $2\mathbf{Z}$ die Abbildung f an, die durch $f(x) := 2x$ definiert ist. Der Rest ist einfach:

Ist f eine Abbildung von \mathbf{Z} nach $2\mathbf{Z}$? Ja, denn jeder ganzen Zahl x wird eine gerade Zahl $2x$ zugeordnet.

Ist f surjektiv? Ja, denn eine beliebige gerade Zahl $2x$ hat die ganze Zahl x als Urbild.

Ist f injektiv? Ja, denn zwei verschiedene ganze Zahlen x und y werden auf die verschiedenen Zahlen $2x$ und $2y$ abgebildet. \square

Gleichmächtigkeit von \mathbf{N} und \mathbf{Z}

Die Mengen \mathbf{N} und \mathbf{Z} sind gleichmächtig.

Auch hier müssen wir eine Abbildung f finden, die jeder natürlichen Zahl eine ganze Zahl so zuordnet, dass jede ganze Zahl genau einmal erfasst wird. Es gibt viele Abbildungen, die dies leisten; die einfachste sieht vielleicht so aus:

$$f(0) := 0, f(1) := -1, f(2) := 1, f(3) := -2, f(4) := 2, f(5) := -3, f(6) := 3, \dots$$

Die allgemeine Formel dafür lautet

$$f(2n) := n \text{ (für } n \in \mathbf{N}) \text{ und } f(2n-1) := -n \text{ (für } n \in \mathbf{N}, n \geq 1) .$$

Es ist offensichtlich, dass jede ganze Zahl als Bild vorkommt; also ist f surjektiv. Ferner überzeugt man sich leicht, dass jede ganze Zahl nur ein Urbild hat: Eine natürliche Zahl n hat $2n$ und eine negative ganze Zahl $-n$ hat $2n-1$ als Urbild.

Also ist f eine bijektive Abbildung. \square

Als Übung sollen Sie zeigen, dass \mathbf{N} und \mathbf{Q} gleichmächtig sind (siehe Übungsaufgabe 17).

Unendliche Mengen haben viele charakteristische Eigenschaften; eine davon ist die folgende, die man auch zur Definition unendlicher Mengen hätte verwenden können: Ob von einer unendlichen Menge endlich viele Elemente entfernt wurden, kann man an ihrer Mächtigkeit nicht erkennen. Diese Charakterisierung stammt von Richard Dedekind (1831–1916).

Dedekindsche Beschreibung unendlicher Mengen

Jede unendliche Menge enthält eine echte Teilmenge derselben Mächtigkeit.

Wie folgt dies? Wir betrachten eine beliebige unendliche Menge X . Daraus entfernen wir ein beliebiges Element x ; dann ist $X' := X \setminus \{x\}$ eine echte Teilmenge von X .

Behauptung: X' und X haben dieselbe Mächtigkeit.

Wir machen uns dies für den Fall $X = \mathbf{N}$ und $x = 0$ klar.

Dann ist die Sache einfach; eine gesuchte Abbildung f von $\mathbf{N} = \{0, 1, 2, \dots\}$ nach $\mathbf{N}' = \{1, 2, 3, \dots\}$ ist durch $f(n) := n + 1$ definiert.

Im allgemeinen Fall verwendet man ein ähnliches, aber technisch komplizierteres Argument. (Vergleichen Sie dazu [Hal], Kapitel 15.) \square

Bislang ist unser Prototyp für eine unendliche Menge die Menge \mathbf{N} der natürlichen Zahlen. Gibt es Mengen, die eine noch größere Mächtigkeit haben? Sicher **R**! Gibt es Mengen, die eine noch größere Mächtigkeit als **R** haben? Und Mengen, die eine noch größere Mächtigkeit haben? Bevor wir vor lauter Fragen größenwahnsinnig werden, beantworten wir die Fragen in einem Satz: Ja, hier ist die Freiheit grenzenlos. Es geht immer noch größer!

Satz von der Potenzmenge

Für jede nichtleere Menge X gilt: Die **Potenzmenge** $\mathbf{P}(X)$ (das heißt: die Menge aller Teilmengen von X) ist nicht gleichmächtig zu X .

Da $\mathbf{P}(X)$ immer eine zu X gleichmächtige Teilmenge enthält (nämlich die Menge aller einelementigen Teilmengen $\{x\}$ mit $x \in X$), hat $\mathbf{P}(X)$ also eine größere Mächtigkeit als X .

Diese Tatsache kann man für endliche Mengen ganz einfach durch Abzählen beweisen; dies werden wir im Anschluss auch machen. Der Beweis für den allgemeinen Fall beruht aber auf einem Trick.

Der Beweis erfolgt durch Widerspruch. Angenommen, es gäbe eine bijektive Abbildung f von X auf $\mathbf{P}(X)$.

Dann ist für alle $x \in X$ das Bild $f(x)$ eine Teilmenge von X . Für gewisse $x \in X$ gilt, dass x in $f(x)$ liegt, während für die anderen Elemente $x \notin f(x)$ gilt. (Bisher haben wir nur benutzt, dass f eine Abbildung ist, und haben die Annahme, dass f bijektiv ist, noch gar nicht ausgenutzt!)

Der Trick des Beweises besteht darin, die Subjekte x mit $x \notin f(x)$ genauer unter die Lupe zu nehmen. Wir fassen diese zu einer Menge zusammen:

$$U := \{x \in X \mid x \notin f(x)\}.$$

Dann ist U ganz bestimmt eine Teilmenge von X , also ein Element von $\mathbf{P}(X)$. Da f eine bijektive Abbildung von X nach $\mathbf{P}(X)$ ist, gibt es also ein Element $u \in X$ mit $f(u) = U$.

Für dieses u gilt – wie für jedes Element aus X – die Alternative, in U enthalten zu sein oder nicht. Wir werden aber sofort sehen, dass beide Möglichkeiten auf einen Widerspruch führen, also nicht möglich sind.

Ist $u \in U$, so muss u die definierende Eigenschaft der Menge U erfüllen, nämlich $u \notin f(u) = U$. Also folgte aus $u \in U$ die Tatsache $u \notin U$, offenbar ein Widerspruch.

Daher bleibt nur die zweite Möglichkeit: $u \notin U$. Wegen $U = f(u)$ heißt dies $u \notin f(u)$. Das heißt aber, dass u die definierende Eigenschaft der Menge U erfüllt; daher muss auch $u \in U$

gelten. Wie bitte? Wir haben $u \notin U$ angenommen und daraus messerscharf geschlossen, dass $u \in U$ ist. Was soll der Blödsinn? – Tja, dies zeigt „nur“, dass auch diese Möglichkeit nicht auftreten kann.

Und was heißt das? Ganz einfach: Unsere ursprüngliche Annahme, dass es eine bijektive Abbildung von X nach $\mathbf{P}(X)$ gibt, war falsch. Damit ist der Satz bewiesen. \square

Zugegeben, dieser Beweis ist tricky. Wenn Sie ihn nicht sofort verstanden haben, machen Sie sich nichts draus: Erstens werden wir dieses Argument in der linearen Algebra nicht mehr brauchen, und zweitens wird Ihnen sofort anschließend für den wichtigsten Fall ein anderer Beweis geliefert.

Mächtigkeit der Potenzmenge

Eine endliche Menge mit n Elementen hat genau 2^n Teilmengen. Mit anderen Worten: Die Potenzmenge einer endlichen Menge ist „viel größer“ als die Menge selbst.

Dies *beweisen* wir durch Induktion nach n . Da jede Menge mit nur einem Element genau zwei Teilmengen hat (nämlich \emptyset und sich selbst), ist die Induktionsbasis ($n = 1$) richtig.

Sei nun $n > 1$, und sei die Aussage richtig für $n - 1$. Wir betrachten eine n -elementige Menge X und zeichnen darin ein Element x_0 aus. Damit können wir die Teilmengen von X in zwei Klassen einteilen: In diejenigen Teilmengen von X , die x_0 nicht enthalten, und in diejenigen, die x_0 als Element enthalten. Wir müssen jetzt ausrechnen, wie viele Teilmengen in jeder Klasse liegen.

- Für die erste Klasse ist dies ganz einfach: Diejenigen Teilmengen von X , die x_0 nicht enthalten, sind genau die Teilmengen der $(n - 1)$ -elementigen Menge $X \setminus \{x_0\}$. Nach Induktion gibt es also genau 2^{n-1} Teilmengen in der ersten Klasse.
- Für die zweite Klasse ist es nicht viel schwerer: Jede Teilmenge Y von X , die x_0 enthält, „entspricht eindeutig“ der Teilmenge $Y' := Y \setminus \{x_0\}$ von $X' := X \setminus \{x_0\}$. Also gibt es auch genau 2^{n-1} Teilmengen der zweiten Klasse.

[Zur Übung führen wir die Aussage „entspricht eindeutig“ nochmals eine Ebene formaler aus. Diese Aussage bedeutet, dass die beiden fraglichen Mengen gleichmächtig sind, dass es also eine bijektive Abbildung der Menge aller Teilmengen von X , die x_0 enthalten, auf die Menge aller Teilmengen von X' gibt. Als Kandidat für eine bijektive Abbildung wählen wir die Abbildung f , die für alle $Y \subseteq X$ mit $x_0 \in Y$ durch

$$f(Y) := Y \setminus \{x_0\}$$

definiert ist.

Die Abbildung f ist *surjektiv*, da jede Teilmenge Y' von X' ein Urbild hat, nämlich die Menge $Y := Y' \cup \{x_0\}$.

Die Abbildung f ist auch *injektiv*. Wenn nämlich zwei Teilmengen Y_1 und Y_2 , die x_0 enthalten, verschieden sind, so müssen sie sich außerhalb von x_0 unterscheiden; das heißt

$$Y_1 \setminus \{x_0\} \neq Y_2 \setminus \{x_0\},$$

also

$$f(Y_1) \neq f(Y_2).$$

Also sind diese beiden Mengen gleichmächtig. Das heißt: Die Anzahl der Teilmengen von X , die x_0 enthalten, ist gleich der Anzahl der Teilmengen von X' , also gleich 2^{n-1} .]

Die Anzahl aller Teilmengen von X ist also gleich

$$2^{n-1} + 2^{n-1} = 2 \cdot 2^{n-1} = 2^n.$$

1.5 Die Σ -Notation

Wir werden oft viele Terme ähnlicher Art addieren. Zum Beispiel:

$$1 + 2 + 3 + 4 + \dots + n$$

oder

$$1 + 2 + 4 + 8 + \dots + 2^n$$

oder allgemein

$$a_1 + a_2 + \dots + a_n.$$

Diese Summen kann man auf zwei Arten darstellen: Die erste Möglichkeit besteht darin, die Drei-Punktchen-Schreibweise zu verwenden. Diese haben wir bei den obigen Beispielen angewandt. Diese Schreibweise ist suggestiv und oft unmittelbar verständlich. Ihr Nachteil liegt darin, dass die drei Pünktchen etwas vage sind und das Muster der einzelnen Terme nicht andeuten. Zum Beispiel ist nicht klar, ob

$$1 + 2 + \dots + 2^n$$

eine Summe aus $n + 1$ oder aus 2^n Gliedern ist.

Aus diesem Grund hat man die **Σ -Notation** („sigma“) eingeführt. Dies ist eine abkürzende Schreibweise für eine Summe. Wir definieren

$$\sum_{k=1}^n a_k = a_1 + a_2 + \dots + a_n.$$

Damit können wir die obigen Summen unmissverständlich ausdrücken als

$$\sum_{k=0}^n 2^k \text{ bzw. } \sum_{k=1}^{2^n} k .$$

Die Variable k wird als **Summationsindex** bezeichnet. (Oft werden Sie in der Literatur dafür auch den Buchstaben i oder n finden.) Der Summationsindex muss nicht bei 1 anfangen – und nicht bei n aufhören. Auch Ausdrücke der Form

$$\sum_{k=-3}^5 a_k, \sum_{k=10}^{\infty} b_k, \text{ oder } \sum_{k=-\infty}^{\infty} c_k$$

haben ihren Sinn.

Häufig gibt man den Summationsindex nicht direkt, sondern durch eine Bedingung unter den Σ -Zeichen an. Man schreibt etwa

$$\sum_{0 \leq k \leq n} 2^k \text{ statt } \sum_{k=0}^n 2^k .$$

Der Vorteil dieser Schreibweise liegt in einer sehr hohen Flexibilität. So ist es etwa leicht möglich, zu schreiben

$$\sum_{0 \leq k \leq n, \text{ } k \text{ Primzahl}} \frac{1}{k} .$$

Zum Schluss dieses Abschnitts noch ein wichtiges, scheinbar schwieriges, in Wirklichkeit aber einfaches Thema, nämlich **Doppelsummen**. Diese treten dann auf, wenn man Terme mit zwei Indizes summiert oder wenn man zwei Terme mit je einem Index summiert. Wenden wir uns zunächst dem ersten Fall zu.

$$\sum_{1 \leq k \leq 3, 1 \leq h \leq 3} a_{kh} = a_{11} + a_{12} + a_{13} + a_{21} + a_{22} + a_{23} + a_{31} + a_{32} + a_{33} .$$

Man kann auch schreiben

$$\sum_{1 \leq k \leq 3, 1 \leq h \leq 3} a_{kh} = \sum_{1 \leq k \leq 3} \sum_{1 \leq h \leq 3} a_{kh}$$

Die zweite Form bedeutet, dass man zuerst über h summiert, dann über k („von innen nach außen“). Also

$$\begin{aligned} \sum_{1 \leq k \leq 3} \sum_{1 \leq h \leq 3} a_{kh} &= \sum_{1 \leq k \leq 3} (a_{k1} + a_{k2} + a_{k3}) \\ &= (a_{11} + a_{12} + a_{13}) + (a_{21} + a_{22} + a_{23}) + (a_{31} + a_{32} + a_{33}) . \end{aligned}$$

Entsprechend ergibt sich

$$\begin{aligned}\sum_{1 \leq k \leq 3} \sum_{1 \leq h \leq 3} a_{kh} &= \sum_{1 \leq k \leq 3} (a_{1h} + a_{2h} + a_{3h}) \\ &= (a_{11} + a_{21} + a_{31}) + (a_{12} + a_{22} + a_{32}) + (a_{13} + a_{23} + a_{33}).\end{aligned}$$

Also sind beide Summen gleich. Wenn Assoziativ- und Distributivgesetze gelten, so ist allgemein

$$\sum_{1 \leq k \leq n} \sum_{1 \leq h \leq m} a_{kh} = \sum_{1 \leq h \leq m} \sum_{1 \leq k \leq n} a_{kh}$$

Man nennt dies die **Vertauschung der Summationsreihenfolge**.

Entsprechendes gilt auch dann, wenn wir Produkte $a_k \cdot b_h$ addieren:

$$\sum_{1 \leq k \leq n} \sum_{1 \leq h \leq m} a_k b_h = \sum_{1 \leq h \leq m} \sum_{1 \leq k \leq n} a_k b_h.$$

Dies sieht man, wenn man sich die Summen mit drei Pünktchen ausschreibt:

$$\begin{aligned}\sum_{1 \leq h \leq m} \sum_{1 \leq k \leq n} a_k b_h &= \sum_{1 \leq h \leq m} (a_k b_1 + a_k b_2 + \dots + a_k b_m) \\ &= (a_1 b_1 + a_1 b_2 + \dots + a_1 b_m) + (a_2 b_1 + a_2 b_2 + \dots + a_2 b_m) + \dots \\ &\quad + (a_n b_1 + a_n b_2 + \dots + a_n b_m) \\ &= (a_1 b_1 + a_2 b_1 + \dots + a_n b_1) + (a_1 b_2 + a_2 b_2 + \dots + a_n b_2) + \dots \\ &\quad + (a_1 b_m + a_2 b_m + \dots + a_n b_m) \\ &= \sum_{1 \leq h \leq m} (a_1 b_h + a_2 b_h + \dots + a_n b_h) = \sum_{1 \leq h \leq m} \sum_{1 \leq k \leq n} a_k b_h\end{aligned}$$

Achtung Manchmal schreibt man statt $\sum_{\substack{1 \leq k \leq 3, \\ 1 \leq h \leq 3}}$ auch $\sum_{1 \leq k, h \leq 3}$. Das bedeutet, dass

sowohl k als auch h zwischen 1 und 3 variieren. Man könnte das aber auch so lesen, dass an k nur die Bedingung $1 \leq k$ und an h nur die Bedingung $h \leq 3$ gestellt wird. Bei einer solchen Schreibweise muss dann aus dem Zusammenhang klar sein, was gemeint ist.

1.6 Beweisprinzipien

Grundsätzlich ist jeder mathematische Satz eine wenn-dann-Aussage: Jeder Satz ist so aufgebaut, dass aus gewissen Aussagen (der **Voraussetzung**) eine andere Aussage (die **Behauptung**) mit Hilfe der Gesetze der Logik (und *nur* mit diesen) abgeleitet wird; dies geschieht im **Beweis**. In einer Formel geschrieben hat jede mathematische Aussage die Form $A \Rightarrow B$; dabei ist A die Voraussetzung und B die Behauptung. Die Aufgabe des Beweises ist dann, mit Mitteln der Logik die Gültigkeit der Implikation $A \Rightarrow B$ nachzuweisen.

Dafür gibt es verschiedene Mechanismen. Wir haben für jeden der Basismechanismen schon in diesem Kapitel ein Beispiel gesehen. Wir unterscheiden folgende Beweisarten:

Direkter Beweis Dabei wird aus der Voraussetzung die Behauptung „direkt“ bewiesen. Ein gutes Beispiel dafür ist der Beweis des Satzes über die Invertierbarkeit bijektiver Abbildungen. Ein direkter Beweis sieht im Prinzip wie folgt aus:

Beweis Sei A erfüllt.

Bla, bla, bla.

Also gilt B . □

Beweis durch Kontraposition Einem solchen Beweis liegt die logische Tatsache zugrunde, dass die Implikation „ $A \Rightarrow B$ “ gleichwertig zur Implikation „ $\neg B \Rightarrow \neg A$ “ ist. Statt „Aus A folgt B “, kann man genauso gut die Aussage „Aus nicht- B folgt nicht- A “ zeigen – und ist dann fertig! Daher liest sich ein Beweis durch Kontraposition grundsätzlich so:

Beweis Es gelte die Aussage $\neg B$.

Bla, bla, bla.

Also gilt auch die Aussage $\neg A$. □

Widerspruchsbeweis („indirekter Beweis“) Manchmal zieht man einen Beweis auch in Form eines Widerspruchsbeweises auf: „Angenommen, die Behauptung B wäre falsch; dann (so muss man zeigen) wäre auch die Voraussetzung A falsch“. Bei einem Widerspruchsbeweis hat man eine zusätzliche Aussage, nämlich $\neg B$, zur Verfügung, mit der man arbeiten kann. Ein schönes Beispiel für einen Widerspruchsbeweis ist der Beweis des Satzes über die Potenzmenge. Ein Beweis durch Widerspruch liest sich so:

Beweis Es gelte A . Angenommen, B wäre falsch. Dann gilt $\neg B$.

Bla, bla, bla.

Also ergäbe sich ein Widerspruch. Dieser zeigt, dass die Annahme falsch war. □

Manchmal (nicht immer!) hat man die Wahl zwischen einem direkten Beweis oder einem Beweis durch Widerspruch. In diesen Fällen sollten Sie stets den direkten Weg wählen.

Noch zwei Bemerkungen zu Beweistechniken, also Methoden, mit denen man einen Beweis organisiert. Ich nenne nur zwei.

Ringschlüsse Viele mathematische Sätze sind Äquivalenzaussagen: Verschiedene Behauptungen sind gleichwertig; wenn eine gilt, gelten alle anderen. Wie beweist man einen solchen Satz? Häufig geschieht das durch einen so genannten „Ringschluss“. Wenn die äquivalenten Aussagen die Aussagen (a), (b) und (c) sind, so genügt es, die Implikationen „ $(a) \Rightarrow (b)$ “, „ $(b) \Rightarrow (c)$ “ und „ $(c) \Rightarrow (a)$ “ nachzuweisen. Dann folgt jede Aussage aus jeder anderen.

Induktion Vielleicht die wichtigste Beweistechnik ist das Prinzip der vollständigen Induktion. Üben Sie dies so oft, dass es Ihnen in Fleisch und Blut übergeht. Es lohnt sich! Sie sind eingeladen, mit dieser Übung in Aufgabe 24 zu beginnen.

Bemerkung Information über die Bedeutung von Wörtern wie „Lemma“, „Korollar“, sowie Hinweise auf guten (und schlechten) mathematischen Stil finden Sie in meinem Büchlein „Das ist o.B.d.A. trivial!“ [Beu2].

1.7 Verständnisfragen, Übungen und Tipps

Richtig oder falsch?

Bei den Kästchenaufgaben müssen Sie entweder ein Argument für die Richtigkeit der Aussage oder ein Gegenbeispiel angeben.

Hinweis: Mehr als die Hälfte der „Kästchenaussagen“ (aller Übungen zusammen) ist falsch. Versuchen Sie daher im Zweifelsfall zuerst, ein Gegenbeispiel zu finden. Zu Ihrer Kontrolle sind am Ende des Buches die Antworten (ohne Begründung) aufgelistet.

1. Thema: Implikationen

Seien A, B und C mathematische Aussagen, für die „ $A \Rightarrow B$ “ und „ $B \Rightarrow C$ “ gilt. Welche der folgenden Implikationen sind dann richtig?

- ☐ $A \Rightarrow C$,
- ☐ $B \Rightarrow A$,
- ☐ $C \Rightarrow B$,
- ☐ $C \Rightarrow A$,
- ☐ $\neg A \Rightarrow \neg B$ (mit $\neg A$ wird die Negation der Aussage A bezeichnet),
- ☐ $\neg B \Rightarrow \neg A$,
- ☐ $\neg C \Rightarrow \neg A$,
- ☐ $\neg A \Rightarrow \neg C$.

2. Thema: Mengen

Für je zwei Mengen X und Y gilt:

- ☐ $X \setminus Y = \emptyset \Leftrightarrow X = Y$.
- ☐ $|X \cup Y| = |X| + |Y|$ für alle endlichen Mengen X, Y .
- ☐ $X \cup Y$ endlich $\Rightarrow X, Y$ endlich.
- ☐ $X \cap Y$ endlich $\Rightarrow X, Y$ endlich.
- ☐ Für endliche Mengen X und Y gilt: $|X \cup Y| = |X| + |Y| \Rightarrow X \cap Y = \emptyset$.

3. Thema: Äquivalenzrelationen

Die folgenden Vorschriften definieren eine Äquivalenzrelation auf der Menge der natürlichen Zahlen:

- ☐ $x \sim y \Leftrightarrow x, y$ gerade.
- ☐ $x \sim y \Leftrightarrow x - y$ gerade.

- ☐ $x \sim y \Leftrightarrow x + y$ gerade.
- ☐ $x \sim y \Leftrightarrow x - y$ ungerade.
- ☐ $x \sim y \Leftrightarrow x + y$ ungerade.

4. Thema: Äquivalenzklassen

Sei \sim eine Äquivalenzrelation auf der Menge X .

- ☐ Für jedes $x \in X$ ist die Menge $B(x) := \{y \in X \mid \text{es gilt nicht } y \sim x\}$ eine Äquivalenzklasse von \sim .
- ☐ Wenn \sim nur zwei Äquivalenzklassen hat, so ist $B(x)$ eine Äquivalenzklasse.
- ☐ Wenn es ein $x \in X$ gibt, so dass $B(x)$ eine Äquivalenzklasse ist, so hat \sim nur zwei Äquivalenzklassen.
- ☐ Für alle $x_1, x_2, x_3 \in X$ gilt: Wenn weder $x_1 \sim x_2$ noch $x_2 \sim x_3$ gilt, so gilt auch nicht $x_1 \sim x_3$.
- ☐ Für alle $x_1, x_2, x_3 \in X$ gilt: Wenn weder $x_1 \sim x_2$ noch $x_2 \sim x_3$ gilt, so gilt jedenfalls $x_1 \sim x_3$.

5. Thema: Abbildungen

Eine Abbildung f von X nach Y ist genau dann injektiv, wenn gilt

- ☐ aus $x, x' \in X$ und $x \neq x'$ folgt $f(x) \neq f(x')$,
- ☐ zu jedem $y \in Y$ gibt es höchstens ein $x \in X$ mit $f(x) = y$,
- ☐ zu jedem $x \in X$ gibt es genau ein $y \in Y$ mit $f(x) = y$,
- ☐ sind $x, x' \in X$ mit $f(x) = f(x')$, so ist $x = x'$.

Übungsaufgaben

1. Seien X und Y Mengen, für die gilt

$$\text{für alle } x \in X \text{ gilt } x \notin Y.$$

Folgt daraus $X \neq Y$?

2. In meinem Lieblingssteakrestaurant kann man sich seine Mahlzeit aus folgenden Komponenten selbst zusammenstellen:
 - (a) Hüftsteak, Rumpsteak, Filetsteak, Rib-Eye Steak;
 - (b) Gewicht: 180 g oder 250 g;
 - (c) Beilagen: Folienkartoffeln, Pommes frites, Kroketten, Bratkartoffeln, weißer Langkornreis, Maiskolben, Knoblauchbrot, rote Bohnen, Zwiebelringe, Champignons;
 - (d) Saucen: Kräuterbutter, Pfefferrahmsauce, Sauce nach Art Béarnaise.
Wenn ich jeden Monat einmal dort esse: Wie lange brauche ich, um alle Kombinationen durchzuprobieren?
3. Zeigen Sie, dass die in Abschn. 1.2 angegebenen Äquivalenzrelationen wirklich Äquivalenzrelationen sind.

4. Geben Sie Beispiele von Relationen auf einer Menge X an, die
 - reflexiv, symmetrisch, aber nicht transitiv,
 - symmetrisch, transitiv, aber nicht reflexiv,
 - reflexiv, transitiv, aber nicht symmetrisch,
 - weder reflexiv, noch symmetrisch noch transitivsind.
5. Zeigen Sie im Beweis des Satzes über Äquivalenzklassen, dass $A(y) \subseteq A(x)$ gilt.
6. (a) Zeigen Sie, dass in der gewöhnlichen euklidischen Ebene folgendes gilt: Wenn drei Punkte einer Geraden g den gleichen Abstand zu einer Geraden g' haben, so haben alle Punkte von g den gleichen Abstand von g' .
(b) Seien g und g' zwei Geraden der gewöhnlichen euklidischen Ebene, die nicht parallel sind. Zeigen Sie, dass es zwei Punkte von g gibt, die den gleichen Abstand zu g' haben.
7. Zeigen Sie, dass die folgenden Relationen Äquivalenzrelationen sind:
(a) Zwei natürliche Zahlen sind äquivalent, wenn sie die gleiche Quersumme haben.
(b) Zwei Städte sind äquivalent, wenn man von der einen in die andere per Bahn fahren kann.
8. Welche der folgenden Zuordnungen ist eine Abbildung, welche ist surjektiv, welche injektiv?
 - Mensch \rightarrow Freundin,
 - Mensch \rightarrow Vater,
 - Mensch \rightarrow Handynummer,
 - Mensch \rightarrow Lieblingessen.Machen Sie sich jeweils genau klar, welche Mengen Sie als Definitions- und Bildbereich wählen.
9. Welche der folgenden Abbildungen f von \mathbf{R} in sich sind injektiv, welche sind surjektiv?
 - $f(x) = x^3$,
 - $f(x) = ax^2 + bx + c$ ($a \neq 0$),
 - $f(x) = |x|$,
 - $f(x) = e^x$.
10. Geben Sie Beispiele von Abbildungen $f: X \rightarrow Y$ an, die
 - injektiv und surjektiv,
 - injektiv, aber nicht surjektiv,
 - surjektiv, aber nicht injektiv,
 - weder injektiv noch surjektivsind.
11. Beschreiben Sie die Begriffe „injektiv“, „surjektiv“ und „bijektiv“ anhand der „alternativen Definition“ einer Abbildung. („Eine Abbildung $f \subseteq X \times Y$ ist injektiv, wenn ...“)

12. Zeigen Sie den **Satz über die Bijektivität invertierbarer Abbildungen**:

Sei $f: X \rightarrow Y$ eine Abbildung. Wenn es eine Abbildung $f': Y \rightarrow X$ gibt, so dass gilt

$$f \circ f' = \text{id}_Y \text{ und } f' \circ f = \text{id}_X ,$$

dann ist f bijektiv.

13. Sei f eine Abbildung einer Menge X in eine Menge Y . Zeigen Sie:

(a) Wenn X endlich ist, so gilt:

(b) Sind X und Y endlich und ist $|X| = |Y|$, so gilt

$$f \text{ ist injektiv} \Leftrightarrow f \text{ ist surjektiv.}$$

(c) (**Äquivalenz von Injektivität und Surjektivität**) Wenn f eine Abbildung einer endlichen Menge in sich ist, so gilt:

$$f \text{ ist injektiv} \Leftrightarrow f \text{ ist surjektiv} \Leftrightarrow f \text{ ist bijektiv}$$

(d) Zeigen Sie, dass die Aussage (c) falsch ist, wenn X eine unendliche Menge ist.
[Wählen Sie zum Beispiel $X = \mathbf{Z}$.]

14. Sei f eine Abbildung einer Menge X in sich. Zeigen Sie: Wenn $f \circ f = \text{id}$ gilt, so ist f bijektiv.15. Zeigen Sie, dass die Mengen \mathbf{Z} und $2\mathbf{Z} + 1$ gleichmächtig sind.

16. Zeigen Sie: (a) Es gibt unendlich viele Primzahlen.

(b) \mathbf{Z} und die Menge \mathbf{P} aller Primzahlen sind gleichmächtig.

17. Zeigen Sie: \mathbf{N} und \mathbf{Q} sind gleichmächtig. (Wenn Sie nicht weiterkommen, ziehen Sie das Buch von Halmos [Hal] zu Rate.)18. Lesen Sie [Hal] und machen Sie sich klar, dass \mathbf{N} und die Menge der reellen Zahlen zwischen 0 und 1 nicht gleichmächtig sind.19. Studieren Sie den Beweis des Satzes über Potenzmengen genau, und machen Sie sich klar, dass wir mehr bewiesen haben: Es gibt keine surjektive Abbildung von X nach $\mathbf{P}(X)$.

20. Schreiben Sie die folgende Doppelsumme aus:

$$\sum_{3 \leq i \leq 5, 2 \leq j \leq 6} a_{ij} .$$

21. Eine **binäre Folge** ist eine Folge, deren Elemente nur 0 und 1 sind. Besteht eine solche Folge aus n Komponenten, so spricht man von einer binären Folge der **Länge** n . Wie groß ist die Anzahl der binären Folgen der Länge n ?

[Hinweis: Wenn Sie das nicht schnell sehen, schreiben Sie alle binären Folgen der Längen 2 und 3 auf; dann erhalten Sie eine Vermutung, die Sie dann „nur noch“ beweisen müssen.]

22. Sei X eine n -elementige Menge.
- (a) Geben Sie eine bijektive Abbildung der Menge aller binären Folgen der Länge n auf die Menge aller Teilmengen von X an.
 - (b) Geben Sie damit einen neuen Beweis für die Tatsache $\|\mathbf{P}(X)\| = 2^n$ an.
23. Wie viele Folgen (a_1, a_2, \dots, a_n) der Länge n gibt es, wenn die einzelnen Folgenglieder a_i jeweils genau q Werte annehmen können?
24. Beweisen Sie (mit Induktion):

$$1^3 + 2^3 + 3^3 + \dots + n^3 = (1 + 2 + 3 + \dots + n)^2.$$

Drücken Sie diese Gleichung auch verbal aus (Also etwa: „Die Summe der ersten n hm-hm-Zahlen ist gleich ...“).

25. Was bedeuten die Symbole

$$\prod_{k \leq 10} a_k, \quad \bigcup_{i=5}^{n-1} M_i, \quad \bigcap_{n=0}^{\infty} X_n ?$$


Sie sollten mit folgenden Begriffen umgehen können

Menge, Element, Teilmenge, Durchschnitt, Vereinigung, kartesisches Produkt, Potenzmenge, Relation, Äquivalenzrelation, Äquivalenzklasse, Abbildung, injektiv, surjektiv, bijektiv, invertierbare Abbildung, gleichmächtig

Was sagen Sie dazu?

Achtung! Hier sind Fallen versteckt: Eine falsche Behauptung wird „bewiesen“! Sie sollten jeweils zum einen ein Gegenbeispiel zu der Behauptung konstruieren (um sich zu überzeugen, dass die Behauptung falsch ist) und zum anderen den Fehler im Beweis finden.

Gnutpuaheb. Wenn eine Relation symmetrisch und transitiv ist, ist sie auch reflexiv, also eine Äquivalenzrelation.

Sieweb. Sei \sim eine symmetrische und transitive Relation auf einer Menge X . Sei $x \in X$ beliebig, und sei $x \sim y$. Wegen der Symmetrie ist dann auch $y \sim x$ und aufgrund der Transitivität folgt dann auch $x \sim x$. Also ist \sim reflexiv. 

Alle Strukturen der linearen Algebra bauen auf „Körpern“ auf; diese sind aber nicht der eigentliche Untersuchungsgegenstand der linearen Algebra (dies sind die Vektorräume, die wir im nächsten Kapitel behandeln). Ein „Körper“ ist nicht nur eine Menge, sondern diese Menge trägt zusätzlich eine Struktur: Auf einer Menge sind zwei Operationen (nämlich $+$ und \cdot) erklärt. Grob gesagt, sind Körper algebraische Strukturen, in denen man so rechnen (d. h. addieren und multiplizieren) kann wie mit rationalen oder reellen Zahlen.

2.1 Die Definition

Ein **Körper** besteht aus einer Menge K von Elementen zusammen mit zwei Verknüpfungen $+$ und \cdot , die je zwei Elementen $x, y \in K$ wieder ein Element $x + y$ bzw. $x \cdot y$ von K zuordnen. Damit eine solche Struktur Körper genannt wird, müssen die folgenden drei Gruppen von Gesetzen für alle $x, y, z \in K$ erfüllt sein:

2.1.1 Gesetze der Addition

- *Assoziativität:*

$$(x + y) + z = x + (y + z) .$$

- *Existenz und Eindeutigkeit des neutralen Elements:* Es gibt genau ein Element von K , das wir 0 („Nullelement“) nennen, für das gilt

$$0 + x = x .$$

- *Existenz und Eindeutigkeit inverser Elemente:* Zu jedem x gibt es genau ein Element, das wir $-x$ nennen, für das gilt

$$x + -x = 0 .$$

- *Kommutativität:*

$$x + y = y + x .$$

2.1.2 Gesetze der Multiplikation

- *Assoziativität:*

$$x \cdot (y \cdot z) = (x \cdot y) \cdot z$$

- *Existenz und Eindeutigkeit des neutralen Elements:* Es gibt genau ein vom Nullelement verschiedenes Element, das wir 1 („Einselement“) nennen, für das gilt:

$$1 \cdot x = x \cdot 1 = x .$$

- *Existenz und Eindeutigkeit inverser Elemente:* Zu jedem $x \neq 0$ existiert genau ein Element, das wir x^{-1} nennen, für das gilt:

$$x \cdot x^{-1} = 1 = x^{-1} \cdot x .$$

- *Kommutativität:*

$$x \cdot y = y \cdot x .$$

2.1.3 Distributivgesetz

$$x \cdot (y + z) = x \cdot y + x \cdot z .$$

Wenn Sie nachweisen wollen, dass eine gegebene Struktur ein Körper ist, so müssen Sie (solange keine anderen Charakterisierungen zur Verfügung stehen) *alle* genannten Eigenschaften verifizieren.

Statt $x \cdot y$ werden wir in Zukunft oft einfach xy schreiben.

Um uns an die Definition zu gewöhnen, beweisen wir zwei ganz einfache, aber nützliche Eigenschaften, die für alle Körper gelten und die auf den ersten Blick so unscheinbar sind, dass Sie diese vermutlich übersehen hätten.

1. Multiplikation mit 0

Für jedes x aus K gilt $x \cdot 0 = 0$.

Zum *Beweis* gehen wir wie folgt vor: Wir verwenden zweimal die Tatsache, dass 0 neutrales Element bezüglich der Addition ist, und erhalten

$$x \cdot 0 + 0 = x \cdot 0 = x \cdot (0 + 0) = x \cdot 0 + x \cdot 0 .$$

Wenn wir jetzt auf beiden Seiten $x \cdot 0$ abziehen (das ist eine abkürzende Sprechweise für „ $-x \cdot 0$ addieren“), so ergibt sich $0 = x \cdot 0$. \square

2. (Nullteilerfreiheit)

Sind $x, y \in K$ und gilt $x \neq 0, y \neq 0$, so ist auch $x \cdot y \neq 0$.

Beweis Sei $x \neq 0$ und $xy = 0$. Da $x \neq 0$ ist, existiert das zu x inverse Element x^{-1} . Indem wir die Gleichung $xy = 0$ auf beiden Seiten mit x^{-1} multiplizieren, erhalten wir aufgrund der ersten Eigenschaft

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

also $y = 0$. Damit ist diese Behauptung gezeigt. \square

Auf *eine* Forderung, die unscheinbar formuliert ist, weise ich besonders hin; es handelt sich darum, dass sowohl „+“ als auch „ \cdot “ *Verknüpfungen* auf K sein müssen.

Was bedeutet das? Eine **Verknüpfung** (genauer gesagt: eine „binäre“ Verknüpfung) auf K ist eine Abbildung, die jedem Paar von Elementen von K ein Element von K zuordnet. Ganz vornehm ausgedrückt ist eine Verknüpfung also eine Abbildung von $K \times K$ in K . Die Addition und Multiplikation eines Körpers K sind Abbildungen

$$(x, y) \mapsto x + y \quad \text{und} \quad (x, y) \mapsto x \cdot y.$$

Entscheidend ist, dass das Bild (in unserem Fall also $x + y$ und $x \cdot y$) wieder ein Element von K ist; diese Forderung bezeichnet man auch als **Abgeschlossenheit**.

Um diesen Begriff klar zu machen, betrachten wir einige *Beispiele* von Verknüpfungen und von Abbildungen, die keine Verknüpfungen sind. Als Grundmenge wählen wir die Menge \mathbf{N} der natürlichen Zahlen. Die gewöhnliche Addition und Multiplikation sowie die Exponentiation sind Verknüpfungen, da $x + y, x \cdot y$ und x^y natürliche Zahlen sind, falls x, y natürliche Zahlen sind. Demgegenüber ist aber weder die Subtraktion noch die Division eine Verknüpfung, denn im Allgemeinen ist $x - y$ und x/y keine natürliche Zahl für $x, y \in \mathbf{N}$.

Ein weiteres Beispiel ist folgendes: Wenn K ein Körper ist, so ist das Produkt zweier von Null verschiedener Elemente von K , wie wir wissen, ebenfalls verschieden von Null. Dies kann man auch wie folgt ausdrücken: Die Multiplikation ist auf der Menge $K \setminus \{0\}$ abgeschlossen.

Die ersten *Beispiele* für Körper sind offensichtlich: Sowohl die Menge \mathbf{Q} der rationalen Zahlen als auch die Menge \mathbf{R} und \mathbf{C} der reellen Zahlen bilden, jeweils zusammen mit der gewöhnlichen Addition und Multiplikation, einen Körper. (In diesen Strukturen gelten noch viel mehr arithmetische Gesetze; zum Beispiel kann man bei den rationalen und reellen Zahlen zwischen positiven und negativen Zahlen unterscheiden und beliebig kleine Zahlen betrachten – beides Möglichkeiten, die in Körpern im allgemeinen nicht vorhanden sind.)

Bevor wir weitere Beispiele diskutieren, einige Bemerkungen zur Definition eines Körpers, genauer gesagt einige Bemerkungen dazu, was *nicht* in der Definition eines Körpers steht:

- Wir haben auch für die Multiplikation das Kommutativgesetz gefordert. Dies ist zwar üblich – aber man könnte die elementare Körpertheorie auf weite Strecken auch ohne entwickeln. Wenn in einer algebraischen Struktur alle Axiome eines Körpers gelten, nur das Kommutativgesetz für die Multiplikation nicht, so heißt diese Struktur ein Schiefkörper. Man muss dann auch noch das zweite Distributivgesetz (das bei Körpern aufgrund der Kommutativität der Multiplikation automatisch folgt) fordern, nämlich

$$(x + y) \cdot z = x \cdot z + y \cdot z$$

für alle $x, y, z \in K$. Also: Ein **Schiefkörper** ist eine Menge zusammen mit einer Addition und einer Multiplikation, so dass alle Körperaxiome – mit Ausnahme der Kommutativität der Multiplikation – und beide Distributivgesetze gelten.

Wenn wir die Kommutativität der Multiplikation besonders hervorheben wollen, werden wir auch von einem „kommutativen Körper“ sprechen. Dies bedeutet aber für uns nichts anderes als einfach „Körper“.

Auch Schiefkörper sind hochinteressante algebraische Strukturen; einer der Höhepunkte dieses Kapitels wird die Konstruktion des „Quaternionenschiefkörpers“ sein.

- Mit unseren Axiomen (das heißt: Grundgesetzen) können wir nichts über die „Größe“ (etwa den Absolutbetrag) einzelner Elemente aussagen.

Wir können auch keine Aussage über eine **Anordnung** eines Körpers K (also über eine \leq -Beziehung) machen.

Aus den Körperaxiomen folgen auch keine Axiome über die Stetigkeit der Operationen von K ; wir können also beispielsweise keine Konvergenzaussagen verwenden.

- Man könnte die Körperaxiome auch abschwächen, indem man zum Beispiel nur die Existenz (und nicht die Eindeutigkeit) der neutralen und inversen Elemente fordert. Die Eindeutigkeit kann man nämlich daraus beweisen. Um uns diese Beweise zu ersparen, haben wir ein etwas stärkeres Axiomensystem gewählt. (Wir werden solche Beweise später, in Kap. 9, exemplarisch im Zusammenhang mit der Untersuchung von „Gruppen“ kennen lernen.)

Dieses Kapitel hat zwei größere Abschnitte. Im ersten konstruieren wir wichtige Beispiele von Körpern, im zweiten betrachten wir Automorphismen von Körpern.

2.2 Beispiele von Körpern

Die neben \mathbf{Q} und \mathbf{R} wichtigsten Körper sind der Körper \mathbf{C} der komplexen Zahlen und der Körper $\text{GF}(2)$ aus 0 und 1. Zuerst behandeln wir \mathbf{C} .

2.2.1 Der Körper der komplexen Zahlen

Wir erhalten die komplexen Zahlen aus den reellen Zahlen, indem wir aus einer reellen Zahl zwei machen. Eine **komplexe Zahl** ist ein Paar $z = (a, b)$ reeller Zahlen; man nennt

$$\mathbf{C} = \{(a, b) \mid a, b \in \mathbf{R}\} (= \mathbf{R} \times \mathbf{R}) .$$

den Körper (noch sollten wir vorsichtig sein und nur sagen: die Menge) der komplexen Zahlen.

Man kann komplexe Zahlen addieren und multiplizieren. Dies geschieht dadurch, dass man diese Operationen auf die entsprechenden Operationen in \mathbf{R} zurückführt: Seien $z = (a, b)$ und $z' = (a', b')$ zwei komplexe Zahlen. Dann ist

$$z + z' := (a, b) + (a', b') := (a + a', b + b') .$$

Man sagt auch, die Addition sei **komponentenweise** definiert.

Die Multiplikation ist *nicht* komponentenweise definiert, sondern auf folgende zunächst kompliziert erscheinende Art und Weise:

$$z \cdot z' = (a, b) \cdot (a', b') := (aa' - bb', ab' + a'b') .$$

Wir werden zeigen, dass \mathbf{C} mit dieser Addition und Multiplikation einen Körper bildet.

Um zu sehen, dass die Multiplikation nicht völlig sinnlos ist, zeigen wir die Eindeutigkeit und Existenz eines Einselements. Dazu setzen wir das neutrale Element (bezüglich der Multiplikation) mit $e = (x, y)$ an und berechnen x und y . Zunächst nutzen wir aus, dass e die Zahl $(1, 0)$ neutralisieren muss; das heißt

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

Nach Definition der Multiplikation folgt daraus

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

Das bedeutet, dass $x = 1$ und $y = 0$ sein muss. Also: Wenn es überhaupt ein Element e gibt, das alle Elemente von \mathbf{C} neutralisiert, so muss $e = (1, 0)$ sein. Damit ist die Eindeutigkeit gezeigt. Die Existenz folgt leicht aus der Definition der Multiplikation; es ist nämlich

$$z \cdot e = (a \cdot 1 - b \cdot 0, a \cdot 0 + 1 \cdot b) = (a, b) = z$$

und

$$e \cdot z = (1 \cdot a - 0 \cdot b, 1 \cdot b + a \cdot 0) = (a, b) = z .$$

Also ist e tatsächlich das neutrale Element.

Bevor wir die übrigen Körperaxiome nachweisen, stellen wir eine alternative Darstellung der komplexen Zahlen vor.

Zunächst schreiben wir statt $(a, 0)$ einfach a . Damit können wir die reellen Zahlen als Teil von \mathbb{C} auffassen. Außerdem ergeben sich unmittelbar die folgenden Regeln zur Multiplikation einer reellen Zahl r mit einer komplexen Zahl (a, b) :

$$r \cdot (a, b) = (r, 0) \cdot (a, b) = (r \cdot a, r \cdot b)$$

und

$$(a, b) \cdot r = (a, b) \cdot (r, 0) = (ar, br) = (ra, rb) = r \cdot (a, b) .$$

Insbesondere ist

$$r \cdot (1, 0) = (r, 0) = r \text{ und } r \cdot (0, 1) = (0, r) .$$

Nun kommt der entscheidende Trick: Wir definieren i als $i := (0, 1) \in \mathbb{C}$. Aufgrund der Definition der Multiplikation ergibt sich

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

Diese Vereinbarung macht es möglich, dass wir eine beliebige komplexe Zahl $z = (a, b)$ wie folgt ausdrücken können:

$$z = (a, b) = a \cdot (1, 0) + b \cdot (0, 1) = a + b \cdot i = a + ib .$$

Daher schreiben wir statt (a, b) in Zukunft einfach $a + ib$.

Damit können wir das Produkt zweier komplexer Zahlen $z = a + ib, z' = a' + ib'$ berechnen:

$$z \cdot z' = (a, b) \cdot (a', b') = (aa' - bb', ab' + a'b) = aa' - bb' + i(ab' + a'b) .$$

Andererseits kann man das Produkt $(a + bi) \cdot (a' + b'i)$ „einfach ausrechnen“ und erhält:

$$\begin{aligned} (a + bi)(a' + b'i) &= aa' + (ab' + ba')i + bi \cdot b'i = aa' + bb'i^2 + (ab' + ba')i \\ &= aa' - bb' + (ab' + ba')i = z \cdot z' . \end{aligned}$$

Die Regel zur Multiplikation komplexer Zahlen wird dadurch denkbar einfach: *Man stelle die komplexen Zahlen in der Form $a + ib$ dar und rechne mit solchen Zahlen „ganz normal“, unter Beachtung der Tatsache $i^2 = -1$.* (In Übungsaufgabe 1 sollen Sie sich überzeugen, dass jedes Gleichheitszeichen in obiger Gleichungskette zu Recht besteht.)

Man nennt a den **Realteil** und b den **Imaginärteil** der komplexen Zahl $z = a + ib$. Die komplexe Zahl i heißt die **imaginäre Einheit** von \mathbb{C} . Komplexe Zahlen sind gleich, wenn

ihre Realteile und ihre Imaginärteile gleich sind. Die reellen Zahlen sind offenbar genau die komplexen Zahlen, deren Imaginärteil gleich Null ist.

Die Einführung des Symbols i geht auf Leonhard Euler (1707–1783) zurück.

Nun weisen wir alle Körperaxiome für \mathbb{C} nach.

Gesetze der Addition Da die Addition komponentenweise erklärt ist, ergeben sich alle Gesetze aus den entsprechenden Gesetzen von \mathbb{R} : Das Nullelement ist $(0, 0)$; das zu (a, b) inverse Element ist $(-a, -b)$.

Gesetze der Multiplikation Wir haben bereits nachgewiesen, dass $(1, 0)$ das Einselement ist. Wir müssen jetzt noch nachweisen, dass jedes Element $(a, b) \neq (0, 0)$ ein Inverses hat; das bedeutet, dass es genau ein Element (a', b') gibt mit $(a + bi) \cdot (a' + b'i) = 1$, also

$$(a, b) \cdot (a', b') = (1, 0) .$$

Das bedeutet

$$1 = (a + bi) \cdot (a' + b'i) = aa' - bb' + (ab' + a'b)i .$$

Also muss

$$1 = aa' - bb'$$

und

$$0 = ab' + a'b$$

gelten. Daraus folgt

$$abb' = a^2a' - a \quad \text{und} \quad abb' = -a'b^2 ,$$

also

$$a'(a^2 + b^2) = a .$$

Das heißt

$$a' = \frac{a}{a^2 + b^2} , \quad b' = \frac{-b}{a^2 + b^2} .$$

Somit ist

$$z' = \left(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2} \right)$$

das multiplikative Inverse von $z = (a, b) \neq (0, 0)$. (Beachten Sie, dass in jedem Fall $a^2 + b^2 \neq 0$ ist. Warum? Wegen $z \neq (0, 0)$ ist $a \neq 0$ oder $b \neq 0$ (oder beides). Unabhängig davon, ob $a \neq 0$ positiv oder negativ ist, ist a^2 stets positiv. Also ist $a^2 + b^2$ die Summe der beiden nichtnegativen Zahlen a^2 und b^2 , von denen mindestens eine verschieden von Null ist. Daher ist $a^2 + b^2 > 0$ und insbesondere also $a^2 + b^2 \neq 0$.)

Die Assoziativ- und die Distributivgesetze sollen Sie in den Übungen nachrechnen (siehe Übungsaufgabe 2).

Damit haben wir insgesamt nachgewiesen, dass \mathbf{C} ein Körper ist: der **Körper der komplexen Zahlen**. Man nennt i die imaginäre Einheit; dies ist eine komplexe Zahl, für die $i^2 = -1$ gilt.

Zur Geschichte der komplexen Zahlen lesen wir in H. Heusers *Lehrbuch der Analysis I* das folgende:

Komplexe Zahlen verdanken ihr Leben einem Manne, den seine Mutter (wie er selbst berichtet) abtreiben wollte; der sich dann zu einem Wüstling, Streithansl, magisch-mystischen Mathematiker und europaweit gefeierten Arzt entwickelte; ein Mann, der als Student Rektor der Universität Padua und als Greis Insasse des Gefängnisses von Bologna war; der sich erdreistete, das Horoskop Jesu zu stellen und in seinem Buch „Über das Würfelspiel“ Betrugsanleitungen zu geben, und der nebenbei auch noch die „Cardanische Aufhängung“ erfand: Hieronimo Cardano (1501–1576), ein vollblütiger Sohn der italienischen Renaissance. In seiner *Ars magna sive de regulis algebraicis* („Die große Kunst oder über die algebraischen Regeln“, Nürnberg 1545) führt ihn die unverfängliche Aufgabe, eine Strecke der Länge 10 so in zwei Stücke zu zerlegen, dass das aus ihnen gebildete Rechteck die Fläche 40 hat, zu der quadratischen Gleichung $x(10 - x) = 40$ und zu ihren absurden Lösungen $x_{1,2} := 5 \pm \sqrt{-15}$, absurd, weil man aus negativen Zahlen keine (reellen) Quadratwurzeln ziehen kann. Aber nun geschieht etwas Entscheidendes: Cardano setzt die „geistigen Qualen“, die ihm diese Gebilde bereiten, beiseite und findet durch keck-formales Rechnen, dass tatsächlich $x_1 + x_2 = 10$ und $x_1 x_2 = 40$ ist. Sein ironischer Kommentar: „So schreitet der arithmetische Scharfsinn voran, dessen Ergebnis ebenso subtil wie nutzlos ist“. Die „komplexen“ (zusammengesetzten) Ausdrücke $\alpha + \sqrt{-\beta}$ oder $\alpha + i\sqrt{\beta}$ mit der „imaginären Einheit“ $i = \sqrt{-1}$ sind dann nicht mehr aus der Mathematik verschwunden, so sehr sie auch als schein- und gespensterhaft empfunden wurden. Denn sie lieferten nicht nur „Lösungen“ aller quadratischen und kubischen Gleichungen – und zwar solche, die erbaulicherweise den vertrauten Wurzelsätzen des Francois Vieta (1540–1603) genügten –, vielmehr ergab unverdrossenes (und unverstandenes) Rechnen mit diesen windigen „Zahlen“ sogar Sätze „im Reellen“.

2.2.2 Der Quaternionenschiefkörper

Die Mathematiker haben sich (vor allem im letzten Jahrhundert) gefragt, ob man den Prozess der Erweiterung der reellen Zahlen zu den komplexen wiederholen kann. Lange wurde – ohne Erfolg – damit experimentiert, auf der Menge von Tripeln reeller Zahlen eine sinnvolle Multiplikation zu definieren. Der Entdecker der Quaternionen, William Rowan Hamilton (1805–1865), beschreibt die verzweifelten Szenen in einem Brief an seinen Sohn wie folgt:

Every morning, on my coming down to breakfast, you asked me: “Well, Papa, can you multiply triplets?” Whereto I was always obliged to reply, with a sad shake of the head: “No, I can only add and subtract them”.

Schließlich entdeckte Hamilton, nachdem er 13 Jahre lang unermüdlich danach gesucht hatte, wie man auf der Menge aller 4-Tupel (Quadrupel) reeller Zahlen eine Multiplikation so definieren kann, dass man damit wenigstens einen Schiefkörper erhält. Auf Hamilton

geht auch die Bezeichnung Quaternionen für die Elemente dieses Körpers zurück (lateinisch: quattuor: vier). Übrigens stammt Name Quaternion – aus der Bibel; in Apostelgesch. 12, 4 lesen wir in der lateinischen Ausgabe: *[Herodes] misit [Petrum] in carcerem, tradens quattuor quaternionibus militum custodiendum ...* Auf englisch: *[Herodes] put [Peter] in prison, and delivered him to four quaternions of soldiers to keep him ...*

Der Quaternionenschiefkörper wird zu Ehren Hamiltons heute mit \mathbf{H} bezeichnet. Hamilton beschreibt seine Entdeckung äußerst plastisch. Voller Befriedigung berichtet er in einem Brief an seinen Sohn:

On the 16th of October, 1843, – which happened to be a Monday, and a Council day of the Royal Irish Academy – I was walking in to attend and preside, and your mother was walking with me, along the Royal Canal, to which she had perhaps driven; and although she talked with me now and then, yet an under-current of thought was going on in my mind, which gave at last a result, whereof it is not too much to say that I felt at once the importance. An electric circuit seemed to close; and a spark flashed forth, the herald (as I foresaw, immediately) of many long years to come of definitely directed thought and work, by myself it spared, and at all events on the part of others, if I should even be allowed to live long enough distinctly to communicate the discovery. Nor should I resist the impulse – unphilosophically as it may have been – to cut with a knife on a stone at Brougham Bridge, as we passed it, the fundamental formula with the symbols i, j, k ; namely,

$$i^2 = j^2 = k^2 = ijk = -1$$

which contains the Solution of the Problem, but of course the inscription, has long since mouldered away.

Hamilton war besessen von den Quaternionen: Als er im Jahre 1865 starb, gab es bereits 150 Veröffentlichungen über Quaternionen – von denen Hamilton selbst 109 geschrieben hatte. In seinem Nachlass fand man 60 (in Worten: sechzig) Buchmanuskripte zur Mathematik der Quaternionen. Man kann heute aber sicher sagen, dass Hamilton und seine Anhänger die Bedeutung der Quaternionen viel zu hoch eingeschätzt haben.

Wir beschreiben nun den **Quaternionenschiefkörper \mathbf{H}** auf eine sehr übersichtliche Art und Weise, die sich an der Beschreibung der komplexen Zahl in der Form $z = a + ib$ orientiert.

Wir führen dazu drei neue **imaginäre Einheiten** ein, die wir i, j und k nennen. Die Elemente von \mathbf{H} (also die **Quaternionen**) sind alle Ausdrücke der Form

$$h = a + ib + jc + kd \quad \text{mit} \quad a, b, c, d \in \mathbf{R}. \quad (2.1)$$

Die Summe zweier Quaternionen h und $h' = a' + ib' + jc' + kd'$ ist – wie bei den komplexen Zahlen – komponentenweise definiert:

$$h + h' := a + a' + i(b + b') + j(c + c') + k(d + d').$$

Damit ergeben sich die Additionsgesetze von \mathbf{H} wieder ganz einfach aus denen von \mathbf{R} : Das Nullelement ist $0 = 0 + i \cdot 0 + j \cdot 0 + k \cdot 0$, das *additive Inverse* („Negative“) des Elements

$h = a + ib + jc + kd$ ist

$$-h := -a + i(-b) + j(-c) + k(-d) .$$

Um die Multiplikation in \mathbf{H} zu definieren, definieren wir zuerst die Multiplikationsregeln für die imaginären Einheiten. Es soll sein

$$i^2 := -1, j^2 := -1, k^2 := -1, i \cdot j := k, j \cdot k := i, k \cdot i := j . \quad (2.2)$$

Ferner sollen die folgenden Rechenregeln für die Einheiten i, j und k gelten. Zum einen fordern wir das Assoziativgesetz für die imaginären Einheiten, also

$$j \cdot (j \cdot k) = (j \cdot j) \cdot k, k \cdot (k \cdot i) = (k \cdot k) \cdot i, i \cdot (i \cdot j) = (i \cdot i) \cdot j, i \cdot (j \cdot k) = (i \cdot j) \cdot k, \dots$$

zum zweiten soll jede der imaginären Einheiten i, j, k mit jeder reellen Zahl r vertauschbar sein; das heißt:

$$i \cdot r = r \cdot i, j \cdot r = r \cdot j, k \cdot r = r \cdot k \quad \text{für alle } r \in \mathbf{R} .$$

Daraus ergeben sich alle anderen Gesetze! Zum Beispiel: Sind damit alle Produkte der imaginären Einheiten erklärt? Ja, denn es ergibt sich:

$$j \cdot i = j \cdot (j \cdot k) = (j \cdot j) \cdot k = -1 \cdot k = -k .$$

Entsprechend folgt

$$k \cdot j = -i \text{ und } i \cdot k = -j .$$

Daraus ergibt sich sofort, dass \mathbf{H} in keinem Fall ein *kommutativer* Körper werden kann; denn es ist ja bereits

$$i \cdot j = k \neq -k = j \cdot i .$$

Damit \mathbf{H} zumindest ein Schiefkörper werden kann, müssen wir das Produkt beliebiger Quaternionen erklären. Das ist im Prinzip einfach: Wir multiplizieren die Quaternionen aus und bringen das Produkt auf die Form (2.1), indem wir die Gleichungen (2.2) benutzen.

Sollen wir das einmal mit zwei allgemeinen Quaternionen $h = a + ib + jc + kd$ und $h' = a' + ib' + jc' + kd'$ machen? Ja? Dann holen wir tief Luft und fangen an

$$\begin{aligned} h \cdot h' &= (a + ib + jc + kd) \cdot (a' + ib' + jc' + kd') \\ &= aa' + iab' + jac' + kad' + iba' + i^2bb' + ijb'c' + ikbd' \\ &\quad + jca' + jicb' + j^2cc' + jkcd' + kda' + kidxb' + kjdc' + k^2dd' \\ &= aa' - bb' - cc' - dd' + i(ab' + ba' + cd' - dc') \\ &\quad + j(ac' - bd' + ca' + db') + k(ad' + bc' - cb' + da') . \end{aligned} \quad (2.3)$$

Frage Mal ehrlich, hätten Sie weiter gelesen, wenn ich mit dieser Multiplikationsregel angefangen hätte?

Nun zu den Körperaxiomen. Was ist das Einselement? Nach den Erfahrungen mit \mathbf{C} ist es verführerisch, das Element $1 (= 1 + i \cdot 0 + j \cdot 0 + k \cdot 0)$ zu probieren. Geben wir dieser Verführung nach; es ist:

$$1 \cdot h = 1 \cdot (a + ib + jc + kd) = a + ib + jc + kd = h$$

und

$$h \cdot 1 = (a + ib + jc + kd) \cdot 1 = a + ib + jc + kd = h.$$

Was ist das zu $h = a + ib + jc + kd$ inverse Element? Wir setzen dies als $h' = a' + ib' + jc' + kd'$ an und werten die Gleichung $h \cdot h' = 1$ aus. Gemäß (2.3) ergeben sich daraus die folgenden vier Gleichungen in den Unbekannten a', b', c', d' :

$$\begin{aligned} aa' - bb' - cc' - dd' &= 1, \\ ab' + ba' + cd' - dc' &= 0, \quad ac' - bd' + ca' + db' = 0, \\ ad' + bc' - cb' + da' &= 0. \end{aligned}$$

Daraus ergibt sich nach einigen Versuchen und nach höchstens dreimaligem Verrechnen die erstaunlich einfache Beziehung

$$h' = \left(\frac{a}{a^2 + b^2 + c^2 + d^2}, \frac{-b}{a^2 + b^2 + c^2 + d^2}, \frac{-c}{a^2 + b^2 + c^2 + d^2}, \frac{-d}{a^2 + b^2 + c^2 + d^2} \right).$$

Wir werden in Kap. 4 eine Methode kennen lernen, die solche Gleichungssysteme automatisch löst.

Nun zur Assoziativität der Multiplikation. Wir haben gefordert, dass das Assoziativgesetz jedenfalls für die Einheiten gilt. Nun kommt der Trick: Da wir die Multiplikation über die Multiplikation der Symbole i, j, k definiert haben, folgt die Assoziativität von \mathbf{H} aus der Assoziativität der imaginären Einheiten!

Das ist Mathematik: Wir haben das Problem der Verifizierung unendlich vieler Gleichungen darauf reduziert, überschaubar viele Gleichungen zu verifizieren!

In Übungsaufgabe 5 sollen Sie einen ähnlichen Trick für den Nachweis der *Distributivgesetze entwickeln*.

Insgesamt haben wir jetzt gezeigt, dass die Quaternionen einen echten Schiefkörper bilden! Die Konstruktion des Quaternionenschiefkörpers ist ein Stück klassischer Mathematik, das zur mathematischen Allgemeinbildung gehört. Studieren Sie es entsprechend gründlich.

Man könnte eine ganze Vorlesung über diesen Körper halten. Wir machen hier nur noch zwei kleine Bemerkungen.

1. Der Körper \mathbf{R} der reellen Zahlen ist in \mathbf{H} enthalten. Die reellen Zahlen sind nämlich genau die Quaternionen $a + ib + jc + kd$, bei denen die Koeffizienten b, c, d der imaginären Einheiten Null sind.
2. Auch der Körper \mathbf{C} der komplexen Zahlen ist in \mathbf{H} enthalten. Die komplexen Zahlen sind genau die Quaternionen $h = a + ib + jc + kd$, bei denen die Koeffizienten von j und k verschwinden.

Weitere Information über die Geschichte der Quaternionen und ihre mathematischen Eigenschaften findet man in [Ebbl], Kap. 6.

2.2.3 Einige endliche Körper

Zwar ist die Menge \mathbf{Z} aller ganzen Zahlen kein Körper (warum?), aber man kann aus \mathbf{Z} eine äußerst wichtige Klasse von Körpern gewinnen. Die Grundlage hierfür ist der **euklidische Algorithmus**.

Division mit Rest

Sei a eine ganze Zahl und b eine natürliche Zahl mit $b \geq 1$. Dann gibt es eindeutig bestimmte ganze Zahlen q und r mit folgenden Eigenschaften

$$a = q \cdot b + r \text{ und } 0 \leq r < b .$$

Man nennt r den **Rest**, der bei Division von a durch b entsteht.

Wir werden diesen Satz hier nicht beweisen; Sie sind im ersten Projekt des Kapitels 6 eingeladen, einen Beweis zu liefern. □

Einige *Beispiele*:

- $a = 13, b = 3$: $13 = 4 \cdot 3 + 1$, also $q = 4, r = 1$;
- $a = -13, b = 3$: $-13 = (-5) \cdot 3 + 2$, also $q = -5, r = 2$.

Für uns ist der Rest einer Division besonders wichtig; deshalb hat er eine wichtige Bezeichnung erhalten. Man bezeichnet ihn mit

$$a \bmod b \quad (\text{gesprochen „}a \text{ modulo } b\text{“}).$$

Tab. 2.1 Addition und Multiplikation in \mathbb{Z}_5

$+_5$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

\cdot_5	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

Tab. 2.2 Addition und Multiplikation in \mathbb{Z}_6

$+_6$	0	1	2	3	4	5
0	0	1	2	3	4	5
1	1	2	3	4	5	0
2	2	3	4	5	0	1
3	3	4	5	0	1	2
4	4	5	0	1	2	3
5	5	0	1	2	3	4

\cdot_6	0	1	2	3	4	5
0	0	0	0	0	0	0
1	0	1	2	3	4	5
2	0	2	4	0	2	4
3	0	3	0	3	0	3
4	0	4	2	0	4	2
5	0	5	4	3	2	1

Beispiele $13 \bmod 3 = 1$, $-13 \bmod 3 = 2$.

Wir schreiben

$$a \equiv b \pmod{n} \quad (\text{gesprochen „}a \text{ kongruent } b \text{ modulo } n\text{“}),$$

falls $a \bmod n = b \bmod n$ ist, d. h. falls a und b bei Division durch n den gleichen Rest ergeben.

Mit diesem Hilfsmittel können wir zu jeder natürlichen Zahl n eine algebraische Struktur \mathbb{Z}_n mit Addition und Multiplikation erklären, indem wir definieren:

\mathbb{Z}_n besteht aus den ganzen Zahlen $0, 1, \dots, n-1$; die Addition $+_n$ in \mathbb{Z}_n ist definiert durch

$$a +_n b := (a + b) \bmod n;$$

schließlich ist die Multiplikation \cdot_n in \mathbb{Z}_n definiert durch

$$a \cdot_n b := (a \cdot b) \bmod n.$$

In \mathbb{Z}_n werden also alle durch n teilbaren ganzen Zahlen mit 0 identifiziert; alle Zahlen, die bei Division durch n den Rest 1 ergeben, mit 1 usw. Die Addition und Multiplikation erfolgt „modulo n “. Wir schreiben den Index n an das Plus- und Malzeichen, um es nicht mit den Plus- und Malzeichen für ganze Zahlen zu verwechseln. Wenn keine Verwechslungsgefahr mehr besteht, werden wir den Index großzügig weglassen. Wir betrachten nun zwei Beispiele, nämlich \mathbb{Z}_5 und \mathbb{Z}_6 ; wir geben die Addition und die Multiplikation jeweils in einer Tabelle an (Tab. 2.1 und 2.2).

Was erkennen wir an diesen Beispielen? Sind diese Strukturen Körper? Schon auf den ersten Blick erkennen wir, dass die additiven Strukturen zwar sehr ähnlich, die multiplikativen Strukturen aber grundsätzlich verschieden sind. Bei \mathbb{Z}_5 ist zunächst kein Grund

zu erkennen, weshalb diese Struktur kein Körper sein sollte. Aber \mathbf{Z}_6 ? Hier sieht die Multiplikationstabelle schon sehr merkwürdig aus. Zwar ist 1 offenbar ein Einselement, aber das Element 2 hat kein Inverses (denn es gibt kein Element x mit $2 \cdot x = 1$)! Also kann \mathbf{Z}_6 bestimmt kein Körper sein.

Wir beantworten die Frage, ob überhaupt und wenn ja, welche \mathbf{Z}_n Körper sind, in drei Schritten.

1. Etappe

Für jede natürliche Zahl n erfüllt \mathbf{Z}_n alle Axiome eines kommutativen Körpers – bis möglicherweise auf die Existenz eines multiplikativen Inversen.

Dies ergibt sich daraus, dass sich die arithmetischen Gesetze von \mathbf{Z} auf \mathbf{Z}_n übertragen. Ganz einfach sind die Existenz der neutralen Elemente und des negativen Elements einzusehen: Da $a + 0 = a$ ist, ist auch $a +_n 0 = (a + 0) \bmod n = a + 0 = a$. Ebenso folgt $a \cdot 1 \bmod n = a$ für alle Elemente a von \mathbf{Z}_n . Das zu $a \in \mathbf{Z}_n$ negative Element ist $n - a$; denn es ist

$$a +_n (n - a) = (a + (n - a)) \bmod n = n \bmod n = 0.$$

Auch Assoziativ- und Distributivgesetz in \mathbf{Z}_n folgen aus den entsprechenden Gesetzen in \mathbf{Z} . Zum Beispiel ist für Elemente a, b, c von \mathbf{Z}_n :

$$\begin{aligned} (a +_n b) +_n c &= (a + b) \bmod n +_n c = ((a + b) \bmod n + c) \bmod n \\ &= (a + b + c) \bmod n = \dots = a +_n (b +_n c). \end{aligned}$$

Der Nachweis der Assoziativität der Multiplikation und des Distributivgesetzes ist Thema der Übungsaufgabe 11. \square

Der Beweis der ersten Etappe ist ein Beispiel für ein wichtiges Prinzip in der Mathematik, das *Homomorphieprinzip*: Eigenschaften gehen von einer ‚großen‘ Struktur (in unserem Fall \mathbf{Z}) durch eine geeignete Abbildung (einen „Homomorphismus“) auf eine kleine Struktur (in unserem Fall \mathbf{Z}_n) über. Wir werden solche Homomorphiephänomene in Abschn. 2.3 und ausführlich in Kap. 5 studieren. Es ist nicht erstaunlich, dass die „kleine“ Struktur entsprechende Eigenschaften wie die „große“ hat. Es ist aber zunächst nicht einsichtig, weshalb \mathbf{Z}_n (unsere „kleine“ Struktur) ein Körper sein soll (also zusätzliche Eigenschaften haben soll).

Der nächste Schritt gibt dafür auch keine zusätzlichen Indizien her.

2. Etappe

Wenn n eine zusammengesetzte ganze Zahl ist, also $n = ab$ mit $a > 1$ und $b > 1$, dann ist \mathbf{Z}_n bestimmt kein Körper.

Dies ergibt sich ohne große Schwierigkeiten: Ist $n = a \cdot b$ mit $a > 1$ und $b > 1$, so sind a und b Elemente von \mathbf{Z}_n mit

$$a \cdot_n b = (a \cdot b) \mod n = 0.$$

Dies ist aber in einem Körper unmöglich, da das Produkt je zweier von Null verschiedener Elemente ein von Null verschiedenes Element ergibt (Nullteilerfreiheit). \square

Damit können höchstens die Strukturen \mathbf{Z}_n Körper sein, für die n eine Primzahl ist. (Eine **Primzahl** ist eine ganze Zahl $p > 1$, die nur von 1 und p geteilt wird.)

Nun die Überraschung:

3. Etappe (Existenz von Körpern mit Primzahlordnung)

Wenn p eine Primzahl ist, dann ist \mathbf{Z}_p ein kommutativer Körper mit genau p Elementen.

Beweis Es ist klar, dass \mathbf{Z}_p genau p Elemente hat. Also ist nur zu zeigen, dass jedes Element $a \neq 0$ ein multiplikatives Inverses hat. Im Gegensatz zu den vorigen Untersuchungen werden wir das zu a inverse Element nicht explizit angeben, sondern nur seine Existenz zeigen. Wir brauchen dazu folgenden Hilfssatz über ganze Zahlen (den wir hier nicht beweisen):

Teilt die Primzahl p ein Produkt $a \cdot b$ ganzer Zahlen a und b , so teilt p mindestens eine der Zahlen a oder b .

Zum Beispiel folgt aus der Tatsache, dass eine Primzahl p die Zahl 35 teilt, dass p eine der Zahlen 5 oder 7 teilt, also dass $p = 5$ oder $p = 7$ ist. (Wenn p keine Primzahl ist, dann gilt diese Aussage nicht: 4 teilt $60 = 6 \cdot 10$, aber 4 teilt weder 6 noch 10.) Wir werden diesen Hilfssatz hier *nicht* beweisen, aber in Kap. 6 in allgemeinerem Rahmen ausführlicher darauf eingehen.

Nun ans Werk: Sei p eine Primzahl, sei a eine natürliche Zahl mit $1 \leq a < p$. Es ist zu zeigen, dass es eine natürliche Zahl $a' < p$ gibt mit $a \cdot_p a' = 1$, das heißt

$$a \cdot a' \mod p = 1.$$

Dazu betrachten wir die Produkte

$$0 \cdot_p a, 1 \cdot_p a, 2 \cdot_p a, 3 \cdot_p a, \dots, (p-1) \cdot_p a. \quad (*)$$

Behauptung Diese Zahlen sind paarweise verschieden „modulo p “.

Angenommen, es wäre $h \cdot_p a \bmod p = k \cdot_p a \bmod p$ mit $h \neq k$. Das bedeutet, dass $h \cdot a$ und $k \cdot a$ den gleichen Rest bei Division durch p haben. Also ist $h \cdot a - k \cdot a = (h - k) \cdot a$ durch p teilbar. Da p das Produkt $(h - k) \cdot a$ teilt, muss p einen der Faktoren teilen.

Kann p die Zahl a teilen? Nein. Denn a ist kleiner als p . Also muss p die Zahl $h - k$ teilen. Ferner liegt diese Zahl zwischen $-(p - 1)$ und $+(p - 1)$ (denn es ist $h \leq p - 1$ und $k \geq 0$). Die einzige Zahl in diesem Intervall, die durch p teilbar ist, ist aber – die Zahl 0. Daher muss $h - k = 0$, also $h = k$ sein: ein Widerspruch!

Was wissen wir jetzt? Die Elemente von $(\mathbb{Z}_p \setminus \{0\}, *)$, die verschieden von $0 (= 0 \cdot_p a)$ sind, sind verschiedene Elemente von $\mathbb{Z}_p \setminus \{0\}$. Da dies $p - 1$ Elemente sind, und da $\mathbb{Z}_p \setminus \{0\}$ nur $p - 1$ Elemente hat, müssen das alle Elemente von $\mathbb{Z}_p \setminus \{0\}$ sein!

Was nützt uns dies? Viel! Daraus folgt nämlich, dass jedes Element aus $\mathbb{Z}_p \setminus \{0\}$ eine Darstellung der Form $(*)$ hat. Insbesondere hat die Zahl 1 eine solche Darstellung. Daher muss es eine Zahl $h \in \{1, \dots, p - 1\}$ geben mit

$$1 = h \cdot_p a.$$

Dann ist aber dieses Element $h \in \mathbb{Z}_p \setminus \{0\}$ invers zu a . □

2.2.4 Konstruktion eines Körpers mit vier Elementen

Wir stellen uns vor, dass es einen Körper K mit genau vier Elementen gibt. Wir werden sukzessiv die Elemente und die Operationen von K bestimmen. Dadurch wird dieser Körper so klar vor uns stehen, dass wir von seiner Existenz überzeugt sind.

Ein Körper K mit vier Elementen enthält – wie jeder Körper – die Elemente $0, 1, 1 + 1$ (das wir 2 taufen können), $1 + 1 + 1, \dots$. Diese Elemente müssen aber nicht notwendig verschieden sein. Im Gegenteil: Da K nur endlich viele Elemente hat, muss irgendwann

$$\underbrace{1 + 1 + \dots + 1}_{n \text{ mal}} = \underbrace{1 + 1 + \dots + 1}_{n \text{ mal}}$$

mit $n > m$ sein. Daraus ergibt sich

$$\underbrace{1 + 1 + \dots + 1}_{(n-m) \text{ mal}} = 0$$

Insbesondere ist 0 eine Summe von Einsen. Indem wir 1 subtrahieren, erhalten wir

$$\underbrace{1 + 1 + \dots + 1}_{(n-m-1) \text{ mal}} = -1$$

Insbesondere halten wir die erstaunliche Tatsache fest, dass das Element -1 von der Form $1 + 1 + \dots + 1$ ist.

Tab. 2.3 Die Addition in $\text{GF}(4)$

+	0	1	a	$a+1$
0	0	1	a	$a+1$
1	1	0	$a+1$	a
a	a	$a+1$	0	1
$a+1$	$a+1$	a	1	0

Wir wissen, dass \mathbb{Z}_4 kein Körper ist. Das bedeutet, dass K – falls ein solcher Körper überhaupt existiert! – nicht nur aus den Elementen $0, 1, 1+1, 1+1+1, \dots$ bestehen kann. Es muss also ein Element a von K geben, das nicht von der Form $0, 1, 1+1, \dots$ ist. Welche Elemente muss K enthalten?

Sicherlich – wie jeder Körper – die Elemente 0 und 1 . Außerdem a , also auch $a+1$. Kann $a+1$ eines der Elemente sein, die wir schon aufgelistet haben? Nein: Wäre $a+1 = a$, so folgte $1 = 0$; wäre $a+1 = 1$, so folgte $a = 0$; wäre $a+1 = 0$, so folgte $a = -1$, also von der Form $1+1+\dots+1$: In jedem Fall ein Widerspruch.

Also besteht der Körper K mit vier Elementen genau aus den Elementen $0, 1, a$ und $a+1$. Damit kann man die Additions- und Multiplikationstabelle leicht aufstellen.

Zunächst zur *Addition*: Da die Addition kommutativ ist, gilt $1+a = a+1$. Als nächstes folgt $1+1 = 0$, denn sowohl $1+1 = 1$ als auch $1+1 = a$ als auch $1+1 = a+1$ führen auf einen Widerspruch; also muss $1+1$ das vierte Element, also gleich 0 sein. Daraus folgt auch

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

und

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

Damit ergeben sich alle anderen Summen:

$$(a+1) + 1 = a = 1 + (a+1) \text{ und } (a+1) + a = a + a + 1 = 1.$$

Die Additionstabelle sieht also wie folgt aus (Tab. 2.3).

Man sieht mit einem Blick, dass 0 das Nullelement ist und dass jedes Element genau eine additive Inverse hat. Mit etwas mehr Mühe zeigt man auch das Assoziativgesetz.

Nun zur *Multiplikation*: Alle Produkte, in denen 0 oder 1 als Faktor vorkommt, sind klar. Wir müssen also nur noch $a \cdot a$, $(a+1) \cdot a$, $a \cdot (a+1)$ und $(a+1) \cdot (a+1)$ bestimmen. Was kann $a \cdot (a+1)$ sein? Weder 0 (sonst wäre $a=0$ oder $a+1=0$), noch a (sonst wäre $a+1=1$), noch $a+1$ (sonst wäre $a=1$). Also muss $a \cdot (a+1) = 1$ sein. Entsprechend ergibt sich $(a+1) \cdot a = 1$. Daraus folgt

$$a^2 = a(a+1) + a = 1 + a = a+1 \text{ und } (a+1)^2 = (a+1)a + a+1 = 1 + a+1 = a.$$

Somit hat die Multiplikationstabelle folgendes Aussehen (Tab. 2.4).

Tab. 2.4 Die Multiplikation in $\text{GF}(4)$

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

Auch bei dieser Tabelle sieht man unschwer, dass jedes von Null verschiedene Element genau ein multiplikatives Inverses hat (denn in jeder von 0 verschiedenen Zeile und Spalte kommt das Element 1 genau einmal vor). Wie üblich sind das Assoziativ- und die Distributivgesetze zwar prinzipiell sehr einfach, in Wahrheit aber relativ mühsam nachzuweisen.

Wenn dies alles geleistet ist, dann haben wir bewiesen: Es gibt einen endlichen Körper mit genau 4 Elementen.

Schlussbemerkungen Endliche Körper mit q Elementen werden oft mit \mathbb{F}_q oder $\text{GF}(q)$ bezeichnet. („GF“ steht für „Galoisfeld“ nach Evariste Galois (1811–1832). Das Wort „Feld“ steht dabei für „Körper“; im Englischen heißt ein (mathematischer) Körper bis heute „field“.)

Evariste Galois hat (um das mindeste zu sagen) ein äußerst interessantes Leben geführt. Sie sollten nicht versäumen, in einem Buch über Geschichte der Mathematik (etwa [Wu-ßA]) den Roman seines Lebens nachzulesen.

Wir haben uns klargemacht, dass es endliche Körper $\text{GF}(p)$ gibt, wenn p eine Primzahl ist. In der Algebra zeigt man, dass es einen endlichen Körper $\text{GF}(q)$ genau dann gibt, wenn q eine Primzahlpotenz, also von der Form $q = p^n$ mit p Primzahl und n natürliche Zahl, ist. Jeder solche Körper ist bis auf Isomorphie (siehe den folgenden Abschnitt) eindeutig bestimmt. Das heißt: Für jede Primzahlpotenz q gibt es genau einen Körper mit q Elementen. In den Übungen (siehe Übungsaufgabe 14) wird $\text{GF}(9)$ konstruiert werden.

2.3 Automorphismen von Körpern

Es hat sich in der Mathematik als außerordentlich nützlich erwiesen, zu jeder Struktur die zugehörigen strukturerhaltenden Abbildungen (Homomorphismen, Automorphismen) zu betrachten. Wir untersuchen die Automorphismen von einigen der in Abschn. 2.2 konstruierten Körper; dies dient hauptsächlich dem Zweck, an substantiellem, aber technisch nicht zu schwierigem Stoff Mathematik zu üben.

2.3.1 Die Definitionen

Seien K und L Körper; wir bezeichnen in beiden Körpern die Addition mit $+$ und die Multiplikation mit \cdot .

Ein **Homomorphismus** von K nach L ist eine Abbildung f von K nach L , für die

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

für alle $x, y \in K$ gilt und folgende Eigenschaft erfüllt ist:

$$f(1) \neq 0.$$

Anschaulich bedeutet dies: Die arithmetische Struktur von K (das heißt „ $x + y$ “, „ $x \cdot y$ “) wird durch f auf die arithmetische Struktur von L (also „ $f(x) + f(y)$ “, „ $f(x) \cdot f(y)$ “) übertragen.

Das Konzept der Homomorphie ist zentral in der Mathematik; es tritt bei allen Strukturen (algebraisch, geometrisch, topologisch, ...) auf. Wir werden es später im Rahmen der linearen Abbildungen ausführlich studieren.

Jeder Homomorphismus zwischen Körpern ist automatisch injektiv (siehe Übungsaufgabe 21). Ein Homomorphismus zwischen Körpern heißt ein **Isomorphismus**, falls er bijektiv ist. Wenn es einen Isomorphismus zwischen zwei Strukturen gibt, sagt man, sie sind **isomorph**.

Isomorphe Strukturen sind „strukturgleich“; das bedeutet, dass sie – bis auf eventuelle andere Namen für die Elemente und Verknüpfungen – gleich sind. Für die Isomorphie zweier Strukturen gebraucht man häufig das Symbol \cong .

Ein Isomorphismus einer Struktur auf sich selbst heißt **Automorphismus**. Jede Struktur hat mindestens einen Automorphismus, nämlich die identische Abbildung; diesen bezeichnet man auch als den **trivialen** Automorphismus.

In diesem Abschnitt werden wir die Automorphismen einiger der in Abschn. 2.2 konstruierten Körper studieren. Es ist leider aber so, dass die meisten dieser Körper nur den trivialen Automorphismus haben. Wir werden bis zu \mathbb{C} vorstoßen müssen, um einen – allerdings sehr wichtigen – nichttrivialen Automorphismus zu sehen.

2.3.2 Der Körper der rationalen Zahlen

In diesem Abschnitt beweisen wir den folgenden Satz.

Starrheit von \mathbb{Q}

Der Körper der rationalen Zahlen besitzt nur den trivialen Automorphismus.

Dazu überlegen wir uns zunächst zwei ganz einfache Hilfstatsachen, nämlich die Invarianz der neutralen Elemente und die Invarianz der negativen Elemente („Invarianz“ bedeutet Unveränderlichkeit).

Invarianz der neutralen Elemente

Jeder Automorphismus f eines beliebigen Körpers K führt das Nullelement 0 und Einselement 1 in sich über. Das heißt, es ist $f(0) = 0$ und $f(1) = 1$.

Zum Beweis schreiben wir

$$0 = 0 + 0$$

(dies folgt aus $a = a + 0$ für alle $a \in K$). Wenn wir auf beide Seiten f anwenden, ergibt sich

$$f(0) = f(0 + 0) = f(0) + f(0) .$$

Nun subtrahieren wir auf beiden Seiten $f(0)$ und erhalten

$$0 = f(0) - f(0) = f(0) + f(0) - f(0) = f(0) ,$$

d. h. $f(0) = 0$.

Damit folgt auch $f(1) \neq 0$. (Denn sonst wäre $f(1) = 0 = f(0)$, und f wäre nicht injektiv.) Nun wenden wir den entsprechenden Trick auf die Gleichung $1 = 1 \cdot 1$ an:

$$f(1) = f(1 \cdot 1) = f(1) \cdot f(1) .$$

Da $f(1) \neq 0$ ist, existiert $f(1)^{-1}$, und wir erhalten

$$1 = f(1) \cdot f(1)^{-1} = f(1) \cdot f(1) \cdot f(1)^{-1} = f(1) ,$$

also $f(1) = 1$. □

Die zweite einfache Tatsache ist die folgende:

Invarianz der negativen Elemente

Jeder Automorphismus f eines beliebigen Körpers K erfüllt

$$f(-a) = -f(a) \text{ für alle } a \in K .$$

Insbesondere ist stets

$$f(-1) = -f(1) = -1.$$

Es ist zu zeigen, dass $f(-a) + f(a) = 0$ ist:

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

Nun zu \mathbf{Q} : Sei f ein beliebiger Automorphismus des Körpers \mathbf{Q} . Wir müssen zeigen, dass f jedes Element von \mathbf{Q} fest lässt. Dazu schalten wir einen Schritt vor:

Invarianz der ganzen Zahlen

Jeder Automorphismus f von \mathbf{Q} lässt jede ganze Zahl fest:

Beweis Sei zunächst n eine natürliche Zahl. Dann ist

$$n = \underbrace{1 + 1 + \dots + 1}_{n \text{ mal}}.$$

Da $f(1) = 1$ ist, ergibt sich daraus

$$f(n) = f(1 + 1 + \dots + 1) = f(1) + f(1) + \dots + f(1) = 1 + 1 + \dots + 1 = n.$$

Wenn $-n$ eine negative ganze Zahl ist, dann ist

$$-n = (-1) + (-1) + \dots + (-1) \quad (n \text{ mal}).$$

Da $f(-1) = -f(1) = -1$ ist, folgt

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

Daraus folgt, dass jede rationale Zahl $q = r/s$ festbleibt:

$$f(q) = f(r/s) = f(r \cdot s^{-1}) = f(r) \cdot f(s^{-1}) = f(r) \cdot f(s)^{-1} = r \cdot s^{-1} = q,$$

da $r, s \in \mathbf{Z}$.

(Die Gleichung $f(s^{-1}) = f(s)^{-1}$ ist Gegenstand von Übungsaufgabe 16.)

Damit haben wir gezeigt, dass der Körper \mathbf{Q} nur den trivialen Automorphismus besitzt; man sagt dazu auch: \mathbf{Q} ist **starr**.

2.3.3 Der Körper der reellen Zahlen

Nun zeigen wir, dass auch der Körper \mathbf{R} starr ist.

Dazu stellen wir uns wieder einen beliebigen Automorphismus f des Körpers \mathbf{R} vor und zeigen, dass f jede reelle Zahl auf sich abbildet. Genau so wie im vorigen Abschnitt zeigt man, dass f jede rationale Zahl wieder auf eine rationale Zahl abbildet (das Bild von r/s ist $f(r)/f(s)$). Also wissen wir, dass f einen Automorphismus von \mathbf{Q} „induziert“. Das Ergebnis des vorigen Abschnitts sagt also, dass f alle rationalen Zahlen festlassen muss. Daher brauchen wir uns nur noch zu überzeugen, dass f an keiner irrationalen Zahl „wackeln“ kann.

Der Trick besteht darin, eine irrationale Zahl zwischen rationalen Zahlen so einzuschachteln, dass sie nicht aus diesem Intervall entweichen kann. Dazu müssen wir nachweisen, dass f die Ordnungsrelation „ \leq “ respektiert:

Invarianz der Ordnungsrelation

Wenn $a < b$ ist, dann gilt $f(a) < f(b)$.

Das zunächst unlösbar scheinende Problem hierbei ist, dass nicht zu sehen ist, wie man f auf eine *Ungleichung* anwenden kann. Man könnte natürlich „ $a < b$ “ so in eine Gleichung verwandeln, dass man „ $a + c = b$ (mit positivem c)“ schreiben kann. Wenn man darauf f anwendet, erhält man

$$f(a) + f(c) = f(b) .$$

Da man aber nicht kontrollieren kann, ob $f(c)$ positiv oder negativ ist, nützt einem diese Gleichung wenig. Also muss man anders vorgehen. Hier kommt der folgende Trick zur Anwendung:

Aus $a < b$ folgt $b - a > 0$. Da jede positive reelle Zahl eine reelle Quadratwurzel hat, gibt es eine reelle Zahl r mit $b - a = r^2$. Damit sind wir in der Lage, mit f zu operieren:

$$f(b) - f(a) = f(b - a) = f(r^2) = f(r) \cdot f(r) > 0 ,$$

da das Quadrat einer jeden reellen Zahl nichtnegativ ist.

Das bedeutet

$$f(b) > f(a) .$$

Hieraus ergibt sich nun leicht, dass \mathbf{R} starr ist: Angenommen, es gäbe eine reelle Zahl r mit $f(r) \neq r$. Sei o. B. d. A. $r < f(r)$. („O. B. d. A.“ ist eine Abkürzung für „ohne Beschränkung der Allgemeinheit“; dies bedeutet in unserem Fall, dass der Fall „ $f(r) < r$ “ genauso funktioniert.)

Wir wählen eine rationale Zahl q mit

$$r < q < f(r) .$$

Wegen der Invarianz der Ordnungsrelation folgt aus $r < q$ aber

$$f(r) < f(q) = q .$$

Dies ist ein Widerspruch. Also bleibt jede reelle Zahl unter f fest. Somit gilt:

Starrheit von \mathbb{R}

Der Körper der reellen Zahlen besitzt nur den trivialen Automorphismus.

Nun kommen wir endlich zu einem Körper mit einem nichttrivialen Automorphismus.

2.3.4 Konjugiert-komplexe Zahlen

Der Körper \mathbb{C} der komplexen Zahlen hat zwar – man glaubt es kaum – überabzählbar viele Automorphismen (siehe etwa [Hun, S. 317, Exercise 6]), *ein* Automorphismus spielt aber eine ganz besondere Rolle; diesen behandeln wir hier.

Für eine komplexe Zahl $z = a + ib$ sei

$$\bar{z} = a - ib$$

Diese Zahl heißt die zu z **konjugiert-komplexe Zahl**.

Satz über konjugiert-komplexe Zahlen

Die Abbildung f von \mathbb{C} in sich, die definiert ist durch

$$f(z) = f(a + ib) := a - ib = \bar{z} ,$$

ist ein Automorphismus von \mathbb{C} .

Beweis Offenbar ist f surjektiv und injektiv. Sind $z = a + ib$ und $z' = a' + ib'$ zwei beliebige komplexe Zahlen, so gilt:

$$f(z + z') = f(a + a' + i(b + b')) = a + a' - i(b + b') = a - ib + a' - ib' = f(z) + f(z') ,$$

und

$$\begin{aligned} f(z z') &= f((a + ib)(a' + ib')) = f(aa' - bb' + i(ab' + a'b)) \\ &= aa' - bb' - i(ab' + a'b) = (a - ib)(a' - ib') = f(z)f(z') . \end{aligned}$$

Hurra! Die Abbildung f ist ein Automorphismus von \mathbf{C} , der nicht alle Elemente fest lässt!

Apropos: Welche Elemente von \mathbf{C} bleiben unter f fest? Das sind diejenigen Elemente, für die gilt

$$f(z) = z,$$

also

$$a - ib = a + ib, \quad \text{d. h.} \quad 0 = 2ib, \quad \text{also} \quad b = 0.$$

Das sind also genau die Elemente von \mathbf{C} , deren Imaginärteil gleich Null ist, also genau die reellen Zahlen.

2.4 Verständnisfragen, Übungen und Tipps

Richtig oder falsch?

1. Thema: Komplexe Zahlen

- ☐ i ist die einzige komplexe Zahl, deren Quadrat gleich -1 ist.
- ☐ 1 ist die einzige komplexe Zahl, die zu sich selbst (multiplikativ) invers ist.
- ☐ Die multiplikative Inverse einer Zahl aus $\mathbf{C} \setminus \mathbf{R}$ ist nicht reell.
- ☐ Das Produkt je zweier Zahlen aus $\mathbf{C} \setminus \mathbf{R}$ ist in $\mathbf{C} \setminus \mathbf{R}$.

2. Thema: Automorphismen von Körpern.

Sei K ein Körper.

- ☐ Die Identität ist immer ein Automorphismus von K .
- ☐ Die Identität ist nie ein Automorphismus von K .
- ☐ Jeder Körper hat nur die Identität als Automorphismus.
- ☐ Jeder Körper hat mindestens zwei Automorphismen, nämlich die Identität und die **Nullabbildung** (das ist diejenige Abbildung, die jedes Element auf 0 abbildet).
- ☐ Kein Automorphismus $\neq \text{id}$ bildet ein Element von K auf sich ab.
- ☐ Jeder Automorphismus von K lässt mindestens zwei Elemente von K fest.
- ☐ $\text{GF}(4)$ hat nur die Identität als Automorphismus.

Übungsaufgaben

1. Überzeugen Sie sich, dass jedes der Gleichheitszeichen in folgender Gleichungskette im Körper \mathbf{C} der komplexen Zahlen zu Recht besteht (das heißt, auf die Definition zurückgeführt werden kann).

$$\begin{aligned} (a + bi)(a' + b'i) &= aa' + ab'i + bia' + bi \cdot b'i \\ &= aa' + bb'i^2 + (ab' + ba')i \\ &= aa' - bb' + (ab' + ba')i = z \cdot z'. \end{aligned}$$

2. Zeigen Sie, dass in \mathbf{C} die Assoziativgesetze und das Distributivgesetz gelten.
3. Berechnen Sie die multiplikativen Inversen der folgenden komplexen Zahlen:

$$5 + 2i, 7 - i, 1 + 2i.$$

4. Zeigen Sie, dass für die Einheiten von \mathbf{H} das Assoziativgesetz gilt, falls man (neben den anderen Festlegungen, die wir in Abschn. 2.2 für die Einheiten getroffen haben) nur die folgenden Gesetze fordert:

$$j \cdot (j \cdot k) = (j \cdot j) \cdot k, k \cdot (k \cdot i) = (k \cdot k) \cdot i, i \cdot (i \cdot j) = (i \cdot i) \cdot j.$$

5. (a) Zeigen Sie: Man kann die Gültigkeit der Distributivgesetze in \mathbf{H} auf die Gültigkeit der Distributivgesetze für die Einheiten zurückführen.
(b) Weisen Sie die Distributivgesetze für die Einheiten von \mathbf{H} nach.
6. Führen Sie den Quaternionenschiefkörper nach der Art und Weise ein, in der wir die komplexen Zahlen eingeführt haben. Also etwa so: Die Quaternionen sind die 4-Tupel (a, b, c, d) reeller Zahlen. Speziell setzen wir fest: $i := (0, 1, 0, 0)$, $j := (0, 0, 1, 0)$, $k := (0, 0, 0, 1)$. Die Addition wird komponentenweise definiert.
(a) Definieren Sie die Multiplikation.
(b) Zeigen Sie, dass man auf diese Weise einen Schiefkörper erhält.
7. Sei n eine natürliche Zahl. Zeigen Sie: Wenn a, b, a', b' natürliche Zahlen sind mit

$$a' \equiv a \pmod{n} \text{ und } b' \equiv b \pmod{n},$$

so gilt auch

$$a' + b' \equiv a + b \pmod{n} \text{ und } a' \cdot b' \equiv a \cdot b \pmod{n}.$$

8. Lösen Sie die folgenden Gleichungen „modulo 11“ (das heißt in \mathbf{Z}_{11}):

$$6 \cdot x = 2, 2 \cdot x + 4 = 9, 3 \cdot x - 9 = 5, 7 \cdot x = 1.$$

9. Welche der folgenden Gleichungen sind „modulo 12“ (das heißt in \mathbf{Z}_{12}) lösbar? Geben Sie gegebenenfalls eine Lösung an:

$$6 \cdot x = 2, 2 \cdot x + 4 = 9, 3 \cdot x - 9 = 5, 7 \cdot x = 1.$$

10. Warum ist \mathbf{Z}_{ab} für $a > 1$ und $b > 1$ kein Körper?
Ist es möglich, dass eine Teilmenge $U \subseteq \mathbf{Z}_{ab}$ (mit der Addition und Multiplikation von \mathbf{Z}_{ab} !) ein Körper ist?
11. Zeigen Sie, dass in \mathbf{Z}_n das Assoziativ- und das Distributivgesetz gelten.

12. Die **Ordnung** eines Elements $a \neq 0$ aus \mathbf{Z}_n ist die kleinste natürliche Zahl i mit $a^i = 1$ (in \mathbf{Z}_n). Bestimmen Sie die Ordnung in folgenden Fällen:

$$a = 5, n = 7; a = 10, n = 17; a = 8, n = 15; a = 2, n = 7.$$

13. Überlegen Sie, dass es keinen Körper mit genau 6 Elementen gibt.
 14. Konstruieren Sie einen Körper, der genau 9 Elemente hat.
 15. Sei K ein Körper, in dem es kein Element a gibt, für das $a^2 = -1$ gilt. Wir definieren auf der Menge $K \times K$ die Addition komponentenweise und eine Multiplikation durch folgende Vorschrift:

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

- (a) Zeigen Sie, dass die so definierte Struktur ein Null- und ein Einselement besitzt.
 (b) Zeigen Sie, dass jedes von 0 verschiedene Element ein multiplikatives Inverses besitzt.
 (c) Ist $K \times K$ mit den oben definierten Verknüpfungen ein Körper?
 16. Zeigen Sie: Wenn f ein Automorphismus eines Körpers K ist, dann gilt $f(k^{-1}) = f(k)^{-1}$ für alle $k \in K \setminus \{0\}$.
 17. Sei K ein endlicher Körper.
 (a) Dann gibt es eine natürliche Zahl n derart, dass

$$n \cdot 1 = \underbrace{1 + 1 + \dots + 1}_{n \text{ mal}} = 0.$$

gilt.

- (b) Sei p die kleinste natürliche Zahl mit $p \cdot 1 = 0$. Zeigen Sie: p ist eine Primzahl.
 [Man nennt p die **Charakteristik** des Körpers K]
 18. (a) Sei p eine Primzahl. Zeigen Sie: \mathbf{Z}_p hat die Charakteristik p .
 (b) Sei K ein Körper mit Charakteristik p . Dann enthält K einen Körper, der isomorph zu \mathbf{Z}_p ist.
 (c) Sei K ein Körper der Charakteristik p , und sei K nicht isomorph zu \mathbf{Z}_p . Dann gilt $|K| \geq p^2$.
 19. Bestimmen Sie alle Automorphismen der Körper $\text{GF}(2)$, $\text{GF}(3)$, $\text{GF}(4)$.
 20. Bestimmen Sie alle Automorphismen des Körpers $\text{GF}(p)$ ($= \mathbf{Z}_p$), wobei p eine Primzahl ist.
 21. Sei $f: K \rightarrow L$ ein Homomorphismus des Körpers K in den Körper L . Zeigen Sie:
 (a) Das Element $0 \in K$ ist das einzige Element, das auf $0 \in L$ abgebildet wird.
 (b) Die Abbildung f ist injektiv.

Projekt: Die Gaußsche Zahlenebene

Ein „Projekt“ ist eine große Übungsaufgabe. Sie sollten versuchen, nachdem Sie den Stoff verstanden haben, ein Projekt zu bearbeiten. Nehmen Sie sich dazu eine Woche

Zeit, lösen Sie pro Tag eine Teilaufgabe. Versuchen Sie, Ihre Erkenntnisse sauber aufzuschreiben.

Erschrecken Sie nicht vor der Länge der Aufgabenstellung. Diese dient nur dazu, Ihnen zu helfen und Ihnen genau zu sagen, was Sie im nächsten Teilschritt tun sollten.

Carl Friedrich Gauß (1777–1855), einer der bedeutendsten Mathematiker aller Zeiten, hat mit seiner geometrischen Interpretation der komplexen Zahlen wesentlich dazu beigetragen, dass man allgemein an die Existenz der komplexen Zahlen „glaubte“. (Diese Interpretation war allerdings schon zuvor von John Wallis (1616–1703) erahnt und von Caspar Wessel (1745–1818) präzise formuliert worden.)

Der erste Schritt ist ganz einfach: Wir identifizieren die komplexe Zahl $a + ib$ mit dem Punkt $P = (a, b)$ der kartesischen Ebene (siehe Abb. 2.1). Zum Beispiel wird die Zahl 0 mit dem Koordinatenursprung $O = (0, 0)$ identifiziert. Umgekehrt nennt man die kartesische Ebene, bei der jeder Punkt als komplexe Zahl interpretiert wird, die **Gaußsche Zahlenebene** (Abb. 2.1).

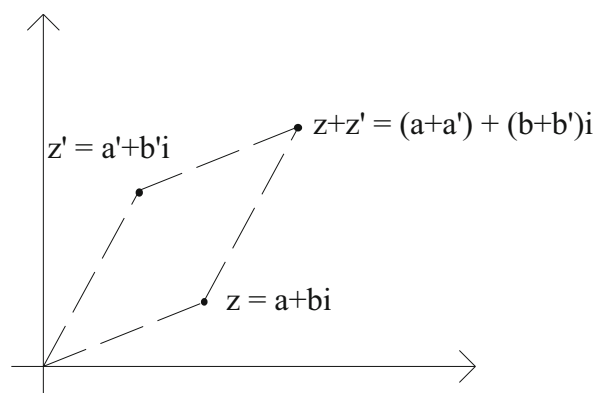


Abb. 2.1 Die Gaußsche Zahlenebene

Die Frage ist jetzt, ob sich die Addition und die Multiplikation im Körper \mathbb{C} geometrisch deuten lassen. Wir setzen für dieses Projekt einige elementare Kenntnisse der ebenen Geometrie voraus. Als einführende und vertiefende Literatur empfehle ich [Cox], Kap. 9.

1. Zeigen Sie, dass die Summe der Elemente $a + ib, c + id \in \mathbb{C}$ der Summe der Punkte (a, b) und (c, d) entspricht. Die **Summe** zweier Punkte P und Q ist dabei erklärt als der vierte Punkt des Parallelogramms, das die Punkte O, P und Q als Ecken hat.

Nun zur Multiplikation. Die Multiplikation eines Punktes mit einer reellen Zahl ist einfach zu beschreiben.

2. Auf welchen Punkt wird ein Punkt A abgebildet, wenn er mit 2 multipliziert wird? Gibt es einen Punkt, der bei dieser Abbildung unverändert bleibt? (Einen solchen Punkt nennt man **Fixpunkt**.) Welche Geraden bleiben fest?

Eine **Streckung** ist eine Abbildung der Ebene \mathbf{R}^2 in sich, die Geraden in Geraden überführt, genau einen Fixpunkt P (das **Zentrum**) hat, so dass jede Gerade durch diesen Fixpunkt in sich überführt wird.

3. Zeigen Sie: Die Abbildung der Gaußschen Zahlenebene in sich, die durch Multiplikation mit einer reellen Zahl $\neq 0$ beschrieben wird, ist eine Streckung mit Zentrum O .

Nun zur Multiplikation mit echt komplexen Zahlen.

4. Zeichnen Sie ein Dreieck mit den Eckpunkten $A = (a, a')$, $B = (b, b')$ und $C = (c, c')$ in die Ebene (wählen Sie konkrete Werte für a, a', b, b', c, c'). Multiplizieren Sie jeden Punkt mit i . (Das heißt: Bilden Sie die Punkte $A^* = (a + ia')i$, $B^* = \dots$) Beschreiben Sie die geometrische Operation, die das Ausgangsdreieck vollzogen hat.
5. Welchen Effekt beobachten Sie, wenn Sie die Punkte A, B, C mit der komplexen Zahl $z = \cos 45^\circ + i \cdot \sin 45^\circ$ multiplizieren.
6. Zeigen Sie: Die Multiplikation mit der komplexen Zahl $z = \cos \varphi + i \cdot \sin \varphi$ beschreibt eine Drehung um den Ursprung um den Winkel φ .
Für welchen Winkel φ erhält man eine Punktspiegelung?
7. Beschreiben Sie die geometrische Operation, die durch die Multiplikation mit einer komplexen Zahl $c + id$ beschrieben wird, indem Sie diese als Hintereinanderausführung einer Drehung und einer Streckung („Drehstreckung“) beschreiben.

Es gibt noch weitere geometrisch interessante Abbildungen, zum Beispiel die Spiegelungen an einer Achse. Kann man auch diese in der Gaußschen Zahlenebene darstellen? Nichts leichter als das:

8. Zeigen Sie: Die Abbildung $\sigma: z \rightarrow \bar{z}$ beschreibt eine Spiegelung an der x -Achse. (Dafür müssen Sie zeigen, dass σ Punkte in Punkte, Geraden in Geraden überführt, dass jeder Punkt der x -Achse festbleibt und dass jede Gerade senkrecht zur x -Achse in sich überführt wird.)

Sie sollten mit folgenden Begriffen umgehen können

Körper, neutrales Element, inverses Element, Assoziativität, Distributivität, Kommutativität, Schiefkörper, \mathbf{R} , \mathbf{C} , i , imaginäre Einheit, Realteil, Imaginärteil, \mathbf{H} , Quaternionen, $a \bmod b$, \mathbf{Z}_n , $\text{GF}(q)$, Homomorphismus, Isomorphismus, isomorph, Automorphismus, starrer Körper, konjugiert komplex



Nun kommen wir zum zentralen Thema der linearen Algebra, den Vektorräumen. Vektorräume haben sich im Laufe des 19. und 20. Jahrhunderts als eine der allerwichtigsten mathematischen Strukturen herausgestellt, die in praktisch jeder mathematischen Disziplin eine grundlegende Rolle spielen. Deshalb bilden Vektorräume auch zu Recht einen Schwerpunkt in der mathematischen Grundausbildung.

3.1 Die Definition

Jedem Vektorraum liegt ein Körper K zugrunde. Welcher spezielle Körper das ist, wird meistens keine Rolle spielen; deshalb nennen wir den Körper *neutral* K . Wir werden die Elemente von K oft auch **Skalare** nennen. Diese werden so bezeichnet, weil man sich die Körperelemente, insbesondere die reellen Zahlen als Elemente einer „Skala“ vorstellen kann.

Die Hauptsache eines Vektorraums sind aber seine Elemente, die Vektoren. Ein **Vektorraum** über dem Körper K (auch **K -Vektorraum** genannt) besteht aus einer Menge V von Elementen, die wir **Vektoren** nennen, die den folgenden Gesetzen genügt:

1. Verknüpfung von Vektoren Es gibt eine Verknüpfung $+$ auf V , die je zwei Vektoren v und w einen Vektor $v + w$ zuordnet, so dass für alle $u, v, w \in V$ die folgenden Eigenschaften erfüllt sind:

Assoziativität

$$u + (v + w) = (u + v) + w .$$

Existenz des Nullvektors Es gibt einen Vektor, den wir mit o bezeichnen, mit folgender Eigenschaft

$$v + o = v .$$

Existenz negativer Vektoren Zu jedem Vektor v gibt es einen Vektor, den wir $-v$ nennen, mit

$$v + (-v) = o .$$

Kommutativität

$$u + v = v + u .$$

2. Verknüpfung von Skalaren und Vektoren Für jeden Vektor $v \in V$ und jeden Skalar $k \in K$ ist ein Vektor $k \cdot v$ definiert (das Objekt $k \cdot v$ (für das wir auch kurz kv schreiben) soll also ein Element von V sein). Diese Bildung des **skalaren Vielfachen** ist so, dass für alle $h, k \in K$ und für alle Vektoren $v, w \in V$ die folgenden Eigenschaften gelten:

$$\begin{aligned}(k + h)v &= kv + hv , \\ (k \cdot h) \cdot v &= k \cdot (h \cdot v) , \\ 1 \cdot v &= v , \\ k \cdot (v + w) &= k \cdot v + k \cdot w .\end{aligned}$$

Mit diesen Eigenschaften werden wir im Laufe dieses Kurses unübertrieben tausendfach umgehen; es lohnt sich also, sich diese einzuprägen.

Wir werden oft **R**-Vektorräume bzw. Vektorräume über dem Körper **C** betrachten; solche Vektorräume nennt man auch **reelle Vektorräume** bzw. **komplexe Vektorräume**.

Wir überlegen uns zunächst zwei einfache Folgerungen aus den Axiomen.

Eindeutigkeit des Nullvektors

Es gibt genau einen Vektor o mit $v + o = v = o + v$ für alle $v \in V$.

Angenommen, es gäbe ein zweites Element aus V , das wir o' nennen, für das $v + o' = v$ für alle $v \in V$ gilt. Um zu zeigen, dass $o' = o$ ist, lassen wir die beiden gegeneinander antreten: Da die Addition kommutativ ist und da o ein Nullelement ist, gilt

$$o + o' = o' + o = o' .$$

Da o' ein Nullelement ist, gilt aber genauso gut

$$o + o' = o .$$

Zusammen ergibt sich

$$o' = o + o' = o, \text{ also } o' = o .$$

□

Eindeutigkeit der negativen Vektoren

Zu jedem $v \in V$ gibt es genau einen Vektor $-v$ mit $v + (-v) = o$.

Um das zu *beweisen* betrachten wir einen Vektor v' , der ebenfalls $v + v' = o$ erfüllt. Wir addieren zu beiden Seiten der Gleichung $v + v' = o$ den Vektor $-v$ und erhalten

$$-v + (v + v') = -v + o = -v ,$$

also

$$v' = (-v + v) + v' = -v + (v + v') = -v + o = -v .$$

□

Gleich hier einige Bemerkungen. In einem Vektorraum sind zwei Strukturen, die des Körpers K und die der Menge V der Vektoren, miteinander verknüpft. In beiden Strukturen kann man addieren; also gibt es zwei verschiedene „+“-Zeichen: eins für Skalare, eins für Vektoren. Ebenso gibt es zwei Nullelemente: das Nullelement von K und den Nullvektor. Schließlich gibt es auch zwei verschiedene Multiplikationen: die in K und die, die einen Skalar mit einem Vektor verknüpft. Wir werden aber nur ein „+“-Zeichen und nur ein „·“-Zeichen verwenden. Das scheint zunächst eine völlig unbillige Schikane zu sein – ist es aber nicht. Denn in spätestens zwei Seiten würde Ihnen der Zwang zum Gebrauch verschiedener Symbole lästig werden. Glauben Sie mir: Bisher hat jeder Mathematiker diese Hürde überwunden; Sie werden es also auch schaffen!

Dieses Kapitel wird zwei weitere größere Abschnitte haben. Zunächst werden wir uns eine ganze Reihe von Beispielen von Vektorräumen ansehen und dann, in einem relativ langen Abschnitt, die elementare Theorie der Vektorräume entwickeln. Dort werden so nützliche Begriffe wie „Basis“ und „Dimension“ eingeführt, die eine hervorragende Beschreibung von Vektorräumen ermöglichen.

3.2 Beispiele von Vektorräumen

In diesem Abschnitt werden wir einige wichtige Beispiele von Vektorräumen vorstellen. Die Strukturen werden so weit erklärt, dass Sie in den Übungen ohne weiteres verifizieren können, dass es sich wirklich um Vektorräume handelt (siehe Übungsaufgabe 1).

3.2.1 Vektorräume mit Hilfe von Geometrie

Wir legen die euklidische Ebene oder den euklidischen Raum (das heißt die Anschauungsebene oder den Anschauungsraum) zugrunde. Wir stellen uns eine spezielle Sorte von

Abbildungen in der Ebene oder im Raum vor, nämlich die Verschiebungen (auch Translationen genannt). Eine Verschiebung ist eindeutig bestimmt, wenn man von irgendeinem Punkt P sein Bild P' kennt. Oft wird diese Translation durch einen **Pfeil** $\overrightarrow{PP'}$ von P nach P' beschrieben.

Wir nennen zwei Pfeile $\overrightarrow{AA'}$ und $\overrightarrow{BB'}$ **äquivalent**, wenn die Punkte A, A', B, B' ein Parallelogramm bilden, also wenn die Pfeile – anschaulich gesprochen – die gleiche Richtung und die gleiche Länge haben. Man kann nachweisen, dass diese Relation eine Äquivalenzrelation ist. Eine Äquivalenzklasse äquivalenter Verschiebungspfeile nennen wir einen **(Verschiebungs-)vektor**. Ein einzelner Pfeil ist dann ein **Repräsentant** des Vektors. (Der Name „Vektor“ kommt übrigens vom lateinischen Wort „vehere“ = transportieren: Man „transportiert“ P nach P' .)

Oft zeichnen wir einen Punkt der Geometrie aus und nennen ihn O . Die Pfeile mit Anfangspunkt O nennen wir dann **Ortsvektoren**. Jede Verschiebung (also jeder Vektor) wird durch genau einen Ortsvektor repräsentiert

Die Rechenoperationen (das heißt Addition und skalare Multiplikation) für Vektoren (also Pfeilklassen) können wir nun dadurch definieren, dass wir diese für Repräsentanten definieren. Die Addition zweier Vektoren ist definiert als Hintereinanderausführung der beiden Verschiebungen; man kann sie sich veranschaulichen, indem man entsprechende Repräsentanten „aneinanderhängt“ (der zweite Pfeil beginnt am Endpunkt des ersten). Die skalare Multiplikation eines Vektors mit einem Faktor k ist das k -fache der Verschiebung. Dieses Vielfache wird repräsentiert durch einen Pfeil der k -fachen Länge.

Mit diesen Festlegungen kann man alle Vektorraumaxiome nachrechnen.

Dieses geometrische Modell wird uns oft als Veranschaulichung dienen. Aber: Es gibt viel mehr Vektorräume, und manche davon sind ganz andersartig.

Eine Bemerkung zur Bezeichnung von Vektoren: Manche Menschen glauben, ein Buchstabe würde nur dann einen Vektor bezeichnen, wenn ein Pfeil darüber gemalt ist. Dies ist ein Aberglaube. Wir werden diese Verschwendung von Druckerschwärze nicht mitmachen!

3.2.2 Der Vektorraum K^n

Sei K ein Körper und n eine natürliche Zahl. Dann können wir auf dem n -fachen kartesischen Produkt

$$V = K^n = \{(k_1, k_2, \dots, k_n) \mid k_i \in K\}$$

von K mit sich eine Addition erklären durch

$$(k_1, k_2, \dots, k_n) + (h_1, h_2, \dots, h_n) := (k_1 + h_1, k_2 + h_2, \dots, k_n + h_n);$$

die skalare Multiplikation definieren wir ebenfalls komponentenweise:

$$k \cdot (k_1, k_2, \dots, k_n) := (kk_1, kk_2, \dots, kk_n).$$

Dies ergibt einen K -Vektorraum V , der uns oft als Prototyp eines Vektorraums dienen wird.

Zum *Beispiel* ist \mathbf{R}^3 der Vektorraum, der aus allen Tripeln (x, y, z) von reellen Zahlen x, y, z besteht; die Addition und die skalare Multiplikation sind komponentenweise erklärt. Dies ist der mit Abstand wichtigste Vektorraum.

3.2.3 Der Vektorraum aller $m \times n$ -Matrizen

Eine Variante des Vektorraums K^n ist der Vektorraum aller Matrizen einer festen Größe.

Seien m und n natürliche Zahlen. Eine $m \times n$ -**Matrix** über dem Körper K ist ein rechteckiges Schema aus m Zeilen und n Spalten, deren Einträge Elemente aus K sind. Also sieht eine $m \times n$ -Matrix A mit den Elementen $a_{ij} \in K$ ($i = 1, \dots, m, j = 1, \dots, n$) folgendermaßen aus:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

Wir schreiben für eine Matrix A oft auch

$$A = (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n} \quad (a_{ij} \in K).$$

Die Menge aller $m \times n$ -Matrizen über K wird mit $K^{m \times n}$ bezeichnet. Auf dieser Menge kann man eine Addition komponentenweise erklären: Die Summe der Matrizen

$$A = (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n} \text{ und } B = (b_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$$

ist definiert als

$$A + B := (a_{ij} + b_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}.$$

Die skalare Multiplikation ist ebenfalls komponentenweise definiert:

$$k \cdot A := (k \cdot a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}.$$

Auch dies ist ein wichtiger K -Vektorraum.

Wir betrachten zwei *Beispiele*. (a) Zunächst sei $K = \mathbf{R}$ und $m = n = 2$. Dann sind

$$\begin{pmatrix} 3 & -1 \\ 0 & 0,5 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 0,14 & 65,537 \\ 1024 & -18 \end{pmatrix}$$

Matrizen aus $\mathbf{R}^{2 \times 2}$; ihre Summe ist

$$\begin{pmatrix} 3 & -1 \\ 0 & 0,5 \end{pmatrix} + \begin{pmatrix} 0,14 & 65,537 \\ 1024 & -18 \end{pmatrix} = \begin{pmatrix} 3,14 & 65,536 \\ 1024 & -17,5 \end{pmatrix}.$$

(b) Jetzt betrachten wir binäre Matrizen, das heißt $K = \text{GF}(2)$. Sei $m = 2, n = 4$. Dann gilt beispielsweise

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

3.2.4 Der Vektorraum aller unendlichen Folgen

Sei K ein beliebiger Körper. Mit V bezeichnen wir die Menge aller unendlichen Folgen (a_1, a_2, \dots) mit $a_i \in K$.

Entsprechend wie bei K^n (siehe Abschn. 3.2.2) macht man V zu einem Vektorraum, indem man die Operationen komponentenweise erklärt.

3.2.5 Ein Vektorraum unendlicher Folgen

Sei K ein beliebiger Körper. Mit V_∞ bezeichnen wir die Menge aller unendlichen Folgen (a_1, a_2, \dots) mit $a_i \in K$, wobei *höchstens endlich viele Elemente a_i ungleich 0 sind*.

Entsprechend wie in Abschn. 3.2.2 macht man V_∞ zu einem Vektorraum, indem man die Operationen komponentenweise erklärt.

Wir werden später sehen (siehe Kap. 6), dass die Elemente von V_∞ im Grunde nichts anderes als die Polynome mit Koeffizienten aus K sind.

Beachten Sie aber, dass die Vektorräume aus den Abschn. 3.2.4 und 3.2.5 grundsätzlich verschieden von K^n sind; denn hier gibt es keine natürliche Zahl n , so dass sich V oder V_∞ als (Teilmenge von) K^n beschreiben lassen.

Den Vektorraum V_∞ werden wir im Projekt am Ende des Kapitels ausführlich studieren.

3.2.6 Vektorräume von Funktionen

Viele in der Analysis auftretende Klassen von Funktionen von \mathbf{R} nach \mathbf{R} bilden einen \mathbf{R} -Vektorraum. Etwa

- die Menge aller Funktionen von \mathbf{R} in sich,
- die Menge aller stetigen Funktionen von \mathbf{R} in sich,
- die Menge aller n -mal differenzierbaren Funktionen von \mathbf{R} in sich.

Dabei werden die **Summe** zweier Funktionen f und g sowie das **Produkt** einer Funktion f mit einer reellen Zahl r wie folgt definiert:

$$(f + g)(r) := f(r) + g(r) \text{ und } (a \cdot f)(r) := a \cdot f(r) \quad \text{für alle } r \in \mathbf{R}.$$

Die obigen Mengen bilden einen \mathbf{R} -Vektorraum (siehe Übungsaufgabe 22).

3.2.7 Lösungen eines Gleichungssystems

Wir betrachten die Gleichung

$$ax + by = 0$$

in den Unbekannten x und y mit $a, b \in \mathbf{R}$. Für $b \neq 0$ ist die Lösungsmenge dieser Gleichung

$$\{(x, -ax/b) | x \in \mathbf{R}\}.$$

Man überzeugt sich leicht, dass die Menge dieser Paare einen \mathbf{R} -Vektorraum bildet.

Eine entsprechende Aussage gilt auch für ein System aus zwei oder mehr Gleichungen. Für drei Gleichungen sollen Sie das in Übungsaufgabe 3 verifizieren. Die Lösbarkeit von Gleichungssystemen werden wir (vom theoretischen und vom praktischen Gesichtspunkt aus) in Abschn. 4.1 detailliert behandeln.

3.2.8 Teilmengen einer Menge

Sei X eine beliebige Menge. Die Vektoren seien die Teilmengen von X . (Nach dem Satz über die Mächtigkeit der Potenzmenge aus Abschn. 1.4 gibt es also genau 2^n Vektoren, falls X eine endliche Menge mit n Elementen ist.) Als Summe zweier Vektoren definieren wir die symmetrische Differenz der entsprechenden Teilmengen:

$$Y + Z := Y \Delta Z := (Y \cup Z) \setminus (Y \cap Z).$$

Als Körper wählen wir $K = \text{GF}(2) = \{0, 1\}$ und definieren die möglichen skalaren Multiplikationen durch

$$1 \cdot Y := Y,$$

$$0 \cdot Y := \emptyset.$$

Damit wird V zu einem $\text{GF}(2)$ -Vektorraum. (Der Nullvektor ist die leere Menge; jede Teilmenge Y von X ist zu sich selbst invers; denn es ist

$$Y \Delta Y = (Y \cup Y) - (Y \cap Y) = Y \setminus Y = \emptyset.)$$

3.2.9 Körper als Vektorräume

Sei K ein Körper und L ein Körper oder ein Schiefkörper, der eine **Erweiterung** von K ist. Das bedeutet, dass $K \subseteq L$ gilt und dass die Operationen von L „eingeschränkt auf K “ die Operationen auf K ergeben. Das wiederum heißt, dass die Verknüpfungen auf K „dieselben“ wie die auf L sind. (Hier haben die Mathematiker manchmal Bauchweh, da – ganz

streng gesehen – zu einer Verknüpfung auch die Angabe des Definitions- und Bildbereichs gehört; die Addition auf K und L sind also in jedem Fall verschiedene Verknüpfungen.) Wenn L eine Erweiterung des Körpers K ist, nennt man K auch einen **Teilkörper** (oder **Unterkörper**) von L .

Es ist klar, dass jeder Körper ein Teilkörper von sich selbst ist. Weitere Beispiele: \mathbf{Q} ist ein Teilkörper von \mathbf{R} , \mathbf{R} ein Teilkörper von \mathbf{C} , \mathbf{C} ein Teilkörper von \mathbf{H} und $\text{GF}(2)$ ein Teilkörper von $\text{GF}(4)$.

Wenn L eine Erweiterung von K ist, können wir einen K -Vektorraum dadurch konstruieren, dass wir kurzerhand die Elemente von L zu Vektoren erklären. Dann sind alle Vektorraumaxiome „trivial“ erfüllt, wir vergessen nämlich nur, dass man auch Elemente von $L \setminus K$ miteinander multiplizieren kann.

Zum Beispiel wird so \mathbf{C} zu einem \mathbf{R} -Vektorraum. Wie geht das? Ganz einfach: Kann man Vektoren addieren? Ja, das ist die Summe zweier komplexer Zahlen. – Kann man einen Vektor mit einem Skalar multiplizieren? Natürlich, das ist einfach das Produkt einer komplexen mit einer reellen Zahl.

Umgekehrt wird auch ein Schuh draus: Man hätte den Körper der komplexen Zahlen einfach wie folgt einführen können: *Wir betrachten den Vektorraum $\mathbf{C} = \mathbf{R}^2$ und definieren darauf zusätzlich ein Produkt durch*

$$(a, b) \cdot (a', b') := (aa' - bb', ab' + a'b) .$$

Dann wären schon eine ganze Reihe von Eigenschaften „automatisch“ erfüllt (d. h., sie würden aus der Vektorraumtheorie folgen), und man müsste „nur noch“ die Eigenschaften der Multiplikation untersuchen.

Ganz entsprechend hätte man den Quaternionenschiefkörper dadurch einführen können, dass man gesagt hätte: *Auf dem Vektorraum $\mathbf{H} := \mathbf{R}^4$ definieren wir das folgende Produkt ...* (siehe Übungsaufgabe 8).

Diese Beispiele zeigen, dass es Vektorräume in Hülle und Fülle und für jeden Geschmack gibt. Nun machen wir uns daran, Vektorräume systematisch zu untersuchen. Das Ziel ist, wie stets in der Mathematik, *gute Beschreibungen* der betrachteten Objekte zu erhalten. Insbesondere wollen wir einen *Überblick über alle Vektorräume* erhalten.

3.3 Elementare Theorie der Vektorräume

In diesem Abschnitt sei V stets ein Vektorraum über dem Körper K . Das zentrale Instrument zur Untersuchung der Struktur eines Vektorraums ist der Begriff der „Linearkombination“; dieser liegt allen anderen Begriffen, wie etwa „linear abhängig“ und „linear unabhängig“, zugrunde.

3.3.1 Der Begriff der Basis

Wir betrachten eine Folge v_1, v_2, \dots, v_n von Vektoren des Vektorraums V . (Dabei erlauben wir sogar, dass gleiche Vektoren auftauchen, dass also $v_i = v_j$ für $i \neq j$ gilt.) Eine **Linearkombination** von v_1, v_2, \dots, v_n ist ein Vektor der Form

$$k_1 v_1 + k_2 v_2 + \dots + k_n v_n \text{ mit } k_i \in K.$$

Beispiele Wenn u, v, w Vektoren eines reellen Vektorraums sind, dann sind die Vektoren $5u + 3v - 7w$ oder $7/2u - v + 17w$ Linearkombinationen dieser Vektoren.

Wir betrachten auch Linearkombinationen von unendlichen Mengen $\{v_1, v_2, \dots\}$ von Vektoren. In diesem Fall müssen wir bei der Definition scharf aufpassen. Ein Vektor v wird **Linearkombination** der Menge $\{v_1, v_2, \dots\}$ genannt, wenn es eine *endliche* Teilmenge von $\{v_1, v_2, \dots\}$ gibt, so dass v eine Linearkombination dieser Teilmenge ist. Mit anderen Worten: Der Vektor v ist eine Linearkombination von $\{v_1, v_2, \dots\}$, falls es endlich viele Vektoren $v_{i_1}, v_{i_2}, \dots, v_{i_n}$ aus $\{v_1, v_2, \dots\}$ und Skalare $k_{i_1}, k_{i_2}, \dots, k_{i_n}$ gibt mit

$$v = k_{i_1} v_{i_1} + k_{i_2} v_{i_2} + \dots + k_{i_n} v_{i_n}.$$

Eine Linearkombination kann gleich dem Nullvektor o sein; sie ist bestimmt dann gleich o , wenn alle $k_i = 0$ sind. Manchmal heißt diese Linearkombination die **triviale** Darstellung des Nullvektors.

Man nennt v_1, v_2, \dots linear unabhängig, wenn die triviale Linearkombination die einzige ist, die den Nullvektor liefert: Die Vektoren v_1, v_2, \dots heißen **linear unabhängig**, falls gilt

$$k_1 v_1 + k_2 v_2 + \dots = o \Rightarrow k_1 = 0, k_2 = 0, \dots$$

Der Begriff „linear unabhängig“ ist der wichtigste – und schwierigste – Begriff der linearen Algebra. Nehmen Sie ihn von vornherein ernst.

Die Folge v_1, v_2, \dots von Vektoren heißt **linear abhängig**, falls sie nicht linear unabhängig ist. Das bedeutet:

Die Vektoren v_1, v_2, \dots, v_n sind genau dann linear abhängig, falls es Skalare k_1, k_2, \dots, k_n aus K gibt, von denen nicht alle Null sind, mit

$$k_1 v_1 + k_2 v_2 + \dots + k_n v_n = o.$$

Mit anderen Worten: Die Vektoren v_1, v_2, \dots sind linear abhängig, falls mit ihnen eine nichttriviale Linearkombination des Nullvektors möglich ist.

Wir wiederholen diese Begriffe nochmals für unendliche Mengen: *Eine unendliche Menge von Vektoren ist linear abhängig, wenn es eine endliche linear abhängige Teilmenge gibt; sie ist linear unabhängig, wenn jede endliche Teilmenge linear unabhängig ist.*

Wie kann man eine Menge linear abhängiger Vektoren erkennen?

Kann ein einzelner Vektor v linear abhängig sein? Dies ist genau dann der Fall, wenn es ein Körperelement $k \neq 0$ gibt mit $kv = 0$. Dies gilt offenbar genau dann, wenn v der Nullvektor ist. Wir haben damit gezeigt: *Der Nullvektor ist der einzige Vektor, der linear abhängig ist.*

Für Mengen aus mindestens zwei Vektoren gilt:

Charakterisierung der linearen Abhängigkeit

Sei v_1, v_2, \dots, v_n eine Folge linear abhängiger Vektoren ($n \geq 2$). Dann gibt es (mindestens) einen Vektor v_i , der sich als Linearkombination der anderen Vektoren darstellen lässt:

$$v_i = \sum_{j \neq i} h_j v_j .$$

Beweis Nach Definition der linearen Abhängigkeit gibt es $k_1, k_2, \dots, k_n \in K$ mit

$$k_1 v_1 + k_2 v_2 + \dots + k_n v_n = 0$$

so, dass mindestens einer der Koeffizienten k_i ungleich Null ist. Dann folgt

$$k_i v_i = -k_1 v_1 - k_2 v_2 - \dots - k_{i-1} v_{i-1} - k_{i+1} v_{i+1} - \dots - k_n v_n ,$$

also, da $k_i \neq 0$,

$$v_i = h_1 v_1 + h_2 v_2 + \dots + h_{i-1} v_{i-1} + h_{i+1} v_{i+1} + \dots + h_n v_n$$

mit $h_j := -k_j/k_i$. □

Damit können wir die ersten Beispiele linear unabhängiger Mengen von Vektoren finden:

- Ein einzelner Vektor ist genau dann linear unabhängig, wenn er nicht der Nullvektor ist.
- Zwei Vektoren sind genau dann linear unabhängig, wenn keiner ein skalares Vielfaches des anderen ist.

Vorsicht Aus der Tatsache, dass die Vektoren v_1, v_2, \dots, v_n linear abhängig sind, folgt im Allgemeinen nicht, dass (zum Beispiel) der erste Vektor v_1 eine Linearkombination von v_2, \dots, v_n ist.

Zum Beispiel ist der Vektor $v_1 = (1, 2)$ aus \mathbf{R}^2 keine Linearkombination der Vektoren $v_2 = (1, 1)$ und $v_3 = (7, 7)$.

Aus der linearen Abhängigkeit folgt lediglich, dass es ein v_i mit obiger Eigenschaft gibt; oft kann man allerdings „o.B.d.A.“ annehmen, dass man v_1 als Linearkombination der übrigen Vektoren darstellen kann. Das „ohne Beschränkung der Allgemeinheit“ bedeutet hier,

dass man die Vektoren so in v_1', v_2', \dots, v_n' umbenennen kann (Umnummerierung), dass tatsächlich v_1' eine Linearkombination von v_2', \dots, v_n' ist. Anschließend tut man dann so, wie wenn man von vornherein umnummeriert hätte und schreibt also statt v_1', \dots, v_n' wieder v_1, \dots, v_n . (Auch das sieht komplizierter aus als es in Wirklichkeit ist; Sie werden sich bald daran gewöhnt haben!)

Eine wichtige Anwendung von linear unabhängigen Vektoren besteht im Koeffizientenvergleich.

Koeffizientenvergleich

Seien v_1, v_2, \dots, v_n linear unabhängige Vektoren. Dann gilt: Aus

$$k_1 v_1 + k_2 v_2 + \dots + k_n v_n = h_1 v_1 + h_2 v_2 + \dots + h_n v_n$$

folgt, dass die Koeffizienten der v_i gleich sind; das bedeutet: $k_1 = h_1, k_2 = h_2, \dots, k_n = h_n$.

Der Beweis ergibt sich fast unmittelbar aus der Definition. Aus

$$k_1 v_1 + k_2 v_2 + \dots + k_n v_n = h_1 v_1 + h_2 v_2 + \dots + h_n v_n$$

folgt zunächst

$$(k_1 - h_1)v_1 + (k_2 - h_2)v_2 + \dots + (k_n - h_n)v_n = 0.$$

Da die Vektoren v_1, v_2, \dots, v_n linear unabhängig sind, müssen nach Definition die Koeffizienten $k_i - h_i$ der v_i gleich Null sein; das heißt:

$$k_1 - h_1 = 0, k_2 - h_2 = 0, \dots, k_n - h_n = 0.$$

Also ist tatsächlich $k_1 = h_1, k_2 = h_2, \dots, k_n = h_n$. □

Die folgenden Begriffe spielen beim Aufbau der Vektorräume eine entscheidende Rolle.

Für eine Menge $\{v_1, v_2, \dots\}$ von Vektoren sei

$$\langle \{v_1, v_2, \dots\} \rangle$$

die Menge aller Linearkombinationen, die mit v_1, v_2, \dots, v_n gebildet werden können. Wenn die Menge der Vektoren endlich ist, so bedeutet dies:

$$\langle \{v_1, v_2, \dots, v_n\} \rangle := \{k_1 v_1 + k_2 v_2 + \dots + k_n v_n \mid k_1, k_2, \dots, k_n \in K\}$$

Man nennt $\langle \{v_1, v_2, \dots\} \rangle$ das **Erzeugnis** der Vektoren v_1, v_2, \dots und schreibt dafür auch kurz $\langle v_1, v_2, \dots \rangle$.

Zum *Beispiel* ist $\langle v \rangle$ die Menge aller skalaren Vielfachen des Vektors v .

Ein weiteres Beispiel ist das folgende: Das Erzeugnis der Vektoren $(1, 0, 0)$ und $(1, 1, 1)$ von \mathbb{R}^3 ist

$$\langle (1, 0, 0), (1, 1, 1) \rangle = \{x \cdot (1, 0, 0) + y \cdot (1, 1, 1) \mid x, y \in \mathbb{R}\} = \{(x + y, y, y) \mid x, y \in \mathbb{R}\}.$$

Hat $\langle v_1, v_2, \dots \rangle$ eine „wilde“ Struktur? Nein, denn es gilt: $\langle v_1, v_2, \dots \rangle$ ist ein K -Vektorraum (der im Allgemeinen kleiner als V ist).

Warum ist dies so? Wir müssen dazu „einfach“ die Vektorraumaxiome nachweisen. Dies sollen Sie in Übungsaufgabe 20 auch tun. Da prinzipiell sehr viele Axiome zu zeigen sind, lohnt es sich (für Sie), sich vorher zu überlegen, welche Axiome man in der jetzigen Situation wirklich beweisen muss und welche nicht. Der Trick besteht darin, nur ganz wenige Axiome nachzuweisen und sich zu überlegen, dass die anderen daraus folgen. Diese wenigen Axiome sind die folgenden:

- Die *Differenz zweier Elemente* aus $\langle v_1, v_2, \dots \rangle$ liegt wieder in $\langle v_1, v_2, \dots \rangle$ (Abgeschlossenheit bezüglich der Subtraktion),
- das *skalare Vielfache jedes Elements* aus $\langle v_1, v_2, \dots \rangle$ liegt wieder in $\langle v_1, v_2, \dots \rangle$ (Abgeschlossenheit bezüglich der skalaren Multiplikation),
- der *Nullvektor* liegt in $\langle v_1, v_2, \dots \rangle$.

Behauptung Es reicht, diese unscheinbaren Eigenschaften zu zeigen.

Wie das? Warum gilt zum Beispiel das Assoziativgesetz bezüglich der Addition oder die Distributivgesetze in $\langle v_1, v_2, \dots \rangle$? Ganz einfach: Weil jeder Vektor aus $\langle v_1, v_2, \dots \rangle$ auch ein Vektor aus V ist und in V die entsprechenden Gesetze gelten!

Warum liegt mit jedem Vektor $v \in \langle v_1, v_2, \dots \rangle$ auch $-v$ in $\langle v_1, v_2, \dots \rangle$? Auch das ist nicht schwer: Nach der dritten und ersten Eigenschaft liegt mit v auch $0 - v = -v$ in $\langle v_1, v_2, \dots \rangle$.

Warum ist die Summe zweier Vektoren $v, w \in \langle v_1, v_2, \dots \rangle$ in $\langle v_1, v_2, \dots \rangle$? Auch diese Eigenschaft folgt einfach: Wir haben uns soeben klargemacht, dass auch $-w$ in $\langle v_1, v_2, \dots \rangle$ liegt. Nach der ersten Eigenschaft liegt also auch der Vektor $v - (-w) = v + w$ in $\langle v_1, v_2, \dots \rangle$.

Die restlichen Axiome folgen ähnlich einfach. Dass es reicht, nur die drei obigen Eigenschaften zu zeigen, wird im **Unterraumkriterium** ausgedrückt (vergleichen Sie dazu Übungsaufgabe 19).

Wenn Sie also in Übungsaufgabe 20 nachweisen, dass $\langle v_1, v_2, \dots, v_n \rangle$ ein Unterraum ist, so beweisen Sie bitte *nur* die drei oben aufgeführten Eigenschaften. Damit machen Sie sich (und eventuellen Korrektoren) das Leben wesentlich einfacher! \square

Man nennt eine Teilmenge W von V , die (bezüglich der auf V definierten Verknüpfungen) selbst wieder einen K -Vektorraum bildet, einen **Unterraum** von V . Die obige Aussage lautet also: $\langle v_1, v_2, \dots, v_n \rangle$ ist ein Unterraum von V .

Jeder Vektorraum V hat die **trivialen** Unterräume V und $\{0\}$.

Wenn $\langle v_1, v_2, \dots \rangle = V$ gilt, so nennt man $\{v_1, v_2, \dots\}$ ein **Erzeugendensystem** von V . Ein Erzeugendensystem von V ist also eine Menge E von Vektoren, so dass jeder Vektor von V eine Linearkombination von Vektoren aus E ist.

Wir werden uns sehr häufig für spezielle, nämlich die „kleinsten“ Erzeugendensysteme, die so genannten „Basen“ interessieren:

Eine Menge von Vektoren aus V heißt eine **Basis** von V , falls sie ein linear unabhängiges Erzeugendensystem ist.

Um gemäß der Definition nachzuweisen, dass eine Menge B von Vektoren eine Basis ist, muss man also zwei Eigenschaften beweisen: B muss linear unabhängig und ein Erzeugendensystem sein. Übrigens: Der Plural von *Basis* heißt **Basen**.

Noch wissen wir nicht, ob ein beliebiger Vektorraum V eine Basis hat. Ein trivialer Fall lässt sich sofort lösen und zwar durch eine Festsetzung, die auf Ihre Gefühle wenig Rücksicht nimmt: Der Vektorraum $V = \{0\}$, der nur aus dem Nullvektor besteht, hat eine Basis, nämlich die leere Menge.

Ein zweites, wichtiges Beispiel ist der Vektorraum K^n . In diesem betrachten wir die Vektoren

$$e_1 = (1, 0, 0, \dots, 0), e_2 = (0, 1, 0, \dots, 0), \dots, e_n = (0, 0, \dots, 0, 1) .$$

Man nennt diese Vektoren auch die **Einheitsvektoren** des K^n . Dann ist die Menge $B = \{e_1, e_2, \dots, e_n\}$ eine Basis von K^n (vergleiche Übungsaufgabe 15).

Die Bedeutung der Basen kommt in folgendem fundamentalen Satz zum Ausdruck.

Eindeutigkeitseigenschaft einer Basis

Jeder Vektor lässt sich eindeutig als Linearkombination der Elemente einer Basis darstellen.

Warum? Sei $\{v_1, v_2, \dots\}$ eine Basis von V . Wir betrachten einen beliebigen Vektor v aus V . Da die Basis $\{v_1, v_2, \dots\}$ ein Erzeugendensystem ist, ist v eine Linearkombination von Elementen aus $\{v_1, v_2, \dots\}$. Also ist nur die *Eindeutigkeit* zu zeigen.

Dazu nehmen wir an, es gäbe Skalare k_1, \dots, k_n und h_1, \dots, h_n aus K mit

$$v = k_1 v_1 + \dots + k_n v_n \text{ und } v = h_1 v_1 + \dots + h_n v_n .$$

Daraus ergibt sich

$$0 = v - v = (k_1 v_1 + \dots + k_n v_n) - (h_1 v_1 + \dots + h_n v_n) = (k_1 - h_1) v_1 + \dots + (k_n - h_n) v_n .$$

Da die Vektoren v_1, \dots, v_n linear unabhängig sind, kann es sich hierbei nur um die triviale Darstellung des Nullvektors handeln. Also sind die Koeffizienten $k_1 = h_1, \dots, k_n = h_n$ alle Null, und das bedeutet $k_1 = h_1, \dots, k_n = h_n$. \square

Aus dieser Eindeutigkeitseigenschaft einer Basis können wir in einem Spezialfall schon folgern, dass *alle Basen gleich viele Elemente haben*. Der Spezialfall ist der, dass V ein **endlicher** Vektorraum ist, also ein Vektorraum mit nur einer endlichen Anzahl von Vektoren. Wenn V der Nullvektorraum ist, so hat V nur eine Basis, nämlich die leere Menge; insbesondere haben alle Basen gleich viele Elemente. Also können wir o.B.d.A. annehmen, dass V einen Vektor $v \neq o$ enthält. Dann muss auch der Körper K endlich sein. Denn V enthält dann auch die Vektoren kv mit $k \in K$. Also enthält V für jedes Körperelement einen Vektor, d. h. $|K| \leq |V| < \infty$.

Sei $|K| = q$ die Anzahl der Elemente in K . Wir beweisen den Satz über die

Anzahl der Vektoren eines endlichen Vektorraums

Sei V ein endlicher Vektorraum über einem Körper K mit q Elementen. Hat V eine Basis aus n Vektoren, so besteht V insgesamt aus genau q^n Vektoren. Insbesondere hat jede Basis genau n Elemente.

Wir betrachten dazu eine Basis $\{v_1, v_2, \dots, v_n\}$. Wie viele Linearkombinationen kann man aus diesen Vektoren bilden? Jeder der Koeffizienten kann jeden der q Werte aus K annehmen. Also gibt es genau $q \cdot q \cdot \dots \cdot q = q^n$ Linearkombinationen. Da $\{v_1, v_2, \dots, v_n\}$ eine Basis ist, werden dadurch alle Vektoren aus V erfasst; wegen der Eindeutigkeitseigenschaft wird jeder Vektor nur einmal dargestellt. Also ist

$$|V| = q^n.$$

Wenn $\{w_1, w_2, \dots, w_m\}$ eine andere Basis ist, so folgt entsprechend

$$|V| = q^m,$$

also muss $m = n$ sein. Somit haben alle Basen gleichviel Elemente. \square

Beachten Sie: Noch wissen wir nicht, ob ein Vektorraum im Allgemeinen überhaupt eine Basis hat! Dies nachzuweisen, muss natürlich unser nächstes Ziel sein. Eine einfache Prozedur, um eine Basis zu konstruieren, ist die folgende: Man startet mit einem Vektor $v_1 \neq o$. Dann wählt man sich einen Vektor v_2 , der kein Vielfaches von v_1 ist. Dann wählt man einen Vektor v_3 , der keine Linearkombination von v_1 und v_2 ist, usw. Man konstruiert also sukzessiv möglichst große Mengen linear unabhängiger Vektoren. Dass dieses Vorgehen zum Ziel (d. h. zu einer Basis) führt, ist der Inhalt der folgenden Aussage.

Basisergänzungssatz

Sei B eine Menge von Vektoren des Vektorraums V . Dann gilt: Genau dann ist B eine Basis von V , wenn B eine maximale linear unabhängige Menge ist.

Dabei heißt eine Menge B linear unabhängiger Vektoren **maximal**, wenn es keinen Vektor v aus $V \setminus B$ gibt, so dass $B \cup \{v\}$ immer noch linear unabhängig ist.

Hierfür *müssen wir zwei Richtungen beweisen*: Zunächst setzen wir voraus, dass B eine Basis ist. Wir müssen zeigen, dass B linear unabhängig ist ... (Hmhm? Folgt das nicht unmittelbar aus der Definition einer Basis? Also können wir das schon mal abhaken!) ... und dass B maximal ist. Dafür müssen wir wirklich etwas arbeiten: Angenommen, es gäbe einen Vektor $v \notin B$, so dass $B \cup \{v\}$ linear unabhängig ist. Da B ein Erzeugendensystem ist, ist

$$v = k_1 v_1 + k_2 v_2 + \dots$$

Damit erhalten wir die folgende nichttriviale Linearkombination des Nullvektors mit Vektoren aus $B \cup \{v\}$:

$$0 = v - k_1 v_1 - k_2 v_2 - \dots$$

Dies ist ein Widerspruch, da $B \cup \{v\}$ als linear unabhängig angenommen war.

Nun zur anderen Richtung. Wir setzen voraus, dass B eine maximale linear unabhängige Menge von Vektoren ist. Es ist zu zeigen, dass B eine Basis ist. Das bedeutet (a), dass B linear unabhängig ist (o.k., das steht schon in der Voraussetzung) und (b), dass B ein Erzeugendensystem ist.

Das zeigen wir so: Sei v ein beliebiger Vektor aus V . Wir müssen zeigen, dass v eine Linearkombination von Vektoren aus B ist: Angenommen, das wäre nicht der Fall. Dann ist $v \notin B$. Also wäre $B \cup \{v\}$ eine Menge von Vektoren, die immer noch linear unabhängig ist. Dies ist ein Widerspruch zur Maximalität von B . \square

Genauso einfach können Sie (in Übungsaufgabe 14) folgende Tatsache nachweisen:

Basisauswahlsatz

Sei B eine Menge von Vektoren von V . Dann gilt: Genau dann ist B eine Basis, wenn B ein minimales Erzeugendensystem ist. \square

Nun können wir einfach zeigen, dass V eine Basis besitzt. Dazu setzen wir ab jetzt voraus, dass V **endlich erzeugbar** ist; das bedeutet, dass es ein Erzeugendensystem aus einer endlichen Menge von Vektoren gibt. (Man kann alle folgenden Überlegungen auch für Vektorräume durchführen, die nicht endlich erzeugbar sind; dann muss man das **Zornsche Lemma** („Jede nach oben beschränkte Kette einer halbgeordneten Menge besitzt ein

maximales Element“) voraussetzen. Meiner Erfahrung nach gehen diese technischen Überlegungen an den meisten Studienanfängern vorbei.)

Satz über die Existenz einer Basis

Sei V ein endlich erzeugbarer Vektorraum. Dann besitzt V eine Basis, sogar eine endliche Basis!

Und dies folgt so: Nach Voraussetzung hat V ein endliches Erzeugendensystem E . Wir nehmen aus E so lange Vektoren weg, bis ein minimales Erzeugendensystem übrig bleibt:

Ist E bereits ein minimales Erzeugendensystem, so sind wir fertig. Ansonsten gibt es einen Vektor $v_0 \in E$, auf den man verzichten kann. Das heißt, dass auch $E_1 := E \setminus \{v_0\}$ ein Erzeugendensystem ist. Ist E_1 eine Basis, so sind wir fertig. Ansonsten gibt es einen Vektor $v_1 \in E_1$ derart, dass auch $E_2 := E_1 \setminus \{v_1\}$ ein Erzeugendensystem ist. Ist E_2 minimal, so sind wir fertig. Ansonsten ...

Da E endlich ist, muss dieser Prozess nach endlich vielen Schritten ein minimales Erzeugendensystem E_n liefern. Nach dem vorigen Satz ist E_n eine Basis. \square

Bemerkung Ich werde immer wieder mit der Äußerung konfrontiert, eine Basis habe drei Definitionen; eine Basis sei nämlich ein minimales Erzeugendensystem, eine maximale linear unabhängige Menge und (manchmal auch „oder“) ein linear unabhängiges Erzeugendensystem. Oft wird dann immerhin noch hinzugefügt, dass diese Definitionen alle gleichwertig seien.

In dieser Aussage steckt zwar ein richtiger Kern (den man erkennen kann, wenn man guten Willens ist), andererseits ist sie derart schief, dass ich sie einfach nicht unkommentiert stehen lassen kann:

1. Eine Basis wird – wie jeder mathematische Begriff – durch eine Definition festgelegt. Natürlich gibt es Beschreibungen, die dazu äquivalent sind; dies wird dann aber in einem Satz formuliert, den man beweisen muss. Wir haben das im Basisergänzungs- und im Basisauswahlsatz gemacht.
2. Man kann bei einem Aufbau der linearen Algebra jede der Beschreibungen einer Basis als Definition wählen. Dies ist keine mathematische, sondern eine didaktische Entscheidung.
3. Ich habe mich dafür entschieden, die Beschreibung einer Basis als linear unabhängiges Erzeugendensystem als Definition zu wählen, weil hier die beiden wesentlichen Eigenschaften einer Basis besonders deutlich werden.

Wenn nicht ausdrücklich anders gesagt, werden wir von nun an die folgende Voraussetzung machen:

- **Vereinbarung** Sei V stets ein *endlich erzeugbarer* Vektorraum.

3.3.2 Der Steinitzsche Austauschsatz

Das in obigem Satz präsentierte Verfahren ist in der Praxis wenig hilfreich zur Konstruktion einer Basis. Wir brauchen also effizientere Verfahren. Ein noch drängenderes Problem ist das folgende: Es ist bei unseren derzeitigen Kenntnissen noch vorstellbar, dass *Sie* eine Basis eines Vektorraums V finden, die 5555 Elemente hat, *ich* jedoch eine Basis desselben Vektorraums zu finden in der Lage bin, die 1.000.000 Elemente hat. Dies ist noch vorstellbar – aber es ist nicht wahr. Ein wichtiges Etappenziel ist zu zeigen, dass alle Basen gleich viele Elemente haben. Zu diesem Ziel führt der folgende Hilfssatz.

Austauschlemma

Sei B eine Basis des Vektorraums V . Dann gilt: Zu jedem Vektor $w \neq o$ aus V gibt es einen Vektor $v \in B$ derart, dass

$$B' := (B \setminus \{v\}) \cup \{w\}$$

ebenfalls eine Basis von V ist.

Man sagt dazu auch, dass man in B den **Vektor** v durch w **ersetzen** kann.

Wie folgt dieser Satz? Da B ein Erzeugendensystem ist, gibt es Vektoren $v_1, v_2, \dots, v_n \in B$ mit

$$w = k_1 v_1 + k_2 v_2 + \dots + k_n v_n .$$

Da $w \neq o$ ist, muss mindestens ein $k_j \neq 0$ sein.

Sei v_j so bestimmt, dass der Koeffizient k_j in obiger Linearkombination von w ungleich Null ist. Wir definieren nun

$$B' = (B \setminus \{v_j\}) \cup \{w\} .$$

(Es ist also $v = v_j$ in der Aussage des Austauschlemmas.)

Wir müssen zeigen, dass B' eine Basis von V ist. Zuerst überzeugen wir uns, dass B' ein Erzeugendensystem ist. Dazu brauchen wir uns nur zu vergewissern, dass v_j eine Linearkombination der Vektoren aus B' ist; denn dann liegen im Erzeugnis von B' alle Vektoren der Basis B ; somit ist

$$V = \langle B \rangle \subseteq \langle B' \rangle .$$

Da natürlich $\langle B' \rangle \subseteq V$ gilt (B' ist ja eine Teilmenge von V), folgt zusammen $\langle B' \rangle = V$.

Aus der Darstellung von w ergibt sich

$$k_j v_j = w - \sum_{i \neq j} k_i v_i .$$

Da $k_j \neq 0$ ist, folgt daraus

$$v_j = \frac{1}{k_j} w - \sum_{i \neq j} \frac{k_i}{k_j} v_i$$

Nun zeigen wir noch, dass B' linear unabhängig ist. Angenommen, B' wäre linear abhängig. Dann gäbe es Skalare h_i ($i=0, 1, \dots, n, i \neq j$) mit

$$h_0 w - \sum_{i \neq j} h_i v_i = 0,$$

wobei mindestens ein $h_i \neq 0$ ist. Dann muss insbesondere $h_0 \neq 0$ sein (sonst hätten wir eine nichttriviale Linearkombination des Nullvektors aus Vektoren von B). Daher können wir durch h_0 dividieren und w einsetzen.

$$0 = w - \sum_{i \neq j} \frac{h_i}{h_0} v_i = \sum_i k_i v_i - \sum_{i \neq j} \frac{h_i}{h_0} v_i.$$

Daraus folgt

$$0 = k_j v_j + \sum_{i \neq j} \left(k_i - \frac{h_i}{h_0} \right) v_i$$

Da B linear unabhängig ist, folgt daraus insbesondere $k_j = 0$: ein Widerspruch. \square

Ergänzung zum Austauschlemma

Als Vektor v kann jeder Vektor v_j aus B gewählt werden, der einen von 0 verschiedenen Koeffizienten k_j in der Linearkombination

$$w = \sum_{i=1}^n k_i v_i$$

hat. \square

Beispiel Die Menge $B = \{v_1 = (0, 1, 1), v_2 = (1, 0, 1), v_3 = (1, 1, 0)\}$ ist eine Basis des Vektorraums \mathbf{R}^3 . Ist $w = (1, -1, 0)$, so gilt

$$w = -v_1 + v_2$$

Also kann man v_1 oder v_2 durch w ersetzen, nicht aber v_3 . (In der Tat sind v_1, v_2 und w linear abhängig.)

Nun können wir bereits einen Vektor in eine Basis einbauen; das bedeutet insbesondere, dass jeder von Null verschiedene Vektor in einer Basis von V enthalten ist.

Ist auch jede Menge von Vektoren in einer Basis enthalten? In der Form ist die Frage bestimmt zu unvorsichtig gestellt; denn jede Teilmenge einer Basis besteht aus linear unabhängigen Vektoren. Die richtige Frage lautet also: Ist jede Menge linear unabhängiger Vektoren zu einer Basis ergänzbar? Dass die Antwort darauf positiv ist, ist die wichtigste Aussage dieses Abschnitts:

Der Austauschsatz von Steinitz

Sei $B = \{v_1, \dots, v_n\}$ eine endliche Basis von V , und sei $C = \{w_1, \dots, w_m\}$ eine Menge linear unabhängiger Vektoren. Dann gibt es eine Menge von $n-m$ Vektoren in B , o.B.d.A. v_{m+1}, \dots, v_n , so dass

$$\{w_1, \dots, w_m, v_{m+1}, \dots, v_n\}$$

eine Basis von V ist.

Man kann also v_1, \dots, v_m durch die Vektoren w_1, \dots, w_m ersetzen. Insbesondere kann jede linear unabhängige Menge zu einer Basis ergänzt werden.

Insbesondere kann man in K^n jede linear unabhängige Menge von Vektoren sogar mit Einheitsvektoren zu einer Basis ergänzen.

Bevor wir den Austauschsatz von Steinitz beweisen, notieren wir einige wichtige Folgerungen.

1. Folgerung

Sei $B = \{v_1, \dots, v_n\}$ eine endliche Basis von V , und sei $C = \{w_1, \dots, w_m\}$ eine Menge linear unabhängiger Vektoren. Dann gilt $m \leq n$, das heißt $|C| \leq |B|$.

Beweis Nach dem Austauschsatz von Steinitz kann C durch $n-m$ Vektoren aus B zu einer Basis ergänzt werden. Also muss $n-m$ eine nichtnegative ganze Zahl sein. Das heißt $m \leq n$.

□

2. Folgerung

Jede linear unabhängige Menge eines endlich erzeugbaren Vektorraums, insbesondere jede Basis, hat nur endlich viele Elemente.

Zum *Beweis* dieser Folgerung nehmen wir an, es gäbe eine linear unabhängige Menge C_∞ mit unendlich vielen Elementen. Sei B eine endliche Basis mit n Elementen. In C_∞ gäbe es eine endliche Teilmenge C mit $n+1$ Elementen. Nach der 1. Folgerung wäre dann

$$n + 1 = |C| \leq |B| = n ,$$

was offensichtlich falsch ist. □

Daher können wir in der Voraussetzung des Steinitzschen Austauschsatzes das Wort „endliche“ in „Sei B eine endliche Basis von $V \dots$ “ weglassen, da diese Voraussetzung nach unserer allgemeinen Vereinbarung erfüllt ist.

3. Folgerung

Je zwei Basen von V haben gleich viele Elemente.

Seien dazu B_1 und B_2 Basen von V . Da nach der 2. Folgerung B_1 und B_2 endlich sind, kann sowohl B_1 als auch B_2 die Rolle von B – und entsprechend B_2 und B_1 die Rolle von C – in der 1. Folgerung übernehmen. Es folgt

$$|B_2| \leq |B_1| \text{ und } |B_1| \leq |B_2| ,$$

also $|B_1| = |B_2|$. □

Die Zahl der Elemente einer Basis eines Vektorraums V ist dessen wichtigster Parameter. Man nennt ihn die **Dimension** von V und schreibt für diese Zahl $\dim(V)$. Da die Dimension eines endlich erzeugbaren Vektorraums eine nichtnegative ganze Zahl ist, heißen solche Vektorräume auch **endlich-dimensional**. Wenn nicht anders gesagt, ist jeder auftretende Vektorraum endlich-dimensional.

4. Folgerung

Sei $\dim(V) = n$. Dann gilt: Jede linear unabhängige Menge von Vektoren hat höchstens n Elemente. Wenn eine linear unabhängige Menge genau n Elemente hat, ist sie eine Basis.

Der erste Teil folgt direkt aus der 1. Folgerung.

Zum zweiten Teil: Angenommen, es gäbe eine linear unabhängige Menge C mit n Vektoren, die keine Basis ist. Dann kann C keine maximale linear unabhängige Menge sein. Also könnte sie erweitert werden zu einer linear unabhängigen Menge mit $n+1$ Vektoren. Dies steht im Widerspruch zum ersten Teil der Folgerung. □

5. Folgerung

Jede Basis eines Unterraums von V hat höchstens n Elemente. Mit anderen Worten: Die Dimension eines Unterraums ist höchstens so groß wie die Dimension des Gesamttraums. \square

6. Folgerung

Jede Basis eines Unterraums von V kann zu einer Basis von V ergänzt werden.

Beide Folgerungen ergeben sich direkt aus der Tatsache, dass eine Basis eines Unterraums eine Menge linear unabhängiger Vektoren von V ist. \square

7. Folgerung

Seien U_1 und U_2 Unterräume von V mit $U_1 \subseteq U_2$. Dann ist $\dim(U_1) \leq \dim(U_2)$. Gleichheit gilt genau dann, wenn $U_1 = U_2$ ist.

Dies folgt aus Folgerung 5, wenn man U_2 die Rolle von V spielen lässt. \square

Sie sehen: der Steinitzsche Austauschsatz gibt uns ganz direkt entscheidende Erkenntnisse über die Struktur eines Vektorraums. Außerdem werden wir später noch weitere Folgerungen daraus ziehen. Nun müssen wir diesen Satz aber *beweisen*.

Seien die Bezeichnungen so gewählt wie im Austauschsatz von Steinitz beschrieben. Wir beweisen den Satz durch Induktion nach m .

Sei zunächst $m = 1$. Dann folgt die Behauptung direkt aus dem Austauschlemma.

Sei nun $m > 1$, und die Aussage sei richtig für $m-1$. Wir betrachten eine linear unabhängige Menge $C = \{w_1, \dots, w_m\}$ aus genau m Vektoren.

Sei $C' := C \setminus \{w_m\}$. Dann besteht C' nur aus den $m-1$ Elementen w_1, \dots, w_{m-1} , und wir können daher die Induktionsvoraussetzung auf C' anwenden.

Daher gibt es eine Menge $\{v_m, \dots, v_n\}$ von Vektoren aus B derart, dass

$$\{w_1, \dots, w_{m-1}, v_m, \dots, v_n\}$$

eine Basis von V ist. Wir müssen noch den Vektor w_m in die Basis

$$\{w_1, \dots, w_{m-1}, v_m, \dots, v_n\}$$

„einbauen“. Dazu müssen wir einen der Vektoren v_m, \dots, v_n durch w_m ersetzen. Um das zu tun, stellen wir w_m als Linearkombination der Vektoren der neuen Basis dar:

$$w_m = k_1 w_1 + \dots + k_{m-1} w_{m-1} + k_m v_m + \dots + k_n v_n .$$

Angenommen, alle k_i mit $i \geq m$ wären Null. Dann wäre

$$w_m = k_1 w_1 + \dots + k_{m-1} w_{m-1}$$

also wäre $\{w_1, \dots, w_{m-1}, w_m\}$ linear abhängig: ein Widerspruch.

Also gibt es ein $i \geq m$ mit $k_i \neq 0$, sei o. B. d. A. $k_m \neq 0$. Dann ist nach dem Austauschlemma auch

$$\{w_1, \dots, w_{m-1}, w_m, v_{m+1}, \dots, v_n\}$$

eine Basis.

Damit ist der Austauschsatz von Steinitz bewiesen. □

Wir beenden diesen Abschnitt mit einer Folgerung, die sich oft als ausgesprochen nützlich erweist.

Seien U_1 und U_2 zwei Unterräume von V . Dann heißt U_1 ein **Komplement** von U_2 , falls

- der Durchschnitt von U_1 und U_2 „so klein wie möglich“ ist:

$$U_1 \cap U_2 = \{0\}$$

und

- das Erzeugnis von U_1 und U_2 „so groß wie möglich“ ist:

$$\langle U_1, U_2 \rangle = V .$$

Wir nennen dann U_1 und U_2 auch **komplementär**. (Offenbar gilt: Ist U_1 ein Komplement von U_2 , so ist auch U_2 ein Komplement von U_1 .)

8. Folgerung

Jeder Unterraum von V hat einen komplementären Unterraum.

Zum *Beweis* betrachten wir einen beliebigen Unterraum U_1 von V . Sei $B_1 = \{v_1, \dots, v_m\}$ eine Basis von U_1 . Nach der 6. Folgerung kann B_1 zu einer Basis

$$B = \{v_1, \dots, v_m, v_{m+1}, \dots, v_n\}$$

von V ergänzt werden. Definiere

$$U_2 := \langle v_{m+1}, \dots, v_n \rangle .$$

Behauptung U_1 und U_2 sind komplementär.

„ $U_1 \cap U_2 = \{o\}$ “: Sei $v \in U_1 \cap U_2$. Dann ist

$$v = k_1 v_1 + \dots + k_m v_m \in U_1$$

und

$$v = k_{m+1} v_{m+1} + \dots + k_n v_n \in U_2$$

Also ist

$$k_1 v_1 + \dots + k_m v_m - k_{m+1} v_{m+1} - \dots - k_n v_n = v - v = o .$$

Da B eine Basis ist, müssen alle $k_i = 0$ sein; also ist $v = o$. Das bedeutet: Der Nullvektor ist der einzige Vektor in $U_1 \cap U_2$.

„ $\langle U_1, U_2 \rangle = V$ “: In $\langle U_1, U_2 \rangle$ liegen bestimmt die erzeugenden Vektoren v_1, \dots, v_m von U_1 und die erzeugenden Vektoren v_{m+1}, \dots, v_n von U_2 . Also ist B in $\langle U_1, U_2 \rangle$ enthalten, und damit gilt $\langle U_1, U_2 \rangle = V$. \square

Beispiele

- (a) Im \mathbf{R}^2 sind je zwei verschiedene 1-dimensionale Unterräume komplementär.
- (b) Im \mathbf{R}^3 ist jeder 2-dimensionale Unterraum zu jedem nicht in ihm enthaltenen 1-dimensionalen Unterraum komplementär.

Es hält sich hartnäckig das Gerücht, dass ein Unterraum eines Vektorraums *nur ein* Komplement habe. Das ist fast immer völlig falsch. Betrachten wir dazu etwa den Vektorraum \mathbf{R}^2 . Dann hat die „ y -Achse“ $Y = \{(0, y) \mid y \in \mathbf{R}\}$ als ein Komplement zum Beispiel die „ x -Achse“ $X = \{(x, 0) \mid x \in \mathbf{R}\}$; aber es gibt viele andere Komplemente, etwa den Unterraum $Z = \{(x, x) \mid x \in \mathbf{R}\}$ – ja *jeder* 1-dimensionale Unterraum $\neq Y$ von \mathbf{R}^2 ist ein Komplement von Y . (Vergleichen Sie hierzu auch Übungsaufgabe 27.)

In der Aussage des folgenden Satzes steckt die Bedeutung komplementärer Unterräume.

Zerlegung eines Vektors bezüglich komplementärer Unterräume

Seien U_1, U_2 komplementäre Unterräume. Dann lässt sich jeder Vektor $v \in V$ eindeutig als Summe eines Vektors $u_1 \in U_1$ und eines Vektors $u_2 \in U_2$ schreiben:

$$v = u_1 + u_2 \text{ für eindeutige } u_1 \in U_1, u_2 \in U_2 .$$

Die Existenz von u_1 und u_2 ergibt sich direkt aus der Definition des Erzeugnisses: Da U_1 und U_2 den Vektorraum V erzeugen, kann jeder Vektor aus V als Linearkombination von Vektoren aus U_1 und U_2 , also als Summe eines Vektors aus U_1 und eines Vektors aus U_2 dargestellt werden.

Der Nachweis der *Eindeutigkeit* beruht auf einem *Trick*, den es sich zu merken lohnt.

Sei $v = u_1 + u_2$ und $v = w_1 + w_2$ mit $u_1, w_1 \in U_1$ und $u_2, w_2 \in U_2$. Dann ist

$$u_1 + u_2 = v = w_1 + w_2$$

also

$$u_1 - w_1 = w_2 - u_2.$$

In dieser Gleichung ist die linke Seite die Differenz zweier Vektoren aus U_1 , also liegt sie in U_1 ; die rechte Seite ist Differenz zweier Vektoren aus U_2 , also ist sie in U_2 . Die Gleichung stellt also einen Vektor dar, der sowohl in U_1 als auch in U_2 liegt; dieser muss der Nullvektor sein, da $U_1 \cap U_2 = \{0\}$ ist.

Somit folgt $u_1 = w_1$ und $u_2 = w_2$. Das bedeutet, dass u_1 und u_2 eindeutig bestimmt sind. \square

In einem Vektorraum der Dimension n heißen die Unterräume der Dimension $n-1$ **Hyper-ebenen**. Das ist eine relativ ungefährliche Bezeichnung. Ich warne aber davor, die Vektoren grundsätzlich als „Punkt“ und die 1-dimensionalen Unterräume grundsätzlich als „Geraden“ zu (denken und) bezeichnen!

Es gibt Modelle für gewisse geometrische Strukturen („affine Räume“), in denen die Punkte tatsächlich die Vektoren eines Vektorraums und gewisse Geraden die 1-dimensionalen Unterräume sind (siehe Abschn. 4.1). Es gibt jedoch auch Geometrien, die ganz anders aus Vektorräumen konstruiert werden. Und es gibt Anwendungen von Vektorräumen, in denen die Begriffe „Punkte“ und „Geraden“ gar keinen Sinn machen. Der entscheidende Fortschritt moderner mathematischer Begriffsbildung besteht ja gerade darin, dass die Begriffe unabhängig von konkreten Modellen sind!

Dies soll kein Schock sein und Sie keinesfalls davon abhalten, sich etwas vorzustellen – im Gegenteil! Um mathematische Beweise zu finden, muss man oft sehr konkrete Vorstellungen entwickeln. Der \mathbf{R}^3 ist ein wichtiger Vektorraum, den Sie sich gut vorstellen können müssen. Aber: Der \mathbf{R}^3 ist nicht der einzige Vektorraum! Lernen Sie, sich auch andere Vektorräume vorzustellen, etwa den Vektorraum aus Abschn. 3.2.8 mit der symmetrischen Differenz oder einen Vektorraum, der aus Lösungen eines Gleichungssystems besteht (Abschn. 4.2), oder einen Code (Abschn. 4.3), oder ...

Eine Schlussbemerkung zu diesem Abschnitt. Da wir uns zu Beginn sehr anstrengen mussten, um die Existenz einer Basis nachzuweisen, könnten Sie zu der Ansicht gelangen, eine Basis sei ein besonders rares Objekt. Dieser Eindruck ist falsch! Wir zeigen später, dass fast jede zufällig gewählte n -elementige Menge von Vektoren eine Basis ist.

3.3.3 Der Dimensionssatz

Sind zwei Unterräume eines Vektorraums gegeben, so wird (und muss) man häufig die Dimension ihres Erzeugnisses berechnen. Diese Dimension hängt nicht nur von den Dimensionen der einzelnen Unterräume, sondern auch von deren Durchschnitt ab: Je kleiner der Durchschnitt, desto größer das Erzeugnis. Das Modell für den Dimensionssatz ist die Summenformel für die Mächtigkeit der Vereinigung zweier endlicher Mengen (vgl. Abschn. 1.1):

$$|X \cup Y| = |X| + |Y| - |X \cap Y|.$$

Im Dimensionssatz wird die Mächtigkeit durch die Dimension des entsprechenden Unterraums ersetzt (das ist das richtige Maß für die Größe eines Unterraums). Da auch der Durchschnitt zweier Unterräume ein Unterraum ist (siehe Übungsaufgabe 28), kann man mit Hilfe der Dimension auch die Größe des Durchschnitts messen. Die Vereinigung von Mengen muss dabei allerdings durch das Erzeugnis von Unterräumen ersetzt werden. (*Achtung!* Die mengentheoretische Vereinigung zweier Unterräume ist nämlich nur in den seltensten Fällen wieder ein Unterraum – vergleichen Sie hierzu Übungsaufgabe 29).

Dimensionssatz

Seien U_1 und U_2 Unterräume des Vektorraums V . Dann gilt

$$\dim(\langle U_1, U_2 \rangle) = \dim(U_1) + \dim(U_2) - \dim(U_1 \cap U_2).$$

Der *Beweis* beruht auf folgender Idee:

Wir wählen eine Basis $B = \{v_1, \dots, v_m\}$ von $U_0 := U_1 \cap U_2$ und ergänzen diese einerseits zu einer Basis

$$B_1 = \{v_1, \dots, v_m, v_{m+1}, \dots, v_r\}$$

von U_1 , andererseits zu einer Basis

$$B_2 = \{v_1, \dots, v_m, w_{m+1}, \dots, w_t\}$$

von U_2 . Dann betrachten wir die Menge

$$C := \{v_1, \dots, v_m, v_{m+1}, \dots, v_r, w_{m+1}, \dots, w_t\} = B_1 \cup B_2$$

von Vektoren. Für die Anzahl der Elemente von C berechnet man

$$|C| = r + t - m = \dim(U_1) + \dim(U_2) - \dim(U_1 \cap U_2).$$

Wenn wir zeigen könnten, dass C eine Basis von $\langle U_1, U_2 \rangle$ ist, dann hätten wir den Dimensionssatz bewiesen! Also ans Werk:

Zunächst ist klar, dass C den ganzen Raum $\langle U_1, U_2 \rangle$ erzeugt, denn es ist $C = B_1 \cup B_2$, und B_1 erzeugt U_1 , und B_2 erzeugt U_2 .

Es ist etwas schwieriger, die *lineare Unabhängigkeit* von C nachzuweisen. Wir betrachten eine nichttriviale Linearkombination, die den Nullvektor liefert:

$$k_1 v_1 + \dots + k_m v_m + k_{m+1} v_{m+1} + \dots + k_r v_r + h_{m+1} w_{m+1} + \dots + h_t w_t = 0 .$$

Daraus folgt

$$-k_1 v_1 - \dots - k_m v_m - k_{m+1} v_{m+1} - \dots - k_r v_r = h_{m+1} w_{m+1} + \dots + h_t w_t \in U_1 \cap U_2 = U_0 .$$

Also gibt es Skalare a_1, \dots, a_m mit

$$h_{m+1} w_{m+1} + \dots + h_t w_t = a_1 v_1 + \dots + a_m v_m .$$

Da aber die Menge $\{v_1, \dots, v_m, w_{m+1}, \dots, w_t\}$ linear unabhängig ist, folgt daraus $h_{m+1} = \dots = h_t = 0$.

Dann ist die Ausgangs-Linearkombination aber eine Linearkombination von Vektoren aus B_1 und somit sind alle Koeffizienten gleich 0.

Also ist C linear unabhängig. Damit ist der Dimensionssatz bewiesen. \square

Wir werden diesen Satz sehr häufig anwenden. Zwei einfache Folgerungen seien jetzt schon gezogen:

Folgerung 1

Sei H eine Hyperebene, und sei U ein t -dimensionaler Unterraum von V . Dann gilt

$$U \subseteq H \text{ oder } \dim(U \cap H) = t - 1 .$$

Wegen $H \subseteq \langle H, U \rangle \subseteq V$ ist nämlich

$$n - 1 = \dim(H) \leq \dim(U, H) \leq \dim(V) = n .$$

Also ist $\dim(U, H) = n - 1$ oder $\dim(U, H) = n$. Im ersten Fall ist $U \subseteq H$. Im zweiten Fall ergibt sich mit dem Dimensionssatz

$$\dim(U \cap H) = \dim(U) + \dim(H) - \dim(U, H) = t + n - 1 - n = t - 1 .$$

\square

Folgerung 2

Sind U_1, U_2 komplementäre Unterräume, so gilt

$$\dim(U_1) + \dim(U_2) = n$$

Wegen $\dim(U_1 \cap U_2) = 0$ und $\dim(U_1, U_2) = n$ folgt dies direkt aus der Dimensionsformel. \square

3.3.4 Faktorräume

Faktorräume sind meiner Erfahrung nach ein Schreckgespenst für die Studierenden – obwohl es sich um ein wichtiges und nicht zu schwieriges Konzept handelt. In vielen Situationen, in denen man einen Unterraum U eines Vektorraums V betrachtet, kann man die relevanten Überlegungen auf die „Struktur außerhalb von U “ reduzieren. Das adäquate Instrument hierzu ist der Faktorraum V/U .

Doch wir fangen ganz langsam an. Wir betrachten einen Unterraum U des K -Vektorraums V . Für einen Vektor $v \in V$ bezeichnen wir mit

$$v + U := \{v + u \mid u \in U\}$$

die **Nebenklasse** von U durch v . Man nennt v einen **Repräsentanten** der Nebenklasse $v + U$.

Wir erhalten also die Nebenklasse $v + U$ („ v plus U “), indem wir v zu jedem Vektor aus U addieren; die so erhaltene Menge von Vektoren ist $v + U$.

Was für ein Objekt ist eine Nebenklasse? Sicherlich eine Menge von Vektoren. Ist eine Nebenklasse ein Unterraum? Probieren wir's: Wenn die Nebenklasse $v + U$ ein Unterraum sein soll, muss der Nullvektor in ihr enthalten sein. Das heißt, es muss ein $u \in U$ geben mit

$$0 = v + u.$$

Das ist gleichwertig damit, dass $v = -u \in U$ ist.

Das bedeutet: Wenn $v + U$ ein Unterraum ist, muss der Repräsentant v in U liegen.

Sei umgekehrt $v \in U$. Dann ist

$$v + U = U.$$

(Denn sicher ist $v + U \subseteq U$. Sei umgekehrt $u \in U$. Dann ist

$$u = v + (u - v) \in v + U,$$

da $u - v \in U$ ist.).

Also haben wir gezeigt: *Eine Nebenklasse $v+U$ ist genau dann ein Unterraum (und dann gleich U), wenn ihr Repräsentant v in U liegt.* \square

In der Regel ist eine Nebenklasse also *kein* Unterraum. Welche Rolle spielen die Repräsentanten? Ist „der“ Repräsentant einer Nebenklasse eindeutig oder gibt es viele? Wenn ja, welches sind die anderen Repräsentanten einer Nebenklasse? Die Frage ist also: Gibt es Vektoren $v' \neq v$ mit

$$v' + U = v + U?$$

Auch diese Frage können wir durch einfache Analyse beantworten: Wegen $o \in U$ ist $v' = v' + o \in v' + U$. Wenn also $v' + U = v + U$ ist, so muss

$$v' \in v' + U = v + U$$

gelten. Also muss es ein $u_0 \in U$ geben mit $v' = v + u_0$. Das bedeutet $v' - v = u_0 \in U$.

Umgekehrt: Sei $v' - v \in U$. Dann definieren wir $u_0 := v' - v$ und schreiben

$$v' = v + u_0 \in v + U,$$

also

$$v' + U = \{v' + u \mid u \in U\} = \{v + u_0 + u \mid u \in U\} \subseteq \{v + u_1 \mid u_1 \in U\} = v + U.$$

Aber dieses Argument kann man auch für die andere Inklusion benutzen: Wegen $v = v' - u_0 \in v' + U$ folgt entsprechend

$$v + U \subseteq v' + U,$$

also $v + U = v' + U$.

Damit haben wir das **Kriterium für die Gleichheit von Nebenklassen** bewiesen:

$$v + U = v' + U \Leftrightarrow v' - v \in U.$$

\square

Merken Sie sich dieses Kriterium; wir werden es nicht nur in diesem Buch zig mal verwenden, sondern es besteht eine relativ hohe Wahrscheinlichkeit dafür, dass Sie es in einer Klausur oder einer mündlichen Prüfung wissen müssen.

Man kann diese Aussage auch anders ausdrücken: *Die Repräsentanten einer Nebenklasse sind genau die Elemente dieser Nebenklasse.*

Beweis Wenn $v' \in v + U$ gilt, so ist $v' = v + u$ mit $u \in U$, also $v' - v \in U$; nach obiger Aussage ist v' also ein Repräsentant. Wenn umgekehrt v' ein Repräsentant ist, so ist $v' - v = u \in U$, also $v' = v + u \in v + U$. \square

Nebenklassen sind also außerordentlich demokratische Institutionen; jedes Mitglied ist in gleicher Weise fähig, die Gesamtheit zu repräsentieren.

Beispiel Sei $V = \mathbf{R}^2$, und sei $U = \langle (1, 1) \rangle$. Was ist $(1, 0) + \langle (1, 1) \rangle$? Ganz einfach:

$$(1, 0) + \langle (1, 1) \rangle = \{(1, 0) + (x, x) | x \in \mathbf{R}\} = \{(1 + x, x) | x \in \mathbf{R}\}.$$

Bislang haben wir *eine einzige* Nebenklasse betrachtet; nun studieren wir das *Zusammenspiel aller* Nebenklassen bezüglich eines Unterraums U . Zuerst stellen wir die Frage: Können sich zwei Nebenklassen X und Y überlappen? Kann es passieren, dass X und Y einen gemeinsamen Vektor z enthalten?

Ist $X = v + U$ und $Y = w + U$, so müssten die folgenden Gleichungen gelten

$$z = v + u \text{ und } z = w + u' \text{ mit } u, u' \in U.$$

Dann ist $v + u = w + u'$, also $v - w = u' - u \in U$; nach obigem Kriterium bedeutet dies $v + U = w + U$, also $X = Y$.

Bemerkung Wir haben damit bewiesen, dass die Äquivalenzklassen der folgendermaßen definierten Relation \sim

$$x \sim y : \Leftrightarrow x - y \in U$$

eine Partition von V bilden. (Vergleichen sie dazu den Abschn. 1.2.)

Damit haben wir folgende bemerkenswerte Tatsache gezeigt:

Lemma über verschiedene Nebenklassen

Verschiedene Nebenklassen sind disjunkt.

Genauer gilt: Die Menge der Vektoren von V ist die disjunkte Vereinigung der Nebenklassen von U .

Man darf sich die Menge der Nebenklassen nach U als „Parallelschar“ vorstellen. □

Wie viele Nebenklassen von U gibt es in V ? Wir beantworten diese Frage nur für Vektorräume über endlichen Körpern, wo die Frage eigentlich auch nur Sinn macht. Dazu müssen wir zuerst die Frage behandeln, wie viele Vektoren eine Nebenklasse besitzt.

Anzahl der Vektoren einer Nebenklasse

Sei V ein endlicher Vektorraum über einem Körper K mit q Elementen, und sei U ein k -dimensionaler Unterraum von V . Dann besteht jede Nebenklasse $v + U$ von V nach U aus genau q^k Vektoren.

Wir machen uns zunächst den einfachsten Fall, nämlich die Nebenklasse U , klar. Hier ist die Situation im Grunde schon geklärt: U ist ein k -dimensionaler Unterraum von V , also ein k -dimensionaler Vektorraum über einem endlichen Körper mit q Elementen. Nach dem Satz über die Anzahl der Elemente eines endlichen Vektorraums hat U also q^k Elemente.

Wir müssen demnach nur noch zeigen, dass jede Nebenklasse $v+U$ gleichviel Elemente wie U hat. Dazu genügt es, eine bijektive Abbildung von U auf $v+U$ anzugeben. Nichts einfacher als das: Sei f die Abbildung, die durch

$$f(u) := v + u \text{ für alle } u \in U$$

definiert ist. Direkt aus der Definition folgt, dass f surjektiv ist. Fast genauso schnell folgt, dass f injektiv ist: Seien $u, u' \in U$ mit $f(u) = f(u')$. Das heißt $v+u = v+u'$, also $u = u'$.

Also ist f bijektiv, und daher hat $v+U$ gleichviel Elemente wie U , also genau q^k . \square

Anzahl der Nebenklassen

Sei V ein endlicher n -dimensionaler Vektorraum über einem Körper K mit q Elementen, und sei U ein k -dimensionaler Unterraum von V . Dann gibt es genau q^{n-k} Nebenklassen von U in V .

Wir wissen, dass V die disjunkte Vereinigung der Nebenklassen von U ist und dass jede Nebenklasse genau q^k Elemente hat. Da V genau q^n Vektoren besitzt, folgt die Behauptung. \square

Wir definieren: Die Menge aller Nebenklassen von U in V heißt der **Faktorraum von V nach U** ; dieser wird mit V/U (gesprochen „ V nach U “) bezeichnet:

$$V/U = \{v + U \mid v \in V\}.$$

Wie erhält man den Faktorraum V/U ? Man bildet sämtliche Nebenklassen $v+U$; die Menge all dieser Nebenklassen ist V/U . [Beachten Sie: In der Menge aller Nebenklassen kommt jede Nebenklasse nur einmal vor; auch wenn sie von verschiedenen Repräsentanten dargestellt werden kann.]

Warum heißt diese Menge „Raum“? Der Verdacht liegt nahe, dass dies deswegen so ist, weil V/U wieder ein Vektorraum ist. Dies ist zunächst schwer vorstellbar, aber – es ist richtig:

Satz vom Faktorraum

Der Faktorraum V/U eines K -Vektorraums V nach einem Unterraum U ist ein K -Vektorraum.

Vor dem Beweis dieses Satzes die wichtigste Regel im Umgang mit Faktorräumen: *Auch die Elemente eines Faktorraums sind zunächst einfach Vektoren!* Obwohl seine Elemente nicht v heißen, sondern merkwürdige (und umständlich anmutende) Namen wie zum Beispiel $v+U$ tragen, sind sie – auf einer gewissen Ebene – einfach Vektoren, und zwar des Vektorraums V/U . Nur manchmal, wenn man's ganz genau wissen will, muss man sozusagen eine Ebene tiefer hinabsteigen, und sich kurz klar machen, was diese Vektoren „in Wirklichkeit“ sind (nämlich Nebenklassen von V); diese tiefere Ebene betritt man aber nur, wenn es nicht anders geht – und verlässt sie so schnell wie möglich wieder.

Ich möchte Ihnen das an einer vergleichbaren Situation des täglichen Lebens verdeutlichen: Ein normaler Verein, wie etwa der Sportverein Ihres Heimatortes, hat als Mitglieder normale („natürliche“) Personen. Es gibt aber auch Vereine, deren Mitglieder keine natürlichen Personen sind, sondern wieder Vereine; dazu gehört zum Beispiel der Deutsche Fußball Bund, dessen Mitglieder die einzelnen Fußballvereine sind. Man nennt solche Mitglieder oft auch „juristische Personen“; das bedeutet, dass diese im Regelfall so behandelt werden wie natürliche Personen und man nur manchmal auf die eigentliche Natur dieser juristischen Personen Rücksicht nehmen muss.

Es ist immer gut, diese beiden Ebenen genau zu unterscheiden. Machen Sie sich in den folgenden Beweisen jeweils klar, auf welcher Ebene Sie sich befinden.

Zum Beweis des Satzes vom Faktorraum müssen wir einfach die Vektorraumaxiome nachweisen. Also zuerst die Axiome der Addition. Dazu müssen wir ... ??? Wie bitte? Halt! Keiner verlässt das Lokal!

Was soll denn überhaupt $+$ und \cdot sein? Wie addiert man eine Nebenklasse zu einer anderen? Zwei riesige Goliaths, wo wir doch nur Vektoren, also bescheidene Davids, addieren können??? – Aber vielleicht kommen wir ja damit aus; wir probieren's einfach:

Seien X und Y Elemente von V/U , also Nebenklassen nach U . Es sei

$$X + Y := x + y + U \text{ für } x \in X, y \in Y.$$

Wir halten die Luft an und betrachten diese Definition. Leistet sie das Gewünschte? Ja, denn je zwei Nebenklassen X und Y wird als Summe wieder eine Nebenklasse zugeordnet. Wie berechne ich die Summe $X+Y$? Dazu muss ich ein $x \in X$ und ein $y \in Y$ wählen, die beiden addieren (das ist möglich, da x und y Vektoren sind) und die zugehörige Nebenklasse bilden.

Alles klar? Na, dann versuchen Sie's doch mal: Sie nehmen ein Element aus X ... Was, Sie haben nicht *mein* x gewählt, sondern ein *eigenes* x' ! Dürfen Sie das überhaupt? – Tatsächlich, das ist nicht verboten, und was nicht verboten ist, ist erlaubt. Nun wählen Sie wahrscheinlich, ich seh's schon kommen, ja wirklich, Sie wählen auch ein $y' \in Y$ mit $y' \neq y$. Dann bilden Sie die Nebenklasse $x' + y' + U$ und behaupten, das sei die Summe von X und Y .

Wer von uns hat nun Recht? – Eines ist jedenfalls klar: Wenn die Definition einen Sinn machen soll, dann haben wir beide recht, denn der Wert einer Summe darf bestimmt nicht davon abhängen, *wer* sie ausrechnet. Mathematisch gesprochen ist also unsere gemeinsa-

me Aufgabe zu zeigen, dass die Nebenklasse $x + y + U$ unabhängig von der Auswahl der Vektoren $x \in X$ und $y \in Y$ ist. Noch präziser ausgedrückt: Wir müssen zeigen, dass

$$x + y + U = x' + y' + U$$

ist für $x, x' \in X$ und $y, y' \in Y$.

Wenn die Aufgabe einmal so präzise gestellt ist, ist sie auch nicht mehr schwer zu lösen. Da x und x' aus derselben Nebenklasse sind, ist $u := x' - x$ ein Element aus U . Genauso folgt $y' - y =: w \in U$. Damit ergibt sich

$$x' + y' = x + u + y + w = x + y + u + w,$$

also

$$(x' + y') - (x + y) = x' - x + y' - y = u + w \in U.$$

Indem wir zum dritten Mal das Kriterium für die Gleichheit von Nebenklassen anwenden (wo waren die anderen beiden Anwendungen?), erhalten wir

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

also die Behauptung.

Uff! Mit viel Geschick haben wir diese Klippe umschifft. Die nächste ist zwar schon in Sicht, aber die kann uns nicht mehr schrecken. Wir machen kurzen Prozess:

Für eine Nebenklasse X und ein Element $k \in K$ sei

$$k \cdot X := k \cdot x + U \quad \text{für } x \in X.$$

Wieder muss man zeigen, dass diese Definition unabhängig von der Auswahl des speziellen Repräsentanten $x \in X$ ist: Seien also x und x' zwei Elemente aus X . Dann ist $x' - x \in U$, also auch $k(x' - x) \in U$, da U ein Unterraum ist. Aus $kx' - kx \in U$ folgt nun wieder

$$kx' + U = kx + U.$$

Übrigens: Wenn man nachweisen muss, dass eine Definition unabhängig von der Auswahl der Repräsentanten ist, sagt man dazu auch, man zeigt, dass die Vorschrift **wohldefiniert** ist.

Es könnte scheinen, dass wir erst am Anfang, ja vielleicht noch vor Beginn des Nachweises der Vektorraumaxiome sind. Dies ist zwar formal richtig, aber die Axiome ergeben sich nun völlig routinemäßig (Übungsaufgabe 31), weil man nun alle wünschenswerten Eigenschaften von V/U auf die (nach Voraussetzung gegebenen) Eigenschaften von V zurückführen kann. Wir halten hier nur eine, besonders wichtige Tatsache fest: Der Nullvektor des Faktorraums V/U ist die Nebenklasse $U (= o + U)$. \square

Nun wissen wir, dass V/U ein Vektorraum ist. Aus Abschn. 3.3.2 wissen wir, dass jeder Vektorraum eine Dimension hat. Welche Dimension hat der Vektorraum V/U ? Die Antwort darauf gibt die folgende Aussage, die oft auch als **Zweiter Dimensionssatz** bezeichnet

wird:

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

Zum *Beweis* müssen wir eine Basis von V/U mit der „richtigen“ Anzahl von Elementen angeben. Sei dazu $B_0 = \{v_1, \dots, v_s\}$ eine Basis von U , die wir zu einer Basis $B = \{v_1, \dots, v_s, v_{s+1}, \dots, v_n\}$ von V ergänzen.

Behauptung Dann ist $B^* := \{v_{s+1} + U, \dots, v_n + U\}$ eine Basis von V/U .

Wenn wir das gezeigt haben, sind wir fertig; denn dann ist

$$\dim(V/U) = |B^*| = n - s = \dim(V) - \dim(U) .$$

Der Nachweis, dass B^* eine Basis von V/U ist, besteht aus relativ einfachem Rechnen mit Nebenklassen. Zunächst zeigen wir, dass B^* *linear unabhängig* ist: Sei

$$k_{s+1}(v_{s+1} + U) + \dots + k_n(v_n + U) = U .$$

(Erinnern Sie sich daran, dass U das Nullelement von V/U ist.) Daraus ergibt sich

$$k_{s+1}v_{s+1} + \dots + k_nv_n + U = U ,$$

also nach dem Kriterium über die Gleichheit von Nebenklassen

$$k_{s+1}v_{s+1} + \dots + k_nv_n \in U .$$

Dann ist dieser Vektor aber auch eine Linearkombination von v_1, \dots, v_s , und es folgt $k_{s+1} = \dots = k_n = 0$, da B eine Basis ist.

Nun zeigen wir noch, dass B^* den Vektorraum V/U *erzeugt*. Sei dazu $v+U$ ein beliebiges Element von V/U . Wir stellen zunächst den Repräsentanten v als Linearkombination der Vektoren aus B dar:

$$v = k_1v_1 + \dots + k_sv_s + k_{s+1}v_{s+1} + \dots + k_nv_n .$$

Da

$$k_1v_1 + \dots + k_sv_s$$

ein Element von U ist, folgt

$$v + U = k_{s+1}v_{s+1} + \dots + k_nv_n + U .$$

Also ist $v+U$ tatsächlich eine Linearkombination der Elemente von B^* . □

Schlussbemerkung Ich hoffe, dass Sie meine Begeisterung über Faktorräume teilen. Dennoch sehe ich mich genötigt, Ihnen mitzuteilen, dass es (zur Not) auch ohne geht. Man

kann in der Theorie der Vektorräume alles, was man mit Faktorräumen ausdrücken kann, auch mit Komplementen ausdrücken. Dies sei noch kurz angedeutet.

Sei U ein Unterraum von V , und sei W ein Komplement von U in V . Dann schneidet jede Nebenklasse von U den Unterraum W in genau einem Vektor.

Sei dazu $X = v + U$ eine beliebige Nebenklasse. Wenn $w \in W$ und $u \in U$ die Vektoren mit $v = w + u$ sind, dann ist

$$X = v + U = w + u + U = w + U.$$

Also schneidet X den Unterraum W zumindest in dem Vektor w .

Angenommen, es gäbe noch einen Vektor $w' \in W \cap X$. Da $w', w \in X$ sind, liegt $w' - w$ in U ; da andererseits w' und w in dem Unterraum W enthalten sind, muss $w' - w$ aus W sein. Zusammen folgt

$$w' - w \in U \cap W = \{0\},$$

also $w' = w$.

Damit ist die Abbildung

$$f : X \rightarrow X \cap W$$

eine Bijektion, mit der man V/U auf W strukturtreu übertragen kann (und umgekehrt). In Kap. 5 werden wir solche „Vektorraumisomorphismen“ genau studieren.

Wenn man ohne Faktorraum V/U auskommen möchte, kann man also stets den Unterraum W benutzen. (Ein vielleicht eher theoretisches Problem dabei ist, dass man ein neues, eigentlich „unnötiges“ Objekt einführen muss, das auch nicht eindeutig bestimmt ist.) ... und spätestens, wenn man die entsprechende Struktur bei Gruppen (Kap. 9) einführen möchte, führt kein Weg mehr an der Bildung der Faktorstruktur vorbei!

3.4 Zur Geschichte der linearen Algebra

Die Geburt der modernen linearen Algebra kann man vergleichsweise genau datieren: In den Jahren 1843 und 1844 fanden zwei wichtige Ereignisse statt, die beide für die lineare Algebra entscheidend waren. Das erste Ereignis kennen wir schon: Die Entdeckung der Quaternionen durch William Rowan Hamilton am 16. Oktober 1843. Übrigens geht auch die Bezeichnung „Vektor“ auf Hamilton zurück.

Das zweite Ereignis fand im Jahre 1844 statt und war das Erscheinen des Buches *Die lineale Ausdehnungslehre dargestellt und durch Anwendungen auf die übrigen Zweige der Mathematik, wie auch auf die Statik, Mechanik, die Lehre vom Magnetismus und die Krytallonomie erläutert* von Hermann Graßmann, Lehrer an der Friedrich-Wilhelms Schule in Stettin.

Hermann Günther Graßmann (1809–1877) war ein ungewöhnlicher Mann. Auf den ersten Blick ist an seiner Biographie allerdings nichts Außergewöhnliches zu erkennen: Er studierte in Berlin Theologie und klassische Philologie (wahrscheinlich hat er nie eine

mathematische Vorlesung gehört), wurde 1831 Lehrer in Stettin (dem heutigen Szczecin), wo er bis zu seinem Tode blieb. Aber seine Interessen waren äußerst vielseitig: Graßmann arbeitete zeitweilig in der Redaktion der „Norddeutschen Zeitung“, sammelte pommersche Volkslieder und war Vorsitzender des „Pommerschen Hauptvereins für die Evangelisierung Chinas“. Bis heute kann man in zwei völlig unterschiedlichen Wissenschaftsgebieten seine Nachwirkung feststellen: In der Sprachwissenschaft – und in der Mathematik.

Graßmanns sprachwissenschaftliche Studien kulminierten in der Übersetzung der *Rig-veda*, einer Sammlung altindischer religiöser Texte und Lieder, und der Herausgabe eines dazugehörigen Wörterbuches; für diese Leistung wurde ihm 1876 die Ehrendoktorwürde der Universität Tübingen verliehen.

Die für uns wichtigste Leistung Graßmanns besteht aber in der Mathematik. In seiner Ausdehnungslehre hat er bereits den Begriff eines n -dimensionalen Vektorraums, und zwar den eines „allgemeinen“ Vektorraums (also nicht nur des K^n) ausdrücklich vorgestellt. Er schreibt:

Es geht darum, die sinnlichen Anschauungen der Geometrie zu allgemeinen, logischen Begriffen zu erweitern und zu vergeistigen ... Ich sage, eine GröÙe a sei aus den GröÙen b, c, \dots durch die Zahlen β, γ, \dots abgeleitet, wenn $a = \beta b + \gamma c + \dots$. Dabei seien β, γ, \dots reelle Zahlen.

Die GröÙen a, b, c, \dots stehen zueinander in einer Zahlbeziehung, wenn irgend eine sich aus den anderen numerisch berechnen lässt ... Einheit nenne ich jede GröÙe, welche dazu dienen soll, um aus ihr eine Reihe von GröÙen abzuleiten. Ein System von Einheiten nenne ich jeden Verein von GröÙen, welche in keiner Zahlbeziehung zueinander stehen und welche dazu dienen sollen, um aus ihnen durch beliebige Zahlen andere GröÙen abzuleiten. Die algebraischen GröÙen heißen auch extensive GröÙen.

Für extensive GröÙen gelten die Fundamentalformeln:

$$\begin{aligned} a + b &= b + a \\ a + (b + c) &= (a + b) + c \\ a + b - b &= a ; \end{aligned}$$

a, b, c sind GröÙen, α, β reelle Zahlen:

$$\begin{aligned} \alpha a &= a \alpha \\ \alpha(\beta a) &= (\alpha \beta) a \\ \alpha(a + b) &= \alpha a + \alpha b \\ a(\alpha + \beta) &= \alpha a + a \beta \\ 1a &= a . \end{aligned}$$

Die Gesamtheit der GröÙen, welche aus einer Reihe von GröÙen a_1, a_2, \dots, a_n numerisch ableitbar sind, nenne ich das aus jenen GröÙen ableitbare Gebiet n -ter Stufe, wenn jene GröÙen von erster Stufe sind und sich das Gebiet nicht aus weniger als n solchen GröÙen ableiten lässt. Jedes Gebiet n -ter Stufe kann aus n (ihm angehörenden) GröÙen erster Stufe, die in keiner Zahlbeziehung zueinander stehen, abgeleitet werden, und zwar aus beliebigen n solchen GröÙen des Gebiets.

Graßmanns Werk hätte der Beginn der modernen linearen Algebra sein können – aber seine Erkenntnisse blieben fast ohne Wirkung. Seine Darstellung war „äußerst schwer zugänglich, ja fast unlesbar“. Seine (mindestens für Mathematiker) verwirrenden philosophischen Gedankengänge taten ein Übriges: Fast die gesamte Auflage der „linealen Ausdehnungslehre“ wurde wieder eingestampft.

Der Vektorraumbegriff, wie wir ihn heute benutzen, wurde (allerdings nur über dem Körper der reellen Zahlen) erstmals von Hermann Weyl (1885–1955) in seinem Buch *Raum – Zeit – Materie. Vorlesungen über allgemeine Relativitätstheorie* 1917 vorgestellt. Dort heißt es:

1. Vektoren

Je zwei Vektoren a und b bestimmen eindeutig einen Vektor $a+b$ als ihre „Summe“; eine Zahl λ und ein Vektor a bestimmen eindeutig einen Vektor λa , das „ λ -fache von a “ (Multiplikation). Diese Operationen genügen den folgenden Gesetzen:

a) Addition.

- i. $a + b = b + a$ (kommutatives Gesetz).
- ii. $(a + b) + c = a + (b + c)$ (assoziatives Gesetz).
- iii. Sind a und c irgend zwei Vektoren, so gibt es einen und nur einen Vektor x , für welchen die Gleichung $a + x = c$ gilt. Er heißt die Differenz $c - a$ von c und a (Möglichkeit der Subtraktion).

b) Multiplikation

- i. $(\lambda + \mu)a = (\lambda a) + (\mu a)$ (erstes distributives Gesetz).
- ii. $\lambda(\mu a) = (\lambda\mu)a$ (assoziatives Gesetz).
- iii. $1a = a$.
- iv. $\lambda(a + b) = (\lambda a) + (\lambda b)$ (zweites distributives Gesetz).

Dimensionsaxiom: Es gibt n linear unabhängige Vektoren, aber je $n+1$ sind voneinander linear abhängig.

Spätestens seit B. L. van der Waerdens Buch *Moderne Algebra* 1936 hat sich der Begriff des Vektorraums über einem beliebigen Körper eingebürgert (van der Waerden betrachtet sogar Vektorräume über Schiefkörpern!) und ist heute aus keiner Mathematikausbildung mehr wegzudenken.

Weitere Informationen zur Geschichte der Linearen Algebra finden Sie in [Cro], [Mäd] und [Scho], Kap. 13. das Werk Graßmanns wird ausführlich in [Zad] vorgestellt.

3.5 Verständnisfragen, Übungen und Tipps

Richtig oder falsch?

1. Thema: Definition eines Vektorraums

- ☐ Man kann je zwei Vektoren eines Vektorraums addieren.
- ☐ Man kann einen Vektor v durch einen Vektor w dividieren, falls $w \neq o$ ist.

- ☐ Jeder Vektorraum hat ein eindeutiges Nullelement.
- ☐ Jeder Vektorraum hat ein eindeutiges Einselement.
- ☐ Wenn V ein K -Vektorraum ist, dann ist $\{v+w \mid v \in V, w \in V\} = V$.
- ☐ Wenn V ein K -Vektorraum ist, dann ist $\{v+w \mid v \in V, w \in V\} = V \times V$.
- ☐ Für alle u, v, w eines Vektorraums V gilt $u \cdot (v \cdot w) = (u \cdot v) \cdot w$.
- ☐ \mathbf{R}^n besteht aus allen n -Tupeln reeller Zahlen.
- ☐ \mathbf{R}^n besteht aus n -Tupeln von Vektoren.

Welche der folgenden Mengen sind Vektorräume, wenn man die Addition und die Multiplikation komponentenweise definiert?

- ☐ Die Menge aller reellen Folgen der Länge 1000,
- ☐ die Menge aller endlichen reellen Folgen,
- ☐ die Menge aller unendlichen reellen Folgen,
- ☐ die Menge aller unendlichen reellen Folgen, die nur endlich viele von 0 verschiedene Komponenten haben,
- ☐ die Menge aller unendlichen reellen Folgen, die unendlich viele von 0 verschiedene Komponenten haben,
- ☐ die Menge aller unendlichen reellen Folgen, die nur endlich viele von 1 verschiedene Komponenten haben.

2. Thema: Erzeugnis und lineare Unabhängigkeit

- ☐ Die Vektoren v_1, v_2, \dots, v_n sind linear unabhängig, wenn eine Linearkombination von v_1, v_2, \dots, v_n den Nullvektor ergibt.
- ☐ Die Vektoren v_1, v_2, \dots, v_n sind linear unabhängig, wenn ihre Summe 0 ist.
- ☐ Der Nullvektor ist nur durch die triviale Linearkombination darstellbar.
- ☐ Der Nullvektor ist stets durch die triviale Linearkombination darstellbar.
- ☐ Ist B ein Erzeugendensystem eines Vektorraums V , so ist jeder Vektor durch mindestens/genau/höchstens eine Linearkombination der Vektoren aus B darstellbar.
- ☐ Ist B eine Menge linear unabhängiger Vektoren eines Vektorraums V , so ist jeder Vektor durch mindestens/genau/höchstens eine Linearkombination der Vektoren aus B darstellbar.

3. Thema: Dimension

Welche der folgenden Objekte haben eine Dimension?

- ☐ Ein Vektor,
- ☐ eine Linearkombination,
- ☐ eine Basis,
- ☐ ein Unterraum.
- ☐ Welche der folgenden Zahlen kommen als Dimension eines Vektorraums in Frage:
-1, 0, 1, 17.5, 2001, ∞ ?

4. Thema: Basen

Sei V ein Vektorraum.

- ☐ Jede Teilmenge einer Menge linear unabhängiger Vektoren ist linear unabhängig.
- ☐ Jede Teilmenge einer Menge linear abhängiger Vektoren ist linear abhängig.
- ☐ Wenn eine Basis von V unendlich ist, sind alle Basen von V unendlich.

- ☐ Wenn eine Basis von V endlich ist, sind alle Basen von V endlich.
- ☐ Wenn V ein unendliches Erzeugendensystem hat, sind alle Basen von V unendlich.
- ☐ Es gibt eine Basis des \mathbf{R}^3 aus Vektoren der Form (x, x, x) .
- ☐ Jede Basis des \mathbf{R}^3 besteht aus Vektoren der Form (x, x, x) .
- ☐ Jeder Vektor der Form (x, x, x) kann zu einer Basis von \mathbf{R}^3 ergänzt werden.

5. Thema: Unterräume

Sei U ein Unterraum des Vektorraums V . Dann gilt für alle $u, u' \in V$:

- ☐ $u, u' \notin U \Rightarrow u+u' \notin U$,
- ☐ $u, u' \notin U \Rightarrow u+u' \in U$,
- ☐ $u \notin U, u' \in U \Rightarrow u+u' \notin U$.

6. Thema: Nebenklassen

Sei U ein Unterraum des Vektorraums V .

- ☐ Jede Nebenklasse ist ein Unterraum.
- ☐ Jeder Unterraum ist eine Nebenklasse.
- ☐ Die leere Menge ist eine Nebenklasse.
- ☐ Es ist möglich, dass eine Nebenklasse von U in V in einer anderen enthalten ist.
- ☐ Zu jedem Unterraum gehören mindestens zwei Nebenklassen.
- ☐ Jeder Unterraum von \mathbf{R}^3 hat unendlich viele Nebenklassen.

7. Thema: Faktorraum

Der Faktorraum von V nach U ist

- ☐ eine Menge von Vektoren aus U ,
- ☐ ein Unterraum von V ,
- ☐ ein Vektorraum,
- ☐ eine Nebenklasse von U ,
- ☐ eine Menge von Nebenklassen.
- ☐ Aus $v+U = U$ folgt $v = o$, weil $o+U = U$ gilt.
- ☐ Aus $v+U = U$ folgt $v = o$, weil $v+U = o+U$ ist.
- ☐ Aus $v+U = U$ folgt $v = o$, weil man auf beiden Seiten $-U$ addieren kann.
- ☐ Aus $v+U = U$ folgt $v \in U$.

Übungsaufgaben

1. Weisen Sie für mindestens zwei Beispiele aus Abschn. 3.2 die Vektorraumaxiome explizit nach.
2. Wie können Addition und skalare Multiplikation definiert werden, so dass die folgenden Mengen zu Vektorräumen werden?

$$V_1 = \{(x_1, x_2, \dots, x_n) \mid x_i \in \mathbf{R}\},$$

$$V_2 = \{(x_1, x_2, \dots, x_n) \mid x_i \in \{0, 1\}\},$$

$$V_3 = \{f \mid f \text{ stetige Abbildung auf } (a, b)\}.$$

3. Zeigen Sie, dass die Menge der Lösungen (x, y, z) des reellen Gleichungssystems

$$5x - 3y + 21z = 0$$

$$3x + 7y + 12z = 0$$

$$x - 30y + 6z = 0$$

einen Vektorraum bilden. Wie viele „Freiheitsgrade“ hat dieser Vektorraum?

4. Welche der folgenden Teilmengen U von \mathbf{R}^n ist ein Vektorraum?

☐ $U = \{(x_1, x_2, \dots, x_n) \in \mathbf{R}^n \mid x_1 = x_2 = \dots = x_n\},$

☐ $U = \{(x_1, x_2, \dots, x_n) \in \mathbf{R}^n \mid x_1 = 1\},$

☐ $U = \{(x_1, x_2, \dots, x_n) \in \mathbf{R}^n \mid x_1 = 0\},$

☐ $U = \{(x_1, x_2, \dots, x_n) \in \mathbf{R}^n \mid x_1^2 = 0\},$

☐ $U = \{(x_1, x_2, \dots, x_n) \in \mathbf{R}^n \mid x_1^2 - x_2^2 = 0\}.$

5. Zeigen Sie: (a) Ist g eine Gerade der euklidischen Ebene \mathbf{R}^2 durch den Nullpunkt, so ist g ein Vektorraum.

(b) Sei E eine Ebene des \mathbf{R}^3 durch den Nullpunkt. Dann ist E ein Vektorraum.

6. Zeigen Sie: Ist g eine Gerade der euklidischen Ebene \mathbf{R}^2 , die nicht durch den Nullpunkt geht, so ist g kein Vektorraum.

7. Wir stellen uns eine Balkenwaage vor, die man an drei Stellen mit Gewichten belasten kann; diese Stellen seien 20 cm links vom Aufhängungspunkt und 5 bzw. 10 cm rechts vom Aufhängungspunkt.

(a) Geben Sie zwei Gewichtssätze an, bei denen die Waage im Gleichgewicht ist.

(b) Zeigen Sie, dass die Menge aller Gewichtssätze, bei denen die Waage im Gleichgewicht ist, bezüglich komponentenweiser Verknüpfungen ein reeller Vektorraum ist. (Dabei muss man allerdings auch negative Gewichte zulassen.)

(c) Welche Dimension hat dieser Vektorraum?

8. Definieren Sie den Quaternionenschiefkörper dadurch, dass Sie formulieren: Auf dem Vektorraum $\mathbf{H} := \mathbf{R}^4$ definieren wir das folgende Produkt ...

Machen Sie sich klar, wie viel einfacher der Nachweis der Körperaxiome wird, wenn man die elementare Theorie der Vektorräume voraussetzt.

9. Im Folgenden werden jeweils einige Vektoren des Vektorraums $V = \mathbf{R}^3$ angegeben. Entscheiden Sie, ob es Vektoren von \mathbf{R}^3 gibt, die nicht im Erzeugnis der angegebenen Vektoren liegen, und geben Sie gegebenenfalls einen solchen Vektor an.

(a) $(1, 0, 0), (1, 1, 1);$

(b) $(1, 0, 0), (0, 1, 0), (1, 1, 1);$

(c) $(1, 0, 0), (0, -1, 0), (1, 1, -1);$

(d) $(3, 4, 7), (1, 0, 3), (0, 4, -2), (3, 8, 5);$

(e) $(0.00000000001, 0, 0), (0, 0.00000000001, 0), (0, 0, 0.00000000001);$

(f) $(\pi, e, 0), (e, \pi, 0), (0, \pi, e).$

10. Geben Sie mindestens drei Basen des \mathbf{R}^3 an.

11. Zeigen Sie: Je drei Vektoren der Menge $\{(1, x, x^2) \mid x \in K\}$ von K^3 sind linear unabhängig.

12. Beweisen Sie folgende Aussagen. Sei V ein K -Vektorraum.

(a) Für alle $k \in K$, $v \in V$ gilt

$$k \cdot o = o \quad \text{und} \quad (-k) \cdot v = -(k \cdot v) = k \cdot (-v)$$

[Überlegen Sie zunächst, in welchen Strukturen das Minuszeichen was bedeutet.]

(b) Für alle $k_1, k_2 \in K$ und alle $v \in V$ mit $v \neq o$ gilt:

$$k_1 \cdot v = k_2 \cdot v \Leftrightarrow k_1 = k_2.$$

13. Zeigen Sie: Wenn v_1, \dots, v_n Elemente eines K -Vektorraums V sind, dann gilt:

(a) Die Summe je zweier Vektoren aus $\langle v_1, \dots, v_n \rangle$ liegt wieder in $\langle v_1, \dots, v_n \rangle$.

(b) Das Produkt eines Skalars k mit einem Vektor aus $\langle v_1, \dots, v_n \rangle$ liegt wieder in $\langle v_1, \dots, v_n \rangle$.

14. Beweisen Sie. Sei B eine Menge von Vektoren eines Vektorraums V . Dann gilt: Genau dann ist B eine Basis von V , wenn B ein minimales Erzeugendensystem ist.

15. Sei K ein Körper, und sei $n \in \mathbb{N}$.

(a) Zeigen Sie, dass die Menge der Einheitsvektoren $e_1 = (1, 0, 0, \dots, 0)$, $e_2 = (0, 1, 0, \dots, 0)$, \dots , $e_n = (0, 0, \dots, 0, 1)$ eine Basis von K^n ist.

(b) Welche Dimension hat K^n ?

(c) Gibt es eine Basis $\{v_1, v_2, \dots, v_n\}$ von K^n , bei der v_i genau i von 0 verschiedene Komponenten hat?

16. Welche Dimension hat der K -Vektorraum $K^{m \times n}$? Geben Sie eine Basis dieses Vektorraums an.

17. Sei V der $\text{GF}(2)$ -Vektorraum aller Teilmengen einer n -elementigen Menge X (vergleichen Sie dazu Abschn. 3.2.8).

(a) Geben Sie eine Basis von V an.

(b) Gibt es eine Basis von V , deren Elemente Teilmengen von X mit mehr als einem Element sind?

(c) Welche Dimension hat V ?

18. Sei V der $\text{GF}(2)$ -Vektorraum aller Teilmengen einer n -elementigen Menge X (vergleichen Sie Abschn. 3.2.8).

(a) Sei W die Menge aller Teilmengen von X , die eine *gerade* Anzahl von Elementen besitzen. Zeigen Sie: W ist ein Unterraum von V .

(b) Sei Y eine Teilmenge von X *ungerader* Mächtigkeit. Zeigen Sie, dass W zusammen mit Y bereits ganz V erzeugt. Ist W eine Hyperebene von V ?

(c) Wie viele Elemente hat W ? Wie viele Elemente von V liegen außerhalb von W ? Interpretieren Sie diese Tatsache als Aussage über die Teilmengen einer endlichen Menge.

19. Beweisen Sie folgendes (außerordentlich nützlich!) **Unterraumkriterium**: Sei U eine Teilmenge eines K -Vektorraums V . Wenn gilt

- $U \neq \emptyset$,
- für alle $k \in K$ und $u \in U$ gilt $k \cdot u \in U$,
- für alle $u, u' \in U$ ist $u - u' \in U$,

dann ist U ein Unterraum von V .

20. Zeigen Sie, dass das Erzeugnis $\langle v_1, \dots, v_n \rangle$ von Vektoren v_1, \dots, v_n eines Vektorraums V ein Unterraum von V ist.
21. Sei $V = \mathbf{R}^3$.
- (a) Zeigen Sie, dass die Menge ein Unterraum von V ist.
 - (b) Geben Sie eine Basis von U an.
 - (c) Ergänzen Sie diese Basis zu einer Basis von V .
22. Sei V die Menge aller Abbildungen von \mathbf{R} nach \mathbf{R} .
- (a) Zeigen Sie: Wenn man eine Addition und eine skalare Multiplikation auf V wie folgt definiert:

$$(f+g)(r) := f(r) + g(r), (a \cdot f)(r) := a \cdot f(r) \quad \text{für alle } r \in \mathbf{R} (f, g \in V, a \in \mathbf{R}),$$

so wird V zu einem \mathbf{R} -Vektorraum.

- (b) Wir definieren:

$$W := \{f \in V \mid f(1) = 0 \quad \text{und} \quad f(-1) = 0\}.$$

Zeigen Sie: W ist ein Unterraum von V .

23. Sei V ein Vektorraum der Dimension n . Zeigen Sie: Jedes Erzeugendensystem C besteht aus mindestens n Vektoren; Gleichheit gilt genau dann, wenn C eine Basis von V ist.
24. Seien U und U' Unterräume eines K -Vektorraums V .
- (a) Zeigen Sie: Die Menge

$$U + U' := \{u + u' \mid u \in U, u' \in U'\}.$$

(die **Summe** von U und U') ist ein Unterraum von V .

- (b) Gilt $U + U' = \langle U, U' \rangle$?

25. Seien U und U' Unterräume eines K -Vektorraums V . Wir definieren:

$$U - U' := \{u - u' \mid u \in U, u' \in U'\}.$$

Welche der folgenden Aussagen gilt dann

- (a) $U - U' = \{0\}$,
- (b) $U - U' = \emptyset$,
- (c) $U - U' = U + U'$?

26. Zeigen Sie: Seien U und U' komplementäre Unterräume des Vektorraums V . Wenn B eine Basis von U und B' eine Basis von U' ist, dann ist $B \cup B'$ eine Basis von V .

27. Bestimmen Sie diejenigen Unterräume eines Vektorraums, die nur ein Komplement haben.
28. Zeigen Sie: Der Durchschnitt beliebig vieler Unterräume eines Vektorraums V ist wieder ein Unterraum von V .
29. Seien U und U' Unterräume des Vektorraums V .
- (a) Untersuchen Sie, wann die mengentheoretische Vereinigung $U \cup U'$ (also nicht das Erzeugnis!) von U und U' ein Unterraum von V ist.
- (b) Wann ist $U \cup U' = \langle U, U' \rangle$?
30. Sei U ein Unterraum des K -Vektorraums V . Zeigen Sie: Das mengentheoretische Komplement von U in V (siehe Abschn. 1.1) ist nie ein Unterraum.
31. Zeigen Sie, dass für jeden Unterraum U eines Vektorraums V die Struktur V/U ein Vektorraum ist.
32. Sei $\{v_1, v_2\}$ eine Basis eines 2-dimensionalen \mathbf{R} -Vektorraums V . Für welche reellen Zahlen s, t ist die Menge $\{w_1, w_2\}$ mit

$$w_1 := sv_1 + v_2,$$

$$w_2 := v_1 + tv_2$$

ebenfalls eine Basis von V ?

33. (a) Sei $\{v_1, v_2, v_3\}$ eine Basis eines 3-dimensionalen \mathbf{R} -Vektorraums V . Zeigen Sie, dass dann auch die Menge $\{c_1, c_2, c_3\}$ eine Basis ist mit

$$c_1 := 3v_1 + v_2 + 2v_3,$$

$$c_2 := v_1 + v_2 + v_3$$

$$c_3 := v_1 + v_2 + 2v_3.$$

- (b) Gilt diese Aussage auch noch, wenn man als K den Körper $\text{GF}(2)$ oder $\text{GF}(3)$ wählt?

34. (a) Zeigen Sie, dass die Menge $B := \{(1, 2, 3, 4), (2, 0, 1, -1), (-1, 0, 0, 1), (0, 2, 3, 0)\}$ eine Basis des Vektorraums \mathbf{R}^4 ist.
- (b) Ergänzen Sie die Menge $\{(0, 4, 5, 9), (3, 3, 3, 3)\}$ durch Vektoren aus B zu einer Basis von \mathbf{R}^4 .
35. Welche Dimension hat der von den Vektoren $(1, 2, t+2), (-1, t+1, t), (0, t, 1)$ erzeugte Unterraum von \mathbf{R}^3 ($t \in \mathbf{R}$)?
[Hinweis: Achtung!]
36. Sei V die Menge aller unendlichen Folgen (a_1, a_2, a_3, \dots) reeller Zahlen mit der Eigenschaft $a_i = a_{i-2} + a_{i-1}$ für $i \geq 3$.
- (a) Zeigen Sie, dass V zusammen mit der komponentenweisen Addition und Skalarmultiplikation ein Vektorraum ist.
- (b) Welche Dimension hat V ?
- (c) Können Sie eine Basis von V angeben?

37. Eine reelle $n \times n$ -Matrix wird ein **magisches Quadrat der Ordnung n** genannt, wenn die Summen der Elemente in jeder Zeile und jeder Spalte gleich sind.
- (a) Zeigen Sie, dass die Menge aller magischen Quadrate der Ordnung n einen Vektorraum bilden.
- (b) Bestimmen Sie die Dimension des Vektorraums der magischen Quadrate der Ordnung 3. Geben Sie eine Basis dieses Vektorraums an.
[Hinweis: Es gibt eine Basis aus Matrizen, die in jeder Zeile und jeder Spalte nur eine Eins (und sonst Nullen) enthalten.]
- (c) Versuchen Sie, (b) zu verallgemeinern.
38. (a) Informieren Sie sich in einem Mathematikerlexikon oder einem Buch zur Geschichte der Mathematik über Ernst Steinitz.
- (b) Lesen Sie, was K. Jänich in seinem Buch über Lineare Algebra zur historischen Einordnung des „Austauschsatzes von Steinitz“ sagt.
39. Studieren Sie den Textauszug aus Graßmanns „Ausdehnungslehre“. Welche unserer Begriffe erkennen Sie wieder? (Erzeugnis, linear unabhängig, Dimension, ...?)
40. Vergleichen Sie die Definition eines Vektorraums nach Hermann Weyl mit der Definition in Abschn. 3.1.

Projekt: Der unendlichdimensionale Vektorraum V_∞

Wir haben Vektorräume, die keine endliche Basis haben (so genannte **unendlichdimensionale** Vektorräume) von den Betrachtungen dieses Kapitels weitgehend ausgeschlossen. Die Existenz solcher Vektorräume ist aber nicht zu leugnen, und besonders wichtig ist der Vektorraum aller unendlichen Folgen aus Elementen eines Körpers, in denen nur endlich viele Elemente verschieden von 0 sind (siehe Abschn. 3.2.5).

Sie sind eingeladen, diesen Vektorraum selbständig zu untersuchen. Dabei werde ich Ihnen einige Tipps geben, insbesondere dadurch, dass ich Ihnen realistische Arbeitsziele in einer vernünftigen Reihenfolge nenne. Wenn Sie dieses Projekt erfolgreich bearbeitet haben, kennen Sie sich mit der grundlegenden Theorie der Vektorräume aus!

Wir betrachten also den Vektorraum V_∞ aller unendlichen Folgen mit Elementen aus einem Körper K , bei denen nur endlich viele Elemente $\neq 0$ sind. Um die Bezeichnungen einzuführen, beschreiben wir diesen Vektorraum nochmals auf andere Weise:

$$V_\infty = \{(k_1, k_2, \dots) \mid k_i \in K, \text{ nur endlich viele } k_i \neq 0\}.$$

Die Endlichkeitsbedingung kann man auch so formulieren: Für jeden Vektor $v = (k_1, k_2, \dots)$ aus V_∞ gibt es eine natürliche Zahl n , so dass $k_i = 0$ für alle $i > n$ gilt.

0 Machen Sie sich nochmals völlig klar, dass V_∞ ein K -Vektorraum ist.

Ihr erstes großes Ziel ist zu zeigen, dass V_∞ eine Basis hat. Dafür bietet sich ein Kandidat an, nämlich die Menge B^* , die aus den Einheitsvektoren e_i besteht. Dabei ist e_i der Vektor, der an der i -ten Stelle eine 1 hat und ansonsten aus Nullen besteht; zum Beispiel ist

$$e_0 = (1, 0, 0, \dots), e_1 = (0, 1, 0, 0, \dots), e_2 = (0, 0, 1, 0, 0, \dots).$$

(Beachten sie, dass wir aus Gründen, die später – in Kap. 6 – klar werden, hier mit 0 zu zählen beginnen.)

Um zu zeigen, dass B^* eine Basis ist, müssen wir den Begriff der Linearkombination noch etwas genauer fassen. Wir sagen, ein Vektor v ist eine **Linearkombination** einer Menge $M = \{v_1, v_2, \dots\}$ von Vektoren, wenn es *endlich viele* Vektoren aus M gibt, so dass v eine Linearkombination dieser Vektoren ist.

Wenn man das genauer aufschreiben möchte, so schreibt man dafür am einfachsten (glauben Sie mir!) wie folgt:

$$v = \sum_{j=1}^n k_{ij} v_{ij}.$$

Was soll das bedeuten? Das bedeutet, dass nur eine endliche Menge von Indizes eine Rolle spielt, nämlich die Menge $\{i_1, i_2, \dots, i_n\}$. Bei der Darstellung von v spielen nur Vektoren mit diesen Indizes mit; der Vektor v_{ij} wird dabei mit dem Koeffizienten k_{ij} multipliziert.

Nochmals anders ausgedrückt. Man wählt zunächst eine endliche Folge i_1, i_2, \dots, i_n von Indizes (also von nichtnegativen ganzen Zahlen). Dass diese Folge hier üblicherweise mit i_1, i_2, \dots, i_n und nicht mit a_1, a_2, \dots, a_n bezeichnet wird, darf Sie nicht stören. Dann wählt man die Vektoren aus M mit diesen Indizes und multipliziert diese mit irgendwelchen Skalaren.

Man sagt, dass eine Menge M den Vektorraum V_∞ **erzeugt**, wenn jeder Vektor aus V eine Linearkombination endlich vieler Elemente aus M ist. Ferner heißt M **linear unabhängig**, wenn der Nullvektor keine nichttriviale Linearkombination von M ist.

Damit können Sie die ersten Schritte des Projekts bewältigen.

1. Beweisen Sie, dass B^* eine Basis ist.
2. Beweisen Sie das **Austauschlemma** bezüglich der Standardbasis B^* . Sei v ein von 0 verschiedener Vektor von V_∞ . Dann gibt es einen Vektor $e_i \in B^*$, so dass man e_i durch v ersetzen kann, das heißt, dass die Menge

$$(B^* \setminus \{e_i\}) \cup \{v\}$$

ebenfalls eine Basis von V_∞ ist.

3. Formulieren und beweisen Sie den **Steinitzschen Austauschsatz** bezüglich der Basis B^* .
4. Zeigen Sie, dass der Unterraum $V_i := \langle e_i, e_{i+1}, \dots \rangle$ „isomorph“ zu V_∞ ist.
5. Sei U ein endlichdimensionaler Unterraum von V_∞ . Zeigen Sie, dass V_∞/U „isomorph“ zu V_∞ ist.

Bemerkung: Der Isomorphiebegriff wird in Kap. 5 ausführlich behandelt.

Wir nennen zwei Unterräume U_1 und U_2 von V_∞ **komplementär**, wenn $\langle U_1, U_2 \rangle = V_\infty$ und $U_1 \cap U_2 = \emptyset$ ist.

6. Zeigen Sie: Sind U_1 und U_2 komplementäre Unterräume von V_∞ , so hat jeder Vektor v von V_∞ eine eindeutige Darstellung als $v = u_1 + u_2$ mit $u_1 \in U_1$ und $u_2 \in U_2$.
7. Zeigen sie: Jeder endlichdimensionale Unterraum von V_∞ hat einen komplementären Unterraum.
Eine **Hyperebene** von V_∞ ist ein Unterraum $H \neq V_\infty$, für den es einen Vektor v gibt mit $\langle H, v \rangle = V_\infty$.
8. Zeigen Sie: Ist H eine Hyperebene des Vektorraums V_∞ , so gilt $\langle H, v \rangle = V_\infty$ für alle $v \in V_\infty \setminus H$.
9. Sei H eine Hyperebene von V_∞ , und sei U ein t -dimensionaler Unterraum von V_∞ . Dann gilt: Ist U kein Unterraum von H , so ist $\dim(U \cap H) = t - 1$.

Sie sollten mit folgenden Begriffen umgehen können

Vektorraum, Vektor, Skalar, linear abhängig, linear unabhängig, K^n , Linearkombination, Unterraum, Erzeugendensystem, Basis, Dimension, Komplement, Hyperebene, Nebenklasse, Faktorraum, wohldefiniert



In diesem Kapitel behandeln wir drei wichtige Anwendungen der bisher entwickelten Vektorraumtheorie. Diese Anwendungen scheinen sehr verschieden zu sein. Wenn man aber genauer hinschaut, erkennt man, dass sie zum Teil eng zusammenhängen.

4.1 Lineare Gleichungssysteme

4.1.1 Begriffe und Fragen

Ein **lineares Gleichungssystem** in den **Unbekannten** x_1, \dots, x_n besteht aus m linearen Gleichungen der Form

$$\begin{aligned} a_{11} \cdot x_1 + a_{12} \cdot x_2 + \dots + a_{1n} \cdot x_n &= b_1 \\ a_{21} \cdot x_1 + a_{22} \cdot x_2 + \dots + a_{2n} \cdot x_n &= b_2 \\ &\dots \\ a_{m1} \cdot x_1 + a_{m2} \cdot x_2 + \dots + a_{mn} \cdot x_n &= b_m . \end{aligned}$$

Dabei sind die Elemente a_{ij} und b_i Elemente eines Körpers K . Man nennt ein solches Gleichungssystem **homogen**, falls $b_1 = \dots = b_m = 0$ ist und sonst **inhomogen**.

In der Theorie der linearen Gleichungssysteme fragt man nach **Lösungen** eines vorgelegten Gleichungssystems; darunter versteht man ein n -Tupel (k_1, \dots, k_n) von Körperelementen derart, dass gilt:

$$\begin{aligned} a_{11} \cdot k_1 + a_{12} \cdot k_2 + \dots + a_{1n} \cdot k_n &= b_1 \\ a_{21} \cdot k_1 + a_{22} \cdot k_2 + \dots + a_{2n} \cdot k_n &= b_2 \\ &\dots \\ a_{m1} \cdot k_1 + a_{m2} \cdot k_2 + \dots + a_{mn} \cdot k_n &= b_m . \end{aligned} \tag{4.1}$$

Die zentralen Fragen der Theorie der linearen Gleichungssysteme sind die folgenden:

- Hat ein vorgelegtes lineares Gleichungssystem *überhaupt eine Lösung*?
- *Wie viele Lösungen* hat ein lineares Gleichungssystem?
- Wie kann man eine (alle) *Lösung(en) konstruieren*?

Auf alle diese Fragen werden wir eine Antwort geben. Zunächst werden wir jedoch einen Exkurs über Matrizen machen; denn die Darstellung von linearen Gleichungssystemen mit Hilfe von Matrizen ist deutlich übersichtlicher.

4.1.2 Exkurs über Matrizen

Der Begriff einer $m \times n$ -Matrix über einem Körper K wurde bereits in Abschn. 3.2 eingeführt. Dort haben wir auch die Addition von Matrizen und die Multiplikation einer Matrix mit einem Körperelement erklärt, und zwar komponentenweise. Hier geht es darum, das *Produkt* zweier Matrizen zu definieren. Alle im Folgenden betrachteten Matrizen seien über einem Körper K definiert.

Ich falle mit der Tür ins Haus, werde Ihnen aber anschließend helfen, sich von dem Schock zu erholen.

- Seien $A = (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$ und $B = (b_{uv})_{1 \leq u \leq r, 1 \leq v \leq s}$ Matrizen. Damit ein **Produkt** von A und B überhaupt erklärt werden kann, muss die Anzahl der Spalten von A gleich der Anzahl der Zeilen von B sein. Sei also $n = r$.
- In diesem Fall ergibt sich als Produkt eine $m \times s$ -Matrix $C = (c_{iv})_{1 \leq i \leq m, 1 \leq v \leq s}$, deren Elemente wie folgt definiert sind:

$$c_{iv} = \sum_{j=1}^n a_{ij} b_{jv}.$$

Wie sollen wir uns diese Formel jemals merken? Wie kann man sich das vorstellen? Geht's nicht noch komplizierter? Ausgeburt eines perversen Mathematikerhirns!??

Gemach, gemacht! Zum einen erweist sich gerade diese Definition der Multiplikation von Matrizen als die einzig wahre; zum anderen verspreche ich Ihnen, dass Sie in spätestens zehn Minuten den Mechanismus der Matrizenmultiplikation verstanden haben.

Um das Element c_{iv} auszurechnen (und nur darum geht es!), muss man die i -te Zeile von A und die v -te Spalte von B betrachten. Tun wir das:

$$(a_{i1} \quad a_{i2} \quad \cdots \quad a_{in}) \cdot \begin{pmatrix} b_{1v} \\ b_{2v} \\ \vdots \\ b_{rv} \end{pmatrix}.$$

Nun kommt der entscheidende Trick: Wie viele Elemente hat die i -te Zeile von A ? Nun, genauso viele, wie die Anzahl der Spalten von A beträgt, also genau n . Entsprechend ist die Anzahl der Elemente der v -ten Spalte von B gleich der Anzahl r der Zeilen von B . Da $n = r$ ist, können wir uns diese beiden Objekte (die i -te Zeile von A und die v -te Spalte von B) übereinander gelegt vorstellen.

$$\begin{pmatrix} a_{i1} & a_{i2} & \dots & a_{in} \\ b_{1v} & b_{2v} & \dots & b_{nv} \end{pmatrix}.$$

Wenn wir soweit sind, multiplizieren wir die übereinander liegenden Elemente und addieren die Produkte; dies ergibt c_{iv} !!! In der folgenden Schemazeichnung sind entsprechende Elemente durch das gleiche Symbol dargestellt; in der Ergebnismatrix ist in symbolischer Weise das Element c_{iv} eingetragen.

$$\begin{pmatrix} - & - & - & - \\ \clubsuit & \diamond & \heartsuit & \spadesuit \\ - & - & - & - \\ - & - & - & - \\ - & - & - & - \end{pmatrix} \cdot \begin{pmatrix} - & - & - & - & \clubsuit & - \\ - & - & - & - & \diamond & - \\ - & - & - & - & \heartsuit & - \\ - & - & - & - & \spadesuit & - \end{pmatrix} = \begin{pmatrix} - & - & - & - & - & - \\ - & - & - & - & \clubsuit\clubsuit + \diamond\diamond + \heartsuit\heartsuit + \spadesuit\spadesuit & - \\ - & - & - & - & - & - \\ - & - & - & - & - & - \\ - & - & - & - & - & - \end{pmatrix}.$$

Nun noch ein *Beispiel*, und Sie haben alles verstanden! Was ist

$$\begin{pmatrix} 1 & 0 & 5 \\ 3 & 2 & 7 \end{pmatrix} \cdot \begin{pmatrix} 2 & 3 \\ 4 & 10 \\ -1 & 2 \end{pmatrix}?$$

Das ist einfach: Aber zunächst muss man fragen: Geht das überhaupt? Ja, es geht, denn die Anzahl der Spalten der ersten Matrix A ist 3, und dies ist auch die Anzahl der Zeilen der zweiten Matrix B . Welches sind die Dimensionen der Produktmatrix C ? Da A zwei Zeilen hat, hat auch C zwei Zeilen, und da B zwei Spalten hat, gilt dies auch für C . Das Ergebnis ist also eine 2×2 -Matrix.

Unsere Aufgabe ist es, die Koeffizienten c_{ij} zu bestimmen. Ans Werk! Was ist c_{11} ? Dieser Wert ergibt sich, indem man die erste Zeile von A mit der ersten Spalte von B multipliziert, das heißt:

$$c_{11} = 1 \cdot 2 + 0 \cdot 4 + 5 \cdot (-1) = -3.$$

Auch bei der Berechnung des nächsten Koeffizienten c_{12} erkennt man an den Indizes, was man zu tun hat: Die 1 an der ersten Stelle sagt, dass man die Zeile Nummer 1 der ersten Matrix mit der Spalte Nummer 2 der zweiten Matrix (das sagt die 2 an der zweiten Stelle) multiplizieren muss:

$$c_{12} = 1 \cdot 3 + 0 \cdot 10 + 5 \cdot 2 = 13.$$

Nun ist klar, wie der Hase läuft: c_{21} ergibt sich nach der Melodie „Zeile Nummer 2 der ersten Matrix mal Spalte Nummer 1 der zweiten Matrix“:

$$c_{21} = 3 \cdot 2 + 2 \cdot 4 + 7 \cdot (-1) = 7.$$

Auch der letzte Koeffizient ist nicht mehr schwer: c_{22} ist „Zeile Nummer 2 der ersten mal Spalte Nummer 2 der zweiten“, also

$$c_{22} = 3 \cdot 3 + 2 \cdot 10 + 7 \cdot 2 = 43.$$

Als Ergebnis erhalten wir also

$$\begin{pmatrix} -3 & 13 \\ 7 & 43 \end{pmatrix}.$$

So einfach ist das!

Noch ein Beispiel? Gut! Kann man auch Vektoren miteinander multiplizieren? Natürlich – wenn die unumgängliche Regel *Anzahl der Spalten der ersten gleich Anzahl der Zeilen der zweiten* gilt. Das bedeutet, dass man zum Beispiel eine 1×4 -Matrix (in unserer bisherigen Sprache hätten wir das wahrscheinlich einen Zeilenvektor der Länge 4 genannt) mit einer 4×1 -Matrix (also einem Spaltenvektor der Länge 4) multiplizieren kann, und eine 1×1 -Matrix, also ... – ein Körperelement, erhält. Versuchen wir's:

$$(1 \quad 0 \quad 6 \quad 6) \cdot \begin{pmatrix} 1 \\ 7 \\ 8 \\ 9 \end{pmatrix} = (1 \cdot 1 + 0 \cdot 7 + 6 \cdot 8 + 6 \cdot 9) = (103).$$

Man kann aber auch einen Spaltenvektor mit einem Zeilenvektor multiplizieren – und es ergibt sich eine „richtige Matrix“:

$$\begin{pmatrix} 1 \\ 7 \\ 8 \\ 9 \end{pmatrix} \cdot (1 \quad 0 \quad 6 \quad 6) = \begin{pmatrix} 1 \cdot 1 & 1 \cdot 0 & 1 \cdot 6 & 1 \cdot 6 \\ 7 \cdot 1 & 7 \cdot 0 & 7 \cdot 6 & 7 \cdot 6 \\ 8 \cdot 1 & 8 \cdot 0 & 8 \cdot 6 & 8 \cdot 6 \\ 9 \cdot 1 & 9 \cdot 0 & 9 \cdot 6 & 9 \cdot 6 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 6 & 6 \\ 7 & 0 & 42 & 42 \\ 8 & 0 & 48 & 48 \\ 9 & 0 & 54 & 54 \end{pmatrix}.$$

Haben Sie's jetzt verstanden? Klar, wusst' ich doch! ... und wie lange hat's gedauert? Gut, ich gebe zu, etwas länger haben wir gebraucht, aber der Zweck heiligt die Mittel.

Eigentlich brauchen wir in diesem Abschnitt nicht die Multiplikation von Matrizen im allgemeinen, sondern nur die Multiplikation einer Matrix mit einem Vektor. Dies formulieren wir explizit.

Sei $A = (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$ eine $m \times n$ -Matrix mit $a_{ij} \in K$, und sei k ein Spaltenvektor. Wenn dieser die Länge n hat, so kann das Produkt $A \cdot k$ gebildet werden. Bevor wir dies tun, noch eine kleine technische Bemerkung. Wenn wir einen Spaltenvektor explizit angeben wollen, brauchen wir in der Regel viel Platz, und es ist nicht sehr übersichtlich;

denn unsere gesamten Schreib- und Lesegewohnheiten sind zeilenorientiert. Daher **transponieren** wir den Spaltenvektor k und erhalten aus ihm den Zeilenvektor k^T (lies „ k transponiert“). Formal sieht das so aus:

$$\begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix}^T := (k_1, \quad k_2, \quad \dots, \quad k_n) .$$

Umgekehrt wird aus einem Zeilenvektor $(k_1, k_2, \dots, k_n)^T$ ein Spaltenvektor $k = (k_1, k_2, \dots, k_n)^T$ der Länge n . Dann gilt

$$A \cdot k = (a_{ij}) \cdot (k_1, \dots, k_n)^T = \left(\sum_{j=1}^n a_{1j}k_j, \quad \sum_{j=1}^n a_{2j}k_j, \quad \dots, \quad \sum_{j=1}^n a_{mj}k_j \right)^T .$$

Beachten Sie, dass das Getüm auf der rechten Seite dem ersten Anschein zum Trotz ein Spaltenvektor ist!

Damit haben wir die folgende äußerst Platz sparende Schreibweise für das Gleichungssystem (4.1) erhalten, nämlich

$$Ax = b . \quad (4.2)$$

Dies zeigt doch schon, dass die zunächst völlig willkürlich erscheinende Vorschrift zur Multiplikation von Matrizen ihren Sinn hat.

In diesem Abschnitt werden wir noch zwei Eigenschaften der Matrizenmultiplikation nachrechnen, die vor allem im nächsten Kapitel von entscheidender Bedeutung sein werden.

Additivität der Matrizenmultiplikation

Sei A eine $m \times n$ -Matrix über K . Dann gilt für je zwei Spaltenvektoren x und y der Länge n mit Elementen aus K :

$$A \cdot (x + y) = A \cdot x + A \cdot y .$$

Dieser Satz sagt in Worten: Man kann entweder zwei Vektoren zuerst addieren und dann die Matrix A mit der Summe multiplizieren oder die Matrix A zunächst mit jedem Vektor einzeln multiplizieren und die Ergebnisse addieren – und es kommt dasselbe heraus!

Der *Beweis* ist nicht schwierig: Man muss nur die Elemente von A und x und y bezeichnen und die Geduld aufbringen, beide Seiten auszuixen:

Sei $A = (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$, und seien $x = (x_1, \dots, x_n)^T$ und $y = (y_1, \dots, y_n)^T$. Dann ist $x + y = (x_1 + y_1, \dots, x_n + y_n)^T$, also nach (4.2)

$$A \cdot (x + y) = \left(\sum_{j=1}^n a_{1j}(x_j + y_j), \sum_{j=1}^n a_{2j}(x_j + y_j), \dots, \sum_{j=1}^n a_{mj}(x_j + y_j) \right)^T.$$

Andererseits ergibt sich die rechte Seite als

$$\begin{aligned} A \cdot x + A \cdot y &= \left(\sum_{j=1}^n a_{1j}x_j, \sum_{j=1}^n a_{2j}x_j, \dots, \sum_{j=1}^n a_{mj}x_j \right)^T \\ &\quad + \left(\sum_{j=1}^n a_{1j}y_j, \sum_{j=1}^n a_{2j}y_j, \dots, \sum_{j=1}^n a_{mj}y_j \right)^T. \end{aligned}$$

Da die berechneten Spaltenvektoren offenbar übereinstimmen, ist die Aussage bewiesen. \square

Homogenität der Matrizenmultiplikation

Sei A eine $m \times n$ -Matrix über K . Dann gilt für jeden Spaltenvektor x der Länge n mit Elementen aus K und jedes Element $k \in K$:

$$A \cdot (kx) = (kA) \cdot x = k \cdot (Ax).$$

Dieser Satz sagt in Worten: Man kann entweder einen Vektor zuerst mit einem Körperelement multiplizieren und dann die Matrix A mit dem Ergebnis multiplizieren oder die Matrix A zunächst mit dem Körperelement multiplizieren und das Ergebnis mit dem Vektor multiplizieren – und es kommt dasselbe heraus!

Der *Beweis* ist genauso schwierig wie der obige und wird in Übungsaufgabe 5 behandelt. \square

Es könnte scheinen, als ob man mit Matrizen so „wie mit Zahlen“ rechnen könnte. Das gilt nur zum Teil. Wir nennen zum Schluss dieses Abschnitts zwei Eigenschaften der Matrizenmultiplikation, die sich fundamental von Körpereigenschaften unterscheiden.

Das Produkt zweier von der Nullmatrix verschiedener Matrizen kann sehr wohl die Nullmatrix sein; man sagt dazu auch, die Matrizenmultiplikation ist nicht nullteilerfrei. Dafür gibt es viele Beispiele; man muss nur Matrizen mit „genügend vielen“, „gut verteilten“ Nullen wählen. Wählt man etwa

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 3 & 7 \\ 0 & 0 \end{pmatrix},$$

so sieht man unschwer, dass $A \cdot B$ gleich der Nullmatrix ist.

Die Matrizenmultiplikation ist nicht kommutativ: Hierfür können wir dieselben Matrizen wählen; es ergibt sich

$$A \cdot B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 3 & 7 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 3 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 7 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = B \cdot A$$

eine Matrix, die offenbar verschieden von der Nullmatrix ist.

4.1.3 Lösbarkeit von linearen Gleichungssystemen

Wir bezeichnen die Menge aller Lösungen des Gleichungssystems

$$Ax = b \tag{4.2}$$

mit $L(A, b)$.

Zunächst beweisen wir zwei einfache Aussagen über die Menge aller Lösungen des Gleichungssystems (4.2).

Lösungsraum eines homogenen Systems

Die Menge $L(A, 0)$ der Lösungen eines homogenen Systems bildet einen Unterraum des Vektorraums $V = K^n$.

Um das zu sehen, wenden wir das Unterraumkriterium an:

- $L(A, 0)$ ist nicht leer, da der Vektor $(0, 0, \dots, 0)^T$ eine Lösung des homogenen Systems ist (triviale Lösung).
- Seien $x = (x_1, \dots, x_n)^T$ und $y = (y_1, \dots, y_n)^T$ zwei Lösungen. Dann ist

$$A(x + y) = Ax + Ay = 0 + 0 = 0 ;$$

also ist mit x und y auch $x + y$ in $L(A, 0)$.

- Schließlich sei $x = (x_1, \dots, x_n)^T \in L(A, 0)$, und sei $k \in K$ beliebig. Dann ist

$$A(kx) = k \cdot Ax = k \cdot 0 = 0 .$$

Somit ist gezeigt, dass $L(A, 0)$ ein Unterraum von V ist. □

Wir werden sehen, dass ein inhomogenes Gleichungssystem nicht unter allen Umständen eine Lösung besitzen muss. Aber *wenn* es eine Lösung besitzt, kann man *alle Lösungen* gut beschreiben.

Lösungsraum eines inhomogenen Systems

Wenn ein inhomogenes System $Ax = b$ (mindestens) eine Lösung besitzt, so ist $L(A, b)$ eine Nebenklasse des Unterraums $L(A, 0)$. Genauer gilt: Ist k_0 eine Lösung des inhomogenen Systems, so ist

$$L(A, b) = k_0 + L(A, 0) .$$

Zuerst zeigen wir „ $L(A, b) \subseteq k_0 + L(A, 0)$ “: Sei dazu k' ein beliebiges Element von $L(A, b)$, also eine beliebige Lösung von $Ax = b$. Dann ist

$$Ak_0 = b \quad \text{und} \quad Ak' = b ,$$

also

$$A(k_0 - k') = Ak_0 - Ak' = b - b = 0 .$$

Daher liegt der Vektor $k_0 - k'$ in $L(A, 0)$. Nach dem Kriterium über Gleichheit von Nebenklassen bestimmen also k_0 und k' dieselbe Nebenklasse von $L(A, 0)$. Also ist $k' \in k_0 + L(A, 0)$.

Um die Inklusion „ $k_0 + L(A, 0) \subseteq L(A, b)$ “ zu zeigen, betrachten wir eine beliebige Lösung $h = (h_1, \dots, h_n)^T$ des homogenen Systems $Ax = 0$. Es ist also $Ah = 0$ und damit

$$A(h + k_0) = A \cdot h + A \cdot k_0 = 0 + b = b .$$

Also ist $h + k_0$ ebenfalls eine Lösung von $Ax = b$. Dies zeigt $h + k_0 \in L(A, b)$.

Zusammen folgt $L(A, b) = k_0 + L(A, 0)$ und damit alle Behauptungen des Satzes. \square

Man drückt obigen Satz traditionell oft wie folgt aus: *Man erhält die allgemeine Lösung des inhomogenen Systems, indem man zu einer speziellen Lösung des inhomogenen Systems die allgemeine Lösung des homogenen Systems addiert.*

Wir haben die Lösungsmenge eines inhomogenen Systems linearer Gleichungen beschrieben – unter der Voraussetzung, dass überhaupt eine Lösung existiert! Als nächstes müssen wir die Frage klären, wann eine Lösung existiert und wie man das erkennen kann. Hierbei spielt der Begriff des Rangs einer Matrix die entscheidende Rolle.

Sei M eine $m \times n$ -Matrix über dem Körper K . Wir fassen die Spalten von M als Vektoren aus K^m auf. Dann erzeugen die Spalten einen Unterraum von K^m ; die Dimension dieses Unterraums wird der **(Spalten-)Rang** von M genannt und mit $\text{Rang}(M)$ bezeichnet.

Kurz: Der Rang von M ist die Dimension des Erzeugnisses der Spaltenvektoren. Man kann dies auch anders ausdrücken: Da die Spalten von M einen Unterraum von K^m erzeugen, findet man in diesem Erzeugendensystem eine Basis dieses Unterraums, und zwar findet man sie so, dass man eine maximale Menge linear unabhängiger Vektoren sucht.

Damit können wir den Rang einer Matrix auch so beschreiben: *Der Rang von M ist die Maximalzahl linear unabhängiger Spalten von M .*

Zum Beispiel haben die folgenden Matrizen über \mathbf{R}

$$\begin{pmatrix} 2 & 0 & 3 \\ 0 & 1 & 6 \\ 8 & -4 & 17 \end{pmatrix}, \quad \begin{pmatrix} 2 & 0 & 2 \\ 0 & 1 & 1 \\ 8 & -4 & 4 \end{pmatrix}, \quad \begin{pmatrix} 2 & 2 & 2 \\ 0 & 0 & 0 \\ 8 & 8 & 8 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

den Rang 3, 2, 1 und 0.

Die **Einheitsmatrix** E_n hat offenbar den Rang n ; dabei ist

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

Die Matrix E_n hat also auf der **Hauptdiagonale** Einsen und sonst nur Nullen.

Nun kehren wir zu dem Gleichungssystem (4.2) zurück. Wir nennen A die zu (4.2) gehörige Matrix und nennen die Matrix $A|b$, die aus A dadurch entsteht, dass als letzte Spalte noch b angehängt wird, die **erweiterte Matrix** des Gleichungssystems (4.2).

Kriterium für die Lösbarkeit eines linearen Gleichungssystems

Genau dann ist das lineare Gleichungssystem $Ax=b$ lösbar, wenn $\text{Rang}(A) = \text{Rang}(A|b)$ ist, wenn also der Rang durch Hinzufügen der Spalte b nicht größer wird.

Der *Beweis* dieses Kriteriums ist nicht schwer: Wenn A und $A|b$ den gleichen Rang haben, dann erzeugen die Spalten von $A|b$ denselben Vektorraum wie die Spalten von A . Also kann man die Spalte b durch eine Linearkombination der Spalten von A ausdrücken. Was bedeutet dies? Dies bedeutet, dass es Körperelemente k_1, k_2, \dots, k_n gibt, so dass

$$s_1 k_1 + s_2 k_2 + \dots + s_n k_n = b$$

ist, wobei s_1, s_2, \dots, s_n die Spalten von A sein sollen. Wenn man diese Gleichung komponentenweise liest, ergibt sich

$$\begin{aligned} a_{11}k_1 + a_{12}k_2 + \dots + a_{1n}k_n &= b_1 \\ a_{21}k_1 + a_{22}k_2 + \dots + a_{2n}k_n &= b_2 \\ &\vdots \\ a_{m1}k_1 + a_{m2}k_2 + \dots + a_{mn}k_n &= b_m. \end{aligned}$$

Das heißt nichts anderes, als dass $(k_1, k_2, \dots, k_n)^T$ eine Lösung von (4.1) ist.

Umgekehrt möge das Gleichungssystem (4.1) eine Lösung $(k_1, k_2, \dots, k_n)^T$ haben. Dann ist b linear abhängig von den Spalten s_1, s_2, \dots, s_n von A . Das bedeutet, dass die Spalten von $A \mid b$ keinen größeren Vektorraum erzeugen können als die Spalten von A . Mit anderen Worten: $\text{Rang}(A \mid b) = \text{Rang}(A)$. \square

Die nächste Frage, die wir beantworten, ist die Frage nach der *Anzahl der Lösungen* eines homogenen linearen Gleichungssystems. Die wirkliche Frage hierbei ist, was das richtige Maß für die Anzahl aller Lösungen ist. Naiv könnte man natürlich sagen „Anzahl ist Anzahl“. Dagegen spricht aber, dass man dann über \mathbf{R} nur unterscheiden kann, ob ein lineares Gleichungssystem eine oder unendlich viele Lösungen hat. Wir wollen aber auch zwischen „verschiedenen Unendlichkeiten“ unterscheiden. Als besseres Maß für die Anzahl der Lösungen bietet sich die *Dimension des Lösungsraums* an: Auch wenn es (etwa über \mathbf{R}) unendlich viele Lösungen gibt, so kann man doch noch unterscheiden, ob der Lösungsraum die Dimension 1, 2, ... hat. Diese Dimension kann man auch ausrechnen:

Dimension des Lösungsraums

Sei $Ax = 0$ ein homogenes lineares Gleichungssystem, wobei A eine $m \times n$ -Matrix ist. Dann gilt

$$\dim(\mathbf{L}(A, 0)) = n - \text{Rang}(A) .$$

Wir werden dies in Abschn. 5.3 mit Hilfe linearer Abbildungen beweisen.

Dies sind sehr brauchbare Sätze über die Existenz und die Anzahl von Lösungen eines linearen Gleichungssystems. Man muss dazu allerdings den Rang bestimmen. Wir haben diesen über die Spalten einer Matrix definiert. Bei einer Matrix sind aber Spalten und Zeilen *de jure* gleichberechtigt, und es kann einem jedenfalls niemand verbieten, einen „Zeilenrang“ zu definieren. Es wird sich herausstellen, dass in der idealen Welt der Mathematik Zeilen und Spalten auch *de facto* gleichberechtigt sind. Dies müssen wir aber beweisen.

Wie man den **Zeilenrang** einer Matrix M über K zu definieren hat, ist klar: Das ist die Dimension des K -Vektorraums, der von den Zeilen von M erzeugt wird. Alternativ: Der Zeilenrang von M ist die Maximalzahl linear unabhängiger Zeilen. Nun der entscheidende Satz:

Zeilenrang = Spaltenrang

Etwas solider ausgedrückt: Für jede Matrix M ist ihr Zeilenrang gleich ihrem Spaltenrang.

Meist verwendet man den Satz so: Die Maximalzahl linear unabhängiger Zeilen von M ist gleich dem Rang von M .

Die Aussage dieses Satzes hat nicht nur ästhetischen Wert, sondern sie wird wirklich gebraucht. Manchmal tritt nämlich ganz natürlich der Spaltenrang auf (wie etwa beim Kriterium für die Lösbarkeit eines linearen Gleichungssystems), und manchmal erweist sich der Zeilenrang als besonders nützlich (etwa, wenn man lineare Gleichungssysteme effektiv lösen möchte.)

Der *Beweis* dieses Satzes ist nicht ganz offensichtlich. Um uns gut ausdrücken zu können, nennen wir eine Spalte (bzw. eine Zeile) **überflüssig**, wenn sie aus den anderen Spalten (bzw. Zeilen) linear kombiniert werden kann. Die Idee des Beweises besteht darin, so viele überflüssige Zeilen und Spalten wie möglich wegzulassen. Wir müssen dabei aufpassen, dass uns der Prozess nicht außer Kontrolle gerät. Durch Weglassen einer überflüssigen Zeile ändert sich natürlich der Zeilenrang nicht. Überraschenderweise gilt aber auch:

Durch Weglassen einer überflüssigen Spalte ändert sich der Zeilenrang nicht.

Zum *Beweis* dieser *Zwischenbehauptung* stellen wir uns vor, (o. B. d. A.) die letzte Spalte sei überflüssig. Dann kann man auch in jeder Zeile (und jeder Linearkombination von Zeilen) die letzte Komponente aus den vorigen linear kombinieren. Man verwendet nämlich einfach dieselben Koeffizienten, die man von der linearen Abhängigkeit der letzten Spalte von den vorigen erhalten hat. Als *Beispiel* können wir uns vorstellen, dass die letzte Komponente die Summe der übrigen Koeffizienten ist. Dann gilt

$$s_1 + s_2 + \dots + s_{n-1} = s_n ,$$

wenn s_i die i -te Spalte unserer Matrix ist; entsprechend gilt für die i -te Komponente

$$a_{i1} + a_{i2} + \dots + a_{i,n-1} = a_{in} \quad \text{für } i = 1, \dots, m .$$

Betrachten wir nun eine Linearkombination von Zeilen von M , die Null ergibt. Wir streichen aus jeder dieser Zeilen die letzte Komponente und nennen das Ergebnis eine „verstümmelte“ Zeile. Wir betrachten die Linearkombination mit denselben Koeffizienten der verstümmelten Zeilen. Dann muss trivialerweise auch diese Null sein.

Aber es gilt auch die Umkehrung: Wenn eine Linearkombination $\sum k_i z_i'$ der verstümmelten Zeilen z_i' Null ist, so ist auch die entsprechende Linearkombination $\sum k_i z_i$ der Originalzeilen z_i gleich Null: Betrachten wir das Beispiel, dass die letzte Komponente von z_i einfach die Summe der übrigen Komponenten ist. In diesem Fall ergibt sich

$$\sum_i k_i a_{ij} = 0 \quad \text{für } j = 1, \dots, n-1 ,$$

also

$$\sum_i k_i a_{in} = \sum_i k_i (a_{i1} + \dots + a_{i,n-1}) = \sum_i k_i a_{i1} + \dots + \sum_i k_i a_{i,n-1} = 0 ,$$

da die entsprechende Linearkombination der verstümmelten Zeilen komponentenweise Null ergibt. (Sie sind aufgefordert, dieses Argument im Allgemeinen durchzuführen, also dann, wenn die letzte Komponente nicht einfach die Summe, sondern eine allgemeine Linearkombination der ersten Komponenten ist; siehe Übungsaufgabe 6.)

Deshalb ändert sich der Zeilenrang durch Weglassen der überflüssigen letzten Spalte nicht, und damit ist die Zwischenbehauptung bewiesen.

Nun kehren wir zum eigentlichen Beweis zurück: Wir wissen jetzt, dass Weglassen einer überflüssigen Spalte weder den Spalten- noch den Zeilenrang ändert. Ebenso ändert das Weglassen von überflüssigen Zeilen weder den Spalten- noch den Zeilenrang. Tun wir also, was wir nicht lassen können und lassen überflüssige Spalten und Zeilen solange weg, bis keine Spalte und keine Zeile mehr überflüssig ist. Wir haben dadurch eine Matrix N erhalten, von der wir folgendes wissen

- Der Spaltenrang von N ist gleich dem Spaltenrang von M ;
- der Zeilenrang von N ist gleich dem Zeilenrang von M ;
- der Spaltenrang von N ist gleich der Anzahl der Spalten von N ;
- der Zeilenrang von N ist gleich der Anzahl der Zeilen von N .

Wir müssen uns jetzt also nur noch überlegen, dass N genauso viele Spalten wie Zeilen hat (man sagt dazu auch, dass N eine **quadratische Matrix** ist).

Warum hat N höchstens so viele Zeilen wie Spalten? Hätte N mehr Zeilen als Spalten, so wäre insbesondere die Maximalzahl z linear unabhängiger Zeilen größer als die Anzahl s der Spalten von N . Wir hätten also $z > s$ linear unabhängige s -Tupel: Dass das nicht geht, wissen wir aber schon längst.

Entsprechend folgt, dass die Anzahl der Spalten von N nicht größer als die Anzahl der Zeilen von N ist, und damit ist der Satz endgültig bewiesen. \square

4.1.4 Der Gaußsche Algorithmus

Zunächst stellen wir ein Instrument zur Behandlung von Matrizen vor, das in seiner praktischen Bedeutung kaum überschätzt werden kann, nämlich die elementaren Spalten- und Zeilenumformungen.

Elementare Zeilen- und Spaltenumformungen

Das Ziel der so genannten elementaren Umformungen ist es, eine Matrix einer Schönheitsoperation zu unterziehen. Das soll bedeuten, dass sich durch eine solche Operation zwar das äußere Ansehen der Matrix völlig ändert, dass aber das Wesen der Matrix erhalten bleibt, ja dass das Wesen der Matrix durch die Schönheitsoperation eigentlich erst richtig zum Vorschein kommt.

Sei M eine $m \times n$ -Matrix über dem Körper K . Unter einer **elementaren Zeilenumformung** versteht man einen der folgenden Prozesse:

Typ 1: Vertauschung zweier Zeilen,

Typ 2: Multiplikation einer Zeile mit einem Körperelement $\neq 0$,

Typ 3: Addition eines beliebigen skalaren Vielfachen einer Zeile zu einer anderen (!) Zeile.

Ganz analog definiert man die drei Typen von **elementaren Spaltenumformungen**.

Mit Hilfe dieser elementaren Umformungen kann man das Aussehen von Matrizen erheblich verändern. Schauen wir uns ein Beispiel an:

$$\begin{pmatrix} 70 & 190 & 660 \\ 7 & 20 & 66 \\ -14 & -10 & -99 \end{pmatrix} \rightarrow \begin{pmatrix} 70 & 190 & 660 \\ 70 & 200 & 660 \\ -14 & -10 & -99 \end{pmatrix} \rightarrow \begin{pmatrix} 70 & 190 & 660 \\ 0 & 10 & 0 \\ -14 & -10 & -99 \end{pmatrix} \rightarrow$$

$$\rightarrow \begin{pmatrix} 0 & 140 & 165 \\ 0 & 10 & 0 \\ -14 & -10 & -99 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 0 & 165 \\ 0 & 10 & 0 \\ -14 & -10 & -99 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Machen sie sich bei jedem Pfeil klar, welcher Typ von elementarer Zeilenumformung angewandt wurde.

Man könnte den Eindruck gewinnen, als ob durch diese Umformungen Matrizen ununterscheidbar würden. Wir müssen uns also überlegen, ob bei elementaren Umformungen das Wesen einer Matrix verändert wird, oder ob es sich wirklich „nur“ um Schönheitsoperationen handelt.

Es ändert sich nicht alles:

Invarianz des Rangs einer Matrix bei elementaren Umformungen

Bei einer elementaren Umformung bleibt der Rang einer Matrix erhalten.

Betrachten wir o. B. d. A. elementare Zeilenumformungen. Diese haben die Eigenschaft, dass der Vektorraum, der von den Zeilen erzeugt wird, vor und nach der Umformung derselbe ist. Wenn Sie's nicht glauben, so überzeugen Sie sich (zum Beispiel in Übungsaufgabe 12), dass bei jedem Typ von Umformung eine Basis in eine Basis übergeht. \square

Die erste Anwendung von elementaren Umformungen besteht darin, den Rang einer Matrix zu bestimmen. Wir versuchen, durch elementare Spalten- und Zeilenumformungen eine vorgegebene Matrix auf die folgende Gestalt zu bringen:

$$\left(\begin{array}{ccc|c} a_{11} & & & \\ & a_{22} & * & \\ & & \ddots & \\ & 0 & & \ddots & \\ & & & & a_{rr} \\ \hline & & & 0 & \end{array} \right) \quad (4.3)$$

mit $a_{ii} \neq 0$. Dabei soll die 0 unterhalb des Striches bedeuten, dass dort nur Nullen stehen; der Stern bedeutet, dass uns die Elemente, die in diesem Bereich der Matrix stehen, nicht kümmern.

Wir beweisen nun zwei Aussagen. Die erste sagt, dass wir aus einer Matrix der Gestalt (4.3) den Rang leicht ablesen können, die zweite, dass wir die Gestalt (4.3) leicht erreichen können.

1. Eine Matrix der Gestalt (4.3) hat den Rang r .

Das folgt ganz einfach: Die Nullzeilen im unteren Bereich der Matrix sind überflüssig, und die ersten r Zeilen z_1, z_2, \dots, z_r sind linear unabhängig:

Ist $z_1 k_1 + z_2 k_2 + \dots + z_r k_r = 0$, so folgt daraus zunächst $k_1 a_{11} = 0$, also $k_1 = 0$. Damit muss zwangsläufig $k_2 a_{22} = 0$, also $k_2 = 0$ sein. Sie sehen schon: Sukzessiv erhält man $k_1 = 0, k_2 = 0, \dots, k_r = 0$.

Also ist der Zeilenrang, und damit der Rang der betrachteten Matrix gleich r . \square

2. Man kann durch elementare Spalten- und Zeilenumformungen jede Matrix in die Form (4.3) bringen.

Wir geben dazu ein Verfahren an: Sei die Ausgangsmatrix bereits in folgende Gestalt gebracht

$$\left(\begin{array}{cccc|c} a_{11} & & & & \\ & a_{22} & & * & \\ & & \ddots & & * \\ & 0 & & \ddots & \\ & & & & a_{kk} \\ \hline & & & & 0 \\ & & & & B \end{array} \right),$$

wobei a_{11}, \dots, a_{kk} alle ungleich Null sein sollen.

Ist B die Nullmatrix, so haben wir unser Ziel bereits erreicht. Wenn B nicht die Nullmatrix ist, so besitzt B ein von Null verschiedenes Element, das man durch Zeilen- und Spaltenvertauschungen (Typ 1) an die Position $(k+1, k+1)$ bringen kann, ohne den „linken oberen Quadranten“ zu verändern:

Durch Operationen vom Typ 3 kann man jetzt folgende Gestalt erreichen:

$$\left(\begin{array}{cccc|c} a_{11} & & & & \\ & a_{22} & & * & \\ & & \ddots & & * \\ & 0 & & a_{kk} & \\ & & & & a_{k+1,k+1} \\ \hline & & & & 0 \\ & & & & B' \end{array} \right).$$

Indem man das Verfahren fortführt, solange es geht, erhält man schließlich die Gestalt (4.3). Beachten Sie noch, dass der Fall $k=0$ bedeutet, dass wir mit einer beliebigen Matrix beginnen können, nachdem wir ein von Null verschiedenes Element in die obere linke Ecke gebracht haben. \square

Beispiel

$$\begin{pmatrix} 1 & 1 & 2 & 2 \\ 0 & 3 & 7 & 3 \\ 5 & -4 & 0 & 2 \\ 7 & 0 & 11 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 1 & 2 & 2 \\ 0 & 3 & 7 & 3 \\ 0 & -9 & -10 & -8 \\ 0 & -7 & -3 & -13 \end{pmatrix} \\ \rightarrow \begin{pmatrix} 1 & 1 & 2 & 2 \\ 0 & 3 & 7 & 3 \\ 0 & 0 & 11 & 1 \\ 0 & 0 & 40/3 & -6 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 1 & 2 & 2 \\ 0 & 3 & 7 & 3 \\ 0 & 0 & 11 & 1 \\ 0 & 0 & 0 & -238/33 \end{pmatrix}.$$

Also hat die Ausgangsmatrix (über \mathbf{R}) den Rang 4.

Der Gaußsche Algorithmus basiert auf den folgenden beiden einfach einzusehenden Beobachtungen:

Invarianz des Lösungsraums bei elementaren Zeilenumformungen

Sei $A \mid b$ die erweiterte Matrix eines linearen Gleichungssystems. Wir wenden darauf elementare Zeilenumformungen an und erhalten ein Gleichungssystem mit erweiterter Matrix $A' \mid b'$. Dann gilt:

$$\mathbf{L}(A', b') = \mathbf{L}(A, b).$$

Beweis Vertauschung von Zeilen ändert die Lösungsmenge nicht, denn diese hängt nicht davon ab, in welcher Reihenfolge wir die zu lösenden Gleichungen hinschreiben. Wenn man eine Zeile von $A \mid b$ mit einem Skalar k multipliziert, ändert sich eine Lösung nicht, denn man multipliziert ja auch den Eintrag im Vektor b mit k . Entsprechend behandelt man elementare Zeilenumformungen vom Typ 3; siehe Übungsaufgabe 14. \square

Invarianz des Lösungsraums bei elementaren Spaltenvertauschungen

Wenn wir die Matrix $A \mid b$ durch eine Spaltenvertauschung zu $A' \mid b$ verändern, stimmen $\mathbf{L}(A, b)$ und $\mathbf{L}(A', b)$ zwar immer noch „im Prinzip“ überein, aber es könnte sein, dass die Unbekannten anders nummeriert sind.

Beweis Eine Vertauschung von Spalten entspricht einer Umnummerierung der Unbekannten. Man muss, wenn man das Gleichungssystem entsprechend umformt, also über die Umbenennungen der Unbekannten Buch führen. \square

Zunächst stellen wir den Gaußschen Algorithmus in einem Spezialfall vor.

Gaußscher Algorithmus I Sei A eine $n \times n$ -Matrix über dem Körper K mit $\text{Rang}(A) = n$, sei $b \in K^n$.

Wir betrachten das lineare Gleichungssystem $Ax = b$. Das Verfahren zur Lösung arbeitet in n Schritten.

1. Schritt Man beginnt mit der Matrix $A | b$. Da $\text{Rang}(A) = n$ ist, gibt es in der ersten Spalte ein von Null verschiedenes Element. Durch Zeilenvertauschung erreicht man, dass dieses in der ersten Zeile steht. Danach addiert man ein geeignetes Vielfaches der ersten Zeile zu jeder anderen Zeile und erreicht damit, dass alle Elemente in der ersten Spalte, die unterhalb des ersten Elements liegen, Null sind.

k-ter Schritt Vor diesem finden wir eine Matrix folgender Gestalt vor:

$$M_{k-1} = \left(\begin{array}{cccc|c} a'_{11} & & & & * \\ & a'_{22} & & 0 & * \\ & & \ddots & & \vdots \\ & 0 & & \ddots & \vdots \\ & & & & a'_{k-1,k-1} & * \\ \hline & & & 0 & & C \end{array} \right)$$

mit $a'_{11}, a'_{22}, \dots, a'_{k-1,k-1} \neq 0$.

Zunächst ist klar, dass in der ersten Spalte der Matrix C (also in der k -ten Spalte von M_{k-1} in den Zeilen k, \dots, n) ein von Null verschiedenes Element vorhanden ist. (Andernfalls wäre die k -te Spalte von M_{k-1} linear abhängig von den ersten $k-1$ Spalten, und A hätte nicht den Rang n .)

Durch Vertauschungen der letzten $n - k$ Zeilen bringen wir ein solches Element in die k -te Zeile (also an die Stelle (k, k)), und erreichen wieder, dass unterhalb und oberhalb dieses Elements nur Nullen stehen.

Nach dem $(n-1)$ -ten Schritt haben wir eine Matrix M_n folgender Gestalt erhalten:

$$M_n = \left(\begin{array}{cccc|c} a'_{11} & & & & b'_1 \\ & a'_{22} & & 0 & b'_2 \\ & & \ddots & & \vdots \\ & 0 & & a'_{n-1,n-1} & b'_{n-1} \\ & & & & a'_{nn} & b'_n \end{array} \right)$$

mit $a'_{11}, a'_{22}, \dots, a'_{nn} \neq 0$.

Daraus ergibt sich die Lösung von $Ax = b$ denkbar einfach:

$$x_i = \frac{b'_i}{a'_{ii}} \quad \text{für } i = 1, \dots, n.$$

Wir erhalten daraus eine eindeutige Lösung! Dies bedeutet:

Eindeutige Lösung eines linearen Gleichungssystems

Wenn A eine $n \times n$ -Matrix vom Rang n ist, so hat das Gleichungssystem $Ax = b$ genau eine Lösung. \square

Beispiel Um das Gleichungssystem

$$-x_1 + 2x_2 + x_3 = -2$$

$$3x_1 - 8x_2 - 2x_3 = 4$$

$$x_1 + 4x_3 = -2$$

zu lösen, starten wir mit folgender Matrix:

$$A|b = \begin{pmatrix} -1 & 2 & 1 & -2 \\ 3 & -8 & -2 & 4 \\ 1 & 0 & 4 & -2 \end{pmatrix}$$

und formen diese sukzessiv um:

$$1. \text{ Schritt: } M_1 = \begin{pmatrix} -1 & 2 & 1 & -2 \\ 0 & -2 & 1 & -2 \\ 0 & 2 & 5 & -4 \end{pmatrix}.$$

$$2. \text{ Schritt: } M_2 = \begin{pmatrix} -1 & 0 & 2 & -4 \\ 0 & -2 & 1 & -2 \\ 0 & 0 & 6 & -6 \end{pmatrix}.$$

$$3. \text{ Schritt: } M_3 = \begin{pmatrix} -1 & 0 & 0 & -2 \\ 0 & -2 & 0 & -1 \\ 0 & 0 & 6 & -6 \end{pmatrix}.$$

Daraus lesen wir ohne jede Schwierigkeit ab: $x_1 = 2$, $x_2 = 1/2$, $x_3 = -1$.

Nun lösen wir lineare Gleichungssysteme im Allgemeinen.

Gaußscher Algorithmus II Für ein beliebiges lineares Gleichungssystem (mit m Gleichungen und n Unbekannten) verfährt man wie folgt:

- Zunächst führt man das erste Gaußsche Verfahren durch, wie wenn nichts wäre, also solange es geht. Das Verfahren stoppt – zunächst –, wenn wir bei einer Matrix folgender

Gestalt angelangt sind:

$$\left(\begin{array}{ccc|c|c} a'_{11} & 0 & & * & \\ & a'_{22} & & * & * \\ & & \ddots & \vdots & \\ & 0 & & a'_{k'k'} & * \\ \hline & & & 0 & \\ & 0 & & \vdots & * \\ & & & 0 & \end{array} \right).$$

- Jetzt führen wir eine Spaltenvertauschung durch, um in einer der Zeilen $k' + 1, \dots, m$ in die $(k' + 1)$ -te Spalte ein von Null verschiedenes Element zu bekommen. Bei diesem Schritt muss man aufpassen, denn er bedeutet eine Umnummerierung der Unbekannten!

Notieren Sie, was Sie gemacht haben!!!

Damit können wir das normale Gaußsche Verfahren weiterlaufen lassen, als ob nichts gewesen wäre.

- Schließlich erhält man eine Matrix der folgenden Gestalt:

$$\left(\begin{array}{ccc|c|c} a'_{11} & 0 & & * & b'_1 \\ & a'_{22} & & & b'_2 \\ & & \ddots & & \vdots \\ & 0 & & a'_{kk} & \\ \hline & & & 0 & \\ & 0 & & & \vdots \\ & & & & b'_m \end{array} \right).$$

mit $a'_{11}, a'_{22}, \dots, a'_{kk} \neq 0$.

Hieran kann man ablesen, ob das Gleichungssystem $Ax = b$ lösbar ist oder nicht:

- Wenn auch nur eines der Elemente b'_{k+1}, \dots, b'_m ungleich Null ist, ist das Gleichungssystem nicht lösbar. (Dies sieht man direkt oder durch Anwendung des Rangkriteriums.)
- Wenn andererseits $b'_{k+1} = \dots = b'_m = 0$ gilt, so ist das Gleichungssystem lösbar.

Dies sieht man am einfachsten so: Wenn man für die letzten $n - k$ Unbekannten *irgendwelche* Körperelemente einsetzt, erhält man ein Gleichungssystem der Form $a'_{11}x_1 = c_1, \dots, a'_{kk}x_k = c_k$, das man ohne weiteres lösen kann. Damit sieht man auch, dass der Lösungsraum (des entsprechenden homogenen Systems) die Dimension $n - k$ hat.

4.2 Affine Geometrie

Ein Anstoß zur Behandlung von Vektorräumen war die Geometrie. Wir hatten bereits die „natürliche“ Identifizierung von Punkten und Vektoren (über die Translation, die den

Abb. 4.1 Das Parallelenaxiom

festen Punkt O in den Punkt P überführt) erwähnt. Dieser Zusammenhang soll jetzt deutlicher dargestellt werden. Wir beginnen mit der – zugegebenermaßen etwas länglichen – Definition eines „affinen Raumes“. Weshalb man solche Geometrien „affin“ nennt, darauf will ich hier nicht eingehen. Ich möchte Sie aber daran erinnern, dass das Vorbild für die affinen Räume unsere Anschauungsebene beziehungsweise der 3-dimensionale Anschauungsraum ist.

4.2.1 Affine Räume

Eine **Geometrie** besteht aus einer Menge P , deren Elemente wir **Punkte** nennen, einer Menge G , deren Elemente wir **Geraden** nennen, und einer Relation I (**Inzidenz**) die festlegt, wann ein Punkt mit einer Geraden inzidiert und wann nicht. Inzidiert ein Punkt P mit der Geraden g , so schreiben wir dafür $P I g$ und sagen dazu „ P liegt auf g “, ... (Sie machen keinen Fehler, wenn Sie sich die Inzidenz als mengentheoretisches Enthaltensein vorstellen; für theoretische Zwecke ist es aber besser, von einer allgemeinen „Inzidenz“ zu reden.)

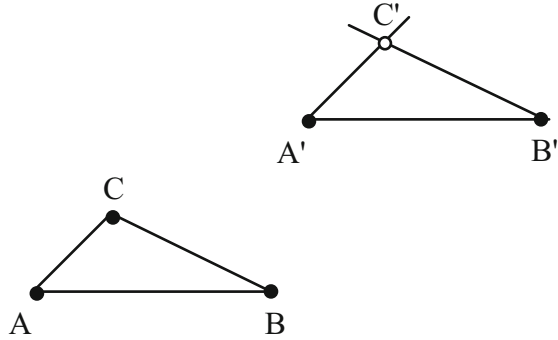
Eine solche Geometrie heißt **affiner Raum**, falls die folgenden drei Axiome gelten:

- (1) *Verbindungsaxiom*: Je zwei verschiedene Punkte P und Q inzidieren mit genau einer Geraden, die wir mit PQ bezeichnen.
- (2) *Parallelenaxiom*: Es gibt eine Äquivalenzrelation \parallel („parallel“) auf der Menge der Geraden, so dass jeder Punkt auf genau einer Geraden aus jeder Äquivalenzklasse liegt. Das bedeutet: Die Menge der Geraden ist in Klassen („Parallelscharen“) aufgeteilt; man nennt Geraden in derselben Klasse **parallel**. Es gilt: Zu jeder Geraden g und zu jedem Punkt P gibt es genau eine Gerade h durch P mit $h \parallel g$ (siehe Abb. 4.1).
- (3) *Dreiecksaxiom*: Seien A, B, C drei Punkte, die nicht auf einer gemeinsamen Geraden liegen. Seien A' und B' Punkte mit $A'B' \parallel AB$. Dann treffen sich die Parallelen zu AC durch A' und zu BC durch B' in einem Punkt C' (siehe Abb. 4.2).

Eine Struktur mit diesen Eigenschaften ist ein affiner Raum! Überzeugen Sie sich, dass die Anschauungsebene die Axiome für einen affinen Raum erfüllt. Machen Sie sich insbesondere klar, wie die Parallelscharen aussehen.

Bemerkung Man nennt das Dreiecksaxiom oft auch **Tamaschke-Axiom**, da O. Tamaschke [Tama] als erster die Bedeutung dieses Axioms erkannte.

Abb. 4.2 Das Dreiecksaxiom



Unsere Aufgabe ist es, affine Räume mit Vektorräumen in Verbindung zu bringen. Das geschieht durch die folgende zentrale Definition:

Sei V ein K -Vektorraum. Mit $\mathbf{A}(V)$ bezeichnen wir folgende Geometrie:

Die *Punkte* von $\mathbf{A}(V)$ sind die Vektoren von V ; die *Geraden* von $\mathbf{A}(V)$ sind alle Nebenklassen $v + U$ mit $v \in V$ und $\dim(U) = 1$; die *Inzidenz* von $\mathbf{A}(V)$ ist mengentheoretisches Enthaltensein.

Bevor wir im allgemeinen nachweisen, dass $\mathbf{A}(V)$ ein affiner Raum ist, machen wir uns klar, dass mit dieser Konstruktion im Falle $\dim(V) = 2$ die Geometrie der kartesischen Ebene modelliert wird.

Sei $V = \mathbf{R}^2$. Dann sind die Punkte die Vektoren von V , also die Paare (x, y) mit $x, y \in \mathbf{R}$. Das ist uns allen wohl bekannt. Ebenso wohl bekannt ist uns, dass eine Gerade durch eine Gleichung der Form $y = mx + b$ beschrieben wird. Werden dadurch ganz andere Objekte beschrieben oder ist das eine verkappte Form der „Nebenklassen-Geraden“? Überlegen wir uns dazu zunächst, welche Punktmenge von V durch die Gleichung $y = mx + b$ beschrieben wird. Im einfachsten Fall $b = 0$ sind dies Punkte der Form $y = mx$; also ist die Menge der Punkte dieser Geraden gleich

$$\{(x, mx) | x \in \mathbf{R}\} = \{x \cdot (1, m) | x \in \mathbf{R}\} = \langle (1, m) \rangle.$$

Also ist diese Gerade ein Unterraum, nämlich der 1-dimensionale Unterraum, der von dem Vektor $(1, m)$ erzeugt wird.

Nun trauen wir uns auch zu, die Gerade mit der Gleichung $y = mx + b$ als Nebenklasse zu erkennen. Die Menge der Punkte auf dieser Geraden ist gleich

$$\{(x, mx + b) | x \in \mathbf{R}\} = \{(0, b) + x(1, m) | x \in \mathbf{R}\} = (0, b) + \langle (1, m) \rangle.$$

Das heißt: Die Gerade mit der Gleichung $y = mx + b$ ist die Nebenklasse $(0, b) + \langle (1, m) \rangle$.

Es gibt eine weitere Sorte von Geraden, die in der analytischen Geometrie der Schule immer separat behandelt werden müssen; das sind Geraden mit der Gleichung $x = c$. Auch diese Geraden werden durch Nebenklassen beschrieben; es gilt nämlich

$$\{(c, y) | y \in \mathbf{R}\} = \{(c, 0) + y(0, 1) | y \in \mathbf{R}\} = (c, 0) + \langle (0, 1) \rangle.$$

Nun überlegen wir uns noch, wie sich der vertraute Parallelismusbegriff in die Sprache der Vektorräume übersetzt. Wir erinnern uns an die Schulgeometrie: Zwei Geraden sind parallel, wenn sie die gleiche Steigung haben. Wie kann man „gleiche Steigung haben“ in der Sprache der Nebenklassen ausdrücken? Seien $y = mx + b$ und $y = mx + b'$ die Gleichungen von Geraden mit gleicher Steigung m . Dann sind die zugehörigen Nebenklassen die folgenden

$$(0, b) + \langle (1, m) \rangle \quad \text{und} \quad (0, b') + \langle (1, m) \rangle .$$

Das bedeutet: Zwei Geraden $v + U$ und $v' + U'$ sind parallel, falls $U = U'$ ist.

Wir zeigen jetzt, dass $\mathbf{A} = \mathbf{A}(V)$ auch im Allgemeinen ein *affiner Raum* ist. Dazu weisen wir der Reihe nach die Axiome (1), (2) und (3) nach.

(1) *Verbindungsaxiom* Seien P und Q zwei verschiedene Punkte von $\mathbf{A}(V)$. Nach Definition gibt es zwei Vektoren v, w mit $P = v, Q = w$. Wie sehen die Nebenklassen nach 1-dimensionalen Unterräumen aus, die v und w enthalten? Wir bezeichnen eine solche Nebenklasse mit $X = x + U$.

Wegen $v, w \in X$ ist $v - w \in U$. Wegen $v \neq w$ ist $v - w \neq 0$. Also ist $v - w$ ein von Null verschiedener Vektor des 1-dimensionalen Unterraums U . Daher erzeugt $v - w$ den Unterraum U :

$$U = \langle v - w \rangle .$$

Als Repräsentant einer Nebenklasse können wir – wie wir wissen – jeden Vektor der Nebenklasse wählen. Also ist zum Beispiel

$$X = w + \langle v - w \rangle .$$

Eine andere Möglichkeit wäre, $X = v + \langle v - w \rangle$ zu schreiben. In jedem Fall ist die Nebenklasse durch v und w eindeutig bestimmt. Das bedeutet: *Die Gerade durch v und w ist die Nebenklasse*

$$w + \langle v - w \rangle .$$

Dies ist eine **Methode zur Berechnung der Geraden durch zwei Punkte**.

(2) *Parallelenaxiom* Beim Nachweis des Axioms (2) müssen wir zunächst die Parallelrelation definieren. Wir orientieren uns am Beispiel der kartesischen Ebene und definieren für zwei Geraden $v + U$ und $v' + U'$:

$$v + U \parallel v' + U' :\Leftrightarrow U = U' .$$

Damit ergibt sich, dass \parallel eine Äquivalenzrelation ist. (Wir haben die Parallelität auf eine Gleichheit, nämlich auf die Gleichheit der „begleitenden Unterräume“, zurückgeführt. Da die Gleichheitsrelation eine Äquivalenzrelation ist, ergibt sich daraus, dass auch die Parallelität eine Äquivalenzrelation ist.)

Es bleibt zu zeigen, dass durch jeden Punkt von $\mathbf{A}(V)$ genau eine Gerade einer jeden Parallelenschar geht. Sei also U ein 1-dimensionaler Unterraum, und sei v ein beliebiger Punkt von $\mathbf{A}(V)$. Dann geht mindestens eine zu U parallele Gerade durch v , nämlich $v + U$. Da je zwei Nebenklassen bezüglich desselben Unterraums disjunkt sind, kann andererseits durch v höchstens eine zu U parallele Gerade gehen.

Die **Methode zur Bestimmung von Parallelen** ist also die folgende: *Die Parallele zu $v + U$ durch den Punkt w ist die Gerade $w + U$.*

(3) *Dreiecksaxiom* Dieses Axiom weisen wir durch im Prinzip einfache Rechnung nach. Wie „einfach“ die Rechnung wird (oder ob sie überhaupt einfach wird) hängt in entscheidender Weise vom Geschick ab, die Bezeichnungen günstig zu wählen.

Zunächst modellieren wir die Ausgangssituation. Die Punkte A, B, C werden durch die Vektoren u, v, w dargestellt. Damit können wir nach der Methode zur Berechnung der Geraden durch zwei Punkte die Verbindungsgeraden der drei Punkte bestimmen:

$$AB = u + \langle v - u \rangle, BC = w + \langle v - w \rangle, CA = w + \langle u - w \rangle.$$

Sei A' durch den Vektor u' dargestellt. Da die Gerade $A'B'$ parallel zu AB ist, hat $A'B'$ nach der Methode zur Bestimmung von Parallelen die Form

$$A'B' = u' + \langle v - u \rangle.$$

Insbesondere hat B' die Form $u' + k(v - u)$. Damit können wir die Parallele b' zu AC durch A' und die Parallele a' zu BC durch B' bestimmen:

$$b' = u' + \langle u - w \rangle, a' = u' + k(v - u) + \langle v - w \rangle.$$

Wir müssen entscheiden, ob die Geraden b' und a' einen Schnittpunkt haben. Mit einem etwas geübten Auge (das durch die gute Wahl der Bezeichnungen wesentlich unterstützt wird), erkennt man, dass dies der Punkt

$$u' - ku + kw$$

ist.

Somit ist auch das Dreiecksaxiom (3) erfüllt, und $\mathbf{A}(V)$ ist tatsächlich ein affiner Raum.

4.2.2 Unterräume

Es ist verführerisch, auch höherdimensionale Unterräume zu definieren: Die **Unterräume** von $\mathbf{A}(V)$ (auch **affine Unterräume** genannt) sind genau die Nebenklassen von Unterräumen des Vektorraums V ; wenn der Unterraum U von V die Dimension t hat, so sagt man,

dass auch der affine Unterraum $v + U$ von $\mathbf{A}(V)$ die **Dimension** t hat. Wir nennen einen affinen Unterraum der Dimension 2 eine **Ebene** von $\mathbf{A}(V)$.

Wir haben es hier also mit zwei Sorten von Unterräumen zu tun: Mit denen von V und mit denen von $\mathbf{A}(V)$. Unglücklicherweise haben diese auch miteinander viel zu tun. Ich werde immer sagen, welche Sorte von Unterraum gerade betrachtet wird. Gehe hin und tue desgleichen!

Ein t -dimensionaler Unterraum des affinen Raums $\mathbf{A}(V)$ entsteht also auf folgende Weise: Man nehme einen t -dimensionalen Unterraum U von V und einen Vektor $v \in V$; man bilde die Nebenklasse $v + U$, und fertig ist der affine Unterraum.

Tragen die affinen Unterräume ihren Namen zu Recht? Über das Wörtchen „affin“ wollen wir auch jetzt nicht streiten, aber das Wort „Unterraum“ hat doch eine gewisse Bedeutung. Wir werden nicht „wilde“ Punktmengen beispielsweise als „Ebene“ bezeichnen, sondern gewisse Eigenschaften von „Ebenen“ verlangen. Welche? Meiner Ansicht nach sollten die Ebenen einer Geometrie folgende Eigenschaften haben:

- Für jede einzelne Ebene gilt: Eine Gerade hat ein ganz bestimmtes Verhältnis zu einer Ebene: Sie muss sie überhaupt nicht oder in nur einem Punkt treffen, oder sie muss ganz in der Ebene enthalten sein. Mit anderen Worten: Enthält eine Gerade zwei Punkte einer Ebene, so liegt jeder Punkt dieser Geraden in der Ebene.
- Für das System aller Ebenen gilt: Durch je drei Punkte, die nicht auf derselben Geraden liegen, geht genau eine Ebene.

Wir verifizieren die entsprechenden Eigenschaften für beliebige affine Unterräume.

Abgeschlossenheit eines affinen Unterraums gegenüber Bildung von Geraden

Jeder Punkt einer Geraden, die zwei Punkte eines affinen Unterraums enthält, liegt in diesem Unterraum.

Um dies nachweisen zu können, müssen wir Bezeichnungen einführen (Benenne und beherrsche!): Sei $v + U$ ein Unterraum von $\mathbf{A}(V)$, und seien $v + u_1, v + u_2$ zwei verschiedene Punkte von $v + U$. Wir wissen, wie man die Gerade durch diese beiden Punkte berechnet; sie ist

$$v + u_1 + \langle (v + u_2) - (v + u_1) \rangle = v + u_1 + \langle u_2 - u_1 \rangle .$$

Sei jetzt u_0 ein beliebiger Punkt dieser Geraden; es ist also $u_0 = v + u_1 + k(u_2 - u_1)$. Da U ein Unterraum von V ist und da $u_1, u_2 \in U$ sind, ist auch $u_1 + k(u_2 - u_1)$ in U . Somit ist

$$u_0 = v + u_1 + k(u_2 - u_1) \in v + U .$$

□

Die zweite Eigenschaft soll die Eigenschaft verallgemeinern, dass durch je drei nichtkollineare (das heißt nicht auf derselben Geraden liegende) Punkte genau eine Ebene geht. Am

einfachsten wäre dazu zu sagen „durch je drei Punkte geht genau eine Ebene“, aber dies ist leider nicht immer richtig; denn die drei Punkte könnten ja zufällig auf einer gemeinsamen Geraden liegen, und dann gehen viele Ebenen durch diese drei Punkte. Deshalb muss man das Kleingedruckte (nichtkollinear) auch beachten.

Auch im allgemeinen hätte man gerne, dass je $t + 1$ Punkte genau einen t -dimensionalen affinen Unterraum bestimmen. Auch hier werden wir also eine kleingedruckte Vorsichtsregel haben müssen. Nun können wir die Aussage formulieren:

Erzeugung von affinen Unterräumen

Durch je $t + 1$ Punkte, die nicht in einem gemeinsamen $(t - 1)$ -dimensionalen affinen Unterraum liegen, geht genau ein t -dimensionaler affiner Unterraum.

Zum Beweis betrachten wir Punkte v_0, v_1, \dots, v_t von $\mathbf{A}(V)$, die nicht in einem gemeinsamen $(t - 1)$ -dimensionalen Unterraum von $\mathbf{A}(V)$ liegen. Zunächst übersetzen wir diese geometrische Eigenschaft in eine Eigenschaft von V :

Behauptung Die Vektoren $v_1 - v_0, v_2 - v_0, \dots, v_t - v_0$ sind linear unabhängig.

(Wenn nicht, so wären sie in einem Unterraum W des Vektorraums V einer Dimension $t - 1$ enthalten. Dann lägen aber die fraglichen Punkte in dem Unterraum $v_0 + W$ von $\mathbf{A}(V)$. Dies ist ein Widerspruch, da $v_0 + W$ die Dimension $t - 1$ hat.)

Dieses Argument werden wir gleich nochmals benutzen.

Existenz eines t -dimensionalen Unterraums durch v_0, v_1, \dots, v_t : Aufgrund obiger Behauptung spannen $v_1 - v_0, v_2 - v_0, \dots, v_t - v_0$ einen t -dimensionalen Unterraum U des Vektorraums V auf. Dann liegen die Punkte v_0, v_1, \dots, v_t in dem t -dimensionalen affinen Unterraum $v_0 + U$.

Eindeutigkeit des t -dimensionalen Unterraums durch v_0, v_1, \dots, v_t : Seien $v + U$ und $v' + U'$ affine Unterräume der Dimension t durch v_0, v_1, \dots, v_t . Wir wissen, dass die Vektoren

$$v_1 - v_0, v_2 - v_0, \dots, v_t - v_0$$

in U und in U' enthalten sein müssen. Da diese Vektoren linear unabhängig sind, ist

$$U = \langle v_1 - v_0, v_2 - v_0, \dots, v_t - v_0 \rangle = U'.$$

Da die Nebenklassen $v + U$ und $v' + U$ gemeinsame Vektoren (nämlich v_0, v_1, \dots, v_t) enthalten, müssen sie gleich sein. \square

Wir bemerken zum Abschluss dieses Abschnitts, dass man in den Grundlagen der Geometrie umgekehrt vorgeht: Man startet mit einem („abstrakten“) affinen Raum \mathbf{A} , also einer Geometrie, die den Axiomen (1), (2), (3) genügt, und beweist dann, dass es einen Vektorraum V gibt mit $\mathbf{A} = \mathbf{A}(V)$. Dieser Fundamentalsatz gilt dann, wenn der affine Raum \mathbf{A}

mindestens die Dimension 3 hat; im Falle der Dimension 2 gibt es Gegenbeispiele (vergleichen Sie hierzu [BeuRo]).

4.3 Codierungstheorie

Die Codierungstheorie ist eine ganz angewandte Wissenschaft, die auf den ersten Blick gar nichts mit linearer Algebra zu tun hat. In Wirklichkeit ist die lineare Algebra aber ein unentbehrliches Mittel zur Analyse und Konstruktion von Codes.

Was ist das Problem? Bei jeder Übertragung oder Speicherung von Daten treten mehr oder weniger zufällig Fehler auf. Das passiert Menschen (ich bin zum Beispiel immer wieder erstaunt, wie viele Fehler ich in einem Manuskript unterbringen kann, ohne dass ich sie entdecke), aber auch Computer sind nicht unfehlbar! Deshalb braucht man Mechanismen, die solche Fehler automatisch entdecken und korrigieren. Eine Utopie? Lassen Sie sich überraschen!

Die Codierungstheorie hat sich in den letzten Jahrzehnten zu einem Gebiet von großer praktischer und theoretischer Bedeutung entwickelt. Daraus wurde eine umfangreiche Theorie entwickelt, von der wir hier die Grundlagen darstellen. Der interessierte Leser sei auf die Literatur (etwa [HeQu], [Hill] oder die „Bibel“ der Codierungstheorie [MWSI]) verwiesen. Beispiele für fehlererkennende Codes (Strichcode, ISBN-Nummern usw.) sind in [Beu1] und [Schu] eingängig dargestellt. Der Akzent in diesem Abschnitt liegt darauf, das Problem der Codierungstheorie klar herauszuarbeiten und dann zu zeigen, wie dies mit Hilfe der Methoden der linearen Algebra (mindestens) prinzipiell gelöst werden kann. In der Darstellung lehnen wir uns an [BeuRo] an.

4.3.1 Grundlegende Begriffe

Die Codierungstheorie geht von folgender Situation aus: Ein **Sender** will einem **Empfänger** gewisse **Daten** übermitteln. Diese Daten werden über einen **Kanal** übertragen, der die Nachrichten allerdings nicht fehlerfrei transportiert; es können zufällige Fehler vorkommen. Diese treten in der Regel aus physikalischen Gründen auf. Ein typisches Beispiel eines Kanals, bei dem solche „Störungen“ besonders deutlich werden, ist ein Funkkanal; wohl jeder hat sich schon bei schlechtem Radio- oder Fernsehempfang über „atmosphärische Störungen“ geärgert.

Genauer betrachtet sieht das Kommunikationsmodell wie folgt aus (siehe Abb. 4.3). Der Sender **codiert** einen Datensatz d zu einer **Nachricht** c (**Codewort**); diese wird über den Kanal geschickt. Der Empfänger versucht durch **Decodieren** zu erkennen, ob Fehler aufgetreten sind, und den Datensatz wieder zu rekonstruieren.

Die Fehler, die mit den Mitteln der Codierungstheorie behandelt werden, sind *zufällige* Fehler, also keine Veränderungen, die von einem gutmeinenden oder böswilligen Mitmenschen gezielt verursacht werden. Die Fehler, die wir behandeln, sind Veränderungen

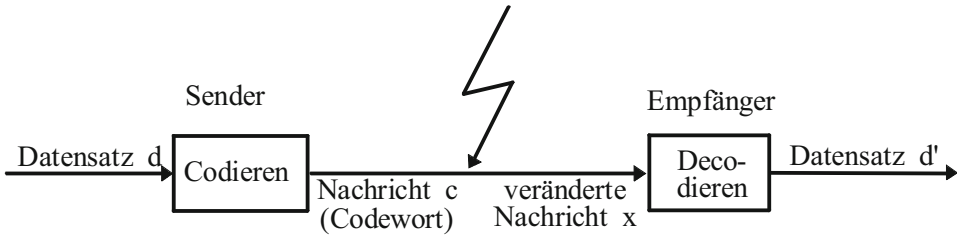


Abb. 4.3 Codierung und Decodierung von Daten

von Zeichen. Keine Fehler in diesem Sinne sind also z. B. Verlust oder Hinzufügen von Zeichen.

Unser erstes Ziel ist es, die Eigenschaft, Fehler korrigieren zu können, zu präzisieren.

In diesem Abschnitt ist eine **Nachricht** stets ein binäres n -Tupel, also ein Element der Menge $V := \{0,1\}^n$. Meist werden wir $\{0, 1\}$ nicht nur als Menge, sondern als den Körper $\text{GF}(2)$ mit zwei Elementen und damit V als den Vektorraum $V = \text{GF}(2)^n$ auffassen.

Unsere Vorstellung kann also so beschrieben werden, dass der Kanal zu dem gesendeten Vektor c (der „Nachricht“) einen **Fehlervektor** e addiert und der Empfänger den Vektor $x = c + e$ erhält. Die Aufgabe des Empfängers ist es dann, x zu decodieren, das heißt den Fehlervektor zu bestimmen, um aus x wieder c zu rekonstruieren.

Der zentrale Begriff der Codierungstheorie ist der des Hamming-Abstands.

Seien $v = (v_1, \dots, v_n)$, $w = (w_1, \dots, w_n) \in V$. Der **Abstand** $d(v, w)$ von v und w ist die Anzahl der Stellen, an denen sich v und w unterscheiden:

$$d(v, w) = |\{i \mid v_i \neq w_i\}|.$$

Oft wird d auch als **Hamming-Abstand** bezeichnet; dies geschieht zu Ehren eines der Gründerväter der Codierungstheorie Richard W. Hamming. Wir überzeugen uns zunächst, dass d den Namen „Abstand“ zu Recht trägt.

Die Funktion d ist eine Metrik auf V .

Dazu müssen wir die folgenden definierenden Eigenschaften einer Metrik nachweisen:

- (1) Da $d(v, w)$ eine Anzahl ist, ist $d(v, w) \geq 0$; ferner gilt $d(v, w) = 0$ genau dann, wenn sich v und w an keiner Stelle unterscheiden, also wenn sie gleich sind.
- (2) *Symmetrie:* Offenbar gilt $d(v, w) = d(w, v)$.
- (3) Die *Dreiecksungleichung* nachzuweisen, ist etwas kniffliger: Seien $u, v, w \in V$; es ist zu zeigen

$$d(u, w) \leq d(u, v) + d(v, w).$$

Wir können o. B. d. A. annehmen, dass sich u und w genau an den ersten $a = d(u, w)$ Stellen unterscheiden. Unter diesen a Stellen mögen b sein, an denen sich v und w unterscheiden (also u und v übereinstimmen); ferner gebe es c Stellen außerhalb der ersten

a Stellen, an denen sich v von w unterscheidet. Natürlich ist dann $d(v, w) = b + c$.

	$\overbrace{\hspace{10em}}^a$										
u	x	x	x	x	x	x	x	x	x	x	x
w	o	o	o	o	o	x	x	x	x	x	x
v	x	x	x	o	o	o	o	o	x	x	x
	$\underbrace{\hspace{5em}}_b$					$\underbrace{\hspace{6em}}_c$					

Daraus erkennt man $d(u, v) = a - b + c$. Es ergibt sich

$$d(u, v) + d(v, w) = a - b + c + b + c = a + 2c \geq a = d(u, w) .$$

□

Es wird sich zeigen, dass zur Beschreibung von Codes später die Kugeln bezüglich der Hamming-Metrik von Nutzen sein werden.

Sei $v \in V$, und sei r eine nichtnegative ganze Zahl. Dann heißt

$$S_r(v) := \{x \in V \mid d(x, v) \leq r\}$$

die **Kugel vom Radius r um den Mittelpunkt v** .

Nun sind wir in der Lage, einen fehlerkorrigierenden Code zu definieren. Man könnte zunächst einen „Code“ definieren; dies ist einfach eine beliebige Teilmenge C von $\{0, 1\}^n$. Offenbar ist dies kein sehr starker Begriff. Deshalb gehen wir sofort weiter.

Sei t eine natürliche Zahl. Eine Teilmenge C von $V = \{0, 1\}^n$ heißt ein **t -fehlerkorrigierender Code**, falls für je zwei verschiedene Elemente $v, w \in C$ gilt

$$d(v, w) \geq 2t + 1 .$$

Anders gesagt: $C \subseteq V$ ist ein t -fehlerkorrigierender Code, wenn der **Minimalabstand**

$$d(C) := \min\{d(c, c') \mid c, c' \in C, c \neq c'\}$$

von C mindestens $2t + 1$ ist. Wir nennen die Elemente eines Codes auch **Codewörter**.

Um die dieser Definition zugrunde liegende Vorstellung überzeugend erklären zu können, brauchen wir einen kleinen Hilfssatz.

Lemma über Hammingkugeln

Sei C ein t -fehlerkorrigierender Code. Dann gilt:

- (a) Zu jedem Vektor $v \in V$ gibt es höchstens ein Element $c \in C$ mit $d(v, c) \leq t$.
- (b) Die Kugeln $S_t(c)$ mit $c \in C$ sind paarweise disjunkt.

Dies ist nicht schwer zu zeigen.

- (a) *Angenommen*, es gäbe zwei verschiedene Elemente $c, c' \in C$ und einen Vektor $v \in V$ mit $d(v, c) \leq t$ und $d(v, c') \leq t$. Wegen der Dreiecksungleichung folgte daraus $d(c, c') \leq d(c, v) + d(v, c') \leq 2t$, im Widerspruch zu $d(C) \geq 2t + 1$.
- (b) folgt direkt aus (a). □

Jetzt können wir auch erläutern, warum ein t -fehlerkorrigierender Code seinen Namen zu Recht trägt: Als gesendete Nachrichten werden nur Codewörter zugelassen. Wenn während der Übertragung eines Codewortes c höchstens t Fehler auftreten, so hat der empfangene Vektor x höchstens den Abstand t zu c . Nach dem Lemma über Hammingkugeln gibt es nur ein Codewort, das einen Abstand $\leq t$ zu x hat. Der Empfänger **decodiert** x zu c .

Hier ist die Vorstellung der Kugeln besonders hilfreich: Die Tatsache, dass bei der Übertragung von c höchstens t Fehler auftreten, bedeutet, dass der empfangene Vektor jedenfalls noch in $S_t(c)$ liegt. Da nach obigem Lemma je zwei Kugeln um Codewörter disjunkt sind, kann der empfangene Vektor decodiert werden, und zwar zu dem Codewort, welches der Mittelpunkt der Kugel ist, in der x liegt.

Bemerkung Wenn pro Codewort mehr als t Fehler auftreten, wird der empfangene Vektor im Allgemeinen nicht korrekt decodiert. In der Praxis wird man so vorgehen, dass man zunächst abschätzt, wie fehleranfällig der Kanal ist, dann die Zahl t entsprechend wählt und schließlich einen t -fehlerkorrigierenden Code konstruiert.

Wir können nun das **Ziel der Codierungstheorie** klar formulieren. Das Ziel ist es, Codes zu konstruieren, die

- einen großen Minimalabstand (und damit gute Fehlerkorrektureigenschaften) haben und
- für die es einen effizienten Decodieralgorithmus gibt.

Wir beschließen diesen Abschnitt mit dem Beispiel eines Codes, der noch relativ klein ist, aber doch deutlich macht, dass unser bisheriges Instrumentarium noch nicht ausreichend ist. Dieser Code wird uns mehrfach als Beispiel dienen.

Die folgenden 16 Vektoren aus $V = \{0,1\}^7$ bilden einen 1-fehlerkorrigierenden Code:

0	0	0	0	0	0	0	1	1	1	1	1	1	1
0	0	0	1	1	1	0	1	1	1	0	0	0	1
0	0	1	0	1	1	1	1	1	0	1	0	0	0
0	0	1	1	0	0	1	1	1	0	0	1	1	0
0	1	0	0	1	0	1	1	0	1	1	0	1	0
0	1	0	1	0	1	1	1	0	1	0	1	0	0
0	1	1	0	0	1	0	1	0	0	1	1	0	1
0	1	1	1	1	1	0	0	1	0	0	0	1	1

In Übungsaufgabe 20 sollen Sie diesen Code untersuchen.

4.3.2 Lineare Codes

Unser bisheriger naiver Ansatz liefert Lösungen, die weit entfernt von jeder Praktikabilität sind. Das fängt beim Speichern des Codes an (man muss jedes Codewort abspeichern), geht über die Bestimmung des Minimalabstands (man muss je zwei Codewörter vergleichen, hat also quadratischen Aufwand in $|C|$) und endet schließlich bei den Decodieralgorithmen (bei jedem empfangenen Vektor muss man alle Codewörter untersuchen).

Das Zauberwort für den entscheidenden Schritt in Richtung praktischer Anwendung heißt „lineare Codes“, und dies bedeutet nichts anderes als Anwendung der linearen Algebra.

Ein Code $C \subseteq V$ heißt **linear**, falls C ein *Unterraum* des Vektorraums V (und nicht nur eine *Teilmenge* der Menge V) ist. In diesem Fall hat C auch eine Dimension; diese wird oft mit dem Buchstaben k bezeichnet.

Ein erster Vorteil linearer Codes ist sofort einsichtig: Um mit C arbeiten zu können, braucht man nur eine **Basis** von C zu kennen. Sei c_1, \dots, c_k eine Basis eines linearen Codes C . Dann heißt die $k \times n$ -Matrix G , deren i -te Zeile der Basisvektor c_i ist, eine **Generator-matrix** von C .

Da der Speicheraufwand für eine Generatormatrix k Vektoren sind, bedeutet dies eine enorme Ersparnis im Vergleich zu den 2^k Vektoren, aus denen C besteht.

Beispiel Eine Generatormatrix des Codes aus Beispiel 4.3.1 ist

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

Auch die Bestimmung des Minimalabstands ist bei linearen Codes viel leichter als im allgemeinen Fall. Um das einsehen zu können, brauchen wir folgende Definition.

Das **Gewicht** $w(x)$ eines Vektors $x \in V$ ist die Anzahl der von 0 verschiedenen Stellen von x . Mit anderen Worten:

$$w(x) = d(x, 0).$$

Das **Minimalgewicht** $w(C)$ des Codes C ist definiert als

$$w(C) := \min\{w(c) \mid c \in C, c \neq 0\}.$$

Lemma über das Minimalgewicht

Sei C ein linearer Code. Dann gilt

$$d(C) = w(C).$$

Mit anderen Worten: Den Minimalabstand eines linearen Codes kann man einfach dadurch berechnen, dass man das Minimalgewicht bestimmt.

Beweis Für jeden Code, der den Nullvektor enthält, gilt

$$d(C) = \min\{d(c, c') \mid c, c' \in C, c \neq c'\} \leq \min\{d(c, 0) \mid c \in C, c \neq 0\} = w(C).$$

Für die umgekehrte Richtung ist zu zeigen, dass es ein Codewort c_0 vom Gewicht $d(C)$ gibt. Seien $c, c' \in C$ mit $d(c, c') = d(C)$. Dann gilt:

$$w(c - c') = d(c - c', 0) = d(c, c') = d(C).$$

Da C linear ist, ist $c_0 := c - c' \in C$. Damit ist alles gezeigt. \square

Um den Minimalabstand, und damit die Fehlerkorrekturqualität von C zu bestimmen, muss man also nur das Minimalgewicht ausrechnen; dazu braucht man höchstens $|C|$ Schritte.

Zur Erklärung, wie man mit einem linearen Code decodiert, müssen wir etwas ausholen.

Sei $C \subseteq V$ ein Code. Der zu C **duale Code** C^\perp ist wie folgt definiert:

$$C^\perp := \{v \in V \mid c \cdot v = 0 \text{ für alle } c \in C\};$$

dabei ist das **innere Produkt** $c \cdot v$ der Vektoren $c = (c_1, \dots, c_n)$ und $v = (v_1, \dots, v_n)$ erklärt durch

$$c \cdot v = c_1 v_1 + c_2 v_2 + \dots + c_n v_n.$$

Wenn $c \cdot v = 0$ ist, so sagt man auch, dass c und v **orthogonal** sind (oder *senkrecht aufeinander stehen*).

Vorsicht! Dieses innere Produkt sieht so ähnlich aus wie ein „Skalarprodukt“ (siehe Kap. 10). Das hier betrachtete innere Produkt hat aber sehr merkwürdige Eigenschaften: C und C^\perp schneiden sich im Allgemeinen keineswegs nur im Nullvektor; unser Beispielformat C erfüllt sogar $C^\perp \subseteq C$. Dieses Phänomen kommt hauptsächlich daher, dass wir nicht über dem Körper \mathbf{R} , sondern über \mathbf{Z}_2 arbeiten.

Bemerkung „Dualer“ Code ist eine schlechte Bezeichnung, die nichts, aber auch gar nichts mit dem „dualen Vektorraum“ aus Kap. 5 zu tun hat.

Dimensionsformel für den dualen Code

Ist C ein linearer Code der Dimension k , so ist C^\perp ein Untervektorraum von V der Dimension $n - k$.

Beweis Unabhängig davon, ob C linear ist oder nicht, ist klar, dass C^\perp ein Unterraum von V ist. Es ist zu zeigen, dass C^\perp die Dimension $n - k$ hat.

Dazu betrachten wir eine Generatormatrix G mit den Zeilen c_1, \dots, c_k von C . Dann gilt:

$$C^\perp = \{v \in V \mid c_i \cdot v = 0, i = 1, \dots, k\}.$$

Wenn man die vorige Zeile unschuldig betrachtet, erkennt man, dass nach denjenigen Vektoren $v = (v_1, \dots, v_n) \in V$ gefragt ist, die Lösungen des homogenen Gleichungssystems mit der Koeffizientenmatrix G sind. In Abschn. 4.1.3 haben wir uns klar gemacht, dass die Dimension des Lösungsraums gleich $n - \text{Rang}(G)$ ist. Da die Zeilen von G eine Basis von C bilden, hat G den Rang k . Also gilt $\dim(C^\perp) = n - k$. \square

Man kann die Bildung des dualen Codes natürlich iterieren; dass man dadurch nichts Neues erhält (und also die Iteration nicht durchzuführen braucht), sagt der folgende Satz.

Satz vom „Bidualcode“

Sei C ein linearer Code. Dann ist

$$C^{\perp\perp} = C.$$

Beweis Zunächst machen wir uns klar, dass $C \subseteq C^{\perp\perp}$ gilt: $C^{\perp\perp}$ besteht aus all den Vektoren, die orthogonal zu allen Vektoren aus C^\perp sind; dazu gehören aber bestimmt die Vektoren aus C , da C^\perp ja die Menge derjenigen Vektoren ist, die orthogonal zu jedem Vektor aus C sind. (Lesen Sie diesen Satz nochmals ganz langsam, dann werden Sie ihn verstehen!)

Wenn wir die Dimensionsformel auf C^\perp anwenden, erhalten wir

$$\dim(C^{\perp\perp}) = n - \dim(C^\perp) = n - (n - k) = k = \dim(C).$$

Zusammen folgt $C^{\perp\perp} = C$. \square

Sei $C \subseteq V$ ein linearer Code. Eine Matrix H , deren Zeilen eine Basis des dualen Codes C^\perp bilden, heißt eine **Kontrollmatrix** von C .

Da C^\perp die Dimension $n - k$ hat, hat jede Kontrollmatrix von C genau $n - k$ Zeilen und n Spalten.

Um zu sehen, dass decodieren bei linearen Codes einfacher ist als bei nichtlinearen Codes, ist der Begriff des Syndroms eines Vektors wichtig. Sei H eine Kontrollmatrix des

linearen Codes $C \subseteq V$. Für jeden Vektor $v \in V$ definieren wir sein **Syndrom** als

$$s(v) := v \cdot H^T,$$

wobei H^T die zu H transponierte Matrix ist. (Das ist die Matrix, die aus H durch Vertauschen von Zeilen und Spalten entsteht.)

Ein Syndrom ist also ein binärer Vektor der Länge $n - k$.

Mit Hilfe einer Kontrollmatrix und des Syndroms kann man einen linearen Code hervorragend beschreiben:

Satz über die Kontrollmatrix

Ist C ein linearer Code mit Kontrollmatrix H , so gilt

$$C = \{v \in V \mid s(v) = 0\}.$$

Beweis Sei $v \in V$ beliebig. Dann gilt

$$\begin{aligned} s(v) &= 0 \\ \Leftrightarrow v \cdot H^T &= 0 \\ \Leftrightarrow v &\text{ ist orthogonal zu allen Vektoren einer Basis von } C^\perp \\ \Leftrightarrow v &\in C^{\perp\perp} \\ \Leftrightarrow v &\in C \text{ wegen } C^{\perp\perp} = C. \end{aligned}$$

□

Für die Syndromdecodierung ist die folgende Beobachtung entscheidend, die sagt, dass das Syndrom $s(v)$ nur von der Nebenklasse abhängt, in der v liegt.

Sei H eine Kontrollmatrix eines linearen Codes $C \subseteq V$. Für alle Vektoren $v, w \in V$ gilt

$$s(v) = s(w) \Leftrightarrow v + C = w + C.$$

Beweis Sei $v, w \in V$ beliebig. Dann gilt

$$\begin{aligned} s(v) &= s(w) \\ \Leftrightarrow v \cdot H^T &= w \cdot H^T \\ \Leftrightarrow v \cdot H^T - w \cdot H^T &= 0 \\ \Leftrightarrow (v - w) \cdot H^T &= 0 \\ \Leftrightarrow v - w &\in C \\ \Leftrightarrow v + C &= w + C. \end{aligned}$$

□

Nun können wir beschreiben, wie man mit Hilfe eines allgemeinen linearen Codes decodieren kann. Dazu machen wir scheinbar einen Umweg: Wir überlegen uns, was die möglichen Fehler mit den Nebenklassen des Codes C zu tun haben.

Sei $C \subseteq V$ ein linearer Code. Dann ist C ein Unterraum von V , wir können also von den Nebenklassen von C sprechen. Ein Vektor heißt **Anführer** einer Nebenklasse von C , wenn er unter allen Vektoren dieser Nebenklasse minimales Gewicht hat.

Im Allgemeinen sind Nebenklassenanführer nicht eindeutig bestimmt. Jedoch gilt folgende Tatsache, welche die Bedeutung dieses Konzeptes klar macht:

Eindeutigkeit der Nebenklassenanführer

Sei $C \subseteq V$ ein linearer t -fehlerkorrigierender Code. Dann gilt:

- (a) Jeder Vektor von V vom Gewicht $\leq t$ ist Anführer einer Nebenklasse.
- (b) Die Anführer von Nebenklassen, die einen Vektor vom Gewicht $\leq t$ enthalten, sind eindeutig bestimmt.

Wir *beweisen* (a) und (b) gemeinsam. Sei v ein Vektor vom Gewicht $\leq t$. Betrachte einen beliebigen Vektor $v' \in v + C$ mit $v' \neq v$. Es ist zu zeigen, dass v' mindestens das Gewicht $t + 1$ hat.

Da v und v' in derselben Nebenklasse von C sind, ist $v - v' \in C$. Da $v \neq v'$ ist, gilt $v - v' \neq 0$, also $w(v - v') \geq 2t + 1$ nach Definition eines t -fehlerkorrigierenden Codes. Daraus folgt

$$\begin{aligned} 2t + 1 &\leq w(v - v') = d(v - v', 0) = d(v, v') \\ &\leq d(v, 0) + d(0, v') = w(v) + w(v') \leq t + w(v') , \end{aligned}$$

also $w(v') \geq t + 1$. □

Damit haben wir einen Decodieralgorithmus: Man bestimme zunächst die Nebenklasse von C , in der der empfangene Vektor x liegt. Wenn höchstens t Fehler aufgetreten sind, dann hat der Fehlervektor das gleiche Syndrom wie x ; also ist der Fehlervektor der Anführer der Nebenklasse, in der x liegt. Man addiere den Anführer zu x und erhält das Codewort zurück. Wenn mehr als t Fehler aufgetreten sind, kann man auch mit Hilfe des Anführers der Nebenklasse $x + C$ korrigieren; allerdings ist man dann nicht sicher, ob richtig decodiert wurde.

Diese Prozedur kann noch etwas geschickter organisiert werden:

Syndrom-Decodierung

Sei $C \subseteq V$ ein linearer Code, der t -fehlerkorrigierend ist. Man erstellt eine Liste der Nebenklassenanhänger und der zugehörigen Syndrome. Für einen empfangenen Vektor x

- berechnet man das Syndrom $s(x)$,
- sucht dies in der Liste der Syndrome,
- stellt den zugehörigen Nebenklassenanhänger e fest und
- decodiert x zu $c := x + e$.

Wenn zum Beispiel *kein* Fehler passiert ist, so ist das Syndrom von x der Nullvektor, die zugehörige Nebenklasse also C ; dann ist der Anhänger der Nullvektor von C , und x wird richtig zu $x + o = x$ decodiert.

Die Eindeutigkeit der Nebenklassenanhänger garantiert, dass mit der Syndrom-Decodierung richtig decodiert wird, wenn höchstens t Fehler auftreten.

Zur Illustration der Syndrom-Decodierung greifen wir das *Beispiel* aus Abschn. 4.3.1 auf. Zunächst müssen wir eine Kontrollmatrix bestimmen. Wir können einfach verifizieren, dass

$$H = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{pmatrix}$$

eine Kontrollmatrix von C ist. (Dazu muss man nur nachweisen, dass jede Zeile von H auf jeder Zeile einer Generatormatrix von C senkrecht steht und dass die Zeilen linear unabhängig sind; in Übungsaufgabe 24 werden Sie eingeladen, diese Arbeit zu tun.)

Die Nebenklassenanhänger sind die Vektoren 0000000, 0000001, 0000010, ... (Dies können Sie sich zum Beispiel dadurch klarmachen, dass Sie ausrechnen, dass die angegebenen Vektoren in verschiedenen Nebenklassen liegen. Nach dem Satz über die Anzahl der Nebenklassen (Abschn. 3.3) gibt es aber auch nur acht Nebenklassen von C .)

Die Liste der Nebenklassenanhänger mit ihren Syndromen ist also die folgende:

Nebenklassenanhänger	Syndrom
0000000	000
0000001	111
0000010	011
0000100	101
0001000	110
0010000	001
0100000	010
1000000	100

Wird beispielsweise der Vektor $x = 0010001$ empfangen, so berechnet man sein Syndrom $s(x) = 110$. Danach bestimmt man aus der Liste den Fehlervektor $e = 0001000$; als Codewort ergibt sich

$$c = x + e = 0010001 + 0001000 = 0011001 .$$

4.4 Verständnisfragen, Übungen und Tipps

Richtig oder falsch?

1. Thema: Lineare Gleichungssysteme

- ☐ Jedes lineare Gleichungssystem hat mindestens eine Lösung.
- ☐ Jedes homogene lineare Gleichungssystem hat mindestens eine Lösung.
- ☐ Jedes homogene lineare Gleichungssystem hat mindestens zwei Lösungen.
- ☐ Jedes homogene lineare Gleichungssystem über \mathbf{R} , das mindestens zwei Lösungen hat, hat unendlich viele.
- ☐ Jedes inhomogene lineare Gleichungssystem hat höchstens eine Lösung.

2. Thema: Matrizenmultiplikation

Alle in dieser Aufgabe auftretenden Matrizen seien über einem Körper K definiert; A und B seien zwei solche Matrizen.

- ☐ Man kann je zwei Matrizen miteinander multiplizieren.
- ☐ Man kann jede Matrix mit sich selbst multiplizieren.
Man kann das Produkt $A \cdot B$ von A und B bilden, wenn die folgenden Zahlen gleich sind:
 - ☐ Die Anzahl der Zeilen von A und die Anzahl der Spalten von B .
 - ☐ Die Anzahl der Spalten von A und die Anzahl der Zeilen von B .
 - ☐ Die Anzahl der Spalten von A und die Anzahl der Spalten von B .
 - ☐ Die Anzahl der Zeilen von A sowie die Anzahl der Zeilen und Spalten von B .
- ☐ Wenn $A \cdot B$ definiert ist, so ist auch $B \cdot A$ definiert.
- ☐ Wenn $A \cdot B$ und $B \cdot A$ definiert ist, dann ist $A \cdot B = B \cdot A$.
- ☐ Wenn $A \cdot B$ und $B \cdot A$ definiert ist, dann ist $A \cdot B \neq B \cdot A$.
- ☐ Wenn A und B verschieden von der Nullmatrix sind, dann ist auch $A \cdot B$ verschieden von der Nullmatrix.
- ☐ Wenn A verschieden von der Nullmatrix ist, dann ist auch $A \cdot A$ verschieden von der Nullmatrix.

3. Thema: Affine Unterräume

- ☐ Jeder Untervektorraum ist ein affiner Unterraum.
- ☐ Jeder affine Unterraum ist ein Untervektorraum.
- ☐ Manche affinen Unterräume sind Untervektorräume.

4. Thema: Codes

- ☐ Wenn ein Code den Nullvektor enthält, ist er linear.

- ☐ Wenn ein Code genau 2^k ($k \in \mathbb{N}$) Elemente hat, ist er linear.
- ☐ Ist $w(C) = d(C)$, so ist C linear.
- ☐ Ein t -fehlerkorrigierender Code hat das Minimalgewicht $\geq 2t + 1$.
- ☐ Jedes Codewort eines t -fehlerkorrigierenden Codes hat Gewicht $2t + 1$.

Übungsaufgaben

1. Berechnen Sie

$$\begin{pmatrix} 2 & 0 & 1 & 4 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 0 \\ 1 \\ 4 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 2 \\ 0 \\ 1 \\ 4 \end{pmatrix} \cdot \begin{pmatrix} 2 & 0 & 1 & 4 \end{pmatrix}.$$

2. Machen Sie sich klar, dass für jede $n \times n$ -Matrix A gilt

$$A \cdot E_n = A = E_n \cdot A,$$

wobei E_n die $n \times n$ -Einheitsmatrix ist.

3. Zeigen Sie, dass die Matrizenmultiplikation homogen ist, dass also für jede $m \times n$ -Matrix A über K , jeden Spaltenvektor x der Länge n mit Elementen aus K und jedes Element $k \in K$ gilt:

$$A \times (kx) = (kA) \times x = k \times (Ax).$$

4. Beweisen Sie den Satz „Zeilenrang = Spaltenrang“ im Allgemeinen.
5. Berechnen Sie die **Potenzen** der folgenden Matrix A (das heißt die Matrizen A, A^2, A^3, A^4, \dots):

$$A := \begin{pmatrix} 0 & \cdots & 0 & 0 & 1 \\ 0 & \cdots & 0 & 1 & 0 \\ 0 & \cdots & 1 & 0 & 0 \\ \cdots & & & \cdots & \\ 1 & 0 & \cdots & 0 & 0 \end{pmatrix}$$

Die Matrix A hat als Elemente auf der **Nebendiagonale** Einsen und sonst nur Nullen.

6. Welchen Rang haben die folgenden reellen Matrizen?

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 & 0 \\ 4 & 5 & 6 & 0 \\ 7 & 8 & 9 & 0 \\ 10 & 11 & 12 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 & 0 \\ 7 & 2 & 0 & 3 \\ 1 & -1 & 3 & -8 \\ -1 & 5 & 3 & 8 \end{pmatrix}.$$

7. Sei A eine $m \times n$ -Matrix und B eine $n \times s$ -Matrix über dem Körper K . Zeigen sie:

$$\text{Rang}(A \times B) \leq \text{Rang}(A) .$$

(Hinweis: Machen Sie sich klar, dass der von den Spalten von $A \cdot B$ erzeugte Unterraum in dem von den Spalten von A erzeugten Unterraum von K^m enthalten ist.)

8. Bestimmen Sie den Rang der folgenden reellen Matrizen A , B und $A \cdot B$:

$$A = \begin{pmatrix} 3 & 5 & 1 & 4 \\ 2 & -1 & 1 & 1 \\ 8 & 9 & 3 & 9 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 1 & 6 & 6 \\ 3 & 1 & 1 & -1 \\ 5 & 2 & 7 & 5 \\ -2 & 4 & 3 & 2 \end{pmatrix} .$$

9. Bestimmen Sie den Rang der folgenden reellen $n \times n$ -Matrix:

$$\begin{pmatrix} a & b & b & \dots & b \\ b & a & b & \dots & b \\ b & b & a & \dots & b \\ \dots & \dots & \dots & \dots & \dots \\ b & b & b & \dots & a \end{pmatrix} .$$

Dabei sind a und b verschiedene reelle Zahlen.

[Hinweis: Subtrahieren Sie zunächst die erste Zeile von allen anderen Zeilen.]

10. Zeigen Sie: Bei einer elementaren Umformung ändert sich der Rang einer Matrix nicht.
11. Sei M eine $n \times n$ -Matrix über dem Körper K . Zeigen Sie: Wenn M den Rang n hat, dann kann man M allein durch elementare *Zeilen*vertauschungen in die Einheitsmatrix verwandeln.
- [Hinweis: Beginnen Sie mit der 1. Spalte, betrachten Sie dann die 2. Spalte usw.]
12. Zeigen Sie, dass sich der Lösungsraum eines linearen Gleichungssystems bei elementaren Zeilenumformungen vom Typ 3 nicht ändert.
13. Lösen Sie das folgende lineare Gleichungssystem über \mathbf{R} :

$$-2x + 3y + 2z + 4w = 0$$

$$x - y + 2z + 3w = 0$$

$$2x + y + 2z - 2w = 0 .$$

14. Stellen Sie fest, ob das folgende reelle lineare Gleichungssystem lösbar ist und bestimmen Sie gegebenenfalls die Lösungsmenge:

$$\begin{aligned}x_1 + 2x_2 + 3x_3 &= 1 \\4x_1 + 5x_2 + 6x_3 &= 2 \\7x_1 + 8x_2 + 9x_3 &= 3 \\5x_1 + 7x_2 + 9x_3 &= 4.\end{aligned}$$

15. Lösen Sie das folgende lineare Gleichungssystem über \mathbf{R} :

$$\begin{aligned}2x_1 + x_2 + x_3 &= a_1 \\5x_1 + 4x_2 - 5x_3 &= a_2 \\3x_1 + 2x_2 - x_3 &= a_3,\end{aligned}$$

wobei

(a) $a_1 = 5, a_2 = -1, a_3 = 3$,

(b) $a_1 = 1, a_2 = -1, a_3 = 1$

zu wählen sind.

16. Bestimmen Sie $a, b \in \mathbf{R}$ so, dass das folgende lineare Gleichungssystem über \mathbf{R} lösbar ist:

$$\begin{aligned}2x_1 + x_2 + x_3 &= -1 \\5x_1 + 4x_2 - 5x_3 &= a \\3x_1 + 2x_2 - x_3 &= b.\end{aligned}$$

17. Sei $A \in K^{n \times n}$. Zeigen Sie: Wenn $\text{Rang}(A) = n$ ist, so hat das lineare Gleichungssystem

$$Ax = b$$

unabhängig von der Gestalt von b genau eine Lösung.

18. Sei V ein K -Vektorraum. Für einen *affinen* Unterraum U definieren wir

$$\mathbf{L} = \mathbf{L}(U) := \{x - y \mid x, y \in U\}.$$

Zeigen Sie, dass $\mathbf{L}(U)$ ein Unterraum des Vektorraums V ist.

19. Sei V ein K -Vektorraum. Wir nennen eine Linearkombination $k_1 v_1 + k_2 v_2 + \dots + k_n v_n$ eine **affine Linearkombination**, falls $k_1 + k_2 + \dots + k_n = 1$ ist.

Zeigen Sie: Jede affine Linearkombination von Vektoren eines affinen Unterraums U liegt ebenfalls in U .

20. Zeigen Sie, dass der folgende Code 1-fehlerkorrigierend ist:

0	0	0	0	0	0	0	1	1	1	1	1	1	1
0	0	0	1	1	1	0	1	1	1	0	0	0	1
0	0	1	0	1	1	1	1	1	0	1	0	0	0
0	0	1	1	0	0	1	1	1	0	0	1	1	0
0	1	0	0	1	0	1	1	0	1	1	0	1	0
0	1	0	1	0	1	1	1	0	1	0	1	0	0
0	1	1	0	0	1	0	1	0	0	1	1	0	1
0	1	1	1	1	0	0	1	0	0	0	0	1	1

21. Zeigen Sie, dass die Matrix H , deren Spalten die sämtlichen von Null verschiedenen binären r -Tupel sind, den Rang r hat.

22. Ein Code $C \subseteq \{0, 1\}^n$ heißt **s -fehlererkennend**, falls kein Codewort aus C durch Hinzufügen von höchstens s Fehlern wieder zu einem Codewort aus C wird.

(a) Machen Sie sich klar, dass dieser Name zu Recht besteht.

(b) Zeigen Sie: Ein Code C ist genau dann s -fehlererkennend, wenn $d(C) \geq s + 1$ ist.

23. Sei C die folgendermaßen definierte Teilmenge von $\{0, 1\}^n$:

$$C := (a_1, \dots, a_n) \mid a_1 + \dots + a_n \text{ ist gerade}.$$

Zeigen Sie: C ist ein linearer 1-fehlererkennender Code mit Minimalgewicht 2.

24. Zeigen Sie, dass die Matrix

$$H = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{pmatrix}$$

eine Kontrollmatrix des Codes aus Abschn. 4.3.1 ist.

Projekt: Die Hamming-Codes

Die Hamming Codes gehören zu den wichtigsten (und einfachsten!) Codes. Sie sind wie folgt definiert. Man wählt sich eine natürliche Zahl r . Dann gibt es genau $2^r - 1$ vom Nullvektor verschiedene binäre r -Tupel. Wenn wir diese (in beliebiger Reihenfolge) als Spalten in eine Matrix schreiben, erhalten wir eine binäre $r \times (2^r - 1)$ -Matrix H , deren Spalten sämtliche von 0 verschiedene binäre r -Tupel sind. Nun definieren wir den **Hamming-Code** der Länge $n = 2^r - 1$ durch

$$\text{Ham}(r) := \{c = (c_1, \dots, c_n) \in \{0, 1\}^n \mid c \cdot H^T = 0\}.$$

Das bedeutet genau diejenigen Vektoren c sind Codewörter des Hamming-Codes $\text{Ham}(r)$, für die $c \times H^T$ der Nullvektor der Länge r ist.

In diesem Projekt sei stets r eine natürliche Zahl und H eine $r \times (2^r - 1)$ -Matrix, deren Spalten die sämtlichen von 0 verschiedenen binären r -Tupel sind.

1. Berechnen Sie sämtliche Codewörter von $\text{Ham}(3)$.
2. Gibt es in $\text{Ham}(r)$ Codewörter vom Gewicht 1, 2 oder 3? Also gilt $d(\text{Ham}(r)) \geq \dots$
[Angenommen, es gäbe ein Codewort vom Gewicht 1 oder 2. Schließen Sie daraus, dass eine bzw. zwei Zeilen von H linear abhängig sein müssten.]
3. Zeigen Sie: $d(\text{Ham}(r)) \leq 3$. Zusammen mit dem zweiten Schritt folgt, dass $\text{Ham}(r)$ ein 1-fehlerkorrigierender Code ist.
4. Das obige Argument kann man verallgemeinern: Sei C ein beliebiger linearer Code, und sei H eine Kontrollmatrix von C . Dann gilt:
Wenn je $d-1$ Spalten von H linear unabhängig sind, so gilt $d(C) \geq d$.

Als nächstes sollen Sie zeigen, dass $\text{Ham}(r)$ ein „perfekter“ Code ist. Ein t -fehlerkorrigierender Code $C \subseteq \{0,1\}^n$ heißt **perfekt**, falls es zu jedem Vektor v aus $\{0,1\}^n$ mindestens ein Codewort c gibt mit $d(v, c) \leq t$. Mit anderen Worten: C ist perfekt, wenn die Hammingkugeln $S_t(c)$ vom Radius t um die Codewörter c den gesamten Raum $\{0,1\}^n$ ausfüllen. Perfekte Codes sind also diejenigen, bei denen die Hammingkugeln $S_t(c)$ so dicht wie nur denkbar gepackt sind: Jeder Vektor liegt in (genau) einer solchen Kugel.

5. Sei C ein binärer Code der Länge n .
 - (a) Dann enthält eine Hammingkugel $S_1(c)$ genau $n + 1$ Vektoren.
 - (b) Eine Hammingkugel $S_1(c)$ eines Hamming-Codes $\text{Ham}(r)$ enthält also genau ... Vektoren.
6. (**Hamming-Schranke**). Sei C ein binärer 1-fehlerkorrigierender Code der Länge n . Zeigen Sie

$$|C| \leq 2^n / (n + 1) .$$

Gleichheit gilt genau dann, wenn C perfekt ist.

7. Zeigen Sie: Die Hamming-Codes $\text{Ham}(r)$ sind perfekte 1-fehlerkorrigierende Codes.

Nun kommen wir zur Decodierung der Hamming-Codes. Diese ist ganz besonders schön und einfach.

8. Sei v ein binärer Vektor der Länge $2^r - 1$, der aus einem Codewort des Hamming-Codes $\text{Ham}(r)$ dadurch hervorgeht, dass genau ein Fehler passiert. Berechnen Sie das Syndrom von v .

Wir erinnern uns an die Definition von H : Als Matrix H konnten wir *irgendeine* $r \times (2^r - 1)$ -Matrix wählen, deren Spalten die von Null verschiedenen binären r -Tupel sind; die Ordnung der Spalten spielte bei unseren bisherigen Überlegungen keine Rolle. Nun interpretieren wir die Spalten von H als binäre Darstellung der Zahlen $1, \dots, 2^r - 1$

und ordnen die Spalten so an, dass die i -te Spalte s_i die Zahl i darstellt. Zum Beispiel ist die letzte Spalte diejenige, die nur aus Einsen besteht.

9. Sei v ein binärer Vektor der Länge $2^r - 1$, der aus einem Codewort des Hamming-Codes $\text{Ham}(r)$ dadurch hervorgeht, dass an genau einer Stelle ein Fehler passiert. Zeigen Sie: Das Syndrom von v ist die Binärdarstellung der Zahl, welche die Stelle angibt, an der der Fehler auftrat.

Schließlich studieren wir die Frage, ob man aus den (perfekten!) Hamming-Codes noch bessere Codes machen kann. Das Stichwort hierfür heißt Erweiterung. Wir verlängern jedes Codewort aus $\text{Ham}(r)$ um eine Stelle; auch diese Stelle wird mit 0 oder 1 besetzt und zwar so, dass die Gesamtzahl der Einsen in jedem Codewort gerade ist. Der so entstehende binäre Code heißt $\text{Ham}(r)^*$ und wird **erweiterter Hamming-Code** genannt.

10. Listen Sie die Codewörter von $\text{Ham}(3)^*$ auf.
 11. Zeigen Sie: $\text{Ham}(r)^*$ ist ein linearer Code mit Minimalgewicht 4.
 12. Konstruieren Sie aus einer Kontrollmatrix H für $\text{Ham}(r)$ eine Kontrollmatrix H^* für $\text{Ham}(r)^*$.

Sie sollten mit folgenden Begriffen umgehen können:

Geometrie, affiner Raum, $\mathbf{A}(V)$, affiner Unterraum, Gerade, Ebene, Hyperebene, lineares Gleichungssystem, homogen, inhomogen, Matrizenmultiplikation, erweiterte Matrix eines linearen Gleichungssystems, Rang, Spaltenrang, Zeilenrang, elementare Umformungen, Gaußscher Algorithmus, Code, Hamming-Abstand, Fehlerkorrigierender Code, linearer Code, Generatormatrix, Kontrollmatrix, Syndrom, Syndromdecodierung

Was sagen Sie dazu?

Gnutpuaheb. *Alle Punkte der Anschauungsebene liegen auf einer gemeinsamen Geraden.*

Sieweb. Es genügt zu zeigen, dass je endlich viele Punkte auf einer gemeinsamen Geraden liegen. Seien also P_1, \dots, P_n endlich viele Punkte. Wir zeigen durch Induktion nach n , dass diese auf einer gemeinsamen Geraden liegen.

Induktionsanfang: $n = 1$. Trivialerweise liegt ein einziger Punkt auf einer (gemeinsamen) Geraden.

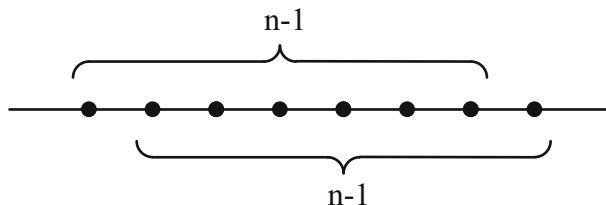


Abb. 4.4 Alle (?) Punkte auf einer gemeinsamen Geraden

Nun zum Induktionsschluss: Sei die Aussage richtig für $n-1$. Seien P_1, \dots, P_n Punkte der Anschauungsebene. Nach Induktion liegen sowohl P_1, \dots, P_{n-1} als auch P_2, \dots, P_n auf einer gemeinsamen Geraden. Also liegen auch $P_1, P_2, \dots, P_{n-1}, P_n$ auf einer gemeinsamen Geraden (siehe Abb. 4.4).



... es ist eine Textaufgabe: Ramses kauft 2 Eulen und 3 Jbisse. Er bezahlt mit 1 Sack Feuerbohnen. Wie viele Schalen Vogelfutter muss der Händler ihm noch dazugeben?

Bei jeder mathematischen Struktur ist es äußerst wichtig, die Struktur erhaltenden Abbildungen, die so genannten Homomorphismen, zu studieren. Dies hat folgende Gründe:

- Wir müssen feststellen können, ob zwei Strukturen, in unserem Fall also zwei Vektorräume „im wesentlichen gleich“ (das heißt „isomorph“) sind. Damit kann man auch feststellen, durch welche Daten ein Vektorraum bestimmt ist. Zum Beispiel kann man sich fragen, ob ein Vektorraum schon durch den zugrunde liegenden Körper und die Dimension „im wesentlichen“ eindeutig bestimmt ist. Wir werden diese Frage mit „ja“ beantworten.
- Innerhalb eines gegebenen Vektorraums wollen wir feststellen können, ob zwei Objekte desselben Typs durch einen Automorphismus ineinander überführbar sind. Es wird sich zeigen, dass je zwei Basen durch einen Vektorraumautomorphismus aufeinander abgebildet werden können. Dies impliziert dann, dass wir – wenn notwendig – ohne den Vektorraum zu ändern, eine bestimmte Basis o. B. d. A. auswählen können!
- Durch einen Homomorphismus wird ein gegebener Vektorraum V wieder auf einen Vektorraum abgebildet. Kann man eine Übersicht über alle so erhaltenen Vektorräume bekommen? Der Homomorphiesatz wird darauf eine Antwort geben.

5.1 Definitionen und grundlegende Eigenschaften

Nun zur Sache: Wir betrachten zwei Vektorräume V und W , die über demselben Körper K definiert sind. Eine Abbildung $f: V \rightarrow W$ wird **linear** genannt, falls für alle $v, v' \in V$ und alle $k \in K$ gilt:

$$\text{Additivität: } f(v + v') = f(v) + f(v'),$$

$$\text{Homogenität: } f(k \cdot v) = k \cdot f(v).$$

Man spricht auch von einer **linearen Abbildung** oder, vollkommen gleichbedeutend, von einem **Homomorphismus**.

Durch die Additivität und die Homogenität werden die Operationen von V (Addition und Skalarmultiplikation) auf W übertragen, genauer gesagt: auf den Teil von W , der von f erfasst wird; diesen Teil nennen wir das **Bild** von f :

$$\text{Bild}(f) := \{w \in W \mid \text{es gibt } v \in V \text{ mit } f(v) = w\}.$$

Als erste Übung im Umgang mit linearen Abbildungen zeigen wir, dass $\text{Bild}(f)$ *wieder ein Vektorraum ist*.

Wir zeigen genauer, dass $\text{Bild}(f)$ *ein Unterraum von W ist*. Und dies beweisen wir, indem wir das Unterraumkriterium anwenden.

Ist der Nullvektor (von W) in $\text{Bild}(f)$? Ja, denn der Nullvektor o_V von V wird auf den Nullvektor o_W von W abgebildet:

$$f(o_V) + f(o_V) = f(o_V + o_V) = f(o_V).$$

Wenn wir auf beiden Seiten den Vektor $f(o_V)$ ($\in W$) abziehen, ergibt sich $f(o_V) = o_W$. Daraus folgt $o_W \in \text{Bild}(f)$.

Ist das skalare Vielfache eines jeden Vektors $w \in \text{Bild}(f)$ in $\text{Bild}(f)$? Sei dazu v ein Urbild von w ; dann gilt für alle $k \in K$:

$$k \cdot w = k \cdot f(v) = f(k \cdot v) \in \text{Bild}(f).$$

Ähnlich ergibt sich, dass *mit je zwei Vektoren w, w' auch die Differenz $w - w'$ in $\text{Bild}(f)$ liegt*: Seien dazu v und v' Vektoren aus V mit $f(v) = w$ und $f(v') = w'$. Dann folgt

$$w - w' = f(v) - f(v') = f(v - v') \in \text{Bild}(f).$$

Mit Hilfe des Unterraumkriteriums ergibt sich also, dass $\text{Bild}(f)$ ein Unterraum von W ist. □

Nun haben wir schon mit dem Studium linearer Abbildungen begonnen, ohne uns durch Beispiele von der Existenz solcher Objekte überzeugt zu haben. Wir eilen, dieses Versäumnis auszubügeln. Zunächst zwei ganz einfache Beispiele:

- (0) Die **Nullabbildung** (das ist die Abbildung, die jeden Vektor von V auf den Nullvektor von W abbildet) ist eine lineare Abbildung.
 - (1) Die **identische Abbildung** id_V auf V (das ist die Abbildung, die jedes Element von V auf sich abbildet) ist eine lineare Abbildung von V in sich.
- Fast genauso einfach ist das folgende Beispiel:

- (2) Sei V ein Vektorraum über dem kommutativen Körper K , und sei $k_0 \in K$. Dann ist die Abbildung f , die jeden Vektor v auf $k_0 \cdot v$ abbildet, eine lineare Abbildung von V in sich. Es ist nämlich

$$f(v + v') = k_0 \cdot (v + v') = k_0 \cdot v + k_0 \cdot v' = f(v) + f(v')$$

und

$$f(kv) = k_0 \cdot (kv) = (k_0 k) \cdot v = (kk_0) \cdot v = k \cdot (k_0 v) = k \cdot f(v).$$

Für $k_0 = 1$ erhalten wir die identische Abbildung und für $k_0 = 0$ die Nullabbildung auf V ; insofern sind (0) und (1) Spezialfälle von (2).

Wegen dieser einfachen Beispiele würde es sich wirklich nicht lohnen, lineare Abbildungen zu studieren. Deshalb betrachten wir jetzt noch ein interessanteres Beispiel, das man auf vielfältige Weise verallgemeinern kann (siehe Übungsaufgabe 2).

- (3) Wir definieren eine lineare Abbildung f von \mathbf{R}^2 nach \mathbf{R}^2 durch

$$f : (a, b) \mapsto (3a + b, a - b).$$

Dass f linear ist, zeigt man so: Seien (a, b) und (a', b') Vektoren aus \mathbf{R}^2 , und sei $k \in \mathbf{R}$. Die Additivität folgt so

$$\begin{aligned} f((a, b) + (a', b')) &= f((a + a', b + b')) = (3(a + a') + b + b', a + a' - (b + b')) \\ &= (3a + b + 3a' + b', a - b + a' - b') \\ &= (3a + b, a - b) + (3a' + b', a' - b') = f(a, b) + f(a', b'). \end{aligned}$$

Die Homogenität ergibt sich ganz ähnlich:

$$\begin{aligned} f(k(a, b)) &= f(ka, kb) = (3ka + kb, ka - kb) = (k(3a + b), k(a - b)) \\ &= k \cdot (3a + b, a - b) = k \cdot f(a, b). \end{aligned}$$

Dies sind bei weitem nicht alle lineare Abbildungen, aber die Abbildung (3) ist schon ein typisches Beispiel einer linearen Abbildung. Wir werden später weitere Beispiele konstruieren. Ja, unser Ziel ist sogar, *alle* linearen Abbildungen genau zu beschreiben.

Offenbar kann man an den Beispielen schon eine ganz grobe Einteilung der linearen Abbildungen vornehmen. Insbesondere geben wir zwei wichtigen Teilklassen einen Namen.

Eine lineare Abbildung von V nach W heißt **Isomorphismus**, falls sie injektiv und surjektiv ist. Wenn es einen Isomorphismus eines Vektorraums V in einen Vektorraum W gibt, dann nennt man V und W **isomorph** und schreibt dafür $V \cong W$. Ein Isomorphismus von V nach W heißt **Automorphismus**, falls $W = V$ ist; ein Automorphismus ist also eine umkehrbare lineare Abbildung eines Vektorraums in sich. Manchmal wird eine lineare

Abbildung von V in sich auch **Endomorphismus** genannt. (In dieser Terminologie ist ein Automorphismus also ein umkehrbarer Endomorphismus.)

Die Menge aller Automorphismen des Vektorraums V bekommt auch einen Namen, nämlich $GL(V)$. (Dies ist die Abkürzung für „general linear group“ – in der Tat bilden die Elemente von $GL(V)$ eine „Gruppe“ – mehr dazu in Kap. 9.)

Warum heißt eine lineare Abbildung eigentlich „linear“? „Linear“ hat etwas mit „Linie“, also „Gerade“ zu tun. Tatsächlich gilt der folgende Satz.

Invarianz der Geraden unter einer linearen Abbildung

Jeder Automorphismus f von V bildet jede Gerade des affinen Raums $A(V)$ wieder auf eine Gerade ab.

Zum *Beweis* dieser Tatsache betrachten wir eine Gerade von $A(V)$. Was ist eine Gerade von $A(V)$? Das wissen wir spätestens seit Abschn. 4.2: Jede Gerade des affinen Raums $A(V)$ hat die Form $v + U$, wobei U ein 1-dimensionaler Unterraum von V ist, das heißt $U = \langle u \rangle$ mit $u \in V \setminus \{0\}$. Wir müssen zeigen, dass die Menge der Bilder der Elemente von $v + U$ wieder eine Gerade bilden. Wenn wir raten müssten, welche Gerade dies sein könnte, dann würde uns als sinnvoller Vorschlag wahrscheinlich

$$f(v) + f(U)$$

einfallen. (Dabei ist $f(U) := \{f(u) \mid u \in U\}$.)

Mal schauen, ob das stimmt: Sei $v + ku$ ein beliebiger Punkt der Geraden $v + U$. Dann ist

$$f(v + ku) = f(v) + f(ku) = f(v) + k \cdot f(u) \in f(v) + \langle f(u) \rangle.$$

Also liegt das Bild jedes Punktes der Geraden $v + U$ in der Menge $f(v) + \langle f(u) \rangle$. Ferner ist $f(u) \neq 0$, da f bijektiv ist; also ist $f(v) + \langle f(u) \rangle$ tatsächlich eine Gerade.

Sei umgekehrt $f(v) + k'f(u)$ ein beliebiger Punkt der hypothetischen Bildgeraden. Dann ist $v + k'u$ ein Urbild dieses Punktes.

Also werden die Punkte von $v + \langle u \rangle$ genau auf die Punkte von $f(v) + \langle f(u) \rangle$ abgebildet. Das heißt: Geraden („Linien“) gehen in Geraden über. \square

Unser erstes großes Ziel ist es, alle linearen Abbildungen von V nach W zu beschreiben. Dazu ist folgende Aussage von entscheidender Bedeutung:

Beschreibung linearer Abbildungen durch die Bilder einer Basis

Sei $\{v_1, \dots, v_n\}$ eine Basis von V . Dann gibt es zu beliebigen Vektoren $w_1, \dots, w_n \in W$ genau eine lineare Abbildung $f: V \rightarrow W$ mit $f(v_i) = w_i$ ($i = 1, \dots, n$).

Mit anderen Worten: Man kann sich für die Bilder einer Basis von V beliebige Vektoren wünschen; wir nennen die w_i daher auch „Wunschvektoren“. Durch die Vorgabe der Wunschvektoren ist eine lineare Abbildung dann eindeutig bestimmt.

Den Beweis dieser Behauptung zerlegen wir in zwei Teilbehauptungen. Wir werden jeweils eine etwas allgemeinere Aussage beweisen. Zuerst zeigen wir die Existenz von f .

- (a) Ist $\{v_1, \dots, v_r\}$ eine linear unabhängige Menge von Vektoren, so gibt es für beliebige Wunschvektoren $w_1, \dots, w_r \in W$ mindestens eine lineare Abbildung $f: V \rightarrow W$ mit $f(v_i) = w_i$ für alle $i \in \{1, \dots, r\}$.

Dies folgt so: Wir ergänzen $\{v_1, \dots, v_r\}$ zu einer Basis $\{v_1, \dots, v_n\}$ von V und wählen $w_{r+1}, \dots, w_n \in W$ beliebig. (Wenn wir wollten, könnten wir zum Beispiel $w_{r+1} = \dots = w_n = 0$ wählen.)

Um $f(v)$ für einen beliebigen Vektor $v \in V$ zu definieren, stellen wir v als Linearkombination $v = k_1 v_1 + k_2 v_2 + \dots + k_n v_n$ dar. Wir definieren dann

$$f(v) := k_1 w_1 + k_2 w_2 + \dots + k_n w_n .$$

Das heißt, wir „ersetzen“ jedes v_i durch w_i .

Behauptung: f ist eine lineare Abbildung.

Additivität: Sei $u = h_1 v_1 + h_2 v_2 + \dots + h_n v_n$. Dann ist

$$\begin{aligned} f(v + u) &= f((k_1 + h_1)v_1 + (k_2 + h_2)v_2 + \dots + (k_n + h_n)v_n) \\ &= (k_1 + h_1)w_1 + (k_2 + h_2)w_2 + \dots + (k_n + h_n)w_n \\ &= (k_1 w_1 + k_2 w_2 + \dots + k_n w_n) + (h_1 w_1 + h_2 w_2 + \dots + h_n w_n) = f(v) + f(u) . \end{aligned}$$

Homogenität: Sei $k \in K$ beliebig. Dann ist

$$\begin{aligned} f(kv) &= f(kk_1 v_1 + kk_2 v_2 + \dots + kk_n v_n) = kk_1 w_1 + kk_2 w_2 + \dots + kk_n w_n \\ &= k(k_1 w_1 + k_2 w_2 + \dots + k_n w_n) = k \cdot f(v) . \end{aligned}$$

Nun zeigen wir die Eindeutigkeit von f :

- (b) Ist $\{v_1, \dots, v_s\}$ ein Erzeugendensystem von V , so gibt es für beliebige Wunschvektoren $w_1, \dots, w_s \in W$ höchstens eine lineare Abbildung f von V nach W mit $f(v_i) = w_i$ für $i = 1, \dots, s$.

Dies sehen wir so: Seien f und g lineare Abbildungen von V in W mit $f(v_i) = w_i = g(v_i)$ ($i = 1, \dots, s$). Wir müssen zeigen, dass dann f und g nicht nur auf den verschwindend wenigen Vektoren v_1, \dots, v_s dasselbe bewirken, sondern auf jedem Vektor von V .

Wir betrachten dazu einen beliebigen Vektor $v \in V$. Es ist zu zeigen, dass $f(v) = g(v)$ gilt.

Da $\{v_1, \dots, v_s\}$ ein Erzeugendensystem ist, kann man v als Linearkombination der v_i darstellen; sei

$$v = k_1 v_1 + k_2 v_2 + \dots + k_s v_s .$$

Daraus ergibt sich

$$\begin{aligned} f(v) &= f(k_1 v_1 + k_2 v_2 + \dots + k_s v_s) = k_1 f(v_1) + k_2 f(v_2) + \dots + k_s f(v_s) \\ &= k_1 w_1 + k_2 w_2 + \dots + k_s w_s = k_1 g(v_1) + k_2 g(v_2) + \dots + k_s g(v_s) \\ &= g(k_1 v_1 + k_2 v_2 + \dots + k_s v_s) = g(v) . \end{aligned}$$

Dies zeigt $f = g$. □

Die Bedeutung dieser Aussage liegt darin, dass man eine lineare Abbildung durch ganz wenige Daten beschreiben kann, nämlich durch eine Basis von V und die Bilder der Vektoren dieser Basis.

Der obige Satz liefert zunächst „nur“ eine äußerst effiziente Beschreibung einer linearen Abbildung. Kann man an dieser Beschreibung auch Eigenschaften der linearen Abbildung ablesen? (Es könnte ja sein, dass man eine zwar sehr platz sparende Beschreibung gefunden hat, diese aber zu nichts anderem nütze ist.) Die Antwort ist „ja“:

Charakterisierung einer linearen Abbildung durch die Bilder der Vektoren einer Basis

Sei $f: V \rightarrow W$ eine lineare Abbildung; sei $\{v_1, v_2, \dots, v_n\}$ eine Basis von V , und sei $w_i := f(v_i)$ für $i = 1, 2, \dots, n$. Dann gilt:

- (a) Die Vektoren w_1, w_2, \dots, w_n sind genau dann linear unabhängig, wenn f injektiv ist.
- (b) Genau dann ist $\{w_1, w_2, \dots, w_n\}$ ein Erzeugendensystem von W , wenn f surjektiv ist.
- (c) Genau dann ist $\{w_1, w_2, \dots, w_n\}$ eine Basis von W , wenn f bijektiv, also ein Isomorphismus, ist.

Zum Beweis müssen wir nur (a) und (b) zeigen.

- (a) Zunächst setzen wir voraus, dass die Vektoren w_1, \dots, w_n linear unabhängig sind. Seien v und v' Vektoren aus V mit $f(v) = f(v')$. Wir zeigen $v = v'$.

Dazu stellen wir v und v' als Linearkombination der Basis $\{v_1, \dots, v_n\}$ dar:

$$v = k_1 v_1 + k_2 v_2 + \dots + k_n v_n \text{ und } v' = k'_1 v_1 + k'_2 v_2 + \dots + k'_n v_n .$$

Es folgt

$$f(v) = k_1 w_1 + k_2 w_2 + \dots + k_n w_n \text{ und } f(v') = k'_1 w_1 + k'_2 w_2 + \dots + k'_n w_n .$$

Wegen $f(v) = f(v')$ ist also

$$k_1 w_1 + k_2 w_2 + \dots + k_n w_n = k'_1 w_1 + k'_2 w_2 + \dots + k'_n w_n .$$

Da w_1, w_2, \dots, w_n linear unabhängig sind, folgt durch Koeffizientenvergleich

$$k_1 = k'_1, k_2 = k'_2, \dots, k_n = k'_n ,$$

also $v = v'$.

Sie sind eingeladen, die Umkehrung als Übungsaufgabe 4 zu zeigen.

- (b) Sei wiederum zunächst $\{w_1, \dots, w_n\}$ ein Erzeugendensystem von W . Wir müssen zeigen, dass es für ein beliebiges $w \in W$ ein $v \in V$ gibt mit $f(v) = w$.

Da w_1, \dots, w_n ein Erzeugendensystem von W ist, können wir w (nicht notwendig eindeutig) schreiben als

$$w = k_1 w_1 + k_2 w_2 + \dots + k_n w_n .$$

Dann ist offenbar

$$v := k_1 v_1 + k_2 v_2 + \dots + k_n v_n$$

ein Urbild von w .

Wieder sind Sie eingeladen, die Umkehrung als Übung zu zeigen (Übungsaufgabe 5). \square

Wir ziehen aus obiger Charakterisierung interessante und wichtige Folgerungen.

Folgerung 1

„Wenn einmal unabhängig, dann immer unabhängig; wenn einmal Erzeugendensystem, dann immer Erzeugendensystem.“ Das heißt: Wenn die Bilder einer Basis von V unter f linear unabhängig sind (bzw. ein Erzeugendensystem bilden), dann sind die Bilder jeder Basis von V unter f linear unabhängig (bzw. bilden ein Erzeugendensystem).

Der *Beweis* ist, wie es sich für eine Folgerung gehört, einfach: Sei B eine Basis, deren Bilder linear unabhängig sind. Dann ist f nach dem vorigen Satz injektiv. Angenommen, es gäbe eine Basis B^* , deren Bilder nicht linear unabhängig sind; dann wäre f nicht injektiv. Zusammen ergibt sich ein Widerspruch.

Die anderen Behauptungen folgen entsprechend. \square

Folgerung 2

Seien V und W Vektorräume über dem Körper K , welche die gleiche Dimension n haben. Dann gibt es zu jeder Basis $\{v_1, \dots, v_n\}$ von V und jeder Basis $\{w_1, \dots, w_n\}$ von W (genau) eine lineare Abbildung $f: V \rightarrow W$ mit $f(v_i) = w_i$ ($i = 1, \dots, n$). Diese lineare Abbildung ist ein Isomorphismus.

Beweis Dass es – bei fester Reihenfolge der Basisvektoren – genau eine lineare Abbildung f mit den gewünschten Eigenschaften gibt, folgt aus dem Satz über die Beschreibung linearer Abbildungen durch Wunschvektoren. Dass f bijektiv ist, ergibt sich aus Teil (c) des vorigen Satzes. \square

Folgerung 3

Je zwei Basen eines Vektorraums V können durch genau eine lineare Abbildung ineinander überführt werden; diese lineare Abbildung ist ein Automorphismus. Man nennt solche linearen Abbildungen auch **Basistransformationen**. \square

Folgerung 4 (Fundamentalsatz für endlichdimensionale Vektorräume)

Je zwei K -Vektorräume derselben Dimension sind isomorph.

Zum *Beweis* wähle man in zwei n -dimensionalen K -Vektorräumen jeweils eine Basis. Nach Folgerung 2 ist die Abbildung, die die eine Basis in die andere überführt, ein Isomorphismus. \square

Mit Folgerung 4 ist eine der eingangs gestellten Fragen beantwortet: Ja, jeder Vektorraum ist durch seinen Körper und seine Dimension eindeutig bestimmt!

5.2 Darstellung von linearen Abbildungen durch Matrizen

Wenn man mit linearen Abbildungen konkret rechnen will (oder muss), dann bietet es sich an, ihre Darstellung durch Matrizen zu benutzen. Das werden wir in diesem Abschnitt erklären.

Sei stets $f: V \rightarrow W$ eine lineare Abbildung eines n -dimensionalen K -Vektorraums V in einen m -dimensionalen K -Vektorraum W .

Es gibt nicht *die* Matrix, die eine lineare Abbildung f beschreibt, sondern viele, und zwar zu jeder Basis von V und jeder Basis von W genau eine.

Wir wählen also eine Basis $B = \{v_1, \dots, v_n\}$ von V und eine Basis $C = \{w_1, \dots, w_m\}$ von W und halten diese Bezeichnungen fest. Der entscheidende Punkt ist, dass wir *die Bilder der Elemente v_j aus B als Linearkombinationen der Basis C ausdrücken*:

$$f(v_j) = a_{1j}w_1 + a_{2j}w_2 + \dots + a_{mj}w_m \quad (j = 1, \dots, n)$$

mit $a_{ij} \in K$.

Dadurch werden $m \cdot n$ Körperelemente a_{ij} definiert; diese fassen wir in der Matrix A zusammen:

$$A = {}_B M_C(f) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

Jede so gewonnene Matrix ${}_B M_C(f)$ wird eine **Darstellungsmatrix** der linearen Abbildung f genannt. (Dabei deutet das „ M “ in „ ${}_B M_C(f)$ “ nur an, dass eine Matrix gebildet wird.)

Wie kann man sich eine solche Darstellungsmatrix A vorstellen? Die Merkregel lautet: *Die Koeffizienten des Bildes von v_j sind die Einträge in der j -ten Spalte von A .*

Beispiel Sei f die Abbildung von \mathbf{R}^2 in sich, die durch folgende Vorschrift definiert ist:

$$f : (x, y) \mapsto (3x - 2y, x + y).$$

Wir unterscheiden verschiedene Wahlen der Basen B und C .

- Sei zunächst $B = C = \{(1, 0), (0, 1)\}$. Dann ist

$${}_B M_C(f) = {}_B M_B(f) = \begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix}.$$

- Sei jetzt $B = \{(1, 0), (0, 1)\}$ und $C = \{(3, 1), (-2, 1)\}$. Dann ist

$${}_B M_C(f) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

- Ist schließlich $B = \{(5, 8), (-1, 1)\}$ und $C = \{(0, 1), (1, 1)\}$, so gilt

$${}_B M_C(f) = \begin{pmatrix} 14 & 5 \\ -1 & -5 \end{pmatrix},$$

da

$$f(5, 8) = (-1, 13) = 14 \cdot (0, 1) + (-1) \cdot (1, 1)$$

und

$$f(-1, 1) = (-5, 0) = 5 \cdot (0, 1) + (-5) \cdot (1, 1)$$

ist.

Bemerkung Wie kommt man auf die Koeffizienten? Ganz einfach; dies kann man stur ausrechnen. Wir machen uns das am Beispiel von $f(5, 8)$ klar.

Durch einfaches Einsetzen sieht man $f(5, 8) = (-1, 13)$. Wir müssen diesen Vektor als Linearkombination von $(0, 1)$ und $(1, 1)$ darstellen. Dazu machen wir folgenden „Ansatz“:

$$(-1, 13) = x \cdot (0, 1) + y \cdot (1, 1).$$

Dies führt dann auf die Gleichungen $-1 = y$ und $13 = x + y$, die elementar aufzulösen sind.

Wir beobachten, dass umgekehrt *jede Matrix aus $K^m \times^n$ Darstellungsmatrix einer linearen Abbildung ist*:

Wir wählen dazu wieder Basen $B = \{v_1, \dots, v_n\}$ von V und $C = \{w_1, \dots, w_m\}$ von W . Wenn $A = (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$ gegeben ist, so definieren wir die Bilder der Basisvektoren v_j wie folgt:

$$f(v_j) := a_{1j}w_1 + a_{2j}w_2 + \dots + a_{mj}w_m \quad (j = 1, \dots, n).$$

Nach den Ergebnissen des letzten Abschnitts (Beschreibung einer linearen Abbildung durch die Bilder einer Basis) ist dadurch die lineare Abbildung f eindeutig definiert; es ist klar, dass f bezüglich B und C die Darstellungsmatrix $A = (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$ hat.

Wir fassen diese Überlegungen zusammen:

Darstellungssatz

Sei $f: V \rightarrow W$ eine lineare Abbildung. Sei $B = \{v_1, \dots, v_n\}$ eine Basis von V und $C = \{w_1, \dots, w_m\}$ eine Basis von W . Dann wird durch

$$f(v_j) = a_{1j}w_1 + a_{2j}w_2 + \dots + a_{mj}w_m \quad (j = 1, \dots, n)$$

eine Matrix $A = (a_{ij}) \in K^m \times^n$ definiert. Umgekehrt gehört zu jeder Matrix $A = (a_{ij})$ aus $K^m \times^n$ (bei festen Basen B und C) genau eine lineare Abbildung f ; diese wird durch

$$f(v_j) := a_{1j}w_1 + a_{2j}w_2 + \dots + a_{mj}w_m \quad (j = 1, \dots, n)$$

definiert.

□

Ferner gilt:

Darstellungsmatrix der Identität

Die Identität auf V hat bezüglich jeder Wahl einer Basis $B = \{v_1, \dots, v_n\}$ von V die Einheitsmatrix als Darstellungsmatrix:

$${}_B M_B(\text{id}) = E_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}.$$

Dies folgt direkt aus der Tatsache, dass für alle j die Beziehung $\text{id}(v_j) = v_j = 1 \cdot v_j$ gilt. \square

Nun diskutieren wir ausführlich einen konkreten Fall, nämlich den Fall von linearen Abbildungen des Vektorraums K^n in den Vektorraum K^m .

Man kann jede $m \times n$ -Matrix als lineare Abbildung von K^n in K^m auffassen. Dazu stellen wir die Vektoren von K^n bzw. K^m als Spaltenvektoren dar:

$$K^n = \left\{ \begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix} \mid k_1, \dots, k_n \in K \right\}.$$

Dann induziert (das heißt „ist“) jede $m \times n$ -Matrix $A = (a_{ij})$ über folgende Vorschrift eine lineare Abbildung von K^n in K^m : Der Vektor

$$v = \begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix}$$

wird abgebildet auf den Vektor

$$A \cdot v := \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix} = \begin{pmatrix} a_{11} \cdot k_1 + a_{12} \cdot k_2 + \dots + a_{1n} \cdot k_n \\ a_{21} \cdot k_1 + a_{22} \cdot k_2 + \dots + a_{2n} \cdot k_n \\ \vdots \\ a_{m1} \cdot k_1 + a_{m2} \cdot k_2 + \dots + a_{mn} \cdot k_n \end{pmatrix}.$$

Wir machen uns dies an einem *Beispiel* klar. Sei $n = m = 3$, und sei

$$A = \begin{pmatrix} 3 & -2 & 7 \\ 1 & 2 & 0 \\ -1 & 4 & 8 \end{pmatrix}$$

eine reelle 3×3 -Matrix. Dann wird durch A eine lineare Abbildung von \mathbf{R}^3 in sich definiert, die wie folgt operiert

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto A \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3x - 2y + 7z \\ x + 2y \\ -x + 4y + 8z \end{pmatrix}.$$

Wir haben uns überlegt, dass jede Matrix A als lineare Abbildung interpretiert werden kann; zuvor hatten wir auch gezeigt, dass jede lineare Abbildung eine Darstellungsmatrix hat. Wie sieht eine Darstellungsmatrix zu A aus? Gibt es vielleicht eine Darstellungsmatrix, die „etwas mit A zu tun“ hat? Oder sogar eine, die gleich A ist? Ist also A Darstellungsmatrix von sich selbst? Die Antwort lautet – wie könnte es anders sein – ja:

Darstellungsmatrix einer als lineare Abbildung interpretierten Matrix

Sei A eine Matrix aus $K^{m \times n}$, die wir als lineare Abbildung K^n nach K^m auffassen. Ist $B = \{e_1, e_2, \dots, e_n\}$ bzw. $C = \{e_1, e_2, \dots, e_m\}$ die Basis aus Einheitsvektoren des K^n bzw. des K^m , so gilt ${}_B M_C(f) = A$.

Das bedeutet, dass A bezüglich der Basen aus Einheitsvektoren sich selbst als Darstellungsmatrix hat.

Beweis Sei $A = (a_{ij})$. Dann wird $e_1 \in K^n$ auf den Vektor abgebildet, der in der i -ten Komponente das Element a_{i1} . $1 = a_{i1}$ stehen hat. Mit anderen Worten: e_1 wird auf die erste Spalte von A abgebildet.

Natürlich ist am ersten Einheitsvektor nichts Besonderes: auch im Allgemeinen gilt, dass e_j auf die j -te Spalte von A abgebildet wird.

Daraus folgt, dass die Darstellungsmatrix der linearen Abbildung A bezüglich der Basen B und C gleich A ist. \square

Neben der effizienten Darstellung einer linearen Abbildung hat die Matrizendarstellung einen weiteren Vorteil, nämlich dass man scheinbar unzusammenhängende Begriffe in Verbindung bringen kann. Als erstes Beispiel nennen wir folgenden Satz, mit dem man die Dimension von $\text{Bild}(f)$ ausrechnen kann.

Rang einer linearen Abbildung

Wenn A eine Darstellungsmatrix der linearen Abbildung $f: V \rightarrow W$ ist, so gilt

$$\dim(\text{Bild}(f)) = \text{Rang}(A).$$

Insbesondere ist $\text{Rang}(A)$ unabhängig von der Wahl der Darstellungsmatrix A von f . Man nennt diese Zahl daher auch den **Rang** von f .

Dies ist nicht schwer *einzusehen*: Sei $B = \{v_1, v_2, \dots, v_n\}$ die Basis von V , die zur Bildung der Darstellungsmatrix A verwendet wurde, also etwa $A = {}_B M_C(f)$.

Wir überlegen uns zunächst, dass *der von den Spalten s_1, s_2, \dots, s_n von A aufgespannte Unterraum U von K^m isomorph ist zu dem Unterraum W' von W , der von den Bildern $f(v_1), f(v_2), \dots, f(v_n)$ der Basisvektoren erzeugt wird*. (Ein Isomorphismus wird durch die Abbildung

$$s_i \mapsto f(v_i)$$

hergestellt; vergleiche Übungsaufgabe 12.)

Eine andere Weise, dies auszudrücken, ist die folgende: Die Koordinatenvektoren s_1, s_2, \dots, s_n spannen einen zu $\langle f(v_1), f(v_2), \dots, f(v_n) \rangle = \text{Bild}(f)$ isomorphen Vektorraum auf. Nach Definition des Rangs einer Matrix ist also

$$\text{Rang}(A) = \dim(U) = \dim(\text{Bild}(f)) .$$

□

Als Korollar aus diesem Satz ergibt sich das folgende Lemma: *Genau dann ist eine Menge $s_{i_1}, s_{i_2}, \dots, s_{i_k}$ von Spalten von A linear unabhängig, wenn die Vektoren $f(v_{i_1}), f(v_{i_2}), \dots, f(v_{i_k})$ linear unabhängig sind*. □

Wir betrachten jetzt noch den wichtigen Spezialfall, dass die Vektorräume V und W dieselbe Dimension n haben. Dann gilt:

Charakterisierung von Isomorphismen

Eine lineare Abbildung f eines n -dimensionalen K -Vektorraums V in einen n -dimensionalen K -Vektorraum W ist genau dann ein Isomorphismus, wenn für beliebige Basen B von V und C von W die Matrix $A = {}_B M_C(f)$ den Rang n hat.

Beweis Sei zunächst $\text{Rang}(A) < n$. Dann sind die Spalten von A linear abhängig. Nach dem vorigen Satz sind also die Bilder der Basis B linear abhängig; somit ist f nicht injektiv, insbesondere also kein Isomorphismus.

Sei andererseits $\text{Rang}(A) = n$. Dann sind nach dem vorigen Satz die Bilder der Basis B linear unabhängig, erzeugen also einen Unterraum der Dimension n von W . Da $\dim(W) = n$ ist, sind die Bilder von B eine Basis von W . Nach der Charakterisierung einer linearen Abbildung durch die Bilder einer Basis ist f bijektiv. □

Im letzten Teil dieses Abschnitts beschäftigen wir uns mit folgendem wichtigen Problem: Wir können (manchmal) zwei lineare Abbildungen miteinander verknüpfen, nämlich durch Hintereinanderausführung, und wir können (manchmal) zwei Matrizen miteinander verknüpfen, nämlich durch Multiplikation. Passen diese beiden Prozesse zusammen? Antwort: Ja – unter gewissen, selbstverständlichen Vorsichtsmaßnahmen.

Darstellungsmatrix des Produkts von linearen Abbildungen

Seien V_1, V_2, V_3 Vektorräume über dem Körper K ; seien $g: V_1 \rightarrow V_2$ und $f: V_2 \rightarrow V_3$ lineare Abbildungen. Sei B_i eine Basis von V_i ($i = 1, 2, 3$), und sei $A = {}_{B_1}M_{B_2}(g)$, $B = {}_{B_2}M_{B_3}(f)$. Dann ist

$${}_{B_1}M_{B_3}(f \circ g) = BA.$$

Kurz: Hintereinanderausführung von linearen Abbildungen entspricht der Multiplikation von Matrizen.

Beweis Sei $A = (a_{ji})$ die Darstellungsmatrix von g , und sei $B = (b_{kj})$ die Darstellungsmatrix von f .

Wir definieren die Vektoren u_i, v_j und w_k durch $B_1 = \{u_i \mid i = 1, \dots, \dim(V_1)\}$, $B_2 = \{v_j \mid j = 1, \dots, \dim(V_2)\}$ und $B_3 = \{w_k \mid k = 1, \dots, \dim(V_3)\}$.

Um die Darstellungsmatrix $C = (c_{ki})$ von fg zu erhalten, müssen wir $fg(u_i)$ ausrechnen und diesen Vektor als Linearkombination der Vektoren aus B_3 darstellen: Da

$$g(u_i) = \sum_{j=1}^{\dim(V_2)} a_{ji} v_j$$

ist, folgt

$$\begin{aligned} fg(u_i) &= f\left(\sum_{j=1}^{\dim(V_2)} a_{ji} v_j\right) = \sum_{j=1}^{\dim(V_2)} a_{ji} f(v_j) = \sum_{j=1}^{\dim(V_2)} a_{ji} \left(\sum_{k=1}^{\dim(V_3)} b_{kj} w_k\right) \\ &= \sum_{j=1}^{\dim(V_2)} \sum_{k=1}^{\dim(V_3)} a_{ji} b_{kj} w_k = \sum_{k=1}^{\dim(V_3)} \left(\sum_{j=1}^{\dim(V_2)} b_{kj} a_{ji}\right) w_k = \sum_{k=1}^{\dim(V_3)} c_{ki} w_k. \end{aligned}$$

Dies ist bereits die Behauptung. \square

Eine $n \times n$ -Matrix M mit Elementen aus dem Körper K heißt **invertierbar** (oder **umkehrbar**, manchmal auch **regulär**), falls es eine Matrix M' gibt mit $MM' = E_n = M'M$. Wir schreiben dann M^{-1} für die Inverse M' von M . In der folgenden Aussage beschreiben wir die wichtige Klasse der invertierbaren Matrizen.

Korollar 1 (Charakterisierung invertierbarer Matrizen)

Eine $n \times n$ -Matrix ist genau dann invertierbar, wenn ihr Rang gleich n ist.

Beweis Sei A eine $n \times n$ -Matrix über K . Wir betrachten einen n -dimensionalen Vektorraum V über K und eine Basis B von V . Die Matrix A stellt bezüglich der Basis B eine lineare Abbildung f dar; das heißt $A = {}_B M_B(f)$.

Nun zum eigentlichen Beweis. Zunächst setzen wir voraus, dass A invertierbar ist. Dann existiert die Matrix A^{-1} ; diese stellt ebenfalls bezüglich der Basis B eine lineare Abbildung f' dar (das heißt $A^{-1} = {}_B M_B(f')$). Nach dem obigen Satz ist dann AA^{-1} die Darstellungsmatrix (bezüglich der Basis B) der Abbildung $f \cdot f'$; da AA^{-1} die Einheitsmatrix ist, muss $f \cdot f'$ die Identität sein. Ebenso zeigt man, dass $f' \cdot f$ die Identität ist. Daraus folgt, dass $f' = f^{-1}$ ist. Also ist f invertierbar, das heißt ein Isomorphismus. Daraus folgt aber, wie wir bereits wissen, dass $\text{Rang}(A) = n$ ist. (Vergleichen Sie dazu den Satz über die Invertierbarkeit bijektiver Abbildungen aus Abschn. 1.3)

Sei umgekehrt $\text{Rang}(A) = n$. Dann ist f ein Isomorphismus, und es existiert also der zu f inverse Automorphismus f^{-1} . Sei A' die Darstellungsmatrix von f^{-1} bezüglich B . Dann folgt, dass sowohl AA' als auch $A'A$ Darstellungsmatrizen der Identität sind. Also ist sowohl AA' als auch $A'A$ eine Einheitsmatrix; d. h. A ist invertierbar. \square

Korollar 2 (Transformation der Darstellungsmatrix bei Basiswechsel)

Sei $f: V \rightarrow W$ eine lineare Abbildung. Seien B und B' Basen von V , C und C' Basen von W . Seien $A = {}_B M_C(f)$ und $A' = {}_{B'} M_{C'}(f)$ die entsprechenden Darstellungsmatrizen. Dann gibt es eine invertierbare Matrix T und eine invertierbare Matrix S mit

$$A' = TAS.$$

Beweis Sei $S = {}_{B'} M_B(\text{id}_V)$ die Darstellungsmatrix der Identität auf V bezüglich der Basen B' und B , und sei $T = {}_C M_{C'}(\text{id}_W)$ die Darstellungsmatrix der Identität auf W bezüglich der Basen C und C' . Dann ist nach dem eben bewiesenen Satz TAS die Darstellungsmatrix der linearen Abbildung $\text{id}_W \circ f \circ \text{id}_V = f$ bezüglich der Basen B' und C' . Das heißt

$$TAS = A'.$$

Korollar 3 (Ähnlichkeit von Darstellungsmatrizen)

Sei f eine lineare Abbildung eines n -dimensionalen K -Vektorraums V in sich, und seien B, B' Basen von V . Seien $A = {}_B M_B(f)$ und $A' = {}_{B'} M_{B'}(f)$. Dann gibt es eine

invertierbare Matrix S mit

$$A' = S^{-1}AS.$$

Beweis Sei $S = {}_{B'}M_B(\text{id})$ die Darstellungsmatrix der Identität bezüglich der Basen B' und B , und sei $T = {}_BM_{B'}(\text{id})$ die Darstellungsmatrix der Identität bezüglich der Basen B und B' . Dann ist $TAS = A'$ nach Korollar 2. Ferner gilt

$$ST = {}_{B'}M_B(\text{id}) \cdot {}_BM_{B'}(\text{id}) = {}_BM_B(\text{id}) = E.$$

Entsprechend zeigt man $TS = E$. Also ist $T = S^{-1}$. □

Man nennt zwei $n \times n$ -Matrizen M und M' **ähnlich**, falls es eine invertierbare $n \times n$ -Matrix S gibt mit $M' = S^{-1}MS$.

Man kann leicht nachweisen (siehe Übungsaufgabe 16), dass die Ähnlichkeit von Matrizen eine Äquivalenzrelation ist.

Mit dieser Begriffsbildung kann man Korollar 3 also auch so aussprechen:

Ähnlichkeit von Darstellungsmatrizen

Je zwei Darstellungsmatrizen derselben linearen Abbildung eines Vektorraums in sich sind ähnlich. □

5.3 Der Homomorphiesatz

Wir haben bereits den Unterraum $\text{Bild}(f)$ von W , der zu einer linearen Abbildung f von V nach W gehört, eingehend studiert. Eine viel wichtigere Rolle spielt allerdings – entgegen dem ersten Augenschein – ein anderer zu f gehöriger Unterraum, nämlich der so genannte „Kern“ von f .

Sei stets $f: V \rightarrow W$ eine lineare Abbildung des K -Vektorraums V in den K -Vektorraum W . Die Menge

$$\text{Kern}(f) := \{v \in V \mid f(v) = 0\}$$

heißt der **Kern** von f .

Beispiel: Ist f die lineare Abbildung von \mathbf{R}^2 in sich, die durch $f(x, y) := (x, 0)$ definiert ist, so besteht $\text{Kern}(f)$ aus allen Vektoren der Form $(0, y)$.

1. Beobachtung

$\text{Kern}(f)$ ist ein Unterraum von V .

Beweis Wir wenden das Unterraumkriterium an.

- Wegen $f(o) = o$ ist $o \in \text{Kern}(f)$, also ist $\text{Kern}(f) \neq \emptyset$.
- Seien $v \in \text{Kern}(f)$ und $k \in K$. Dann gilt

$$f(k \cdot v) = k \cdot f(v) = k \cdot o = o;$$

somit ist mit $v \in \text{Kern}(f)$ auch $k \cdot v \in \text{Kern}(f)$.

- Seien $v, v' \in \text{Kern}(f)$. Dann ist

$$f(v - v') = f(v + (-1)v') = f(v) + f((-1) \cdot v') = f(v) + (-1) \cdot f(v') = o + o = o;$$

also ist $v - v' \in \text{Kern}(f)$. □

Wann ist der Kern am größten? Größer als V kann $\text{Kern}(f)$ nicht werden; $\text{Kern}(f) = V$ gilt offenbar genau dann, wenn jedes Element von V auf o abgebildet wird, wenn also f die Nullabbildung ist – ein ganz besonders trivialer Fall. Wann ist der Kern am kleinsten? Kleiner als $\{o\}$ kann er nicht werden. Ob und wann diese Grenze erreicht wird, sagt der folgende Satz.

2. Beobachtung

Genau dann ist $\text{Kern}(f) = \{o\}$, wenn f injektiv ist.

Beweis Sei zunächst $\text{Kern}(f) = \{o\}$, und seien $v, v' \in V$ mit $f(v) = f(v')$. Dann folgt

$$o = f(v) - f(v') = f(v - v'),$$

also $v - v' \in \text{Kern}(f)$. Da $\text{Kern}(f)$ nur aus dem Nullvektor besteht, folgt $v - v' = o$, also ist $v = v'$. Also ist f injektiv.

Umgekehrt setzen wir voraus, dass f injektiv ist. Da schon o auf o abgebildet wird, darf wegen der Injektivität von f kein von o verschiedener Vektor von V auf o abgebildet werden. Also besteht $\text{Kern}(f)$ nur aus dem Nullvektor. □

Nun nähern wir uns dem Homomorphiesatz. Da $\text{Kern}(f)$ ein Unterraum von V ist, können wir (das kann uns keiner verbieten) den Faktorraum $V/\text{Kern}(f)$ bilden. Aber jeder darf

uns *fragen*, was das denn soll! Die überraschende Antwort gibt der Homomorphiesatz: Auf diese Weise entsteht ein Vektorraum, und zwar einer, den wir schon kennen, nämlich $\text{Bild}(f)$!!!

Homomorphiesatz

Sei $f: V \rightarrow W$ eine lineare Abbildung. Dann gilt

$$V/\text{Kern}(f) \cong \text{Bild}(f) .$$

Genauer gesagt: Die Abbildung $\varphi: V/\text{Kern}(f) \rightarrow \text{Bild}(f)$, die durch

$$\varphi(v + \text{Kern}(f)) := f(v)$$

definiert ist, ist ein Isomorphismus von $V/\text{Kern}(f)$ nach $\text{Bild}(f)$.

Bevor wir den Homomorphiesatz beweisen, schauen wir uns seine Aussage nochmals andersherum an:

$$\text{Bild}(f) \cong V/\text{Kern}(f) .$$

Von diesem Standpunkt aus erkennt man folgendes: Der Vektorraum $\text{Bild}(f)$, also ein Raum, der meilenweit von uns entfernt liegen kann, ist strukturgleich zu $V/\text{Kern}(f)$, also einem Raum, der „innerhalb von V “ zu finden ist. Das bedeutet: Um alle möglichen Bilder einer linearen Abbildung von V nach irgendwohin zu finden, muss man nur in V nachschauen. (Ob Sie das als Hinweis auf die Gültigkeit des biedermeierlichen Lebensprinzips „Wozu in die Ferne schweifen? Sieh, das Gute liegt so nah!“ gelten lassen, müssen Sie selbst entscheiden.)

Genug der Philosophie, jetzt *beweisen wir den Homomorphiesatz*. Wir zeigen, dass φ eine *bijektive, lineare Abbildung* ist, und zwar zeigen wir diese drei Eigenschaften von „rechts nach links“.

- *Ist φ überhaupt eine Abbildung?* Mit anderen Worten: Ist φ wohldefiniert? Dies müssen wir nachweisen, da φ mit Hilfe von Repräsentanten einer Nebenklasse definiert ist. Seien also v und v' Repräsentanten derselben Nebenklasse nach $\text{Kern}(f)$. Das heißt

$$v + \text{Kern}(f) = v' + \text{Kern}(f) .$$

Wir müssen zeigen, dass daraus $\varphi(v + \text{Kern}(f)) = \varphi(v' + \text{Kern}(f))$, das heißt $f(v) = f(v')$ folgt. Wegen $v + \text{Kern}(f) = v' + \text{Kern}(f)$ ergibt sich mit dem Kriterium über die Gleichheit von Nebenklassen $v - v' \in \text{Kern}(f)$. Dies bedeutet $f(v - v') = 0$. Daraus folgt $f(v) = f(v')$.

Jetzt, da nachgewiesen ist, dass φ tatsächlich eine Abbildung ist, können wir schmarotzerhaft ausnutzen, dass φ über Repräsentanten definiert ist: Die restlichen Eigenschaften können wir beweisen, „ohne genau hinzusehen“:

- **Linearität von φ :** Seien $v, v' \in V$. Dann ist

$$\begin{aligned}\varphi[(v + \text{Kern}(f)) + (v' + \text{Kern}(f))] &= \varphi[v + v' + \text{Kern}(f)] = f(v + v') \\ &= f(v) + f(v') = \varphi(v + \text{Kern}(f)) + \varphi(v' + \text{Kern}(f)) .\end{aligned}$$

Die Homogenität folgt ähnlich einfach.

- Schließlich zur *Bijektivität* von φ . Da die Surjektivität von φ klar ist (denn die Bilder von φ und die Bilder von f stimmen überein), ist nur noch die Injektivität fraglich: Seien $v + \text{Kern}(f)$ und $v' + \text{Kern}(f)$ verschiedene Nebenklassen. Hätten diese dasselbe Bild unter φ , so wäre $f(v) = f(v')$. Daraus ergäbe sich $v - v' \in \text{Kern}(f)$, also $v + \text{Kern}(f) = v' + \text{Kern}(f)$, ein Widerspruch. \square

Die Bedeutung eines Satzes kann man auch daran messen, ob er angenehme Folgen hat. Hier sind drei äußerst angenehme Folgen des Homomorphiesatzes:

Dimensionsformel für lineare Abbildungen

Sei $f: V \rightarrow W$ eine lineare Abbildung. Sei $\dim(V) = n$. Dann gilt:

$$\dim(\text{Kern}(f)) + \dim(\text{Bild}(f)) = n .$$

Beweis Aus dem Homomorphiesatz und der zweiten Dimensionsformel ergibt sich

$$\dim(\text{Bild}(f)) = \dim(V / \text{Kern}(f)) = \dim(V) - \dim(\text{Kern}(f)) = n - \dim(\text{Kern}(f)) .$$

\square

Damit kann man die Dimension des Lösungsraums eines homogenen Gleichungssystems bestimmen (erinnern Sie sich an Abschn. 4.1.3). Das geht wie folgt:

Dimension des Lösungsraums

Sei $Ax = 0$ ein homogenes lineares Gleichungssystem, wobei A eine $m \times n$ -Matrix ist. Dann gilt

$$\dim(\mathbf{L}(A, 0)) = n - \text{Rang}(A) .$$

Beweis Wir fassen A als lineare Abbildung von K^n nach K^m auf und wenden die Dimensionsformel für lineare Abbildungen an. Nach Definition von $L(A, 0)$ gilt $\text{Kern}(A) = L(A, 0)$. Ferner ist $\text{Bild}(A)$ gleich dem Erzeugnis der Spalten von A . (Denn das Bild des i -ten Einheitsvektors ist die i -te Spalte von A .) Mit der Dimensionsformel folgt daraus

$$n = \dim(\text{Kern}(f)) + \dim(\text{Bild}(f)) = \dim(L(A, 0)) + \text{Rang}(A) .$$

□

Äquivalenz von Injektivität und Surjektivität

Sei f eine lineare Abbildung eines endlichdimensionalen Vektorraums V in einen Vektorraum W . Wenn $\dim(V) = \dim(W)$ ist, so gilt

$$f \text{ injektiv} \Leftrightarrow f \text{ surjektiv} .$$

Zum Nachweis der Bijektivität einer linearen Abbildung von V nach W braucht man also nur ihre Injektivität oder ihre Surjektivität zu zeigen.

Die Voraussetzung dieses Satzes ist insbesondere für eine lineare Abbildung eines endlichdimensionalen Vektorraums in sich erfüllt.

Beweis Es gilt

$$f \text{ injektiv} \Leftrightarrow \text{Kern}(f) = \{0\} \Leftrightarrow \dim(\text{Bild}(f)) = n \Leftrightarrow f \text{ surjektiv} .$$

□

Eine abschließende *Bemerkung*: Es ist im Allgemeinen keineswegs so, dass φ der einzige Isomorphismus zwischen $V/\text{Kern}(f)$ und $\text{Bild}(f)$ ist. Denn wir wissen: Wenn zwei Vektorräume isomorph sind, so ist jede lineare Abbildung, die eine Basis des einen auf eine Basis des anderen abbildet, ein Isomorphismus. Da es im Allgemeinen Myriaden von Basen gibt, gibt es im Allgemeinen also auch Myriaden von Isomorphismen. (Vergleichen Sie dazu auch Übungsaufgabe 15.)

5.4 Der Dualraum

Nun betrachten wir nicht mehr nur eine oder zwei lineare Abbildungen, sondern alle linearen Abbildungen einer gewissen Art. Damit erobern wir uns eine höhere Abstraktionsebene und erhalten dementsprechend höhere Einsichten.

Sei stets V ein n -dimensionaler K -Vektorraum. Wir betrachten lineare Abbildungen von V in den 1-dimensionalen K -Vektorraum K (vergleichen Sie dazu Abschn. 3.2.9).

Eine **Linearform** von V ist eine lineare Abbildung von V in K . Eine Linearform f ordnet also jedem Vektor v ein Körperelement $f(v)$ so zu, dass gilt

$$f(v + v') = f(v) + f(v') \quad \text{und} \quad f(kv) = k \cdot f(v) \quad \text{für alle} \quad v, v' \in V, k \in K.$$

Beispiel Sei $V = \mathbf{R}^3$. Die Abbildung f von \mathbf{R}^3 in \mathbf{R} , die einen Vektor (x, y, z) auf $x+y+z$ abbildet, ist eine Linearform.

Der Kern dieser Abbildung ist der Unterraum

$$\text{Kern}(f) = \{(x, y, -x - y) \mid x, y \in \mathbf{R}\}.$$

Der Kern hat also die Dimension 2, ist daher „ziemlich groß“. Dies ist ein Phänomen, das auch im Allgemeinen gilt:

Kern einer Linearform

Sei f eine Linearform von V , die nicht die Nullabbildung ist. Dann hat $\text{Kern}(f)$ die Dimension $n-1$.

Da f nicht die Nullabbildung ist, hat $\text{Bild}(f)$ mindestens die Dimension 1; da K aber 1-dimensional ist, folgt $\dim(\text{Bild}(f)) = 1$. Nach der Dimensionsformel für lineare Abbildungen ergibt sich daraus $\dim(\text{Kern}(f)) = n-1$. \square

Wir bezeichnen die Menge aller Linearformen von V mit V^* und nennen V^* den zu V **dualen Raum** oder den **Dualraum** von V . (Dies hat höchstens indirekt mit dem in Abschn. 4.3 definierten „dualen Code“ zu tun!)

Dieser Name ist eine Verpflichtung:

Satz vom Dualraum

Die Menge V^* ist bezüglich der folgendermaßen definierten Verknüpfungen $f+g$ und $k \cdot f$

$$(f + g)(v) := f(v) + g(v) \quad \text{und} \quad (k \cdot f)(v) := k \cdot f(v)$$

ein K -Vektorraum der Dimension n .

Beweis Der Nachweis der Vektorraumeigenschaften ist einfach und wird Ihnen als Übung (siehe Übungsaufgabe 20) überlassen.

Wie bitte? Sie protestieren? Das soll so eine üble k.o.-Formulierung sein („... ist einfach, wird als Übungsaufgabe ...“), die den Studierenden den letzten Mut nimmt?? Ich würde meine Prinzipien verletzen???

Nein, nein, nein!!! Das ist wirklich einfach! Soll ich's Ihnen beweisen? Gut. Also, was müssen wir tun? Richtig, wir müssen die Vektorraumaxiome nachprüfen. Der Reihe nach:

Ist die Summe zweier Linearformen wieder eine Linearform? Ist das skalare Vielfache einer Linearform wieder eine Linearform?

Ist die Addition assoziativ?

Was ist das Nullelement? (Das ist einfach: Die Nullabbildung ist eine Linearform; also wird es wohl schon die sein.)

Was ist das „Negative“ einer Linearform f ? Auch das ist nicht schwierig; dies ist die Abbildung $-f$, die definiert ist durch $(-f)(v) := -(f(v))$.

Ist die Addition kommutativ?

... und jetzt kommen noch die Gesetze der Skalarmultiplikation ...

Da hab ich mich aber gründlich verschätzt, das sind ja mindestens drei Übungsaufgaben! Ich mache Ihnen ein Kompromissangebot: Weisen Sie einfach nur die Hälfte der Eigenschaften nach.

Noch sind wir aber nicht fertig, denn wir müssen noch zeigen, dass V^* die Dimension n hat. Dazu braucht man einen kleinen Trick: Zu jeder Basis $B = \{v_1, \dots, v_n\}$ von V erhalten wir auf die folgende Weise eine Basis $B^* = \{v_1^*, \dots, v_n^*\}$ von V^* :

$$v_i^*(v_j) := \delta_{ij}.$$

Dabei hat das so genannte **Kroneckersymbol** δ_{ij} die Werte 0 oder 1, und zwar nimmt δ_{ij} den Wert 1 genau dann an, wenn $i = j$ ist. Die Vorschrift $v_i^*(v_j) := \delta_{ij}$ bedeutet also, dass v_i^* dem Vektor v_i den Wert 1 zuordnet und allen anderen Basisvektoren v_j den Wert 0. (Beachten Sie, dass man v_i^* – wie jede lineare Abbildung – nur auf den Vektoren der Basis B zu definieren braucht.) Die Menge B^* heißt die zu B **duale Basis**. (Das Symbol δ_{ij} ist nach Leopold Kronecker, 1823–1891, benannt.)

Wenn wir gezeigt haben, dass B^* eine Basis von V^* ist, haben wir nachgewiesen, dass V^* die Dimension n hat.

- B^* ist linear unabhängig: Sei $k_1 v_1^* + k_2 v_2^* + \dots + k_n v_n^* = 0$. Das bedeutet, dass $k_1 v_1^* + k_2 v_2^* + \dots + k_n v_n^*$ die Nullabbildung ist. Dann macht diese Linearform insbesondere all die Vektoren v_1, v_2, \dots, v_n zu Null. Daher gilt für alle $i = 1, \dots, n$:

$$0 = [k_1 v_1^* + k_2 v_2^* + \dots + k_n v_n^*](v_i) = k_1 v_1^*(v_i) + k_2 v_2^*(v_i) + \dots + k_n v_n^*(v_i) = k_i \cdot 1.$$

Also ist $k_i = 0$. Somit sind die Vektoren aus B^* linear unabhängig. (Machen Sie sich jedes Gleichheitszeichen in der obigen Gleichung explizit klar!)

- B^* ist ein Erzeugendensystem von V^* : Sei f eine beliebige Linearform. Wir wissen, dass f durch Vorgabe der Werte von f auf den Vektoren aus B eindeutig festgelegt ist. Sei also

$$f(v_i) := a_i \quad (i = 1, \dots, n).$$

Dann folgt $f = a_1 v_1^* + a_2 v_2^* + \dots + a_n v_n^*$, da die Linearformen auf beiden Seiten der behaupteten Gleichung auf den Vektoren aus B übereinstimmen. \square

Der obige Satz sagt insbesondere, dass V und V^* isomorph sind und dass die folgendermaßen definierte Abbildung φ ein Isomorphismus ist:

$$\varphi(k_1 v_1 + \dots + k_n v_n) := k_1 v_1^* + \dots + k_n v_n^* .$$

Sei $\{v_1, v_2, \dots, v_n\}$ eine Basis von V , und sei $\{v_1^*, v_2^*, \dots, v_n^*\}$ die zugehörige duale Basis. Ist $v = k_1 v_1 + \dots + k_n v_n \in V$, so bezeichnen wir die Linearform $k_1 v_1^* + \dots + k_n v_n^*$ mit v^* .

Wir können leicht das Bild eines beliebigen Vektors v unter einer beliebigen Linearform f ausrechnen: Dazu stellen wir v als Linearkombination der v_i und f als Linearkombination der v_j^* dar:

$$v = k_1 v_1 + k_2 v_2 + \dots + k_n v_n, f = a_1 v_1^* + a_2 v_2^* + \dots + a_n v_n^* .$$

Dann rechnet man ohne Schwierigkeiten nach, dass gilt

$$f(v) = a_1 k_1 + a_2 k_2 + \dots + a_n k_n .$$

Ihnen mag vielleicht die rechte Seite der letzten Gleichung wie ein „Skalarprodukt“ vorkommen, Sie werden aber wahrscheinlich Ihrer Vorstellung nicht trauen, weil die linke Seite ja wie etwas ganz anderes aussieht. Im Schlusskapitel 10 werden wir uns aber klarmachen, dass Ihre Vorstellung – wie könnte es anders sein – richtig ist! Durch Linearformen erhält man tatsächlich ein „Skalarprodukt“.

Natürlich kann man das Spiel mit dem Dualraum beliebig lange weiterspielen: Man kann (wenn man möchte) den Dualraum des Dualraums definieren (das sind die linearen Abbildungen von V^* nach K), ihm einen Namen geben (man nennt ihn den **Bidualraum** von V (bi: zwei) und bezeichnet ihn mit V^{**}) usw. Da $\dim(V^*) = \dim(V)$ ist, folgt auch $\dim(V^{**}) = \dim(V^*) = \dim(V)$. Also sind V^{**} und V isomorph. Es stellt sich aber heraus, dass V^{**} und V in besonders „natürlicher“ Weise isomorph zueinander sind. Das beweisen wir noch – um daraus die Berechtigung abzuleiten, später nur noch V und V^* zu betrachten.

Um nachzuweisen, dass V und V^{**} isomorph sind, müssen wir jedem Vektor $v \in V$ ein Element $v^{**} \in V^{**}$ bijektiv zuordnen. Was ist v^{**} ? Antwort: v^{**} muss ein Element aus V^{**} sein, also eine Abbildung, die jeder Linearform $f \in V^*$ ein Element aus K zuordnet. Das heißt $v^{**}(f)$ muss ein Element aus K sein.

Nach dieser Analyse, was zu tun ist, können wir uns jetzt fragen, was die naheliegendste Möglichkeit ist, dem Symbol $v^{**}(f)$ ein Körperelement zuzuordnen. Nach kurzem Nachdenken und ein bisschen Mut zur Einfachheit (eine Grundvoraussetzung für jeden Mathematiker!) fällt auch Ihnen bestimmt nichts Besseres ein, als zu definieren:

$$v^{**}(f) := f(v) .$$

Klar! $f(v)$ ist ja aus K , da f eine Linearform ist.

Jetzt kann uns nichts und niemand mehr aufhalten, den zugehörigen Satz zu formulieren und zu beweisen:

„Natürlicher Isomorphismus“ des Bidualraums

Sei V ein (endlichdimensionaler) Vektorraum. Dann ist die folgendermaßen definierte Abbildung φ ein Isomorphismus von V auf V^{**} :

$$\varphi(v) := v^{**},$$

wobei die Abbildung v^{**} durch $v^{**}(f) := f(v)$ definiert ist.

Beweis Wir müssen zeigen, dass $\varphi(v)$ tatsächlich in V^{**} liegt und dass φ linear und bijektiv ist.

„ $v^{**} = \varphi(v)$ liegt in V^{**} “: Wir haben uns oben schon klargemacht, dass v^{**} tatsächlich eine Abbildung von V^{**} nach K ist. Wir müssen also noch zeigen, dass v^{**} linear ist.

Dazu betrachten wir zwei Linearformen f und g sowie ein Körperelement k . Es ist zu zeigen, dass $v^{**}(f+g) = v^{**}(f) + v^{**}(g)$ und $v^{**}(kf) = kv^{**}(f)$ gilt. Dies folgt so:

$$v^{**}(f+g) = (f+g)(v) = f(v) + g(v) = v^{**}(f) + v^{**}(g)$$

und

$$v^{**}(kf) = (kf)(v) = k \cdot f(v) = k \cdot v^{**}(f).$$

(Machen Sie sich bei jedem Gleichheitszeichen klar, warum dies gilt!)

„ φ ist linear“: Seien $v, w \in V$, und sei $k \in K$. Wir zeigen die Gleichheit von $\varphi(v+w)$ und $\varphi(v) + \varphi(w)$, indem wir zeigen, dass diese beiden Abbildungen aus V^{**} eine Linearform f gleich abbilden:

$$\begin{aligned} \varphi(v+w)(f) &= (v+w)^{**}(f) = f(v+w) = f(v) + f(w) = v^{**}(f) + v^{**}(w) \\ &= \varphi(v) + \varphi(w). \end{aligned}$$

Entsprechend zeigt man die Gleichheit von $\varphi(kv)$ und $k\varphi(v)$:

$$\varphi(kv)(f) = (kv)^{**}(f) = f(kv) = k \cdot f(v) = k \cdot v^{**}(f) = k \cdot \varphi(v)(f).$$

„ $\varphi(v)$ ist bijektiv“: Wegen der Äquivalenz von Injektivität und Surjektivität einer linearen Abbildung müssen wir nur zeigen, dass φ injektiv ist. Dazu genügt es, folgendes zu zeigen: Ist $v \neq 0$, so ist auch $\varphi(v) = v^{**} \neq 0$. Sei also v ein von Null verschiedener Vektor von V . Wir müssen zeigen, dass es eine Linearform f gibt mit $v^{**}(f) \neq 0$. Dazu ergänzen wir v zu einer Basis B von V und definieren eine Linearform f dadurch, dass wir festlegen: f soll auf allen Vektoren von B den Wert 1 annehmen. Dann ist

$$v^{**}(f) = f(v) = 1.$$

Damit ist alles gezeigt. □

5.5 Verständnisfragen, Übungen und Tipps

Richtig oder falsch?

1. Thema: Lineare Abbildungen

Sei f eine lineare Abbildung von \mathbf{R}^m in \mathbf{R}^n . Dann gilt:

- ☐ $n \leq m$,
- ☐ $m \leq n$,
- ☐ $m = n$.
- ☐ Nur der Nullvektor wird auf den Nullvektor abgebildet.
- ☐ Es gibt genau drei lineare Abbildungen von \mathbf{R}^3 nach \mathbf{R}^3 .
- ☐ Es gibt unendlich viele lineare Abbildungen von \mathbf{R}^m nach \mathbf{R}^n .
- ☐ Es gibt unendlich viele lineare Abbildungen von \mathbf{R} nach \mathbf{R} .

2. Thema: Lineare Abbildungen

Sei f eine lineare Abbildung eines Vektorraums V in einen Vektorraum W . Dann gilt:

- ☐ $f(o) = o$.
- ☐ $f(1) = 1$.
- ☐ $f(-v) = -f(v)$ für alle $v \in V$.
- ☐ $f(-v) = -f \cdot v$ für alle $v \in V$.
- ☐ $f(kv) = f(k) + f(v)$ für alle $k \in K$ und alle $v \in V$.
- ☐ $f(k \cdot v) = f(k) \cdot f(v)$ für alle $k \in K$ und alle $v \in V$.
- ☐ $\text{Bild}(f) = W$.

3. Thema: Kern einer linearen Abbildung

Sei U der Kern einer linearen Abbildung $f: V \rightarrow W$. Dann gilt:

- ☐ $U = \{w \in W \mid f(w) = o\}$.
- ☐ $U = \{w \in V \mid f(w) = o\}$.
- ☐ $U = \{f(v) \mid v = o\}$.
- ☐ $U = \{v \mid v = o\}$.
- ☐ $U = \{v \in V \mid f(v) = 1\}$.
- ☐ Ist f eine Projektion aus W_0 auf U , so ist $\text{Kern}(f) = \text{Bild}(f)$.
- ☐ Ist f eine Projektion aus W_0 auf U , so sind $\text{Kern}(f)$ und $\text{Bild}(f)$ komplementär.

4. Thema: Rang einer linearen Abbildung

Der Rang einer linearen Abbildung $f: V \rightarrow W$ ist gleich:

- ☐ $\dim(V)$,
- ☐ $\dim(W)$,
- ☐ $\dim(W) - \dim(V)$,
- ☐ $\dim(\text{Bild}(f))$,
- ☐ $\dim(\text{Kern}(f))$,
- ☐ der Rang einer geeigneten Darstellungsmatrix von f ,
- ☐ der Rang jeder Darstellungsmatrix von f .

5. Thema: Linearformen

Eine Linearform des K -Vektorraums V ist

- ☐ eine Abbildung von V in V ,
- ☐ eine Abbildung von K in V ,
- ☐ eine Abbildung von V in K ,
- ☐ eine Abbildung von K in K ,
- ☐ der Übergang von v auf v^* ,
- ☐ ein Vektor aus V ,
- ☐ ein Körperelement.

Übungsaufgaben

1. Seien V und W Vektorräume über dem Körper K , und sei $w_0 \in W$. Zeigen Sie: Die folgendermaßen definierte Abbildung

$$f(v) := w_0$$

von V nach W ist genau dann eine lineare Abbildung, wenn $w_0 = 0$ ist.

2. Sind die folgenden Abbildungen $f: \mathbb{R}^3 \rightarrow \mathbb{R}^2$, $g: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $h: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ lineare Abbildungen? Wenn nein, warum nicht?

$$f: (a, b, c) \mapsto (ab, a + b),$$

$$g: (a, b) \mapsto (3a + 1, 4b + a + 1),$$

$$h: (a, b) \mapsto (a + 3b, b - 3a, 0).$$

3. Sei f eine lineare Abbildung des K -Vektorraums V in den K -Vektorraum W .
 - (a) Seien v_1, v_2, v_3 Vektoren aus V . Zeigen Sie: $f(v_1 + v_2 + v_3) = f(v_1) + f(v_2) + f(v_3)$.
 - (b) Seien v_1, \dots, v_s Vektoren aus V und k_1, \dots, k_s Elemente von K . Dann gilt:

$$f\left(\sum_{i=1}^s k_i v_i\right) = \sum_{i=1}^s k_i f(v_i).$$

4. Sei $f: V \rightarrow W$ eine lineare Abbildung, sei $\{v_1, \dots, v_n\}$ eine Basis von V und seien die Vektoren $w_i \in W$ definiert durch $w_i := f(v_i)$. Zeigen Sie: Wenn f injektiv ist, so sind die Vektoren w_1, \dots, w_n linear unabhängig.
5. Sei $f: V \rightarrow W$ eine lineare Abbildung, sei $\{v_1, \dots, v_n\}$ eine Basis von V und seien die Vektoren $w_i \in W$ definiert durch $w_i := f(v_i)$. Zeigen Sie: Wenn f surjektiv ist, so ist $\{w_1, \dots, w_n\}$ ein Erzeugendensystem von W .
6. Geben Sie sämtliche linearen Abbildungen von $\text{GF}(2)^2 \rightarrow \text{GF}(2)^2$ an.
7. Machen Sie sich klar, dass der Ableitungsoperator aus der Analysis eine lineare Abbildung ist. (Legen Sie zunächst Definitions- und Bildbereich fest.)

8. Sei $f: \mathbf{R}_3 \rightarrow \mathbf{R}_3$ die folgendermaßen definierte Abbildung:

$$f: (x, y, z) \mapsto (x + 2y + z, y + z, -x + 3y + 4z).$$

Zeigen Sie, dass f eine lineare Abbildung ist. Bestimmen Sie $\text{Kern}(f)$ und $\text{Bild}(f)$.

9. Bestimmen Sie Kern und Bild der linearen Abbildung $f: \mathbf{R}_3 \rightarrow \mathbf{R}_3$, die durch folgende Angaben definiert ist:

$$f(1, 0, 0) := (-1, 1, 3),$$

$$f(0, 1, 0) := (0, 6, 3),$$

$$f(0, 0, 1) := (2, 4, -3).$$

10. Sei $f: \mathbf{R}_3 \rightarrow \mathbf{R}_3$ die folgendermaßen definierte Abbildung:

$$f: (x, y, z) \mapsto (x - y + z, -6y + 12z, -2x + 2y - 2z).$$

(a) Zeigen Sie: f ist eine lineare Abbildung.

(b) Berechnen Sie die Darstellungsmatrizen ${}_B M_C(f)$ von f bezüglich folgender Basen:

$$B = C = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$$

und

$$B = C = \{(-1, 0, 1), (-1, 2, 1), (-2, 0, 4)\}.$$

11. Sei $f: \mathbf{R}_3 \rightarrow \mathbf{R}_3$ die folgendermaßen definierte Abbildung von \mathbf{R}_3 nach \mathbf{R}_3 :

$$f: (x, y, z) \mapsto (-5x - 18y - 24z, 4x + 13y + 16z, -2x - 6y - 7z).$$

Bestimmen Sie die Darstellungsmatrizen ${}_B M_B(f)$ und ${}_{B'} M_{B'}(f)$ mit

(a) $B = \{(1, 0, 0), (0, 1, 1), (-1, 0, 1)\}$,

und

(b) $B' = \{(3, -1, 0), (-1, -1, 1), (-3, 2, -1)\}$.

(c) Berechnen Sie ${}_B M_{B'}(\text{id})$ und ${}_{B'} M_B(\text{id})$.

(d) Verifizieren Sie, dass gilt

$${}_{B'} M_{B'}(f) = {}_B M_{B'}(\text{id}) \cdot {}_B M_B(f) \cdot {}_{B'} M_B(\text{id}).$$

Bestimmen Sie eine Basis von $\text{Kern}(f)$ und eine Basis von $\text{Bild}(f)$. Welche Dimension hat $\mathbf{R}^3 / \text{Kern}(f)$?

12. Sei A die Darstellungsmatrix einer linearen Abbildung f . Zeigen Sie: Der von den Spalten von A aufgespannte Unterraum von K_m ist isomorph zu $\text{Bild}(f)$.
13. Sei V ein n -dimensionaler Vektorraum, und sei $\{v_1, v_2, \dots, v_n\}$ eine Basis von V . Sei f die durch

$$f(v_i) = a \cdot v_i + \sum_{i \neq j} v_j$$

definierte lineare Abbildung von V in sich. Bestimmen Sie die Dimension von $\text{Bild}(f)$.

14. Sei V ein beliebiger Vektorraum einer Dimension ≥ 2 über einem beliebigen Körper. Geben Sie eine lineare Abbildung von V in sich an, die nicht injektiv ist und eine, die nicht surjektiv ist.
15. Geben Sie diejenigen Vektorräume V an, in denen es *genau eine* bijektive lineare Abbildung von V in sich gibt. Für welche Vektorräume V und W gibt es *nur einen* Isomorphismus von V nach W .
16. Zeigen Sie, dass Ähnlichkeit eine Äquivalenzrelation auf der Menge der $n \times n$ -Matrizen ist.
17. Sei $V = K^{3 \times 3}$ der Vektorraum aller 3×3 -Matrizen über einem Körper K . Zeigen Sie, dass die Abbildung $f: V \rightarrow K$, die durch

$$f \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} := a_{11} + a_{22} + a_{33}$$

definiert ist, eine lineare Abbildung ist. Geben Sie eine Basis von $\text{Kern}(f)$ an.

(Man nennt die Summe der Diagonalelemente einer quadratischen Matrix M die **Spur** von M ; obige Abbildung heißt auch die **Spurabbildung**.)

18. Sei $V = K^{2 \times 2}$ der Vektorraum aller 2×2 -Matrizen über einem Körper K . Zeigen Sie, dass die Abbildung $f: V \rightarrow K$, die durch

$$f \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} := a_{11} + a_{12} + a_{21} + a_{22}$$

definiert ist, eine lineare Abbildung ist.

Geben Sie eine Basis von $\text{Kern}(f)$ an.

19. Sei $V = K^{3 \times 3}$ der Vektorraum aller 3×3 -Matrizen über einem Körper K .
(a) Zeigen Sie, dass die Abbildung $f: V \rightarrow K^3$, die durch

$$f \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} := \begin{pmatrix} a_{11} + a_{22} + a_{33} \\ a_{12} + a_{23} + a_{31} \\ a_{13} + a_{21} + a_{32} \end{pmatrix}$$

definiert ist, eine lineare Abbildung ist.

(b) Bestimmen Sie die Dimension von $\text{Bild}(f)$ und von $\text{Kern}(f)$.

(c) Geben Sie eine Basis von $\text{Kern}(f)$ an.

20. Zeigen Sie, dass der Dualraum V^* eines K -Vektorraums V seinerseits ein K -Vektorraum ist.

Projekt: $\text{Hom}(V, W)$

In diesem Projekt soll die Menge aller linearen Abbildungen (Homomorphismen) eines K -Vektorraums V in einen K -Vektorraum W untersucht werden. Diese Menge wird mit

$\text{Hom}(V, W)$ bezeichnet:

$$\text{Hom}(V, W) = \{f : V \rightarrow W \mid f \text{ ist lineare Abbildung}\}.$$

1. Zeigen Sie, dass $\text{Hom}(V, W)$ ein K -Vektorraum ist. Definieren Sie dazu zunächst die Addition und die Skalarmultiplikation.
2. Zeigen Sie: Sei $\dim(V) = n$, $\dim(W) = m$. Dann hat der Vektorraum $\text{Hom}(V, W)$ die Dimension $n \cdot m$. Geben Sie eine Basis von $\text{Hom}(V, W)$ an.
Da zu jeder linearen Abbildung von V nach W eine $m \times n$ -Matrix gehört, und da $K^{m \times n}$ ebenfalls die Dimension mn hat, könnte man vermuten, dass dieser Raum isomorph zu $\text{Hom}(V, W)$ ist. Dies ist tatsächlich der Fall:
3. Zeigen Sie: $\text{Hom}(V, W)$ ist isomorph zu $K^{m \times n}$.
Geben Sie einen Isomorphismus an.
Gibt es einen Isomorphismus, den Sie als „natürlich“ bezeichnen würden?

Sie sollten mit folgenden Begriffen umgehen können

Lineare Abbildung, Homomorphismus, Bild, Kern, Projektion, Darstellungsmatrix, Darstellungssatz, Basistransformation, Homomorphiesatz, Dimensionsformel für lineare Abbildungen, Linearform, Dualraum, duale Basis.



Polynomringe spielen in der *Algebra* eine ausgesprochene Hauptrolle. In der *Linearen Algebra* kommen sie nur an einer, allerdings entscheidenden Stelle ins Spiel, nämlich bei der Frage der Diagonalisierbarkeit linearer Abbildungen (Kap. 8). Dort treten das „charakteristische Polynom“ und das „Minimalpolynom“ auf. Insbesondere zur Definition und zum Studium des Minimalpolynoms braucht man gründliche Kenntnisse über Polynomringe. Außerdem behandeln wir Polynomringe auch deswegen relativ ausführlich, weil sie ein grundlegendes (und sehr schönes) Stück Mathematik sind.

6.1 Ringe

Der Begriff des Rings verallgemeinert den Begriff des Körpers. Ein **Ring** besteht aus einer Menge R von Elementen zusammen mit zwei Verknüpfungen $+$ und \cdot , die je zwei Elementen $x, y \in R$ wieder ein Element $x + y$ bzw. $x \cdot y$ von R zuordnen. Damit eine solche Struktur Ring genannt wird, müssen die folgenden drei Gruppen von Gesetzen für alle Elemente $x, y, z \in R$ erfüllt sein:

6.1.1 Gesetze der Addition

- *Assoziativität:*

$$(x + y) + z = x + (y + z) .$$

- *Existenz und Eindeutigkeit des neutralen Elements:* Es gibt genau ein Element von R , das wir 0 („Nullelement“) nennen, für das gilt

$$0 + x = x .$$

- *Existenz und Eindeutigkeit inverser Elemente:* Zu jedem x gibt es genau ein Element, das wir $-x$ nennen, für das gilt

$$x + -x = 0 .$$

- *Kommutativität:*

$$x + y = y + x .$$

6.1.2 Gesetz der Multiplikation

- *Assoziativität:*

$$x \cdot (y \cdot z) = (x \cdot y) \cdot z$$

6.1.3 Distributivgesetze

$$x \cdot (y + z) = x \cdot y + x \cdot z ,$$

$$(x + y) \cdot z = x \cdot z + y \cdot z .$$

Es ist klar, dass jeder Körper ein Ring ist. Was aber fehlt einem Ring zum Körper? Vor allem zwei Dinge: Die Existenz eines Einselements, also eines neutralen Elements bezüglich der Multiplikation, und die Existenz inverser Elemente. Während die meisten Ringe ein Einselement besitzen (vergleichen Sie die Beispiele, die wir gleich anschließend behandeln), erweist sich die Existenz (genauer gesagt: die Nichtexistenz) multiplikativ inverser Elemente als Hauptdefizit von Ringen. (Man kann das natürlich auch anders sehen: Gerade das macht Ringe so interessant!)

Eine Bemerkung zur Namengebung: Die Bezeichnung „Ring“ für einen Ring soll nicht etwas „Rundes“ assoziieren, sondern einen Zusammenschluss von Elementen zu einem Ganzen. Diese Bedeutung des Wortes „Ring“ ist in der deutschen Sprache fast vollständig ausgestorben; sie ist nur noch in altertümlichen Vereinsbezeichnungen wie „Deutscher Ring“, „Weißer Ring“, „RCDS“, ... zu finden.

Wir betrachten zunächst die wichtigsten Beispiele von Ringen. Das erste Beispiel ist uns schon längst vertraut (vergleichen Sie hierzu Abschn. 2.2.3).

Die Menge \mathbf{Z} der ganzen Zahlen ist zusammen mit der gewöhnlichen Addition und Multiplikation ein Ring. Dieser Ring hat ein Einselement (nämlich die Zahl 1) und ist **kommutativ** (das bedeutet, dass auch die Multiplikation kommutativ ist).

Sei n eine natürliche Zahl. Mit \mathbf{Z}_n bezeichnen wir die Struktur, die aus der Menge der Zahlen $0, 1, \dots, n-1$ und den Operationen $+_n$ und \cdot_n besteht, die wie folgt definiert sind:

$$a +_n b = a + b \mod n ,$$

$$a \cdot_n b = a \cdot b \mod n .$$

Man nennt \mathbf{Z}_n den **Restklassenring modulo n** . Diese Strukturen sind die zweite wichtige Klasse von Ringen:

Die Menge \mathbf{Z}_n ist zusammen mit $+_n$ und \cdot_n ein Ring. Dieser Ring hat ein Einselement und ist kommutativ. (Diese Tatsache ergibt sich ebenso wie die erste durch einfaches Nachrechnen; siehe Übungsaufgabe 1.)

Wir geben eine weitere wichtige Klasse von Ringen an. Sei $K^{n \times n}$ die Menge aller $n \times n$ -Matrizen über dem Körper K mit der üblichen Matrizenaddition und -multiplikation als Verknüpfung.

Satz über Matrizenringe

Die Menge $K^{n \times n}$ ist zusammen mit der Matrizenaddition und -multiplikation ein Ring; dieser hat ein Einselement, ist aber für $n \geq 2$ nicht kommutativ.

Dass $K^{n \times n}$ ein Ring ist, ergibt sich ganz einfach (siehe Übungsaufgaben 2 und 3). Um zu zeigen, dass dieser Ring nicht kommutativ ist, genügt es, zwei Matrizen A und B anzugeben mit $AB \neq BA$; im Fall $n = 2$ sieht man etwa

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

□

Die (für uns) wichtigste Klasse von Ringen sind jedoch die Polynomringe; deshalb behandeln wir diese auch in einem eigenen Abschnitt.

6.2 Was ist eigentlich x ?

Jeder glaubt zu wissen, was ein Polynom ist. Zum Beispiel ist $x^2 - 1$ ein Polynom. Aber: Was bedeutet dabei eigentlich das Symbol x ? Es heißt, das sei eine „Unbekannte“; manche nennen x auch „Variable“. In der Schule heißt es auch, x sei ein „Platzhalter“. Als Mathematiker muss man aber doch fragen: Was ist eigentlich x in Wirklichkeit? Jedenfalls muss man auf eine solche Frage gefasst sein – und eine Antwort darauf wissen. Eine Zahl? Ein Vektor? Eine Folge? Eine Matrix? Eine Menge? Irgendetwas muss es doch sein!

Wir werden in diesem Abschnitt den mystischen Schleier von x entfernen. Anschließend werden Sie genau sagen können, was x ist – und von da an genauso mit Polynomen rechnen können wie bisher – aber guten Gewissens!

Wir beginnen mit einer Struktur, die scheinbar gar nichts mit Polynomen zu tun hat.

Sei stets K ein kommutativer Körper. Wir betrachten die Menge V_∞ aller unendlichen Folgen (a_0, a_1, a_2, \dots) mit Elementen aus K , die *nur endlich viele Komponenten* $\neq 0$ haben. (Vergleichen Sie dazu den Abschn. 3.2.5 und das Projekt am Ende von Kap. 3.)

Sie erinnern sich sicher, dass dieser Vektorraum zwar unendliche Dimension hat, wir ihn aber gut beschreiben können. Insbesondere erinnern wir uns daran, dass die Menge

$$B = \{e_i | i \in \mathbf{N}\}$$

eine Basis von V_∞ ist, wobei e_i die Folge ist, die nur an der i -ten Stelle eine 1 hat und sonst aus Nullen besteht.

Wir führen jetzt zusätzlich eine Multiplikation auf V_∞ ein. Seien

$$f = (a_0, a_1, a_2, \dots, a_n, 0, \dots) \quad \text{und} \quad g = (b_0, b_1, b_2, \dots, b_m, 0, \dots)$$

Elemente von V_∞ . Das **Produkt** $f \cdot g$ ist definiert als die Folge $f \cdot g = (c_0, c_1, c_2, \dots)$ mit

$$c_k = a_0 b_k + a_1 b_{k-1} + \dots + a_{k-1} b_1 + a_k b_0 \left(= \sum_{i=0}^k a_i b_{k-i} \right).$$

Bevor Sie sich über diese Multiplikationsvorschrift zu wundern anfangen, beweisen wir schnell einen Satz, der eine erste Rechtfertigung für diese Multiplikationsvorschrift ist.

Satz über den Polynomring

Der Vektorraum V_∞ ist mit der komponentenweisen Addition und der eben definierten Multiplikation ein kommutativer Ring; das Einselement ist die Folge $(1, 0, 0, \dots)$.

Beweis Da die Addition komponentenweise definiert ist, ergeben sich die Gesetze der *Addition* für V_∞ aus denen von K .

Zu den Gesetzen der *Multiplikation*: Dazu zeigen wir zunächst, dass $f \cdot g$ tatsächlich wieder in V_∞ liegt: Ist nämlich $k > m + n$, so treten in der Summe

$$c_k = \sum_{i=0}^k a_i b_{k-i}$$

nur Summanden $a_i b_j$ auf mit $i > n$ oder $j > m$. Also ist $a_i = 0$ oder $b_j = 0$. Das heißt: Wenn $k > m + n$ ist, so gilt $c_k = 0$. Also hat die Folge $f \cdot g$ an höchstens $m + n$ Stellen einen von Null verschiedenen Wert, und somit ist $f \cdot g \in V_\infty$.

Der Nachweis des Assoziativgesetzes ist zwar (technisch gesehen) völlig trivial, da man letztlich alles auf die Assoziativität der Basiselemente zurückführen kann, aber außerordentlich mühsam; wir ersparen uns daher diese Tortur; ich empfehle Ihnen dies nicht als Übung.

Schließlich zeigen wir noch die Distributivgesetze. Seien dazu $f = (a_0, a_1, a_2, \dots)$, $g = (b_0, b_1, b_2, \dots)$ und $h = (c_0, c_1, c_2, \dots)$ Elemente von V_∞ . Dann ist die k -te Komponente d_k von $f(g + h)$ gleich

$$d_k = \sum_{i=0}^k a_i (b_{k-i} + c_{k-i}) .$$

Da die k -te Komponente e_k von $f \cdot g + f \cdot h$ gleich

$$e_k = \sum_{i=0}^k a_i b_{k-i} + \sum_{i=0}^k a_i c_{k-i}$$

ist, ist die Behauptung $f \cdot (g + h) = f \cdot g + f \cdot h$ bewiesen.

Entsprechend zeigt man das andere Distributivgesetz. □

Nun betrachten wir spezielle Elemente von V_∞ . Die einfachsten Elemente sind die Folgen der Form $(a, 0, \dots)$ mit $a \in K$. Für diese gilt

$$(a, 0, \dots) \cdot f = a \cdot f \quad \text{und} \quad f \cdot (a, 0, \dots) = a \cdot f \quad \text{für alle} \quad f \in V_\infty .$$

Warum? Sei $f = (b_0, b_1, b_2, \dots)$. Dann ist die k -te Komponente c_k von $(a, 0, \dots) \cdot f$ gleich

$$c_k = \sum_{i=0}^k a_i b_{k-i} = a b_k .$$

Entsprechend folgt $f \cdot (a, 0, \dots) = a \cdot f$. □

Aufgrund dieser Tatsache können wir die Folge $(a, 0, \dots)$ mit dem Körperelement a identifizieren. Dies soll hinfort auch geschehen: $a = (a, 0, 0, \dots)$.

Insbesondere halten wir fest, dass das Element $1 = (1, 0, 0, \dots)$ ein neutrales Element bezüglich der Multiplikation, also ein Einselement ist.

Aber viel interessanter sind die Folgen des Typs e_i , also neben $1 = (1, 0, 0, \dots)$ die Folgen $e_1 = (0, 1, 0, \dots)$, $e_2 = (0, 0, 1, 0, \dots)$ usw. Was passiert, wenn wir diese miteinander multiplizieren? Was ist $e_s \cdot e_t$? Was ist e_1^2 , e_1^3 ?

Dies können wir einfach ausrechnen: Sei $e_s \cdot e_t = (c_0, c_1, c_2, \dots)$. Um die Koeffizienten c_k zu bestimmen, müssen wir die Produkte $a_i b_j$ mit $i + j = k$ berechnen. Da in e_s bzw. e_t nur der Koeffizient a_s bzw. b_t verschieden von Null ist, ist jedenfalls $c_k = 0$, falls $k \neq s + t$ ist. Im Fall $k = s + t$ ist auch nur ein Summand ungleich Null, nämlich $a_s \cdot b_t = 1 \cdot 1 = 1$. Somit ergibt sich $e_s \cdot e_t = e_{s+t}$.

Insbesondere ist

$$e_1^2 = e_2, e_1^3 = e_2 \cdot e_1 = e_3, \dots, e_1^n = e_n .$$

Wenn wir noch $e_1^0 := e_0$ setzen, ergibt sich, dass die Menge

$$B = \{e_1^n \mid n = 0, 1, 2, \dots\}$$

unsere wohlbekannte Basis von V_∞ ist.

Jetzt kommt's! Wir definieren einfach

$$x := e_1 .$$

Das ominöse x ist also nicht mehr als eine Abkürzung für eine spezielle Folge, nämlich

$$x = (0, 1, 0, 0, \dots) .$$

Die obigen Überlegungen sagen, dass die Menge $\{x^0 (= 1), x, x^2, x^3, \dots\}$ eine Basis des Vektorraums V_∞ ist. Daher können wir jedes Element $f = (a_0, a_1, a_2, \dots, a_n, 0, \dots)$ von V_∞ eindeutig als Linearkombination dieser Basis darstellen; es ergibt sich

$$f = a_0 \cdot x^0 + a_1 \cdot x + a_2 \cdot x^2 + \dots + a_n \cdot x^n .$$

Das kommt Ihnen sicher schon sehr vertraut vor!

Nun holen wir tief Luft und definieren ganz mutig: Der Vektorraum V_∞ wird ab jetzt mit $K[x]$ bezeichnet und heißt der **Polynomring** in der **Unbestimmten** x . Jedes Element von $K[x]$ wird ein **Polynom** (über K) genannt. Wenn

$$f = a_0 \cdot x^0 + a_1 x + a_2 x^2 + \dots + a_n x^n$$

ein Polynom ist, so heißen die Körperelemente a_0, a_1, \dots, a_n die **Koeffizienten** des Polynoms f . Statt $a_0 \cdot x^0$ schreiben wir auch einfach a_0 . Die Folge $(0, 0, \dots)$ wird als **Nullpolynom** bezeichnet. Ist

$$f = a_0 \cdot x^0 + a_1 \cdot x + a_2 \cdot x^2 + \dots + a_n \cdot x^n$$

ein Polynom, und ist $a_n \neq 0$, so heißt n der **Grad** des Polynoms f ; wir schreiben $n = \text{Grad}(f)$; das Nullpolynom hat den **Grad** $-\infty$.

Polynome sind also – wie gewohnt – Ausdrücke wie etwa $x^2 - 1$, $7x^5 + 3x^2 + x + 1$. Die Multiplikationsregel für Folgen in V_∞ überträgt sich ganz einfach auf die (vertraute!) Vorschrift zur Multiplikation von Polynomen: Das Produkt der Polynome

$$f = a_0 + a_1 \cdot x + a_2 \cdot x^2 + \dots + a_n \cdot x^n \quad \text{und} \quad g = b_0 + b_1 \cdot x + b_2 \cdot x^2 + \dots + b_n \cdot x^n$$

ist ein Polynom $h = c_0 + c_1 x + c_2 x^2 + \dots$ mit

$$c_k = \sum_{i=0}^k a_i b_{k-i} .$$

Statt $f \cdot g$ schreiben wir oft einfach auch fg .

Machen wir uns das an einem *Beispiel* nochmals klar: Sei $f = x^3 + 2x^2 + 2x + 1$ und $g = x - 3$. Dann kann man $f \cdot g = (x^3 + 2x^2 + 2x + 1) \cdot (x - 3)$ so ausrechnen, dass man zunächst f

mit x und dann f mit -3 multipliziert und dann die beiden Zwischenergebnisse zusammenzählt. Das bedeutet:

$$f \cdot g = (x^3 + 2x^2 + 2x + 1) \cdot (x - 3) = (x^3 + 2x^2 + 2x + 1) \cdot x + (x^3 + 2x^2 + 2x + 1) \cdot (-3) = \\ [x^4 + 2x^3 + 2x^2 + x] + [-3x^3 - 6x^2 - 6x - 3] = x^4 - x^3 - 4x^2 - 5x - 3.$$

Man tut also so, als ob man mit dem Symbol x „einfach so“ rechnen könnte. Das sollten Sie nach den bisherigen Überlegungen zwar nicht mehr in aller Unschuld, aber dafür umso überzeugter tun.

Auf eine kleine Finesse möchte ich Sie hinweisen: Bei der Berechnung des ersten Summanden geht alles seinen ordentlichen Gang:

$$(x^3 + 2x^2 + 2x + 1) \cdot x = x^3 \cdot x + 2x^2 \cdot x + 2x \cdot x + 1 \cdot x = x^4 + 2x^3 + 2x^2 + x,$$

... aber wenn Sie bei der Berechnung des zweiten Summanden ...

$$(x^3 + 2x^2 + 2x + 1) \cdot (-3) = x^3 \cdot (-3) + 2x^2 \cdot (-3) + 2x \cdot (-3) + 1 \cdot (-3)$$

... (scheinbar) genau aufpassen, so stellen Sie sich an dieser Stelle vielleicht die kritische Frage, ob (und wenn ja warum) Sie die -3 nach vorne ziehen dürfen. Mit anderen Worten: Warum dürfen Sie x mit jedem Körperelement vertauschen? Weil Sie das schon immer so gemacht haben? Weil dies nutzlose Haarspaltereien sind und Sie sich weigern, dabei mitzumachen? Weil K kommutativ ist? – Nichts von alledem! Die Antwort lautet: Weil wir das so definiert haben! *Das wollen Sie genau wissen?* Bitte schön: Was ist x ? Nach Definition ist x die Folge $(0, 1, 0, 0, \dots)$. Was ist also das Produkt eines beliebigen Körperelements k (zum Beispiel $k = -3$) mit x ? Ganz einfach: Die Folge $(0, k, 0, 0, \dots)$. Wie lautet diese Folge in Polynomschreibweise? Noch einfacher: $k \cdot x$. – Sehen Sie: $x \cdot k = k \cdot x$.

Bemerkung Bis zu dieser Stelle hätte man Polynomringe auch allgemeiner entwickeln können, und zwar so, dass man anstelle des Körpers K einen kommutativen Ring R zugrundelegt. Dann hätte man die Elemente von $R[x]$ definiert als Folgen (r_0, r_1, \dots) , bei denen höchstens endlich viele Komponenten ungleich Null sind; die Addition hätte man komponentenweise erklärt und die Multiplikation auf die uns nun schon vertraute, nur scheinbar komplizierte Weise. Dann hätten wir beweisen können, dass $R[x]$ ein kommutativer Ring ist. Wieder hätte man x definiert als die Folge $(0, 1, 0, 0, \dots)$; dann lässt sich auch jedes Element f von $R[x]$ eindeutig in der Form

$$f = r_0 + r_1x + r_2x^2 + \dots + r_nx^n$$

(für eine natürliche Zahl n) schreiben (siehe Übungsaufgabe 7).

Ein solches Vorgehen ist sinnvoll, wenn man auch Polynome in mehreren Veränderlichen studieren möchte. Wenn man zum Beispiel den Polynomring in der Unbestimmten

x und y über einem Körper K konstruieren will, so konstruiert man zunächst $K[x]$, fasst dann $K[x]$ als Ring R auf und bildet $R[y] = K[x][y]$ und nennt dieses Gebilde $K[x, y]$ (siehe Übungsaufgabe 8).

Ab jetzt müssen wir aber voraussetzen, dass wir einen Polynomring über einem Körper K haben. Schon der nächste Satz ist im Allgemeinen für Ringe falsch.

Gradformel für Polynome

Seien f und g Polynome vom Grad n und m aus $K[x]$, wobei K ein Körper ist. Dann hat $f \cdot g$ den Grad $n + m$.

Beweis Seien f und g wie oben, aber verschieden vom Nullpolynom, und sei $h = f \cdot g = c_0 + c_1x + c_2x^2 + \dots$ das Produkt von f und g . Ist $k > m + n$, so ist

$$c_k = \sum_{i=0}^k a_i b_{k-i}$$

gleich Null, da in jedem Summanden $a_i b_j$ entweder $i > n$ oder $j > m$ ist (oder beides), und also $a_i b_j = 0$ ist. Also hat $h = fg$ höchstens den Grad $n + m$. Da sich aber $c_{n+m} = a_n b_m \neq 0$ ergibt, hat fg tatsächlich den Grad $n + m$. Beachten Sie, dass wir $a_n b_m \neq 0$ schließen konnten, weil $K \setminus \{0\}$ bezüglich der Multiplikation abgeschlossen ist („Nullteilerfreiheit“; vergleichen Sie hierzu Abschn. 2.1.3).

Ist f oder g das Nullpolynom, so ist auch fg das Nullpolynom, und die Behauptung gilt ebenfalls. \square

Invertierbare Polynome

Die einzigen invertierbaren Polynome aus $K[x]$ sind die Polynome vom Grad 0, also die konstanten Polynome $a \in K \setminus \{0\}$.

Beweis Sei f ein invertierbares Polynom. Das bedeutet, dass es ein Polynom g gibt mit $f \cdot g = 1$. Da das konstante Polynom 1 den Grad 0 hat, müssen nach der Gradformel auch die Polynome f und g einen Grad ≤ 0 haben. Da f nicht das Nullpolynom sein kann, muss f also den Grad 0 haben. \square

Wir haben jetzt noch einen Begriff zu klären, der nicht ganz einfach ist, mit dem wir aber vertraut sein müssen. Es handelt sich darum, dass man gewohnt ist, in Polynome konkrete Werte „einsetzen“ zu dürfen. Was soll das bedeuten? Zunächst scheint das ganz einfach zu sein: Statt x schreibt man das, was man einsetzen möchte.

An dieser Stelle sollte sich Ihr Gewissen zu Wort melden. Sie wissen genau, was x ist, nämlich eine unendliche Folge von Elementen aus K . Es kann doch nicht angehen, dass man anstelle dieses komplexen Objekts einfach eine Zahl schreibt (und dann die Augen zumacht und hofft, alles würde gut gehen)! Das ist ja schlimmer als lechts und rinks zu verwechseln!

Man möchte aber Elemente in Polynome einsetzen, viel schlimmer: Wir möchten nicht nur Elemente aus K („Zahlen“) einsetzen, sondern auch viel komplexere Objekte, zum Beispiel Matrizen. Was passiert dabei? Kann das gut gehen?

Auch diesen Einsetzungsprozess kann man so exakt und nüchtern beschreiben, dass Sie anschließend wieder munter einsetzen können, ohne nachts schweißgebadet aufzuwachen, und nicht zu wissen, ob Sie etwas getan haben, was „man nicht tut“.

Sei K wie bisher ein kommutativer Körper, und sei R ein Ring mit Einselement 1, der K enthält, so dass die Einschränkung der Addition und Multiplikation von R auf K die Addition und Multiplikation von K ergibt. (Man sagt dann auch, K sei ein **Unterring** von R .)

Beispiele

(a) $R = K$ ist ein triviales Beispiel für diese Situation.

(b) Sei $R = K^{n \times n}$ der Ring der $n \times n$ -Matrizen mit Einträgen aus K . Ist K in R enthalten? Ja! Es ist so, dass R einen zu K isomorphen Unterring enthält, nämlich die Menge K' der Matrizen $k \cdot E_n$ ($k \in K$). Wenn man die Körperelemente k mit der Matrix $k \cdot E_n$ identifiziert (es kommt uns ja nicht darauf an, wie die Objekte heißen), kann man tatsächlich sagen, dass K ein Unterring von R ist.

Dieser Ring enthält auch ein Einselement, nämlich die Einheitsmatrix.

Sei $f = a_0x^0 + a_1x + a_2x^2 + \dots + a_nx^n \in K[x]$ ein Polynom. Für jedes Element $r \in R$ definieren wir

$$f(r) = a_01 + a_1r + a_2r^2 + \dots + a_nr^n.$$

Man sagt, dass $f(r)$ durch **Einsetzen** von r in f entsteht.

Achtung Das Polynom $f \in K[x]$ dient sozusagen nur als „Muster“ zur Bildung von $f(r)$; bei der Bildung von $f(r)$ kommen nur Addition und Multiplikation von R vor!

Machen wir uns dies an einem *Beispiel* klar. Sei f das Polynom $x^2 + x - 6 \in \mathbf{R}[x]$, und sei M die folgende Matrix

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

Was ist $f(M)$? Dies berechnet sich wie folgt:

$$\begin{aligned} f(M) &= \begin{pmatrix} 2 & 1 \\ 0 & -3 \end{pmatrix} \cdot \begin{pmatrix} 2 & 1 \\ 0 & -3 \end{pmatrix} + \begin{pmatrix} 2 & 1 \\ 0 & -3 \end{pmatrix} - \begin{pmatrix} 6 & 0 \\ 0 & 6 \end{pmatrix} = \begin{pmatrix} 4 & -1 \\ 0 & 9 \end{pmatrix} \\ &\quad + \begin{pmatrix} -4 & 1 \\ 0 & -9 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

Für uns ist folgende Erkenntnis die entscheidende:

Satz vom Einsetzungshomomorphismus

Sei r ein Element des Rings R . Wenn r mit jedem Element von K vertauschbar ist (d. h. $kr = rk$ – und also auch $kr^2 = r^2k$ usw. – für alle $k \in K$), dann gilt für alle Polynome $f, g \in K[x]$:

$$(f + g)(r) = f(r) + g(r) ,$$

$$(f \cdot g)(r) = f(r) \cdot g(r) .$$

(Man sagt dann auch, dass die Abbildung

$$\varphi : K[x] \rightarrow R : f \mapsto f(r)$$

ein **Ringhomomorphismus** ist und nennt ihn den **Einsetzungshomomorphismus**.)

Beweis Seien $f = a_0x^0 + a_1x + a_2x^2 + \dots + a_nx^n$ und $g = b_0x^0 + b_1x + b_2x^2 + \dots + b_nx^n$ (da bei darf a_n oder b_n gleich Null sein). Zunächst zeigen wir die Additivität:

$$\begin{aligned} (f + g)(r) &= (a_0 + b_0) + (a_1 + b_1)r + (a_2 + b_2)r^2 + \dots + (a_n + b_n)r^n \\ &= a_0 + a_1r + a_2r^2 + \dots + a_nrn + b_0 + b_1r + b_2r^2 + \dots + b_nr^n = f(r) + g(r) . \end{aligned}$$

Nun zum Produkt $f \cdot g$: Der Koeffizient c_k von x^k in $f \cdot g$ ist

$$c_k = \sum_{i=0}^k a_i b_{k-i}$$

Also ergibt sich

$$\begin{aligned} (f \cdot g)(r) &= a_0b_0 \cdot 1 + (a_0b_1 + a_1b_0) \cdot r + \dots + \left(\sum_{i=0}^k a_i b_{k-i} \right) \cdot r^k + \dots \\ &= a_0b_0 \cdot 1 + a_0b_1 \cdot r + a_1b_0 \cdot r + \dots + \sum_{i=0}^k a_i b_{k-i} r^k + \dots \end{aligned}$$

Andererseits ist

$$f(r) \cdot g(r) = a_0b_0 \cdot 1 + a_0(b_1 \cdot r) + (a_1 \cdot r)b_0 + \dots + \sum_{i=0}^k a_i \cdot r^i b_{k-i} \cdot r^{k-i} + \dots$$

Da r mit jedem Element von K , insbesondere also mit allen a_i und b_j , vertauschbar ist, folgt $(fg)(r) = f(r) \cdot g(r)$. \square

Bemerkung Die Voraussetzung, dass jedes Element von R mit jedem Element von K vertauschbar sein muss, ist im wichtigsten Anwendungsfall $R = K^{n \times n}$ erfüllt. Denn jede Matrix der Form $k \cdot E_n$ ist mit jeder $n \times n$ -Matrix vertauschbar (siehe Übungsaufgabe 6).

6.3 Polynomdivision

Wir haben gesehen, dass ein Polynom im Allgemeinen keine multiplikative Inverse hat. Das bedeutet, dass man in den meisten Fällen nicht durch ein Polynom teilen kann. Man möchte aber trotzdem ein Polynom durch ein anderes dividieren! Überraschenderweise geht das auch gut – allerdings bleibt bei der Division ein Rest! Dies wird in folgendem Satz präzise formuliert. Dieser Satz ist die Grundlage für alle unsere weiteren Untersuchungen über Polynomringe.

Polynomdivision

Seien f und g Polynome über dem Körper K . Wenn g nicht das Nullpolynom ist, dann gibt es eindeutig bestimmte Polynome q, r aus $K[x]$ mit

$$f = qg + r \quad \text{und} \quad \text{Grad}(r) < \text{Grad}(g).$$

Beweis Wir müssen zeigen, dass es solche Polynome gibt (Existenz) und dass es nur ein solches Paar von Polynomen gibt (Eindeutigkeit).

Zunächst zeigen wir die *Existenz* von q und r . Dies geschieht durch Induktion nach $n = \text{Grad}(f)$. Sei $m = \text{Grad}(g)$.

Ist $n < m$, so kann man q als Nullpolynom und $r = f$ wählen. Dann erfüllen q und r die geforderten Bedingungen.

Seien nun n und m natürliche Zahlen mit $n \geq m$. Indem man g mit x^{n-m} multipliziert, erhält man ein Polynom vom Grad n . Indem man dieses noch mit einem geeigneten Körperelement a multipliziert, erhält man ein Polynom $g' (= a \cdot x^{n-m} \cdot g)$ vom Grad n , das bei x^n den gleichen Koeffizienten wie f hat. Daher ist

$$h = f - g'$$

ein Polynom, dessen Grad kleiner als n ist. Somit können wir auf h die Induktionsannahme anwenden: Es gibt also Polynome q' und r mit

$$h = q'g + r \quad \text{und} \quad \text{Grad}(r) < \text{Grad}(g).$$

Indem wir die Definition von h einsetzen, erhalten wir

$$f = h + g' = q'g + r + a \cdot x^{n-m} \cdot g = (q' + a \cdot x^{n-m}) \cdot g + r = qg + r,$$

wobei

$$q = q' + a \cdot x^{n-m}$$

ist. Damit haben wir Polynome q und r gefunden, die die gewünschte Gleichung erfüllen.

Als zweiten Schritt beweisen wir die *Eindeutigkeit* von q und r . Seien q, r sowie q', r' Polynome mit

$$f = qg + r \quad \text{und} \quad \text{Grad}(r) < \text{Grad}(g)$$

und

$$f = q'g + r' \quad \text{und} \quad \text{Grad}(r') < \text{Grad}(g).$$

Daraus ergibt sich

$$(q - q')g = qg - q'g = r' - r.$$

Da die rechte Seite dieser Gleichung ein Polynom ist, dessen Grad kleiner als $\text{Grad}(g)$ ist, muss auch die linke Seite einen Grad kleiner als $\text{Grad}(g)$ haben. Da die linke Seite aber ein Vielfaches von g ist, kann dies nur dann der Fall sein, wenn das fragliche Vielfache von g das Nullpolynom ist.

Also ist $r - r'$ das Nullpolynom, und somit ist auch $q' - q$ das Nullpolynom. Das heißt $r = r'$ und $q' = q$. \square

Dieser Satz wird uns zu zwei Zwecken dienen. Einerseits werden wir damit all das beweisen können, was wir über „Nullstellen“ von Polynomen wissen müssen; andererseits werden wir ihn dazu benutzen, eine für uns entscheidende Strukturaussage über Polynomringe zu beweisen; hierbei geht es um einen bedeutenden Klassifikationssatz für so genannte „Ideale“ von Ringen.

Sei $K[x]$ der Polynomring über dem kommutativen Körper K . Sei $f \in K[x]$. Man nennt ein Körperelement k eine **Nullstelle** von f , falls $f(k) = 0$ ist.

Lemma über die Nullstellen eines Polynoms

Sei k eine Nullstelle des Polynoms $f \in K[x]$. Dann gibt es ein Polynom $q \in K[x]$ mit $f = q \cdot (x - k)$.

Beweis Wir dividieren f durch das Polynom $x - k$:

$$f = q \cdot (x - k) + r \quad \text{mit} \quad \text{Grad}(r) < \text{Grad}(g) = 1.$$

Daher muss r den Grad 0 haben oder das Nullpolynom sein. Wenn r das Nullpolynom ist, steht die Behauptung da.

Wäre r nicht das Nullpolynom, so könnten wir durch Einsetzen schließen, dass $r(k) = 0$ ist. Denn nach Voraussetzung ist $f(k) = 0$. Auch wenn wir k in das Polynom $x - k$ einsetzen, erhalten wir 0; formal ausgedrückt heißt dies $(x - k)(k) = 0$. Der Satz über den Einsetzungshomomorphismus liefert damit

$$r(k) = [f - q \cdot (x - k)](k) = f(k) - q(k) \cdot (x - k)(k) = 0 - q(k) \cdot 0 = 0.$$

Das einzige konstante Polynom mit einer Nullstelle ist aber das Nullpolynom; somit ist $f = q \cdot (x - k)$, wie behauptet. \square

Das obige Lemma heißt oft auch der **Satz von Ruffini** (nach Paolo Ruffini (1765–1822)).

In Analogie zu der Situation in \mathbf{Z} sagen wir, dass ein Polynom g ein Polynom $f \in K[x]$ **teilt**, wenn es ein Polynom $h \in K[x]$ gibt mit $f = g \cdot h$.

Sei k eine Nullstelle des Polynoms $f \in K[x]$. Dann heißt die größte natürliche Zahl v , so dass $(x - k)^v$ das Polynom f teilt, die **Vielfachheit der Nullstelle** k von f .

Ist v die Vielfachheit der Nullstelle k von f , so kann man f also schreiben als

$$f = (x - k)^v \cdot q$$

für ein geeignetes Polynom $q \in K[x]$.

Ein Polynom mit Hilfe seiner Nullstellen so zu beschreiben, wie Sie es gewohnt sind, ist nicht ohne. Dazu brauchen wir einige Hilfsmittel. Das erste Hilfsmittel ist ein Lemma, das auf den französischen Mathematiker Étienne Bézout (1730–1783) zurückgeht.

Lemma von Bézout

Seien f und g Polynome aus $K[x]$. Wenn es kein Polynom vom Grad ≥ 1 gibt, das sowohl f als auch g teilt (man sagt dazu auch, dass f und g **teilerfremd** sind), dann gibt es Polynome f^* und g^* mit

$$f \cdot f^* + g \cdot g^* = 1$$

Zum *Beispiel* sind die Polynome $f = x^4 + x^3 + x^2 + x + 1$ und $g = x^2 + x + 1$ teilerfremd. Und in der Tat ergibt sich mit $f^* := -x$ und $g^* := x^3 + 1$

$$f \cdot f^* + g \cdot g^* = (x^4 + x^3 + x^2 + x + 1)(-x) + (x^2 + x + 1)(x^3 + 1) = 1.$$

Nun zum *Beweis* des Lemmas von Bézout. Sei o. B. d. A. $\text{Grad}(f) \geq \text{Grad}(g)$. Wir können ebenfalls o. B. d. A. voraussetzen, dass g nicht das Nullpolynom ist. [Wenn dies so wäre, so

müsste f den Grad 0 haben. (Denn f teilt das Nullpolynom, also g ; da f und g teilerfremd sind, hat f also einen Grad < 1 .) Dann ist f ein Körperelement $\neq 0$, also innerhalb von K invertierbar. Die Behauptung folgt jetzt mit $f^* := f^{-1}$ und $g^* := g$.]

Wir beweisen die Behauptung durch Induktion nach $\text{Grad}(g)$. Da g nicht das Nullpolynom ist, gilt $\text{Grad}(g) \geq 0$.

Sei zunächst $\text{Grad}(g) = 0$. Dann ist g ein von 0 verschiedenes Element des Körpers K ; also gibt es in K ein zu g inverses Element g^* (nämlich g^{-1}). Mit $f^* := 0$ folgt dann

$$f \cdot f^* + g \cdot g^* = f \cdot 0 + g \cdot g^{-1} = 1.$$

Sei jetzt $\text{Grad}(g) \geq 1$, und sei die Aussage des Lemmas richtig für alle Polynome g' mit $\text{Grad}(g') < \text{Grad}(g)$. Nach dem Satz über Polynomdivision gibt es Polynome q und r mit

$$f = qg + r \quad \text{und} \quad \text{Grad}(r) < \text{Grad}(g)$$

Da $\text{Grad}(r) < \text{Grad}(g)$ ist, ist es verführerisch, auf die Polynome g und r die Induktionsvoraussetzung anzuwenden. Das dürfen wir aber nur, wenn diese Polynome auch die Voraussetzungen des Lemmas erfüllen! Welche Voraussetzungen? – Eben die, dass die Polynome teilerfremd sind. Das rechnen wir einfach nach: Angenommen, es gäbe ein Polynom h vom Grad ≥ 1 , das g und r teilt. Dann würde h auch $qg + r$, und also f teilen. Dies ist ein Widerspruch zur Voraussetzung.

Jetzt kann uns aber nichts mehr hindern, die Induktionsvoraussetzung anzuwenden: Es gibt also Polynome g' und r' mit

$$g \cdot g' + r \cdot r' = 1.$$

Zusammen folgt die Behauptung

$$1 = g \cdot g' + r \cdot r' = g \cdot g' + (f - qg)r' = f \cdot r' + g \cdot (g' - qr') = f \cdot f^* + g \cdot g^*$$

mit $f^* := r'$ und $g^* := g' - qr'$. □

Das Lemma von Bézout hat eine interessante und in der Algebra äußerst wichtige Konsequenz (siehe dazu auch Übungsaufgaben 20 und 21). Wir nennen ein Polynom $f \in K[x]$ **irreduzibel**, falls sein Grad mindestens 1 ist und falls jedes Polynom aus $K[x]$, das f teilt, entweder Grad 0 hat oder ein skalares Vielfaches von f ist. (Die irreduziblen Polynome entsprechen den Primzahlen in \mathbb{Z} .) Aus dem Lemma von Bézout ergibt sich die folgende Tatsache:

Invertierbarkeit modulo eines irreduziblen Polynoms

Sei f ein irreduzibles Polynom aus $K[x]$. Dann gibt es zu jedem Polynom g aus $K[x]$ mit $\text{Grad}(g) < \text{Grad}(f)$ ein Polynom g^* aus $K[x]$, so dass $g \cdot g^* - 1$ ein Vielfaches von f ist. (Man sagt dazu auch, dass $g \cdot g^*$ „kongruent 1 modulo f “ ist.) \square

Die Bedeutung dieses Satzes ist Ihnen wahrscheinlich nicht unmittelbar klar; wenn Sie mehr wissen wollen, empfehle ich Ihnen Übungsaufgabe 21.

Je mehr Nullstellen eines Polynoms man kennt, desto besser kann man es beschreiben:

Satz über die Nullstellen eines Polynoms (René Descartes, 1596–1650)

Seien k_1, k_2, \dots, k_s paarweise verschiedene Nullstellen mit Vielfachheiten v_1, v_2, \dots, v_s des Polynoms $f \in K[x]$. Dann kann man f schreiben als

$$f = (x - k_1)^{v_1} \cdot (x - k_2)^{v_2} \cdot \dots \cdot (x - k_s)^{v_s} \cdot q$$

für ein geeignetes Polynom $q \in K[x]$.

Das wichtigste Korollar aus diesem Satz ist der

Satz über die Anzahl der Nullstellen

Ein Polynom vom Grad n über einem kommutativen Körper hat höchstens n verschiedene Nullstellen.

Der Beweis hierfür ist klar: Wir zerlegen f so weit wie möglich in Linearfaktoren. Sei also

$$f = (x - k_1)^{v_1} \cdot (x - k_2)^{v_2} \cdot \dots \cdot (x - k_s)^{v_s} \cdot q$$

mit einem Polynom q , das keine Nullstelle hat. Wir betrachten jetzt das Polynom

$$f^* = (x - k_1)^{v_1} \cdot (x - k_2)^{v_2} \cdot \dots \cdot (x - k_s)^{v_s}.$$

Da f^* die Faktoren $(x - k_1)^{v_1}, (x - k_2)^{v_2}, \dots, (x - k_s)^{v_s}$ hat, ist der Grad von f^* genau $v_1 + v_2 + \dots + v_s$. Insbesondere gilt für die Anzahl s aller Nullstellen von f :

$$s \leq v_1 + v_2 + \dots + v_s = \text{Grad}(f^*) \leq \text{Grad}(f) = n.$$

Also hat f höchstens n Nullstellen. \square

Der obige Satz über die Nullstellen eines Polynoms wird durch Induktion nach s *bewiesen*.

Für $s = 1$ folgt die Behauptung aus der Definition der Vielfachheit der Nullstelle k_1 .

Sei nun $s > 1$, und seien k_1, \dots, k_{s-1}, k_s paarweise verschiedene Nullstellen von f . Nach Induktion ist der Satz richtig für k_1, \dots, k_{s-1} . Also gibt es natürliche Zahlen v_1, \dots, v_{s-1} und ein Polynom $g \in K[x]$ mit

$$f = (x - k_1)^{v_1} \cdot \dots \cdot (x - k_{s-1})^{v_{s-1}} \cdot g.$$

Jetzt setzen wir die Nullstelle k_s in f ein. Da das Einsetzen von k_s ein Homomorphismus ist, folgt

$$\begin{aligned} 0 = f(k_s) &= [(x - k_1)^{v_1} \cdot \dots \cdot (x - k_{s-1})^{v_{s-1}} \cdot g](k_s) \\ &= (x - k_1)^{v_1}(k_s) \cdot \dots \cdot (x - k_{s-1})^{v_{s-1}}(k_s) \cdot g(k_s). \end{aligned}$$

Da

$$(x - k_1)^{v_1}(k_s) \cdot \dots \cdot (x - k_{s-1})^{v_{s-1}}(k_s) = (k_s - k_1)^{v_1} \cdot \dots \cdot (k_s - k_{s-1})^{v_{s-1}} \neq 0$$

ist, muss $g(k_s) = 0$ gelten.

Also gibt es ein Polynom q mit $g = q \cdot (x - k_s)^{v_s}$. Daraus folgt die Behauptung. \square

6.4 Ideale von $K[x]$

Nun wenden wir uns einer äußerst wichtigen Strukturaussage über Polynomringe zu.

Sei R ein kommutativer Ring. Eine Teilmenge I von R heißt ein **Ideal** von R , falls die folgenden Eigenschaften gelten:

- $I \neq \emptyset$,
- (*Additivität*) Wenn $a, b \in I$ sind, so sind auch $a + b \in I$ und $a - b \in I$,
- (*magnetische Anziehungseigenschaft*) Wenn $a \in I$ und $r \in R$ ist, so ist auch $r \cdot a \in I$.

Gibt es heute noch Ideale? – In der Algebra jedenfalls sind Ideale billig zu haben:

Jeder Ring hat die **trivialen** Ideale $\{0\}$ und R .

Ein nichttriviales Beispiel eines Ideals von $K[x]$ ist die Menge aller Polynome, bei denen das Absolutglied gleich Null ist. Dieses Ideal I kann in einer Formel so beschrieben werden:

$$I = \{x \cdot f \mid f \in K[x]\}.$$

Diese Konstruktion kann erheblich verallgemeinert werden: Man nehme irgendwelche Elemente a_1, a_2, a_3, \dots eines Ringes R , der ein Einselement 1 hat, und betrachte alle „Linearkombinationen mit Elementen aus R “, das heißt die Menge

$$I(a_1, a_2, a_3, \dots) = \{r_1 a_1 + r_2 a_2 + r_3 a_3 + \dots \mid r_1, r_2, r_3, \dots \in R\}.$$

Wir können uns leicht überzeugen, dass dies ein Ideal ist: Wegen $a_1 \in I(a_1, a_2, a_3, \dots)$ (dazu braucht man das Einselement) ist diese Menge nicht leer.

Auch die zweite Eigenschaft ist erfüllt, denn die Summe zweier Elemente

$$r_1 a_1 + r_2 a_2 + r_3 a_3 + \dots + s_1 a_1 + s_2 a_2 + s_3 a_3 + \dots \in I(a_1, a_2, a_3, \dots)$$

ist

$$(r_1 + s_1)a_1 + (r_2 + s_2)a_2 + (r_3 + s_3)a_3 + \dots,$$

also wieder ein Element von $I(a_1, a_2, a_3, \dots)$.

Sei schließlich $r \in R$ beliebig. Dann ist mit $r_1 a_1 + r_2 a_2 + r_3 a_3 + \dots$ auch

$$r \cdot (r_1 a_1 + r_2 a_2 + r_3 a_3 + \dots) = (rr_1)a_1 + (rr_2)a_2 + (rr_3)a_3 + \dots$$

in $I(a_1, a_2, a_3, \dots)$.

Also ist $I(a_1, a_2, a_3, \dots)$ tatsächlich ein Ideal; man nennt es das von a_1, a_2, a_3, \dots **erzeugte** Ideal.

Dies sind offenbar „einfache“ (oder „schöne“) Ideale; die schönsten Ideale sind die von einem einzigen Element erzeugten Ideale. Ist $I = I(a)$, so nennt man I das von a erzeugte **Hauptideal**. Ein Hauptideal besteht also nur aus den Vielfachen eines einzigen Elements; dies sind die Ideale, die man am einfachsten beschreiben kann. In der Literatur findet man häufig auch die Schreibweise (a) für $I(a)$.

Hier stellen sich verschiedene Fragen: Kann man erkennen, ob ein Ideal ein Hauptideal ist? Ist es vielleicht so, dass die meisten Ideale Hauptideale sind? Vielleicht (verwegene Frage!) sind überhaupt alle Ideale Hauptideale? Die Antwort auf diese Frage ist im Allgemeinen ein klares „Nein“ – in dem uns interessierenden Fall ist die Antwort jedoch „ja“. Die überraschende Tatsache ist die folgende:

Ideale eines Polynomrings

Jedes Ideal eines Polynomrings $K[x]$ in einer Unbestimmten ist ein Hauptideal!

Dieser Satz sagt also insbesondere folgendes: Für jedes Ideal $I(f_1, f_2, f_3, \dots)$ von $K[x]$, das von beliebig vielen Polynomen f_1, f_2, f_3, \dots erzeugt wird, gibt es ein Polynom g mit

$$I(f_1, f_2, f_3, \dots) = I(g).$$

Zum Beispiel kann man sich fragen: Was ist $I(x^3 - 1, x^4 - 1)$? Die Antwort ist: Dies ist das Hauptideal, das von $x - 1$ erzeugt wird: $I(x^3 - 1, x^4 - 1) = I(x - 1)$. (Vergleichen Sie dazu die Übungsaufgaben 24 bis 27.)

Schöner („idealer“) geht's nicht mehr!

Eigentlich dürfen wir uns darüber aber erst dann richtig freuen, wenn wir diese Tatsache *bewiesen* haben. Dies muss jetzt auch geschehen. Es zeigt sich, dass dieser Satz nur der Satz über Polynomdivision in einer raffinierten Kostümierung ist.

Dazu betrachten wir ein beliebiges Ideal I . Wir müssen zeigen, dass es ein Polynom g gibt mit $I = I(g)$. Wir können o. B. d. A. voraussetzen, dass $I \neq \{0\}$ ist, denn das „Nullideal“ $\{0\}$ wird von dem Nullpolynom 0 erzeugt.

Daher können wir voraussetzen, dass I Polynome vom Grad ≥ 0 enthält.

Der Trick des Beweises besteht darin, ein Polynom in I zu betrachten, das nicht das Nullpolynom ist, und unter allen diesen Polynomen kleinsten Grad hat. Genauer gesagt: Wir wissen, dass die Menge der Grade ≥ 0 der in I liegenden Polynome nicht leer ist; also hat diese Menge nichtnegativer ganzer Zahlen ein kleinstes Element m . Sei g ein Polynom vom Grad m in I .

Behauptung g erzeugt I .

Dazu müssen wir zeigen, dass es für jedes Polynom $f \in I$ ein Polynom $q \in K[x]$ gibt mit $f = q \cdot g$.

Um die Behauptung nachzuweisen, wenden wir Polynomdivision an: Es gibt jedenfalls Polynome q und r mit

$$f = qg + r \quad \text{und} \quad \text{Grad}(r) < \text{Grad}(g).$$

Nach Definition eines Ideals ist mit g auch $(-q)g$ in I ; mit f und $-qg$ ist auch $f - qg$ in I . Das bedeutet, dass $r = f - qg$ ein Element des Ideals I ist!

Da aber $\text{Grad}(r) < \text{Grad}(g)$ ist und da der Grad von g minimal unter allen nichtnegativen Graden der Polynome in I ist, folgt daraus, dass r einen negativen Grad haben muss. Somit ist r das Nullpolynom, und es folgt in der Tat

$$f = qg.$$

□

Bemerkung Es ist klar, dass ein Ideal von $K[x]$ im allgemeinen von verschiedenen Polynomen erzeugt werden kann; denn zum Beispiel wird jedes Ideal $I(g)$ von $K[x]$ auch von dem Polynom kg erzeugt ($k \in K$). Wenn man sich allerdings auf **normierte** erzeugende Polynome beschränkt (also solche, bei denen der höchste Koeffizient gleich 1 ist), so kann man sagen: *Jedes vom Nullideal verschiedene Ideal von $K[x]$ wird von einem eindeutig bestimmten normierten Polynom erzeugt.*

6.5 Verständnisfragen, Übungen und Tipps

Richtig oder falsch?

1. Thema: Polynome

Welcher der folgenden Ausdrücke ist ein Polynom aus $\mathbf{R}[x]$?

☐ $x^{1000000000000000000000000}$

☐ $x + 5y$

☐ $x + x^{-1}$

☐ $x^2 - x - 1 = 0$

☐ $y = x^2$

☐ $\sum_{i=0}^n a_i \cdot x^i$

☐ $\sum_{i=0}^{\infty} a_i \cdot x^i$

☐ 0

☐ 2001

☐ -3,14

☐ -3,14x

☐ $x^{3,14}$

☐ $1 + x + 1/2x^2 + 1/6x^3 + 1/24x^4 + \dots + 1/n! x^n + \dots$

2. Thema: Einsetzungshomomorphismus

☐ In ein Polynom aus $K[x]$ darf man nur Elemente aus K einsetzen.

☐ Obwohl ein Polynom f aus $K[x]$ nur über K definiert ist, darf man in f auch Elemente, die nicht in K liegen, einsetzen.

3. Thema: Ideale von $K[x]$

Sei K ein Körper.

☐ Es gibt ein Ideal von $K[x]$, das kein Element enthält.

☐ Es gibt ein Ideal von $K[x]$, das genau ein Element enthält.

☐ Es gibt ein Ideal von $K[x]$, das genau zwei Elemente enthält.

☐ Es gibt ein Ideal von $K[x]$, das alle Elemente von $K[x]$ enthält.

☐ Es gibt ein Ideal $\neq K[x]$ von $K[x]$, das unendlich viele Elemente enthält.

Übungsaufgaben

1. Zeigen Sie, dass \mathbf{Z}_n ein kommutativer Ring mit einem Einselement ist.

2. Seien $A, B, C \in \mathbf{K}^{2 \times 2}$. Zeigen Sie

$$(A \cdot B) \cdot C = A \cdot (B \cdot C)$$

(Assoziativität der Multiplikation) und

$$(A + B) \cdot C = A \cdot C + B \cdot C$$

(Distributivgesetz).

3. Zeigen Sie, dass die Menge $K^{n \times n}$ aller $n \times n$ -Matrizen über dem Körper K zusammen mit der Matrizenaddition und -multiplikation einen Ring bildet.
4. Zeigen Sie, dass die Multiplikation von Matrizen aus $K^{n \times n}$ für $n \geq 2$ nicht kommutativ ist.
5. Begründen Sie, weshalb die Menge, die aus den invertierbaren $n \times n$ -Matrizen und der Nullmatrix besteht, für $n \geq 2$ zusammen mit der Matrizenaddition und -multiplikation *keinen* Körper bildet.
6. Zeigen Sie, dass jede Matrix der Form $k \cdot E_n$ mit jeder Matrix aus $K^{n \times n}$ vertauschbar ist.
7. Sei R ein kommutativer Ring. Zeigen Sie
 - (a) $R[x]$ ist ein kommutativer Ring.
 - (b) Wenn R ein Einselement (neutrales Element bezüglich der Multiplikation) hat, so hat auch $R[x]$ ein Einselement.
8. Sei K ein Körper. Definieren Sie $K[x, y, z]$ und zeigen Sie, dass $K[x, y, z]$ ein kommutativer Ring mit Einselement ist.
9. Bestimmen Sie Polynome $q, r \in R[x]$ so, dass

$$x^7 + x^5 + x^3 + 1 = (x^3 + x + 1) \cdot q + r \quad \text{mit} \quad \text{Grad}(r) < 3$$

gilt.

10. Bestimmen Sie Polynome $q, r \in \mathbb{Z}_2[x]$ so, dass

$$x^7 + x^5 + x^3 + 1 = (x^3 + x + 1) \cdot q + r \quad \text{mit} \quad \text{Grad}(r) < 3$$

gilt.

11. Sie Polynome $q, r \in \mathbb{Z}_3[x]$, so dass gilt:

$$x^6 + x^4 + x^2 + 2x = (2x^2 + x + 1) \cdot q + r \quad \text{und} \quad \text{Grad}(r) < 2.$$

12. Bestimmen Sie die Nullstellen des reellen Polynoms $x^3 - 19x + 30$.
13. Ist das Polynom $x^3 - 2x + 4$
 - (a) als Polynom über \mathbb{R} ,
 - (b) als Polynom über \mathbb{Q}
 in Linearfaktoren zerlegbar?
14. Sie die Polynome $x^4 + 1$, $x^4 + x^3 + x + 1$ und $x^5 + 1$ aus $\mathbb{Z}_2[x]$ in irreduzible Polynome:
15. $f \in K[x]$ ein Polynom vom Grad n .
 - (a) Angenommen, Sie kennen n Nullstellen k_1, \dots, k_n von f . Können Sie daraus die Koeffizienten von f bestimmen? Und wenn Sie wissen, dass das Polynom normiert ist?
 - (b) Angenommen, Sie kennen an $n+1$ Stellen x_0, x_1, \dots, x_n die Werte y_0, y_1, \dots, y_n von f (das heißt $y_i = f(x_i)$). Können Sie daraus die Koeffizienten von f bestimmen? [Wenn Sie diese Aufgabe lösen, haben Sie die nach Joseph Louis Lagrange

(1736–1813) benannte Interpolationsformel gefunden. Leider hat sie Lagrange schon vor Ihnen gefunden ...]

16. Zeigen Sie (ohne den Satz über die Nullstellen eines Polynoms zu benutzen!), dass die einzige Nullstelle des Polynoms $(x - a)^v$ mit $a \in K$ das Element a ist.
17. Zeigen Sie: Zu jedem Ideal $I \neq \{0\}$ von $K[x]$ gibt es *genau ein* normiertes Polynom g mit $I = I(g)$.
18. Sei K ein Körper. Zeigen Sie:
 - (a) Wenn es ein Polynom aus $K[x]$ gibt, das nicht das Nullpolynom ist, das aber $f(k) = 0$ für alle $k \in K$ erfüllt, so ist K endlich.
 - (b) Wenn K endlich ist, so gibt es ein Polynom $f \in K[x]$ mit $f(k) = 0$ für alle $k \in K$.
19. Sei U ein Unterraum des Vektorraums $K[x]$. Zeigen Sie: Genau dann ist U ein Ideal des Rings $K[x]$, wenn gilt:

$$f \in U \Rightarrow x \cdot f \in U.$$

20. Sei I ein Ideal des Ringes R . Eine **Nebenklasse von R nach I** ist eine Menge der Form $r + I := \{r + i \mid i \in I\}$. Der **Faktorring** von R nach I ist die Menge R/I aller Nebenklassen zusammen mit der induzierten Addition und Multiplikation.
 - (a) Definieren Sie die Addition und Multiplikation auf R/I präzise.
 - (b) Zeigen Sie, dass die Addition und Multiplikation auf R/I wohldefiniert sind.
 - (c) Zeigen Sie, dass R/I ein Ring ist.
21. Sei f ein Polynom aus $K[x]$. Zeigen Sie:
 - (a) $K[x]/I(f)$ ist ein Ring. [Hinweis: Benutzen Sie die vorige Aufgabe.]
 - (b) Wenn f irreduzibel ist, dann ist $K[x]/I(f)$ ein Körper. [Benutzen Sie die Invertierbarkeit modulo eines irreduziblen Polynoms.]
 - (c) Wenn es ein über $\text{GF}(p)$ irreduzibles Polynom vom Grad n gibt, so gibt es einen Körper mit p^n Elementen.
22. (a) Zeigen Sie: Sei z eine komplexe Zahl. Dann ist die Menge aller Polynome f aus $\mathbb{R}[x]$ mit $f(z) = 0$ ein Ideal von $\mathbb{R}[x]$.
 (b) Geben Sie im Fall $z = i$ und im Fall $z = 1 + i$ ein erzeugendes Element dieses Ideals an.
23. Sei K ein Körper, der als Teilkörper in einem Körper L enthalten ist, und sei $l \in L$. Zeigen Sie: Die Menge aller Polynome $f \in K[x]$ mit $f(l) = 0$ ist ein Ideal von $K[x]$.
24. Seien f, g, h Polynome aus $K[x]$. Man sagt, dass h ein **Teiler** von f ist, falls es ein Polynom $q \in K[x]$ gibt mit $h \cdot q = f$. Das Polynom h wird ein **größter gemeinsamer Teiler** von f und g genannt, wenn die folgenden beiden Eigenschaften erfüllt sind:
 - h ist sowohl ein Teiler von f als auch ein Teiler von g .
 - Wenn ein Polynom t sowohl f als auch g teilt, so teilt t auch das Polynom h .
 Entwickeln Sie mit Hilfe der Polynomdivision einen Algorithmus („euklidischer Algorithmus“) zur Berechnung eines größten gemeinsamen Teilers zweier Polynome.

25. Zeigen Sie: Wenn f und g zwei Polynome aus $K[x]$ sind, so gilt

$$I(f, g) = I(h),$$

wobei h ein größter gemeinsamer Teiler von f und g ist.

26. Bestimmen Sie ein Polynom, welches das von $x^7 + x^5 + x^3 + 1$ und $x^3 + x + 1$ erzeugte Ideal von $\text{GF}(2)[x]$ erzeugt.
27. Bestimmen Sie ein Polynom, welches das von $x^2 - 1$, $x^3 + 2x^2 + 2x + 1$ und $x^4 + x^3 + x + 1$ erzeugte Ideal von $R[x]$ erzeugt.
28. Zeigen Sie: Wenn f und g zwei teilerfremde Polynome aus $K[x]$ sind (das bedeutet, dass f und g das Körperelement 1 als größten gemeinsamen Teiler haben), dann ist das von f und g erzeugte Ideal $I(f, g)$ gleich $K[x]$.

Projekte

Projekt A: Der Ring \mathbb{Z}

In diesem ersten Projekt dieses Kapitels sind Sie eingeladen, die bekannteste algebraische Struktur, nämlich den Ring \mathbb{Z} der ganzen Zahlen zu studieren. Sie sollen die Division mit Rest verstehen und damit beweisen, dass auch \mathbb{Z} ein **Hauptidealring** ist; das bedeutet, dass jedes Ideal von \mathbb{Z} ein Hauptideal ist.

Wir beginnen ganz behutsam. Seien a und b zwei ganze Zahlen. Wir sagen, dass a ein **Teiler** von b ist, wenn es eine ganze Zahl c gibt mit $ac = b$. Wir schreiben in diesem Fall $a \mid b$. Zum Beispiel gilt $6 \mid 24$, $-4 \mid 8$, $3 \mid -15$ und $2001 \mid 0$. (Beachten Sie, dass $c = 0$ erlaubt ist.)

1. Seien a , b , und b' ganze Zahlen. Zeigen Sie die folgenden Aussagen:

$$\begin{aligned} a \mid b \quad \text{und} \quad a \mid b' &\Rightarrow a \mid b + b', \\ a \mid b \quad \text{und} \quad a \mid b' &\Rightarrow a \mid b - b', \\ a \mid b &\Rightarrow a \mid b \cdot z \quad \text{für alle} \quad z \in \mathbb{Z} \end{aligned}$$

Insbesondere gilt also $a \parallel 0$ für alle $a \in \mathbb{Z}$.

2. Seien a und b ganze Zahlen. Zeigen Sie:

$$a \mid b \quad \text{und} \quad b \mid a \Rightarrow a = \pm b.$$

3. (**Division mit Rest**). Seien a und b ganze Zahlen mit $b > 0$. Zeigen Sie, dass es eindeutig bestimmte ganze Zahlen q und r gibt mit folgenden Eigenschaften:

$$a = b \cdot q + r \quad \text{und} \quad 0 \leq r < b.$$

[Sie müssen sowohl die Existenz als auch die Eindeutigkeit der Zahlen q und r nachweisen!]

Ein entscheidender Begriff ist der des größten gemeinsamen Teilers zweier ganzer Zahlen. Seien a und b zwei ganze Zahlen. Dann heißt eine ganze Zahl g ein **größter gemeinsamer Teiler** von a und b , falls g die folgenden Aussagen erfüllt:

- $g \geq 0$.
- g teilt a und g teilt b .
- Jede ganze Zahl, die sowohl a als auch b teilt, teilt auch g .

Zum *Beispiel* ist 6 ein größter gemeinsamer Teiler von 24 und 42. (Die ganzen Zahlen, die 24 und 42 teilen, sind $-6, -3, -2, -1, 1, 2, 3, 6$; jede dieser Zahlen teilt 6.)

Beachten Sie, dass wir derzeit noch überhaupt nicht wissen, ob je zwei ganze Zahlen einen größten gemeinsamen Teiler haben und ob ein solcher eindeutig bestimmt ist. In den nächsten Schritten werden wir uns davon überzeugen.

4. (**Eindeutigkeit des größten gemeinsamen Teilers**) Zeigen Sie: Je zwei ganze Zahlen haben höchstens einen größten gemeinsamen Teiler.
[Nehmen Sie an, dass g und g' größte gemeinsame Teiler der Zahlen a und b sind. Zeigen Sie $g \parallel g'$ und $g' \parallel g$, und benutzen Sie Teilschritt 2.]

Wenn zwei ganze Zahlen a und b den größten gemeinsamen Teiler g haben, so schreiben wir dafür auch $g = \text{ggT}(a, b)$.

5. Seien a und b ganze Zahlen mit $b > 0$. Zeigen Sie: Wenn q und r ganze Zahlen sind mit

$$a = b \cdot q + r \quad \text{und} \quad 0 \leq r < b,$$

so ist der größte gemeinsame Teiler von a und b auch der größte gemeinsame Teiler von b und r und umgekehrt.

6. Sei a eine natürliche Zahl. Zeigen Sie: Die Zahlen a und 0 haben genau einen größten gemeinsamen Teiler, nämlich a .
7. (**Euklidischer Algorithmus zur Berechnung des größten gemeinsamen Teilers**) Benutzen Sie den Teilschritt 5, um einen größten gemeinsamen Teiler zweier vorgegebener ganzer Zahlen a und b zu berechnen.
8. Berechnen Sie mit Hilfe dieses Verfahrens den größten gemeinsamen Teiler Ihrer Telefonnummer und Ihrer Matrikelnummer (alternativ: den größten gemeinsamen Teiler der alten und der neuen Postleitzahl Ihres Heimatorts.)
9. (**Lemma von Bézout**). Seien a und $n \geq 2$ ganze Zahlen. Zeigen Sie: Genau dann hat a eine **multiplikative Inverse modulo n** (das bedeutet, dass es eine ganze Zahl a' mit $aa' \bmod n = 1$ gibt), wenn a und n den größten gemeinsamen Teiler 1 haben. (Ganze Zahlen mit größtem gemeinsamen Teiler 1 nennt man auch **teilerfremd**.)
10. Sei n eine natürliche Zahl mit $n \geq 1$. Zeigen Sie: Die Menge \mathbb{Z}_n^* (das heißt die Menge aller invertierbaren Elemente des Rings \mathbb{Z}_n) besteht aus genau den natürlichen Zahlen zwischen 1 und n , die teilerfremd zu n sind.

11. Zeigen Sie (mit Hilfe der „Division mit Rest“), dass jedes Ideal von \mathbf{Z} ein Hauptideal ist.
12. Sei p eine Primzahl. Mit \mathbf{Q}_p bezeichnen wir die Menge derjenigen rationalen Zahlen, deren Nenner nicht durch p teilbar ist. Genauer gesagt gilt:

$$\mathbf{Q}_p = \{m/n \mid m, n \in \mathbf{Z}, n \neq 0, \text{ggT}(m, n) = 1, p \text{ teilt nicht } n\}.$$

Zeigen Sie:

- (a) \mathbf{Q}_p ist ein Ring.
 (b) Die Menge

$$I = \{m/n \mid m, n \in \mathbf{Z}, n \neq 0, \text{ggT}(m, n) = 1, p \text{ teilt nicht } n, p \text{ teilt } m\}$$

ist ein Ideal von \mathbf{Q}_p .

Projekt B: Der Ring $\mathbf{H}[x]$

Das Ziel dieses zweiten Projektes ist es zu untersuchen, inwieweit sich die Überlegungen dieses Kapitels auf Polynomringe $K[x]$ übertragen lassen, wenn K ein *Schiefkörper*, also ein *nichtkommutativer Körper* ist. Wir werden dies am Beispiel des Quaternionenschiefkörpers durchführen.

Nach diesen Einleitungsworten stellen Sie sich vermutlich auf eine äußerst stumpfsinnige Folge von Aufgaben ein und sind versucht, überhaupt nicht mehr weiter zu lesen. Tun Sie das nicht! Denn es kann sich nicht um eine langweilige wörtliche Übertragung der Aussagen und Beweise dieses Kapitels handeln. Warum? Weil hier eine Überraschung auf Sie wartet. Es ist nämlich so, dass der Satz, den Sie vermutlich schon in der Schule gelernt haben „Ein Polynom vom Grad n hat höchstens n Nullstellen“ hier falsch ist. Wie bitte? Ja, es gibt tatsächlich ein Polynom vom Grad 2, das mehr als 2, nämlich sogar 6 Nullstellen hat. Dies ist das harmlose Polynom $f = x^2 + 1 \in \mathbf{H}[x]$. Dieses Polynom hat nicht nur (wie in \mathbf{C}) die Nullstellen i und $-i$, sondern auch j und $-j$, sowie k und $-k$. Es ist sogar noch viel schlimmer: Dieses Polynom hat sogar unendlich viele Nullstellen (siehe Teilschritt 11).

Verrückt! Das muss irgendwie damit zu tun haben, dass \mathbf{H} nicht kommutativ ist. Aber woran liegt's? Wo steckt der Fehler? Welcher Schluss ist bei Polynomringen über nichtkommutativen Körpern nicht mehr richtig? Es ist ja keineswegs so, dass alles falsch ist – im Gegenteil: Das meiste ist richtig. Sie sollen im Folgenden herauspräparieren, was richtig ist und was nicht mehr gilt. Sie sollen also genau die Stelle finden, an der sich der Polynomring über \mathbf{H} von den Polynomringen über kommutativen Körpern unterscheidet.

Wir fangen ganz von vorne an. Die Definition von $\mathbf{H}[x]$, die Definition der Multiplikation von Polynomen und des Grads eines Polynoms bietet keine Überraschungen und läuft genau so wie im kommutativen Fall. Ferner ergibt sich genau so, dass $\mathbf{H}[x]$ ein Ring mit 1 ist – der allerdings nichtkommutativ ist, da ja schon die Teilmenge \mathbf{H} der konstanten Polynome nichtkommutativ ist.

1. Zeigen Sie, dass in $\mathbf{H}[x]$ die Gradformel gilt:

$$\text{Grad}(fg) = \text{Grad}(f) + \text{Grad}(g).$$

2. Zeigen Sie, dass für Polynome $f, g \in \mathbf{H}[x]$, $g \neq 0$ die Polynomdivision gilt: Das heißt: Es gibt eindeutig bestimmte Polynome $q, r \in \mathbf{H}[x]$ mit

$$f = qg + r \quad \text{und} \quad \text{Grad}(r) < \text{Grad}(g).$$

Die Definition eines Teilers können Sie vermutlich selbständig vornehmen: Wir sagen, ein Polynom $f \in \mathbf{H}[x]$ **teilt** ein Polynom $g \in \mathbf{H}[x]$, wenn es ein $f^* \in \mathbf{H}[x]$ gibt mit $f \cdot f^* = g$.

3. Zeigen Sie, dass $\mathbf{H}[x]$ ein Hauptidealring ist.
 4. Zeigen Sie: Ist a eine Nullstelle des Polynoms $f \in \mathbf{H}[x]$, so gibt es ein Polynom g mit $f = (x-a)g$.
 5. Zeigen Sie das Lemma von Bézout: Seien f und g Polynome aus $\mathbf{H}[x]$. Wenn f und g teilerfremd sind, so gibt es Polynome $f^*, g^* \in \mathbf{H}[x]$ mit $1 = f \cdot f^* + g \cdot g^*$.

Bis hierher ging alles gut! Das kann aber nur heißen, dass die kritische Stelle noch kommen muss. Wir rufen uns das Argument in Erinnerung, mit dem wir bewiesen haben, dass ein Polynom n -ten Grades höchstens n Nullstellen haben kann. Die entscheidende Überlegung bestand darin, dass sich ein Polynom $f \in K[x]$, das die Nullstellen a_1, a_2, \dots, a_s hat, als $f = (x-a_1)(x-a_2) \dots (x-a_s)g$ schreiben lässt – jedenfalls wenn K kommutativ ist! Dies liegt wiederum daran, dass die Zerlegung eines Polynoms in irreduzible Polynome im Wesentlichen eindeutig ist. Beide Tatsachen sind über Schiefkörpern falsch!

6. Machen Sie sich klar, dass sich das Polynom $x^2 + 1 \in \mathbf{H}[x]$ auf mindestens drei Weisen als Produkt (irreduzibler) Polynome vom Grad 1 schreiben lässt.
 7. Gilt in \mathbf{H} die binomische Formel (also das wohlbekannte $(a+b)^2 = a^2 + 2ab + b^2$)? [Probieren Sie für a und b die Einheiten i, j, k .]
 8. Überlegen Sie sich, dass für Polynome $f, g \in \mathbf{H}[x]$ die folgende Aussage im allgemeinen *nicht* gilt:
 Ist h irreduzibel und gilt $h \mid fg$, so gilt $h \mid f$ oder $h \mid g$.
 9. Analysieren Sie genau, welchen Schluss des Beweises des Satzes vom Einsetzungshomomorphismus man in $\mathbf{H}[x]$ *nicht* nachmachen kann.
 10. Zeigen Sie, dass das Polynom $f = x^2 + 1 \in \mathbf{H}[x]$ unendlich viele Nullstellen in \mathbf{H} hat.

Sie sollten mit folgenden Begriffen umgehen können

Ring, Einselement, kommutativer Ring, \mathbf{Z}_n , $K^{n \times n}$, x , $K[x]$, Polynom, Grad, Einsetzungshomomorphismus, Polynomdivision, Nullstelle, Ideal, Hauptideal, erzeugendes Element eines Ideals.



STATISTIKER LIEGEN MEIST DANESEN

Determinanten sind ein äußerst wichtiges Instrument zur Untersuchung von Matrizen und linearen Abbildungen. Vielleicht denken Sie, Determinanten seien einfach, weil Sie wissen, dass die Determinante einer 2×2 -Matrix gleich $ad - bc$ ist. Es ist jedoch so, dass man Determinanten auf zwei Weisen einführen kann, und beide sind ein Schock für die Studierenden, die den Determinanten zum ersten Mal Aug in Aug gegenüberstehen. Ich habe mich dafür entschieden, mit dem mathematisch gewichtigeren Schock zu beginnen.

Noch eine Vorbemerkung: Die Theorie der Determinanten ist zwar begrifflich nicht besonders tiefsinnig, aber ziemlich diffizil und teilweise auch technisch. Man muss sehr häufig sehr genau hinschauen; mitunter spielt sich die Mathematik sozusagen nur auf der Ebene der Indizes ab. Daher gibt es zwei Möglichkeiten, sich mit Determinanten zu beschäftigen. Sie können sich entweder die Sätze anschauen und versuchen, sich deren Bedeutung anhand von Beispielen klar zu machen. Oder Sie müssen sich mit Geduld und Scharfblick wappnen – und werden dann ohne weitere Schwierigkeiten auch die Details verstehen.

7.1 Die Determinantenfunktion

Sei K stets ein (kommutativer) Körper.

Eine Abbildung $\det: K^{n \times n} \rightarrow K$, die jeder $n \times n$ -Matrix M über K ein Element aus K zuordnet, das wir $\det(M)$ nennen, heißt eine **Determinantenfunktion**, falls sie die folgenden drei Eigenschaften hat:

(D1) Die Abbildung \det ist *linear* (das heißt *additiv* und *homogen*) in *jeder Zeile*. Das bedeutet: Für jedes $i \in \{1, \dots, n\}$ gilt

$$\det \begin{pmatrix} z_1 \\ \vdots \\ z_i + z_i' \\ \vdots \\ z_n \end{pmatrix} = \det \begin{pmatrix} z_1 \\ \vdots \\ z_i \\ \vdots \\ z_n \end{pmatrix} + \det \begin{pmatrix} z_1 \\ \vdots \\ z_i' \\ \vdots \\ z_n \end{pmatrix}$$

und

$$\det \begin{pmatrix} z_1 \\ \vdots \\ k \cdot z_i \\ \vdots \\ z_n \end{pmatrix} = k \cdot \det \begin{pmatrix} z_1 \\ \vdots \\ z_i \\ \vdots \\ z_n \end{pmatrix}.$$

(Dabei sind z_1, z_2, \dots, z_n die Zeilen einer $n \times n$ -Matrix.)

(D2) Ist $\text{Rang}(M) < n$, so ist $\det(M) = 0$.

(D3) $\det(E_n) = 1$.

Man nennt $\det(M)$ die **Determinante** von M .

Determinanten sind also nur für **quadratische** Matrizen, das heißt für Matrizen, bei denen die Anzahl der Zeilen gleich der Anzahl der Spalten ist, definiert.

Es erhebt sich sofort die Frage: Was soll das? Wie kommt man darauf? Was kann man damit anfangen?

All diese Fragen werden später in diesem Kapitel beantwortet werden. Unser erstes Ziel ist es, die *Existenz* einer solchen Abbildung nachzuweisen. Wir werden sogar zeigen, dass es für jedes n *genau eine* solche Abbildung gibt.

Als einführendes *Beispiel* machen wir uns folgendes klar:

Determinante einer 2×2 -Matrix

Die Abbildung \det , die auf der Menge der 2×2 -Matrizen definiert ist durch

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} := ad - bc,$$

ist eine Determinantenfunktion.

Beweis (D1): Für jedes $k \in K$ gilt

$$\det \begin{pmatrix} k \cdot a & k \cdot b \\ c & d \end{pmatrix} = kad - kbc = k(ad - bc) = k \cdot \det \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

und

$$\begin{aligned}\det \begin{pmatrix} a+a' & b+b' \\ c & d \end{pmatrix} &= (a+a')d - (b+b')c = ad - bc + a'd - b'c \\ &= \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} + \det \begin{pmatrix} a' & b' \\ c & d \end{pmatrix}.\end{aligned}$$

Entsprechend zeigt man die Linearität in der zweiten Zeile. Somit gilt (D1).

(D2) ist einfach: Wenn die Matrix nur aus Nullen besteht, erhält man als Ergebnis Null. Wenn der Rang 1 ist, so ist eine Zeile ein Vielfaches der anderen, und als Wert von \det ergibt sich $a \cdot kb - b \cdot ka = 0$.

Da sich (D3) durch bloßes Hinsehen ergibt, ist \det eine Determinantenfunktion. \square

Wir nähern uns den Determinantenfunktionen, indem wir uns klarmachen, wie sich die Determinante einer Matrix bei elementaren Zeilenumformungen ändert.

Invarianz der Determinantenfunktion gegenüber elementaren Zeilenumformungen

- (a) Verwandelt man M durch Multiplikation einer Zeile mit einem Element $k \in K$ in M' , so ist $\det(M') = k \cdot \det(M)$.
- (b) Verwandelt man die $n \times n$ -Matrix M durch Addition eines Vielfachen einer Zeile zu einer anderen Zeile aus M in die Matrix M' , so ist $\det(M') = \det(M)$.
- (c) Verwandelt man M durch Vertauschen zweier Zeilen in M' , so ist $\det(M') = -\det(M)$.

Bemerkung In der Sprache von Kap. 4 handelt es sich hierbei um „elementare Zeilenumformungen des Typs 2, 3, und 1“ (Vergleichen Sie dazu den Abschn. 4.1.4).

Beweis

- (a) ist nichts anderes als die Homogenität in der entsprechenden Zeile.
- (b) Es werde das k -fache der i -ten Zeile zur j -ten Zeile von M addiert, um M' zu erhalten. Sei M^* die Matrix, die aus M entsteht, indem die j -te Zeile durch das k -fache der i -ten ersetzt wird. Dann folgt aus (D1), dass $\det(M') = \det(M) + \det(M^*)$ gilt. Da in M^* die i -te und die j -te Zeile linear abhängig sind, folgt wegen (D2) $\det(M^*) = 0$. Mit (D1) ergibt sich also

$$\det(M') = \det(M) + \det(M^*) = \det(M).$$

- (c) Wir wenden nochmals einen ähnlichen Trick an: Die Matrix M' möge aus M durch Vertauschen der i -ten und j -ten Zeile hervorgegangen sein.

Wenn man in M die j -te Zeile zur i -ten addiert, erhält man eine Matrix M_i . Entsprechend erhält man eine Matrix M_j' , indem man in M' die j -te Zeile (also die i -te Zeile von M) zur i -ten Zeile addiert. Dann ist nach (b) $\det(M_i) = \det(M)$ und $\det(M_j') = \det(M')$. Die Matrizen M_i und M_j' unterscheiden sich nur in der j -ten Zeile, da in beiden Matrizen in der i -ten Zeile die Summe der i -ten und j -ten Zeile von M steht.

Wenn wir mit N diejenige Matrix bezeichnen, die aus M dadurch entsteht, dass wir sowohl die i -te als auch die j -te Zeile durch die Summe der i -ten und j -ten Zeile ersetzen, ergibt sich aufgrund von (D1)

$$\det(N) = \det(M_i) + \det(M_j') = \det(M) + \det(M') .$$

Da N zwei gleiche Zeilen besitzt, ist $\det(N) = 0$; daraus ergibt sich die Behauptung. \square

Nun beweisen wir die

Eindeutigkeit der Determinantenfunktion

Für jedes $n \in \mathbb{N}$ gibt es höchstens eine Determinantenfunktion.

Beweis Seien \det und \det' Abbildungen, die jeder $n \times n$ -Matrix ein Element aus K so zuordnen, dass (D1), (D2) und (D3) erfüllt ist. Wir müssen zeigen, dass für alle $n \times n$ -Matrizen M gilt

$$\det(M) = \det'(M) .$$

Dazu benötigen wir folgende Hilfsaussage: Wenn die $n \times n$ -Matrizen N und M durch elementare Zeilenumformungen auseinander hervorgehen, so gilt

$$\det(M) = \det'(M) \Rightarrow \det(N) = \det'(N) .$$

(Da sich elementare Zeilenumformungen durch elementare Zeilenumformungen rückgängig machen lassen, folgt dies aus der Invarianz gegenüber elementaren Zeilenumformungen.)

Sei M eine beliebige $n \times n$ -Matrix. Ist $\text{Rang}(M) < n$, so ist nach (D2) $\det(M) = 0 = \det'(M)$. Sei also $\text{Rang}(M) = n$. Dann kann man M durch elementare Zeilenumformungen in die Einheitsmatrix E_n verwandeln (vgl. dazu Übungsaufgabe 13 aus Kap. 4). Nach (D3) ist $\det(E_n) = 1 = \det'(E_n)$. Aufgrund der Hilfsaussage folgt also $\det(M) = \det'(M)$.

Somit ist $\det' = \det$. \square

Wir notieren ein wichtiges Korollar:

Determinante einer regulären Matrix

Sei M eine $n \times n$ -Matrix. Dann sind folgende Aussagen gleichwertig:

- (a) M ist regulär,
- (b) $\text{Rang}(M) = n$,
- (c) $\det(M) \neq 0$.

Beweis Im Grunde haben wir schon alles bewiesen (wie es bei einem „Korollar“ auch sein soll):

- Dass eine $n \times n$ -Matrix genau dann regulär (also invertierbar) ist, wenn sie den Rang n hat, wissen wir schon längst. Also gilt „(a) \Leftrightarrow (b)“.
- Die Richtung „(c) \Rightarrow (b)“ ist nichts anderes als die Forderung (D2).
- Die Umkehrung, also „(b) \Rightarrow (c)“ wurde in obigem Beweis mitbewiesen. □

Ich bin mir bewusst, dass ich Ihr (mathematisches) Hauptbedürfnis noch nicht befriedigt habe, nämlich die *Existenz* einer Determinantenfunktion zu zeigen. Dies soll jetzt geschehen, und zwar dadurch, dass wir uns eine kompliziert scheinende Formel für die Determinante klar machen. Diese Formel geht auf Gottfried Wilhelm Leibniz (1646–1716) zurück, der als Mathematiker und Philosoph gleichermaßen bedeutend war.

Bei diesem zweiten Schock spielen Permutationen eine entscheidende Rolle. Daher müssen wir uns zunächst mit einigen grundlegenden Eigenschaften von Permutationen vertraut machen.

7.2 Permutationen

Eine bijektive Abbildung einer endlichen Menge X in sich wird eine **Permutation** von X genannt. Wir werden stets voraussetzen, dass X mindestens zwei Elemente hat.

Anzahl aller Permutationen

Sei X eine endliche Menge mit n Elementen. Dann gibt es genau $n!$ Permutationen von X .

Beweis Sei $X = \{x_1, x_2, \dots, x_n\}$. Um eine Permutation π festzulegen, müssen wir die Bilder $\pi(x_1), \pi(x_2), \dots, \pi(x_n)$ festlegen. Für $\pi(x_1)$ haben wir n Möglichkeiten, nämlich alle Elemente von X . Wenn $\pi(x_1)$ festgelegt ist, gibt es für $\pi(x_2)$ noch $n-1$ Möglichkeiten, nämlich alle Elemente von X außer $\pi(x_1)$ usw. Für das Bild des vorletzten Elements x_{n-1} gibt es noch

zwei Möglichkeiten, nämlich alle Elemente von X außer $\pi(x_1), \pi(x_2), \dots, \pi(x_{n-2})$. Für das Bild des letzten Elements bleibt nur noch eine Möglichkeit. Also gibt es insgesamt

$$n \cdot (n-1) \cdot \dots \cdot 2 \cdot 1$$

Möglichkeiten. Diese Zahl wird mit $n!$ (sprich „ n Fakultät“) abgekürzt. Da es $n!$ Möglichkeiten gibt, eine Permutation von X festzulegen, gibt es auch genau $n!$ Permutationen von X . \square

Offenbar ist es für das Studium von Permutationen unerheblich, *welche* n -elementige Menge wir betrachten. Daher werden wir uns das Leben einfach machen und in der Regel die Menge $X = \{1, 2, \dots, n\}$ zugrunde legen. (Das heißt nur, dass wir statt x_i einfach i schreiben.)

Die Menge aller Permutationen der Menge $\{1, 2, \dots, n\}$ bezeichnet man mit S_n . Diese Struktur wird die **symmetrische Gruppe** auf n Elementen genannt. Das liegt daran, dass S_n bezüglich der Hintereinanderausführung von Permutationen (**Produkt** von Permutationen) eine „Gruppe“ bildet. Mehr dazu in Kap. 9.

Offenbar ist das Produkt zweier Permutationen wieder eine Permutation. Wenn wir das Produkt $\pi\psi$ zweier Permutationen π und ψ berechnen wollen, müssen wir daran denken, dass wir das Produkt *von rechts nach links lesen* („erst ψ , dann π “). Dies kommt daher, dass wir die Permutationen *von links schreiben* (wir schreiben $\pi(i)$ und nicht $i\pi$).

Wie kann man eine einzelne Permutation beschreiben? Dafür gibt es prinzipiell zwei Weisen. Zum einen kann man eine Permutation π – wie jede Abbildung – dadurch beschreiben, dass man das Bild jedes Elementes angibt. Bei Permutationen hat sich dazu folgende Schreibweise eingebürgert: In einem rechteckigen Schema mit zwei Zeilen und n Spalten (also einer $2 \times n$ -Matrix) schreibt man in die erste Zeile die Elemente von X (meist in der natürlichen Reihenfolge); in der zweiten Zeile steht unter jedem Element $i \in X$ das Bild $\pi(i)$.

Ein *Beispiel*: Die Schreibweise

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 4 & 5 & 6 & 1 & 2 & 8 & 7 \end{pmatrix}$$

bedeutet, dass die Permutation π das Element 1 auf 3, 2 auf 4, 3 auf 5, 4 auf 6, 5 auf 1, 6 auf 2, 7 auf 8 und 8 auf 7 abbildet.

Im Allgemeinen sieht dieses Verfahren also wie folgt aus:

$$\pi = \begin{pmatrix} 1 & 2 & \dots & i & \dots & n-1 & n \\ \pi(1) & \pi(2) & \dots & \pi(i) & \dots & \pi(n-1) & \pi(n) \end{pmatrix}.$$

Um die andere Darstellungsform einführen zu können, brauchen wir noch den Begriff einer zyklischen Permutation.

Sei π eine Permutation der Menge X . Wir sagen, dass π das Element $i \in X$ fest lässt, falls $\pi(i) = i$ ist. Man nennt i dann auch einen **Fixpunkt** von π .

Eine Permutation π von X wird ein **Zyklus** (oder **zyklische Permutation**) genannt, falls – grob gesprochen – die Elemente, die von π bewegt werden, zyklisch vertauscht werden. Genauer gesagt: Eine Permutation π heißt zyklisch, falls es ein $i \in X$ und eine natürliche Zahl k gibt, so dass die folgenden drei Bedingungen gelten:

- (1) $\pi^k(i) = i$,
- (2) die Elemente $i, \pi(i), \pi^2(i), \dots, \pi^{k-1}(i)$ sind paarweise verschieden,
- (3) jedes Element, das verschieden von $i, \pi(i), \pi^2(i), \dots, \pi^{k-1}(i), \pi^k(i) (= i)$ ist, wird von π fest gelassen.

Die kleinste natürliche Zahl k mit obiger Eigenschaft wird die **Länge** des Zyklus π genannt. Ein Zyklus der Länge k heißt auch k -Zyklus. Wir schreiben dann

$$\pi = (i \pi(i) \pi^2(i) \dots \pi^{k-1}(i)).$$

Kompliziert? Nein, ein *Beispiel* macht die Sache sofort klar: Betrachten wir folgenden Zyklus der Länge 5:

$$\pi = (3 \ 7 \ 5 \ 1 \ 2) \in \mathbf{S}_7.$$

Daraus lesen wir ab, dass die Permutation π die Elemente 3, 7, 5, 1, 2 zyklisch vertauscht und alle anderen Elemente (das heißt 4 und 6) fest lässt. Also gilt:

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 3 & 7 & 4 & 1 & 6 & 5 \end{pmatrix}.$$

Offenbar ist die Zykelschreibweise wesentlich effizienter als die zunächst eingeführte Beschreibung von Permutationen. Daher möchte man nicht nur eine zyklische, sondern *jede* Permutation ähnlich effizient schreiben. Dass dies möglich ist, drückt der folgende Satz aus.

Darstellung einer Permutation als Produkt disjunkter Zyklen

Jede Permutation kann als Produkt zyklischer Permutationen geschrieben werden, von denen keine zwei ein Element gemeinsam haben.

Das heißt: Zu jedem $\pi \in \mathbf{S}_n$ gibt es zyklische Permutationen $\zeta_1, \dots, \zeta_s \in \mathbf{S}_n$, so dass folgende Eigenschaften erfüllt sind:

- $\pi = \zeta_1 \cdot \zeta_2 \cdot \dots \cdot \zeta_s$,
- kein Element aus X , das als Komponente in ζ_i vorkommt, kommt in ζ_j vor ($i, j = 1, \dots, s, i \neq j$). (Das bedeutet: Wenn ein Element $x \in X$ in einem Zyklus ζ_i „vorkommt“, so wird x von jedem anderen Zyklus ζ_j ($j \neq i$) fest gelassen.)

Wir demonstrieren dies zunächst an einem *Beispiel*: Wir betrachten die Permutation

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 4 & 5 & 6 & 1 & 2 & 8 & 7 \end{pmatrix} \in \mathbf{S}_8.$$

Um die Zyklen zu erhalten, verfolgen wir den „Weg“ der einzelnen Elemente: Das Element 1 wird auf 3 abgebildet, 3 wird auf 5 und 5 wieder auf 1 abgebildet; der erste Zyklus ist also

$$\zeta_1 = (1\ 3\ 5).$$

Nun betrachten wir ein Element, das noch nicht auftauchte, zum Beispiel das Element 2. Dieses wird auf 4, 4 wird auf 6 und 6 wieder auf 2 abgebildet. Daher lautet der zweite Zyklus

$$\zeta_2 = (2\ 4\ 6).$$

Gibt es noch ein Element, das nicht erfasst wurde? Ja, das Element 7. Dieses wird mit 8 vertauscht; also lautet der dritte (und letzte) Zyklus

$$\zeta_3 = (7\ 8).$$

Die gesamte Permutation π lässt sich also schreiben als

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 4 & 5 & 6 & 1 & 2 & 8 & 7 \end{pmatrix} = (1\ 3\ 5)(2\ 4\ 6)(7\ 8) = \zeta_1 \cdot \zeta_2 \cdot \zeta_3.$$

Nach diesem Beispiel ist völlig klar, wie der *Beweis* verläuft: Man wählt sich ein Element i und schreibt dieses an den Anfang eines Klammersausdrucks; als nächstes schreibt man das Element $\pi(i)$ hin, dann $\pi^2(i)$, dann $\pi^3(i)$, usw. Dies macht man so lange, bis man das erste Mal eine natürliche Zahl k findet, für die $\pi^k(i) = i$ gilt. Dies schreibt man nicht mehr hin, sondern schließt die Klammer. Man hat damit den ersten Zyklus $(i\ \pi(i)\ \pi^2(i)\ \dots\ \pi^{k-1}(i))$ erhalten. Danach wählt man ein noch nicht betrachtetes Element j und konstruiert den j enthaltenden Zyklus usw. \square

Wenn ein Zyklus der Länge 1 auftritt, lassen wir diesen (ohne Gefahr eines Missverständnisses) meist weg. Statt $(1\ 2)(3\ 6)(4)(5)$ schreiben wir also einfach $(1\ 2)(3\ 6)$.

Zum Schluss dieses Abschnitts noch zwei *Bemerkungen*.

1. Man kann natürlich auch nichtdisjunkte Zyklen miteinander multiplizieren. Dabei muss man allerdings teuflisch aufpassen. Was ist $(1\ 2)(2\ 3)$? Wir überlegen dies für jedes Element einzeln.

Was passiert mit 1? Im rechten Zyklus wird 1 nicht bewegt, im linken auf 2 abgebildet.

Was geschieht mit 2? Dieses Element wird im ersten (rechten!) Zyklus auf 3 abgebildet, und mit der 3 tut sich dann gar nichts mehr.

Und worauf wird das Element 3 abgebildet? Im rechten Zyklus auf 2, und die 2 wird im linken Zyklus auf 1 abgebildet. Insgesamt ergibt sich also

$$(1\ 2)(2\ 3) = (1\ 2\ 3) .$$

Machen Sie sich klar, dass $(2\ 3)(1\ 2) = (1\ 3\ 2)$ gilt, und also $(1\ 2)(2\ 3) \neq (2\ 3)(1\ 2)$ ist. Berechnen Sie zur Übung

$$(1\ 2)(2\ 3)(3\ 4) \text{ und } (3\ 4)(2\ 3)(1\ 2) .$$

2. Ein Produkt aus paarweise disjunkten Zyklen ist schon deswegen besonders angenehm, weil man dabei die Reihenfolge der Zyklen vertauschen kann. (Klar: Wenn ein Element x in einem Zyklus bewegt wird, liegt das Bild in demselben Zyklus und kommt also in keinem anderen mehr vor. Also wird jedes Element nur ein einziges Mal „angefasst“. Daraus ergibt sich, dass die Reihenfolge der Zyklen keine Rolle spielt.)

7.3 Gerade und ungerade Permutationen

Für viele Zwecke, insbesondere für die Leibnizsche Determinantenformel ist es wichtig, die Einteilung aller Permutationen in so genannte „gerade“ und „ungerade“ Permutationen zu kennen. Diese erhält man mit Hilfe von Transpositionen.

Eine Permutation, die nur zwei Elemente vertauscht, heißt eine **Transposition**.

Zum Beispiel ist $\tau = (5\ 8)$ diejenige Transposition, die 5 mit 8 vertauscht, aber alle anderen Elemente fest lässt.

Eine Transposition, die zwei aufeinander folgende Elemente (also i und $i + 1$) vertauscht, heißt eine **Nachbartransposition**.

Darstellung von Permutationen durch Transpositionen

Man kann jede Permutation als Produkt von Transpositionen, ja sogar als Produkt von Nachbartranspositionen darstellen.

Beweis Da wir jede Permutation als Produkt von Zyklen darstellen können, müssen wir die Aussage nur für Zyklen beweisen. Wir reduzieren das, was wir zu zeigen haben, weiter: Der Zyklus $(x_1\ x_2\ x_3\ \dots\ x_k)$ lässt sich jedenfalls als Produkt von Transpositionen schreiben:

$$(x_1x_2x_3\ \dots\ x_k) = (x_1x_2)(x_2x_3)\ \dots\ (x_{k-2}x_{k-1})(x_{k-1}x_k) .$$

Beachten Sie, dass die Vereinbarung über die Schreibweise von Abbildungen impliziert, dass wir *Produkte von rechts nach links* lesen.

Es genügt also zu zeigen, dass jede Transposition ein Produkt von Nachbartranspositionen ist. Sei also $\tau = (i j)$ eine Transposition. O. B. d. A. ist $i < j$. Dann gilt

$$\tau = (i j) = (j-1 j)(j-2 j-1) \dots (i+1 i+2)(i i+1)(i+1 i+2) \dots (j-1 j) .$$

Um das zu sehen, muss man nur genau lesen: Das Element j wird schrittweise über $j-1, \dots, i+1$ zu i hingeführt und dann nicht mehr bewegt. Symmetrisch dazu wird i schrittweise über $i+1, \dots, j-1$ bis zu j geführt. Jedes Element k zwischen i und j wird einerseits (in der „rechten Hälfte“) auf $k+1$ abgebildet, um dann andererseits (in der „linken Hälfte“) wieder auf k zurückgeführt zu werden. \square

Als *Beispiel* betrachten wir die Permutation $\pi = (3 \ 5 \ 8)$. Diese lässt sich gemäß obiger Vorschrift wie folgt als Produkt von Transpositionen schreiben:

$$\pi = (3 \ 5 \ 8) = (3 \ 5)(5 \ 8) .$$

Die Transpositionen $(5 \ 8)$ und $(8 \ 3)$ lassen sich wie folgt als Produkt von Nachbartranspositionen schreiben:

$$(3 \ 5) = (4 \ 5)(3 \ 4)(4 \ 5) \text{ und } (5 \ 8) = (7 \ 8)(6 \ 7)(5 \ 6)(6 \ 7)(7 \ 8) .$$

Damit ergibt sich schließlich

$$\pi = (3 \ 5 \ 8) = (3 \ 5) \cdot (5 \ 8) = (4 \ 5)(3 \ 4)(4 \ 5) \cdot (7 \ 8)(6 \ 7)(5 \ 6)(6 \ 7)(7 \ 8) .$$

Als Korollar halten wir folgende Aussage fest.

Darstellung einer Transposition durch eine ungerade Anzahl von Nachbartranspositionen

Jede Transposition kann als Produkt einer ungeraden Anzahl von Nachbartranspositionen geschrieben werden.

In obigem Beweis steht, dass die Transposition $(i j)$ geschrieben werden kann als Produkt

$$(i j) = (j-1 j)(j-2 j-1) \dots (i+1 i+2)(i i+1)(i+1 i+2) \dots (j-1 j) .$$

Wie viele Nachbartranspositionen sind dies? (Solche Fragen gehören – für mich – zu den schwierigsten der Mathematik.) Relativ einfach ist zu sehen, dass es sich um eine ungerade Anzahl handeln muss, denn zu beiden Seiten der Nachbartransposition

$(i+1)$ steht dieselbe Menge von Nachbartranspositionen. Durch genauestes Hinschauen bekommt man auch die genaue Anzahl heraus: Es handelt sich um genau $2(j-i-1)+1$ Nachbartranspositionen. \square

Wenn Sie Schwierigkeiten mit dieser allgemeinen Darstellung haben, dann empfehle ich Ihnen, sich die Sache an einem Beispiel, etwa im Falle $\tau = (5\ 8)$ konkret aufzuschreiben; dann wird es Ihnen wie Schuppen von den Augen fallen.

Man kann eine gegebene Permutation auf viele Weisen als Produkt von Transpositionen oder Nachbartranspositionen schreiben; die Darstellung ist weit davon entfernt eindeutig zu sein. Nur eine kleine Tatsache ist dabei unveränderbar: Wenn Sie es schaffen, eine Permutation als Produkt einer *geraden* Anzahl von Transpositionen zu schreiben, werde sogar *ich* es nicht schaffen, diese als Produkt einer *ungeraden* Anzahl von Transpositionen zu schreiben. Dies müssen wir uns noch klar machen – aber dazu bedarf es einiger Vorbereitungen.

Für eine beliebige Permutation $\pi \in S_n$ nennt man ein Paar (i, j) von Elementen der Menge $X = \{1, 2, \dots, n\}$ mit $i < j$ aber $\pi(i) > \pi(j)$ einen **Fehlstand** von π . Ein Fehlstand von π ist also ein Zahlenpaar, dessen Ordnung durch π umgedreht wird. Die Anzahl der Fehlstände einer Permutation π bezeichnen wir mit $f(\pi)$.

Wir machen uns diesen Begriff an folgendem *Beispiel* klar: Sei

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix}.$$

Um die Anzahl der Fehlstände von π zu finden, müssen wir alle Paare (i, j) mit $i < j$ und $i, j \in \{1, 2, 3, 4\}$ betrachten:

(i, j)	(1, 2)	(1, 3)	(1, 4)	(2, 3)	(2, 4)	(3, 4)
$(\pi(i), \pi(j))$	(2, 4)	(2, 1)	(2, 3)	(4, 1)	(4, 3)	(1, 3)
$\pi(i) < \pi(j)$?	ja	nein	ja	nein	nein	ja

Jeder Nein-Eintrag in der letzten Zeile obiger Tabelle deutet auf einen Fehlstand hin; die Anzahl der Fehlstände von π ist also gleich 3.

Nun bestimmen wir die Anzahl von Fehlständen einer Transposition.

Fehlstände einer Transposition

Sei $\tau = (i\ j)$ eine Transposition aus S_n . Sei $i < j$. Dann gilt für die Anzahl $f(\tau)$ der Fehlstände von τ :

$$f(\tau) = 2(j - i - 1) + 1.$$

Insbesondere hat jede Transposition eine ungerade Anzahl von Fehlständen. Ferner hat jede Nachbartransposition genau einen Fehlstand.

Beweis Genau die folgenden Paare sind Fehlstände von τ :

$$(i, x) \text{ mit } i+1 \leq x \leq j-1; (y, j) \text{ mit } i+1 \leq y \leq j-1 \text{ und } (i, j).$$

Dies sind $2(j-i-1) + 1$ Fehlstände. Die Aussage über Nachbartranspositionen ergibt sich unmittelbar, da in diesem Fall $j = i+1$, also $j-i-1 = 0$ ist. \square

Eine Permutation aus S_n heißt **gerade**, falls die Anzahl ihrer Fehlstände gerade ist; sie heißt **ungerade**, wenn sie eine ungerade Anzahl von Fehlständen hat. Für ein $\pi \in S_n$ definieren wir:

$\text{sig}(\pi) = 1$, falls π eine gerade Permutation ist

$\text{sig}(\pi) = -1$, falls π eine ungerade Permutation ist.

Man nennt $\text{sig}(\pi)$ das **Signum** der Permutation π .

Das nächste Lemma ist einfach, aber entscheidend. Es beschreibt den Effekt, den die „Multiplikation“ einer beliebigen Permutation mit einer Nachbartransposition hat.

Effekt einer Nachbartransposition

Sei $\pi \in S_n$ eine beliebige Permutation, und sei τ eine Nachbartransposition aus S_n . Dann gilt

$$f(\tau\pi) = f(\pi) \pm 1.$$

Betrachten wir zum *Beispiel* $\tau = (4\ 5)$, $\pi = (3\ 4\ 5\ 8)$, also $\tau\pi = (4\ 5)(3\ 4\ 5\ 8) = (3\ 5\ 8)$. Dann ist $f(\pi) = 7$ und $f(\tau\pi) = 8$, denn die Fehlstände der Permutation π sind die Paare $(3, 8), (4, 8), (5, 6), (5, 7), (5, 8), (6, 8), (7, 8)$, und die Fehlstände von $\tau\pi = (3\ 5\ 8)$ sind die Paare $(3, 4), (3, 8), (4, 8), (5, 6), (5, 7), (5, 8), (6, 8)$ und $(7, 8)$.

Zum *Beweis* machen wir uns folgendes klar: Sei $\tau = (i\ i+1)$ eine Nachbartransposition. Wenn das Paar $(i, i+1)$ ein Fehlstand von π ist, dann hebt die Multiplikation von π mit τ diesen Fehlstand auf; also ist in diesem Fall $f(\tau\pi) = f(\pi) - 1$. Wenn $(i, i+1)$ aber kein Fehlstand von π ist, so kommt bei $\tau\pi$ außer den Fehlständen von π noch der Fehlstand $(i, i+1)$ hinzu; also ist $f(\tau\pi) = f(\pi) + 1$. \square

Die nächste Folgerung wird für die weiteren Überlegungen außerordentlich nützlich sein.

Satz über Nachbartranspositionen

Sei π eine Permutation. Dann sind die folgenden Aussagen äquivalent:

- (a) π ist eine gerade Permutation.
- (b) Jede Darstellung von π als ein Produkt von Nachbartranspositionen besteht aus einer geraden Anzahl von Nachbartranspositionen.

- (c) Es gibt eine Darstellung von π als Produkt einer geraden Anzahl von Nachbartranspositionen.

Beweis „ $(a) \Rightarrow (b)$ “: Sei π eine gerade Permutation. Angenommen, es gibt eine Darstellung von π als Produkt einer ungeraden Anzahl von Nachbartranspositionen. Wir fassen zwei Tatsachen zusammen:

- Da die Identität keinen Fehlstand hat, besitzt sie insbesondere eine gerade Anzahl von Fehlständen (0 ist eine gerade Zahl!). Also ist die Identität eine gerade Permutation.
- Wenn man eine gerade Permutation (also zum Beispiel die Identität) mit einer ungeraden Anzahl von Nachbartranspositionen multipliziert, erhält man eine ungerade Permutation. (Nach dem vorigen Satz ändert sich die Anzahl der Fehlstände bei Multiplikation mit zwei Nachbartranspositionen um -2 , 0 oder 2 .)

Aus diesen beiden Tatsachen ergibt sich, dass das Produkt π einer ungeraden Anzahl von Nachbartranspositionen eine ungerade Permutation ist, ein Widerspruch.

„ $(b) \Rightarrow (c)$ “ ist trivial. (Dieses Wort darf man nur dann verwenden, wenn die Sache wirklich sonnenklar ist; in diesem Fall ist dies so: Schauen Sie sich einfach die Aussagen (b) und (c) nochmals an!)

„ $(c) \Rightarrow (a)$ “: Sei π eine Permutation, die sich als Produkt einer geraden Anzahl von Nachbartranspositionen schreiben lässt. Wir schließen ähnlich wie im ersten Beweisteil:

Wenn man eine gerade Permutation (also zum Beispiel die Identität) mit einer geraden Anzahl von Nachbartranspositionen multipliziert, erhält man wieder eine gerade Permutation.

Daraus ergibt sich, dass das Produkt π einer geraden Anzahl von Nachbartranspositionen eine gerade Anzahl von Fehlständen hat, also eine gerade Permutation ist.

Damit ist der Satz bewiesen. \square

Satz über gerade Permutationen

Sei $\pi \in S_n$ eine beliebige Permutation. Dann sind folgende Aussagen gleichwertig:

- (a) π ist eine gerade Permutation,
- (b) π ist Produkt einer geraden Anzahl von Nachbartranspositionen,
- (c) π ist Produkt einer geraden Anzahl von Transpositionen.

Wegen des obigen Satzes über Nachbartranspositionen brauchen wir nur noch die Äquivalenz von (b) und (c) zu beweisen: Trivialerweise folgt (c) aus (b) .

„ $(c) \Rightarrow (b)$ “: Da jede Transposition nach dem Satz über die Darstellung einer Transposition durch eine ungerade Anzahl von Nachbartranspositionen als Produkt einer ungeraden

den Anzahl von Nachbartranspositionen dargestellt werden kann, ist jedes Produkt einer geraden Anzahl von Transpositionen auch ein Produkt einer geraden Anzahl von Nachbartranspositionen.

Damit ist bereits alles gezeigt. \square

Analog erhalten wir folgende Aussage.

Satz über ungerade Permutationen

Sei $\pi \in S_n$ eine beliebige Permutation. Dann sind folgende Aussagen gleichwertig:

- (a) π ist eine ungerade Permutation.
- (b) π ist Produkt einer ungeraden Anzahl von Nachbartranspositionen.
- (c) π ist Produkt einer ungeraden Anzahl von Transpositionen. \square

Wir machen uns die Aussage der beiden Sätze nochmals an einem *Beispiel* klar. Wenn eine Permutation Produkt von zum *Beispiel* 1024 Transpositionen ist, kann sie keinesfalls als Produkt von 2001 Transpositionen oder irgendeiner ungeraden Zahl von Transpositionen dargestellt werden.

Aus den obigen Sätzen ergibt sich auch folgende Aussage:

Signum der inversen Permutation

Für jede Permutation $\pi \in S_n$ gilt: π ist genau dann gerade, wenn π^{-1} gerade ist. Mit anderen Worten:

$$\text{sig}(\pi) = \text{sig}(\pi^{-1}).$$

Beweis Da die Identität $\text{id} = \pi \cdot \pi^{-1}$ gerade ist, müssen π und π^{-1} entweder beide Produkt einer geraden Anzahl oder beide Produkt einer ungeraden Anzahl von Transpositionen sein. \square

Wir bezeichnen die Menge aller geraden Permutationen aus S_n mit A_n ; diese Menge wird oft die **alternierende Gruppe** genannt. Dies liegt daran, dass A_n bezüglich der Hintereinanderausführung von Permutationen eine „Gruppe“ bildet (siehe Kap. 9). Wesentliche Eigenschaften von A_n werden in folgendem Satz zusammengefasst:

Satz über die alternierende Gruppe

Sei n eine natürliche Zahl. Dann gilt:

- (a) Das Produkt je zweier Permutationen aus A_n liegt wieder in A_n . Mit anderen Worten: A_n ist bezüglich der Hintereinanderausführung von Permutationen abgeschlossen.
- (b) Für jede Permutation $\pi \in S_n \setminus A_n$ gilt

$$S_n = A_n \cup \pi A_n ,$$

wobei $\pi A_n := \{\pi\alpha \mid \alpha \in A_n\}$ ist.

- (c) Ist $n \geq 2$, so gilt für die Mächtigkeit von A_n

$$|A_n| = \frac{n!}{2} .$$

Genau die Hälfte aller Permutationen ist also gerade.

Beweis

- (a) Nach dem Satz über gerade Permutationen ist A_n die Menge aller Permutationen, die Produkt einer geraden Anzahl von Transpositionen sind. Es ist klar, dass das Produkt zweier solcher Permutationen wieder ein Produkt einer geraden Anzahl von Transpositionen ist, also in A_n liegt.
- (b) Da die Inklusion „ $A_n \cup \pi A_n \subseteq S_n$ “ trivial ist, ist nur zu zeigen, dass eine beliebige Permutation $\psi \in S_n$ in A_n oder in πA_n liegt. Wir können o. B. d. A. annehmen, dass ψ nicht in A_n liegt. Also ist ψ eine ungerade Permutation. Da auch π (und damit π^{-1}) eine ungerade Permutation ist, ist $\pi^{-1}\psi$ ein Produkt zweier Permutationen, die beide ein Produkt einer ungeraden Anzahl von Transpositionen sind. Also ist $\pi^{-1}\psi$ ein Produkt einer geraden Anzahl von Transpositionen. Daher liegt $\pi^{-1}\psi$ in A_n . Somit gibt es ein $\alpha \in A_n$ mit

$$\pi^{-1}\psi = \alpha .$$

Es folgt

$$\psi = \pi\alpha \in \pi A_n .$$

- (c) Offenbar haben A_n und πA_n gleich viele Elemente. (Man macht sich nämlich leicht klar, dass die Abbildung $\alpha \rightarrow \pi\alpha$ bijektiv ist.) Da A_n aus den geraden und πA_n aus ungeraden Permutationen besteht, sind A_n und πA_n disjunkt. Da nach (b) jede Permutation in A_n oder πA_n liegt und die Gesamtzahl der Permutationen gleich $n!$ ist, folgt die Behauptung. \square

Wir bemerken, dass für jede ungerade Permutation π aus S_n gilt

$$\pi A_n = A_n \pi ;$$

denn sowohl πA_n als auch $A_n \pi$ bestehen aus ungeraden Permutationen (sind also Teilmengen von $S_n \setminus A_n$) und beide Mengen haben genau $|A_n|$ viele Elemente.

Also gilt auch

$$S_n = A_n \cup A_n \pi$$

für jede ungerade Permutation π aus S_n .

Eine Aussage, die im Beweis von Teil (b) des Satzes über die alternierende Gruppe implizit vorkam, heben wir explizit hervor:

Signumsformel

Für je zwei Permutationen $\pi, \psi \in S_n$ gilt

$$\text{sig}(\pi) \cdot \text{sig}(\psi) = \text{sig}(\pi\psi) .$$

Der *Beweis* ergibt sich unmittelbar aus der Definition des Signums: Eine Permutation hat genau dann Signum 1 (und sonst -1), wenn sie gerade, wenn sie also Produkt einer geraden Anzahl von Transpositionen ist. \square

7.4 Die Leibnizsche Determinantenformel

Nun zur angekündigten Formel – die den zweiten Schock dieses Kapitels darstellt:

Leibnizsche Determinantenformel

Sei n eine natürliche Zahl. Durch die folgende Vorschrift \det wird eine Determinantenfunktion definiert. Sei $M = (m_{ij})_{1 \leq i, j \leq n}$ eine $n \times n$ -Matrix mit Elementen aus K . Dann ist

$$\det(M) = \sum_{\pi \in S_n} \text{sig}(\pi) m_{1,\pi(1)} \cdot m_{2,\pi(2)} \cdot \dots \cdot m_{n,\pi(n)} .$$

Auf den ersten Blick scheint es, als ob hier die mathematische Vernunft ein Monster geboren hätte. Die folgenden Spezialfälle sollen Ihnen vor Augen führen, dass dem nicht so ist. Durch die Behandlung dieser Spezialfälle werden Sie freundlichst eingeladen, die Formel wenigstens zu *lesen*.

1. Auch gemäß der Leibnizschen Formel ist $\det(E_n) = 1$. Mit anderen Worten: die Leibnizsche Funktion genügt jedenfalls der Bedingung (D3).

Warum? Theoretisch müssen wir für alle Permutationen die Produkte ausrechnen und diese aufsummieren. In Wirklichkeit müssen wir dafür aber nur diejenigen Permutationen π berücksichtigen, für die

$$\text{sig}(\pi) m_{1,\pi(1)} m_{2,\pi(2)} \cdots m_{n,\pi(n)}$$

verschieden von Null ist. Dazu müssen all die Elemente

$$m_{1,\pi(1)}, m_{2,\pi(2)}, \dots, m_{n,\pi(n)}$$

verschieden von Null sein. Da M die Einheitsmatrix ist, folgt also

$$\pi(1) = 1, \pi(2) = 2, \dots, \pi(n) = n;$$

also ist π die Identität ε . Somit ist

$$\det(E_n) = \text{sig}(\varepsilon) \cdot m_{11} \cdot m_{22} \cdots m_{nn} = 1.$$

□

2. Die Determinante einer 2×2 -Matrix

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

berechnet sich als

$$\det(M) = ad - bc.$$

Beweis. Es gibt genau zwei Permutationen der Menge $\{1, 2\}$, nämlich die Identität ε und die Transposition $\tau = (1\ 2)$. Daher ergibt sich gemäß der Leibnizschen Formel, wenn wir kurzfristig die Matrix M als

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}$$

schreiben:

$$\begin{aligned} \det(M) &= \text{sig}(\varepsilon) m_{1,\varepsilon(1)} m_{2,\varepsilon(2)} + \text{sig}(\tau) m_{1,\tau(1)} m_{2,\tau(2)} = 1 \cdot m_{11} m_{22} + (-1) \cdot m_{12} m_{21} \\ &= ad - bc. \end{aligned}$$

□

3. Regel von Sarrus. Ist $M = (m_{ij})$ eine 3×3 -Matrix, so gilt

$$\begin{aligned} \det(M) &= \det \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{pmatrix} \\ &= m_{11} m_{22} m_{33} + m_{12} m_{23} m_{31} + m_{13} m_{21} m_{32} \\ &\quad - m_{31} m_{22} m_{13} - m_{32} m_{23} m_{11} - m_{33} m_{21} m_{12}. \end{aligned}$$

(Man muss die drei „Diagonalen“ in Richtung der Hauptdiagonalen mit Plus und die drei „Diagonalen“ in Richtung der Nebendiagonalen mit Minus versehen.)

Beweis. Dies können wir wie folgt einsehen: Es gibt sechs Permutationen der Menge $\{1, 2, 3\}$, nämlich die Identität ϵ , die beiden 3-Zyklen $(1\ 2\ 3)$ und $(1\ 3\ 2)$ sowie die drei Transpositionen $(1\ 2)$, $(2\ 3)$ und $(1\ 3)$. Also kann man die Determinante der 3×3 -Matrix M gemäß der Leibnizschen Formel wie folgt ausrechnen:

$$\begin{aligned} \det(M) &= \det \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{pmatrix} \\ &= \text{sig}(\epsilon) \cdot m_{11}m_{22}m_{33} + \text{sig}(1\ 2\ 3) \cdot m_{12}m_{23}m_{31} + \text{sig}(1\ 3\ 2) \cdot m_{13}m_{21}m_{32} \\ &\quad + \text{sig}(1\ 2) \cdot m_{12}m_{21}m_{33} + \text{sig}(2\ 3) \cdot m_{11}m_{23}m_{32} + \text{sig}(1\ 3) \cdot m_{13}m_{22}m_{31} . \quad \square \end{aligned}$$

Diese Regel ist nach dem französischen Mathematiker Pierre-Frédéric Sarrus (1798–1858) benannt.

Schließlich halten wir noch die wichtigste Folgerung aus der Leibnizschen Formel fest:

4. Für jede natürliche Zahl n gibt es genau eine Determinantenfunktion. Jede Determinantenfunktion lässt sich mit der Leibnizschen Formel ausdrücken.

Beweis. Nach dem Satz über die Eindeutigkeit der Determinantenfunktion gibt es für jedes n höchstens eine Determinantenfunktion. Der obige Satz sagt, dass die Leibnizsche Formel eine Determinantenfunktion beschreibt. Also wird durch die Leibnizsche Formel jede Determinantenfunktion beschrieben. \square

Um nun aber endlich zu *beweisen*, dass die Leibnizsche Formel tatsächlich eine Determinantenfunktion definiert, gehen wir wie folgt vor. Wir definieren für jede $n \times n$ -Matrix M :

$$L(M) = \sum_{\pi \in S_n} \text{sig}(\pi) m_{1,\pi(1)} \cdot m_{2,\pi(2)} \cdots m_{n,\pi(n)} .$$

Dann ist zu zeigen, dass die Abbildung L die Eigenschaften (D1), (D2) und (D3) hat. Dann ist L eine Determinantenfunktion, und nach dem Eindeutigkeitsatz folgt also $L = \det$.

(D1) Sei $M = (m_{ij}) \in K^{n \times n}$, und sei $k \in K$. Dann gilt für jede Permutation $\pi \in S_n$:

$$\begin{aligned} &m_{1,\pi(1)} \cdots m_{i-1,\pi(i-1)} \cdot (m_{i,\pi(i)} + m'_{i,\pi(i)}) \cdot m_{i+1,\pi(i+1)} \cdots m_{n,\pi(n)} \\ &= m_{1,\pi(1)} \cdots m_{i-1,\pi(i-1)} \cdot m_{i,\pi(i)} \cdots m_{n,\pi(n)} \\ &\quad + m_{1,\pi(1)} \cdots m_{i-1,\pi(i-1)} \cdot m'_{i,\pi(i)} \cdots m_{n,\pi(n)} . \end{aligned}$$

Damit folgt die *Additivität* auch für die Summe dieser Ausdrücke über alle Permutationen. Das bedeutet, dass die Abbildung L in der i -ten Zeile additiv ist ($i = 1, \dots, n$).

Ähnlich zeigt man die *Homogenität*:

$$\begin{aligned} & m_{1,\pi(1)} \cdot \dots \cdot m_{i-1,\pi(i-1)} \cdot k \cdot m_{i,\pi(i)} \cdot m_{i+1,\pi(i+1)} \cdot \dots \cdot m_{n,\pi(n)} \\ &= k \cdot m_{1,\pi(1)} \cdot \dots \cdot m_{i-1,\pi(i-1)} \cdot m_{i,\pi(i)} \cdot m_{i+1,\pi(i+1)} \cdot \dots \cdot m_{n,\pi(n)} . \end{aligned}$$

Also ist L auch in der i -ten Zeile homogen.

(D2) Sei $\text{Rang}(M) < n$. Dann sind die Zeilen von M linear abhängig. Wir versuchen zunächst, das Problem zu reduzieren, das heißt auf einfachere Matrizen zurückzuführen.

Sei o. B. d. A. die erste Zeile z_1 eine Linearkombination der anderen: $z_1 = \sum_{i=2}^n k_i z_i$

Aufgrund der schon bewiesenen Eigenschaft (D1) gilt dann

$$L \begin{pmatrix} \sum_{i=2}^n k_i z_i \\ z_2 \\ z_3 \\ \vdots \\ z_n \end{pmatrix} = k_2 \cdot L \begin{pmatrix} z_2 \\ z_2 \\ z_3 \\ \vdots \\ z_n \end{pmatrix} + k_3 \cdot L \begin{pmatrix} z_3 \\ z_2 \\ z_3 \\ \vdots \\ z_n \end{pmatrix} + \dots + k_n \cdot L \begin{pmatrix} z_n \\ z_2 \\ z_3 \\ \vdots \\ z_n \end{pmatrix} .$$

Auf der rechten Seite wird die Abbildung L jeweils auf Matrizen angewandt, die zwei gleiche Zeilen haben. Wenn wir also zeigen können, dass L jede Matrix mit zwei gleichen Zeilen zu Null macht, haben wir unser Ziel auch erreicht.

Genau dies ist jetzt unsere Behauptung: *Ist M eine Matrix, die zwei gleiche Zeilen hat, so ist $L(M) = 0$.* Zum Nachweis dieser Behauptung können wir o. B. d. A. voraussetzen, dass die *beiden ersten Zeilen von M gleich* sind.

Sei $\tau = (1\ 2)$. Da nach dem Satz über die alternierende Gruppe jede Permutation aus S_n entweder gerade ist oder die Form $\alpha\tau$ hat, wobei α eine gerade Permutation ist, gilt $S_n = A_n \cup A_n\tau$. Also berechnet sich $L(M)$ wie folgt:

$$\begin{aligned} L(M) &= \sum_{\pi \in S_n} \text{sig}(\pi) m_{1,\pi(1)} \cdot m_{2,\pi(2)} \cdot \dots \cdot m_{n,\pi(n)} \\ &= \sum_{\alpha \in A_n} (\text{sig}(\alpha) m_{1,\alpha(1)} \cdot m_{2,\alpha(2)} \cdot \dots \cdot m_{n,\alpha(n)} \\ &\quad + \text{sig}(\alpha\tau) m_{1,\alpha\tau(1)} \cdot m_{2,\alpha\tau(2)} \cdot \dots \cdot m_{n,\alpha\tau(n)}) \\ &= \sum_{\alpha \in A_n} (m_{1,\alpha(1)} \cdot m_{2,\alpha(2)} \cdot \dots \cdot m_{n,\alpha(n)} - m_{1,\alpha(1)} \cdot m_{2,\alpha(2)} \cdot \dots \cdot m_{n,\alpha(n)}) \\ &= \sum_{\alpha \in A_n} (m_{1,\alpha(1)} \cdot m_{2,\alpha(2)} - m_{1,\alpha(2)} \cdot m_{2,\alpha(1)}) m_{3,\alpha(3)} \cdot \dots \cdot m_{n,\alpha(n)} = 0 , \end{aligned}$$

denn aus der Gleichheit der ersten beiden Zeilen von M ergibt sich insbesondere die Gleichheit der beiden Summanden in der Klammer.

Damit ist die Behauptung bewiesen, und also gilt (D2).

Da wir (D3) schon unmittelbar nach der Präsentation der Leibnizschen Formel gezeigt haben, folgt $L = \det$, und der Satz ist vollständig bewiesen. \square

Der Vorteil ist immens: Wann immer wir eine Determinantenfunktion konkret berechnen wollen, können wir uns der Leibnizschen Formel bedienen.

Übrigens: Lassen Sie sich von dem (auf den ersten Blick) fürchterlichen Aussehen dieser (und anderer) Formeln nicht schrecken! Was für uns Menschen auf den ersten Blick so kompliziert aussieht, dass man schon gar nicht anfängt zu lesen, ist oft etwas ganz Konkretes, eine Formel oder eine Vorschrift, die ganz einfach handzuhaben ist und das gewünschte Ergebnis mittels rein mechanischer Operationen liefert. Also: Haben Sie den Mut, solche Formelungetüme (langsam ... und mit Genuss!) zu lesen.

7.5 Wie berechnet man eine Determinante?

In den meisten Fällen empfehlen sich elementare Zeilenumformungen als *die* Methode zur Berechnung der Determinante einer Matrix. Dabei geht man in zwei Schritten vor: Zunächst bringt man eine gegebene Matrix auf Dreiecksgestalt und bestimmt danach die Determinante der Dreiecksmatrix.

Wir tun den zweiten Schritt vor dem ersten.

Determinante einer Dreiecksmatrix

Sei $M \in K^{n \times n}$ eine Matrix der folgenden Gestalt:

$$M = \begin{pmatrix} m_{11} & m_{12} & \cdot s & \cdot s & m_{1n} \\ 0 & m_{22} & & & m_{2n} \\ 0 & 0 & m_{33} & & \vdots \\ 0 & & & \ddots & m_{n-1,n} \\ 0 & 0 & \cdot s & 0 & m_{nn} \end{pmatrix}.$$

(Man nennt M dann auch eine **obere Dreiecksmatrix**.) Dann gilt:

$$\det(M) = m_{11} \cdot m_{22} \cdot \dots \cdot m_{nn}.$$

Kurz: Die Determinante einer Dreiecksmatrix ist das Produkt ihrer Diagonalelemente.

Dies gilt insbesondere für Diagonalmatrizen.

Beweis In der Leibnizschen Formel leisten nur solche Summanden

$$\text{sig}(\pi) \cdot m_{1,\pi(1)} \cdot m_{2,\pi(2)} \cdot \dots \cdot m_{n,\pi(n)}$$

einen Beitrag, für die alle $m_{i,\pi(i)}$ ungleich Null sind.

Da M eine Dreiecksmatrix ist, ist ein Eintrag m_{ij} sicher dann gleich Null, wenn $i > j$ ist (denn dann steht m_{ij} unterhalb der Hauptdiagonale).

Das heißt: Wir brauchen nur die m_{ij} mit $i \leq j$ zu betrachten, also nur diejenigen $m_i, \pi(i)$ mit $i \leq \pi(i)$.

Da dies für alle i gelten muss, heißt dies: Wir brauchen nur solche Permutationen π zu betrachten, für die $1 \leq \pi(1), 2 \leq \pi(2), \dots$ und $n \leq \pi(n)$ gilt.

Welches sind diese Permutationen? Behauptung: Diese Eigenschaft hat nur die Identität! (Wäre π nicht die Identität, so gäbe es ein i mit $i < \pi(i)$. Sei i die kleinste solche Zahl. Sei j das Element mit $\pi(j) = i$. Dann ist $j \neq 1, 2, \dots, i-1$ (sonst wäre ja $\pi(j) = j \neq i$). Ferner ist auch $\pi(j) \neq \pi(i)$ (sonst wäre $\pi(i) = \pi(j) = i$). Also muss $j > i$ sein, und es wäre also $j > i = \pi(j)$, ein Widerspruch zur Voraussetzung.)

Daher zählt bei der Determinantenbildung einer Dreiecksmatrix nur die Identität ϵ , und es folgt also

$$\begin{aligned} \det(M) &= \sum_{\pi \in S_n} \text{sig}(\pi) m_{1,\pi(1)} \cdot m_{2,\pi(2)} \cdot \dots \cdot m_{n,\pi(n)} \\ &= \text{sig}(\epsilon) \cdot m_{11} \cdot m_{22} \cdot \dots \cdot m_{nn} = m_{11} \cdot m_{22} \cdot \dots \cdot m_{nn} . \end{aligned}$$

□

Nun können wir die Determinante einer beliebigen Matrix auf die Determinante einer Dreiecksmatrix zurückführen.

Reduktion der Determinante auf eine Diagonalmatrix

Sei M eine beliebige Matrix aus $K^{n \times n}$. Dann kann man M allein durch Anwendung von elementaren Zeilenumformungen vom Typ 3 (das heißt Addition eines Vielfachen einer Zeile zu einer anderen) in eine Matrix verwandeln, bei der jedes Element unterhalb der Hauptdiagonalen Null ist (**Dreiecksmatrix**). Bei diesem Vorgang ändert sich die Determinante nicht.

Wenn $\det(M) \neq 0$ ist, so kann man die Matrix M durch Anwendung von elementaren Zeilenumformungen vom Typ 3 sogar in eine Diagonalmatrix verwandeln, die dieselbe Determinante wie M hat.

Beweis Das erste Ziel ist, M in eine Dreiecksmatrix zu verwandeln, also eine Matrix, bei der unterhalb der Hauptdiagonalen nur Nullen stehen.

Dies zeigen wir induktiv. Da die Behauptung für 1×1 -Matrizen trivial ist, können wir annehmen, dass $n > 1$ ist, und die Behauptung für $n-1$ richtig ist.

Wir zeigen jetzt, dass man durch elementare Zeilenumformungen vom Typ 3 die erste Spalte so gestalten kann, dass jedes Element, das verschieden vom ersten ist, gleich Null ist: Wenn alle Elemente der ersten Spalte gleich Null sind, sind wir fertig. Wenn andererseits irgendein Element der ersten Spalte ungleich Null ist, so kann man durch eine geeignete

Operation des Typs 3 ein von Null verschiedenes Element an die erste Stelle bringen. Indem man zu den Zeilen $2, \dots, n$ ein geeignetes Vielfaches der ersten Zeile addiert, erreicht man, dass in der Tat alle Elemente der ersten Spalte – mit Ausnahme des ersten – gleich Null sind.

Wir haben also eine Matrix M_1 folgender Gestalt konstruiert:

$$M_1 = \begin{pmatrix} m_{11} & * & \cdots s & * \\ 0 & & & \\ 0 & & M_1' & \\ 0 & & & \end{pmatrix}.$$

Nun wenden wir Induktion auf die $(n-1) \times (n-1)$ -Matrix M_1' an: Durch elementare Umformungen des Typs 3 der Zeilen $2, \dots, n$ kann man M_1' in eine Dreiecksmatrix verwandeln.

Damit haben wir das erste Ziel erreicht. Wir haben eine Matrix M_2 der folgenden Gestalt erhalten:

$$M_2 = \begin{pmatrix} m_{11} & m_{12} & \cdots s & \cdots s & m_{1n} \\ 0 & m_{22} & & & m_{2n} \\ 0 & 0 & m_{33} & & \vdots \\ 0 & & & \ddots & m_{n-1,n} \\ 0 & 0 & \cdots s & 0 & m_{nn} \end{pmatrix}.$$

Aber nun ist auch das zweite Ziel in greifbarer Nähe: Wir behaupten: Wenn $\det(M) \neq 0$ ist, so kann man M_2 durch elementare Zeilenumformungen des Typs 3 in eine Diagonalmatrix verwandeln.

Dazu gehen wir die Spalten von vorne durch. In der ersten Spalte müssen wir offenbar nichts mehr ändern. Da $\det(M_2) = \det(M) \neq 0$ ist, muss auch $m_{22} \neq 0$ sein. Nun addieren wir ein geeignetes Vielfaches der zweiten Zeile zur ersten und erreichen damit, dass an der Stelle $(1, 2)$ eine Null steht.

Jetzt wissen Sie bestimmt, wie's weitergeht. Daher führen wir den Rest auch gar nicht mehr aus, sondern schreiben guten Gewissens ein „usw.“

Damit sind wir fertig. Denn wir müssen nur noch festhalten, dass sich nach dem Satz über die Invarianz der Determinantenfunktion gegenüber elementaren Zeilenumformungen bei elementaren Zeilenumformungen des Typs 3 die Determinante nicht ändert. \square

In vielen Fällen wird die „Entwicklung nach einer Zeile oder Spalte“ angewandt. Für eine $n \times n$ -Matrix M bezeichnen wir mit M_{ij} diejenige $(n-1) \times (n-1)$ -Matrix, die dadurch entsteht, dass man aus M die i -te Zeile und die j -te Spalte streicht. Wenn zum Beispiel

$$M = \begin{pmatrix} 1 & 2 & 5 \\ 3 & 0 & -1 \\ 2 & 4 & 9 \end{pmatrix}$$

ist, so gilt etwa

$$M_{11} = \begin{pmatrix} 0 & -1 \\ 4 & 9 \end{pmatrix}, \quad M_{31} = \begin{pmatrix} 2 & 5 \\ 0 & -1 \end{pmatrix} \quad \text{und} \quad M_{22} = \begin{pmatrix} 1 & 5 \\ 2 & 9 \end{pmatrix}.$$

Die Idee besteht jetzt darin, die Berechnung der Determinante einer „großen“ $n \times n$ -Matrix M auf die Berechnung der Determinanten der „kleineren“ $(n-1) \times (n-1)$ -Matrizen M_{ij} zurückzuführen. Genauer gesagt gilt der folgende Satz:

Entwicklung nach einer Zeile

Sei $M = (m_{ij})_{1 \leq i, j \leq n}$ eine Matrix aus $K^{n \times n}$. Dann gilt für alle $i \in \{1, 2, \dots, n\}$:

$$\det(M) = \sum_{j=1}^n (-1)^{i+j} m_{ij} \cdot \det(M_{ij})$$

Mit anderen Worten: Um die Determinante zu berechnen, geht man die i -te Zeile Element für Element durch, multipliziert jeweils ein Element mit der entsprechenden Unterdeterminante und addiert diese Produkte mit alternierendem Vorzeichen.

Wir beweisen hier die Entwicklung nach einer Zeile nur im (technisch) einfachsten Fall, nämlich die Entwicklung nach der ersten Zeile. Diese lautet:

Entwicklung nach der ersten Zeile

Sei $M = (m_{ij})_{1 \leq i, j \leq n}$ eine Matrix aus $K^{n \times n}$. Dann gilt:

$$\det(M) = \sum_{j=1}^n (-1)^{1+j} m_{1j} \cdot \det(M_{1j}).$$

Mit anderen Worten: Um die Determinante zu berechnen, geht man die erste Zeile Element für Element durch, multipliziert jeweils ein Element mit der entsprechenden Unterdeterminante und addiert diese Produkte mit alternierendem Vorzeichen, wobei man mit + beginnt.

Vor dem Beweis berechnen wir als *Beispiel* die Determinante der obigen Matrix:

$$\begin{aligned} \det(M) &= \begin{vmatrix} 1 & 2 & 5 \\ 3 & 0 & -1 \\ 2 & 4 & 9 \end{vmatrix} = 1 \cdot \det \begin{pmatrix} 0 & -1 \\ 4 & 9 \end{pmatrix} - 2 \cdot \det \begin{pmatrix} 3 & -1 \\ 2 & 9 \end{pmatrix} + 5 \cdot \det \begin{pmatrix} 3 & 0 \\ 2 & 4 \end{pmatrix} \\ &= 1 \cdot 4 - 2 \cdot 29 + 5 \cdot 12 = 6. \end{aligned}$$

Der Beweis erfolgt durch Induktion nach n .

Im Fall $n = 1$ ist die Aussage trivial, und auch im Falle $n = 2$ kann man sich von der Richtigkeit der Aussage leicht überzeugen (siehe Übungsaufgabe 7).

Sei nun $n > 1$ und die Aussage richtig für $n-1$. Wir zeigen, dass die Abbildung

$$\Delta : M \rightarrow \sum_{j=1}^n (-1)^{1+j} m_{1j} \cdot \det(M_{1j})$$

die Bedingungen (D1), (D2) und (D3) erfüllt, also eine (und damit *die*) Determinantenfunktion ist. (Beachten Sie, dass die Funktion \det auf der rechten Seite auf Matrizen mit $n-1$ Zeilen und Spalten operiert.)

(D1) Wir zeigen nur die *Additivität*. Sei $z_i' = (m_{i1}', \dots, m_{in}')$ eine weitere „ i -te Zeile“ von M .

1. Fall: $i = 1$. Dann folgt

$$\begin{aligned} \Delta \begin{pmatrix} z_1 + z_1' \\ z_2 \\ \vdots \\ z_n \end{pmatrix} &= \sum_{j=1}^n (-1)^{1+j} (m_{1j} + m_{1j}') \cdot \det(M_{1j}) \\ &= \sum_{j=1}^n (-1)^{1+j} m_{1j} \cdot \det(M_{1j}) + \sum_{j=1}^n (-1)^{1+j} m_{1j}' \cdot \det(M_{1j}) \\ &= \Delta \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix} + \Delta \begin{pmatrix} z_1' \\ z_2 \\ \vdots \\ z_n \end{pmatrix}. \end{aligned}$$

2. Fall: $i > 1$. Sei o. B. d. A. $i = 2$. Dann folgt

$$\Delta \begin{pmatrix} z_1 \\ z_2 + z_2' \\ z_3 \\ \vdots \\ z_n \end{pmatrix} = \sum_{j=1}^n (-1)^{1+j} m_{1j} \cdot \det(M_{1j}) = \Delta \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \\ z_n \end{pmatrix} + \Delta \begin{pmatrix} z_1 \\ z_2' \\ z_3 \\ \vdots \\ z_n \end{pmatrix},$$

da die Determinantenfunktion \det in der ersten Zeile der $(n-1) \times (n-1)$ -Matrizen M_{1j} additiv ist.

Die *Homogenität* ergibt sich entsprechend (siehe Übungsaufgabe 8).

(D2) Sei $\text{Rang}(M) < n$. Dann sind die *Spalten* von M linear abhängig. (Natürlich sind auch die *Zeilen* von M linear abhängig, aber es wird sich zeigen, dass es günstiger ist, die

Spalten zu betrachten.) Wir können uns – genauso wie im Beweis dafür, dass die Leibnizsche Formel eine Determinantenfunktion beschreibt – überlegen, dass wir o. B. d. A. annehmen können, dass M zwei gleiche Spalten hat. Ferner können wir o. B. d. A. annehmen, dass die *ersten beiden Spalten von M gleich* sind. (Wenn es nicht die ersten beiden Spalten sein sollten, müssten wir nur komplizierter nummerieren.)

Jetzt betrachten wir die einzelnen Summanden der Entwicklungsformel. Zunächst zu den ersten beiden Summanden:

1. Behauptung

$$m_{11} \cdot \det(M_{11}) = m_{12} \cdot \det(M_{12}) .$$

Das bedeutet, dass sich in der Entwicklungsformel die ersten beiden Summanden gegenseitig wegheben, denn der erste wird mit $+1$, der zweite mit -1 multipliziert. (Diese Behauptung ergibt sich ganz einfach: Da die erste Spalte von M gleich der zweiten ist, folgt zunächst $m_{11} = m_{12}$.

Außerdem ist $M_{11} = M_{12}$. Denn die erste Spalte von M_{11} ist gleich der (verkürzten) zweiten Spalte von M , und die erste Spalte von M_{12} ist gleich der (verkürzten) ersten Spalte von M ; da die beiden Spalten von M gleich sind, folgt die Behauptung.)

Nun betrachten wir die restlichen Summanden der Entwicklungsformel:

2. Behauptung

Für $j > 2$ gilt:

$$m_{1j} \cdot \det(M_{1j}) = 0 .$$

(Die ergibt sich noch einfacher: Die Matrix M_{1j} enthält die (verkürzten) beiden ersten Spalten von M . Da diese gleich sind, enthält M_{1j} also zwei gleiche Spalten. Daher ist $\det(M_{1j}) = 0$, und damit erst recht $m_{1j} \cdot \det(M_{1j}) = 0$.)

(D3) Diese Eigenschaft ist die einfachste: Da die erste Zeile der Einheitsmatrix E_n nur an der ersten Stelle eine Eins und sonst nur Nullen enthält, reduziert sich die Entwicklungsformel auf

$$\Delta(E_n) = 1 \cdot \det(E_{n-1}) .$$

Nach Induktion folgt also $\Delta(E_n) = 1$.

Damit ist alles gezeigt. \square

Sie werden sich fragen, ob man die Determinante einer Matrix auch durch Entwicklung nach einer *Spalte* berechnen kann. Die Antwort darauf ist „ja“:

Entwicklung nach einer Spalte

Sei $M = (m_{ij})_{1 \leq i, j \leq n}$ eine Matrix aus $K^{n \times n}$. Dann gilt für alle $j \in \{1, 2, \dots, n\}$:

$$\det(M) = \sum_{i=1}^n (-1)^{i+j} m_{ij} \cdot \det(M_{ij}) .$$

Mit anderen Worten: Um die Determinante zu berechnen, geht man die j -te Spalte Element für Element durch, multipliziert jeweils ein Element mit der entsprechenden Unterdeterminante und addiert diese Produkte mit alternierendem Vorzeichen.

Wie können wir diesen Satz *beweisen*? Den Trick kennen wir schon. Wenn wir zum Beispiel die Entwicklung nach der ersten Spalte beweisen wollten, würden wir eine Funktion Γ wie folgt definieren

$$\Gamma: M \rightarrow \sum_{i=1}^n (-1)^{1+i} m_{i1} \cdot \Gamma(M_{i1})$$

und zeigen, dass Γ den Bedingungen (D1), (D2) und (D3) genügt. Dann folgt mit dem Eindeutigkeitssatz, dass $\Gamma = \det$ ist. Diesen Beweis wollen wir hier nicht durchführen.

Sie können überprüfen, ob Sie den Beweis für die Entwicklung nach der ersten Zeile verstanden haben, indem Sie die Entwicklung nach der ersten Spalte beweisen (siehe Übungsaufgabe 9). \square

Achtung! Es ist ein verbreiteter Aberglaube, dass durch die Entwicklung nach einer Zeile oder nach einer Spalte die Berechnung einer Determinante rechentechnisch *leichter* werde. Das ist nicht richtig; denn man muss auch in diesem Fall (im Allgemeinen) alle $n!$ Permutationen betrachten.

Wenn man Glück (und den „richtigen Blick“) hat, kennt man manchmal schon einige der Unterdeterminanten. Besonders schön sind die Fälle, in denen man „sieht“, dass einige der Unterdeterminanten gleich Null sind. In der Regel lohnt sich die Berechnung einer Determinante mit Hilfe der Entwicklung nach einer Zeile oder Spalte nur, wenn es eine Zeile oder Spalte mit vielen Nullen gibt.

Die Entwicklung nach einer Zeile oder Spalte ist allerdings für theoretische Zwecke äußerst nützlich, da man mit diesem Instrument die Möglichkeit hat, Sätze mit Induktion nach der Spaltenzahl n zu beweisen.

Wir beschließen die Überlegungen dieses Abschnitts mit der Berechnung der Determinante der „transponierten“ Matrix von M . Sei M eine $m \times n$ -Matrix. Die zu M **transponierte** Matrix M^T entsteht dadurch, dass man die Zeilen als Spalten liest und umgekehrt. (Dann ist M^T eine $n \times m$ -Matrix.) Wenn zum *Beispiel*

$$M = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 0 \\ A & B & C & D & E \end{pmatrix}$$

ist, so gilt

$$M^T = \begin{pmatrix} 1 & 6 & A \\ 2 & 7 & B \\ 3 & 8 & C \\ 4 & 9 & D \\ 5 & 0 & E \end{pmatrix}.$$

Das können wir natürlich noch genauer sagen: Sei $M = (m_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$. Dann ist die zu M transponierte Matrix M^T gleich

$$M^T = (m_{ji})_{1 \leq j \leq n, 1 \leq i \leq m}.$$

(Die Bezeichnungen für die zu M transponierte Matrix sind nicht einheitlich. Man findet für M^T auch Bezeichnungen wie M' , M^* , M^t oder tM . Ich habe mich für die Bezeichnung M^T entschieden, weil ich glaube, dass man sich diese leicht merken kann.)

Wenn M eine quadratische Matrix ist, so ist auch M^T eine quadratische Matrix. Man kann also nach der Determinante von M^T fragen. Der folgende Satz gibt darauf eine einfache Antwort.

Determinante der transponierten Matrix

Sei $M \in K^{n \times n}$. Dann gilt

$$\det(M^T) = \det(M).$$

Der Beweis erfolgt durch Induktion nach n und ist nicht schwer.

Im Fall $n = 1$ ist die Aussage trivial; denn für jede 1×1 -Matrix M gilt sogar $M^T = M$.

Sei nun $n > 1$ und die Aussage richtig für $n-1$. Wir rechnen $\det(M)$ dadurch aus, dass wir $\det(M)$ nach der ersten Zeile entwickeln:

$$\det(M) = \sum_{j=1}^n (-1)^{1+j} m_{1j} \cdot \det(M_{1j})$$

Umgekehrt berechnen wir $\det(M^T)$ durch Entwicklung nach der ersten Spalte:

$$\det(M^T) = \sum_{j=1}^n (-1)^{j+1} m_{j1}^T \cdot \det(M_{j1}^T) = \sum_{j=1}^n (-1)^{j+1} m_{1j} \cdot \det(M_{j1}^T).$$

Nun betrachten wir die $(n-1) \times (n-1)$ -Matrizen M_{1j} und M_{j1}^T . Die Matrix M_{j1}^T entsteht dadurch, dass wir die j -te Zeile und die erste Spalte von M^T streichen, während M_{1j} dadurch entsteht, dass wir die erste Zeile und die j -te Spalte von M streichen. Daraus folgt, dass M_{1j} und M_{j1}^T transponierte Matrizen sind. Nach Induktion ist also

$$\det(M_{1j}) = \det(M_{j1}^T),$$

... und aus den obigen Formeln lesen wir nun die Behauptung $\det(M) = \det(M^T)$ ab. \square

Eine nützliche Folgerung ist, dass wir eine Determinante nicht nur mit Hilfe von elementaren Zeilenumformungen, sondern auch durch elementare Spaltenumformungen berechnen können:

Invarianz der Determinantenfunktion gegenüber elementaren Spaltenumformungen

Für jede Matrix $M \in K^{n \times n}$ gilt:

- (a) Verwandelt man M durch Multiplikation einer Spalte mit einem Element $k \in K$ in M' , so ist $\det(M') = k \cdot \det(M)$.
- (b) Verwandelt man die $n \times n$ -Matrix M durch Addition eines Vielfachen einer Spalte zu einer anderen Spalte aus M in die Matrix M' , so ist $\det(M') = \det(M)$.
- (c) Verwandelt man M durch Vertauschen zweier Spalten in M' , so ist $\det(M') = -\det(M)$.

Der Beweis ergibt sich aus dem Satz über die Invarianz der Determinantenfunktion gegenüber elementaren Zeilenumformungen aus Abschn. 7.1 und dem Satz über die Determinante der transponierten Matrix. \square

Zur Demonstration der Stärke der entwickelten Methoden, berechnen wir jetzt eine berühmte, auf den ersten Blick sehr kompliziert aussehende Determinante.

Seien $m_1, m_2, \dots, m_n \in K$. Eine Matrix $M \in K^{n \times n}$ der Form

$$M = \begin{pmatrix} 1 & m_1 & m_1^2 & \cdots & s & m_1^{n-1} \\ 1 & m_2 & m_2^2 & \cdots & s & m_2^{n-1} \\ 1 & m_3 & m_3^2 & \cdots & s & m_3^{n-1} \\ \vdots & \vdots & \vdots & & & \vdots \\ 1 & m_n & m_n^2 & \cdots & s & m_n^{n-1} \end{pmatrix}$$

heißt nach dem französischen Mathematiker Alexandre Théophile Vandermonde (1735–1796) **Vandermondesche Matrix**; ihre Determinante heißt **Vandermondesche Determinante**.

Die Aufgabe besteht darin, die Determinante einer beliebigen Vandermondeschen Matrix zu berechnen (und diese als Funktion von m_1, m_2, \dots, m_n auszudrücken). Es bietet sich an, bei der Berechnung der Determinante durch Induktion nach n vorzugehen. Die grundsätzliche Schwierigkeit bei jedem Induktionsbeweis besteht darin, dass man *vorher* wissen muss, was man beweisen will. Man braucht also eine *Vermutung*. Diese wird durch die Behauptung des folgenden Satzes ausgedrückt.

Satz über die Vandermondesche Determinante

Seien $m_1, m_2, \dots, m_n \in K$. Dann gilt

$$\det \begin{pmatrix} 1 & m_1 & m_1^2 & \cdots & m_1^{n-1} \\ 1 & m_2 & m_2^2 & \cdots & m_2^{n-1} \\ 1 & m_3 & m_3^2 & \cdots & m_3^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & m_n & m_n^2 & \cdots & m_n^{n-1} \end{pmatrix} = \prod_{1 \leq i < j \leq n} (m_j - m_i).$$

Zum *Beispiel* ist also

$$\det \begin{pmatrix} 1 & a & a^2 & a^3 \\ 1 & b & b^2 & b^3 \\ 1 & c & c^2 & c^3 \\ 1 & d & d^2 & d^3 \end{pmatrix} = (b-a)(c-a)(d-a)(c-b)(d-b)(d-c).$$

Die Strategie des *Beweises* besteht darin, durch elementare Umformungen eine Zeile herzustellen, in der nur ein Element von Null verschieden ist und dann nach dieser Zeile zu entwickeln. Also ans Werk!

In einer Art *Vorspiel* überzeugen wir uns, dass die Aussage jedenfalls für $n=2$ richtig ist: In diesem Fall ist

$$M = \begin{pmatrix} 1 & m_1 \\ 1 & m_2 \end{pmatrix}$$

und als Determinante ergibt sich unschwer

$$\det(M) = \det \begin{pmatrix} 1 & m_1 \\ 1 & m_2 \end{pmatrix} = 1 \cdot m_2 - m_1 \cdot 1 = m_2 - m_1.$$

Sei nun $n > 2$ und die Aussage des Satzes richtig für $n-1$. Das jetzt folgende Schauspiel gliedert sich in drei Akte:

Erster Akt Zunächst multiplizieren wir die *vorletzte* Spalte mit m_1 und ziehen sie von der letzten ab. Damit erreichen wir zwei Effekte: Zum einen ergibt sich als letztes Element in der ersten Zeile Null (Hurra!) und zweitens ändert sich nach dem Satz über elementare (Spalten-)Umformungen die Determinante dadurch nicht.

Was einmal gut geht, geht auch ein zweites Mal gut: Wir multiplizieren die *vorvorletzte* Spalte mit m_1 und ziehen sie von der *vorletzten* ab. Damit wird auch das *vorletzte* Element der ersten Zeile Null (Hurra, hurra!) und die Determinante hat sich noch immer nicht geändert.

Nun ist klar, wie es weitergeht: Wir führen den entsprechenden Prozess so oft wie möglich durch. Der letzte Durchgang besteht darin, das m_1 -fache der ersten Spalte von der

zweiten abzuziehen. Dadurch wird auch das zweite Element der ersten Zeile Null, und wir haben tatsächlich unsere Wunschzeile erreicht:

$$\det \begin{pmatrix} 1 & m_1 & m_1^2 & \cdots s & m_1^{n-1} \\ 1 & m_2 & m_2^2 & \cdots s & m_2^{n-1} \\ 1 & m_3 & m_3^2 & \cdots s & m_3^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & m_n & m_n^2 & \cdots s & m_n^{n-1} \end{pmatrix}$$

$$= \det \begin{pmatrix} 1 & 0 & 0 & \cdots s & 0 \\ 1 & m_2 - m_1 & m_2^2 - m_1 m_2 & \cdots s & m_2^{n-1} - m_1 m_2^{n-2} \\ 1 & m_3 - m_1 & m_3^2 - m_1 m_3 & \cdots s & m_3^{n-1} - m_1 m_3^{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & m_n - m_1 & m_n^2 - m_1 m_n & \cdots s & m_n^{n-1} - m_1 m_n^{n-2} \end{pmatrix}.$$

Zweiter Akt Wir entwickeln die erhaltene Matrix nach der ersten Zeile. Dies ist in unserem Fall besonders einfach; wir erhalten:

$$\det(M) = 1 \cdot \det \begin{pmatrix} m_2 - m_1 & m_2^2 - m_1 m_2 & \cdots s & m_2^{n-1} - m_1 m_2^{n-2} \\ m_3 - m_1 & m_3^2 - m_1 m_3 & \cdots s & m_3^{n-1} - m_1 m_3^{n-2} \\ \vdots & \vdots & & \vdots \\ m_n - m_1 & m_n^2 - m_1 m_n & \cdots s & m_n^{n-1} - m_1 m_n^{n-2} \end{pmatrix}.$$

Der nächste Schritt ist klar: Aus der ersten Zeile dieser Matrix kann man den Faktor $m_2 - m_1$ herausziehen, aus der zweiten den Faktor $m_3 - m_1, \dots$ und aus der letzten den Faktor $m_n - m_1$. So erhalten wir

$$\det(M) = \dots = (m_2 - m_1) \cdot (m_3 - m_1) \cdot \dots \cdot (m_n - m_1)$$

$$\cdot \det \begin{pmatrix} 1 & m_2 & m_2^2 & \cdots s & m_2^{n-2} \\ 1 & m_3 & m_3^2 & \cdots s & m_3^{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & m_n & m_n^2 & \cdots s & m_n^{n-2} \end{pmatrix}.$$

Dritter und letzter Akt Zu unserer großen Freude erkennen wir jetzt auf der rechten Seite wieder eine Vandermondematrix – und was noch schöner ist: Es handelt sich um eine $(n-1) \times (n-1)$ -Vandermondematrix. Also können wir Induktion anwenden und erhalten:

$$\det(M) = \dots = (m_2 - m_1) \cdot (m_3 - m_1) \cdot \dots \cdot (m_n - m_1) \cdot \prod_{2 \leq i < j \leq n} m_j - m_i$$

$$= \prod_{1 \leq i < j \leq n} m_j - m_i.$$

Damit haben wir tatsächlich alles gezeigt, und der Beweis ist beendet. \square

7.6 Der Multiplikationssatz

Das Hauptergebnis dieses Abschnitts ist das folgende:

Multiplikationssatz für Determinanten

Seien $M, M' \in K^{n \times n}$. Dann gilt

$$\det(MM') = \det(M) \cdot \det(M') .$$

In Worten heißt dies: Die Determinante des Produkts zweier Matrizen ist gleich dem Produkt ihrer Determinanten.

Bevor wir den Satz beweisen, werden wir einige Folgerungen daraus aufschreiben. Alle vorkommenden Matrizen seien aus $K^{n \times n}$.

Folgerungen aus dem Multiplikationssatz

1. **Folgerung.** Das Produkt zweier Matrizen mit Determinante $\neq 0$ hat ebenfalls Determinante $\neq 0$. □
2. **Folgerung.**
 - (a) Das Produkt zweier Matrizen mit Determinante 1 hat ebenfalls Determinante 1.
 - (b) Das Produkt zweier Matrizen mit Determinante 1 oder -1 hat ebenfalls Determinante 1 oder -1 . □
3. **Folgerung.** Für je zwei Matrizen $M, M' \in K^{n \times n}$ gilt

$$\det(MM') = \det(M'M) .$$

□

4. **Folgerung.** Für jede invertierbare Matrix $M \in K^{n \times n}$ gilt

$$\det(M^{-1}) = \det(M)^{-1} .$$

Das bedeutet: Um die Determinante der Matrix M^{-1} auszurechnen, muss man nicht die Matrix M invertieren; wenn man $\det(M)$ kennt, reicht es, dieses Körperelement zu invertieren. □

Zum Beweis des Multiplikationssatzes brauchen wir einige Vorbereitungen. Zunächst überlegen wir uns, wie man eine elementare Zeilenumformung vom Typ 3 durch eine Matrizenmultiplikation ausdrücken kann.

Multiplikation mit Elementarmatrizen

Sei $M \in K^{n \times n}$. Die Matrix M' entstehe aus M durch Addition des a -fachen der j -ten Zeile von M zur i -ten Zeile von M ($i \neq j$). Dann gilt

$$M' = A \cdot M,$$

wobei A die Matrix aus $K^{n \times n}$ ist, die auf der Hauptdiagonale Einsen, an der Stelle (i, j) das Element a und sonst Nullen hat:

$$A = \begin{pmatrix} 1 & 0 & 0 & \cdots & s & 0 \\ 0 & 1 & 0 & & & 0 \\ 0 & 0 & \ddots & a & \vdots & \\ \vdots & & & \ddots & 0 & \\ 0 & 0 & & 0 & 1 \end{pmatrix}.$$

Solche Matrizen werden **Elementarmatrizen** genannt.

Jede Elementarmatrix hat die Determinante 1.

Der *Beweis* ist einfach und erfolgt dadurch, dass wir die Matrix AM berechnen.

Wie erhält man die erste Zeile von AM ? Nichts leichter als das: Man multipliziert den Vektor $(1, 0, 0, \dots, 0)$ der Reihe nach mit den Spalten von M ; offenbar reproduziert man dadurch nur die erste Zeile z_1 von M .

Dieses Argument können wir sofort verallgemeinern: Auch die Zeilen z_2, \dots, z_{i-1} und z_{i+1}, \dots, z_n werden nur reproduziert.

Letzte Frage: Was passiert mit der i -ten Zeile? Da die i -te Zeile von A an der i -ten Stelle eine 1, an der j -ten Stelle das Element a und sonst Nullen hat, ergibt sich als i -te Zeile von AM die Summe $z_i + az_j$.

Insgesamt haben wir damit $AM = M'$ gezeigt.

Da jede Elementarmatrix eine Dreiecksmatrix ist, die auf der Hauptdiagonale nur Einsen hat, ist ihre Determinante gleich 1. \square

Die zu einer Elementarmatrix inverse Matrix ist wieder eine Elementarmatrix. Genauer gesagt gilt: Ist A die Elementarmatrix aus $K^{n \times n}$, bei der an der Stelle (i, j) außerhalb der Hauptdiagonalen das Element $a \neq 0$ steht, so ist A^{-1} diejenige Elementarmatrix, bei der an der Stelle (i, j) das Element $-a$ steht (siehe Übungsaufgabe 13).

Eine wichtige Folgerung aus obigem Lemma ist der folgende Satz.

Satz

Sei $M \in K^{n \times n}$. Wenn $\det(M) \neq 0$ ist, so gibt es ein Produkt S von Elementarmatrizen und eine Diagonalmatrix D mit $\det(D) = \det(M)$, so dass gilt:

$$M = SD.$$

Man kann D so wählen, dass jedes Diagonalelement – bis auf das letzte – gleich 1 ist; dieses letzte Diagonalelement ist dann gleich der Determinante von M .

Beweis Wir verwandeln M durch elementare Zeilenumformungen vom Typ 3 in eine Diagonalmatrix D . Jede solche Zeilenumformung wird durch eine Elementarmatrix beschrieben. Also wird die Verwandlung von M in D durch ein Produkt T von Elementarmatrizen beschrieben. Nach dem Satz über die Multiplikation mit Elementarmatrizen gilt also

$$D = TM.$$

Da die zu einer Elementarmatrix inverse Matrix wieder eine Elementarmatrix ist, hat auch T eine Inverse S , die wieder ein Produkt von Elementarmatrizen ist. Damit gilt $SD = M$.

Der Rest des Beweises ist Übungsaufgabe 14. □

Nun beweisen wir den Multiplikationssatz. Dazu unterscheiden wir zwei Fälle.

1. Fall: $\det(M) = 0$.

Dann ist $\text{Rang}(M) < n$. Wir erinnern uns, dass $\text{Rang}(MM') = \dim(\text{Bild}(MM'))$ ist – wenn man MM' als lineare Abbildung des K^n (dessen Elemente wir als Spaltenvektoren auffassen) ansieht. Da natürlich $\text{Bild}(MM') \subseteq \text{Bild}(M)$ ist, folgt

$$\text{Rang}(MM') = \dim(\text{Bild}(MM')) \leq \dim(\text{Bild}(M)) = \text{Rang}(M) < n.$$

Also ist auch $\det(MM') = 0$.

2. Fall: $\det(M) \neq 0$.

Wir haben in diesem Fall kein Mittel, die Determinante des Produkts zweier Matrizen auszurechnen. Das einzige, was entfernt an eine solche Aufspaltung erinnert, ist die Homogenität: Wir können ein Körperelement aus einer Zeile „herausziehen“. Das ist die Idee: Wir verwandeln eine der beiden Matrizen in ein Körperelement und nützen dann die Homogenität aus.

Bleibt nur die Frage: Wie verwandelt man eine Matrix in ein Körperelement (das auch noch die Determinante der Matrix sein muss)?

Die Antwort ist die folgende: Da $\det(M) \neq 0$ ist, können wir nach dem vorigen Satz M durch elementare Zeilenumformungen vom Typ 3 so in eine Diagonalmatrix D ver-

wandeln, dass alle Diagonalelemente von D gleich 1 sind – nur das letzte ist gleich der Determinante von M , die wir mit d bezeichnen. Das heißt: Es gibt ein Produkt S von Elementarmatrizen, so dass

$$M = SD$$

ist. Dann ist

$$\det(MM') = \det(SDM') = \det(DM'),$$

da SDM' aus DM' durch elementare Zeilenvertauschungen vom Typ 3 hervorgeht. Die ersten $n-1$ Zeilen der Matrix DM' sind die gleichen wie die von M' ; nur die letzte Zeile wurde mit d multipliziert. Wegen der Homogenität folgt somit schließlich

$$\det(MM') = \dots = \det(DM') = d \cdot \det(M') = \det(M) \cdot \det(M').$$

□

7.7 Verständnisfragen, Übungen und Tipps

Richtig oder falsch?

1. Thema: Permutationen

- ☐ Jede Permutation ist zyklisch.
- ☐ Jede Permutation lässt sich als Produkt disjunkter Zyklen schreiben.
- ☐ Jede Permutation lässt sich als Produkt nicht-disjunkter Zyklen schreiben.
- ☐ Jede Permutation lässt sich als Produkt von Zyklen der Länge 5 schreiben.
- ☐ Jede Permutation ist eine Hintereinanderausführung von Fehlständen.

2. Thema: Gerade und ungerade Permutationen

- ☐ Das Produkt zweier gerader Permutationen ist gerade.
- ☐ Das Produkt zweier ungerader Permutationen ist gerade.
- ☐ Das Produkt einer geraden mit einer ungeraden Permutation ist gerade.
- ☐ Wenn n gerade ist, ist jede Permutation aus S_n gerade.

3. Thema: Determinanten

Die Determinante einer $n \times n$ -Matrix ist

- ☐ ein quadratisches Schema,
- ☐ eine Matrix,
- ☐ eine Abbildung,
- ☐ ein Körperelement,
- ☐ ein Vektor,
- ☐ eine Permutation.

Übungsaufgaben

1. Zeigen Sie: Es gibt genau zwei Permutationen aus S_3 , die Zyklen der Länge 3 sind; es gibt genau sechs Permutationen aus S_4 , die durch einen Zyklus der Länge 4 beschrieben werden. Allgemein: Es gibt genau $(n-1)!$ Permutationen aus S_n , die durch einen Zyklus der Länge n beschrieben werden.
2. Zeigen Sie: Die Anzahl der Fehlstände einer Permutation π aus S_n ist gleich der Anzahl der Fehlstände von π^{-1} .
3. Berechnen Sie die Determinanten der folgenden reellen Matrizen:

$$\begin{pmatrix} -1 & 2 & 3 \\ 2 & 1 & 0 \\ 1 & 8 & 9 \end{pmatrix}, \quad \begin{pmatrix} -1 & 0 & 5 \\ 0 & 0 & 7 \\ 3 & 3 & 3 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 6 & 6 \\ 1 & 4 & 9 & 2 \\ 1 & 7 & 8 & 9 \\ 1 & 9 & 9 & 9 \end{pmatrix}.$$

4. Welche Determinanten haben die folgenden reellen Matrizen?

$$\begin{pmatrix} 1 & 3 & 3 \\ 3 & 1 & 3 \\ 3 & 3 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 3 & 3 & 3 \\ 3 & 1 & 3 & 3 \\ 3 & 3 & 1 & 3 \\ 3 & 3 & 3 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & a & 0 & \dots & 0 \\ a & 1 & a & & \vdots \\ 0 & a & \ddots & \ddots & 0 \\ \vdots & & \ddots & 1 & a \\ 0 & \dots & 0 & a & 1 \end{pmatrix}.$$

5. Bestimmen Sie die Determinante der folgenden Matrix aus $K^{n \times n}$:

$$\begin{pmatrix} a & b & b & \dots & b \\ b & a & b & \dots & b \\ b & b & a & & b \\ \vdots & & & \ddots & \vdots \\ b & b & \dots & b & a \end{pmatrix}.$$

6. Eine Matrix P aus $K^{n \times n}$ heißt eine Permutationsmatrix, falls jedes Element von P entweder 0 oder 1 ist, und, falls in jeder Zeile und in jeder Spalte genau eine 1 vorkommt.
Berechnen Sie die Determinante einer beliebigen Permutationsmatrix.
7. Machen Sie sich die Richtigkeit der Entwicklung nach der ersten Zeile in den Fällen $n=2$ und $n=3$ explizit klar.
8. Beweisen Sie, dass die im Beweis der Entwicklung nach der ersten Zeile auftauchende Funktion Δ auch homogen in jeder Zeile ist.

9. Beweisen Sie die Entwicklung einer Determinante nach der ersten Spalte:

Sei $M = (m_{ij})_{1 \leq i, j \leq n}$ eine Matrix aus $K^{n \times n}$. Dann gilt:

$$\det(M) = \sum_{i=1}^n (-1)^{i+1} m_{i1} \cdot \det(M_{i1}) .$$

10. Seien $A, B \in K^{n \times n}$. Zeigen Sie

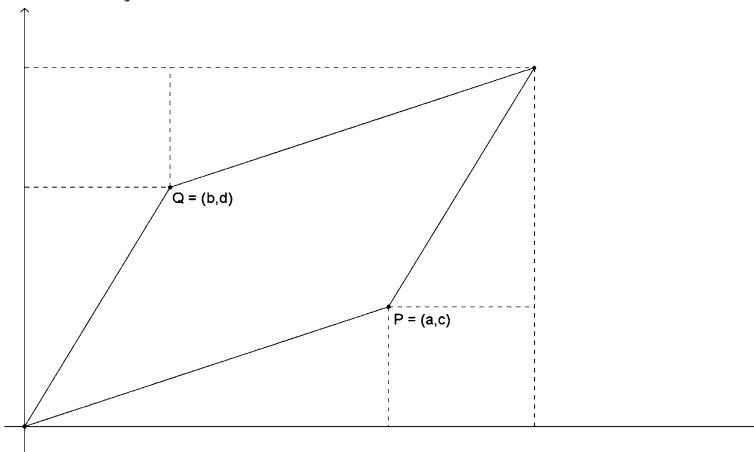
$$(AB)^T = B^T A^T .$$

11. Zeigen Sie, dass das Transponieren von Matrizen eine lineare Abbildung auf der Menge $K^{n \times n}$ ist.
12. Zeigen Sie: Das Produkt zweier reeller $n \times n$ -Matrizen mit positiver Determinante hat ebenfalls eine positive Determinante.
13. Zeigen Sie: Wenn A eine Elementarmatrix aus $K^{n \times n}$ ist, bei der an der Stelle (i, j) außerhalb der Hauptdiagonalen das Element $a \neq 0$ steht, so ist A^{-1} diejenige Elementarmatrix, bei der an der Stelle (i, j) das Element $-a$ steht.
14. Sei M eine $n \times n$ -Matrix, deren Determinante nicht Null ist. Dann gibt es eine Diagonalmatrix D , so dass jedes Diagonalelement – bis auf das letzte – gleich 1 ist; ferner gibt es ein Produkt S von Elementarmatrizen mit $M = SD$.

Projekt: Determinanten und Volumen

1. Berechnen Sie den Flächeninhalt des Parallelogramms, bei dem der Nullpunkt eine Ecke ist und das durch die Vektoren $P = (a, c)$ und $Q = (b, d)$ aufgespannt wird. Wir nennen diese Funktion $A(a, b, c, d)$.

[Tipp: Sie können dazu die in der folgenden Zeichnung angedeutete Flächenzerlegung verwenden.]



2. Kann $A(a, b, c, d) = 0$ sein? Kann $A(a, b, c, d)$ eine negative Zahl sein?
3. Wie ändert sich der Wert der Funktion $A(a, b, c, d)$, wenn der Vektor (a, b) mit einer Zahl $k > 0$ multipliziert wird? Interpretieren Sie das geometrisch.
Was passiert, wenn man (a, b) und (c, d) vertauscht?
4. Formulieren Sie den Zusammenhang der Eigenschaften der Funktion $A(a, b, c, d)$ mit der Determinante.
5. Verallgemeinern Sie den Sachverhalt aus den Aufgabe 1 bis 4 auf den \mathbf{R}^3 .
6. Man kann eine reelle 2×2 -Matrix A auch als lineare Abbildung von \mathbf{R}^2 in sich auffassen. Zeigen Sie: $\det(A)$ ist der Faktor, um den sich die Fläche des Einheitsquadrats bei Anwendung von A ändert.
Interpretieren Sie in diesem Zusammenhang den Multiplikationssatz von Matrizen.

Sie sollten mit folgenden Begriffen umgehen können

Determinantenfunktion, elementare Umformungen, Permutation, Zyklus, Transposition, Nachbartransposition, Fehlstand, gerade und ungerade Permutationen, Signum, Leibnizsche Formel, Regel von Sarrus, Entwicklung nach einer Zeile, transponierte Matrix, Elementarmatrix, Multiplikationssatz

Was sagen Sie dazu?**Gnutpuaheb**

Jede nichtquadratische Matrix ist invertierbar.

Sieweb

Sei M eine nichtquadratische Matrix. Dann hat M keine Determinante. Also ist die Determinante von M nicht Null. Also ist M invertierbar. 



In diesem Kapitel gehen wir der Frage nach, welche linearen Abbildungen eines Vektorraums in sich durch eine „Diagonalmatrix“ dargestellt werden können, also durch eine Matrix, bei der nur auf der Diagonale von Null verschiedene Elemente stehen dürfen. Um diese – nicht ganz einfache – Frage zu beantworten, brauchen wir als wichtige Instrumente die Begriffe Eigenwert, charakteristisches Polynom und Minimalpolynom.

8.1 Einführung

In Kap. 5 haben wir gesehen, dass sich jede lineare Abbildung eines Vektorraums V in einen Vektorraum W durch eine Matrix darstellen lässt. Diese Darstellungsmatrix hängt von der Auswahl einer Basis von V und einer Basis von W ab.

Es stellt sich somit die Frage, ob man durch geeignete Wahl dieser Basen die Darstellungsmatrix möglichst einfach gestalten kann.

Was heißt hier einfach? In diesem Fall ist diese philosophische Frage einfach zu beantworten: Man muss mit Matrizen gut rechnen können! Insbesondere müssen die Multiplikation einer Matrix mit einem Vektor und die Multiplikation zweier Matrizen einfach auszuführen sein.

Wenn wir uns eine einfache Matrix wünschen dürften, was würden wir uns dann wünschen? Am einfachsten sind solche Matrizen zu multiplizieren, die möglichst viele Nullen haben. Es wäre aber weltfern zu erwarten, dass man jede Matrix in eine äquivalente umformen kann, die nur aus Nullen besteht. Das Ideal, von dem man realistisch träumen kann, sind Matrizen, die außerhalb ihrer Diagonalen nur Nullen haben. Diese „Diagonalmatri-

zen“ sind außerordentlich einfach zu multiplizieren; es ist nämlich

$$\begin{pmatrix} k_1 & 0 & \dots & 0 \\ 0 & k_2 & \dots & \dots \\ \dots & \dots & \dots & 0 \\ 0 & \dots & 0 & k_n \end{pmatrix} \cdot \begin{pmatrix} h_1 & 0 & \dots & 0 \\ 0 & h_2 & \dots & \dots \\ \dots & \dots & \dots & 0 \\ 0 & \dots & 0 & h_n \end{pmatrix} = \begin{pmatrix} k_1 \cdot h_1 & 0 & \dots & 0 \\ 0 & k_2 \cdot h_2 & \dots & \dots \\ \dots & \dots & \dots & 0 \\ 0 & \dots & 0 & k_n \cdot h_n \end{pmatrix}.$$

Wir werden das Problem aber nicht in aller Allgemeinheit angehen, sondern sofort zwei Reduktionen o. B. d. A. vornehmen.

- Zum einen betrachten wir nicht lineare Abbildungen eines Vektorraums V in einen beliebigen anderen Vektorraum W , sondern lineare Abbildungen von V in sich. Dies können wir ohne Einschränkung machen: Ist nämlich f eine lineare Abbildung von V in W , so ist $\text{Bild}(f)$ ein Unterraum von W , dessen Dimension nicht größer ist als die Dimension von V . Daher können wir $\text{Bild}(f)$ mit einem Unterraum von V identifizieren; wenn wir eine gute Beschreibung für diesen Fall haben, können wir die Methoden oder die Ergebnisse auch auf die allgemeine Situation übertragen.
- Zum anderen betrachten wir nicht allgemeine Darstellungsmatrizen ${}_B M_{B'}(f)$ bezüglich zweier Basen B und B' , sondern nur Darstellungsmatrizen ${}_B M_B(f)$ bezüglich einer einzigen Basis B . Denn wenn man eine gegebene Basis B umsortieren und die Basis B' frei wählen kann, so kann man erreichen, dass ${}_B M_{B'}(f)$ folgende besonders einfache Form hat (siehe Übungsaufgabe 1):

$${}_B M_{B'}(f) = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & \ddots & & & 0 \\ 0 & & 1 & & 0 \\ \vdots & & & 0 & \vdots \\ 0 & 0 & 0 & \dots & \ddots \end{pmatrix}$$

Dies suggeriert die zentrale Definition dieses Kapitels.

Sei f eine lineare Abbildung des K -Vektorraums V in sich. Dann heißt f **diagonalisierbar**, wenn es eine Basis B von V gibt, so dass in der Darstellungsmatrix ${}_B M_B(f)$ höchstens auf der Diagonalen von Null verschiedene Elemente stehen. Eine solche Matrix heißt auch **Diagonalmatrix**.

Mit anderen Worten: f ist diagonalisierbar, wenn bezüglich einer geeigneten Basis B die Darstellungsmatrix folgende Gestalt hat:

$${}_B M_B(f) = \begin{pmatrix} k_1 & 0 & \dots & 0 \\ 0 & k_2 & \dots & \dots \\ \dots & \dots & \dots & 0 \\ 0 & \dots & 0 & k_n \end{pmatrix}.$$

Wir drücken dieses Phänomen nochmals anders aus: *Die lineare Abbildung f ist genau dann diagonalisierbar, wenn es eine Basis $B = \{v_1, v_2, \dots, v_n\}$ von V gibt, so dass*

$$f(v_1) = k_1 v_1, f(v_2) = k_2 v_2, \dots, f(v_n) = k_n v_n$$

ist für geeignete Körperelemente k_1, k_2, \dots, k_n .

Ist jede lineare Abbildung diagonalisierbar??? Die Antwort lautet – nein! Daraus ergeben sich die zentralen Fragen dieses Kapitels:

- Unter welchen Bedingungen ist eine gegebene lineare Abbildung diagonalisierbar?
- Gibt es ein Verfahren, mit dem man entscheiden kann, ob eine gegebene lineare Abbildung diagonalisierbar ist?
- Gibt es ein Verfahren, mit dem man eine Basis finden kann, bezüglich derer die Darstellungsmatrix einer gegebenen linearen Abbildung eine Diagonalmatrix ist?
- Wenn f nicht diagonalisierbar ist, was dann?

Wir halten im ganzen Kapitel folgende Bezeichnungen fest. Sei V ein n -dimensionaler Vektorraum über einem Körper K , und sei f eine lineare Abbildung von V in sich.

Dieses Kapitel ist nicht ganz einfach, aber zentral in der linearen Algebra. Entsprechend gründlich sollten Sie es studieren.

8.2 Eigenvektoren und Eigenwerte

Das wichtigste Instrument zur Untersuchung der Frage der Diagonalisierbarkeit sind die Begriffe „Eigenvektor“ und „Eigenwert“. Diese Begriffe spielen aber auch unabhängig von der Frage der Diagonalisierbarkeit in der Mathematik und ihren Anwendungen eine zentrale Rolle.

Ein Körperelement $k \in K$ heißt ein **Eigenwert** von f , falls es einen Vektor $v \in V$ gibt mit

$$f(v) = k \cdot v \text{ und } v \neq 0.$$

Ein Vektor $v \in V$ heißt ein **Eigenvektor** von f zum Eigenwert k , falls gilt

$$f(v) = k \cdot v \text{ und } v \neq 0.$$

Bemerkung In der Definition eines Eigenwerts mussten wir fordern, dass v von Null verschieden ist; denn sonst wäre jedes Körperelement ein Eigenwert (da $f(0) = 0 = k \cdot 0$).

Beispiele

- (a) Die Identität hat den Eigenwert 1 und keinen sonst.
- (a)' Eine lineare Abbildung f hat den Eigenwert 1, wenn es einen Vektor $v \neq o$ gibt mit $f(v) = v$.
- (b) Die Nullabbildung hat 0 als Eigenwert und keinen sonst.
- (b)' Eine lineare Abbildung f hat 0 als Eigenwert, wenn es einen Vektor $v \neq o$ gibt mit $f(v) = o$.
- (c) Die lineare Abbildung f von R^2 in sich, die definiert ist durch $f(x, y) := (x, 0)$ hat die Eigenwerte 0 und 1: Jeder Vektor der Form $(0, y)$ hat den Eigenwert 0 und jeder Vektor der Form $(x, 0)$ hat den Eigenwert 1.
- (d) Die Eigenvektoren zum Eigenwert 0 bilden zusammen mit dem Nullvektor genau den Kern von f :
 $\{v \in V | f(v) = o, v \neq 0\} \cup \{o\} = \text{Kern}(f)$.

Mit dem Begriff des Eigenvektors können wir unsere Fragestellung bereits nutzbringend umformulieren.

0. Kriterium zur Diagonalisierbarkeit

Eine lineare Abbildung f von V in sich ist genau dann diagonalisierbar, wenn es eine Basis von V aus Eigenvektoren gibt.

Beweis Wenn V eine Basis aus Eigenvektoren von f besitzt, ist die Darstellungsmatrix von f bezüglich dieser Basis eine Diagonalmatrix.

Umgekehrt: Wenn f diagonalisierbar ist, so gibt es eine Basis $B = \{v_1, v_2, \dots, v_n\}$ von V , so dass die Darstellungsmatrix von f bezüglich B eine Diagonalmatrix ist. Wenn k_1, k_2, \dots, k_n die Elemente auf der Diagonalen dieser Matrix sind, so folgt $f(v_i) = k_i v_i$ für $i = 1, 2, \dots, n$. Also ist jedes v_i ein Eigenvektor, und somit besteht B nur aus Eigenvektoren. \square

Es wird sich als zweckmäßig erweisen, Eigenwerte von Matrizen zu betrachten. Dazu definieren wir: Sei $M \in K^{n \times n}$. Ein Element $k \in K$ heißt ein **Eigenwert** von M , falls k ein Eigenwert einer linearen Abbildung ist, die bezüglich einer Basis durch M definiert ist.

Das bedeutet: Das Element $k \in K$ ist ein Eigenwert von $M = (m_{ij})$, falls es eine Basis $\{v_1, v_2, \dots, v_n\}$ von K^n gibt, so dass die lineare Abbildung f , die durch

$$f(v_j) := m_{1j}v_1 + m_{2j}v_2 + \dots + m_{nj}v_n$$

definiert ist, den Eigenwert k hat.

Kurz: *Eigenwerte einer Matrix sind die Eigenwerte der zugehörigen linearen Abbildung* (vergleichen Sie hierzu Übungsaufgabe 5).

Beispiel Man macht sich ohne große Schwierigkeiten klar, dass die Matrix $M = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ die Eigenwerte 1 und 0 hat.

Satz über die Eigenwerte einer Matrix

Die Eigenwerte einer Matrix M sind genau die Elemente $k \in K$, für die es einen von Null verschiedenen Spaltenvektor x gibt mit $Mx = kx$.

Beweis Übungsaufgabe 2.

Wir beginnen unsere Untersuchungen von Eigenwerten und Eigenvektoren behutsam.

Lemma über verschiedene Eigenwerte

Eigenvektoren zu verschiedenen Eigenwerten sind linear unabhängig.

Beweis Seien v_1, v_2, \dots, v_m Eigenvektoren zu den Eigenwerten k_1, k_2, \dots, k_m , wobei k_1, k_2, \dots, k_m paarweise verschieden sind. Wir zeigen die Behauptung durch Induktion nach m .

Der Fall $m = 1$ ergibt sich aus der Definition eines Eigenwerts, da $v_1 \neq 0$ sein muss.

Sei nun $m > 1$, und sei die Behauptung richtig für $m-1$. Es sei

$$h_1 v_1 + h_2 v_2 + \dots + h_m v_m = 0.$$

Wir haben zu zeigen, dass $h_1 = h_2 = \dots = h_m = 0$ ist. Durch Anwenden von f erhalten wir aus obiger Gleichung

$$0 = f(0) = h_1 f(v_1) + h_2 f(v_2) + \dots + h_m f(v_m) = h_1 k_1 v_1 + h_2 k_2 v_2 + \dots + h_m k_m v_m.$$

Andererseits gilt auch

$$0 = k_m \cdot 0 = k_m \cdot (h_1 v_1 + h_2 v_2 + \dots + h_m v_m) = k_m h_1 v_1 + k_m h_2 v_2 + \dots + k_m h_m v_m.$$

Da K kommutativ ist, ergibt sich zusammen

$$0 = (k_m - k_1)h_1 v_1 + (k_m - k_2)h_2 v_2 + \dots + (k_m - k_{m-1})h_{m-1} v_{m-1}.$$

Nach Induktion folgt daraus

$$(k_m - k_1)h_1 = 0, (k_m - k_2)h_2 = 0, \dots, (k_m - k_{m-1})h_{m-1} = 0.$$

Da nach Voraussetzung $k_m \neq k_i$ ist ($1 \leq i \leq m-1$), erhalten wir $h_1 = h_2 = \dots = h_{m-1} = 0$. Damit folgt auch $h_m = 0$. \square

Aus diesem Lemma ergibt sich der erste Satz, der eine hinreichende Bedingung für die Diagonalisierbarkeit einer linearen Abbildung angibt.

Satz

Wenn eine lineare Abbildung eines n -dimensionalen Vektorraums n verschiedene Eigenwerte hat, so ist sie diagonalisierbar.

Beweis Seien v_1, v_2, \dots, v_n Eigenvektoren zu den n verschiedenen Eigenwerten. Nach obigem Lemma sind dann v_1, v_2, \dots, v_n linear unabhängig; also bilden sie eine Basis. Bezüglich dieser Basis hat f Diagonalgestalt. \square

Beispiel Die Matrix $M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ hat die Eigenwerte 1 und -1 (mit Eigenvektoren $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ und $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$), ist also diagonalisierbar. Eine Diagonalgestalt lautet $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

Der letzte grundlegende Begriff dieses Abschnitts ist der des Eigenraums. Sei $k \in K$. Mit $\text{Eig}(f, k)$ bezeichnen wir die Menge aller Vektoren v mit $f(v) = k \cdot v$:

$$\text{Eig}(f, k) = \{v \in V \mid f(v) = k \cdot v\}.$$

Für einen Eigenwert k nennt man $\text{Eig}(f, k)$ den **Eigenraum** von f zum Eigenwert k .

Im folgenden Hilfssatz fassen wir die wichtigsten Eigenschaften von Eigenräumen zusammen.

Lemma über Eigenräume

Sei $k \in K$ beliebig. Dann gilt:

- (a) $\text{Eig}(f, k)$ ist ein Unterraum von V .
- (b) Genau dann ist k ein Eigenwert von f , wenn $\text{Eig}(f, k) \neq \{0\}$ ist.
- (c) $\text{Eig}(f, k) = \text{Kern}(f - k \cdot \text{id})$.
- (d) Seien k_1, k_2, \dots, k_s (paarweise) verschiedene Elemente von K . Ist B_r eine Basis von $\text{Eig}(f, k_r)$ ($r = 1, 2, \dots, s$), so ist $B = B_1 \cup B_2 \cup \dots \cup B_s$ eine linear unabhängige Menge von V .

Beweis

- (a) Nach dem Unterraumkriterium ist nur die Abgeschlossenheit bezüglich der Addition und der skalaren Multiplikation zu zeigen. Seien also $v, w \in \text{Eig}(f, k)$, und sei $h \in K$. Dann ist

$$f(v + w) = f(v) + f(w) = kv + kw = k(v + w)$$

und

$$f(h \cdot v) = h \cdot f(v) = h \cdot kv = k \cdot hv.$$

Also sind sowohl $v + w$ als auch $h \cdot v$ in $\text{Eig}(f, k)$ enthalten.

- (b) folgt aus der Definition eines Eigenwerts.

- (c) ist nicht viel schwieriger:

$$\begin{aligned} \text{Eig}(f, k) &= \{v \in V \mid f(v) = k \cdot v\} = \{v \in V \mid f(v) - k \cdot v = 0\} \\ &= \{v \in V \mid (f - k \cdot \text{id})(v) = 0\} = \text{Kern}(f - k \cdot \text{id}). \end{aligned}$$

In (d) steckt die wahre Substanz dieses Hilfssatzes: Sei $B_r = \{v_1^{(r)}, v_2^{(r)}, \dots, v_{nr}^{(r)}\}$ ($r = 1, 2, \dots, s$). Sei

$$\sum_{i=1}^{n_1} h_i^{(1)} v_i^{(1)} + \sum_{i=1}^{n_2} h_i^{(2)} v_i^{(2)} + \dots + \sum_{i=1}^{n_s} h_i^{(s)} v_i^{(s)} = 0.$$

Definiere $v_r = \sum_{i=1}^{n_r} h_i^{(r)} v_i^{(r)}$ ($r = 1, 2, \dots, s$). Dann liegen die Vektoren v_r in $\text{Eig}(f, k_r)$, und es ist

$$v_1 + v_2 + \dots + v_s = 0.$$

Das Lemma über verschiedene Eigenwerte impliziert dann $v_1 = v_2 = \dots = v_s = 0$. Also ist

$$\sum_{i=1}^{n_r} h_i^{(r)} v_i^{(r)} = 0$$

für $r = 1, 2, \dots, s$. Da $\{v_1^{(r)}, v_2^{(r)}, \dots, v_{nr}^{(r)}\}$ eine Basis von $\text{Eig}(f, k_r)$ ist, muss also

$$h_1^{(r)} = h_2^{(r)} = \dots = h_{nr}^{(r)} = 0$$

sein. Da dies für alle $r = 1, 2, \dots, s$ gilt, ist der Hilfssatz damit bewiesen. \square

1. Kriterium zur Diagonalisierbarkeit

Eine lineare Abbildung von V in sich ist genau dann diagonalisierbar, wenn die Summe der Dimensionen ihrer Eigenräume mindestens n ist.

Beweis Wenn f diagonalisierbar ist, so gibt es nach dem 0. Kriterium eine Basis B aus Eigenvektoren. Da jeder Eigenvektor in einem Eigenraum liegt, ist das Erzeugnis der Eigenräume mindestens so groß wie das Erzeugnis der Basis B . Also erzeugen die Eigenräume den ganzen Vektorraum V , und somit ist die Summe der Dimensionen der Eigenräume mindestens n .

Umgekehrt möge die Summe der Dimensionen der Eigenräume mindestens n sein. Wir wählen in jedem Eigenraum eine Basis und bilden die Vereinigung B all dieser Basen. Nach dem Hilfssatz über Eigenwerte (d) ist dann B linear unabhängig. Da B nach Voraussetzung mindestens n Elemente hat, muss B eine Basis sein (und genau n Elemente haben). Wiederum nach dem 0. Kriterium ist f also diagonalisierbar. \square

Bemerkung Aus dem Beweis ergibt sich auch, dass die Summe der Dimensionen der Eigenräume gleich n ist; daher könnte man im 1. Kriterium „mindestens“ durch „gleich“ ersetzen.

Damit haben wir ein erstes brauchbares **Verfahren** erarbeitet, um die Diagonalisierbarkeit einer linearen Abbildung f zu überprüfen. Dieses Verfahren geht in drei Schritten vor.

1. *Schritt*: Bestimmung der Eigenwerte von f .
2. *Schritt*: Bestimmung der (Dimensionen der) Eigenräume.
3. *Schritt*: Anwendung des 1. Kriteriums.

Wir demonstrieren dies an folgendem *Beispiel*: Die Matrix $M = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in K^{2 \times 2}$ ist nicht diagonalisierbar.

- (1) Sei $\{v, w\}$ eine Basis von K^2 , und sei f die lineare Abbildung von K^2 in sich, deren Darstellungsmatrix bezüglich $\{v, w\}$ die Matrix M ist. Das bedeutet

$$f(v) = 0 \cdot v + 0 \cdot w = o, \quad f(w) = 1 \cdot v + 0 \cdot w = v.$$

Wir fragen uns, welche $k \in K$ Eigenwerte dieser Abbildung sein können. Wenn k ein Eigenwert ist, muss es einen Vektor $av + bw \neq o$ geben mit

$$f(av + bw) = k(av + bw).$$

$$\text{Wegen } f(av + bw) = a \cdot f(v) + b \cdot f(w) = a \cdot o + b \cdot v = bv$$

$$\text{folgt daraus } bv = f(av + bw) = k(av + bw) = kav + kbw,$$

also $b = ka$ und $0 = kb$. Daraus ergibt sich $k = 0$. (Wäre $k \neq 0$, so müsste $b = 0$, und damit auch $a = 0$ sein, im Widerspruch zu $av + bw \neq o$.) Also ist 0 der einzige Eigenwert von f .

- (2) Der einzige Eigenraum, den wir untersuchen müssen, ist also der Eigenraum $\text{Eig}(f, 0)$ zum Eigenwert 0. Dieser berechnet sich als

$$\begin{aligned} \text{Eig}(f, 0) &= \{av + bw \mid a, b \in K, f(av + bw) = 0(av + bw) = o\} \\ &= \{av + bw \mid a, b \in K, bv = o\} = \{av + bw \mid a, b \in K, b = 0\} \\ &= \{av \mid a \in K\} = \langle v \rangle. \end{aligned}$$

- (3) Nach (1) ist $\text{Eig}(f, 0)$ der einzige nichttriviale Eigenraum; nach (2) hat dieser die Dimension 1. Daher ist die Summe der Dimensionen aller Eigenräume kleiner als n ; nach dem 1. Kriterium ist also f nicht diagonalisierbar.

8.3 Das charakteristische Polynom

Bislang haben wir schon mit Eigenwerten und Eigenvektoren argumentiert, wir wissen aber noch nicht, wie man diese leicht berechnen kann. Die einzige Möglichkeit, die wir bislang haben, ist systematisches Probieren. In diesem Abschnitt wird ein Kalkül vorgestellt, mit dem man Eigenwerte in gewissem Sinn automatisch berechnen kann.

Sei wieder f eine lineare Abbildung des n -dimensionalen K -Vektorraums V in sich. Der folgende Hilfssatz ist entscheidend:

Satz über Eigenwerte einer Matrix

Sei M die Darstellungsmatrix von f bezüglich einer Basis B von V . Dann gilt: Genau dann ist $k \in K$ ein Eigenwert von f , wenn gilt

$$\det(M - kE) = 0.$$

Beweis Wir wissen:

$$\begin{aligned} k &\text{ ist ein Eigenwert von } f \\ \Leftrightarrow \text{Kern}(f - k \cdot \text{id}) &\neq \{0\} \\ \Leftrightarrow f - k \cdot \text{id} &\text{ ist nicht injektiv.} \end{aligned}$$

Nun müssen wir auf Matrizen umschalten. Wie sieht die Darstellungsmatrix (bezüglich B) der linearen Abbildung $f - k \cdot \text{id}$ aus? Da M die Darstellungsmatrix von f und $k \cdot E$ die Darstellungsmatrix von $k \cdot \text{id}$ ist, ist $M - k \cdot E$ die Darstellungsmatrix von $f - k \cdot \text{id}$. Aus Abschn. 5.2 (Charakterisierung von Isomorphismen) wissen wir außerdem, dass eine lineare Abbildung genau dann injektiv ist, wenn der Rang einer Darstellungsmatrix gleich n ist.

Damit können wir obige Äquivalenzen wie folgt fortsetzen:

$$\begin{aligned} k &\text{ ist ein Eigenwert von } f \\ \Leftrightarrow f - k \cdot \text{id} &\text{ ist nicht injektiv} \\ \Leftrightarrow \text{Rang}(M - k \cdot E) &< n \\ \Leftrightarrow \det(M - k \cdot E) &= 0. \end{aligned}$$

□

Sei $M \in K^{n \times n}$. Sei x eine Unbestimmte über K . Dann nennt man das Polynom

$$\chi_M = \det(M - xE) \in K[x]$$

das **charakteristische Polynom** der Matrix M . (χ ist ein griechischer Buchstabe, der „chi“ ausgesprochen wird.)

Das bedeutet: Um χ_M zu berechnen, zieht man in der Matrix M von jedem Eintrag auf der Diagonale den Term x ab, und rechnet von der so erhaltenen Matrix die Determinante formal aus.

Beispiele

- (a) Das charakteristische Polynom der $n \times n$ -Nullmatrix ist $(-1)^n \cdot x^n$.
- (b) Das charakteristische Polynom der $n \times n$ -Einheitsmatrix ist $\chi_{En} = (1-x)^n$.
- (c) Das charakteristische Polynom der 3×3 -Matrix

$$M = \begin{pmatrix} 3 & 2 & -1 \\ 1 & 0 & -4 \\ 3 & 0 & 1 \end{pmatrix}$$

ist

$$\chi_M = (3-x)(-x)(1-x) - 24 - 2(1-x) - 3x = -x^3 + 4x^2 - 4x - 26.$$

Wir wollen „eigentlich“ ein charakteristisches Polynom einer linearen Abbildung (und nicht „nur“ einer Matrix!) definieren. Der einfachste Weg besteht darin, dass wir erklären: Das charakteristische Polynom einer linearen Abbildung ist (nach Definition!) das charakteristische Polynom einer Darstellungsmatrix. Aber Achtung: Hier ist wieder die Fußangel der Wohldefiniertheit verborgen! Denn wir wissen, dass eine lineare Abbildung viele Darstellungsmatrizen hat, und es könnte ja sein, dass verschiedene Darstellungsmatrizen verschiedene charakteristische Polynome haben – und das wäre eine Katastrophe. Also beeilen wir uns zu beweisen, dass diese Katastrophe nicht eintreten kann:

Charakteristisches Polynom einer linearen Abbildung

Die charakteristischen Polynome je zweier Darstellungsmatrizen derselben linearen Abbildung sind gleich.

Beweis Seien M und M' zwei Darstellungsmatrizen derselben linearen Abbildung. Diese werden mit Hilfe zweier Basen gebildet. Ist A die Matrix, welche die entsprechende Basis transformation beschreibt, so ist A invertierbar, und es gilt $M' = AMA^{-1}$. Damit folgt:

$$\begin{aligned} \chi_{M'} &= \chi_{AMA^{-1}} = \det(AMA^{-1} - xE) \\ &= \det(AMA^{-1} - xE \cdot AA^{-1}) \\ &= \det(AMA^{-1} - AxE A^{-1}) \\ &= \det(A(M - xE)A^{-1}) \\ &= \det(A) \cdot \det(M - xE) \cdot \det(A^{-1}) = \det(A) \cdot \chi_M \cdot \det(A)^{-1} = \chi_M. \end{aligned}$$

□

Somit können wir definieren: Sei f eine lineare Abbildung von V in sich. Das **charakteristische Polynom** χ_f von f ist definiert als

$$\chi_f = \chi_M$$

für irgendeine Darstellungsmatrix M von f .

Bemerkung Daraus wird auch der Name klar: χ_f ist charakteristisch für f und nicht nur für M (das heißt: χ_f hängt nur von f und nicht von der Wahl einer Basis ab).

Da $\chi_f \in K[x]$ ist, können wir ein beliebiges $k \in K$ in χ_f einsetzen. Was passiert dabei? Es ergibt sich

$$\chi_f(k) = \det(M - kE).$$

Dies wird im Beweis des folgenden Satzes benutzt:

Satz über das charakteristische Polynom.

Die Nullstellen des charakteristischen Polynoms χ_f sind genau die Eigenwerte von f .

Beweis Sei $k \in K$. Dann gilt:

$$\begin{aligned} k \text{ ist ein Eigenwert von } f \\ \Leftrightarrow \det(M - kE) &= 0 \\ \Leftrightarrow \chi_f(k) &= 0. \end{aligned}$$

□

Damit haben wir einen überzeugenden Kalkül, mit dem wir die Eigenwerte einer linearen Abbildung bestimmen können: Das charakteristische Polynom ist einfach zu bestimmen (es handelt sich um eine Formel!) und dann muss man nur noch die Nullstellen dieses Polynoms bestimmen.

Ein Wermutstropfen fällt jedoch ins Glas: Zwar sieht man die Nullstellen eines so einfachen Polynoms wie $x^2 - 1$ mit bloßem Auge, bei Polynomen größeren Grades, wie etwa $x^5 - 3x^4 + 7x^2 - 5x + 3$ wird's schwieriger. In der numerischen Mathematik lernt man, Nullstellen approximativ zu bestimmen. Die radikale Frage lautet aber: Hat überhaupt jedes Polynom eine Nullstelle? Antwort: Das kommt darauf an, und zwar auf den Körper.

Über \mathbf{R} gibt es Polynome zweiten Grades, die keine (reelle) Nullstelle haben; zum Beispiel $x^2 + 1$ oder $x^2 + 1001$. (Alle Polynome über \mathbf{R} die sich nicht weiter zerlegen lassen, haben Grad 0, 1 oder 2.) Der so genannte **Hauptsatz der Algebra** (der von Gauß bewiesen wurde und den Studierenden heute üblicherweise in der Funktionentheorie präsentiert wird) sagt, dass jedes Polynom über \mathbf{C} in lineare Faktoren (also Faktoren vom Grad 1) zerfällt. Zur komplizierten Geschichte dieses Satzes siehe [Ebbl], Kap. 4.

Wie sieht das charakteristische Polynom aus? Darüber gibt der folgende Satz Auskunft:

Satz über die Koeffizienten des charakteristischen Polynoms

Sei f eine lineare Abbildung von V in sich, und sei $M = (m_{ij})$ eine Darstellungsmatrix von f . Sei $\chi_f = \chi_M$ das charakteristische Polynom. Dann gilt:

(a) χ_f hat den Grad n :

$$\chi_f = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 .$$

(b) Es gilt

$$\begin{aligned} a_n &= (-1)^n \\ a_{n-1} &= (-1)^{n-1} (m_{11} + m_{22} + \dots + m_{nn}), \\ a_0 &= \det(M) . \end{aligned}$$

Bemerkung Man nennt $m_{11} + m_{22} + \dots + m_{nn}$ die **Spur** der Matrix M . Da das charakteristische Polynom unabhängig von der Darstellungsmatrix ist, sagt dieser Satz insbesondere, dass die Spur einer jeden Darstellungsmatrix die gleiche ist.

Beweis

(a) kann man auf viele Arten beweisen. Ein ziemlich technisches Vorgehen wäre, Induktion nach n zu machen und die Matrix nach der 1. Zeile zu entwickeln.

Einfacher ist es, sich die Leibnizsche Formel vor Augen zu führen. Bei der Berechnung der Determinante von $M - xE$ geschieht folgendes: Zu jeder Permutation wird ein Polynom berechnet; die gesuchte Determinante ergibt sich als Summe dieser Polynome. Da jedes zu einer Permutation gehörende Polynom aus jeder Zeile nur einen Faktor hat, hat dieses Polynom einen Grad $\leq n$. Sicherlich gehört zur Identität ein Polynom vom Grad n . Daher hat auch die Determinante (als Summe dieser Polynome) den Grad n .

Auch die unabhängig zu beweisende Aussage (b) impliziert (a).

(b) Wir wenden die eben geschilderte Methode nochmals in verfeinerter Form an, um mit Hilfe der Leibnizschen Formel das Polynom

$$\chi_M = \det(M - xE) = \det \begin{pmatrix} m_{11} - x & m_{12} & \dots & m_{1n} \\ m_{21} & m_{22} - x & \dots & m_{2n} \\ \vdots & \vdots & & \vdots \\ m_{n1} & m_{n2} & \dots & m_{nn} - x \end{pmatrix}$$

zu berechnen. Es ergibt sich

$$\chi_M = (m_{11} - x) \cdot (m_{22} - x) \cdot \dots \cdot (m_{nn} - x) + r ,$$

wobei der erste Summand zur identischen Permutation gehört, und r die Summe ist, die zu den Permutationen aus $S_n \setminus \{\text{id}\}$ gehört. In r treten nur Summanden auf, die höchstens $n-2$ Diagonalelemente enthalten. (Wenn ein Summand Produkt von $n-1$ Diagonalelementen ist, so entspricht er einer Permutation mit mindestens $n-1$ Fixpunkten; dies muss dann die Identität sein.) Daher ist r ein Polynom vom Grad $\leq n-2$. Durch Ausmultiplizieren des Polynoms $(m_{11} - x) \dots (m_{nn} - x)$ erhält man andererseits

$$(m_{11} - x) \dots (m_{nn} - x) = (-1)^n x^n + (-1)^{n-1} (m_{11} + \dots + m_{nn}) x^{n-1} + r_1,$$

wobei r_1 ein Polynom eines Grades $\leq n-2$ ist.

Zusammen folgt

$$\chi_M = (m_{11} - x) \cdot (m_{22} - x) \cdot \dots \cdot (m_{nn} - x) + r = (-1)^n x^n + (-1)^{n-1} (m_{11} + \dots + m_{nn}) x^{n-1} + r_2,$$

wobei r_2 ein Polynom vom Grad $\leq n-2$ ist.

Daraus folgen alle Behauptungen. Die Aussage über die Determinante von M ergibt sich, wenn man 0 in $\chi_f = \chi_M$ einsetzt:

$$\chi_M(0) = \det(M - 0 \cdot E) = \det M.$$

□

Das charakteristische Polynom von 2×2 -Matrizen Mit obigem Satz kann man das charakteristische Polynom einer 2×2 -Matrix ganz einfach bestimmen: Da dies ein Polynom zweiten Grades ist, müssen wir die Koeffizienten von x^2 , von x und das Absolutglied bestimmen. *Koeffizient von x^2 ?* Das ist $(-1)^2$, also 1. *Koeffizient von x ?* Das ist $-\text{Spur}$, also $-(m_{11} + m_{22})$. *Absolutglied?* Das ist die Determinante, also $m_{11}m_{22} - m_{21}m_{12}$. Kurz,

$$\chi_M = x^2 - (m_{11} + m_{22})x + m_{11}m_{22} - m_{21}m_{12}.$$

Ein *Beispiel* gefällig? Sei

$$M = \begin{pmatrix} 2 & 3 \\ 5 & 7 \end{pmatrix}.$$

Und nun geht's schneller als Sie lesen können: *Koeffizient von x^2 ?* Der ist sowieso 1. *Koeffizient von x ?* Der ist $-\text{Spur}$, also -9 . *Absolutglied?* Das ist die Determinante, also $14-15$, also -1 . Das charakteristische Polynom lautet also $x^2 - 9x - 1$.

Bequemer (und schneller) geht's kaum noch ...

Nun bringen wir die Vielfachheit einer Nullstelle des charakteristischen Polynoms mit der Dimension des entsprechenden Eigenraums in Verbindung.

Lemma über die Dimension eines Eigenraums

Sei k eine Nullstelle mit Vielfachheit $v = v(f, k)$ des charakteristischen Polynoms χ_f einer linearen Abbildung f . Dann gilt

$$\dim \operatorname{Eig}(f, k) \leq v(f, k) .$$

Beweis Sei $m := \dim \operatorname{Eig}(f, k)$. Wir betrachten eine Basis $\{v_1, \dots, v_m\}$ von $\operatorname{Eig}(f, k)$ und ergänzen diese zu einer Basis B von V . Dann hat die Darstellungsmatrix M von f bezüglich B die folgende Gestalt:

$$M = {}_B M_B(f) = \begin{pmatrix} k & 0 & 0 & & \\ 0 & \ddots & & & * \\ 0 & 0 & k & & \\ & & & \ddots & \\ 0 & & & & * \end{pmatrix} .$$

Es folgt

$$\chi_f = \chi_M = (k - x)^m \cdot q$$

mit einem geeigneten Polynom q . Also ist die Vielfachheit von k mindestens gleich $m = \dim \operatorname{Eig}(f, k)$. \square

2. Kriterium zur Diagonalisierbarkeit

Die lineare Abbildung f ist genau dann diagonalisierbar, wenn

- (a) χ_f über K vollständig in Linearfaktoren zerfällt, und
- (b) für jeden Eigenwert k gilt

$$\dim(\operatorname{Eig}(f, k)) = v(f, k) .$$

Beweis Sei f diagonalisierbar. Dann hat V nach dem 0. Kriterium eine Basis aus Eigenvektoren; also muss es (mit Vielfachheiten gerechnet) genau n Eigenwerte geben. Insbesondere zerfällt χ_f vollständig über K . Nach dem 1. Kriterium muss die Summe der Dimensionen der Eigenräume gleich n sein. Da auch die Summe der Vielfachheiten gleich n ist, geht dies wegen obigem Lemma nur, falls für jeden Eigenwert k

$$\dim(\operatorname{Eig}(f, k)) = v(f, k)$$

ist.

Seien umgekehrt (a) und (b) erfüllt. Aus (a) ergibt sich, dass die Anzahl der Eigenwerte (mit Vielfachheiten gerechnet) gleich n ist. Wegen (b) muss auch die Summe der Dimensionen der Eigenräume gleich n sein. Nach dem 1. Kriterium folgt also, dass f diagonalisierbar ist. \square

Aus dem obigen Satz ergibt sich ein Verfahren zur Überprüfung der Diagonalisierbarkeit einer linearen Abbildung bzw. einer Matrix.

Schritt 0. Wenn eine lineare Abbildung f auf Diagonalisierbarkeit hin untersucht werden soll, so berechnet man zunächst irgendeine Darstellungsmatrix. Nun hat man in jedem Fall mit einer Matrix M zu tun.

Schritt 1. Man bestimmt das charakteristische Polynom von M . Dies ist ein leichter Schritt.

Schritt 2. Man zerlegt, wenn möglich, das charakteristische Polynom in Linearfaktoren. Damit hat man insbesondere die Eigenwerte bestimmt. Dieser Schritt ist theoretisch sehr schwierig, macht aber in vielen praktischen Fällen wenig Schwierigkeiten.

Schritt 3. Man bestimmt die Dimension der Eigenräume. Dieser Schritt läuft auf die Lösung eines linearen Gleichungssystems hinaus, ist also prinzipiell ebenfalls einfach.

Beispiel Wir betrachten die folgende reelle 3×3 -Matrix

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

Im ersten Schritt bestimmen wir das charakteristische Polynom χ_M

$$\begin{aligned} \chi_M &= \det \begin{pmatrix} -x & 2 & -1 \\ 2 & -1-x & 1 \\ 2 & -1 & 3-x \end{pmatrix} \\ &= [-x \cdot (-1-x)(3-x) + 4 + 2] - [x + 4(3-x) - 2(-1-x)] \\ &= -x^3 + 2x^2 + 4x - 8 = -x^2(x-2) + 4(x-2) = -(x-2)^2(x+2). \end{aligned}$$

Offenbar zerfällt χ_M in Linearfaktoren (Schritt 2).

Im dritten Schritt ist nur zu überprüfen, ob $\text{Eig}(M, 2)$ die Dimension 1 oder 2 hat. Sei $v = (a, b, c)^T$ ein Eigenvektor zum Eigenwert 2 von M . Dann gilt

$$Mv = \begin{pmatrix} 0 & 2 & -1 \\ 2 & -1 & 1 \\ 2 & -1 & 3 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 2a \\ 2b \\ 2c \end{pmatrix},$$

also

$$2b - c = 2a, 2a - b + c = 2b, 2a - b + 3c = 2c.$$

Man sieht unmittelbar, dass dann $b = 0$ sein muss, und es folgt

$$\text{Eig}(M, 2) = \{(a, 0, -2a) \mid a \in \mathbb{R}\}.$$

Da dieser Eigenraum nur die Dimension 1 hat, ist M also *nicht* diagonalisierbar.

8.4 Das Minimalpolynom

In diesem Abschnitt studieren wir die Menge aller Polynome $g \in K[x]$ mit $g(f) = 0$ (Nullabbildung) für eine lineare Abbildung f beziehungsweise $g(M) = 0$ (Nullmatrix) für eine Matrix M . Wir machen uns nochmals klar, was dies bedeutet. Ist

$$g = b_m x^m + b_{m-1} x^{m-1} + \dots + b_1 x + b_0,$$

so ist

$$g(f) = b_m f^m + b_{m-1} f^{m-1} + \dots + b_1 f + b_0 \text{id}$$

und

$$g(M) = b_m M^m + b_{m-1} M^{m-1} + \dots + b_1 M + b_0 E.$$

Wir untersuchen also, für welche Polynome diese Ausdrücke Null werden, und wollen die Menge dieser Polynome möglichst gut beschreiben. Eine Frage ist, ob es überhaupt ein (vom Nullpolynom verschiedenes) Polynom g gibt mit $g(f) = 0$. Dass diese Frage mit „ja“ zu beantworten ist, folgt (neben anderen Aussagen) aus dem berühmten Satz von Cayley und Hamilton. (Man kann sich diese Aussage auch anders klar machen; siehe Übungsaufgabe 17.)

Satz von Cayley-Hamilton

Sei f eine lineare Abbildung eines n -dimensionalen K -Vektorraums in sich. Dann gilt

$$\chi_f(f) = 0.$$

Was bedeutet dies? Wir betrachten zunächst das charakteristische Polynom von f . Sei

$$\chi_f = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0.$$

Dann setzen wir in dieses die Abbildung f ein:

$$\chi_f(f) = a_n f^n + a_{n-1} f^{n-1} + \dots + a_1 f + a_0 \text{id}.$$

Die Behauptung des Satzes von Cayley-Hamilton ist, dass dies die Nullabbildung ist. Um das nachzuweisen, müssen wir also für jeden Vektor $v \in V$ zeigen, dass $\chi_f(f)(v) = o$ ist. Ausgeschrieben bedeutet dies

$$\chi_f(f)(v) = a_n f^n(v) + a_{n-1} f^{n-1}(v) + \dots + a_1 f(v) + a_0 v = o$$

für alle $v \in V$.

Um dies zu zeigen, wählen wir einen Vektor v aus V fest. O. B. d. A. ist $v \neq o$. Der Beweis besteht aus zwei Teilen.

Der erste Schritt besteht darin, dass wir die Existenz eines Unterraums U mit folgenden Eigenschaften nachweisen:

- $v \in U$,
- $f(u) \in U$ für alle $u \in U$. Das bedeutet, dass man f auch als lineare Abbildung von U in sich auffassen kann. Genauer sagt man: Die Einschränkung f_U von f auf U ist eine lineare Abbildung von U in sich. Dann kann man auch das charakteristische Polynom von f_U betrachten; für dieses soll dann gelten:
- $\chi_{f_U}(f)(v) = o$.

Angenommen, wir hätten dies schon gezeigt. Dann können wir den Beweis relativ leicht beenden. Dazu wählen wir eine Basis $\{v_1, v_2, \dots, v_m\}$ von U und ergänzen diese zu einer Basis B von V . Dann hat f bezüglich dieser Basis folgende Darstellungsmatrix:

$$M = \begin{pmatrix} M_1 & * \\ 0 & M_2 \end{pmatrix},$$

wobei M_1 die zu $\{v_1, v_2, \dots, v_m\}$ gehörige Darstellungsmatrix von f_U ist.

In Übungsaufgabe 19 wird gezeigt, dass für eine solche Kästchenmatrix gilt

$$\chi_f = \chi_M = \chi_{M_2} \chi_{M_1} = \chi_{M_2} \chi_{f_U}.$$

Dies ist eine Gleichung von Polynomen. Setzen wir darin f ein, so erhalten wir die folgende Gleichung linearer Abbildungen:

$$\chi_f(f) = \chi_M(f) = [\chi_{M_2} \chi_{f_U}](f) = \chi_{M_2}(f) \chi_{f_U}(f).$$

Wenn wir dies auf den Vektor v anwenden, erhalten wir schließlich die folgende Gleichung von Vektoren:

$$\chi_f(f)(v) = [\chi_{M_2}(f) \chi_{f_U}(f)](v) = \chi_{M_2}(f)[\chi_{f_U}(f)(v)] = \chi_{M_2}(f)[o] = o.$$

Hurra! Genau das mussten wir zeigen! Noch besteht aber kein Anlass, sich bis zur Besinnungslosigkeit zu betrinken, da wir noch fähig sein müssen, die Existenz des Unterraums U nachzuvollziehen.

Dazu betrachten wir die Vektoren

$$u_0 := v, u_1 := f(v) = f(u_0), u_2 := f(f(v)) = f(u_1), u_3 := f(u_2), \dots$$

Da V die Dimension n hat, sind höchstens n dieser Vektoren linear unabhängig. Es gibt also eine natürliche Zahl m , so dass

$$u_0, u_1, \dots, u_{m-1}$$

linear unabhängig ist, aber

$$u_0, u_1, \dots, u_{m-1}, u_m$$

linear abhängig ist. Dann gibt es Skalare k_0, k_1, \dots, k_{m-1} mit

$$u_m = k_0 u_0 + k_1 u_1 + \dots + k_{m-1} u_{m-1}. \quad (*)$$

Wir definieren $U := \langle u_0, u_1, \dots, u_{m-1} \rangle$. Dann ist U ein Unterraum der Dimension m von V , der $v = u_0$ enthält.

Wegen

$$f(u_i) = u_{i+1} \in U \quad (i = 0, 1, \dots, m-2)$$

und

$$f(u_{m-1}) = k_0 u_0 + k_1 u_1 + \dots + k_{m-1} u_{m-1} \in U$$

ist jedes Bild eines Basisvektors in U enthalten. Somit gilt $f(u) \in U$ für alle $u \in U$. Daher ist die Einschränkung f_U von f auf U eine lineare Abbildung von U in sich.

Wie sieht das charakteristische Polynom von f_U aus? Um diese Frage zu beantworten, brauchen wir eine Darstellungsmatrix. Bezüglich der Basis $\{u_0, u_1, \dots, u_{m-1}\}$ hat diese folgende Gestalt:

$$M_1 = \begin{pmatrix} 0 & 0 & & & 0 & k_0 \\ 1 & 0 & 0 & \dots & 0 & k_1 \\ 0 & 1 & 0 & \dots & 0 & k_2 \\ 0 & \dots & 1 & & & \\ \dots & & & 0 & \dots & k_{m-2} \\ 0 & & \dots & 0 & 1 & k_{m-1} \end{pmatrix}.$$

Nach Übungsaufgabe 21 ist daher

$$\chi_{f_U} = \chi_{M_1} = (-1)^m [x^m - k_{m-1}x^{m-1} - k_{m-2}x^{m-2} - \dots - k_1x - k_0].$$

Einsetzen von f liefert die folgende Gleichung von linearen Abbildungen

$$\chi_{f_U}(f) = (-1)^m [f^m - k_{m-1}f^{m-1} - k_{m-2}f^{m-2} - \dots - k_1f - k_0 \text{id}].$$

Um zu sehen, dass die lineare Abbildung auf der rechten Seite den Vektor v auf den Nullvektor abbildet, interpretieren wir (*) neu:

$$\begin{aligned} f^m(v) &= u_m = k_0 u_0 + k_1 u_1 + k_2 u_2 + \dots + k_{m-1} u_{m-1} \\ &= k_0 \text{id}(v) + k_1 f(v) + k_2 f^2(v) + \dots + k_{m-1} f^{m-1}(v) . \end{aligned}$$

Damit ergibt sich

$$\begin{aligned} \chi_{f_U}(f)(v) &= (-1)^m [f^m - k_{m-1} f^{m-1} - k_{m-2} f^{m-2} - \dots - k_1 f - k_0 \text{id}](v) \\ &= (-1)^m [f^m(v) - k_{m-1} f^{m-1}(v) - k_{m-2} f^{m-2}(v) - \dots - k_1 f(v) - k_0 \text{id}(v)] \\ &= (-1)^m o = o . \end{aligned}$$

Und damit ist der Satz von Cayley-Hamilton endgültig bewiesen.

Kompliment für Ihr Durchhaltevermögen! □

Der Mathematiker Arthur Cayley (1821–1895) war ein bedeutender Algebraiker, der als Schöpfer der Matrizenrechnung gilt.

Wir betrachten jetzt die Menge aller Polynome g aus $K[x]$, für die $g(f)$ die Nullabbildung ist:

$$J = J(f) := \{g \in K[x] \mid g(f) = 0\} .$$

Diese Menge nimmt eine Schlüsselstellung ein, weil wir mit ihr das Minimalpolynom von f definieren können. Die Hauptsache ist zunächst die, dass J nicht nur das Nullpolynom umfasst:

Hilfssatz

- (a) Das charakteristische Polynom von f ist in $J(f)$ enthalten.
- (b) $J(f)$ ist ein Ideal von $K[x]$.

Beweis

- (a) Ist genau die Aussage des Satzes von Cayley-Hamilton.
- (b) Wir rechnen die definierenden Eigenschaften eines Ideals nach.

Seien $g, h \in J(f)$. Da das Einsetzen von f ein Homomorphismus ist, folgt

$$(g + h)(f) = g(f) + h(f) = 0 + 0 = 0 .$$

Sei $g \in J(f)$ und $h \in K[x]$ beliebig. Da das Einsetzen von f ein Homomorphismus ist, folgt

$$(gh)(f) = g(f) \cdot h(f) = 0 \cdot h(f) = 0 .$$

Also ist $J(f)$ tatsächlich ein Ideal. □

Nun kommt der entscheidende Satz, mit Hilfe dessen wir das Minimalpolynom einer linearen Abbildung definieren können.

Satz über das Minimalpolynom

- (a) Zu jeder linearen Abbildung f gibt es genau ein Polynom μ_f mit folgenden Eigenschaften:
- μ_f erzeugt $J(f)$; das heißt $J(f) = \{\mu_f g \mid g \in K[x]\}$,
 - der höchste Koeffizient von μ_f ist 1 (man sagt auch: μ_f ist **normiert**).
- (b) Ferner gilt: Das Polynom von μ_f teilt das charakteristische Polynom von f .

Definition Man nennt μ_f das **Minimalpolynom** von f . (μ ist ein griechischer Buchstabe, der „mü“ ausgesprochen wird.). Das Minimalpolynom ist also definiert als das normierte Polynom, das $J(f)$ erzeugt.

Beweis des Satzes über das Minimalpolynom. (a) *Existenz*: Da jedes Ideal von $K[x]$ ein Hauptideal ist, ist auch $J(f)$ ein Hauptideal. Daher gibt es ein Polynom μ_f , so dass $J(f) = \{\mu_f g \mid g \in K[x]\}$ ist. Indem wir μ_f mit einem geeigneten Körperelement multiplizieren, können wir erreichen, dass der höchste Koeffizient von μ_f gleich 1 ist.

Eindeutigkeit Seien μ und μ' zwei Polynome mit den in der Aussage des Satzes beschriebenen Eigenschaften. Da μ und μ' beide in $J(f)$ sind und den gleichen höchsten Koeffizienten haben, ist $\mu - \mu'$ ein Polynom des Ideals $J(f)$, dessen Grad kleiner als der von μ ist. Wegen der Minimalität des Grads von μ folgt daraus, dass $\mu - \mu'$ das Nullpolynom ist. Also ist $\mu' = \mu$.

(b) Nach dem Satz von Cayley-Hamilton ist χ_f in $J(f)$ enthalten. Da μ_f das Ideal $J(f)$ erzeugt, ist χ_f ein Vielfaches von μ_f . \square

Beispiele

- (a) Die Matrix $M = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ hat das charakteristische Polynom $\chi_M = x^2$ und das Minimalpolynom $\mu_M = x$, da M die Nullmatrix ist. Allgemein hat die Nullabbildung stets das Minimalpolynom x .
- (b) Die Matrix $M = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ hat das charakteristische Polynom $\chi_M = x^2$ und auch das Minimalpolynom $\mu_M = x^2$; denn x kann nicht das Minimalpolynom sein.
- (c) Die Matrix $M = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ hat das charakteristische Polynom $\chi_M = (1-x)^2$ und das Minimalpolynom $\mu_M = x-1$; denn $M-E$ ist die Nullmatrix. Allgemein gilt: Die Einheitsmatrix (bzw. identische Abbildung) hat stets das Minimalpolynom $x-1$.

Bemerkung Das Minimalpolynom ist im Allgemeinen längst nicht so einfach zu bestimmen wie das charakteristische Polynom: Wenn man nur die Definition benutzt, ist es fast hoffnungslos. Hier hilft der Satz von Cayley-Hamilton bzw. Teil (b) des Satzes über das Minimalpolynom. Er sagt, dass man nicht mehr alle Polynome testen muss, sondern nur noch die Teiler des charakteristischen Polynoms – und dies sind nur ganz wenige Polynome. Damit hat man eine sehr starke obere Abschätzung für den Aufwand, das Minimalpolynom zu bestimmen.

Wenn zum Beispiel $\chi_f = (x-2)^{10}(x+5)^3(x+1)^2$ ist, so muss man nur $11 \cdot 4 \cdot 3 = 132$ Polynome testen, nämlich die Polynome $(x-2)^i(x+5)^j(x+1)^k$ mit $0 \leq i \leq 10, 0 \leq j \leq 3, 0 \leq k \leq 2$.

Wir wissen, dass das Minimalpolynom das charakteristische Polynom teilt, es kann aber nicht ein beliebiger Teiler des charakteristischen Polynoms sein. Vielmehr gilt folgende „untere Abschätzung“ für das Minimalpolynom:

Satz über die Nullstellen des Minimalpolynoms

Für jede lineare Abbildung gilt: Die Nullstellen des Minimalpolynoms sind genau die Nullstellen des charakteristischen Polynoms. (Die Nullstellen des Minimalpolynoms können aber eine kleinere Vielfachheit haben.)

Beweis Wir müssen nur zeigen, dass jede Nullstelle von χ_f auch eine Nullstelle von μ_f ist. Mit anderen Worten: Wir müssen zeigen, dass jeder beliebige Eigenwert k von f eine Nullstelle von μ_f ist. Sei

$$\mu_f = x^m + a_{m-1}x^{m-1} + \dots + a_1x + a_0.$$

Da $\mu_f(f) = 0$ ist, gilt also

$$0 = \mu_f(f) = f^m + a_{m-1}f^{m-1} + \dots + a_1f + a_0 \text{id}.$$

Wenn wir beide Seiten auf einen Eigenvektor v zum Eigenwert k anwenden, erhalten wir

$$\begin{aligned} 0 &= \mu_f(f)(v) = f^m(v) + a_{m-1}f^{m-1}(v) + \dots + a_1f(v) + a_0 \text{id}(v) \\ &= k^m \cdot v + a_{m-1}k^{m-1} \cdot v + \dots + a_1k \cdot v + a_0 \cdot v = \mu_f(k) \cdot v. \end{aligned}$$

Wegen $v \neq 0$ folgt daraus $\mu_f(k) = 0$. □

Beispiel Oben haben wir gesehen, dass man höchstens 132 Polynome testen muss, um das zu $\chi_f = (x-2)^{10}(x+5)^3(x+1)^2$ gehörige Minimalpolynom zu bestimmen. Der Satz über die Nullstellen des Minimalpolynoms sagt, dass es noch besser geht: Man kann sich auf die Polynome $(x-2)^i(x+5)^j(x+1)^k$ mit $1 \leq i \leq 10, 1 \leq j \leq 3, 1 \leq k \leq 2$ beschränken; dies sind schlappe $10 \cdot 3 \cdot 2 = 60$ Polynome.

Wir beenden diesen Abschnitt mit dem dritten (und letzten) Kriterium zur Diagonalisierbarkeit.

3. Kriterium zur Diagonalisierbarkeit

Eine lineare Abbildung f von V in sich ist genau dann diagonalisierbar, wenn das Minimalpolynom μ_f vollständig in paarweise verschiedene Linearfaktoren zerfällt.

Vor dem Beweis schildern wir ein Verfahren, das auf diesem Kriterium beruht.

Test auf Diagonalisierbarkeit einer linearen Abbildung f bzw. einer Matrix M Der Test erfolgt in drei Schritten, von denen uns die beiden ersten wohlbekannt sind.

1. *Schritt* Überprüfe, ob das charakteristische Polynom vollständig in Linearfaktoren zerfällt. Wenn nein, ist die Abbildung nicht diagonalisierbar.
2. *Schritt* Bestimme die Eigenwerte von f ; seien diese k_1, k_2, \dots, k_m .
3. *Schritt* Überprüfe, ob $(x-k_1) \cdot (x-k_2) \dots (x-k_m)$ das Minimalpolynom von f bzw. M ist. Dazu ist nur zu überprüfen, ob

$$(x - k_1) \cdot (x - k_2) \cdot \dots \cdot (x - k_m)(f) = (f - k_1 \text{id}) \cdot (f - k_2 \text{id}) \cdot \dots \cdot (f - k_m \text{id})$$

bzw.

$$(x - k_1) \cdot (x - k_2) \cdot \dots \cdot (x - k_m)(M) = (M - k_1 E) \cdot (M - k_2 E) \cdot \dots \cdot (M - k_m E)$$

die Nullabbildung bzw. die Nullmatrix ist.

Bemerkung Um dieses Kriterium anzuwenden, muss man nicht die (schwierige!) Aufgabe lösen, das Minimalpolynom zu bestimmen, sondern man muss nur überprüfen, ob ein bestimmtes Polynom das Minimalpolynom ist. Falls die Antwort „ja“ ist, weiß man, dass die lineare Abbildung diagonalisierbar ist; wenn die Antwort allerdings negativ ausfällt, weiß man nicht mehr, als dass die lineare Abbildung nicht diagonalisierbar ist!

Nun zum *Beweis* des Kriteriums. Zunächst sei f als diagonalisierbar vorausgesetzt. Seien k_1, k_2, \dots, k_m die verschiedenen Eigenwerte von f . Es ist zu zeigen, dass folgendes gilt:

$$\mu_f = (x - k_1) \cdot (x - k_2) \cdot \dots \cdot (x - k_m).$$

Da f diagonalisierbar ist, ist nach dem 1. Kriterium die Vereinigung der Basen der Eigenräume eine Basis von V . Mit anderen Worten bedeutet dies, dass sich jeder Vektor $v \in V$ schreiben lässt als

$$v = v_1 + v_2 + \dots + v_m$$

mit eindeutig bestimmten Eigenvektoren v_i zu k_i ($i = 1, \dots, m$).

Sei $g(x) = (x-k_1)(x-k_2) \dots (x-k_m)$. Es ist zu zeigen, dass $g = \mu_f$ ist. Dazu betrachten wir die folgenden Teiler dieses Polynoms: $g_i(x) = (x-k_1)(x-k_2) \dots (x-k_{i-1})(x-k_{i+1}) \dots (x-k_m)$ für $i = 1, \dots, m$. Es gilt also $g(x) = g_i(x)(x-k_i)$.

Wir setzen nun die lineare Abbildung f in das Polynom g ein und wenden die so erhaltene Abbildung auf den Vektor v_i an:

$$g(f)(v_i) = g_i(f)[f - k_i \text{id}](v_i) = g_i(f)[f(v_i) - k_i v_i] = g_i(f)[0] = 0.$$

Jetzt wenden wir die Abbildung $g(f)$ auf den Vektor v an und erhalten

$$g(f)(v) = g(f)[v_1 + v_2 + \dots + v_m] = g(f)[v_1] + g(f)[v_2] + \dots + g(f)[v_m] = 0.$$

Somit ist $g(f)$ die Nullabbildung, also teilt μ_f das Polynom g . Da aber wegen des vorigen Satzes das Polynom $g = (x-k_1) \cdot (x-k_2) \dots (x-k_m)$ bestimmt ein Teiler des Minimalpolynoms ist, folgt in der Tat

$$\mu_f = g = (x - k_1) \cdot (x - k_2) \cdot \dots \cdot (x - k_m).$$

Nun zur Umkehrung. Nach Voraussetzung ist

$$\mu_f = (x - k_1) \cdot (x - k_2) \cdot \dots \cdot (x - k_m)$$

mit verschiedenen Elementen $k_1, k_2, \dots, k_m \in K$. Wir beweisen die Aussage durch Induktion nach n (der Dimension von V).

Ist $n = 1$, so folgt aus $\mu_f = (x-k_1)$, dass f den Eigenwert k_1 hat, und f ist diagonalisierbar.

Sei nun $n > 1$, und sei die Aussage richtig für alle natürlichen Zahlen $m < n$.

Die entscheidende Behauptung ist jetzt die, dass *die beiden Unterräume Kern($f - k_1 \text{id}$) und Bild($f - k_1 \text{id}$) komplementär sind*.

Um dies zu zeigen, weisen wir zunächst nach, dass V von diesen beiden Unterräumen erzeugt wird. Dazu dividieren wir das Polynom $(x-k_2) \dots (x-k_m)$ mit Rest durch $(x-k_1)$; wir erhalten

$$(x - k_2) \cdot \dots \cdot (x - k_m) = q(x - k_1) + r,$$

wobei $\text{Grad}(r) < 1$ sein muss. Also ist r ein Körperelement, das $\neq 0$ ist, da $k_1 \neq k_2, \dots, k_m$ ist. Wir setzen f ein und sehen

$$(f - k_2 \cdot \text{id}) \cdot \dots \cdot (f - k_m \cdot \text{id}) - q(f)(f - k_1 \cdot \text{id}) = r \cdot \text{id}.$$

Sei nun v ein beliebiger Vektor aus V . Wenn wir die eben erhaltene lineare Abbildung auf v anwenden, ergibt sich

$$r \cdot v = (f - k_2 \cdot \text{id}) \cdot \dots \cdot (f - k_m \cdot \text{id})(v) - q(f)(f - k_1 \cdot \text{id})(v).$$

Wenn man auf den ersten Summanden, also auf $(f - k_2 \cdot \text{id}) \dots (f - k_m \cdot \text{id})(v)$, die lineare Abbildung $f - k_1 \cdot \text{id}$ anwendet, erhält man wie oben den Nullvektor; denn nach Voraussetzung ist

$$(f - k_1 \cdot \text{id}) \cdot (f - k_2 \cdot \text{id}) \cdot \dots \cdot (f - k_m \cdot \text{id}) = (x - k_1) \cdot (x - k_2) \cdot \dots \cdot (x - k_m)(f) = \mu_f(f)$$

die Nullabbildung. Also liegt dieser Summand in $\text{Kern}(f - k_1 \text{id})$. Der zweite Summand $q(f)(f - k_1 \cdot \text{id})(v) = (f - k_1 \cdot \text{id})q(f)(v)$ liegt offenbar in $\text{Bild}(f - k_1 \text{id})$.

Also wird V von $\text{Kern}(f - k_1 \text{id})$ und $\text{Bild}(f - k_1 \text{id})$ erzeugt. Da wir schon längst wissen (Dimensionsformel für lineare Abbildungen), dass die Summe der Dimensionen von Kern und Bild gleich n ist, haben wir damit die Zwischenbehauptung bewiesen.

Daraus folgt der Rest relativ einfach. Da k_1 ein Eigenwert ist, ist $\dim(\text{Kern}(f - k_1 \text{id})) > 0$, also $\dim(\text{Bild}(f - k_1 \text{id})) < n$. Daher können wir jetzt die Induktionsvoraussetzung auf $U := \text{Bild}(f - k_1 \text{id})$ anwenden.

Da $f(U) \subseteq U$ gilt, induziert f auch eine lineare Abbildung f_U von U in sich. Ferner ist das Minimalpolynom von f_U ein Teiler des Minimalpolynoms von f . Daraus ergibt sich insbesondere, dass das Minimalpolynom von f_U in paarweise verschiedene Linearfaktoren zerfällt. Nach Induktion ist also f_U diagonalisierbar.

Sei v_1, \dots, v_s eine Basis von U aus Eigenvektoren von f_U , und sei v_{s+1}, \dots, v_n eine Basis von $\text{Kern}(f - k_1 \text{id}) = \text{Eig}(f, k_1)$. Da U und $\text{Kern}(f - k_1 \text{id})$ komplementär sind, ist $v_1, \dots, v_s, v_{s+1}, \dots, v_n$ eine Basis von V aus Eigenvektoren von f . Nach dem 0. Kriterium ist f also diagonalisierbar. \square

In vielen Situation (und noch mehr Übungsaufgaben) wird nach linearen Abbildungen f bzw. Matrizen M gefragt, die einer gewissen Gleichung genügen. Wir betrachten ein *Beispiel*.

Wir fragen nach allen linearen Abbildungen f eines reellen Vektorraums in sich, die der Gleichung $f^2 = 3f$ genügen. Diese Aufgabe kann man im Prinzip dadurch lösen, dass man die definierende Gleichung zunächst wie folgt umschreibt

$$f^2 - 3f = 0$$

und dann beobachtet, dass offenbar genau nach denjenigen linearen Abbildungen f gefragt wird, welche die Nullabbildung ergeben, wenn man sie in das Polynom $x^2 - 3x$ einsetzt.

Das bedeutet, dass das Polynom $x^2 - 3x$ in dem Ideal $J(f)$ liegt; mit anderen Worten: Das Minimalpolynom μ_f teilt das Polynom $x^2 - 3x = x(x - 3)$.

Also gibt es drei Möglichkeiten, nämlich $\mu_f = x$, $\mu_f = x - 3$ und $\mu_f = x(x - 3)$. In jedem Fall zerfällt das Minimalpolynom in verschiedene Linearfaktoren. Also ist f diagonalisierbar; im Fall $\mu_f = x(x - 3)$ muss man nur noch die Vielfachheiten der Eigenwerte 0 und 3 bestimmen; dies kann man zum Beispiel dadurch machen, dass man das charakteristische Polynom ausrechnet.

Abschließende Frage: Wenn nun aber die betrachtete lineare Abbildung nicht diagonalisierbar ist? Was dann? Müssen wir dann die Flinte ins Korn werfen und einfach aufgeben?

Oder kann man da noch etwas machen? Man kann tatsächlich auch „allgemeine“ Matrizen noch sehr gut darstellen; dies geschieht mit Hilfe der „Jordanschen Normalform“. Diese wird in vielen Büchern über lineare Algebra an zentraler Stelle behandelt. Wenn Sie darüber Genaueres wissen wollen, so sollten Sie die entsprechenden Abschnitte in den im Literaturverzeichnis aufgelisteten Büchern über lineare Algebra lesen. Wenn Sie den Stoff bis hierher verstanden haben, wird Ihnen das auch keine Schwierigkeiten bereiten.

8.5 Verständnisfragen, Übungen und Tipps

Richtig oder falsch?

1. Thema: Eigenwerte und Eigenvektoren

Sei f eine lineare Abbildung eines n -dimensionalen K -Vektorraums in sich.

- ☐ Das Nullelement von K kann ein Eigenwert von f sein.
- ☐ Der Nullvektor kann ein Eigenvektor von f sein.
- ☐ Der Nullvektor kann ein Eigenwert von f sein.
- ☐ Der Nullvektor ist der triviale Eigenvektor.
- ☐ Die lineare Abbildung f hat höchstens n verschiedene Eigenwerte.
- ☐ Die lineare Abbildung f kann $n + 1$ Eigenwerte haben.
- ☐ Jede lineare Abbildung hat einen Eigenwert.
- ☐ Jede lineare Abbildung hat einen Eigenvektor.
- ☐ 0 ist der triviale Eigenwert.
- ☐ Hat f die 0 als Eigenwert, so ist f die Nullabbildung.

2. Thema: Eigenvektoren

Sei f eine lineare Abbildung eines K -Vektorraums in sich, für die es einen Vektor $v \neq 0$ gibt mit $f(v) = kv$. Dann ist

- ☐ $-v$ ein Eigenvektor zum Eigenwert k ,
- ☐ $+v$ ein Eigenvektor zum Eigenwert $-k$,
- ☐ $-v$ ein Eigenvektor zum Eigenwert $-k$,
- ☐ $+v$ ein Eigenvektor zum Eigenwert k ,
- ☐ $k = 0$.

3. Thema: Diagonalisierbarkeit

- ☐ Jede Matrix ist diagonalisierbar.
- ☐ Jede $m \times n$ -Matrix ist diagonalisierbar.
- ☐ Eine $m \times n$ -Matrix ist genau dann diagonalisierbar, wenn $m = n$ ist.
- ☐ Jede Matrix hat n Eigenwerte.
- ☐ Wenn eine $n \times n$ -Matrix verschiedene Eigenwerte hat, so ist sie diagonalisierbar.
- ☐ Wenn eine $n \times n$ -Matrix n verschiedene Eigenwerte hat, so ist sie diagonalisierbar.
- ☐ Wenn eine $n \times n$ -Matrix n verschiedene nichttriviale Eigenwerte hat, so ist sie diagonalisierbar.

4. Thema: Das charakteristische Polynom

- ☐ Das charakteristische Polynom zerfällt stets in Linearfaktoren.
- ☐ Jede komplexe $n \times n$ -Matrix ist diagonalisierbar.
- ☐ Jede reelle $n \times n$ -Matrix, deren charakteristisches Polynom zerfällt, ist diagonalisierbar.
- ☐ Jede 2×2 -Matrix über $\text{GF}(2)$ ist diagonalisierbar.
- ☐ Jede komplexe $n \times n$ -Matrix hat mindestens n verschiedene Eigenwerte.
- ☐ Es gibt Körper, über denen manche Polynome in Linearfaktoren zerfallen und manche nicht.
- ☐ Über jedem Körper gibt es Polynome vom Grad 1000, die in Linearfaktoren zerfallen.

5. Thema: Das Minimalpolynom

- ☐ Das Minimalpolynom ist dadurch festgelegt, dass es das charakteristische Polynom teilt.
- ☐ Das charakteristische Polynom teilt das Minimalpolynom.
- ☐ Das Minimalpolynom teilt alle Polynome.
- ☐ Wenn ein Polynom das charakteristische Polynom teilt, ist es das Minimalpolynom.
- ☐ Das Minimalpolynom einer linearen Abbildung kann das Nullpolynom sein.
- ☐ Das Minimalpolynom einer linearen Abbildung kann ein konstantes Polynom sein.
- ☐ Das vom charakteristischen Polynom erzeugte Ideal von $K[x]$ ist in dem vom Minimalpolynom erzeugten Ideal enthalten.
- ☐ Jede Nullstelle des Minimalpolynoms hat Vielfachheit 1.
- ☐ Das Minimalpolynom ist stets irreduzibel.
- ☐ Das Minimalpolynom einer diagonalisierbaren linearen Abbildung ist irreduzibel.

Übungsaufgaben

1. Sei $f: V \rightarrow W$ eine lineare Abbildung, und sei B eine Basis von V . Zeigen Sie: Wenn f injektiv ist, dann gibt es eine Basis B' von W , so dass bei geeigneter Nummerierung von B die Darstellungsmatrix folgende Gestalt hat:

$${}_B M_{B'}(f) = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & \ddots & & & 0 \\ 0 & & 1 & & 0 \\ \vdots & & & 0 & \vdots \\ 0 & 0 & 0 & \dots & \ddots \end{pmatrix}.$$

2. Zeigen Sie: Die Eigenwerte einer $n \times n$ -Matrix M sind genau die Elemente $k \in K$, für die es einen von Null verschiedenen Spaltenvektor x gibt mit $Mx = kx$

3. Bestimmen Sie die Eigenwerte und je einen zugehörigen Eigenvektor der folgendermaßen definierten linearen Abbildung f eines 3-dimensionalen reellen Vektorraums mit Basis $\{v_1, v_2, v_3\}$:

$$f(v_1) := 5/2v_1 + 2v_2 + 1/2v_3$$

$$f(v_2) := 5v_1 + 4v_2 - 2v_3$$

$$f(v_3) := -7/2v_1 - 2v_2 - 3/2v_3.$$

4. Geben Sie eine reelle 2×2 -Matrix an, die keine Eigenwerte hat.
 5. (a) Seien $M, A \in K^{n \times n}$, und sei A regulär. Dann gilt:

$$k \text{ ist ein Eigenwert von } M \Leftrightarrow k \text{ ist ein Eigenwert von } AMA^{-1}$$

- (b) Begründen Sie damit, warum die Definition eines Eigenwerts einer Matrix unabhängig von der Auswahl einer Basis ist.

6. Bestimmen Sie Eigenwerte und Eigenräume der folgenden reellen Matrix:

$$M = \begin{pmatrix} 3 & 4 & -3 \\ 2 & 7 & -4 \\ 3 & 9 & -5 \end{pmatrix}.$$

7. Bestimmen Sie Eigenwerte und Eigenräume der folgenden reellen Matrizen:

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

8. (a) Bestimmen Sie das charakteristische Polynom der Matrix $\begin{pmatrix} 0 & 0 & a \\ 1 & 0 & b \\ 0 & 1 & c \end{pmatrix}$.

- (b) Benützen Sie dieses Ergebnis, um eine 3×3 -Matrix über $\text{GF}(3)$ zu finden, die keine Eigenwerte hat.

9. Sei f eine invertierbare lineare Abbildung. Angenommen, f hat den Eigenwert k . Dann hat auch f^{-1} einen Eigenwert, nämlich ...

10. Zeigen Sie:

- (a) Wenn die Matrix M den Eigenwert k hat, so hat M^2 den Eigenwert k^2 .

- (b) Verallgemeinern Sie diesen Sachverhalt.

- (c) Machen Sie sich durch eine 2×2 -Matrix klar, dass die Umkehrung von (a) nicht gilt.

11. Sei f eine lineare Abbildung eines Vektorraums in sich mit $f^2 = f$. Zeigen Sie, dass 0 und 1 die einzigen möglichen Eigenwerte von f sind.

[Hinweis: Berechnen Sie $f^2(v)$.]

12. Zeigen Sie, dass jede reelle 2×2 -Matrix M der Form $M = \begin{pmatrix} a & b \\ b & a \end{pmatrix}$ diagonalisierbar ist.

Gilt dies auch über beliebigen Körpern K ?

13. Diagonalisieren Sie (falls dies möglich ist) die folgende reelle Matrix

$$M = \begin{pmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{pmatrix}.$$

14. Geben Sie eine Matrix mit ganzzahligen Einträgen an, die als Matrix über \mathbf{R} diagonalisierbar ist, nicht aber als Matrix über \mathbf{Q} .
15. Zeigen Sie: Eine $n \times n$ -Matrix, in der die Summe der Elemente einer jeden Zeile konstant ist („konstante Zeilensumme“), hat mindestens einen Eigenwert. Geben Sie einen zugehörigen Eigenvektor an.

Gilt eine entsprechende Aussage für Matrizen mit konstanter Spaltensumme?

16. Sei M die folgendermaßen definierte 4×4 -Matrix:

$$M = \begin{pmatrix} a & b & c & d \\ d & a & b & c \\ c & d & a & b \\ b & c & d & a \end{pmatrix} \quad (a, b, c, d \in \mathbf{R}),$$

die wir als Matrix aus $\mathbf{C}^{4 \times 4}$ auffassen.

Zeigen Sie, dass M die Eigenwerte $a + b + c + d$, $a - b + c - d$, $a + bi - c - di$, $a - bi - c + di$ hat (wobei i die imaginäre Einheit ist). Geben Sie jeweils einen Eigenvektor an.

17. Sei $M \in K^{n \times n}$. Zeigen Sie, ohne den Satz von Cayley-Hamilton zu verwenden: Es gibt ein Polynom (verschieden vom Nullpolynom) in $K[x]$, das M als Nullstelle hat. [Hinweis: Benutzen Sie, dass die Matrizen E, M, M^2, M^3, \dots im K -Vektorraum aller $n \times n$ -Matrizen nicht linear unabhängig sein können.]

Können Sie etwas über den (kleinstmöglichen) Grad dieses Polynoms aussagen?

18. Sei f eine lineare Abbildung des Vektorraums \mathbf{C}^2 in sich. Zeigen Sie: Wenn $f^2 = 0$ (Nullabbildung) ist, dann gilt $f = 0$, oder f hat eine Darstellungsmatrix $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$.

19. Sei M eine $n \times n$ -Matrix, die sich in folgender **Kästchenform** schreiben lässt:

$$M = \begin{pmatrix} A & * \\ 0 & B \end{pmatrix},$$

wobei A eine $a \times a$ -Matrix und B eine $b \times b$ -Matrix ist. Dann gilt für das charakteristische Polynom $\chi_M = \chi_A \cdot \chi_B$.

20. Suchen (und finden) Sie eine Matrix aus $\mathbf{C}^{4 \times 4}$ mit charakteristischem Polynom $(x-2)^2(x-3)^2$ und Minimalpolynom $(x-2)(x-3)^2$.

21. Bestimmen Sie das charakteristische Polynom der folgenden Matrix $M \in K^{n \times n}$:

$$M = \begin{pmatrix} 0 & 0 & \cdots s & \cdots s & 0 & k_0 \\ 1 & 0 & 0 & \cdots s & 0 & k_1 \\ 0 & 1 & 0 & \cdots s & 0 & k_2 \\ 0 & 0 & 1 & 0 & \cdots s & \cdots s \\ 0 & & \cdots s & 1 & 0 & k_{n-2} \\ 0 & \cdots s & \cdots s & 0 & 1 & k_{n-1} \end{pmatrix}.$$

22. Bestimmen Sie die Minimalpolynome der Matrizen aus den Aufgaben 6 und 7.

23. Bestimmen Sie das Minimalpolynom der folgenden Matrix M :

$$M = \begin{pmatrix} 0 & 0 & \cdots s & 0 & 0 \\ 1 & 0 & 0 & \cdots s & 0 \\ 0 & 1 & 0 & \cdots s & 0 \\ 0 & \cdots s & 1 & 0 & 0 \\ 0 & \cdots s & 0 & 1 & 0 \end{pmatrix}.$$

24. Sei f eine lineare Abbildung von V in sich, und sei v ein Vektor aus V . Zeigen Sie: Die Menge aller Polynome g mit $g(f)(v) = 0$ ist ein Ideal $J(v)$ von $K[x]$. Ferner gibt es ein eindeutig bestimmtes Polynom μ , dessen höchster Koeffizient 1 ist und das alle Polynome g aus $J(v)$ teilt.

Projekt: Drehungen

1. Machen Sie sich klar, dass durch die Matrix

$$D = D(\alpha) = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$$

eine Drehung der reellen Ebene \mathbf{R}^2 um den Winkel α mit dem Punkt $(0, 0)$ als Zentrum beschrieben wird.

2. (a) Bestimmen Sie das charakteristische Polynom der Matrix $D(\alpha)$.
 (b) Für welche Winkel α hat die Matrix $D(\alpha)$ (reelle) Eigenwerte? Interpretieren Sie Ihre Antwort geometrisch.
3. (a) Was ist eine Drehung eines dreidimensionalen Raumes? (Machen Sie sich die Begriffe „Achse“ und „Drehwinkel“ klar.)
 (b) Durch welche Matrizen werden Drehungen im affinen Raum \mathbf{R}^3 beschrieben?
 (c) Welche Eigenwerte und Eigenräume hat eine solche Matrix?
4. Kommentieren Sie folgende Aussage mit Ihrem mathematischen Sachverstand:
 Beim DFB-Pokalspiel Dynamo Dresden gegen FC Bayern München am 9.11.1993 gab es zum Zeitpunkt des Anpfiffs der zweiten Halbzeit mindestens einen Punkt des Fußballs, der an genau der gleichen Stelle war wie zu Beginn der ersten Halbzeit.

Sie sollten mit folgenden Begriffen umgehen können

Diagonalisierbarkeit, Diagonalmatrix, Eigenwert, Eigenvektor, Eigenraum, charakteristisches Polynom, Minimalpolynom

Was sagen Sie dazu?

Gnutpuaheb Es gilt der Satz von Cayley-Hamilton.

Sieweb Wir setzen M in $\det(m - xE)$ ein und erhalten „ganz einfach“:

$$\chi_M(M) = \det(M - xE)(M) = \det(M - ME) = \det(M - M) = \det(0) = 0.$$



UNTER DEM WURZELBAUM

Gruppen spielen in der modernen Mathematik eine wesentliche Rolle, und zwar sowohl in den Grundlagen als auch in den Anwendungen.

Weshalb haben wir dann nicht schon längst Gruppen behandelt? Der Grund ist ganz einfach: Wir sind in den ersten Kapiteln direkt auf die in der linearen Algebra zentralen Strukturen, nämlich Körper und Vektorräume eingegangen. Das hätte man durch Einführung des Begriffs der Gruppe als „Zwischenbegriff“ auch „systematischer“ machen können. Wir haben aber im bisherigen Verlauf der linearen Algebra schon Gruppen in Hülle und Fülle gesehen – sie nur noch nicht so genannt. Deshalb wird nach der Definition eine ganze Latte von Ihnen „schon bekannten“ Beispielen folgen. Dadurch werden wir die entsprechenden Strukturen nochmals „vom höheren Standpunkt aus“ kennen lernen.

9.1 Beispiele von Gruppen

Eine **Gruppe** besteht aus einer Menge G zusammen mit einer Verknüpfung \cdot (die je zwei Elementen aus G wieder ein Element aus G zuordnet), so dass folgende Axiome erfüllt sind:

(G1) *Assoziativität*: Die Verknüpfung \cdot ist assoziativ:

$$(g \cdot h) \cdot k = g \cdot (h \cdot k) \text{ für alle } g, h, k \in G .$$

(G2) *Existenz eines **neutralen Elements***: Es gibt ein Element aus G , das wir e nennen, für das gilt:

$$e \cdot g = g \text{ für alle } g \in G .$$

(G3) *Existenz inverser Elemente*: Für jedes $g \in G$ gibt es ein Element aus G , das wir g^{-1} nennen, für das gilt:

$$g^{-1} \cdot g = e ,$$

wobei e das in (G2) geforderte neutrale Element ist. Wenn man ganz vorsichtig ist, könnte man g^{-1} ein zu g **linksinverses** Element nennen. Wir werden später zeigen, dass g^{-1} auch rechtsinvers ist.

Man sagt oft, dass G eine Gruppe ist; wenn man die Verknüpfung betonen möchte, sagt man, dass G bezüglich der Verknüpfung \cdot eine Gruppe ist und schreibt dann (G, \cdot) .

Wenn zusätzlich das folgende Axiom (G4) gilt, nennt man eine Gruppe G (**kommutativ** oder) **abelsch**:

(G4) *Kommutativität*: Für je zwei Elemente $g, h \in G$ gilt

$$g \cdot h = h \cdot g .$$

In einer Gruppe gibt es also nur eine Verknüpfung, die wir mit \cdot bezeichnet haben. In vielen Situationen wird es allerdings natürlicher sein, die Verknüpfung einer Gruppe additiv zu schreiben. In einer solchen Situation empfiehlt es sich auch, das neutrale Element mit 0 zu bezeichnen und das zu g inverse Element mit $-g$.

Haben Sie's gemerkt? Im letzten Satz habe ich „das“ neutrale Element und „das“ inverse Element geschrieben. Durfte ich das? Klare Antwort: Nein! ... das heißt, ich darf das, wenn ich Ihnen den folgenden Hilfssatz bewiesen habe:

Eindeutigkeit des neutralen Elements und der inversen Elemente

In jeder Gruppe gibt es nur ein neutrales Element; jedes Element einer Gruppe hat genau ein inverses Element.

Der *Beweis* erfolgt in mehreren Schritten, die auch an sich interessant sind. Sei e das nach (G2) existierende neutrale Element.

1. Es gilt $gg^{-1} = e$. Das heißt: Jedes linksinverse Element ist auch „rechtsinvers“. (Sei $h = g^{-1}$. Dann ist

$$\begin{aligned} gg^{-1} &= e(gg^{-1}) = (h^{-1}h)(gg^{-1}) = h^{-1}((hg)g^{-1}) = h^{-1}((g^{-1}g)h) \\ &= h^{-1}(eh) = h^{-1}h = e. \end{aligned}$$

2. Es gilt $ge = g$ für alle $g \in G$. Das heißt: e ist auch „rechtsneutral“.

$$(ge = g(g^{-1}g) = (gg^{-1})g = eg = g.)$$

3. Es gibt nur ein neutrales Element in G .

(Sei e^* ein Element, für das ebenfalls $e^*g = g$ für alle $g \in G$ gilt. Dann ist insbesondere $e^*e = e$. Aus 2. folgt $e^*e = e^*$. Also ist $e^* = e^*e = e$.)

4. Für jedes $g \in G$ gibt es nur ein zu g inverses Element.

(Sei h ein weiteres zu g inverses Element. Dann gilt $hg = e$. Mit 2. und 1. folgt also

$$h = he = h(gg^{-1}) = (hg)g^{-1} = eg^{-1} = g^{-1}.)$$

□

Nun aber zu der angekündigten Liste von *Beispielen von Gruppen*.

9.1.1 Gruppen in bekannten Strukturen

Die erste Sorte von Beispielen beruht auf folgender Idee: Wir kennen viele Strukturen, auf denen zwei Verknüpfungen definiert sind; wenn wir eine dieser Verknüpfungen „vergessen“, erhalten wir – unter kleinen Sicherheitsvorkehrungen – eine Gruppe.

$\mathbf{R} \setminus \{0\}$ ist bezüglich der normalen Multiplikation eine Gruppe; ebenso ist $\mathbf{Q} \setminus \{0\}$ bezüglich der normalen Multiplikation eine Gruppe. Allgemein gilt:

Multiplikative Gruppe eines Körpers

Sei K ein Körper mit Addition $+$ und Multiplikation \cdot . Dann ist die Menge $K^* = K \setminus \{0\}$ bezüglich \cdot (genauer gesagt: bezüglich der Einschränkung von \cdot auf $K \setminus \{0\}$) eine Gruppe. Man nennt K^* die **multiplikative Gruppe des Körpers K** .

Den *Nachweis* dieser Tatsache erhalten Sie einfach dadurch, dass Sie die Definition eines Körpers nachschlagen und diese mit der Definition einer Gruppe vergleichen. □

Additive Gruppe eines Rings

Sei R ein Ring mit Addition $+$ und Multiplikation \cdot . Dann ist die Menge R bezüglich $+$ eine Gruppe. Insbesondere ist jeder Körper K zusammen mit der Addition eine Gruppe (**additive Gruppe des Körpers K**). □

Zum Beispiel ist $(\mathbf{Z}, +)$ eine Gruppe. Auch die Menge der geraden Zahlen bildet eine Gruppe bezüglich der Addition, ebenso die Menge aller durch 3 teilbaren ganzen Zahlen. Allgemein gilt: Für jedes $d \in \mathbf{Z}$ ist $(d\mathbf{Z}, +)$ eine Gruppe.

Aber (\mathbf{Z}, \cdot) ist keine Gruppe, ja nicht einmal $(\mathbf{Z} \setminus \{0\}, \cdot)$.

Auch $(\mathbb{Z}_n, +)$ ist eine Gruppe, während $(\mathbb{Z}_n \setminus \{0\}, \cdot)$ nur eine Gruppe ist, wenn n eine Primzahl ist.

Allgemein definiert man \mathbb{Z}_n^* als die Menge der Elemente aus \mathbb{Z}_n , die teilerfremd zu n sind. Formal heißt das: $\mathbb{Z}_n^* = \{a \in \mathbb{Z}_n \mid \text{ggT}(a, n) = 1\}$. In Übungsaufgabe 3 sind Sie eingeladen zu zeigen, dass \mathbb{Z}_n^* für jede natürliche Zahl $n \geq 1$ eine multiplikative Gruppe ist.

Man hätte einen Körper auch kurz so definieren können: Ein **Körper** besteht aus einer Menge K , auf der zwei Verknüpfungen $+$ und \cdot definiert sind, so dass

- $(K, +)$ eine abelsche Gruppe mit neutralem Element 0 ist,
- $(K \setminus \{0\}, \cdot)$ eine abelsche Gruppe ist und
- ein Distributivgesetz gilt.

Ich empfehle Ihnen, sich ab jetzt die Definition eines Körpers in dieser lapidaren Form zu merken. Entsprechend lautet die Kurzform der Definition eines Rings: Ein **Ring** besteht aus einer additiven abelschen Gruppe R mit Nullelement 0 derart, dass auf $R \setminus \{0\}$ eine assoziative Multiplikation definiert ist, die beide Distributivgesetze erfüllt.

Ist auch in einem Vektorraum eine Gruppe verborgen? Selbstverständlich:

Additive Gruppe eines Vektorraums

Sei V ein Vektorraum über einem Körper K . Dann bildet die Menge V zusammen mit der Vektorraumaddition eine abelsche Gruppe.

Man hätte einen K -Vektorraum also auch kurz und bündig wie folgt definieren können: Ein **Vektorraum** besteht aus einer abelschen Gruppe $(V, +)$ zusammen mit einem Körper K derart, dass die Gesetze der Skalarmultiplikation gelten.

9.1.2 Gruppen aus bekannten Objekten

In der zweiten Kategorie von Beispielen machen wir aus uns bekannten Objekten Gruppen. Die Objekte, die wir studieren, sind lineare Abbildungen und Matrizen.

- Die Menge $K^{m \times n}$ aller $m \times n$ -Matrizen über dem Körper K bilden bezüglich der Matrizenaddition eine Gruppe.
- In ähnlicher Weise zeigt man, dass die Menge aller Abbildungen einer Menge X in eine Gruppe Y bezüglich der Definition der Summe von Abbildungen eine Gruppe ist. Dabei ist die Summe $f + g$ der Abbildungen f und g definiert durch $(f + g)(x) := f(x) + g(x)$ für alle $x \in X$.
- Insbesondere bildet die Menge aller linearen Abbildungen eines Vektorraums V in einen Vektorraum W eine Gruppe bezüglich der gerade definierten „punktweisen“ Addition.

Interessanter wird die Sache, wenn wir als Verknüpfung die Hintereinanderausführung von Abbildungen bzw. die Multiplikation von Matrizen studieren. Wenden wir uns zunächst den Matrizen zu.

Wenn wir zwei Matrizen aus $K^{m \times n}$ miteinander multiplizieren wollen, muss notwendigerweise $m = n$ sein. Da in einer Gruppe jedes Element ein inverses hat, muss in jeder Gruppe G von Matrizen aus $K^{n \times n}$ mit einer Matrix M auch die inverse Matrix M^{-1} in G enthalten sein. Das bedeutet, dass wir uns von vornherein auf $n \times n$ -Matrizen mit Determinante $\neq 0$ beschränken müssen. Es gelten die folgenden Aussagen:

Volle lineare Gruppe

Die Menge aller $n \times n$ -Matrizen über einem Körper K mit Determinante $\neq 0$ bildet bezüglich der Matrizenmultiplikation eine Gruppe; diese wird mit $GL(n, K)$ bezeichnet (**general linear group**). Ist K ein endlicher Körper mit q Elementen, so wird $GL(n, K)$ auch mit $GL(n, q)$ bezeichnet.

Die „volle“ lineare Gruppe heißt so, weil sie aus allen möglichen regulären Matrizen besteht, im Gegensatz zur „speziellen“ linearen Gruppe:

Spezielle lineare Gruppe

Die Menge aller $n \times n$ -Matrizen über einem Körper K mit Determinante 1 bildet bezüglich der Matrizenmultiplikation eine Gruppe; diese wird mit $SL(n, K)$ bezeichnet (**special linear group**). Ist K ein endlicher Körper mit q Elementen, so wird $SL(n, K)$ auch mit $SL(n, q)$ bezeichnet.

Der *Beweis* ergibt sich im Wesentlichen aus dem Multiplikationssatz für Determinanten: Das Produkt zweier Matrizen mit Determinante $\neq 0$ hat wieder eine von 0 verschiedene Determinante; das Produkt zweier Matrizen mit Determinante 1 hat auch Determinante 1. \square

Ganz analog zur ersten Aussage ergibt sich:

Die Menge aller umkehrbaren linearen Abbildungen eines Vektorraums V in sich bildet bezüglich der Hintereinanderausführung eine Gruppe; man bezeichnet diese Gruppe auch mit $GL(V)$. \square

Bemerkung Man kann unschwer zeigen (Übungsaufgabe 21), dass $GL(V)$ und $GL(n, K)$ „isomorph“ sind, wenn V ein n -dimensionaler Vektorraum über dem Körper K ist.

9.1.3 Gruppen aus Permutationen

In diesem letzten Beispielabschnitt behandeln wir einige Gruppen, die von besonderem Interesse sind, aber sich nicht direkt aus den Strukturen ableiten lassen, die in der Linearen Algebra behandelt werden.

Zunächst behandeln wir Permutationen (vergleichen Sie dazu die Abschn. 7.2 und 7.3). Die Menge aller Permutationen einer Menge X bildet bezüglich der Hintereinanderausführung von Abbildungen eine Gruppe; diese wird die **symmetrische Gruppe** auf X genannt. Wenn die Menge X genau n Elemente hat, so wird die symmetrische Gruppe auf X auch mit S_n bezeichnet. Die symmetrische Gruppe auf X ist die größte Gruppe von Permutationen von X . Wir können auch die jeweils zweitgrößte Gruppe beschreiben; es handelt sich dabei um die alternierende Gruppe A_n . Diese besteht aus allen geraden Permutationen. In Abschn. 7.3 haben wir uns auch klargemacht, dass dies genau die Permutationen sind, die sich als Produkt einer geraden Anzahl von Transpositionen schreiben lassen.

Beispiel Wir betrachten die Gruppe S_3 aller Permutationen der Menge $X = \{1, 2, 3\}$. Wir listen die Elemente von S_3 in den beiden uns aus Abschn. 7.2 bekannten Schreibweisen auf:

$$\text{id} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} = (1)(2)(3) ,$$

$$\alpha = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = (123) ,$$

$$\beta = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = (132) ,$$

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = (1)(23) ,$$

$$\tau = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = (13)(2) ,$$

$$\rho = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = (12)(3) .$$

Es folgt $S_3 = \{\text{id}, \alpha, \beta, \sigma, \rho, \tau\}$ und $A_3 = \{\text{id}, \alpha, \beta\}$.

Eine weitere wichtige Klasse von Gruppen erhält man als **Symmetriegruppen** geometrischer Objekte. Eine **Symmetrieabbildung** eines Objekts K ist ein „Automorphismus“ des Raumes (oder der Ebene), die K in sich überführt. Die „Automorphismen“ der Geometrie der reellen Ebene oder des reellen Raums sind dabei all diejenigen Permutationen der Punktmenge, die alle geometrisch relevanten Eigenschaften erhalten, wie zum Beispiel den Abstand. Man kann zeigen, dass die „Automorphismen“ der reellen Ebene, die man üblicherweise „Kongruenzabbildungen“ nennt, genau die Spiegelungen, Drehungen, Verschiebungen und Gleitspiegelungen sind. Wir diskutieren diese – intuitiv klare – Definition

hier nicht weiter, sondern verweisen auf Bücher über Grundlagen der Geometrie. Die Menge aller Symmetrieabbildungen einer Menge K von Punkten wird mit $\text{Sym}(K)$ bezeichnet.

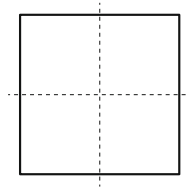
Satz über die Symmetriegruppe

Für jede Menge K von Punkten der Ebene oder des Raums ist $\text{Sym}(K)$ bezüglich der Hintereinanderausführung von Abbildungen eine Gruppe.

Zum *Beweis* müssen wir nur zeigen, dass $\text{Sym}(K)$ eine Untergruppe aller Permutationen der Punktmenge ist. Dies folgt aber daraus, dass da die Menge aller bijektiven Abbildungen des ganzen Raumes in sich assoziativ ist, dass die Identität in $\text{Sym}(K)$ enthalten ist und dass mit jeder Abbildung α auch die inverse Abbildung α^{-1} in $\text{Sym}(K)$ enthalten ist. \square

Wir betrachten zwei Beispiele.

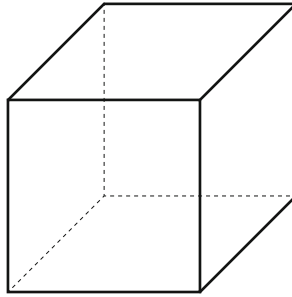
Zunächst betrachten wir die **Gruppe der Ebene, die ein Quadrat in sich überführt (Quadratgruppe)**. Diese besteht aus



- der Identität,
- zwei Spiegelungen an den Diagonalen
- zwei Spiegelungen an den Geraden, die zwei gegenüberliegende Seitenmitten verbinden,
- Drehungen um $\pm 90^\circ$,
- eine Drehung um 180° .

Machen Sie sich klar, dass diese acht Abbildungen wirklich eine Gruppe bilden: Die inversen Elemente sind leicht zu finden. Aber was ist eine Spiegelung an einer Diagonalen mal einer Spiegelung an einer Geraden, die zwei gegenüberliegende Seitenmitten verbindet? Beachten Sie, dass alle möglichen Produkte wieder ein Element der Quadratgruppe sind!

Als weiteres Beispiel betrachten wir die **Gruppe aller Symmetrien eines Würfels (Würfelgruppe)**. Aus welchen Elementen besteht diese Gruppe? Die wenigsten Menschen können diese Frage auf Anhieb beantworten. Aber jeder kann anfangen, eine Liste aufzustellen! Die Gruppe aller Symmetrieabbildungen eines Würfels enthält folgende Abbildungen:



- Die Identität,
- drei Drehungen um 90° um eine Gerade, die zwei gegenüberliegende Flächenmitten verbindet,
- drei Drehungen um -90° um eine Gerade, die zwei gegenüberliegende Flächenmitten verbindet,
- drei Drehungen um 180° um eine Gerade, die zwei gegenüberliegende Flächenmitten verbindet,
- sechs Drehungen um 180° um eine Gerade, die zwei gegenüberliegende Kantenmitten verbindet,
- vier Drehungen um 120° um eine Raumdiagonale,
- vier Drehungen um 240° um eine Raumdiagonale.

Wie kommt man denn darauf? Ganz einfach! Man fängt einfach irgendwie an und listet einige Symmetrieabbildungen auf. An irgendeiner Stelle geht's nicht weiter, d. h. es fällt einem keine weitere Symmetrieabbildung mehr ein. Dann nützt man aus, dass $\text{Sym}(\text{Würfel})$ eine Gruppe sein muss. Man muss sich dann zwei Fragen stellen:

- Ist für jede der bislang aufgeführten Abbildungen α auch α^{-1} aufgeführt?
- Ist für je zwei bislang aufgeführte Abbildungen α und β auch $\alpha\beta$ aufgeführt?

Besonders die zweite Frage erweist sich als außerordentlich hilfreich.

Als Übung sollten Sie versuchen, die Würfelgruppe als Gruppe auf den sechs Flächen bzw. auf den zwölf Kanten des Würfels darzustellen (siehe Übungsaufgabe 8).

9.2 Einfache Strukturaussagen für Gruppen

9.2.1 Untergruppen

Sei G eine Gruppe mit der Verknüpfung \cdot . Eine Teilmenge U von G heißt eine **Untergruppe** von G , falls U zusammen mit der von G auf U induzierten Verknüpfung \cdot ebenfalls eine Gruppe ist.

Beispiele Jede Gruppe G hat die **trivialen** Untergruppen G und $\{e\}$. Nichttriviale Beispiele sind:

- $(n\mathbb{Z}, +)$ ist eine Untergruppe von $(\mathbb{Z}, +)$.
- A_n ist eine Untergruppe von S_n .
- Die Menge $\{-1, 1\}$ ist eine Untergruppe von $(\mathbb{R} \setminus \{0\}, \cdot)$.
- Die Menge $\{1, -1, i, -i\}$ ist eine Untergruppe von $(\mathbb{C} \setminus \{0\}, \cdot)$.
- Die Identität zusammen mit den Drehungen um 90° , 180° und 270° ist eine Untergruppe der Symmetriegruppe eines Quadrats.
- Sie sind eingeladen, sämtliche Untergruppen der S_3 zu bestimmen (siehe Übungsaufgabe 18).

Entsprechend dem Unterraumkriterium gibt es das

Untergruppenkriterium

Sei U eine Teilmenge der Gruppe G . Dann ist U eine Untergruppe von G , falls folgende Bedingungen erfüllt sind:

- $U \neq \emptyset$.
- Wenn $g \in U$ ist, so ist auch $g^{-1} \in U$.
- Wenn $g, h \in U$ sind, so ist auch $gh \in U$.

Bemerkung Die Bedeutung dieses Kriteriums liegt darin, dass man, um eine Menge als Untergruppe nachzuweisen, nur ganz wenig zeigen muss; insbesondere muss man das Assoziativgesetz nicht zeigen.

Beweis des Untergruppenkriteriums Zunächst ist zu zeigen, dass \cdot tatsächlich eine Verknüpfung auf U induziert; dies ist wegen der dritten Forderung erfüllt.

(G1) Die Verknüpfung ist auf U assoziativ, da sie schon auf G assoziativ ist.

(G2) Wegen der ersten Forderung gibt es ein Element g in U . Nach der zweiten Forderung ist dann auch g^{-1} in U , und aus der dritten Forderung ergibt sich, dass damit auch $g \cdot g^{-1} = e$ in U liegt. Daher liegt das neutrale Element von G in U , und dieses wirkt natürlich auch in U als neutrales Element.

(G3) ergibt sich direkt aus der zweiten Forderung. □

Mit Untergruppen einer Gruppe kann man (fast) die gleichen Spielchen machen wie mit Unterräumen eines Vektorraums ... zum Beispiel kann man ... Nebenklassen definieren! Wie das? Ganz einfach: Ist U eine Untergruppe von G , und ist $g \in G$, so ist die Menge

$$gU = \{gu \mid u \in U\}$$

eine **Nebenklasse** von G nach U .

Was Nebenklassen von Vektorräumen recht war, ist Nebenklassen von Gruppen billig:

Satz über Nebenklassen

Sei U eine Untergruppe der Gruppe G .

(a) (**Kriterium für die Gleichheit von Nebenklassen**). Es gilt:

$$gU = hU \Leftrightarrow h^{-1}g \in U.$$

(b) Je zwei verschiedene Nebenklassen sind disjunkt.

(c) Die Abbildung $f: U \rightarrow gU$ mit $f(u) = gu$ ist eine Bijektion.

Beweis

(a) Einerseits gilt

$$\begin{aligned} gU = hU &\Rightarrow g = ge \in hU \Rightarrow g = hu \text{ für ein } u \in U \\ &\Rightarrow h^{-1}g = u \in U (\Leftrightarrow gh^{-1} = (h^{-1}g)^{-1} \in U). \end{aligned}$$

Andererseits folgt aus $u := h^{-1}g \in U$ zum einen $g = hu \in hU$, also $gU \subseteq hU$. Entsprechend folgt auch $h = gu^{-1} \in gU$, also $hU \subseteq gU$, und somit $gU = hU$.

(b) Angenommen, gU und hU haben ein Element x gemeinsam. Dann existieren $u, u' \in U$ mit $gu = x = hu'$. Daraus folgt $h^{-1}g = u'u^{-1} \in U$. Also ist nach (a) $gU = hU$.

(c) Dass f surjektiv ist, folgt aus der Definition von f . Aus $gu = f(u) = f(u') = gu'$ folgt durch Kürzen von g auch $u = u'$. Also ist f injektiv. \square

Wir betrachten jetzt **endliche** Gruppen, das heißt Gruppen mit einer endlichen Anzahl von Elementen. Für eine endliche Gruppe G ist also $|G|$ eine natürliche Zahl, man nennt sie die **Ordnung** von G .

Beispielsweise hat die Quadratgruppe die Ordnung 8, die symmetrische Gruppe S_n hat die Ordnung $n!$ und die alternierende Gruppe auf n Elementen hat die Ordnung $n!/2$.

Eine äußerst wichtige und äußerst fruchtbare Fragestellung in der Gruppentheorie ist, ob man aus der Kenntnis der *Ordnung einer Gruppe Rückschlüsse auf ihre algebraische Struktur* ziehen kann. Das kann man tatsächlich! Zum Beispiel werden wir alle Gruppen bestimmen, deren Ordnung eine Primzahl ist. Ich finde es sehr erstaunlich und bemerkenswert, dass man (in vielen Fällen) die ganze Struktur einer Gruppe durch Angabe einer einzigen Zahl beschreiben kann. (Hier sehen Sie: Mathematik ist die Lehre von den guten Beschreibungen.)

Das entscheidende Werkzeug zur Beschreibung der Struktur einer Gruppe über ihre Parameter ist der einfache, aber fundamentale Satz von Lagrange. Um diesen formulieren zu können, brauchen wir noch eine Definition.

Sei G eine endliche Gruppe, und sei U eine Untergruppe von G . Wir nennen die Anzahl der Nebenklassen von G nach U den **Index** von U in G und bezeichnen diese Zahl mit $[G:U]$.

Zum *Beispiel* ist der Index von A_n in S_n gleich 2.

Satz von Lagrange

Sei G eine endliche Gruppe, und sei U eine Untergruppe von G . Dann gilt

$$|G| = |U| \cdot [G : U] .$$

Bemerkung Die Bedeutung des Satzes liegt *nicht* darin, dass man den Index ausrechnen könnte. Seine Bedeutung liegt darin, dass daraus folgt, dass *die Ordnung einer jeden Untergruppe die Gruppenordnung teilen muss!* Wenn zum Beispiel G die Ordnung 100 hat, so wissen wir nach diesem Satz, dass G nur Untergruppen der Ordnung 1, 2, 4, 5, 10, 20, 25, 50 und 100 haben kann. (Damit ist allerdings nicht gesagt, dass es solche Untergruppen von G wirklich gibt!)

Beweis des Satzes von Lagrange Wir wissen, dass je zwei Nebenklassen disjunkt sind. Da jedes Gruppenelement zu einer Nebenklasse gehört, ist also die Menge der Nebenklassen eine Partition von G .

Ferner gibt es eine bijektive Abbildung von U auf gU ($g \in G$); daher haben je zwei Nebenklassen die gleiche Mächtigkeit, nämlich $|U|$. Daher ist die Menge der Nebenklassen eine Partition von G in Mengen gleicher Mächtigkeit $|U|$. Damit folgt:

$$|G| = \text{Anzahl der Nebenklassen} \cdot \text{Anzahl der Elemente pro Nebenklasse} = [G : U] \cdot |U| .$$

□

Mit dem Satz von Lagrange können wir ganz einfach einen wichtigen Satz der elementaren Zahlentheorie beweisen:

Satz

Sei p eine Primzahl. Wenn -1 ein Quadrat in \mathbb{Z}_p ist, so ist $p-1$ durch 4 teilbar.

Also ist zum *Beispiel* in \mathbb{Z}_3 , \mathbb{Z}_7 , \mathbb{Z}_{11} und \mathbb{Z}_p mit $p = 2^{57885161} - 1$ das Element -1 kein Quadrat. ($p = 2^{57885161} - 1$ ist die größte heute (November 2014) bekannte Primzahl; sie hat 17.425.170 Dezimalstellen und würde ausgeschrieben viele Bücher füllen. Primzahlen der Form $2^n - 1$ nennt man **Mersennesche Primzahlen**. Da $2^{57885161}$ durch 4 teilbar ist, ist $p-1 = 2^{57885161} - 2$ nicht durch 4 teilbar.)

Beweis des Satzes. Sei -1 ein Quadrat in \mathbf{Z}_p . Dann gibt es ein Element $i \in \mathbf{Z}_p$, so dass $i^2 = -1$ ist. Dann ist die Menge

$$U = \{1, -1, i, -i\}$$

abgeschlossen bezüglich der Multiplikation und erfüllt also die Voraussetzungen des Untergruppenkriteriums. Daher ist U eine Untergruppe von \mathbf{Z}_p^* . Nach dem Satz von Lagrange ist also $|U| = 4$ ein Teiler von $|\mathbf{Z}_p^*| = p-1$. \square

Bemerkung Man kann auch zeigen, dass die Umkehrung gilt: Wenn p eine Primzahl ist mit $4 \mid p-1$, so ist -1 ein Quadrat in \mathbf{Z}_p .

9.2.2 Zyklische Gruppen

Das **Erzeugnis** der Elemente g_1, g_2, \dots einer Gruppe G ist definiert als der Durchschnitt aller Untergruppen von G , die g_1, g_2, \dots enthalten; man schreibt dafür $\langle g_1, g_2, \dots \rangle$:

$$\langle g_1, g_2, \dots \rangle = \bigcap \{U \mid U \text{ Untergruppe, die } g_1, g_2, \dots \text{ enthält}\}.$$

Man kann aber das Erzeugnis dieser Elemente auch so beschreiben: $\langle g_1, g_2, \dots \rangle$ besteht aus allen Produkten der Elemente g_1, g_2, \dots und ihrer Inversen; also

$$\begin{aligned} \langle g_1, g_2, \dots \rangle \\ = \{g_1, g_2, \dots, g_1 g_2, g_2 g_1, g_1 g_2 g_1, g_1^{-1} g_2, e, \dots, g_1 g_2 g_3, \dots, g_4 g_5^{-1} g_2 g_1^{-1}, \dots\} \end{aligned}$$

(Machen Sie sich – am besten mit Hilfe des Untergruppenkriteriums – klar, dass diese beiden Darstellungen wirklich die gleiche Menge beschreiben.)

Besonders interessant und wichtig ist das Erzeugnis eines einzelnen Elements: Die von einem Element $g \in G$ erzeugte Untergruppe $\langle g \rangle$ von G besteht aus allen Potenzen von g :

$$\langle g \rangle = \{g^z \mid z \in \mathbf{Z}\}.$$

Man nennt $\langle g \rangle$ die von g erzeugte **zyklische** Untergruppe von G . Im Allgemeinen nennt man jede von einem Element erzeugte Gruppe **zyklisch**.

Bevor wir Beispiele zyklischer Gruppen betrachten, schreiben wir die Definition einer zyklischen Gruppe in die additive Schreibweise um: Eine additiv geschriebene Gruppe $(G, +)$ ist **zyklisch**, wenn es ein Element $g \in G$ gibt, derart, dass G gleich der Menge der (ganzahligen) Vielfachen von g ist:

$$G = \langle g \rangle = \{zg \mid z \in \mathbf{Z}\}.$$

Beispiele

(a) $(\mathbf{Z}, +)$ ist eine unendliche zyklische Gruppe; sie wird von dem Element 1 erzeugt.

- (b) Die Gruppe $(\mathbb{Z}_n, +)$ ist eine zyklische Gruppe der Ordnung n ; sie wird von dem Element 1 erzeugt.

Bemerkung Wir werden später sehen, dass dies „bis auf Isomorphie“ die einzigen zyklischen Gruppen sind.

Die Gruppenordnung ist ein globaler Parameter zur Beschreibung einer Gruppe. Häufig muss man aber auch lokale Parameter betrachten, die die einzelnen Elemente beschreiben:

Die **Ordnung** eines Elements $g \in G$ ist die Ordnung der von g erzeugten Untergruppe von G .

Als einfache Folgerung aus dem Satz von Lagrange halten wir fest:

Satz von Lagrange für Gruppenelemente

Sei g ein Element einer endlichen Gruppe G . Dann teilt die Ordnung von g die Ordnung $|G|$ von G .

Beweis Die Ordnung von g ist die Mächtigkeit der Untergruppe $\langle g \rangle$ von G . □

Man kann die Ordnung eines Elements auch anders beschreiben:

Satz über die Ordnung von Elementen einer Gruppe

Sei g ein Element einer Gruppe G . Dann ist die Ordnung von g unendlich, oder die Ordnung ist die kleinste positive ganze Zahl k mit $g^k = e$.

Beweis Die Ordnung von g möge endlich sein. Dann ist $\langle g \rangle$ eine endliche Gruppe. Sei m die Anzahl der Elemente von $\langle g \rangle$. Dann ist nach Definition m die Ordnung von g .

Sei k die kleinste positive ganze Zahl k mit $g^k = e$. Es ist zu zeigen, dass $k = m$ ist.

„ $k \leq m$ “: Dazu zeigen wir, dass die Elemente $e = g^0, g, g^2, g^3, \dots, g^{k-1}$ verschieden sind: (Wäre $g^i = g^j$ mit $0 \leq i < j < k$, so wäre $g^{j-i} = g^j(g^i)^{-1} = e$ mit $j-i < k$, ein Widerspruch.) Also hat $\langle g \rangle$ mindestens die verschiedenen Elemente $g^0, g^1, g^2, g^3, \dots, g^{k-1}$; damit hat $\langle g \rangle$ mindestens k Elemente; also ist $m \geq k$.

„ $m \leq k$ “: Dazu zeigen wir, dass $G' := \{g^0, g^1, g^2, g^3, \dots, g^{k-1}\}$ eine Untergruppe bildet:

Wir wenden das Untergruppenkriterium an. Sicher ist $G' \neq \emptyset$. Das Inverse eines Elements $g^i \in G'$ ist das Element g^{k-i} ; denn $g^i g^{k-i} = g^k = e$. Da $k-i$ eine nichtnegative ganze Zahl $< k$ ist, ist auch g^{k-i} in G' . Seien schließlich g^i und g^j Elemente von G' . Es ist zu zei-

gen, dass $g^i g^j = g^{i+j}$ in G' ist. Wenn $i+j < k$ ist, ist dies sofort klar. Wenn $i+j \geq k$ ist, ist $i+j = k+i'$ mit $0 \leq i' < k$. In diesem Fall ergibt sich

$$g^{i+j} = g^{k+i'} = g^k g^{i'} = e g^{i'} = g^{i'} \in G'.$$

Somit ist G' eine Untergruppe von G , die g enthält. Nach Definition von $\langle g \rangle$ ist $\langle g \rangle$ in G' enthalten. Also ist die Anzahl m der Elemente von $\langle g \rangle$ höchstens so groß wie die Anzahl k der Elemente von G' . \square

Korollar

Wenn G eine endliche zyklische Gruppe der Ordnung m ist, die von g erzeugt wird, dann gilt

$$G = \{g^0, g^1, g^2, \dots, g^{m-1}\} = \{e, g, g^2, \dots, g^{m-1}\}. \quad \square$$

Als weitere Folgerung halten wir fest:

Kleiner Satz von Fermat

Ist G eine endliche Gruppe, so gilt für alle Elemente $g \in G$:

$$g^{|G|} = e. \quad \square$$

Fermat selbst (Pierre de Fermat, 1601–1665) hat diesen Satz nur für eine spezielle Gruppe ausgesprochen, nämlich für die Gruppe (\mathbb{Z}_p^*, \cdot) für eine Primzahl p . Erst Leonhard Euler (1707–1783) hat den Satz auf \mathbb{Z}_n^* verallgemeinert. In dieser Situation gilt

Satz von Euler

Sei g eine ganze Zahl zwischen 1 und $n-1$. Ist g teilerfremd zu n , so gilt

$$g^{|\mathbb{Z}_n^*|} \bmod n = 1.$$

Insbesondere gilt (**Kleiner Satz von Fermat (Originalfassung)**): Ist p eine Primzahl, so ist

$$g^{p-1} \bmod p = 1. \quad \square$$

Der folgende Satz sagt, dass für *zyklische Gruppen* auch die Umkehrung des Satzes von Lagrange gilt.

Satz über die Untergruppen einer zyklischen Gruppe

Sei G eine zyklische Gruppe der Ordnung n . Ist k ein Teiler von n , so hat G eine Untergruppe der Ordnung k .

Beweis Sei g ein erzeugendes Element von G . Da k ein Teiler von n ist, gibt es eine natürliche Zahl m mit $km = n$.

Behauptung $h := g^m$ ist ein Element der Ordnung k . (Sicher ist $h^k = (g^m)^k = g^{mk} = g^n = e$. Also ist die Ordnung von h höchstens k . Wäre die Ordnung k' von h kleiner als k , so wäre aber auch $g^{mk'} = (g^m)^{k'} = h^{k'} = e$ mit $mk' < mk = n$; also hätte g eine Ordnung $< n$, ein Widerspruch.)

Daher erzeugt h eine Untergruppe $\langle h \rangle$ von G der Ordnung k . □

9.2.3 Der Homomorphiesatz

Seien G und H Gruppen. Eine Abbildung $f: G \rightarrow H$ heißt ein **Homomorphismus** (der Gruppen G und H), falls für alle $g, g' \in G$ gilt

$$f(gg') = f(g)f(g') .$$

Ein **Isomorphismus** ist ein bijektiver Homomorphismus, und ein **Automorphismus** ist ein Isomorphismus einer Gruppe auf sich selbst. Zwei Gruppen G und H heißen **isomorph** (in Zeichen $G \cong H$), falls es einen Isomorphismus von G auf H gibt.

Beispiele

- S_3 und die Symmetriegruppe eines regulären Dreiecks sind isomorphe Gruppen.
- Die Abbildung $\text{sig}: S_n \rightarrow \{-1, 1\}$, die jeder Permutation ihr Signum zuordnet, ist ein Homomorphismus von Gruppen.

Klassifikation der zyklischen Gruppen

Jede zyklische Gruppe ist isomorph zu \mathbb{Z} oder zu \mathbb{Z}_n .

Beweis Sei G eine zyklische Gruppe, die von einem Element g erzeugt wird. Wir unterscheiden zwei Fälle.

1. Fall: $\langle g \rangle$ ist unendlich.

Dann sind die Elemente g^z für $z \in \mathbb{Z}$ alle verschieden und schöpfen nach Definition des Erzeugnisses ganz $\langle g \rangle = G$ aus. Es folgt, dass die Abbildung $f: G \rightarrow \mathbb{Z}$, die durch $f(g^z) := z$ definiert ist, ein Isomorphismus von G in $(\mathbb{Z}, +)$.

2. Fall: $\langle g \rangle$ ist eine endliche Gruppe.

Sei n die Ordnung von $\langle g \rangle$. Dann ist die Abbildung $f: G \rightarrow \mathbb{Z}_n$, die definiert ist durch $f(g^i) = i$, ein Isomorphismus. \square

Sei $f: G \rightarrow H$ ein Homomorphismus der Gruppe G in die Gruppe H . Der **Kern** von f ist definiert als

$$\text{Kern}(f) = \{g \in G \mid f(g) = e\}.$$

Das **Bild** von f besteht aus allen Bildern der Elemente von G :

$$\text{Bild}(f) = \{h \in H \mid \text{es gibt } g \in G \text{ mit } f(g) = h\} = \{f(g) \mid g \in G\}.$$

Zum *Beispiel* ist der Kern der Signumsabbildung $\text{sig}: S_n \rightarrow \{-1, 1\}$ gleich der Menge der geraden Permutationen, also gleich A_n .

Man kann sich leicht überlegen (Übungsaufgabe 25), dass $\text{Kern}(f)$ eine Untergruppe von G und $\text{Bild}(f)$ eine Untergruppe von H ist.

Nun müssen wir noch einen wichtigen Begriff einführen. Dies ist der Begriff des „Normalteilers“. Wir wollen nämlich – analog zu Faktorräumen – auch Faktorgruppen bilden. Es zeigt sich, dass die Menge aller Nebenklassen nach einer Untergruppe nicht notwendig eine Gruppe bildet; dies ist vielmehr nur für eine spezielle Sorte von Untergruppen der Fall, und dies sind die Normalteiler.

Eine Untergruppe N einer Gruppe G heißt ein **Normalteiler** von G , falls für alle $g \in G$ gilt: Die Menge

$$g^{-1}Ng = \{g^{-1}ng \mid n \in N\}$$

ist gleich N .

Kurz: N ist ein Normalteiler, falls für alle $g \in G$ gilt: $g^{-1}Ng = N$.

Beispiele

(0) Jede Untergruppe einer *abelschen* Gruppe ist ein Normalteiler.

(1) Die Menge A_n der geraden Permutationen bildet einen Normalteiler von S_n .

Weitere wichtige Beispiele erhalten wir aus Homomorphismen:

Kern eines Homomorphismus

Sei $f: G \rightarrow H$ ein Homomorphismus von Gruppen. Dann ist $\text{Kern}(f)$ ein Normalteiler von G .

Beweis Dass $\text{Kern}(f)$ eine Untergruppe von G ist, wird in Übungsaufgabe 25 bewiesen.

Sei $h \in \text{Kern}(f)$, also $f(h) = e$. Sei ferner g ein beliebiges Element von G . Dann ist

$$f(g^{-1}hg) = f(g^{-1})f(h)f(g) = f(g^{-1})ef(g) = f(g^{-1})f(g) = f(g)^{-1}f(g) = e.$$

(In Übungsaufgabe 26 sollen Sie zeigen, dass $f(g^{-1}) = f(g)^{-1}$ gilt.)

Also ist tatsächlich $g^{-1}hg \in \text{Kern}(f)$. □

Man kann Normalteiler auf folgende Weise charakterisieren:

Eine Untergruppe N von G ist genau dann ein Normalteiler, falls für alle $g \in G$ gilt

$$gN = Ng.$$

(Dabei ist $gN = \{gn \mid n \in N\}$, und Ng ist entsprechend definiert.)

Dies zeigen wir auf folgende Weise Sei zunächst N ein Normalteiler. Zu zeigen ist $gN \subseteq Ng$ und $Ng \subseteq gN$. Sei $gn \in gN$ beliebig. Da N ein Normalteiler ist, ist $ngn^{-1} \in N$; also gibt es ein $n_1 \in N$ mit $ngn^{-1} = n_1$. Es folgt $gn = n_1g \in Ng$. Entsprechend zeigt man die andere Inklusion.

Nun möge $gN = Ng$ für alle $g \in G$ gelten. Um zu zeigen, dass N ein Normalteiler ist, müssen wir $g^{-1}Ng \subseteq N$ und $N \subseteq g^{-1}Ng$ zeigen.

Sei $g^{-1}ng \in g^{-1}Ng$ beliebig. Wegen $Ng = gN$ gibt es ein $n_1 \in N$ mit $ng = gn_1$. Also ist $g^{-1}ng = n_1 \in N$.

Sei umgekehrt $n \in N$ beliebig. Dann ist auch $n^{-1} \in N$. Wegen $Ng = gN$ gibt es ein $n_1 \in N$ mit $n^{-1}g = gn_1$. Es folgt $n^{-1} = gn_1g^{-1}$, also $n = g^{-1}n_1^{-1}g \in g^{-1}Ng$. □

Sei N ein Normalteiler von G . Dann heißt $G/N = \{gN \mid g \in G\}$ die **Faktorgruppe** von G nach N .

Satz über die Faktorgruppe

Sei N ein Normalteiler der Gruppe G . Dann ist G/N eine Gruppe mit der Verknüpfung

$$(gN) \cdot (hN) := ghN.$$

Beweis Zunächst müssen wir zeigen, dass die Verknüpfung wohldefiniert ist. Seien $g, g', h, h' \in G$ mit $gN = g'N$ und $hN = h'N$. Es ist zu zeigen, dass $ghN = g'h'N$ ist.

Wegen $gN = g'N$ ist $g^{-1}g' = n_1 \in N$. Ebenso ist $n_2 = h^{-1}h' \in N$. Daraus folgt:

$$gh = g'n_1^{-1}h'n_2^{-1} = g'h'n_1'n_2^{-1} \text{ mit } n_1' \in N.$$

Warum? Dies liegt daran, dass N ein Normalteiler ist. Dies bedeutet nämlich, dass für alle g aus G und n aus N gilt $ng = gn'$ mit einem gewissen n' aus N .

Damit unterscheiden sich gh und $g'h'$ nur um das Element $(n_1'n_2)^{-1}$ von N . Damit ist $ghN = g'h'N$. Also ist die Multiplikation auf G/N wohldefiniert.

Die restlichen Eigenschaften sind einfach nachzuweisen und werden Ihnen als Übungsaufgabe überlassen (siehe Übungsaufgabe 29). \square

Korollar

$G/\text{Kern}(f)$ ist eine Gruppe. \square

Es fragt sich, *welche* Gruppe $G/\text{Kern}(f)$ ist. Kennen wir diese Gruppe schon? Die Antwort darauf gibt der ...

Homomorphiesatz für Gruppen

Sei $f: G \rightarrow H$ ein Homomorphismus der Gruppe G in die Gruppe H . Dann gilt

$$G/\text{Kern}(f) \cong \text{Bild}(f).$$

Beweis Sei $N := \text{Kern}(f)$. Wir müssen einen Isomorphismus von G/N auf $\text{Bild}(f)$ definieren. Dazu müssen wir mindestens mal eine Abbildung φ von G/N in $\text{Bild}(f)$ definieren; wenn wir φ so einfach wie wir nur können definieren, so haben wir die Hoffnung, dass φ tatsächlich ein Homomorphismus oder sogar ein Isomorphismus ist.

Ganz analog zum Homomorphiesatz für Faktorräume definieren wir

$$\varphi(gN) := f(g).$$

Es ist klar, dass die Abbildung φ jedenfalls surjektiv ist. Was habe ich da gehört? Die Abbildung φ ? Ist das überhaupt eine Abbildung? Wird jeder Nebenklasse tatsächlich nur ein Element zugeordnet? Wenn ich eine Nebenklasse als gN darstelle und Sie dieselbe Nebenklasse als hN , so erhalte ich als Bild $f(g)$, Sie aber $f(h)$. Es müsste also $f(g) = f(h)$ sein. (Sie erinnern sich: Man nennt dies die Wohldefiniertheit der Abbildung φ .) Aber jetzt gilt:

Gefahr erkannt, Gefahr gebannt!

$$\begin{aligned}
 gN &= hN \\
 &\Leftrightarrow gh^{-1} \in N (= \text{Kern}(f)) \\
 &\Leftrightarrow f(gh^{-1}) = e \\
 &\Leftrightarrow e = f(gh^{-1}) = f(g) \cdot f(h^{-1}) = f(g)f(h)^{-1} \\
 &\Leftrightarrow f(g) = f(h) .
 \end{aligned}$$

Der Rest ist (jetzt!) einfache Routine:

- φ ist ein Homomorphismus: Seien gN und hN zwei beliebige Nebenklassen. Dann gilt:

$$\varphi(gN \cdot hN) = \varphi(ghN) = f(gh) = f(g) \cdot f(h) = \varphi(gN) \cdot \varphi(hN) .$$

- φ ist surjektiv: Dies folgt direkt aus der Definition.
- φ ist injektiv: Seien gN und hN Nebenklassen. Dann ergibt sich ähnlich wie oben:

$$\begin{aligned}
 \varphi(gN) &= \varphi(hN) \\
 &\Leftrightarrow f(g) = f(h) \\
 &\Leftrightarrow f(gh^{-1}) = f(g)f(h)^{-1} = e \\
 &\Leftrightarrow gh^{-1} \in \text{Kern}(f) (= N) \\
 &\Leftrightarrow gN = hN .
 \end{aligned}$$

Damit ist alles gezeigt. □

9.3 Verständnisfragen, Übungen und Tipps

Richtig oder falsch?

1. Thema: Gruppen und Vektorräume

- ☐ Jeder Vektorraum ist eine Gruppe.
- ☐ Jeder Vektorraum ist eine abelsche Gruppe.
- ☐ Jede abelsche Gruppe kann zu einem Vektorraum gemacht werden.
- ☐ Die Menge der linearen Abbildungen eines Vektorraums V in einen Vektorraum W ist bezüglich der punktweisen Addition eine Gruppe.
- ☐ Die Menge der linearen Abbildungen eines Vektorraums in sich ist bezüglich der Hintereinanderausführung eine Gruppe.
- ☐ Die Menge der invertierbaren linearen Abbildungen eines Vektorraums in sich ist bezüglich der Hintereinanderausführung eine Gruppe.

2. Thema: Ordnung von Gruppen und Elementen

- ☐ Zu jeder natürlichen Zahl n gibt es mindestens eine Gruppe der Ordnung n .
- ☐ Zu jeder natürlichen Zahl n gibt es genau eine Gruppe der Ordnung n .
- ☐ Jede unendliche Gruppe enthält mindestens ein Element unendlicher Ordnung.

3. Thema: Untergruppen

- ☐ Jede Gruppe hat eine Untergruppe.
- ☐ Jede Gruppe hat eine nichttriviale Untergruppe.
- ☐ Untergruppen sind Normalteiler.
- ☐ Untergruppen abelscher Gruppen sind Normalteiler.
- ☐ Untergruppen von unendlichen Gruppen sind unendlich.

4. Thema: Isomorphie von Gruppen

- ☐ Je zwei Gruppen gleicher Ordnung sind isomorph.
- ☐ Je zwei zyklische Gruppen gleicher Ordnung sind isomorph.
- ☐ Zu jeder natürlichen Zahl n gibt es mindestens zwei nichtisomorphe Gruppen der Ordnung n .
- ☐ Es gibt – bis auf Isomorphie – nur eine Gruppe mit vier Elementen.

5. Thema: Homomorphiesatz

Sei $f: G \rightarrow H$ ein Gruppenisomorphismus. Aus dem Homomorphiesatz folgt:

- ☐ $\text{Kern}(f)$ und $\text{Bild}(f)$ sind isomorph.
- ☐ $\text{Kern}(f)$ und $\text{Bild}(f)$ sind nie isomorph.
- ☐ Je zwei Gruppen sind isomorph.
- ☐ G ist isomorph zu $\text{Bild}(f)$.
- ☐ G hat eine Untergruppe isomorph zu $\text{Bild}(f)$.
- ☐ G hat eine Faktorgruppe isomorph zu $\text{Bild}(f)$.
- ☐ G hat eine Untergruppe isomorph zu $\text{Kern}(f)$.

Übungsaufgaben

1. Formulieren Sie die Definition eines „Ideals“ mit dem Begriff „Gruppe“.
2. Zeigen Sie: Jede Untergruppe von $(\mathbb{Z}, +)$ ist von der Form $(n\mathbb{Z}, +)$.
3. Machen Sie sich klar, dass \mathbb{Z}_n^* bezüglich der Multiplikation modulo n eine Gruppe bildet. Machen Sie sich die Aussage zunächst am Beispiel $n = 15$ klar.
4. Bestimmen Sie die Symmetriegruppe eines Rechtecks, das kein Quadrat ist. Überzeugen Sie sich durch eine Verknüpfungstabelle, dass die von Ihnen gefundenen Abbildungen wirklich eine Gruppe bilden.
5. Bestimmen Sie die Symmetriegruppe eines Quaders mit den Seitenlängen $a, a, b \neq a$.
6. Bestimmen Sie die Symmetriegruppe eines regulären Tetraeders.
7. Bestimmen Sie die Ordnungen der in den vorigen drei Aufgaben gefundenen Gruppen.
8. Stellen Sie die Würfelgruppe dar
 - (a) als Gruppe von Permutationen auf den sechs Flächen eines Würfels,

- (b) als Gruppe von Permutationen auf den zwölf Kanten eines Würfels.
9. (a) Zeigen Sie, dass jede Untergruppe von S_4 die Ordnung 1, 2, 3, 4, 6, 8, 12 oder 24 haben muss.
(b) Geben Sie zu jeder dieser Ordnungen eine Untergruppe an.
 10. Ist 1 das einzige Element aus \mathbf{Z} , das die zyklische Gruppe $(\mathbf{Z}, +)$ erzeugt?
 11. Bestimmen Sie alle Elemente, die die Gruppe \mathbf{Z}_7^* bzw. \mathbf{Z}_8^* erzeugen.
 12. Zeigen Sie: Die Menge der natürlichen Zahlen zwischen 0 und $n-1$, die ein multiplikatives Inverses modulo n haben, ist gleich der Menge der natürlichen Zahlen zwischen 0 und $n-1$, die mit n den größten gemeinsamen Teiler 1 haben (zwei natürliche Zahlen, deren ggT gleich 1 ist, nennt man auch **teilerfremd**) [Hinweis: Benutzen Sie das Lemma von Bézout.]
 13. Sei p eine Primzahl. Bestimmen Sie $|\mathbf{Z}_p^*|$ und $|\mathbf{Z}_{p^2}^*|$.
 14. Vergleichen Sie das Erzeugnis von Elementen einer Gruppe genau mit dem Erzeugnis von Vektoren eines Vektorraums.
 15. Sei G eine beliebige Gruppe der Ordnung 65.536. Zeigen Sie, dass es in G ein Element der Ordnung 2 gibt (d. h. ein Element g mit $g \neq e$ und $g^2 = e$).
 16. Sei G eine abelsche Gruppe.
(a) Wenn g und h zwei Elemente der Ordnung 2 in G sind, so hat auch gh die Ordnung 2.
(b) Zeigen Sie: Zwei verschiedene Elemente der Ordnung 2 in G erzeugen eine Untergruppe der Ordnung 4.
(c) Zeigen Sie: Sei G eine zyklische Gruppe gerader Ordnung. Dann enthält G genau ein Element der Ordnung 2.
 17. Sei G eine beliebige Gruppe und U eine Untergruppe von G . Zeigen Sie: Wenn U den Index 2 hat, so ist U ein Normalteiler.
 18. Bestimmen Sie sämtliche Untergruppen der S_3 . [Hinweis: Es gibt genau vier nicht-triviale Untergruppen.]
 19. Zeigen Sie: Wenn für jedes Element g einer Gruppe G gilt $g^2 = e$, dann ist G eine abelsche Gruppe.
 20. Sei V ein Vektorraum über einem Körper $K = GF(p)$, dessen Ordnung eine Primzahl p ist. Zeigen Sie: Jedes von Null verschiedene Element von V hat in der additiven Gruppe die Ordnung p .
 21. Sei V ein n -dimensionaler K -Vektorraum. Zeigen Sie, dass $GL(V)$ isomorph zu $GL(n, K)$ ist. [Vergleichen Sie dazu das Projekt in Kap. 5.]
 22. Mit $GL(n, q)$ bezeichnen wir die Gruppe der invertierbaren $n \times n$ -Matrizen mit Elementen aus $GF(q)$. Bestimmen Sie die Ordnung von $GL(n, q)$ für $n = 2$ und $n = 3$.
Zusatzfrage: Können Sie eine allgemeine Formel für $|GL(n, q)|$ angeben?
 23. Zeigen Sie, dass die Menge $SL(n, q)$ der $n \times n$ -Matrizen M über $GF(q)$ mit $\det(M) = 1$ eine Untergruppe von $GL(n, q)$ ist. Bestimmen Sie den Index dieser Untergruppe in $GL(n, q)$.
 24. Geben Sie eine Untergruppe von $GL(n, q)$ an, die genau $q-1$ Elemente hat.

25. Sei $f: G \rightarrow H$ ein Gruppenhomomorphismus. Zeigen Sie: $\text{Kern}(f)$ ist eine Untergruppe von G , $\text{Bild}(f)$ ist eine Untergruppe von H .
26. Sei $f: G \rightarrow H$ ein Gruppenhomomorphismus. Zeigen Sie, dass für jedes Element g von G gilt: $f(g^{-1}) = f(g)^{-1}$.
27. Sei $G := \mathbf{R} \setminus \{-1\}$. Wir definieren auf G eine Operation $*$ wie folgt: $x * y := xy + x + y$.
 (a) Zeigen Sie, dass $(G, *)$ eine Gruppe ist.
 (b) Geben Sie einen Isomorphismus von $(G, *)$ auf (\mathbf{R}^*, \cdot) an.
28. Sei G eine Gruppe, und sei N eine Untergruppe von G . Zeigen Sie
 N ist ein Normalteiler $\Leftrightarrow g^{-1}Ng \subseteq N$ für alle $g \in G$.
29. Sei N ein Normalteiler der Gruppe G . Zeigen Sie, dass G/N eine Gruppe ist.
30. Bestimmen Sie alle Gruppen der Ordnung 1, 2, 3, 4, 5, 6, 7.
 Zusatzfrage: Bestimmen Sie alle Gruppen der Ordnung 8.
31. Sei $K = GF(q)$ ein endlicher Körper. Zeigen Sie, dass die Menge

$$Q := \{x^2 \in K \mid x \in K \setminus \{0\}\}$$

aller „Quadrate von K “ eine Gruppe bezüglich der Multiplikation bildet.

32. Sei K ein endlicher Körper. Betrachten Sie die Abbildung $f: K^* \rightarrow K^*$, die definiert ist durch

$$f(x) := x^2.$$

Ist f ein Gruppenhomomorphismus? (Wenn ja, zwischen welchen Gruppen?)

Ist f ein Körperhomomorphismus?

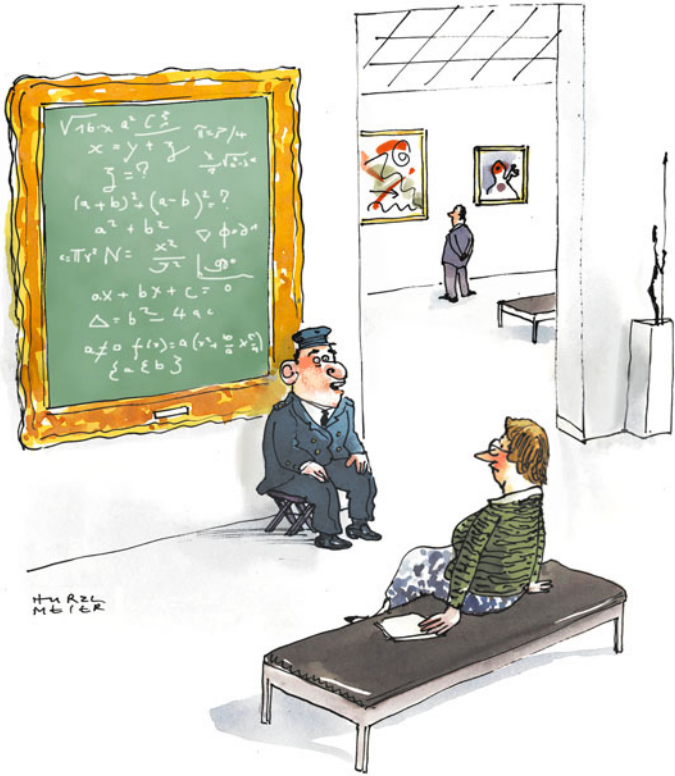
Geben Sie gegebenenfalls den Kern und das Bild von f an.

[Unterscheiden Sie Körper, in denen $1 + 1 = 0$ gilt, von allen anderen.]

33. Bestimmen Sie die Ordnungen folgender geometrischer Abbildungen der affinen Ebenen über \mathbf{R} in sich: Spiegelung, Drehung um den Winkel $360^\circ/n$, Punktspiegelung.

Sie sollten mit folgenden Begriffen umgehen können

Gruppe, abelsche Gruppe, Ordnung einer Gruppe, Untergruppe, Index, Ordnung eines Elements, Normalteiler, Faktorgruppe, Homomorphismus von Gruppen, Homomorphiesatz.



“ Ein früher Gödel, angeblich...”

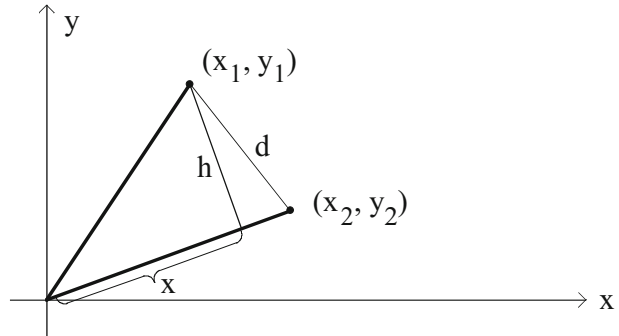
In diesem Kapitel beschäftigen wir uns mit den möglichen (und sinnvollen) Verhältnissen, die zwei Vektoren miteinander haben können. Eine geometrische Betrachtungsweise ist der Winkel, den zwei Vektoren einschließen. Wir betrachten in diesem Kapitel nur solche Verhältnisse von Vektoren, die sich durch ein Körperelement beschreiben lassen. Mit anderen Worten: Wir betrachten also Abbildungen von $V \times V$ in K ; solche Abbildungen nennt man traditionell **Formen**. Unglücklicherweise gibt es einen ganzen Zoo von verschiedenen und zu unterscheidenden Formen. Ich versuche, Ihnen dadurch Orientierung zu verschaffen, dass ich Ihnen zunächst nur eine Hälfte dieses Zoos zeige, nämlich denjenigen Teil, dessen Ziel das Studium der Skalarprodukte in *reellen* Vektorräumen ist. Sie sind eingeladen, die parallele Theorie für Skalarprodukte in *komplexen* Vektorräumen im „Projekt“ dieses Kapitels zu entwickeln.

10.1 Ein Beispiel

Wir beginnen mit einem *Beispiel* aus der Geometrie. Wir legen dazu die kartesische Ebene zugrunde. Vielleicht erinnern Sie sich noch aus der Schule oder sonst woher, was man unter einem „skalaren Produkt“ versteht. Seien $v_1 = (x_1, y_1)$ und $v_2 = (x_2, y_2)$ Vektoren in der reellen Ebene. Dann gibt es – *zwei Definitionen* für das skalare Produkt $\langle v_1, v_2 \rangle$ der Vektoren v_1 und v_2 . Wie bitte? Ganz einfach: Das bedeutet nicht, dass es zwei konkurrierende Definitionen gibt, die zu verschiedenen Strukturen führen. Vielmehr gibt es zwei verschiedene Beschreibungen derselben Sache. Und unsere Aufgabe ist es, die Äquivalenz dieser Beschreibungen nachzuweisen. (Vergleichen Sie hierzu die Diskussion um die „drei Definitionen“ einer Basis in Abschn. 3.3.1.)

Manchmal wird das skalare Produkt definiert als

$$\langle v_1, v_2 \rangle = x_1 x_2 + y_1 y_2 ;$$

Abb. 10.1 Das skalare Produkt

an anderen Stellen liest man

$$\langle v_1, v_2 \rangle = \|v_1\| \cdot \|v_2\| \cdot \cos(\varphi) ,$$

wobei φ der von v_1 und v_2 eingeschlossene Winkel ist.

Ist das das Gleiche? Wenn ja, wäre das sehr schön. Denn mit Hilfe der „ersten Definition“ kann man das skalare Produkt leicht ausrechnen; die „zweite Definition“ zeigt, dass das skalare Produkt sehr genau Längen von Vektoren und Winkel zwischen Vektoren beschreibt.

Wir überlegen uns jetzt, dass die „beiden Definitionen“ den gleichen Wert liefern. Seien dazu $v_1 = (x_1, y_1)$ und $v_2 = (x_2, y_2)$ zwei Vektoren der Länge 1. Das bedeutet $\sqrt{x_1^2 + y_1^2} = 1$ und $\sqrt{x_2^2 + y_2^2} = 1$. (Diese Einschränkung dient ausschließlich dazu, dass wir uns beim Rechnen leichter tun; Sie können aber auch mit Vektoren beliebiger Länge rechnen; siehe Übungsaufgabe 1.) Sei φ der von den Vektoren v_1 und v_2 eingeschlossene Winkel.

Einen einfachen Spezialfall behandeln wir vorab: Wenn $v_1 = v_2$ ist, so gilt

$$\langle v_1, v_2 \rangle = x_1^2 + y_1^2 = \|v_1\|^2 = \|v_1\| \cdot \|v_2\| \cdot \cos(0^\circ) ,$$

da $\cos(0^\circ) = 1$ ist.

Nun betrachten wir den allgemeinen Fall $v_1 \neq v_2$. Mit den Bezeichnungen von Abb. 10.1 gilt dann $\cos(\varphi) = x$. Also ist zu zeigen, dass $x = x_1 x_2 + y_1 y_2$ gilt.

Dazu wenden wir zweimal den Satz des Pythagoras an. Zum einen gilt

$$x^2 + h^2 = 1 ,$$

zum anderen

$$(1 - x)^2 + h^2 = d^2 .$$

Daraus ergibt sich

$$2 - 2x = d^2 .$$

Wir wissen aber auch

$$\begin{aligned} d^2 &= (x_1 - x_2)^2 + (y_1 - y_2)^2 = x_1^2 - 2x_1x_2 + x_2^2 + y_1^2 - 2y_1y_2 + y_2^2 \\ &= x_1^2 + y_1^2 + x_2^2 + y_2^2 - 2x_1x_2 - 2y_1y_2 = 1 + 1 - 2x_1x_2 - 2y_1y_2 . \end{aligned}$$

Zusammen folgt

$$2x = 2 - d^2 = 2 - 2 + 2x_1x_2 + 2y_1y_2 = 2x_1x_2 + 2y_1y_2 ,$$

also

$$\cos(\varphi) = x = x_1x_2 + y_1y_2 .$$

Daher beschreiben beide „Definitionen“ denselben Sachverhalt. Damit ist auch ein Nachteil der ersten Definition behoben; denn diese hängt von der Wahl einer Basis (hier der Basis aus Einheitsvektoren) ab.

An diesem Beispiel erkennen wir verschiedene Eigenschaften, die uns Anlass zu Fragen geben.

- Hat jeder Vektorraum ein „Skalarprodukt“?
- Wir sehen in obigem Beispiel, dass das Skalarprodukt zweier Vektoren genau dann Null ist, wenn diese senkrecht aufeinander stehen. Dies wollen wir in allgemeinem Rahmen nachmachen. Kann man diesen Sachverhalt auf beliebige „Skalarprodukte“ verallgemeinern?
- In unserem Eingangsbeispiel hat die Basis aus Einheitsvektoren die Eigenschaft, dass je zwei Basisvektoren aufeinander senkrecht stehen und jeder Basisvektor die Länge 1 hat. Gibt es für jedes „Skalarprodukt“ eine solche Basis?
- Welche linearen Abbildungen „respektieren“ das Skalarprodukt? Das bedeutet: Welche linearen Abbildungen verändern die Vektoren nur so, dass der Winkel zwischen zwei Vektoren vorher und nachher gleich groß ist? Formaler: Welches sind die linearen Abbildungen f von V in sich mit der Eigenschaft

$$\langle f(v), f(w) \rangle = \langle v, w \rangle \quad \text{für alle } v, w \in V ?$$

10.2 Bilinearformen

Wir werden uns den Skalarprodukten in zwei Stufen nähern. Im diesem Abschnitt studieren wir so genannte Bilinearformen, die aber nur eine Vorform der Skalarprodukte sind, deren Eigenschaften wir ab Abschn. 10.3 genau studieren werden.

Sei stets V ein K -Vektorraum endlicher Dimension n . Eine **Bilinearform** ist eine Abbildung von $V \times V$ in K , die gewissen Eigenschaften genügt. Das Bild des Paares (v, w) von Vektoren unter der Bilinearform bezeichnen wir mit $\langle v, w \rangle$. (Das könnte theoretisch zu

Verwechslungen mit dem von v und w erzeugten Unterraum führen; das wird aber in der Praxis nicht passieren.)

Nun heben wir nochmals an: Eine **Bilinearform** von V ist eine Abbildung

$$\langle , \rangle : V \times V \rightarrow K ,$$

die **bilinear** ist; das bedeutet, dass für alle $v, v', w, w' \in V$ und alle $k \in K$ gilt:

$$\begin{aligned} \langle v + v', w \rangle &= \langle v, w \rangle + \langle v', w \rangle, & \langle k \cdot v, w \rangle &= k \cdot \langle v, w \rangle \quad \text{und} \\ \langle v, w + w' \rangle &= \langle v, w \rangle + \langle v, w' \rangle, & \langle v, kw \rangle &= k \cdot \langle v, w \rangle . \end{aligned}$$

„Bilinear“ bedeutet nichts anderes, als dass die Form in beiden Komponenten linear ist. Mit anderen Worten: Eine Abbildung $\langle , \rangle : V \times V \rightarrow K$ ist eine Bilinearform, falls für alle Vektoren v und w aus V gilt: Die Abbildung $f_v : V \rightarrow K$ mit $f_v(w) := \langle v, w \rangle$ und die Abbildung $g_w : V \rightarrow K$ mit $g_w(v) := \langle v, w \rangle$ sind lineare Abbildungen (also Linearformen; siehe Abschn. 5.4).

Beispiele Wir geben sofort eine riesige Menge von Beispielen an. Dazu wählen wir irgendeine Basis $B = \{v_1, \dots, v_n\}$ von V . Ferner sei $A = (a_{ij})$ eine $n \times n$ -Matrix mit Werten aus K . Wir definieren eine Abbildung \langle , \rangle von $V \times V$ in K dadurch, dass wir für Basisvektoren v_i und v_j definieren

$$\langle v_i, v_j \rangle := a_{ij}$$

und diese Vorschrift „linear fortsetzen“.

Das bedeutet: Wenn

$$v = \sum_{i=1}^n k_i v_i \quad \text{und} \quad w = \sum_{j=1}^n h_j v_j$$

beliebige Vektoren aus V sind, so ist

$$\langle v, w \rangle := \sum_{i,j=1}^n k_i a_{ij} h_j .$$

Wir können diese Vorschrift noch einprägsamer formulieren. Dazu erlauben wir uns eine kleine Schlamperei (vornehm „abuse of notation“ genannt). Wir bezeichnen nämlich den Koordinatenvektor (k_1, k_2, \dots, k_n) des Vektors v (bezüglich der Basis B) ebenfalls mit v ; entsprechend sei mit w auch der Koordinatenvektor von w bezeichnet. Dann ist

$$\langle v, w \rangle = v \cdot A \cdot w^T .$$

Es handelt sich also nur um die Multiplikation der Vektoren mit einer festen Matrix. Nun zeigen wir, dass die so definierte Abbildung \langle , \rangle wirklich eine Bilinearform ist:

Konstruktion von Bilinearformen

Durch die Vorschrift

$$\langle v, w \rangle := \sum_{i,j=1}^n k_i a_{ij} h_j$$

wird eine Bilinearform auf V erklärt. Man nennt $\langle \cdot, \cdot \rangle$ die (bezüglich der Basis B) **zu A gehörige** Bilinearform.

Beweis Offenbar ist es genau so einfach, die Linearität in der ersten Komponente wie die Linearität in der zweiten Komponente nachzuweisen. Deshalb betrachten wir o.B.d.A. nur die erste Komponente. Seien

$$v = \sum_{i=1}^n k_i v_i, v' = \sum_{i=1}^n k'_i v_i \quad \text{und} \quad w = \sum_{j=1}^n h_j v_j.$$

Dann gilt

$$\langle v + v', w \rangle = \sum_{i,j=1}^n (k_i + k'_i) a_{ij} h_j = \sum_{i,j=1}^n k_i a_{ij} h_j + \sum_{i,j=1}^n k'_i a_{ij} h_j = \langle v, w \rangle + \langle v', w \rangle.$$

Also ist $\langle \cdot, \cdot \rangle$ in der ersten Komponente additiv. Ebenso leicht kann man zeigen, dass $\langle \cdot, \cdot \rangle$ in der ersten Komponente auch homogen ist.

Also ist die im Satz definierte Abbildung $\langle \cdot, \cdot \rangle$ wirklich eine Bilinearform. \square

Wenn $\langle \cdot, \cdot \rangle$ eine Bilinearform des n -dimensionalen Vektorraums V ist, so heißt die Matrix $A = (\langle v_i, v_j \rangle)_{1 \leq i, j \leq n}$ die zu $\langle \cdot, \cdot \rangle$ gehörige **Gramsche Matrix**. Diese wird nach dem dänischen Versicherungsmathematiker J rgen Pedersen Gram (1850–1916) benannt. In manchen B chern hei t A auch **Strukturmatrix** von $\langle \cdot, \cdot \rangle$.

Selbstverst ndlich ist es –  hnlich wie bei Darstellungsmatrizen von linearen Abbildungen – eine wichtige Frage, wie durch Wahl einer geeigneten Basis die Gramsche Matrix besonders einfach wird. Noch genauer kann man fragen, unter welchen Umst nden man erreichen kann, dass man eine Gramsche Matrix in Diagonalgestalt erzielen kann.

Besonders einfach zu beschreibende Bilinearformen sind die folgenden. Sei $\{v_1, \dots, v_n\}$ eine Basis von V . F r zwei Vektoren

$$v = \sum_{i=1}^n k_i v_i \quad \text{und} \quad w = \sum_{j=1}^n h_j v_j$$

setzen wir

$$\langle v, w \rangle := f(v, w) := k_1 \varepsilon_1 h_1 + k_2 \varepsilon_2 h_2 + \dots + k_n \varepsilon_n h_n,$$

wobei $\varepsilon_i = \pm 1$ ist. (Vergleichen Sie hierzu auch  bungsaufgabe 8.)

Am einfachsten zu beschreiben ist die **Standardbilinearform**, bei der alle $\varepsilon_i = 1$ sind:

$$\langle v, w \rangle := f(v, w) := k_1 h_1 + k_2 h_2 + \dots + k_n h_n .$$

Mit anderen Worten: Die Gramsche Matrix der Standardbilinearform ist die Einheitsmatrix. Da wir später sehen werden, dass diese Bilinearform auch ein Skalarprodukt ist, sprechen wir auch vom **Standardskalarprodukt**.

Wir werden demnächst einen wichtigen Satz über die Gramsche Matrix beweisen. Dazu brauchen wir aber noch eine Definition.

Sei $\langle \cdot, \cdot \rangle$ eine Bilinearform auf V . Wir sagen, dass zwei Vektoren v, w aus V **senkrecht aufeinander** stehen (**orthogonal** sind), wenn $\langle v, w \rangle = 0$ ist. Wir schreiben dafür auch $v \perp w$.

Man kann sich leicht überlegen, dass der Nullvektor auf allen anderen Vektoren senkrecht steht (siehe Übungsaufgabe 3). Wir nennen die Bilinearform $\langle \cdot, \cdot \rangle$ **nichtausgeartet**, falls der Nullvektor der einzige Vektor ist, der auf jedem Vektor senkrecht steht. Das bedeutet: $\langle \cdot, \cdot \rangle$ ist nichtausgeartet, falls gilt:

$$\langle v_0, w \rangle = 0 \quad \text{für alle } w \in V \Rightarrow v_0 = o .$$

Bemerkung Bei der Definition einer nichtausgearteten Bilinearform haben wir die beiden Komponenten nicht gleichberechtigt behandelt, jedenfalls nicht formal. Man kann jedoch zeigen, dass diese Definition in Wirklichkeit keine Benachteiligung der einen oder anderen Komponente darstellt. Da wir uns im Wesentlichen auf „symmetrische“ Bilinearformen konzentrieren werden, spielt aber diese Finesse für uns ohnedies keine Rolle.

Rang einer Gramschen Matrix

Sei $\{v_1, v_2, \dots, v_n\}$ eine Basis des Vektorraums V , und sei $A = (\langle v_i, v_j \rangle)_{1 \leq i, j \leq n}$ die Gramsche Matrix einer Bilinearform $\langle \cdot, \cdot \rangle$ von V . Dann gilt: Genau dann hat A den Rang n (also Determinante $\neq 0$), wenn die Bilinearform $\langle \cdot, \cdot \rangle$ nichtausgeartet ist.

Beweis Zunächst nehmen wir an, dass A einen Rang $< n$ hat. Wir müssen zeigen, dass $\langle \cdot, \cdot \rangle$ ausgeartet ist. Dazu genügt es, einen von Null verschiedenen Vektor x anzugeben, der auf allen Vektoren der Basis $\{v_1, v_2, \dots, v_n\}$ senkrecht steht.

Wenn A einen Rang $< n$ hat, so sind die Zeilen von A linear abhängig. Daher gibt es Körperelemente k_1, \dots, k_n , so dass die entsprechende Linearkombination der Zeilen von A Null ist. Das bedeutet:

$$\begin{aligned} k_1 \cdot \langle v_1, v_1 \rangle + k_2 \cdot \langle v_2, v_1 \rangle + \dots + k_n \cdot \langle v_n, v_1 \rangle &= 0 , \\ k_1 \cdot \langle v_1, v_2 \rangle + k_2 \cdot \langle v_2, v_2 \rangle + \dots + k_n \cdot \langle v_n, v_2 \rangle &= 0 , \\ &\dots \\ k_1 \cdot \langle v_1, v_n \rangle + k_2 \cdot \langle v_2, v_n \rangle + \dots + k_n \cdot \langle v_n, v_n \rangle &= 0 . \end{aligned}$$

Nun betrachten wir den Vektor

$$x := k_1 v_1 + k_2 v_2 + \dots + k_n v_n .$$

Nach Annahme und Wahl der k_i ist $x \neq o$. Wir behaupten, dass x auf jedem v_j senkrecht steht! Dies müssen wir einfach ausrechnen:

$$\begin{aligned} \langle x, v_j \rangle &= \langle k_1 v_1 + k_2 v_2 + \dots + k_n v_n, v_j \rangle \\ &= k_1 \langle v_1, v_j \rangle + k_2 \langle v_2, v_j \rangle + \dots + k_n \langle v_n, v_j \rangle = 0 \end{aligned}$$

Wegen der obigen Gleichungen. Also steht x tatsächlich auf allen Vektoren aus V senkrecht, und damit ist $\langle \cdot, \cdot \rangle$ ausgeartet.

Der Beweis der Umkehrung verwendet eigentlich die gleichen Argumente: Wir nehmen an, dass die Bilinearform $\langle \cdot, \cdot \rangle$ ausgeartet ist. Dann gibt es einen von Null verschiedenen Vektor x , der auf v_1, v_2, \dots, v_n senkrecht steht. Sei $x = k_1 v_1 + k_2 v_2 + \dots + k_n v_n$. Da $x \neq o$ ist, ist mindestens ein $k_i \neq 0$. Dann gilt für alle $j \in \{1, 2, \dots, n\}$:

$$\begin{aligned} 0 &= \langle x, v_j \rangle = \langle k_1 v_1 + k_2 v_2 + \dots + k_n v_n, v_j \rangle \\ &= k_1 \langle v_1, v_j \rangle + k_2 \langle v_2, v_j \rangle + \dots + k_n \langle v_n, v_j \rangle . \end{aligned}$$

Dies bedeutet aber nichts anderes, als dass die Zeilen der Gramschen Matrix A linear abhängig sind; somit ist der Rang von A kleiner als n . \square

Wir werden nun den erstaunlich engen Zusammenhang zwischen den soeben definierten Bilinearformen und den schon lange bekannten Linearformen herstellen. (Nur, falls Sie das zufällig vergessen haben sollten: Eine Linearform ist eine lineare Abbildung von V in K ; die Menge aller Linearformen nennt man den Dualraum. Vergleichen Sie dazu Abschn. 5.4.) Zunächst zeigen wir, dass jede Bilinearform nicht nur eine, nicht nur zwei, nein: sogar viele Linearformen liefert.

Bilinearformen kommen von Linearformen

Sei $\langle \cdot, \cdot \rangle$ eine Bilinearform auf V . Sei v_0 ein beliebiger Vektor. Wir definieren die Abbildung f_{v_0} von V nach K durch

$$f_{v_0}(w) = \langle v_0, w \rangle \quad \text{für alle } w \in V .$$

Dann ist f_{v_0} eine Linearform von V . Wenn $\langle \cdot, \cdot \rangle$ nicht ausgeartet ist, so ist f_{v_0} genau dann die Nullabbildung, wenn $v_0 = o$ ist.

Entsprechend ist die Abbildung g_{w_0} von V nach K , die durch

$$g_{w_0}(v) = \langle v, w_0 \rangle \quad \text{für alle } v \in V$$

definiert ist, eine lineare Abbildung, die (für den Fall, dass $\langle \cdot, \cdot \rangle$ nichtausgeartet ist) nur dann die Nullabbildung ist, wenn $w_0 = o$ ist.

Beweis Seien $w, w' \in V$. Dann ist

$$f_{v_0}(w + w') = \langle v_0, w + w' \rangle = \langle v_0, w \rangle + \langle v_0, w' \rangle = f_{v_0}(w) + f_{v_0}(w'),$$

da $\langle \cdot, \cdot \rangle$ bilinear ist. Entsprechend folgt für alle $k \in K$

$$f_{v_0}(k \cdot w) = \langle v_0, kw \rangle = k \cdot \langle v_0, w \rangle = k \cdot f_{v_0}(w).$$

Die Abbildung f_{v_0} ist genau dann die Nullabbildung, wenn $\langle v_0, w \rangle = 0$ ist für alle $w \in V$. Da $\langle \cdot, \cdot \rangle$ nichtausgeartet ist, kann dies nur dann der Fall sein, wenn v_0 der Nullvektor ist.

Ähnlich zeigt man die Behauptungen für g_{w_0} . (Achtung: Die Tatsache, dass g_{w_0} genau dann die Nullabbildung ist, wenn $w_0 = o$ ist, ist schwieriger zu zeigen; siehe Übungsaufgabe 9.) \square

Es gilt sogar noch mehr:

Bilinearformen und Abbildungen von V nach V^*

Die Abbildung F von V in den Dualraum V^* , die definiert ist durch

$$F(v) := f_v,$$

ist eine lineare Abbildung; sie ist genau dann injektiv, also ein Isomorphismus, wenn $\langle \cdot, \cdot \rangle$ nichtausgeartet ist.

Entsprechend gilt: Die Abbildung G von V in den Dualraum V^* , die definiert ist durch $G(w) := g_w$, ist eine lineare Abbildung; sie ist genau dann injektiv, also ein Isomorphismus, wenn $\langle \cdot, \cdot \rangle$ nichtausgeartet ist.

Beweis Wir zeigen die *Additivität* der Abbildung F : Seien $v, v' \in V$. Wir müssen beweisen, dass $F(v+v') = f_{v+v'}$ und $F(v) + F(v') = f_v + f_{v'}$ dieselbe Linearform sind. Dazu müssen wir zeigen, dass für alle $w \in V$ die Werte $[F(v+v')](w) = (f_{v+v'})(w)$ und $[F(v) + F(v')](w) = [f_v + f_{v'}](w)$ gleich sind. Und wie so oft stellt sich heraus: Wenn man das Problem erst mal so weit analysiert hat, ist die Lösung ganz einfach:

$$\begin{aligned} [F(v + v')](w) &= (f_{v+v'})(w) = \langle v + v', w \rangle \\ &= \langle v, w \rangle + \langle v', w \rangle = [f_v + f_{v'}](w) = [F(v) + F(v')](w). \end{aligned}$$

Die *Homogenität* von F zeigt man entsprechend.

Es bleibt noch zu zeigen, dass F genau dann injektiv ist, wenn \langle , \rangle nichtausgeartet ist: Die lineare Abbildung F ist genau dann injektiv, wenn ihr Kern nur aus dem Nullvektor besteht. Der Kern von F enthält aber genau dann einen Vektor $v \neq 0$, wenn f_v die Nullabbildung ist. Schließlich gibt es aber genau dann einen Vektor $v \neq 0$ mit $\langle v, w \rangle = f_v(w) = 0$ für alle $w \in V$, wenn \langle , \rangle ausgeartet ist.

Damit ist alles gezeigt. \square

Nun könnten Sie sagen: „Schön und gut; Bilinearformen und erst recht die später studierten Skalarprodukte sind etwas Vertrautes; dass diese merkwürdigerweise Linearformen induzieren, mag vielleicht einen Mathefreak begeistern, aber die ‚normalen‘ Skalarprodukte sind doch etwas ganz anderes!“ Weit gefehlt! Beide Konzepte sind äquivalent: Wir können auch aus jeder Linearform eine Bilinearform konstruieren!

Linearformen induzieren Bilinearformen

Sei F eine lineare Abbildung von V nach V^* . Dann wird durch

$$\langle v, w \rangle := (F(v))(w)$$

eine Bilinearform von V definiert. Das bedeutet: Um $\langle v, w \rangle$ auszurechnen, bestimmt man zuerst die Linearform $F(v)$ und wendet dann diese auf den Vektor w an.

Ferner gilt: Genau dann ist \langle , \rangle nichtausgeartet, wenn F bijektiv ist.

Beweis. Dass \langle , \rangle in der zweiten Komponente linear ist, folgt aus der Tatsache, dass $F(v)$ eine lineare Abbildung von V nach K ist. Die Additivität in der ersten Komponente ergibt sich aus der Additivität der Abbildung F :

$$\begin{aligned} \langle v + v', w \rangle &= (F(v + v'))(w) = (F(v) + F(v'))(w) = (F(v))(w) + (F(v'))(w) \\ &= \langle v, w \rangle + \langle v', w \rangle. \end{aligned}$$

Ganz entsprechend zeigt man die Homogenität in der ersten Komponente.

Die Bilinearform \langle , \rangle ist genau dann ausgeartet, wenn es einen Vektor $v_0 \neq 0$ gibt, so dass $\langle v_0, w \rangle = (F(v_0))(w) = 0$ ist für alle $w \in V$, also genau dann, wenn $F(v_0)$ die Nullabbildung ist. Das bedeutet: Genau dann ist \langle , \rangle ausgeartet, wenn die Linearform F nicht injektiv ist.

Da $\dim(V) = \dim(V^*)$ ist, folgt also auch: Genau dann ist \langle , \rangle nichtausgeartet, wenn F injektiv, also genau dann, wenn F bijektiv ist. \square

Mit diesem Satz können wir noch mehr zeigen. Bislang haben wir jeweils nur *eine* Bilinearform betrachtet. Wir wissen aber, dass es viele gibt. Inwiefern können sich zwei Bilinearformen überhaupt unterscheiden? Können sie „grundsätzlich verschieden“ sein? Oder

sind sie immer „im wesentlichen gleich“? – Der folgende Satz gibt eine Antwort auf diese Fragen.

Verschiedene Bilinearformen von V („Wer eine kennt, kennt alle“)

Sei $\langle \cdot, \cdot \rangle$ eine nichtausgeartete Bilinearform von V .

(a) Sei f eine lineare Abbildung von V in sich. Dann wird durch die Vorschrift

$$[v, w] := \langle f(v), w \rangle$$

eine Bilinearform von V definiert.

(b) Sei $[\cdot, \cdot]$ eine zweite Bilinearform von V . Dann gibt es eine lineare Abbildung φ von V in sich, so dass

$$[v, w] = \langle \varphi(v), w \rangle$$

für alle $v, w \in V$ gilt. Ferner gilt: Genau dann ist φ bijektiv, wenn auch die Bilinearform $[\cdot, \cdot]$ nichtausgeartet ist.

Kurz: Bilinearformen von V unterscheiden sich nur durch eine lineare Abbildung.

Beweis Die Behauptung von (a) ist einfach nachzurechnen; Sie sind eingeladen, dies zu Ihrer Übung in Übungsaufgabe 4 zu tun.

In (b) steckt die Hauptaussage des Satzes.

Wir wissen, dass es eine lineare Abbildung F von V nach V^* gibt mit $\langle v, w \rangle = (F(v))(w)$ für alle $v, w \in V$; da die Bilinearform $\langle \cdot, \cdot \rangle$ nichtausgeartet ist, ist F bijektiv, also invertierbar.

Ebenso gibt es für die Bilinearform $[\cdot, \cdot]$ eine lineare Abbildung F^* von V nach V^* mit $(F^*(v))(w) = [v, w]$ für alle $v, w \in V$.

Nun setzen wir diese Abbildungen zusammen; genauer gesagt: Wir betrachten die Abbildung $\varphi := F^{-1} \cdot F^*$. Als Komposition der linearen Abbildungen F^* und F^{-1} ist φ eine (lineare) Abbildung von V nach V . Ferner ist φ genau dann bijektiv, wenn F^* bijektiv ist, was, wie wir wissen, genau dann der Fall ist, wenn die Bilinearform $[\cdot, \cdot]$ nichtausgeartet ist.

Es bleibt nur zu zeigen, dass man mit Hilfe von φ tatsächlich die Bilinearform $[\cdot, \cdot]$ aus $\langle \cdot, \cdot \rangle$ gewinnen kann. Dazu rechnen wir einfach $\langle \varphi(v), w \rangle$ aus:

$$\langle \varphi(v), w \rangle = (F(\varphi(v)))(w) = (F(F^{-1} \cdot F^*(v)))(w) = (F^*(v))(w) = [v, w].$$

Damit ist auch dieser Satz bewiesen. □

Nun wenden wir uns ernsthaft dem Begriff des Senkrechtstehens zu. Sei dazu stets $\langle \cdot, \cdot \rangle$ eine Bilinearform auf dem Vektorraum V .

Die Menge aller zu einem Vektor v orthogonalen Vektoren bezeichnet man mit v^\perp . Allgemeiner definieren wir für eine beliebige Teilmenge X von V

$$X^\perp := \{v \in V \mid \langle x, v \rangle = 0 \text{ für alle } x \in X\}.$$

Als leichte Übung beweisen wir folgende Aussagen:

Satz über orthogonale Unterräume

Sei $\langle \cdot, \cdot \rangle$ eine Bilinearform. Seien X und Y Teilmengen von V . Dann gelten folgende Aussagen:

- (a) Stets ist X^\perp ein Unterraum von V ; ferner ist $v^\perp = (kv)^\perp$ für jedes von Null verschiedene Körperelement k .
- (b) Wenn $X \subseteq Y$ gilt, dann ist $Y^\perp \subseteq X^\perp$.
- (c) Es gilt $X^\perp = \langle X \rangle^\perp$. (Achtung: Hier kommen die spitzen Klammern in zwei Bedeutungen vor: Zur Bezeichnung der Bilinearform und zur Bezeichnung eines Erzeugnisses!)
- (d) Wenn $\langle \cdot, \cdot \rangle$ symmetrisch ist, so gilt $X \subseteq X^{\perp\perp}$.
- (e) Wenn $\langle \cdot, \cdot \rangle$ symmetrisch ist, so gilt $X^{\perp\perp\perp} = X^\perp$.

Beweis

- (a) Wir wenden das Unterraumkriterium an. Es ist $0 \in X^\perp$ (siehe Übungsaufgabe 3). Seien $v, w \in X^\perp$. Dann gilt für alle $x \in X$:

$$\langle x, v - w \rangle = \langle x, v \rangle - \langle x, w \rangle = 0 - 0 = 0,$$

also $v - w \in X^\perp$. Ähnlich zeigt man, dass mit v auch $k \cdot v \in X^\perp$ ist. Also ist X^\perp ein Unterraum.

Wegen

$$\langle v, w \rangle = 0 \Leftrightarrow \langle k \cdot v, w \rangle = k \cdot \langle v, w \rangle = 0$$

folgt auch die zweite Behauptung.

- (b) Sei $X \subseteq Y$. Dann steht jeder Vektor v , der zu jedem $y \in Y$ orthogonal ist, auch auf jedem $x \in X$ senkrecht. Mit anderen Worten

$$Y^\perp = \{v \in V \mid \langle y, v \rangle = 0 \text{ für alle } y \in Y\} \subseteq \{v \in V \mid \langle x, v \rangle = 0 \text{ für alle } x \in X\} = X^\perp.$$

(c) Wegen $X \subseteq \langle X \rangle$ folgt mit (b)

$$\langle X \rangle^\perp \subseteq X^\perp .$$

Aber jeder Vektor, der senkrecht auf jedem Vektor x des Erzeugendensystems X steht, ist wegen der Bilinearität der Form auch in $\langle X \rangle^\perp$ enthalten.

(d) Sei $x \in X$. Wir müssen zeigen, dass x auch in $X^{\perp\perp}$ liegt. Wegen $x \in X$ gilt für alle $v \in X^\perp$:

$$\langle x, v \rangle = 0 .$$

Da $\langle \cdot, \cdot \rangle$ symmetrisch ist, ergibt sich also

$$\langle v, x \rangle = 0 \quad \text{für alle } v \text{ aus } X^\perp .$$

Nach Definition heißt dies $x \in (X^\perp)^\perp = X^{\perp\perp}$.

(e) Nach (d) ist $X \subseteq X^{\perp\perp}$, also nach (b)

$$X^{\perp\perp\perp} \subseteq X^\perp .$$

Setzt man $Y := X^\perp$, so folgt, wenn man abermals (d) anwendet

$$X^\perp = Y \subseteq Y^{\perp\perp} = X^{\perp\perp\perp} .$$

Zusammen ergibt sich $X^{\perp\perp\perp} = X^\perp$. □

Teil (d) des obigen Satzes sagt, dass für einen Unterraum U von V höchstens die Unterräume U , U^\perp und $U^{\perp\perp}$ verschieden sind. Unser nächstes Ziel ist zu zeigen, dass (für eine nichtausgeartete Bilinearform) sogar $U^{\perp\perp} = U$ gilt; daher sind höchstens U und U^\perp verschieden. Dazu benötigen wir das folgende Ergebnis.

Dimensionsformel für Bilinearformen

Sei $\langle \cdot, \cdot \rangle$ eine Bilinearform eines n -dimensionalen Vektorraums V . Wenn $\langle \cdot, \cdot \rangle$ nichtausgeartet ist, so gilt für alle Unterräume U von V

$$\dim(U) + \dim(U^\perp) = n .$$

Beweis Wir wählen eine Basis $\{v_1, \dots, v_m\}$ von U und ergänzen diese zu einer Basis $\{v_1, \dots, v_m, v_{m+1}, \dots, v_n\}$ von V . Wir müssen zeigen: $\dim(U^\perp) = n - m$.

Der Unterraum U^\perp besteht genau aus den Vektoren

$$x = \sum_{i=1}^n k_i v_i \quad \text{mit} \quad \left\langle \sum_{i=1}^n k_i v_i, v_j \right\rangle = 0 \quad \text{für } j = 1, 2, \dots, m .$$

Das bedeutet

$$k_1 \langle v_1, v_j \rangle + k_2 \langle v_2, v_j \rangle + \dots + k_n \langle v_n, v_j \rangle = 0 \text{ für } j = 1, \dots, m.$$

Dies ist ein homogenes lineares Gleichungssystem in den Unbekannten k_1, k_2, \dots, k_n . Um die Dimension des Lösungsraums bestimmen zu können, müssen wir den Rang der Matrix A des Gleichungssystems kennen. Da die Bilinearform nach Voraussetzung nichtausgeartet ist, hat die Gramsche Matrix den Rang n . Daraus ergibt sich, dass die m Zeilen der Matrix des Gleichungssystems linear unabhängig sind; also hat die Matrix den Rang m .

Daher hat der Lösungsraum U^\perp die Dimension $n - \text{Rang}(A) = n - m$. \square

Daraus ergibt sich ohne Schwierigkeit die folgende Aussage.

Korollar

Sei $\langle \cdot, \cdot \rangle$ eine nichtausgeartete symmetrische Bilinearform des Vektorraums V . Dann gilt für jeden Unterraum U von V .

$$U^{\perp\perp} = U.$$

Beweis Sei $\dim(V) = n$. Wir wenden die Dimensionsformel für Bilinearformen zweimal an und erhalten

$$\dim(U^{\perp\perp}) = n - \dim(U^\perp) = n - (n - \dim(U)) = \dim(U).$$

Aufgrund des Satzes über orthogonale Unterräume ist aber $U \subseteq U^{\perp\perp}$. Zusammen folgt die Behauptung. \square

10.3 Skalarprodukte

Wir nennen eine Bilinearform $\langle \cdot, \cdot \rangle$ eines K -Vektorraums V **symmetrisch**, wenn für alle Vektoren $v, w \in V$ gilt

$$\langle v, w \rangle = \langle w, v \rangle.$$

Konstruktion von symmetrischen Bilinearformen

Sei V ein n -dimensionaler K -Vektorraum, und sei $\{v_1, v_2, \dots, v_n\}$ eine Basis von V . Sei $A = (a_{ij})$ eine $n \times n$ -Matrix mit Einträgen aus K . Dann gilt: Genau dann wird durch

$$\langle v_i, v_j \rangle := a_{ij}$$

eine symmetrische Bilinearform definiert, wenn A eine symmetrische Matrix ist.

Kurz: Eine symmetrische Bilinearform erkennt man daran, dass ihre Gram-Matrix symmetrisch ist.

Beweis Aufgrund des Satzes über die Konstruktion von Bilinearformen wissen wir, dass $\langle \cdot, \cdot \rangle$ jedenfalls eine Bilinearform ist. Es bleibt also zu zeigen, dass $\langle \cdot, \cdot \rangle$ genau dann symmetrisch ist, wenn A eine symmetrische Matrix ist.

Zunächst setzen wir voraus, dass A eine symmetrische Matrix ist. Dann gilt für je zwei beliebige Vektoren

$$v = \sum_{i=1}^n k_i v_i \quad \text{und} \quad w = \sum_{i=1}^n h_i v_i$$

aus V :

$$\langle v, w \rangle = v = \sum_{i,j=1}^n k_i h_j \langle v_i, v_j \rangle = \sum_{i,j=1}^n h_j k_i \langle v_j, v_i \rangle = \langle w, v \rangle.$$

Also ist $\langle \cdot, \cdot \rangle$ symmetrisch.

Sei nun umgekehrt $\langle \cdot, \cdot \rangle$ eine symmetrische Bilinearform. Dann gilt insbesondere für je zwei Vektoren v_i, v_j einer Basis:

$$\langle v_i, v_j \rangle = \langle v_j, v_i \rangle;$$

Somit ist jede Strukturmatrix von $\langle \cdot, \cdot \rangle$ symmetrisch. □

Sei jetzt V ein reeller Vektorraum. Eine Bilinearform $\langle \cdot, \cdot \rangle$ von V heißt **positiv definit**, wenn für alle $v \in V$ gilt

$$\langle v, v \rangle \geq 0,$$

und $\langle v, v \rangle \neq 0$ für $v \neq 0$.

Ein **Skalarprodukt** eines reellen Vektorraums V ist eine symmetrische Bilinearform, die positiv definit ist.

Ein reeller Vektorraum mit Skalarprodukt wird ein **euklidischer Vektorraum** genannt.

Eine *Bemerkung* zur sprachlichen Formulierung: Man könnte den Skalar $\langle v, w \rangle$ das *Skalarprodukt* der Vektoren v und w nennen; hauptsächlich nennt man aber die Abbildung $\langle \cdot, \cdot \rangle$ *Skalarprodukt* (sprachlich besser wäre „Skalarmultiplikation“, aber dieser Begriff hat sich nicht eingebürgert).

Wir notieren schon hier, dass ein Skalarprodukt $\langle \cdot, \cdot \rangle$ jedenfalls eine nichtausgeartete Bilinearform ist. Denn, da $\langle \cdot, \cdot \rangle$ positiv definit ist, steht kein von Null verschiedener Vektor v auf sich selbst senkrecht; also ist v keinesfalls zu allen Vektoren aus V orthogonal. Daher kann $\langle \cdot, \cdot \rangle$ nicht ausgeartet sein.

Auch Skalarprodukte gibt es in Hülle und Fülle. Dies kommt im folgenden Satz zum Ausdruck.

Konstruktion von Skalarprodukten

Sei $\langle \cdot, \cdot \rangle$ eine symmetrische Bilinearform des reellen n -dimensionalen Vektorraums V , und sei A eine Strukturmatrix von $\langle \cdot, \cdot \rangle$. Dann gilt: Genau dann ist $\langle \cdot, \cdot \rangle$ ein Skalarprodukt, wenn für jedes $k \in \{1, 2, \dots, n\}$ die Matrix, die aus den ersten k Elementen in den ersten k Zeilen von A besteht, eine Determinante größer Null hat.

Mit diesem Satz kann man leicht Beispiele von Matrizen angeben, die Skalarprodukte definieren; wir geben einige solche Matrizen in den Fällen $n = 2$ und $n = 3$ an:

$$\begin{pmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{pmatrix}, \begin{pmatrix} 10 & 3 \\ 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix}, \begin{pmatrix} 10 & 3 & \frac{1}{2} \\ 3 & 1 & \frac{1}{10} \\ \frac{1}{2} & \frac{1}{10} & 1 \end{pmatrix}.$$

Die wichtigste Folgerung aus obigem Satz ist aber die folgende Behauptung:

Satz über das Standardskalarprodukt

Die Standardbilinearform eines reellen Vektorraums ist ein Skalarprodukt.

Beweis Die Standardbilinearform hat als Strukturmatrix die Einheitsmatrix. □

Sie sollten diesen Satz als Übungsaufgabe (ohne Rückgriff auf den vorigen Satz) beweisen; siehe Übungsaufgabe 10.

Wir beweisen den obigen Satz über die Konstruktion von Skalarprodukten nur im Spezialfall $n = 2$. Für den allgemeinen Fall siehe etwa Fischer [Fis], S. 217–218.

Sei

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \quad \text{mit} \quad a, b, c \in \mathbf{R}.$$

Es ist zu zeigen, dass die symmetrische Bilinearform mit Gram-Matrix A genau dann positiv definit ist, wenn $a > 0$ und $ac - b^2 > 0$ ist. Sei $\{v_1, v_2\}$ die Basis, bezüglich derer die Matrix A die Gramsche Matrix der Bilinearform $\langle \cdot, \cdot \rangle$ ist.

Zunächst setzen wir voraus, dass $a > 0$ und $ac - b^2 > 0$ gilt. Sei $v = k_1 v_1 + k_2 v_2$ ein beliebiger Vektor aus V . Dann gilt:

$$\begin{aligned}\langle v, v \rangle &= \langle k_1 v_1 + k_2 v_2, k_1 v_1 + k_2 v_2 \rangle \\ &= k_1^2 \langle v_1, v_1 \rangle + 2k_1 k_2 \langle v_1, v_2 \rangle + k_2^2 \langle v_2, v_2 \rangle = k_1^2 a + 2k_1 k_2 b + k_2^2 c \\ &> k_1^2 a + 2k_1 k_2 b + k_2^2 b^2 / a \\ &= \frac{1}{a} (k_1^2 a^2 + 2k_1 k_2 ab + k_2^2 b^2) = \frac{1}{a} (k_1 a + k_2 b)^2 \geq 0\end{aligned}$$

Also ist die symmetrische Bilinearform $\langle \cdot, \cdot \rangle$ positiv definit, und somit ein Skalarprodukt.

Sei nun umgekehrt $\langle v, v \rangle$ positiv für jeden Vektor $v \neq o$ aus V . Da insbesondere $\langle v_1, v_1 \rangle$ positiv ist, gilt $a = \langle v_1, v_1 \rangle > 0$.

Um herauszubekommen, dass auch $ac - b^2 > 0$ gilt, muss man einen Vektor v finden, so dass $\langle v, v \rangle = ac - b^2$ ist. Dafür haben wir zwei Möglichkeiten: *Entweder* Sie nehmen sich ein halbes Stündchen Zeit, probieren verschiedene Ansätze durch und finden schließlich einen solchen Vektor – *oder* Sie lassen mich probieren! ... ich habe tatsächlich schon ein bisschen vorgearbeitet und schlage Ihnen in erster Näherung den Vektor $w := bv_1 - av_2 \neq o$ vor. Jetzt sind Sie wieder dran: Sie müssen das Skalarprodukt $\langle w, w \rangle$ ausrechnen. Sie haben recht: Nichts leichter als das:

$$\begin{aligned}\langle w, w \rangle &= \langle bv_1 - av_2, bv_1 - av_2 \rangle \\ &= b^2 \langle v_1, v_1 \rangle - 2ab \langle v_1, v_2 \rangle + a^2 \langle v_2, v_2 \rangle \\ &= b^2 \cdot a - 2ab \cdot b + a^2 \cdot c = a(ac - b^2) .\end{aligned}$$

Wenn Sie jetzt noch w mit $1/a$ multiplizieren, erhalten Sie einen Vektor v mit $\langle v, v \rangle = ac - b^2$. Mit $\langle v, v \rangle$ muss auch $ac - b^2$ positiv sein. Also ist in der Tat $ac - b^2 > 0$. \square

Wir wenden uns nun wieder dem Begriff der Orthogonalität zu.

Satz vom orthogonalen Komplement

Sei V ein euklidischer Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$. Dann ist für jeden Unterraum U von V der Unterraum U^\perp ein zu U komplementärer Unterraum. Man nennt U^\perp auch das **orthogonale Komplement** von U .

Beweis Wir zeigen, dass sich U und U^\perp nur im Nullvektor schneiden: Sei $v \in U \cap U^\perp$. Da v in U^\perp liegt, steht v also auf jedem Vektor aus U senkrecht, also insbesondere auf v (da v in U liegt). Also gilt $\langle v, v \rangle = 0$. Da $\langle \cdot, \cdot \rangle$ positiv definit ist, ist also $v = o$.

Aufgrund der Dimensionsformel für Unterräume und der Dimensionsformel für Bilinearformen ergibt sich, da $\langle \cdot, \cdot \rangle$ nichtausgeartet ist:

$$\dim(\langle U, U^\perp \rangle) = \dim(U) + \dim(U^\perp) - 0 = \dim(V) .$$

Also sind U und U^\perp tatsächlich komplementäre Unterräume. \square

Obwohl aus geometrischer Sicht auch symmetrische Bilinearformen interessant sind, für die es Vektoren $v \neq o$ gibt mit $\langle v, v \rangle = 0$ (diese führen auf Quadriken; siehe Übungsaufgaben 2 und 8), werden wir uns zunächst auf Skalarprodukte beschränken. Wir haben schon gesehen, dass es viele Skalarprodukte gibt; wir werden aber zeigen, dass es im Wesentlichen nur ein einziges Skalarprodukt gibt. Wir werden nämlich zeigen, dass es zu jedem Skalarprodukt eine Basis gibt, so dass das Skalarprodukt bezüglich dieser Basis das Standardskalarprodukt ist. Das Hauptergebnis dieses Abschnitts wird die Konstruktion einer Basis sein, deren Vektoren senkrecht aufeinander stehen und jeweils die Länge 1 haben. Was „senkrecht“ bedeutet, wissen wir, wir wissen aber noch nicht, was die Länge eines Vektors sein soll. Dies müssen wir zunächst erklären.

Eine **Norm** des euklidischen Vektorraums V ist eine Abbildung $\| \cdot \|$ von V in die nicht-negativen reellen Zahlen mit folgenden Eigenschaften:

- (N1) $\|kv\| = |k| \cdot \|v\|$,
 (N2) $\|v + w\| \leq \|v\| + \|w\|$ (**Dreiecksungleichung**),
 (N3) $\|v\| = 0 \Leftrightarrow v = o$.

Wie kommt man von einem Skalarprodukt (was wir haben) zu einer Norm? Das ist ganz einfach; wir müssen dazu aber zuerst die berühmte Ungleichung von Cauchy und Schwarz beweisen.

In einem Vektorraum mit Skalarprodukt können wir die „euklidische Norm“ definieren:

$$\|v\| := \sqrt{\langle v, v \rangle} \quad \text{für alle } v \in V.$$

Unser Ziel ist zu zeigen, dass dadurch tatsächlich eine Norm erklärt wird.

Ungleichung von Cauchy und Schwarz

Sei V ein reeller Vektorraum mit Skalarprodukt. Dann gilt für alle $v, w \in V$

$$|\langle v, w \rangle| \leq \|v\| \cdot \|w\|;$$

Gleichheit gilt genau dann, wenn v und w linear abhängig sind.

Beweis Im Fall $w = o$ sind beide Seiten gleich Null; daher gilt die Ungleichung mit Gleichheit. Sei also $w \neq o$. Für alle $k \in \mathbb{R}$ gilt

$$\begin{aligned} 0 &\leq \langle v - kw, v - kw \rangle = \langle v, v \rangle - k\langle v, w \rangle - k\langle w, v \rangle + kk\langle w, w \rangle \\ &= \langle v, v \rangle - 2k\langle v, w \rangle + k^2\langle w, w \rangle. \end{aligned}$$

Wegen $w \neq 0$ ist $\langle w, w \rangle \neq 0$. Also können wir $k := \frac{\langle v, w \rangle}{\langle w, w \rangle}$ setzen. Damit erhalten wir

$$0 \leq \langle v, v \rangle - 2 \frac{\langle v, w \rangle}{\langle w, w \rangle} \cdot \langle v, w \rangle + \left(\frac{\langle v, w \rangle}{\langle w, w \rangle} \right)^2 \cdot \langle w, w \rangle.$$

Wenn man mit $\langle w, w \rangle$ multipliziert, ergibt sich

$$0 \leq \langle v, v \rangle \langle w, w \rangle - (\langle v, w \rangle)^2,$$

das heißt

$$|\langle v, w \rangle|^2 \leq \langle v, v \rangle \langle w, w \rangle = \|v\|^2 \cdot \|w\|^2.$$

daraus folgt die Ungleichung.

Wann gilt in dieser Ungleichung Gleichheit? Genau dann, wenn $w = 0$ oder $v - kw = 0$ ist, also genau dann, wenn v und w linear abhängig sind. \square

Die obige Ungleichung wird nach Augustin-Louis Cauchy (1789–1857), Hermann Amandus Schwarz (1843–1921) und manchmal auch nach Viktor Jakowlewitsch Bunjakowski (1804–1889) genannt.

Nun können wir beweisen, dass jedes Skalarprodukt zu einer Norm führt.

Euklidische Norm

Sei $\langle \cdot, \cdot \rangle$ ein Skalarprodukt des reellen Vektorraums V . Dann wird durch

$$\|v\| := \sqrt{\langle v, v \rangle} \quad \text{für alle } v \in V.$$

eine Norm von V definiert.

Zum *Beweis* weisen wir die definierenden Eigenschaften einer Norm nach:

$$(N1) \quad \|kv\| = \sqrt{\langle kv, kv \rangle} = \sqrt{kk \langle v, v \rangle} = \sqrt{kk} \sqrt{\langle v, v \rangle} = |k| \cdot \|v\|.$$

Die Bedingung (N2) folgt aus der Ungleichung von Cauchy-Schwarz.

$$\begin{aligned} \|v + w\|^2 &= \langle v + w, v + w \rangle = \langle v, v \rangle + \langle v, w \rangle + \langle w, v \rangle + \langle w, w \rangle \\ &= \|v\|^2 + 2 \cdot \langle v, w \rangle + \|w\|^2 \\ &\leq \|v\|^2 + 2\|v\| \cdot \|w\| + \|w\|^2 \quad (\text{Cauchy-Schwarz}) \\ &= (\|v\| + \|w\|)^2. \end{aligned}$$

(N3) Ist $v = 0$, so ist $\langle v, v \rangle = 0$, also $\|v\| = 0$. Ist umgekehrt $\|v\| = 0$, so ist $\sqrt{\langle v, v \rangle} = 0$, also $\langle v, v \rangle = 0$ und somit $v = 0$, da ein Skalarprodukt positiv definit ist.

Insgesamt haben wir bewiesen, dass durch $\|v\| := \sqrt{\langle v, v \rangle}$ eine Norm definiert wird. \square

Wir nennen eine Menge $\{v_1, \dots, v_m\}$ von Vektoren von V **orthogonal**, wenn $v_i \perp v_j$ gilt für $i \neq j$. Eine orthogonale Menge heißt **orthonormal**, wenn zusätzlich $\|v_i\| = 1$ für alle $i \in \{1, \dots, m\}$ gilt. Eine orthonormale Basis wird eine **Orthonormalbasis** genannt.

Eine Orthonormalbasis unterscheidet sich also von einer gewöhnlichen Orthonormalbasis durch dreierlei: Zunächst muss der Vektorraum V ein Skalarprodukt besitzen. Dann muss jeder Vektor der Basis die Norm 1 haben, und schließlich müssen je zwei verschiedene Vektoren der Basis senkrecht aufeinander stehen.

Ein *Beispiel* einer Orthonormalbasis fällt jedem sofort ein: Wenn der Vektorraum $V = \mathbf{R}^n$ mit dem Standardskalarprodukt versehen ist, so ist die Basis $\{e_1, e_2, \dots\}$ aus den Einheitsvektoren eine Orthonormalbasis.

Warum sind Orthonormalbasen wichtig? Dazu überlegen wir uns, wie die Strukturmatrix von $\langle \cdot, \cdot \rangle$ bezüglich einer Orthonormalbasis $\{v_1, v_2, \dots, v_n\}$ aussieht. Wegen $\langle v_i, v_i \rangle = 1$ stehen auf der Diagonale nur Einsen, und wegen $\langle v_i, v_j \rangle = 0$ für $i \neq j$ stehen außerhalb der Hauptdiagonalen nur Nullen. Mit anderen Worten: Bezüglich dieser Basis ist $\langle \cdot, \cdot \rangle$ das Standardskalarprodukt!

Das bedeutet: Wenn eine Orthonormalbasis zu einem Skalarprodukt existiert, können wir dieses Skalarprodukt o.B.d.A. als Standardskalarprodukt voraussetzen.

Unsere erste Behauptung zeigt, dass in gewissem Sinne die Orthogonalität die stärkste Eigenschaft ist:

Aus orthogonal mach orthonormal

Sei $\{v_1, \dots, v_m\}$ eine Menge von Vektoren von $V \setminus \{o\}$. Wenn $\{v_1, \dots, v_m\}$ orthogonal ist, so gilt:

- (a) Die Menge $\{w_1 := \frac{v_1}{\|v_1\|}, \dots, w_m := \frac{v_m}{\|v_m\|}\}$ ist orthonormal.
- (b) Die Menge $\{v_1, \dots, v_m\}$ ist linear unabhängig.

Beweis

- (a) Für $w_i = k_i v_i$ mit $k_i \neq 0$ gilt stets $\langle w_i, w_j \rangle = k_i \cdot k_j \cdot \langle v_i, v_j \rangle = k_i k_j \cdot 0 = 0$. Also ist auch die Menge $\{w_1, w_2, \dots, w_m\}$ orthogonal. Da nach Definition $w_i = \frac{v_i}{\|v_i\|}$ ist, und dieser Vektor

die Norm $\left\| \frac{v_i}{\|v_i\|} \right\| = \frac{\|v_i\|}{\|v_i\|} = 1$ hat, ist die Menge $\{w_1, w_2, \dots, w_m\}$ orthonormal.

- (b) Seien $k_1, \dots, k_m \in \mathbf{R}$ mit $k_1 v_1 + \dots + k_m v_m = o$. Wir bilden das Skalarprodukt mit v_i :

$$0 = \langle v_i, k_1 v_1 + \dots + k_m v_m \rangle = k_i \langle v_i, v_i \rangle.$$

Aus $v_i \neq o$ folgt $\langle v_i, v_i \rangle \neq 0$; also muss $k_i = 0$ sein. \square

Bemerkung In Teil (b) haben Sie zum ersten Mal einen Trick gesehen, den es sich zu merken lohnt: *Wenn möglich, bilde das Skalarprodukt mit v_i und prüfe das Ergebnis!* Wir werden diesen Trick noch mehrfach anwenden.

Nun kommen wir zum berühmten Orthonormalisierungssatz von Erhard Schmidt (1876–1959). Vor diesem Satz sollten Sie aber keine zu große Ehrfurcht entwickeln. Der Beweis sieht in vielen Darstellungen sehr kompliziert aus; aber jeder, der den Mut zu einfachen Konzepten hat, kriegt diesen Beweis hin – also auch Sie!

Orthonormalisierungssatz von E. Schmidt

Sei V ein reeller Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$. Dann hat V eine Orthonormalbasis.

Beweis Sei $\dim(V) = n$. Wir zeigen etwas mehr: *Für jedes $m \in \{1, 2, \dots, n\}$ kann jede orthonormale Menge B_{m-1} aus $m-1$ Elementen zu einer orthonormalen Menge B_m mit m Elementen erweitert werden.* Dann ist B_n eine Orthonormalbasis von V .

Für $m = 1$ ist diese Behauptung nicht schwer zu zeigen: Da B_0 die leere Menge ist, müssen wir nur zeigen, dass es eine orthonormale Menge mit einem Element gibt. Man nehme irgendeinen von Null verschiedenen Vektor v_1 und bilde $w_1 := \frac{v_1}{\|v_1\|}$. Dann ist $\{w_1\}$ eine orthonormale Menge.

Wir machen uns den Induktionsschluss zunächst dadurch klar, dass wir zu der orthonormalen Menge $\{w_1\}$ einen Vektor w_2 konstruieren, so dass auch $\{w_1, w_2\}$ eine orthonormale Menge ist. Dazu wählen wir einfach irgendeinen Vektor v_2 , der linear unabhängig von w_1 ist.

Wenn v_2 zufällig senkrecht zu w_1 ist, müssen wir v_2 nur noch durch seine Norm teilen und haben eine orthonormale Menge! Dies ist jedoch ein ganz seltener (und unverdienter) Glücksfall.

Also müssen wir uns mit dem Normalfall herumschlagen, dass v_2 und w_1 nicht senkrecht sind, dass also $\langle v_2, w_1 \rangle \neq 0$ gilt.

Aber ... könnte man dann nicht aus v_2 und w_1 einen Vektor basteln, der senkrecht auf w_1 steht? Wir fragen dazu genauer, wann der Vektor $x := k \cdot w_1 + v_2$ senkrecht auf w_1 steht. Die Frage lautet also noch genauer: Für welche k gilt

$$\langle x, w_1 \rangle = 0, \quad \text{das heißt} \quad \langle k w_1 + v_2, w_1 \rangle = 0 ?$$

Wir verfolgen diesen Ansatz einfach weiter:

$$\begin{aligned} 0 &= \langle k w_1 + v_2, w_1 \rangle \\ \Leftrightarrow 0 &= k \cdot \langle w_1, w_1 \rangle + \langle v_2, w_1 \rangle = k + \langle v_2, w_1 \rangle . \end{aligned}$$

Daraus ergibt sich also

$$k = -\langle v_2, w_1 \rangle .$$

Mit diesem Wert ist $x = kw_1 + v_2$ ein Vektor, der orthogonal zu w_1 ist. Wenn wir jetzt noch

$$w_2 := \frac{x}{\|x\|}$$

setzen, erhalten wir eine orthonormale Menge $\{w_1, w_2\}$.

Der Trick besteht also darin, einen Vektor zu wählen, der linear unabhängig zu den bisher konstruierten ist und aus diesem zusammen mit den bisher konstruierten einen Vektor zu bilden, der orthogonal zu allen bisherigen ist. Das Schöne hierbei ist, dass man nur einen Ansatz machen muss und dann automatisch einen orthogonalen Vektor herausbekommt. So einfach ist das!

Wir diskutieren hier noch den nächsten Fall: Sei $\{w_1, w_2\}$ eine orthonormale Menge, die noch nicht den ganzen Vektorraum V erzeugt. Dann gibt es einen Vektor v_3 , so dass die Menge $\{w_1, w_2, v_3\}$ linear unabhängig ist.

Wir setzen

$$x := k_1 w_1 + k_2 w_2 + v_3$$

und wollen k_1 und k_2 so bestimmen, dass x orthogonal zu w_1 und w_2 ist. Das bedeutet

$$0 = \langle w_1, k_1 w_1 + k_2 w_2 + v_3 \rangle = k_1 \langle w_1, w_1 \rangle + \langle w_1, v_3 \rangle = k_1 + \langle w_1, v_3 \rangle.$$

Also ist $k_1 = -\langle w_1, v_3 \rangle$. Entsprechend folgt $k_2 = -\langle w_2, v_3 \rangle$.

Mit dieser Wahl von k_1 und k_2 haben wir einen Vektor x gefunden, der orthogonal zu w_1 und w_2 ist. Indem wir durch die Norm von x dividieren, erhalten wir eine orthonormale Menge

$$\left\{ w_1, w_2, w_3 := \frac{x}{\|x\|} \right\}.$$

Nun ist Ihnen sicherlich klar, wie man den allgemeinen Fall löst; Sie sind eingeladen, dies in Übungsaufgabe 13 explizit zu tun. \square

Als Korollar aus dem Beweis des Orthonormalisierungssatzes ergibt sich:

Korollar

Jede orthonormale Menge von Vektoren eines euklidischen Vektorraums, insbesondere jeder Vektor v mit $\|v\| = 1$ kann zu einer Orthonormalbasis ergänzt werden. \square

10.4 Orthogonale Abbildungen

Nun studieren wir lineare Abbildungen, die ein Skalarprodukt invariant (das heißt unverändert) lassen. Dieses Vorgehen (Untersuchung derjenigen Abbildungen, die eine zuvor

definierte Struktur invariant lassen) hat sich in der Mathematik als äußerst nützliches Hilfsmittel zur Untersuchung von Strukturen herausgestellt. Implizit ist uns das schon häufig begegnet: Automorphismen von Körpern, lineare Abbildungen von Vektorräumen sind solche strukturerhaltenden Abbildungen.

Sei nun V ein euklidischer Vektorraum, also ein reeller Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$. In dieser Situation sind die linearen Abbildungen von V in sich im Allgemeinen bestimmt nicht die richtigen Abbildungen; denn diese haben mit dem Skalarprodukt nichts zu tun, können also auch keine Information über das Skalarprodukt liefern.

Eine lineare Abbildung f des euklidischen Vektorraums V in sich wird **orthogonal** genannt, wenn für alle $v, w \in V$ gilt

$$\langle f(v), f(w) \rangle = \langle v, w \rangle.$$

Eine lineare Abbildung ist also orthogonal, wenn das Skalarprodukt je zweier Vektoren gleich dem Skalarprodukt der Bilder dieser Vektoren ist. Wenn Sie eine ganz anschauliche Interpretation gern haben: Eine lineare Abbildung ist genau dann orthogonal, wenn sie die Größe des Winkels zwischen Vektoren erhält.

Wenn diese anschauliche Interpretation richtig ist, müsste *zum Beispiel* eine Drehung eine orthogonale Abbildung sein. Das können wir nachprüfen.

Sei $V = \mathbb{R}^2$ (dessen Elemente wir als Spaltenvektoren auffassen), und sei $\langle \cdot, \cdot \rangle$ das Standardskalarprodukt. Eine Drehung um den Winkel φ wird durch die folgende Matrix $D(\varphi)$ beschrieben (vergleichen Sie dazu das Projekt aus Kap. 8).

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

Ein beliebiger Vektor $(x, y)^T$ wird durch die Drehung um den Winkel φ also auf folgenden Vektor abgebildet:

$$\begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \cdot \cos(\varphi) - y \cdot \sin(\varphi) \\ x \cdot \sin(\varphi) + y \cdot \cos(\varphi) \end{pmatrix}.$$

Nun berechnen wir einfach das Skalarprodukt der Bilder der Vektoren $(x_1, y_1)^T$ und $(x_2, y_2)^T$. (... und wenn die Drehung wirklich eine orthogonale Abbildung ist, dann dürfte sich hierbei nichts anderes als das Skalarprodukt $x_1 x_2 + y_1 y_2$ der Urbilder ergeben! Mal sehen:)

$$\begin{aligned} & \left\langle \begin{pmatrix} x_1 \cdot \cos(\varphi) - y_1 \cdot \sin(\varphi) \\ x_1 \cdot \sin(\varphi) + y_1 \cdot \cos(\varphi) \end{pmatrix}, \begin{pmatrix} x_2 \cdot \cos(\varphi) - y_2 \cdot \sin(\varphi) \\ x_2 \cdot \sin(\varphi) + y_2 \cdot \cos(\varphi) \end{pmatrix} \right\rangle \\ &= (x_1 \cdot \cos(\varphi) - y_1 \cdot \sin(\varphi)) (x_2 \cdot \cos(\varphi) - y_2 \cdot \sin(\varphi)) \\ &\quad + (x_1 \cdot \sin(\varphi) + y_1 \cdot \cos(\varphi)) (x_2 \cdot \sin(\varphi) + y_2 \cdot \cos(\varphi)) \\ &= (x_1 x_2 + y_1 y_2) (\cos^2(\varphi) + \sin^2(\varphi)) = x_1 x_2 + y_1 y_2. \end{aligned}$$

Also ist eine Drehung tatsächlich eine orthogonale Abbildung!

Die Frage, die wir uns in diesem Abschnitt vor allem stellen, ist die folgende: Welches sind die orthogonalen Abbildungen? Ist es möglich, alle orthogonalen Abbildungen vernünftig zu beschreiben? Wir werden diese Frage beantworten, indem wir die Darstellungsmatrizen von orthogonalen Abbildungen betrachten. Zuerst untersuchen wir die Eigenwerte einer orthogonalen Abbildung. Da V ein reeller Vektorraum ist, meinen wir mit „Eigenwerten“ stillschweigend „reelle Eigenwerte“ (dass eine lineare Abbildung zusätzlich noch komplexe Eigenwerte haben könnte, versteht sich von selbst).

Eigenwerte einer orthogonalen Abbildung

Sei f eine orthogonale Abbildung des euklidischen Vektorraums V in sich. Dann hat f höchstens die Eigenwerte 1 und -1 . Insbesondere ist f eine umkehrbare lineare Abbildung, und f^{-1} ist ebenfalls eine orthogonale Abbildung.

Beweis Sei k ein Eigenwert von f . Dann gibt es einen Vektor $v \neq 0$ mit

$$f(v) = kv.$$

Also ist

$$\|v\| = \|f(v)\| = \|kv\| = |k| \cdot \|v\|,$$

also ist $|k| = 1$, das heißt $k = \pm 1$. Dies ist die erste Behauptung.

Daraus ergibt sich unmittelbar, dass 0 kein Eigenwert von f ist. Also ist $\text{Kern}(f) = \{0\}$, und somit ist f injektiv. Als lineare Abbildung des endlichdimensionalen Vektorraums V in sich ist f dann auch bijektiv.

Schließlich zeigen wir noch, dass auch f^{-1} orthogonal ist. Dazu betrachten wir zwei beliebige Vektoren $v, w \in V$. Da f orthogonal ist, ergibt sich

$$\langle f^{-1}(v), f^{-1}(w) \rangle = \langle f(f^{-1}(v)), f(f^{-1}(w)) \rangle = \langle v, w \rangle.$$

□

Es wird sich herausstellen, dass das Matrizenäquivalent zu orthogonalen Abbildungen die orthogonalen Matrizen sind:

Eine reelle invertierbare $n \times n$ -Matrix M (also eine Matrix aus $\text{GL}(n, \mathbf{R})$) heißt **orthogonal**, falls man die zu M inverse Matrix dadurch erhält, dass man M transponiert, falls also gilt

$$M^{-1} = M^T.$$

Beispiel. Im Fall $n = 2$ ist die Matrix

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

eine orthogonale Matrix. Es gilt nämlich

$$\begin{aligned} D \cdot D^T &= \begin{pmatrix} \cos(\varphi) & \sin(\varphi) \\ -\sin(\varphi) & \cos(\varphi) \end{pmatrix} \cdot \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix} \\ &= \begin{pmatrix} \cos^2(\varphi) + \sin^2(\varphi) & -\cos(\varphi)\sin(\varphi) + \sin(\varphi)\cos(\varphi) \\ -\sin(\varphi)\cos(\varphi) + \cos(\varphi)\sin(\varphi) & \sin^2(\varphi) + \cos^2(\varphi) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

Ganz entsprechend zeigt man, dass auch $D^T \cdot D$ die Einheitsmatrix ist. Also ist die zu D transponierte Matrix D^T tatsächlich invers zu D .

Es wird sich zeigen, dass die Drehmatrizen „im wesentlichen“ die einzigen orthogonalen Matrizen sind.

Die orthogonalen 2×2 -Matrizen kann man leicht bestimmen; deshalb tun wir das zuerst.

Orthogonale 2×2 -Matrizen

Sei M eine orthogonale 2×2 -Matrix. Dann gibt es eine reelle Zahl φ so dass M die folgende Gestalt hat:

$$M = \begin{pmatrix} \cos(\varphi) & \sin(\varphi) \\ -\sin(\varphi) & \cos(\varphi) \end{pmatrix} \quad \text{oder} \quad M = \begin{pmatrix} \cos(\varphi) & \sin(\varphi) \\ \sin(\varphi) & -\cos(\varphi) \end{pmatrix}.$$

Beweis Wir setzen ganz allgemein an. Sei

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

mit $a, b, c, d \in \mathbf{R}$. Nach Definition einer orthogonalen Matrix gilt $M \cdot M^T = E$. Das bedeutet:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = M \cdot M^T = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} a^2 + b^2 & ac + bd \\ ac + bd & c^2 + d^2 \end{pmatrix},$$

also

$$a^2 + b^2 = 1, \quad c^2 + d^2 = 1 \quad \text{und} \quad ac + bd = 0.$$

Wegen der ersten Gleichung gibt es eine reelle Zahl φ (die wir sogar in dem Intervall $[0, 2\pi)$ wählen können) mit $a = \cos(\varphi)$ und $b = \sin(\varphi)$. Ebenso folgt aus der zweiten Gleichung die Existenz einer reellen Zahl $\psi \in [0, 2\pi)$ mit der Eigenschaft $c = \sin(\psi)$ und $d = \cos(\psi)$. Die dritte Gleichung klärt die Beziehung zwischen φ und ψ ; es gilt nämlich

$$0 = ac + bd = \cos(\varphi)\sin(\psi) + \sin(\varphi)\cos(\psi),$$

und der berühmte Additionssatz für trigonometrische Funktionen sagt, dass die rechte Seite der obigen Gleichung genau gleich

$$\sin(\varphi + \psi)$$

ist. Also ist $\sin(\varphi + \psi) = 0$, und auch eine ganz oberflächliche Kenntnis der trigonometrischen Funktionen zeigt, dass dann $\varphi + \psi$ ein ganzzahliges Vielfaches von π sein muss.

Wegen $\varphi, \psi \in [0, 2\pi)$ ist also $\varphi + \psi$ gleich 0, gleich π , gleich 2π oder gleich 3π . Wenn $\varphi + \psi$ gleich 0 oder gleich 2π ist, so ist $\psi = -\varphi$ oder $\psi = 2\pi - \varphi$, also

$$c = \sin(\psi) = \sin(-\varphi) = -\sin(\varphi) \quad \text{und} \quad d = \cos(\psi) = \cos(-\varphi) = \cos(\varphi).$$

Im anderen Fall ist $\varphi + \psi$ gleich π oder gleich 3π ; also folgt

$$c = \sin(\psi) = \sin(\varphi) \quad \text{und} \quad d = \cos(\psi) = -\cos(\varphi).$$

Damit ist die Behauptung gezeigt. □

Bemerkung Wir werden demnächst („Darstellung orthogonaler Abbildungen eines zweidimensionalen Vektorraums“) den zweiten Fall noch genau behandeln. Eines sei aber jetzt schon verraten: Dieser Fall verspricht mehr als er hält.

Determinante einer orthogonalen Matrix

Jede orthogonale Matrix hat die Determinante 1 oder -1 .

Beweis Sei M eine orthogonale Matrix. Dann folgt mit dem Multiplikationssatz für Determinanten

$$1 = \det(E_n) = \det(MM^{-1}) = \det(MM^T) = \det(M) \det(M^T) = \det(M) \det(M);$$

also ist $\det(M)$ eine (reelle) Wurzel aus 1, also gleich ± 1 . □

Üblicherweise zeichnet man die orthogonalen Matrizen mit Determinante $+1$ speziell aus: Man nennt die Menge aller orthogonalen $n \times n$ -Matrizen (zusammen mit der Multiplikation von Matrizen) die **orthogonale Gruppe** und bezeichnet sie mit $O(n)$. Die Menge aller orthogonalen $n \times n$ -Matrizen mit Determinante $+1$ heißt die **spezielle orthogonale Gruppe** und wird mit $SO(n)$ bezeichnet.

In Übungsaufgabe 17 sollen Sie nachweisen, dass $O(n)$ und $SO(n)$ tatsächlich Gruppen sind.

Der folgende Satz kommt für Sie aus „bezeichnungspsychologischen“ Gründen für Sie sicher nicht unerwartet.

Darstellung orthogonaler Abbildungen

Sei V ein euklidischer Vektorraum, und sei B eine Orthonormalbasis von V . Sei f eine lineare Abbildung von V in sich. Dann gilt:

f ist orthogonal \Leftrightarrow die Darstellungsmatrix ${}_B M_B(f)$ ist eine orthogonale Matrix.

Beweis Sei $B = \{v_1, v_2, \dots, v_n\}$.

Die Abbildung f ist definitionsgemäß genau dann orthogonal, wenn $\langle v, w \rangle = \langle f(v), f(w) \rangle$ gilt für alle $v, w \in V$. Man überlegt sich leicht (siehe Übungsaufgabe 18), dass f genau dann orthogonal ist, wenn $\langle v_i, v_j \rangle = \langle f(v_i), f(v_j) \rangle$ für alle i, j gilt.

Daraus ergibt sich die Beweisstrategie: Wir rechnen $\langle v_i, v_j \rangle$ und $\langle f(v_i), f(v_j) \rangle$ aus und überlegen uns dann, was die Gleichheit dieser beiden Ausdrücke bedeutet.

Da B eine Orthonormalbasis ist, ist $\langle v_i, v_j \rangle = \delta_{ij}$, das heißt $\langle v_i, v_j \rangle = 1$, falls $i = j$ ist, und $\langle v_i, v_j \rangle = 0$ sonst.

Nun zur Berechnung von $\langle f(v_i), f(v_j) \rangle$. Zur Abkürzung setzen wir $A := {}_B M_B(f)$ mit $A = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq n}$. Dann ist nach Definition einer Darstellungsmatrix

$$f(v_i) = a_{1i}v_1 + a_{2i}v_2 + \dots + a_{ni}v_n \quad \text{und} \quad f(v_j) = a_{1j}v_1 + a_{2j}v_2 + \dots + a_{nj}v_n.$$

Also folgt

$$\langle f(v_i), f(v_j) \rangle = a_{1i}a_{1j} + a_{2i}a_{2j} + \dots + a_{ni}a_{nj}.$$

Dies ist ein Körperelement. Ich behaupte, dass dieses Körperelement auch durch folgenden Ausdruck dargestellt werden kann (der zugegebenermaßen nicht vertrauenerweckender aussieht!)

$$e_i \cdot A^T \cdot A \cdot e_j^T,$$

wobei e_i der i -te Einheitsvektor (als Zeilenvektor geschrieben) ist, also derjenige Vektor, der an der i -ten Stelle eine Eins hat und sonst aus Nullen besteht.

Dies zeigen wir dadurch, dass wir diesen letzten Ausdruck einfach ausrechnen. Es ergibt sich

$$\begin{aligned} e_i \cdot A^T \cdot A \cdot e_j^T &= e_i \cdot \begin{pmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \cdot e_j^T \\ &= (a_{1i} \quad a_{2i} \quad \dots \quad a_{ni}) \cdot \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix} = a_{1i} \cdot a_{1j} + a_{2i} \cdot a_{2j} + \dots + a_{ni} \cdot a_{nj}. \end{aligned}$$

Und siehe da! Die beiden Ausdrücke sind gleich. Wir fassen dieses Ergebnis zusammen:

$$\langle f(v_i), f(v_j) \rangle = e_i \cdot A^T \cdot A \cdot e_j^T.$$

Nun setzen wir die beiden berechneten Ausdrücke gleich. Genau dann ist f orthogonal, wenn

$$e_i \cdot A^T \cdot A \cdot e_j^T = \langle f(v_i), f(v_j) \rangle = \langle v_i, v_j \rangle = \delta_{ij}$$

ist.

Aus dieser Gleichung folgt aber die für uns entscheidende Tatsache, dass f genau dann orthogonal ist, wenn $A^T A$ die Einheitsmatrix ist.

Um das einzusehen, lesen Sie die obige Gleichung als

$$e_i \cdot N \cdot e_j^T = \delta_{ij},$$

wobei $N = A^T \cdot A$ ist. Dann ist $e_i \cdot N \cdot e_j^T$ das Element in der i -ten Zeile und j -ten Spalte von N . (Denn $e_i \cdot N$ ist die i -te Spalte von N , und durch die Multiplikation mit e_j^T wird aus dieser die j -te Komponente ausgewählt.) Also sagt die obige Gleichung: Das Element in der i -ten Zeile und j -ten Spalte von N ist gleich δ_{ij} ; das bedeutet aber nichts anderes, als dass $A^T \cdot A$ ($= N$) die Einheitsmatrix ist.

Genau dann ist aber $A^T \cdot A$ die Einheitsmatrix, wenn A^T die Inverse zu A ist. Also: Genau dann ist f orthogonal, wenn $A^{-1} = A^T$ gilt, d. h., wenn A eine orthogonale Matrix ist. \square

Nun können wir alle orthogonalen Abbildungen genau beschreiben, und zwar mit Hilfe einer Darstellungsmatrix. Es wird sich zeigen, dass die Drehmatrizen die wesentlichen Bausteine sind. Wir behandeln zunächst den zweidimensionalen Fall.

Darstellung orthogonaler Abbildungen eines zweidimensionalen Vektorraums

Sei f eine orthogonale lineare Abbildung des zweidimensionalen euklidischen Vektorraums V . Dann gibt es eine Orthonormalbasis B von V , so dass die Darstellungsmatrix ${}_B M_B(f)$ von f bezüglich B folgende Gestalt hat:

$${}_B M_B(f) = \begin{pmatrix} \cos(\varphi) & \sin(\varphi) \\ -\sin(\varphi) & \cos(\varphi) \end{pmatrix} \quad \text{oder} \quad {}_B M_B(f) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Im ersten Fall heißt f eine **Drehung**, im zweiten Fall nennt man f eine **Spiegelung**.

Beweis Sei zunächst B eine beliebige Orthonormalbasis von V . Dann ist $M = {}_B M_B(f)$ eine orthogonale Matrix. Nach dem Satz über orthogonale 2×2 -Matrizen hat M also die fol-

Der Beweis erfolgt durch Induktion nach $n := \dim(V)$.

Für $n = 1$ ist der Satz trivial, und für $n = 2$ haben wir ihn oben bewiesen. Sei nun $n > 3$, und sei die Aussage des Satzes richtig für $n - 1$ und $n - 2$. (Das ist kein Druckfehler! Wir hätten auch schreiben können „für alle $n' < n$ “, aber wir brauchen die Aussage nur für $n - 1$ und $n - 2$.)

Um Induktion anwenden zu können, müssen wir zuerst die folgende entscheidende Aussage beweisen: *Es gibt einen Unterraum W der Dimension 1 oder 2, der durch f auf sich selbst abgebildet wird, das heißt, für den $f(W) = W$ gilt.*

Das folgt so Als erstes müssen wir dazu wissen, dass sich jedes Polynom $g \in \mathbf{R}[x]$ als Produkt von reellen Polynomen schreiben lässt, die höchstens den Grad zwei haben. (Den Beweis hierfür deuten wir nur an. Wir können annehmen, dass g normiert ist. Wir fassen g als Polynom aus $\mathbf{C}[x]$ auf. Der Hauptsatz der Algebra sagt, dass das Polynom dann in Linearfaktoren der Form $x - c$ mit $c \in \mathbf{C}$ zerfällt. Wenn $c \in \mathbf{R}$ ist, so ist $x - c$ ein reelles Polynom vom Grad 1, das g teilt. Sei nun $c = a + bi \in \mathbf{C} \setminus \mathbf{R}$. Da g aber nur reelle Koeffizienten hat, ist mit $x - c$ auch $x - \bar{c}$ ein Linearfaktor von g . Daher ist $(x - c)(x - \bar{c}) = x^2 - 2ax + a^2 + b^2 \in \mathbf{R}[x]$. Insgesamt folgt, dass g sich als Produkt reeller Polynome vom Grad 1 und 2 schreiben lässt.)

Damit ist der Beweis der Zwischenbehauptung aber noch nicht beendet, im Gegenteil: Jetzt geht's erst richtig los: Wir betrachten jetzt das charakteristische Polynom χ_f von f . Wir haben uns gerade überlegt, dass χ_f sich schreiben lässt als Produkt von reellen Polynomen g_i vom Grad 1 oder 2:

$$\chi_f = \pm g_1 \cdot g_2 \cdot \dots \cdot g_k.$$

Wenn zufällig eines der Polynome g_i den Grad 1 hat, also $g = x - a$ mit $a \in \mathbf{R}$, so hat f den Eigenwert a , und jeder Eigenvektor v zu a erzeugt einen 1-dimensionalen Unterraum $\langle v \rangle$, der durch f auf sich abgebildet wird.

Daher müssen wir noch den Fall betrachten, dass alle g_i den Grad 2 haben. Wir zeigen nun, dass es einen Vektor $v \neq o$ gibt, so dass $g_i(f)(v) = o$ für mindestens ein i gilt. Sei dazu $w \neq o$ ein beliebiger Vektor aus V . Wenn zufällig schon $g_1(f)(w) = 0$ ist, sind wir fertig. Wenn nicht, betrachten wir $g_2(f)g_1(f)(w)$. Wenn dieser Vektor gleich o ist, so setzen wir $v := g_1(f)(w)$; es folgt

$$g_2(f)(v) = g_2(f)g_1(f)(w) = o.$$

Wenn auch der Vektor $g_2(f)g_1(f)(w)$ ungleich o ist, so betrachten wir den Vektor $g_3(f)g_2(f)g_1(f)(w)$, usw. Einer dieser Vektoren muss der Nullvektor sein, denn nach dem Satz von Cayley-Hamilton gilt

$$\pm g_1(f) \cdot g_2(f) \cdot \dots \cdot g_k(f) = \pm g_1 \cdot g_2 \cdot \dots \cdot g_k(f) = \chi_f(f) = 0.$$

Sei also $v \neq o$ ein Vektor aus V mit $g_i(f)(v) = o$. Behauptung: Der Unterraum $W := \langle v, f(v) \rangle$ wird durch f in sich abgebildet. Dazu müssen wir nur zeigen, dass $f(f(v))$ wieder in W liegt, dass also $f(f(v))$ eine Linearkombination von v und $f(v)$ ist. Sei dazu

$g_i = x^2 + ax + b$ mit $a, b \in \mathbf{R}$. Dann ist

$$\begin{aligned} 0 &= g_i(f)(v) = (x^2 + ax + b)(f)(v) = f(f(v)) + af(v) + bv, \quad \text{also} \\ f(f(v)) &= -bv - af(v) \end{aligned}$$

– und damit ist die Zwischenbehauptung endgültig und unwiderruflich bewiesen!

Und was nützt uns das Ganze? Gemach, gemacht! Wir atmen kurz durch und machen uns auf zu neuen Taten: Wir betrachten jetzt nicht mehr den (kleinen) Unterraum W , sondern den (großen) Unterraum W^\perp . Wir behaupten:

Es gilt auch $f(W^\perp) = W^\perp$. Um dies zu zeigen, betrachten wir einen beliebigen Vektor x aus W^\perp . Es ist zu zeigen, dass dann auch $f(x)$ auf jedem Vektor $w \in W$ senkrecht steht. Dies rechnen wir einfach aus:

$$\langle f(x), w \rangle = \langle f^{-1}f(x), f^{-1}(w) \rangle = \langle x, f^{-1}(w) \rangle = 0.$$

(Das erste Gleichheitszeichen ergibt sich, weil mit f auch f^{-1} eine orthogonale Abbildung ist; das letzte Gleichheitszeichen folgt aus der Tatsache, dass auch f^{-1} den Unterraum W auf sich abbildet, und somit der Vektor $f^{-1}(w)$ in W liegt.)

Die Einschränkung von f auf W^\perp nennen wir f^* . Dann ist f^* eine orthogonale Abbildung des euklidischen Vektorraums W^\perp in sich. Da W die Dimension 1 oder 2 hat, hat W^\perp die Dimension $n-1$ oder $n-2$. Also können wir (endlich!) Induktion anwenden: Es gibt eine Orthonormalbasis C von W^\perp , so dass $B := C \cup \{w\}$ eine Orthonormalbasis ist, bezüglich der – eventuell nach Umnummerierung – die Matrix ${}_B M_B(f)$ die im Satz angegebene Gestalt hat.

Wir müssen jetzt noch C zu einer Orthonormalbasis von ganz V erweitern. Wenn W die Dimension 1 hat, so wählen wir einen Vektor w der Norm 1 aus W . Dann ist $C \cup \{w\}$ – eventuell nach Umnummerierung – eine Orthonormalbasis, bezüglich der f die im Satz angegebene Gestalt hat.

Sei schließlich $\dim(W) = 2$. Diesen Fall haben wir schon behandelt: Es gibt eine Orthonormalbasis C' von W , so dass die Einschränkung von f auf W eine Darstellungsmatrix hat wie sie im vorangegangenen Satz beschrieben ist. Dann ist $B := C \cup C'$ eine Orthonormalbasis der gesuchten Form. \square

Damit hat dieser Beweis (einer der längsten dieses Buches) doch noch ein Ende gefunden. Kompliment, wenn Sie durchgehalten haben. Der Beweis ist nicht einfach, er verwendet aber Methoden, die Sie sich gründlich ansehen sollten. Aber freuen Sie sich nicht zu früh: Es wartet noch ein Satz mit einem langen Beweis auf Sie!

10.5 ... und eine zweite symmetrische Bilinearform?

Sei V ein euklidischer Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$. Vermutlich wird es noch viele andere Skalarprodukte und noch viel mehr Bilinearformen auf V geben. In diesem Ab-

schnitt untersuchen wir, ob und wie wir mit Hilfe des „festen Skalarprodukts“ $\langle \cdot, \cdot \rangle$ eine andere symmetrische Bilinearform von V beschreiben können. Wir werden schnell sehen, dass dies gleichwertig ist mit der Untersuchung der so genannten „selbstadjungierten“ linearen Abbildungen von V in sich und mit der Untersuchung symmetrischer Matrizen mit reellwertigen Einträgen. Ein Hauptergebnis wird sein, dass man jede solche Matrix diagonalisieren kann.

Sei also $[\cdot, \cdot]$ eine symmetrische Bilinearform auf V . Sei φ die lineare Abbildung von V in sich, die durch folgende Vorschrift definiert ist:

$$[v, w] = \langle \varphi(v), w \rangle$$

für alle $v, w \in V$ (vergleichen Sie dazu Abschn. 10.2).

Beschreibung symmetrischer Bilinearformen durch selbstadjungierte Abbildungen

Die lineare Abbildung φ genügt folgender Gleichung

$$\langle \varphi(v), w \rangle = \langle v, \varphi(w) \rangle$$

für alle $v, w \in V$. Man sagt dazu auch, dass φ **selbstadjungiert** ist.

Umgekehrt gilt: Sei ψ eine selbstadjungierte lineare Abbildung von V in sich. Dann wird durch

$$[v, w] := \langle \psi(v), w \rangle$$

eine symmetrische Bilinearform definiert.

Beweis Seien $v, w \in V$. Da sowohl $[\cdot, \cdot]$ als auch $\langle \cdot, \cdot \rangle$ symmetrisch sind, gilt

$$\langle \varphi(v), w \rangle = [v, w] = [w, v] = \langle \varphi(w), v \rangle = \langle v, \varphi(w) \rangle.$$

Der Nachweis der Umkehrung ist einfach und wird Ihnen als (leichte) Übung empfohlen (siehe Übungsaufgabe 20). \square

Übrigens Machen Sie sich über den Namen „selbstadjungiert“ keine unnötigen Gedanken. Im Allgemeinen kann man die zu einer linearen Abbildung „adjungierte“ lineare Abbildung definieren; dann heißt eine lineare Abbildung „selbstadjungiert“ wenn sie gleich ihrer adjungierten Abbildung ist.

Als nächstes überlegen wir uns, welche Darstellungsmatrix eine selbstadjungierte lineare Abbildung hat.

Darstellungsmatrix einer selbstadjungierten Abbildung

Sei φ eine selbstadjungierte lineare Abbildung des euklidischen Vektorraums V in sich. Dann ist die Darstellungsmatrix ${}_B M_B(\varphi)$ bezüglich jeder Orthonormalbasis B von V eine symmetrische Matrix.

Umgekehrt: Wenn φ bezüglich einer Orthonormalbasis eine symmetrische Darstellungsmatrix hat, dann ist φ selbstadjungiert.

Beweis Den Trick, den wir hier brauchen, kennen wir schon.

Sei $B = \{v_1, v_2, \dots, v_n\}$. Dann gilt (siehe Übungsaufgabe 19): Genau dann ist φ selbstadjungiert, wenn für alle $i, j \in \{1, 2, \dots, n\}$ die folgende Gleichung gilt:

$$\langle \varphi(v_i), v_j \rangle = \langle v_i, \varphi(v_j) \rangle .$$

Wir brauchen die definierende Eigenschaft also nur für die Elemente von B nachzuweisen. Daraus ergibt sich wieder die Beweisstrategie: Berechne $\langle \varphi(v_i), v_j \rangle$ und $\langle v_i, \varphi(v_j) \rangle$ und überlege, wann Gleichheit gilt.

Zur Abkürzung setzen wir $A := {}_B M_B(\varphi)$ mit $A = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq n}$. Weil A eine Darstellungsmatrix ist, gilt $\varphi(v_i) = a_{1i}v_1 + a_{2i}v_2 + \dots + a_{ni}v_n$. Also folgt

$$\begin{aligned} \langle \varphi(v_i), v_j \rangle &= \langle a_{1i}v_1 + a_{2i}v_2 + \dots + a_{ni}v_n, v_j \rangle \\ &= \langle a_{1i}v_1, v_j \rangle + \langle a_{2i}v_2, v_j \rangle + \dots + \langle a_{ji}v_j, v_j \rangle + \dots + \langle a_{ni}v_n, v_j \rangle \\ &= \langle a_{ji}v_j, v_j \rangle = a_{ji} . \end{aligned}$$

Entsprechend ergibt sich aus $\varphi(v_j) = a_{1j}v_1 + a_{2j}v_2 + \dots + a_{nj}v_n$ die Gleichung

$$\langle v_i, \varphi(v_j) \rangle = \langle v_i, a_{1j}v_1 + a_{2j}v_2 + \dots + a_{nj}v_n \rangle = a_{ij} .$$

Insgesamt folgt:

$$\begin{aligned} \varphi &\text{ ist selbstadjungiert} \\ \Leftrightarrow \langle \varphi(v_i), v_j \rangle &= \langle v_i, \varphi(v_j) \rangle \quad \text{für alle } i, j \in \{1, 2, \dots, n\} \\ \Leftrightarrow a_{ji} &= a_{ij} \quad \text{für alle } i, j \in \{1, 2, \dots, n\} \\ \Leftrightarrow A &\text{ ist symmetrisch.} \end{aligned}$$

□

Wir sehen also: Symmetrische Bilinearformen, selbstadjungierte lineare Abbildungen und symmetrische Matrizen beschreiben alle das gleiche Phänomen. Welche Sprache einem persönlich am meisten zusagt, ist zum Teil natürlich eine Geschmacksfrage: Entscheiden Sie sich in freier Wahl für eine dieser drei Sprachen.

Das Hauptergebnis dieses Abschnitts ist der folgende Satz, der sowohl unter dem Namen Spektralsatz als auch unter der traditionellen Bezeichnung Hauptachsentransformation bekannt ist. Wir geben die Aussage dieses Satzes in drei Sprachen wieder.

Hauptachsentransformation

Sei V ein euklidischer Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$.

- (a) Zu jeder selbstadjungierten linearen Abbildung φ von V in sich gibt es eine Orthonormalbasis von V , die aus lauter Eigenvektoren von φ besteht. Insbesondere ist φ diagonalisierbar.
- (b) Zu jeder symmetrischen Bilinearform $[\cdot, \cdot]$ von V gibt es eine Orthonormalbasis, die bezüglich $[\cdot, \cdot]$ eine Orthogonalbasis ist.
- (c) Jede symmetrische reelle Matrix ist diagonalisierbar und hat nur reelle Eigenwerte.

Beweis Der eigentliche Beweis besteht im Nachweis von (a). Wir machen uns zunächst klar, dass (b) und (c) aus (a) folgen.

„(a) \Rightarrow (b)“: Sei $[\cdot, \cdot]$ eine symmetrische Bilinearform von V . Dann wird durch

$$[v, w] = \langle \varphi(v), w \rangle$$

eine selbstadjungierte lineare Abbildung φ von V in sich definiert. Nach (a) gibt es eine Orthonormalbasis $B = \{v_1, v_2, \dots, v_n\}$, die aus lauter Eigenvektoren von φ besteht.

Behauptung: B ist eine Orthogonalbasis von $[\cdot, \cdot]$.

Warum? Da B aus Eigenvektoren von V besteht, gilt:

$$[v_i, v_j] = \langle \varphi(v_i), v_j \rangle = \langle k_i \cdot v_i, v_j \rangle = k_i \cdot \langle v_i, v_j \rangle.$$

Also ist sicher $[v_i, v_j] = 0$ für $i \neq j$.

„(a) \Rightarrow (c)“: Sei A eine symmetrische reelle $n \times n$ -Matrix. Betrachte eine Orthonormalbasis B eines n -dimensionalen euklidischen Vektorraums V . Sei φ die zu A bezüglich der Basis B gehörige lineare Abbildung. Nach dem vorhergehenden Satz ist φ selbstadjungiert. Nach (a) ist φ also diagonalisierbar. Daher zerfällt das charakteristische Polynom von φ über \mathbf{R} vollständig. Da A eine Darstellungsmatrix von φ ist, hat also auch A nur reelle Eigenwerte.

Nun zum *eigentlichen Beweis*. Wir beweisen, wie gesagt, die Aussage (a). Sei φ eine selbstadjungierte Abbildung von V .

Wenn wir wüssten, dass φ einen Eigenvektor v besitzt, wären wir fein raus. Denn dann könnten wir die Aussage relativ einfach mit Hilfe von Induktion nach $n := \dim(V)$ beweisen.

Dazu betrachten wir einen Eigenvektor v von φ (wie gesagt, vorerst ist es ein süßes Geheimnis, ob es solch einen Vektor überhaupt gibt!) und das zu $\langle v \rangle$ orthogonale Komplement $W := \langle v \rangle^\perp$. Zunächst müssen wir uns davon überzeugen, dass φ den Unterraum W in sich abbildet. Sei dazu $w \in W$ beliebig. Wir müssen zeigen, dass dann auch $\varphi(w)$ senkrecht auf v steht. Dies folgt so:

$$\langle \varphi(w), v \rangle = \langle w, \varphi(v) \rangle = \langle w, kv \rangle = k \langle w, v \rangle = 0.$$

(Das erste Gleichheitszeichen gilt, da φ selbstadjungiert ist, das zweite, da v ein Eigenvektor zu einem Eigenwert k ist.) Das bedeutet $\varphi(W) \subseteq W$. Da φ bijektiv ist, folgt daraus auch $\varphi(W) = W$.

Also ist die Einschränkung ψ von φ auf W eine selbstadjungierte lineare Abbildung von W in sich. Da W die Dimension $n - 1$ hat, können wir jetzt Induktion anwenden: Es gibt eine Orthonormalbasis C von W , so dass die Darstellungsmatrix ${}_C M_C(\psi)$ eine Diagonalmatrix ist.

Wenn wir $B := C \cup \{v\}$ setzen, so ist B eine Orthonormalbasis von V mit der Eigenschaft, dass die Darstellungsmatrix ${}_B M_B(\varphi)$ eine Diagonalmatrix ist. Also ist φ diagonalisierbar, und damit ist alles gezei ... – Halt! Keinesfalls ist alles gezeigt. Denn es ist noch die scheinbar kleine Behauptung „jede selbstadjungierte lineare Abbildung hat einen Eigenvektor“ zu beweisen, und das ist gar nicht ohne.

Das geht am besten über den Umweg über komplexe Vektorräume. Diese werden im Projekt dieses Kapitels untersucht, und die gewünschte Aussage wird in Teilschritt 15 ausdrücklich bewiesen. Wem ein Umweg über die Analysis lieber ist, der lese das Lemma auf Seite 207 von [Fis] oder die Aussage $F3$ im zweiten Band von [Lor].

Korollar

Eine symmetrische Bilinearform $[\cdot, \cdot]$ von V ist genau dann ein Skalarprodukt, wenn es eine Orthonormalbasis B gibt, so dass die Gramsche Matrix von $[\cdot, \cdot]$ bezüglich B nur positive Eigenwerte hat. \square

An den obigen Satz schließt sich der Sylverstersche Trägheitssatz (James Joseph Sylvester, 1814–1897) an, der den Reigen der Linearen Algebra beschließen soll.

Sylvesterscher Trägheitssatz

Sei $[\cdot, \cdot]$ eine symmetrische Bilinearform des euklidischen Vektorraums V . Für eine Orthonormalbasis B sei A die Gramsche Matrix von $[\cdot, \cdot]$. Sei s die Anzahl der positiven und t die Anzahl der negativen Eigenwerte von A . Dann hängen s und t nur von $[\cdot, \cdot]$, aber nicht von der Wahl der Basis B ab.

Wir beweisen diesen Satz hier nicht. Sie finden den Beweis in vielen Büchern über Lineare Algebra (zum Beispiel [Fis], [Lor]). Nehmen Sie die Herausforderung an, indem Sie versuchen, den Satz in einem dieser Bücher (a) zu finden (b) zu verstehen und (c) den dort angegebenen Beweis nachzuvollziehen.

10.6 Verständnisfragen, Übungen und Tipps

Richtig oder falsch?

1. Thema: Bilinearformen

- ☐ Sei $\langle \cdot, \cdot \rangle$ eine Bilinearform des Vektorraums V . Dann gilt:
- ☐ Für jeden Vektor v ist $\langle v, v \rangle$ positiv.
- ☐ Für jeden Vektor $v \neq 0$ ist $\langle v, v \rangle$ positiv.
- ☐ Es gibt einen Vektor $v \neq 0$, so dass $\langle v, v \rangle$ positiv ist.
- ☐ Für keinen Vektor $v \neq 0$ gilt $v \perp v$.
- ☐ Wenn $\langle \cdot, \cdot \rangle$ nichtausgeartet ist, gilt für keinen Vektor $v \neq 0$ die Beziehung $v \perp v$.
- ☐ Für jeden Unterraum U von V gilt $U^\perp \cap U = \{0\}$.

2. Thema: Skalarprodukte

- ☐ Das Standardskalarprodukt ist das einzige Skalarprodukt.
- ☐ Ein euklidischer Vektorraum hat nur ein Skalarprodukt.
- ☐ Das orthogonale Komplement eines Unterraums ist eindeutig bestimmt.
- ☐ Wenn $\langle \cdot, \cdot \rangle$ ein Skalarprodukt ist, nimmt $\langle v, v \rangle$ alle Werte aus \mathbf{R} an.
- ☐ Wenn $\langle \cdot, \cdot \rangle$ ein Skalarprodukt ist, nimmt $\langle v, v \rangle$ alle positiven Werte aus \mathbf{R} an.

3. Thema: Orthonormalbasen

Sei V ein euklidischer Vektorraum.

- ☐ V hat genau eine Orthonormalbasis
- ☐ Jede orthonormale Menge kann eindeutig zu einer Orthonormalbasis fortgesetzt werden.
- ☐ Jeder Vektor $\neq 0$ kann zu einer Orthonormalbasis ergänzt werden.
- ☐ Jeder Vektor v mit $\|v\| = 1$ kann zu einer Orthonormalbasis ergänzt werden.

Übungsaufgaben

1. Zeigen Sie, dass die beiden „Definitionen“ aus Abschn. 10.1 auch für Vektoren beliebiger Länge äquivalent sind.
2. Sei $V = K^2$. Dann ist die folgendermaßen definierte Abbildung $\langle \cdot, \cdot \rangle$ eine Bilinearform:

$$\langle (x_1, y_1), (x_2, y_2) \rangle := x_1 y_2 - y_1 x_2.$$

Zeigen Sie, dass $\langle \cdot, \cdot \rangle$ nichtausgeartet ist, dass aber für jeden Vektor v gilt:

$$\langle v, v \rangle = 0.$$

3. Sei $\langle \cdot, \cdot \rangle$ eine Bilinearform des Vektorraums V . Zeigen Sie, dass für alle $v \in V$ gilt:

$$\langle o, v \rangle = 0 \quad \text{und} \quad \langle v, o \rangle = 0.$$

Mit anderen Worten: der Nullvektor steht auf jedem Vektor senkrecht.

4. Sei $\langle \cdot, \cdot \rangle$ eine nichtausgeartete Bilinearform von V . Sei f eine lineare Abbildung von V in sich. Zeigen Sie, dass durch die Vorschrift

$$[v, w] := \langle f(v), w \rangle$$

eine Bilinearform von V definiert wird.

5. Geben Sie eine Gram-Matrix eines Skalarprodukts eines 4-dimensionalen reellen Vektorraums an, deren sämtliche Einträge $\neq 0$ sind.
6. Formulieren und beweisen Sie die Aussage für die zweite Komponente der entsprechenden Bilinearformen des Satzes „Konstruktion von Bilinearformen“ aus Abschn. 10.2.
7. Sei $\langle \cdot, \cdot \rangle$ eine symmetrische Bilinearform des Vektorraums V . Zeigen Sie: Die Menge

$$R := \{v \in V \mid v^\perp = V\}$$

ist ein Unterraum von V . (Man nennt R das **Radikal** der Bilinearform $\langle \cdot, \cdot \rangle$).

8. Sei V ein reeller Vektorraum, und sei $B = \{v_1, v_2, \dots, v_n\}$ eine Basis von V . Für zwei Vektoren

$$v = \sum_{i=1}^n k_i v_i \quad \text{und} \quad w = \sum_{j=1}^n h_j v_j$$

definieren wir

$$\langle v, w \rangle := k_1 h_1 + k_2 h_2 + \dots + k_{n-1} h_{n-1} + k_n h_n.$$

Bestimmen Sie das Radikal von $\langle \cdot, \cdot \rangle$. Welche Dimension hat dieses Radikal?

9. Sei V ein K -Vektorraum, und sei $\langle \cdot, \cdot \rangle$ eine nichtausgeartete Bilinearform auf V . Zeigen Sie, dass die Abbildung g_{w_0} von V nach K , die durch

$$g_{w_0}(v) = \langle v, w_0 \rangle \quad \text{für alle} \quad v \in V$$

definiert ist, genau dann die Nullabbildung ist, wenn $w_0 = o$ ist.

10. Beweisen Sie ohne Rückgriff auf den Satz über die Konstruktion von Skalarprodukten, dass die Standardbilinearform eines reellen Vektorraums ein Skalarprodukt ist.
11. Zeigen Sie: Die Menge der symmetrischen Matrizen aus $K^{n \times n}$ bildet einen K -Vektorraum. Bestimmen Sie die Dimension dieses Vektorraums.
12. Sei V ein euklidischer Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$. Zeigen Sie: Eine lineare Abbildung f von V in sich ist genau dann orthogonal, wenn für alle $v \in V$ gilt

$$\|f(v)\| = \|v\|.$$

13. Führen Sie den Beweis des Orthonormalisierungssatzes zu Ende.
14. Im ersten Teil des Beweises für den Orthonormalisierungssatz haben wir die Fälle $\langle v_2, w_1 \rangle = 0$ und $\langle v_2, w_1 \rangle \neq 0$ unterschieden. War das notwendig?
15. Ergänzen Sie den Vektor $\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{\sqrt{2}}\right)$ zu einer Orthonormalbasis von \mathbf{R}^3 .
16. Ergänzen Sie die Menge $\{(1/2, 1/2, 1/2), (-1/2, 1/2, -1/2, 1/2)\}$ zu einer Orthonormalbasis von \mathbf{R}^4 .
17. Zeigen Sie, dass $O(n)$ und $SO(n)$ (mit der Matrizenmultiplikation als Verknüpfung) Gruppen sind.
18. Sei $B = \{v_1, v_2, \dots, v_n\}$ eine Basis des euklidischen Vektorraums V . Zeigen Sie: Eine lineare Abbildung f von V in sich ist genau dann orthogonal, wenn $\langle v_i, v_j \rangle = \langle f(v_i), f(v_j) \rangle$ für alle $i, j \in \{1, 2, \dots, n\}$ gilt.
19. Sei $B = \{v_1, v_2, \dots, v_n\}$ eine Basis des euklidischen Vektorraums V . Zeigen Sie: Eine lineare Abbildung von V in sich ist genau dann selbstadjungiert, wenn $\langle f(v_i), v_j \rangle = \langle v_i, f(v_j) \rangle$ für alle $i, j \in \{1, 2, \dots, n\}$ gilt.
20. Zeigen Sie: Ist ψ eine selbstadjungierte lineare Abbildung von V in sich, so wird durch

$$[v, w] := \langle \psi(v), w \rangle$$

eine symmetrische Bilinearform definiert.

21. Zeigen Sie die Implikation „(a) \Rightarrow (b)“ im Satz über die Hauptachsentransformation (ohne die Gültigkeit von (a) vorauszusetzen).

Projekt: Skalarprodukte komplexer Vektorräume

Man kann ganz analog zur Theorie der Skalarprodukte über reellen Vektorräumen eine Theorie der Skalarprodukte über komplexen Vektorräumen aufziehen. Diese ist eine natürliche Verallgemeinerung der reellen Situation. Im Projekt dieses Kapitels sind Sie eingeladen, die Theorie der Skalarprodukte über komplexen Vektorräumen schrittweise zu entwickeln. Machen Sie sich stets klar, dass die im Haupttext entwickelte Theorie reeller Vektorräume ein Spezialfall der hier dargestellten komplexen Theorie ist und man die reelle Situation vom höheren Standpunkt aus besser überblicken kann.

Sei stets V ein endlichdimensionaler Vektorraum über dem Körper \mathbf{C} der komplexen Zahlen.

Zunächst müssen wir die Definition einer Bilinearform verallgemeinern. Eine **Semibilinearform** von V ist eine Abbildung $\langle \cdot, \cdot \rangle: V \times V \rightarrow \mathbf{C}$, die in der ersten Komponente linear, aber in der zweiten nur „semilinear“ ist; das bedeutet, dass für alle $v, v', w, w' \in V$ und alle $k \in K$ gilt

$$\begin{aligned} \langle v + v', w \rangle &= \langle v, w \rangle + \langle v', w \rangle \quad \text{und} \quad \langle kv, w \rangle = k \cdot \langle v, w \rangle, \quad \text{sowie} \\ \langle v, w + w' \rangle &= \langle v, w \rangle + \langle v, w' \rangle \quad \text{und} \quad \langle v, k \cdot w \rangle = \bar{k} \langle v, w \rangle \end{aligned}$$

Dabei ist \bar{k} die zu k konjugiert komplexe Zahl (vergleichen Sie dazu Abschn. 2.3.4).

Übrigens: Lassen Sie sich von dem merkwürdigen Wort „semibilinear“ nicht verwirren („semi“ heißt „halb“ und „bi“ bedeutet „zwei“). Diese beiden Operationen „heben

sich aber nicht auf“; vielmehr bezieht sich „bi“ darauf, dass es sich um eine Abbildung von $V \times V$ nach \mathbb{C} handelt, und die Vorsilbe „semi“ deutet an, dass die Abbildung nur *semilinear* (in der zweiten Komponente) ist.

1. Sei $B = \{v_1, v_2, \dots, v_n\}$ eine Basis von V . Sei $A = (a_{ij})$ eine $n \times n$ -Matrix mit Werten aus \mathbb{C} . Wir definieren eine Abbildung f von $V \times V$ in K , dadurch, dass wir für Basisvektoren v_i und v_j definieren

$$\langle v_i, v_j \rangle := a_{ij}$$

und diese Vorschrift „semilinear bzw. linear fortsetzen“. Das bedeutet: Sind

$$v = \sum_{i=1}^n k_i v_i \quad \text{und} \quad w = \sum_{j=1}^n h_j v_j$$

beliebige Vektoren aus V , so ist

$$\langle v, w \rangle := \sum_{i,j=1}^n k_i \bar{h}_j a_{ij}.$$

Zeigen Sie, dass f eine Semibilinearform ist:

Eine Semibilinearform heißt eine **hermitesche Form** (nach dem französischen Mathematiker Charles Hermite, 1822–1901), falls für alle $v, w \in V$ gilt:

$$\langle v, w \rangle = \overline{\langle w, v \rangle}.$$

2. Sei $\langle \cdot, \cdot \rangle$ eine hermitesche Form. Zeigen Sie, dass für jeden Vektor $v \in V$ gilt: $\langle v, v \rangle$ ist eine reelle (und nicht nur eine komplexe) Zahl.
3. Sei $\{v_1, v_2, \dots, v_n\}$ eine Basis von V . Sei $A = (a_{ij})$ eine $n \times n$ -Matrix mit Einträgen aus \mathbb{C} . Dann gilt: Genau dann wird durch die Vorschrift $\langle v_i, v_j \rangle := a_{ij}$ eine hermitesche Form definiert, wenn

$$A = \bar{A}^T \quad (:= (\bar{a}_{ji}))$$

gilt. Eine Matrix mit dieser Eigenschaft heißt auch eine **hermitesche Matrix**.

Wir nennen eine hermitesche Form $\langle \cdot, \cdot \rangle$ ein **Skalarprodukt** von V , falls sie positiv definit ist, falls also

$$\langle v, v \rangle \geq 0$$

gilt für alle $v \in V$ und $\langle v, v \rangle \neq 0$ für $v \neq 0$.

Schließlich nennen wir einen komplexen Vektorraum mit einem Skalarprodukt einen **unitären Vektorraum**.

4. Warum ist die Definition eines Skalarprodukts überhaupt möglich? (Bedenken Sie, dass der Körper \mathbf{C} keine mit Addition und Multiplikation verträgliche Anordnung hat!)
5. Die Einheitsmatrix definiert ein Skalarprodukt von V .

Von nun an sei stets V ein unitärer \mathbf{C} -Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$.

6. Definieren Sie $\|v\| := \sqrt{\langle v, v \rangle}$ für alle $v \in V$ und zeigen Sie die Ungleichung von Cauchy und Schwarz.
7. Zeigen Sie, dass durch die Vorschrift $\|v\| := \sqrt{\langle v, v \rangle}$ für alle $v \in V$ eine Norm auf V definiert wird.
8. Formulieren und beweisen Sie den Orthonormalisierungssatz von E. Schmidt für einen unitären Vektorraum.

Eine lineare Abbildung f von V in sich heißt **unitär**, falls für alle $v, w \in V$ gilt

$$\langle v, w \rangle = \langle f(v), f(w) \rangle.$$

Eine invertierbare Matrix $A \in \mathbf{C}^{n \times n}$ heißt **unitär**, falls gilt

$$A^{-1} = \bar{A}^T.$$

9. Sei f eine unitäre lineare Abbildung von V in sich. Zeigen Sie
 - (a) Jeder Eigenwert von f hat den Betrag 1.
 - (b) Die lineare Abbildung f ist injektiv (also bijektiv).
10. Sei B eine Orthonormalbasis von V . Zeigen Sie: Eine lineare Abbildung f von V in sich ist genau dann unitär, wenn ${}_B M_B(f)$ eine unitäre Matrix ist.
11. Sei f eine unitäre lineare Abbildung von V in sich. Zeigen Sie: Es gibt eine Orthonormalbasis von V , die aus Eigenvektoren von f besteht.
12. Zeigen Sie: Jede unitäre lineare Abbildung von V in sich ist diagonalisierbar.

Eine lineare Abbildung von V in sich heißt **selbstadjungiert**, wenn für alle $v, w \in V$ gilt

$$\langle f(v), w \rangle = \langle v, f(w) \rangle.$$

13. Sei f eine lineare Abbildung von V in sich, und sei B eine Orthonormalbasis von V . Zeigen Sie: Genau dann ist f selbstadjungiert, wenn die Darstellungsmatrix ${}_B M_B(f)$ eine hermitesche Matrix ist.
14. (a) Zu jeder hermiteschen Form $[\cdot, \cdot]$ gibt es eine Orthonormalbasis von V , die auch bezüglich $[\cdot, \cdot]$ eine Orthogonalbasis ist.

- (b) Jede selbstadjungierte lineare Abbildung von V in sich ist diagonalisierbar und hat nur reelle Eigenwerte.
- (c) Jede hermitesche Matrix hat nur reelle Eigenwerte.
15. Sei V ein euklidischer Vektorraum, und sei f eine selbstadjungierte lineare Abbildung von V in sich. Wir betrachten diesen Vektorraum auch als \mathbb{C} -Vektorraum V^* . Das bedeutet: Für irgendeine Orthonormalbasis $B = \{v_1, v_2, \dots, v_n\}$ von V sei

$$V^* := \left\{ \sum_{i=1}^n v_i z_i \mid z_i \in \mathbb{C} \right\}.$$

Zeigen Sie:

- (a) Die Darstellungsmatrix A von f bezüglich B ist (als Matrix über \mathbb{C} betrachtet) eine hermitesche Matrix. Daher definiert A eine selbstadjungierte lineare Abbildung f^* von V^* in sich.
- (b) Die lineare Abbildung f^* hat nur reelle Eigenwerte. Insbesondere hat f einen reellen Eigenvektor aus V .

Sie sollten mit folgenden Begriffen umgehen können

Bilinearform, orthogonal, Gramsche Matrix, symmetrische Bilinearform, positiv definit, Skalarprodukt, euklidischer Vektorraum, Standardskalarprodukt, Orthonormalbasis, orthogonales Komplement, orthogonale Abbildung, orthogonale Matrix, selbstadjungierte lineare Abbildung.

Adieu!

Es wird Zeit, vielleicht höchste Zeit, dass ich mich von Ihnen verabschiede. Wenn Sie die Lineare Algebra bis hierher gut durchgearbeitet und verstanden haben, stehen Sie jetzt auf sicherem Grund, und ich bin sicher, dass Sie auch Ihren weiteren Weg durchs Mathematikstudium erfolgreich beschreiten werden. Zur Vorsicht wünsche ich Ihnen aber auch das dazu nötige Quäntchen Glück.

Die Mathematik ist eine einzigartige Welt für eigene Entdeckungen. Sie können jede Übungsaufgabe, jeden etwas schwierigeren Text als eigenes kleines Forschungsprojekt auffassen. Nach einiger Zeit werden Sie die Lehrbücher hinter sich lassen, selbständig Fragen stellen (nichts schwieriger – und wichtiger – als das!) und Antworten versuchen. Auch wenn viele Ihrer Entdeckungen schon früher von jemand anderem gemacht wurden, sollte das Ihren Genuss nicht trüben und Sie keinesfalls entmutigen! (Es macht Ihnen doch auch nichts aus, wenn Ihr derzeitiges Lieblingsmusikstück schon von Tausenden anderer vorher gehört wurde.)

Es ist so wie in der Erzählung „Der Erfinder“ von Peter Bichsel. Dort wird von einem Erfinder berichtet, der nur Dinge erfindet, die es schon gibt. Am Schluss tröstet uns Peter Bichsel mit folgenden Worten:

Doch er blieb sein Leben lang ein richtiger Erfinder,
denn auch Sachen, die es gibt, zu erfinden, ist schwer,
und nur Erfinder können es.



11.1 Lösungsvektoren der □-Aufgaben

Kapitel 1

1. r f f f f r r f
2. f f r f r
3. f r r f f
4. f r r f f
5. r r f r

Kapitel 2

1. f f r f
2. r f f f f r f

Kapitel 3

1. r f r f r f f r f r f r f f
2. f f f r m h
3. f f f r 0, 1, 2001, ∞
4. r f r r f f f f
5. f f r
6. f r f r f f
7. f f r f r f f f r

Kapitel 4

1. fffrffffffffff
2. frfrf
3. rfr
4. fffrf

Kapitel 5

1. fffffrr
2. rfrffff
3. frffffr
4. fffrfr r
5. ffrffff

Kapitel 6

1. rffffrfr r r r f f
2. fr
3. frfr r

Kapitel 7

1. fr r f f
2. r r f f
3. f f f r f f

Kapitel 8

1. r f f f r f f f f f
2. fr r f f
3. f f f f r q
4. f f f f r r
5. f f f f f r f f f

Kapitel 9

1. r r f r f r
2. r f f
3. r f f r f
4. f r f f
5. f f f f r f

Kapitel 10

1. f f f f f f
2. f f r f r
3. f f f r

(r: richtig, f: falsch, q: Quatsch)

11.2 Tipps zur Lösung der Übungsaufgaben

Lange Zeit habe ich mich dagegen gesträubt, Hinweise zur Lösung der Übungsaufgaben zu veröffentlichen. Mir selbst gegenüber konnte ich das gut begründen: Jeder soll eine Chance haben, die Lösungen selbst zu finden. Jeder soll sich selbst davon überzeugen, dass die Lösung richtig ist – und nicht nur nachschlagen. Viele Aufgaben haben viele Lösungen, „die Musterlösung“ gibt es kaum einmal.

Aber viele Leserinnen und Leser der Linearen Algebra sehen das anders. Ich habe viele begeisterte Briefe erhalten – die alle mit der Klage endeten, dass dieses Buch keine Tipps zur Lösung der Übungsaufgaben enthalte. Sie haben mich überzeugt: Hier sind Lösungshinweise!

Aber: Es sind wirklich nur Tipps und Hinweise, manchmal nur Stichworte. Nicht jede Aufgabe wird gelöst. Manche werden nur teilweise gelöst. Manchmal wird nur das Ergebnis genannt. Und es gibt keine Musterlösungen.

Ehrlich gesagt: Ich hatte noch einen Grund, keine Lösungen von Übungsaufgaben zu veröffentlichen: Ich verrechne mich laufend. Und ich fürchte, dass viele der Hinweise kleinere oder größere Lücken oder Fehler enthalten. Also seien Sie misstrauisch – und trauen Sie im Zweifelsfall Ihrer eigenen Lösung mehr als der hier vorgeschlagenen! (Aber wenn Sie einen wirklich verwirrenden Fehler finden, dann bitte ich Sie, mir diesen mitzuteilen!)

Kapitel 1

1. Typisch mathematischer Trick: Wenn X die leere Menge ist, kann Y alles sein, und die Bedingung gilt immer noch.
Die Antwort ist also „nein“ (weil es ein Gegenbeispiel gibt), obwohl sie in fast allen Fällen „ja“ lautet.
2. Die Anzahl der Möglichkeiten ist $4 \cdot 2 \cdot 10 \cdot 3 = 240$. Ich brauche also 20 Jahre.
3. (a) Die Relation \sim definiert auf der Menge der Bürger von Deutschland durch

$$x \sim y \Leftrightarrow x \text{ und } y \text{ haben ihren ersten Wohnsitz in derselben Stadt}$$

ist eine Äquivalenzrelation. Denn jeder Bürger hat einen Wohnsitz (Reflexivität). Wenn A den gleichen Wohnsitz hat wie B , dann hat auch B den gleichen Wohnsitz

wie A (Symmetrie). Wenn schließlich A und B sowie B und C den gleichen Wohnsitz haben, dann haben alle drei den gleichen Wohnsitz, insbesondere haben A und C den gleichen Wohnsitz (Transitivität).

- (c) Die Gleichheitsrelation $=$ ist eine Äquivalenzrelation. Dies ist die wichtigste Äquivalenzrelation überhaupt, insbesondere da es viele Äquivalenzrelationen gibt, die letztlich über eine Gleichheit definiert sind. Beispiel: Zwei Personen sind äquivalent, wenn sie die gleiche Haarfarbe haben.
- (e) Für jede natürliche Zahl n ist die Relation \sim_n auf der Menge \mathbf{Z} der ganzen Zahlen, die wie folgt definiert ist:

$$x \sim_n y \Leftrightarrow y - x \text{ ist ein ganzzahliges Vielfaches von } n$$

eine Äquivalenzrelation. Wir zeigen nur die Transitivität: Sei $y - x$ ein ganzzahliges Vielfaches von n , also $y - x = kn$ mit $k \in \mathbf{Z}$. Sei ferner auch $y \sim_n z$, also $z - y$ ein ganzzahliges Vielfaches von n , d. h. $z - y = hn$ mit $h \in \mathbf{Z}$. Dann ist $z - x = (z - y) - (y - x) = hn - kn = (h - k)n$, wobei $h - k$ ein Element von \mathbf{Z} ist. Somit gilt auch $x \sim_n z$.

4. Auf der Menge \mathbf{Z} haben die folgendermaßen definierten Relationen die gewünschten Eigenschaften:

$|x - y| \leq 2$: reflexiv, symmetrisch, aber nicht transitiv,

$x - y$ ist eine nichtnegative gerade Zahl: reflexiv, transitiv, aber nicht symmetrisch,

$x - y$ ist eine Primzahl: weder reflexiv, noch symmetrisch noch transitiv.

Um eine Relation zu erhalten, die symmetrisch und transitiv, aber nicht reflexiv ist kann man wie folgt vorgehen: Man nimmt eine Menge M mit Äquivalenzrelation. Zu dieser Menge fügt man ein neues Element ∞ hinzu, das weder mit sich selbst noch mit den Elementen aus M in Relation stehen soll. Dann ist die Relation auf der neuen Menge symmetrisch und transitiv, aber nicht reflexiv.

5. Sei $z \in A(y)$. Das bedeutet $z \sim y$. Aus $x \sim y$ ergibt sich zunächst $y \sim x$, und wegen der Transitivität von \sim folgt auch $z \sim x$. Das heißt $z \in A(x)$.
6. (a) Man zeigt zunächst, dass sich g und g' nicht schneiden (mit anderen Worten: dass g' ganz auf einer Seite von g liegt). Dann kann man zeigen, dass g und g' die (Verlängerung von) gegenüberliegenden Seiten eines Rechtecks sind. Also sind g und g' parallel.
7. (a) Reflexivität: Jede natürliche Zahl hat eine Quersumme. Symmetrie: Wenn die erste Zahl die gleiche Quersumme wie die zweite hat, dann hat auch die zweite Zahl die gleiche Quersumme wie die erste. Transitivität: Wenn die beiden ersten Zahlen die gleiche Quersumme haben und die zweite und die dritte Zahl die gleiche Quersumme haben, dann haben alle Zahlen die gleiche Quersumme, insbesondere also die erste und die dritte.
8. Welche der folgenden Zuordnungen ist eine Abbildung, welche ist surjektiv, welche injektiv?
- Mensch \longrightarrow Freundin: Keine Abbildung, da es Menschen mit mehr als einer Freundin gibt.

Mensch \mapsto Vater: Abbildung, da jeder Mensch nur einen Vater hat. Die Abbildung ist nicht injektiv, da manche Menschen den gleichen Vater haben; sie ist auch nicht surjektiv, da nicht jeder Mensch Vater ist.

Mensch \mapsto Handynummer: Injektive Abbildung, falls jeder der betrachteten Menschen nur eine Handynummer hat.

Mensch \mapsto Lieblingessen: Abbildung, falls jeder Mensch genau ein Lieblingessen hat.

9. $f(x) = x^3$: injektiv und surjektiv
 $f(x) = ax^2 + bx + c$ ($a \neq 0$): weder injektiv noch surjektiv
 $f(x) = |x|$: weder injektiv noch surjektiv
 $f(x) = e^x$: injektiv, aber nicht surjektiv
10. Die Abbildung $f: \{0, 1\} \rightarrow \{0, 1\}$ mit $f(0) = 0$ und $f(1) = 1$ ist injektiv und surjektiv.
 Eine Abbildung $f \subseteq X \times Y$ ist genau dann bijektiv, wenn es zu jedem $y \in Y$ genau ein $y \in Y$ gibt mit $(x, y) \in f$.
 Die Abbildung $f: \{0, 1\} \rightarrow \{a, b, c\}$ mit $f(0) = a$ und $f(1) = b$ ist injektiv, aber nicht surjektiv.
 Die Abbildung $f: \{a, b, c\} \rightarrow \{0, 1\}$ mit $f(a) = 0$ und $f(b) = f(c) = 1$ ist surjektiv, aber nicht injektiv.
 Die Abbildung $f: \{0, 1\} \rightarrow \{a, b, c\}$ mit $f(0) = a$ und $f(1) = a$ ist weder injektiv noch surjektiv.
11. Eine Abbildung $f \subseteq X \times Y$ ist genau dann injektiv, wenn es zu jedem $y \in Y$ höchstens ein $x \in X$ gibt mit $(x, y) \in f$.
 (Oder: ..., wenn es zu jedem $x \in X$ genau ein $y \in Y$ gibt mit $(x, y) \in f$.)
 Eine Abbildung $f \subseteq X \times Y$ ist genau dann surjektiv, wenn es zu jedem $y \in Y$ mindestens ein $x \in X$ gibt mit $(x, y) \in f$.
 Eine Abbildung $f \subseteq X \times Y$ ist genau dann bijektiv, wenn es zu jedem $y \in Y$ genau ein $x \in X$ gibt mit $(x, y) \in f$.
12. Wenn f nicht injektiv wäre, dann gäbe es verschiedene Elemente x, x' aus X mit $f(x) = f(x')$. Dann wäre auch $f' \circ f(x) = f' \circ f(x')$, also $f' \circ f \neq \text{id}_X$.
 Sei $y \in Y$ beliebig. Dann ist $f(y)$ ein Urbild von y ; denn es ist $y = \text{id}_Y(y) = f \circ f'(y) = f(f'(y))$. Also ist f auch surjektiv.
13. (a) Genau dann ist f injektiv, wenn alle Bilder der Elemente von $|X|$ verschieden sind, also genau dann, wenn f genau $|X|$ Bilder hat.
 (b) Sei $X = Y$ endlich. Nach (a) gilt dann: f ist injektiv $\Leftrightarrow f$ hat genau $|X|$ Bilder $\Leftrightarrow f$ ist surjektiv.
 (c) folgt aus (b).
 (d) Die Abbildung $f: \mathbb{Z} \rightarrow \mathbb{Z}$, die definiert ist durch $f(z) := 2z$ ist injektiv, aber nicht surjektiv. Eine Abbildung von \mathbb{Z} nach \mathbb{Z} , die surjektiv, aber nicht injektiv ist, ist die folgendermaßen definierte Abbildung g : Sei z eine ganze Zahl mit n Dezimalstellen. Wenn n gerade ist, sei $g(z)$ die Zahl aus den letzten $n/2$ Stellen; wenn n ungerade ist, sei $g(z)$ die Zahl aus den letzten $(n+1)/2$ Stellen. Dabei soll das Vor-

zeichen erhalten werden, d. h. wenn z positiv ist, soll auch $g(z)$ positiv sein; wenn z negativ ist, soll auch $g(z)$ negativ sein.

14. Wenden Sie Aufgabe 12 an.
15. Die Abbildung $f: \mathbf{Z} \rightarrow 2\mathbf{Z}+1$, definiert durch $f(z) := 2z+1$ ist bijektiv, und somit sind die Mengen gleichmächtig.
16. (a) Beweis nach Euklid (ca. 300 v. Chr.). Angenommen, es gibt nur endlich viele Primzahlen, seien diese p_1, p_2, \dots, p_s . Wir betrachten die Zahl $n := p_1 \cdot p_2 \dots p_s + 1$. Diese Zahl ist – wie jede natürliche Zahl > 1 – durch eine Primzahl teilbar; diese muss eine der Primzahlen p_1, p_2, \dots, p_s sein. Sei p_i ein Teiler von n . Die Zahl p_i teilt aber auch das Produkt $p_1 \cdot p_2 \dots p_s (= n-1)$ da p_i einer der Faktoren ist. Da p_i auch die Differenz der beiden Zahlen teilt, muss p_i ein Teiler von 1 sein: ein Widerspruch.
(b) klar nach (a).
17. Es gibt viele bijektive Abbildungen von \mathbf{N} nach \mathbf{Q} . Die meines Erachtens einfachste kann man wie folgt konstruieren: Man ordnet die (positiven) rationalen Zahlen nach der Größe der Summe aus Zähler und Nenner: Zuerst kommen die Brüche, bei denen diese Summe 2 ist (es gibt nur einen einzigen, nämlich $1/1$). Dann kommen die Brüche, bei denen die Summe 3 ist (das sind $1/2$ und $2/1$). Und so weiter. Innerhalb jeder dieser Gruppe ordnet man die Brüche nach der Größe der Zähler. Damit kann man die positiven rationalen Zahlen der Reihe nach aufschreiben. Dies ist eine Abbildung von \mathbf{N} in die positiven rationalen Zahlen. Also sind diese Mengen gleichmächtig. Den Rest zeigt man wie die Gleichmächtigkeit von \mathbf{N} und \mathbf{Z} .
18. Sie können auch jedes andere Buch über Mengenlehre oder über die mathematische Unendlichkeit lesen, um diese Aufgabe zu lösen.
19. Der entscheidende Trick des Beweises besteht darin, dass die Menge U ein Urbild u hat. Wir benutzten also nur die Surjektivität!
20. $a_{32} + a_{33} + a_{34} + a_{35} + a_{36} + a_{42} + a_{43} + a_{44} + a_{45} + a_{46} + a_{52} + a_{53} + a_{54} + a_{55} + a_{56}$.
21. Die Anzahl der binären Folgen der Länge n ist 2^n .
22. (a) Wir nummerieren die Elemente von $X: X = \{x_1, x_2, \dots, x_n\}$. Wir definieren die Abbildung f von der Menge aller Teilmengen von X in die Menge aller binären Folgen der Länge n wie folgt: Sei Y eine Teilmenge von X . Dann ist $f(Y) = (b_1, b_2, \dots, b_n)$, wobei $b_i = 1$ ist, falls x_i in Y liegt, und 0 sonst.
Diese Abbildung f ist bijektiv.
(b) Da es genau 2^n binäre Folgen der Länge n gibt, gibt es nach (a) auch genau 2^n Teilmengen einer n -elementigen Menge.
23. q^n .
24. Die Summe der ersten n Kubikzahlen ist gleich dem Quadrat der Summe der ersten n natürlichen Zahlen.
Wegen $1+2+\dots+n = n(n+1)/2$, ist nur zu zeigen: $1^3 + 2^3 + 3^3 + \dots + n^3 = n^2(n+1)^2/4$.
Induktionsschritt: $n^2(n+1)^2/4 + (n+1)^3 = \dots = (n+1)^2(n+2)^2/4$.
25. $a_0 \cdot a_1 \cdot a_2 \cdot a_3 \cdot a_4 \cdot a_5 \cdot a_6 \cdot a_7 \cdot a_8 \cdot a_9 \cdot a_{10}$ (falls k eine natürliche Zahl ist).
 $M_5 \cup M_6 \cup M_7 \cup \dots \cup M_{n-1}$.
 $X_0 \cup X_1 \cup X_2 \cup \dots$

Kapitel 2

1. „Distributives Ausrechnen“, $i \cdot b = b \cdot i$, Definition von i^2 , Definition von $z \cdot z'$.
2. Das ist zugegebenermaßen ziemlich mühsam, kann aber grundsätzlich auf wiederholte Anwendung von zwei Operationen ausgeführt werden: (a) Anwenden der Definition der Multiplikation komplexer Zahlen, (b) Ausnutzen der entsprechenden Gesetze in \mathbf{R} .
3. $(5/29 - 2i/29)$, $(7/50 + i/50)$, $(1/5 - 2i/5)$.
4. Zum Beispiel: $i \cdot (j \cdot k) = i \cdot i = -1$ und $(i \cdot j) \cdot k = k \cdot k = -1$.
5. (a) Ausrechnen.
(b) Folgt aus der Definition der Addition und Multiplikation in \mathbf{H} .
6. Bearbeiten Sie diese Aufgabe und erleben Sie, wie durch diese Definition alles klar und einfach wird.
7. Wenn die Zahlen $a' - a$ und $b' - b$ durch n teilbar sind, dann sind auch $a' + b' - (a + b)$ und $a' \cdot b' - a \cdot b$ durch n teilbare ganze Zahlen.
8. $x = 4$ (denn $6 \cdot 4 \bmod 11 = 24 \bmod 11 = 2$), $x = 8$, $x = 1$, $x = 8$.
9. $6x = 2$ ist in \mathbf{Z}_{12} nicht lösbar, da $6x$ nur die Werte 6 und 0 annehmen kann. $2x + 4 = 9$ ist nicht lösbar, da $2x + 4$ nur die Werte 0, 2, 4, 6, 8, 10 annehmen kann. $3x - 9 = 5$ ist ebenfalls nicht lösbar, aber $7x = 1$ ist lösbar.
10. In \mathbf{Z}_{ab} mit $a > 1$ und $b > 1$ sind a und b von Null verschiedene Elemente, deren Produkt Null ergibt.
Sei U eine Teilmenge von \mathbf{Z}_{ab} , die mit der Addition und Multiplikation von \mathbf{Z}_{ab} ein Körper ist. Dann ist 1 ein Element von U , also auch $1+1$, $1+1+1$ usw. Damit erhält man – mit der Addition modulo ab – alle Elemente von \mathbf{Z}_{ab} . Also ist $U = \mathbf{Z}_{ab}$, und somit ist U kein Körper.
11. Man führt das Assoziativ- und das Distributivgesetz in \mathbf{Z}_n auf die Gültigkeit der entsprechenden Gesetze in \mathbf{Z} zurück.
12. $5^2 \bmod 7 = 4$, $5^3 \bmod 7 = 5^2 \cdot 5 \bmod 7 = 4 \cdot 5 \bmod 7 = 6$, $5^4 \bmod 7 = 5^3 \cdot 5 \bmod 7 = 6 \cdot 5 \bmod 7 = 2$, $5^5 \bmod 7 = 5^3 \cdot 5 \bmod 7 = 2 \cdot 5 \bmod 7 = 3$, $5^6 \bmod 7 = 5^5 \cdot 5 \bmod 7 = 3 \cdot 5 \bmod 7 = 1$. Also ist die Ordnung von 5 in \mathbf{Z}_7 gleich 6.
13. (a) Wir verwenden zunächst Aufgabe 17. Angenommen, es gibt einen Körper K mit genau 6 Elementen. Es gibt eine Primzahl p mit $p \cdot 1 = 0$. Es folgt $p \in \{2, 3, 5\}$. Sei $F := \{0, 1, p \cdot 1, \dots, (p-1) \cdot 1\}$.
(b) $p \neq 5$. Sonst gäbe es genau ein Element a , das nicht in F liegt. Dann liegt auch $a+1$ nicht in F . (Sonst $a+1 = n \cdot 1$, also $a = (n-1) \cdot 1 \in F$)
(c) $p \neq 3$. Sonst $F = \{0, 1, 2\}$. Sei a ein Element von K , das nicht in F liegt. Dann ist auch $2 \cdot a$ nicht in F . Also muss es ein Element b aus K geben, das ungleich a und $2a$ ist und nicht in F liegt. Dann liegt $2 \cdot b$ weder in F , noch ist es gleich a oder gleich $2a$, noch gleich b : Widerspruch.
(d) $p \neq 2$. In diesem Fall wäre $F = \{0, 1\}$, und es folgte $K = \{0, 1, a, a+1, b, b+1\}$ für geeignete Elemente a und b . Betrachte die Elemente $a+b$ und $a+b+1$. Beide können weder a , noch b , noch $a+1$, noch $b+1$ sein. Also ist $a+b = 0$ und $a+b+1 = 1$. Aus $a+b = 0$ folgt aber $a = b$ (da $-1 = 1$), ein Widerspruch.

14. Setze $K := \{a+bi \mid a, b \in \mathbb{Z}_3\}$. Definiere Addition und Multiplikation analog wie bei der Konstruktion von \mathbb{C} . Genau so wie dort kann man alle Körperaxiome nachweisen. Nur beim Nachweis der Existenz eines multiplikativ Inversen muss man anders nachweisen, dass für alle $a, b \in \mathbb{Z}_3$ mit $(a, b) \neq (0, 0)$ stets $a + b \neq 0$ ist.
15. (a) Das Nullelement ist $(0, 0)$, das Einselement ist $(1, 0)$.
 (b) Das Element (x', y') mit $x' = x/(x^2+y^2)$ und $y' = -y/(x^2+y^2)$ ist ein multiplikatives Inverses von (x, y) .
 (c) Ja.
16. $f(k^{-1}) \cdot f(k) = f(k^{-1} \cdot k) = f(1) = 1$. Also ist $f(k^{-1})$ die multiplikative Inverse von $f(k)$. Das heißt $f(k^{-1}) = f(k)^{-1}$.
17. (a) Da K endlich ist, passiert es beim Aufaddieren der Eins, dass man irgendwann ein Element erhält, das man schon vorher erhalten hatte. Mit anderen Worten, es gibt positive natürliche Zahlen r und s mit $r < s$ und $r \cdot 1 = s \cdot 1$. Daraus folgt $(s - r) \cdot 1 = 0$. Setze $n := s - r$.
 (b) Angenommen, p wäre keine Primzahl. Dann gäbe es natürliche Zahlen a und b mit $a, b > 1$ und $p = ab$. Also wäre $0 = p \cdot 1 = (ab) \cdot 1 = (a \cdot 1) \cdot (b \cdot 1)$. Also muss $a \cdot 1 = 0$ oder $b \cdot 1 = 0$ gelten. Da $a, b < p$ gilt, widerspricht dies der Minimalität von p .
18. (a) In \mathbb{Z}_p sind alle Elemente $a \cdot 1$ mit $1 \leq a < p$ verschieden von Null, aber $p \cdot 1 = 0$. Also hat \mathbb{Z}_p die Charakteristik p .
 (b) In einem Körper der Charakteristik p ist die Menge $\{a \cdot 1 \mid 0 \leq a < p\}$ ein Körper, der isomorph zu \mathbb{Z}_p ist.
 (c) Sei k ein Element von K , das nicht in dem zu \mathbb{Z}_p isomorphen Teilkörper F liegt. Dann sind die Elemente $a + kb$ mit $a, b \in F$ alle verschieden.
19. Der Körper $\text{GF}(2)$ hat nur die Identität als Automorphismus, denn 0 und 1 müssen in jedem Fall festbleiben. Auch $\text{GF}(3)$ hat nur die Identität als Automorphismus. Dagegen hat $\text{GF}(4)$ zwei Automorphismen.
20. Sei f ein Automorphismus von \mathbb{Z}_p . Da $f(1) = 1$ ist, gilt auch $f(2) = f(1+1) = f(1)+f(1) = 1+1 = 2$. Analog folgt $f(3) = 3, \dots, f(p-1) = p-1$. Also ist $f = \text{id}$.
 Allgemein gilt: Jeder Automorphismus eines Körpers lässt den Primkörper elementweise fest.
21. (a) Angenommen, das Element $a \neq 0$ würde durch f auf 0 abgebildet werden. Dann folgte $f(1) = f(a^{-1} \cdot a) = f(a^{-1}) \cdot f(a) = f(a^{-1}) \cdot 0 = 0$: Widerspruch.
 (b) Sei $f(a) = f(b)$. Dann ist $f(a - b) = f(a) - f(b) = 0$. Nach (a) ist dann $a - b = 0$, also $a = b$.

Kapitel 3

1. Tun Sie das!
2. In den ersten beiden Fällen werden die Addition und die Skalarmultiplikation komponentenweise definiert, im dritten Fall punktweise.
3. Die Behauptung ist bei diesem Gleichungssystem besonders einfach zu zeigen, denn die einzige Lösung ist $x = 0, y = 0, z = 0$. Der Lösungsraum hat keinen Freiheitsgrad.

4. $+, -, +, +, -$

5. (a) Eine Gerade g der euklidischen Ebene \mathbf{R}^2 durch den Nullpunkt ist die Menge der Punkte (x, y) , die einer Gleichung $y = mx + b$ genügen ($m, b \in \mathbf{R}$ fest), d. h.

$$g = \{(x, y) | y = mx + b\}.$$

Dafür kann man die Axiome eines Vektorraums leicht nachweisen.

- (b) Jede Ebene E des \mathbf{R}^3 durch den Nullpunkt besteht aus den Punkten (x, y, z) , die einer Gleichung $ax + by + cz = 0$ genügen. Beweis analog zu (a).
6. Eine Gerade der euklidischen Ebene \mathbf{R}^2 , die nicht durch den Nullpunkt geht, enthält nicht den Nullvektor, kann also kein Unterraum sein.
7. (a) $(0, 0, 0), (1, 2, 1), (1/2, 0, 1), (1, -2, 1), (2000, 2000, 3000), \dots$
 (b) Ein Tripel (a, b, c) ist ein Gewichtssatz, falls gilt $20a = 5b + 10c$. Diese Tripel bilden einen Vektorraum.
 (c) 2
8. Der Nachweis der Körperaxiome vereinfacht sich u. a. dadurch, dass man die gesamten Eigenschaften der Addition „geschenkt bekommt“.
9. Jeder Vektor, der an den beiden letzten Komponenten verschiedene Einträge hat, liegt nicht im Erzeugnis der beiden Vektoren. In allen anderen Fällen erzeugen die angegebenen Vektoren den \mathbf{R}^3 .
10. Das schaffen Sie ohne Hilfe!
11. Sei $a(1, x, x^2) + b(1, y, y^2) + c(1, z, z^2) = (0, 0, 0)$, wobei x, y, z drei verschiedene Elemente $\neq 0$ sind. Dann gilt $a+b+c=0$, $ax+by+cz=0$ und $ax^2+by^2+cz^2=0$. Daraus erhält man $a=b=c=0$.
12. (a) $k \cdot o = k \cdot (o + o) = k \cdot o + k \cdot o$. Wenn man auf beiden Seiten $-k \cdot o$ addiert, ergibt sich $o = k \cdot o$.
 $(-k) \cdot v + k \cdot v = (-k+k) \cdot v = 0 \cdot v = o$. Also ist $(-k) \cdot v$ das Inverse von kv , d. h. $(-k) \cdot v = -kv$.
 $k \cdot (-v) + k \cdot v = k \cdot (-v+v) = k \cdot o = o$. Also ...
 (b) $k_1 \cdot v = k_2 \cdot v \Leftrightarrow (k_1 - k_2) \cdot v = o \Leftrightarrow k_1 - k_2 = 0 \Leftrightarrow k_1 = k_2$.
13. (a) Wenn v und w Linearkombinationen der Vektoren v_1, \dots, v_n sind, dann ist auch $v+w$ eine Linearkombination der Vektoren v_1, \dots, v_n .
14. Sei zunächst B eine Basis. Dann ist B nach Definition ein Erzeugendensystem. Dieses ist auch minimal: Angenommen, für $v \in B$ wäre auch $B \setminus \{v\}$ eine Basis. Dann wäre v eine Linearkombination der anderen Vektoren aus B ; also wäre B nicht linear unabhängig. Umgekehrt möge B ein minimales Erzeugendensystem sein. Wäre B nicht linear unabhängig, dann könnte man einen Vektor v aus B als Linearkombination der anderen darstellen. Dann wäre auch $B \setminus \{v\}$ ein Erzeugendensystem: Widerspruch.
15. (c) Sei v_i der Vektor, der an den ersten i Stellen eine Eins und sonst Nullen hat.
16. $\dim(K^{m \times n}) = mn$. Eine Basis besteht aus den Matrizen, die an genau einer Stelle eine Eins und überall sonst eine Null haben.

17. (a) Die „einfachste“ Basis besteht aus allen 1-elementigen Teilmengen von X , d. h. $B = \{\{x\} \mid x \in X\}$.
- (b) Sei $X = \{x_1, x_2, \dots, x_n\}$. Dann ist $B = \{\{x_1, x_2\}, \{x_2, x_3\}, \dots, \{x_{n-1}, x_n\}, \{x_n, x_1\}\}$ eine Basis von V .
- (c) $\dim(V) = n$.
18. (a) Der Trick besteht darin, sich zu überlegen, dass die symmetrische Differenz zweier Teilmengen mit einer geraden Zahl von Elementen wieder eine Teilmenge mit einer geraden Zahl von Elementen ist.
- (b) Es gibt eine Teilmenge Z gerader Mächtigkeit von X , so dass die symmetrische Differenz von Y und Z eine einelementige Menge ist. Mit dieser kann man mit Hilfe anderer Elemente aus W alle 1-elementigen Teilmengen von X bilden. Also erzeugt Y zusammen mit den Elementen aus W ganz V .
Daher muss W eine Hyperebene von V sein.
- (c) Nach (b) ist W ein $(n-1)$ -dimensionaler Vektorraum über $\text{GF}(2)$; somit hat W genau 2^{n-1} Elemente. Da V ein n -dimensionaler Vektorraum ist, hat V genau 2^n Elemente. Daher liegen außerhalb von W ebenfalls $2^n - 2^{n-1} = 2^{n-1}$ Elemente.
Das bedeutet, dass eine endliche Menge genau so viele Teilmengen gerader wie ungerader Mächtigkeit hat, nämlich jeweils genau 2^{n-1} viele.
19. Wir müssen für U die Vektorraumaxiome nachweisen. Die zweite bzw. dritte Bedingung sagt, dass die Skalarmultiplikation bzw. die Addition von Vektoren auf U abgeschlossen ist. Aus der ersten und dritten Bedingung folgt, dass der Nullvektor in U liegt. Alle anderen Gesetze ergeben sich aus der Tatsache, dass U eine Teilmenge des Vektorraums V ist.
20. Aufgabe 13 und Aufgabe 19.
21. (a) Wenden Sie das Unterraumkriterium an.
- (b) Zum Beispiel ist $B = \{(1, 0, -1), (0, 1, -1)\}$ eine Basis von U .
- (c) Man kann B zum Beispiel durch jeden der Einheitsvektoren zu einer Basis von V ergänzen.
22. (a) Vektorraumaxiome nachrechnen.
- (b) Unterraumkriterium.
23. O. B. d. A. sei C ein minimales Erzeugendensystem. Dann ist C eine Basis. Also enthält C genau n Vektoren. Also hat jedes Erzeugendensystem mindestens n Vektoren. Ferner muss ein Erzeugendensystem aus n Vektoren minimal, also eine Basis sein.
24. (a) Wenden sie das Unterraumkriterium an.
- (b) Es gilt $U+U' = \langle U, U' \rangle$. Denn $U+U'$ ist ein Unterraum, der U und U' enthält und in $\langle U, U' \rangle$ enthalten ist.
25. (a) gilt nur dann, wenn $U = \{o\} = U'$ ist. (b) nein. (c) ja.
26. $B \cup B'$ ist ein Erzeugendensystem, denn $\langle B \cup B' \rangle = \langle U, U' \rangle = V$. Bleibt zu zeigen, dass $B \cup B'$ linear unabhängig ist: Sei v eine Linearkombination aus B und v' eine Linearkombination aus B' mit $v + v' = o$. Dann ist $v = -v'$ ein Vektor aus $U \cap U'$, also $v = v' = o$. Der Rest folgt aus der Tatsache, dass B und B' Basen sind.
27. Der Nullraum $\{o\}$ und der gesamte Vektorraum haben nur ein Komplement.

28. Wenden Sie das Unterraumkriterium an.
29. (a) Die mengentheoretische Vereinigung $U \cup U'$ ist nur dann ein Unterraum, wenn U in U' oder U' in U enthalten ist.
(b) Siehe (a).
30. Da der Unterraum U den Nullvektor enthält, kann das Komplement den Nullvektor nicht enthalten. Also ist das Komplement von U kein Unterraum.
31. Führen Sie die geforderten Eigenschaften von V/U auf die entsprechenden Eigenschaften von V zurück.
32. Die notwendige und hinreichende Bedingung lautet $st - 1 \neq 0$.
33. (a) Der Ansatz $a \cdot c_1 + b \cdot c_2 + c \cdot c_3 = o$ liefert schnell $a = b = c = 0$. Also sind die Vektoren c_1, c_2, c_3 linear unabhängig, somit bilden sie eine Basis.
(b) Im Fall $K = \text{GF}(2)$ ist $c_1 = c_3$, also bilden die Vektoren keine Basis, während im Fall $K = \text{GF}(3)$ die Vektoren $c_1 = v_2 + 2v_3$, $c_2 = v_1 + v_2 + v_3$ und $c_3 = v_1 + v_2 + 2v_3$ eine Basis bilden.
34. (a) Man rechnet nach, dass $B := \{(1, 2, 3, 4), (2, 0, 1, -1), (-1, 0, 0, 1), (0, 2, 3, 0)\}$ linear unabhängig ist.
(b) $(0, 4, 5, 9)$ kann durch den ersten oder den dritten Vektor ersetzt werden ...
35. Offenbar sind die Vektoren $(1, 2, t+2)$ und $(0, t, 1)$ linear unabhängig. Also hat das Erzeugnis mindestens die Dimension 2. Genau dann ist $(-1, t+1, t)$ eine Linearkombination von $(1, 2, t+2)$ und $(0, t, 1)$, wenn $t = 1$ oder $t = -3/2$ gilt. In diesen Fällen ist die Dimension des Erzeugnisses also 2. in allen anderen Fällen 3.
36. (a) Unterraumkriterium.
(b) Da die Folgenglieder a_3, a_4, a_5, \dots durch Vorgabe von a_1 und a_2 bestimmt sind, gilt $\dim(V) = 2$.
(c) $B = \{(1, 1, 2, 3, 5, 8, \dots), (0, 1, 1, 2, 3, 5, \dots)\}$ ist eine Basis von V .
Bemerkung: Die erste Folge ist die Folge der **Fibonacci-Zahlen**.

Kapitel 4

1. $(2 \cdot 2 + 0 \cdot 0 + 1 \cdot 1 + 4 \cdot 4) = (21),$

$$\begin{pmatrix} 4 & 0 & 2 & 8 \\ 0 & 0 & 0 & 0 \\ 2 & 0 & 1 & 4 \\ 8 & 0 & 4 & 16 \end{pmatrix}.$$

2. Man erhält das Element c_{ij} der Matrix $A \cdot E_n$, indem man die i -te Zeile von A mit der j -ten Spalte von E_n multipliziert. Da in der j -ten Spalte von E_n an der j -ten Stelle eine Eins und an den anderen Stellen Nullen stehen, ist das Ergebnis $a_{ij} \cdot 1 = a_{ij}$.
3. In allen Fällen ist die i -te Komponente gleich $\sum_j a_{ij} \cdot kx_j = k \cdot \sum a_{ij}x_j$.
4. Das ist zwar mathematisch genau gleich schwer, aber technisch komplizierter. Sie müssen also sorgfältig arbeiten (leserlich schreiben usw.).

5. Die geraden Potenzen sind gleich der Einheitsmatrix.
6. Die Ränge sind 2, 3, 4.
7. Die erste Spalte von $A \cdot B$ ist gleich $s_1 \cdot b_{11} + s_2 \cdot b_{21} + \dots + s_n \cdot b_{n1}$, wobei s_1, s_2, \dots, s_n die Spalten von A sind.
8. $\text{Rang}(A) = 2$, $\text{Rang}(B) = 3$, $\text{Rang}(AB) = 2$. (Aufgrund der vorigen Aufgabe ist $\text{Rang}(AB) \leq 2$, und man sieht unmittelbar, dass AB nicht den Rang 0 oder 1 hat.)
9. Die Matrix hat den Rang n .
10. Eine Menge linear unabhängiger Zeilen bleibt linear unabhängig, wenn man zwei Zeilen vertauscht (Typ 1), wenn man eine Zeile mit $a \neq 0$ multipliziert (Typ 2) oder wenn man das a -fache einer Zeile zu einer anderen addiert ($a \neq 0$).
11. Da M den Rang n hat, gibt es ein von Null verschiedenes Element in der ersten Spalte. Dann kann man allein durch elementare *Zeilen*vertauschungen erreichen, dass die erste Spalte oben eine 1 und sonst nur Nullen enthält. Für die anderen Spalten argumentiert man entsprechend.
12. Wenn (k_1, k_2, \dots, k_n) eine Lösung für die Zeilen z_1 und z_2 ist, dann ist (k_1, k_2, \dots, k_n) auch eine Lösung für az_1 und z_2 . Für $a \neq 0$ gilt auch die Umkehrung.
13. Die Lösungen bestehen aus allen reellen Vielfachen von $(22, 14, -19, 10)$.
14. Zwar hat das Gleichungssystem aus den ersten drei Zeilen die unendlich viele Lösungen $(k, -2k, k + 1/3)$. Jedoch: Wenn man diese in die letzte Gleichung einsetzt, erhält man einen Widerspruch!
15. Die Koeffizientenmatrix hat den Rang 2. Im Fall (a) gibt es eine Lösung, im Fall (b) nicht.
16. Das Gleichungssystem ist genau dann lösbar, wenn $a = 3b+2$ ist.
17. Existenz einer Lösung: Zum einen gilt: $\text{Rang}(A \ b) \leq n$, da A nur n Zeilen hat. Zum anderen ist $\text{Rang}(A \ b) \geq \text{Rang}(A) = n$. Also ist $\text{Rang}(A \ b) = \text{Rang}(A)$, und damit ist das Gleichungssystem lösbar.
18. Sei $U = v + W$, wobei W ein Unterraum des Vektorraums V ist. Dann ist $L(U) = W$.
19. Sei $v_i = v + w_i \in U = v + W$, wobei W ein Unterraum von V ist.
Dann gilt für jede affine Linearkombination: $k_1 v_1 + k_2 v_2 + \dots + k_n v_n = k_1(v+w_1) + k_2(v+w_2) + \dots + k_n(v+w_n) = (k_1+k_2+\dots+k_n)v + k_1 w_1 + k_2 w_2 + \dots + k_n w_n = 1 \cdot v + k_1 w_1 + k_2 w_2 + \dots + k_n w_n \in v + W = U$.
20. Dies kann man entweder durch direkten Vergleich von je zwei Codewörtern nachweisen, oder man kann feststellen, dass der Code linear ist, und sehen, dass sein Minimalgewicht 3 ist.
21. Die Matrix H hat höchstens den Rang r , da sie nur r Zeilen hat. Andererseits hat sie mindestens den Rang r , da sie als Spalten die r Einheitsvektoren der Länge r enthält.
22. (a) Wenn nur s Fehler passieren, kann kein Codewort in ein anderes Codewort transformiert werden. Der Empfänger, der die Strategie hat, nur Codeworte zu akzeptieren, nimmt daher nie ein „falsches“ Codewort an.
(b) Wenn der Code s -fehlererkennend ist, dann muss sein Minimalabstand größer als s sein, denn sonst könnte durch höchstens s Fehler ein Codewort in ein verschiedenes überführt werden. Wenn umgekehrt $d(C) \geq s+1$ ist, dann können keine

zwei verschiedene Codewörter durch s oder weniger Fehler ineinander überführt werden.

23. C ist ein Unterraum (Unterraumkriterium!), also ein linearer Code. Er enthält nach Definition keine Vektoren vom Gewicht 1, aber zum Beispiel den Vektor $(1, 1, 0, \dots, 0)$. Also hat C Minimalgewicht 2 und ist daher 1-fehlererkennend.
24. Entweder können Sie für jedes angegebene Codewort c nachweisen, dass $c \cdot H^T = 0$ gilt, oder Sie können alle Vektoren v mit $v \cdot H^T = 0$ berechnen und diese und mit der angegebenen Liste vergleichen.

Kapitel 5

1. Wenn $w_0 = o$ ist, ist f die Nullabbildung, also eine lineare Abbildung. Wenn $w_0 \neq o$ ist. Dann ist zum Beispiel $f(2v) = f(v + v) = f(v) + f(v) = w_0 + w_0 \neq w_0 (= f(2v))$.
2. Nein, nein, ja.
3. (a) $f(v_1 + v_2 + v_3) = f((v_1 + v_2) + v_3) = f(v_1 + v_2) + f(v_3) = f(v_1) + f(v_2) + f(v_3)$.
4. Sei $k_1 w_1 + \dots + k_n w_n = o$. Dann ist $o = k_1 w_1 + \dots + k_n w_n = k_1 f(v_1) + \dots + k_n f(v_n) = f(k_1 v_1 + \dots + k_n v_n)$. Da f injektiv ist, muss wegen $f(o) = o$ auch $k_1 v_1 + \dots + k_n v_n = o$ sein. Da $\{v_1, \dots, v_n\}$ eine Basis ist, folgt $k_1 = \dots = k_n = 0$.
5. Sei $w \in W$ beliebig. Da f surjektiv ist, gibt es ein $v \in V$ mit $w = f(v)$. Sei $v = k_1 v_1 + \dots + k_n v_n$. Dann ist $w = f(v) = f(k_1 v_1 + \dots + k_n v_n) = k_1 w_1 + \dots + k_n w_n$. Also ist $\{w_1, \dots, w_n\}$ ein Erzeugendensystem von W .
6. Die Elemente der Basis $\{(1, 0), (0, 1)\}$ können auf beliebige Elemente von $\text{GF}(2)^2$ abgebildet werden. Da $\text{GF}(2)^2$ genau 4 Elemente besitzt, gibt es genau $4 \cdot 4 = 16$ lineare Abbildungen von $\text{GF}(2)^2$ nach $\text{GF}(2)^2$.
7. Es gilt $(f + g)' = f' + g'$ und $(k \cdot f)' = k \cdot f'$.
8. Dass f linear ist, ist klar bzw. folgt durch direktes Nachrechnen. $\text{Kern}(f) = \langle (1, -1, 1) \rangle$, $\text{Bild}(f) = \langle (1, 0, -1), (2, 1, 3) \rangle$.
9. $\text{Kern}(f) = \langle (2, -1, 1) \rangle$. $\text{Bild}(f)$ wird von je zwei Bildern der Basisvektoren aufgespannt.
10. (a) Nachrechnen.
(b) Im ersten Fall ist die Darstellungsmatrix gleich

$${}_B M_C(f) = \begin{pmatrix} 1 & -1 & 1 \\ 0 & -6 & 12 \\ -2 & 2 & -2 \end{pmatrix}.$$

Im zweiten Fall werden die Basisvektoren auf $(0, 12, 0)$, $(-2, 0, 4)$, $(2, 48, -4)$ abgebildet.

11. (a)

$${}_B M_B(f) = \begin{pmatrix} -11 & -84 & -36 \\ 4 & 29 & 12 \\ -6 & -42 & -17 \end{pmatrix}.$$

12. Durch Anwenden der Matrix A auf die Einheitsvektoren von K^m erhält man ein Erzeugendensystem von $\text{Bild}(f)$. Da dies gleichzeitig ein Erzeugendensystem für den von den Spalten von A aufgespannten Unterraum von K^m ist, folgt die Behauptung.
13. Es gibt zwei kritische Werte von a . Falls $a = 1$ ist, ist $\text{Bild}(f)$ der 1-dimensionale Unterraum $\langle v_1 + v_2 + \dots + v_n \rangle$. Im Fall $a = n-1$ ist $\text{Bild}(f) = \{k_1 v_1 + k_2 v_2 + \dots + k_n v_n \mid k_1 + k_2 + \dots + k_n = 0\}$ und $\dim(\text{Bild}(f)) = n-1$. In allen anderen Fällen ist $\text{Bild}(f) = V$.
14. Sei $\{v_1, v_2, \dots\}$ eine Basis von V . Dann wird durch $f(v_1) := v_1$ und $f(v_i) := 0$ für $i \geq 2$ eine lineare Abbildung definiert, die weder injektiv noch surjektiv ist.
15. Der Nullraum besitzt nur eine lineare Abbildung in sich, diese ist bijektiv. Jeder Vektorraum einer Dimension ≥ 2 besitzt mindestens eine nichtidentische bijektive Abbildung; man kann z. B. die Elemente einer Basis permutieren. Betrachten wir schließlich Vektorräume der Dimension 1. Wenn der Körper ein Element $a \neq 0, 1$ besitzt, kann man jeden Vektor $v \neq 0$ auf av abbilden. Dies ergibt eine nichtidentische lineare Abbildung. Mit anderen Worten: Nur der Nullraum und der 1-dimensionale Vektorraum über \mathbb{Z}_2 besitzt nur eine bijektive lineare Abbildung in sich.
16. Reflexivität: $S = \text{id}$. Symmetrie: $S = S^{-1}$. Transitivität: Seien $M' = S^{-1}MS$ und $M'' = T^{-1}M'T$. Dann gilt $M'' = T^{-1}M'T = T^{-1}(S^{-1}MS)T = (T^{-1}S^{-1})M(ST) = (ST)^{-1}M(ST)$.
17. Man rechnet leicht nach, dass f eine lineare Abbildung ist. Da offenbar $\text{Bild}(f) = K$ ist, also die Dimension 1 hat, gilt $\dim(\text{Kern}(f)) = 9 - 1 = 8$. Man erhält eine Basis von $\text{Kern}(f)$, indem man zunächst die Diagonale gleich Null setzt, und jeweils einen der Einträge außerhalb 1 und alle anderen gleich Null wählt; das ergibt schon sechs Vektoren. Die beiden anderen erhält man, indem man alle Elemente gleich Null setzt – bis auf a_{11} und a_{22} bzw. a_{11} und a_{33} , von denen man jeweils einen Wert 1 und den anderen -1 setzt.
18. Man rechnet einfach nach, dass f eine lineare Abbildung ist. Da offenbar $\text{Bild}(f) = K$ ist, also die Dimension 1 hat, besitzt $\text{Kern}(f)$ die Dimension $4 - 1 = 3$. Die folgenden Vektoren bilden eine Basis von $\text{Kern}(f)$:

$$\begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

19. (a) Nachrechnen.
- (b) Offenbar ist $\text{Bild}(f) = K^3$, denn schon, indem man $a_{22}, a_{33}, a_{23}, a_{31}, a_{21}$ und a_{32} Null setzt, kann man jeden Vektor aus K^3 erreichen. Also ist $\dim(\text{Bild}(f)) = 3$, und somit $\dim(\text{Kern}(f)) = 9 - 3 = 6$.

(c) Die folgenden Vektoren bilden eine Basis von $\text{Kern}(f)$:

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

20. Man kann durch Anwenden der Definition einer Linearform alle geforderten Eigenschaften auf die entsprechenden Eigenschaften von K zurückführen.

Kapitel 6

1. Man führt die nachzuweisenden Eigenschaften von \mathbf{Z}_n auf die entsprechenden Eigenschaften von \mathbf{Z} zurück.
2. Benennen Sie die Elemente von A , B und C und rechnen Sie jeweils die linke und rechte Seite aus.
3. Die Gesetze der Addition sind schon längst bekannt, da $n \times n$ -Matrizen über K einen Vektorraum bilden. Die (wenigen) Gesetze der Multiplikation und die Distributivgesetze muss man eben nachrechnen.
4. Für fast alle Matrizen A, B aus $K^{n \times n}$ mit $n \geq 2$ gilt $AB \neq BA$.
5. Die Summe zweier invertierbarer Matrizen kann eine nichtinvertierbare Matrix sein, die nicht gleich der Nullmatrix ist. Beispiel:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}.$$

6. Die Elemente an der Stelle (i, j) der Matrizen $(k \cdot E_n) \cdot A$ und $A \cdot (k \cdot E_n)$ sind $k \cdot a_{ij}$ bzw. $a_{ij} \cdot k$.
7. (a) Der Beweis funktioniert wörtlich wie bei $K[x]$.
(b) Das Einselement von R ist auch das Einselement von $R[x]$.
8. Formal gilt $K[x, y, z] = ((K[x])[y])[z]$ (der Polynomring in der Unbestimmten z , dessen Koeffizienten Elemente von $K[x, y]$ sind). Wenden Sie zweimal Übungsaufgabe 7 an.
9. $q = x^4 - x + 1$, $r = x^2$.
10. Vergleiche vorige Aufgabe.
11. $q = 2x^4 + 2x^3 + x + 1$, $r = 2x + 1$.
12. $x^3 - 19x + 30 = (x - 2)(x - 3)(x + 5)$.
13. $x^3 - 2x + 4 = (x + 2)(x^2 - 2x + 2)$. Da $x^2 - 2x + 2 = (x - 2)^2 + 1 > 0$ ist für alle $x \in \mathbf{R}$, zerfällt das Polynom weder über \mathbf{Q} noch über \mathbf{R} .
14. $x^4 + 1 = (x + 1)^4$, $x^4 + x^3 + x + 1 = (x + 1)^2(x^2 + x + 1)$, $x^5 + 1 = (x + 1)(x^4 + x^3 + x^2 + x + 1)$.

15. (a) Wir wissen, dass dann $f = a(x - k_1) \dots (x - k_n)$ ist (a unbekannt). Wenn f normiert ist, ist $a = 1$, und man erhält die Koeffizienten durch einfaches Ausmultiplizieren.
 (b) Schlagen Sie in der Literatur (z. B. in Büchern über Numerische Mathematik) die Interpolationsformel von Lagrange nach. Meist wird diese nur über \mathbf{R} bewiesen. Machen Sie sich klar, dass sie über jedem Körper funktioniert.
16. Wenn man b in $(x - a)^v$ einsetzt, erhält man $(b - a)^v$. Dieses Produkt ist nur dann gleich Null, wenn einer der Faktoren Null ist, also wenn $b - a = 0$ gilt.
17. Angenommen, es gäbe zwei normierte Polynome g und h , die I erzeugen. Dann müssen diese den gleichen Grad haben (da g ein Vielfaches von h und h ein Vielfaches von g sein muss). Dann ist $g - h$ ein Polynom kleineren Grades als g , das in I liegt. Also ...
18. (a) Da f nur endlich viele Nullstellen hat, muss K endlich sein.
 (b) Sei $K = \{0, 1, k_3, \dots, k_n\}$ endlich. Dann ist $f = x(x - 1)(x - k_3) \dots (x - k_n)$ ein Polynom, das alle Elemente von K als Nullstellen hat.
19. Es ist nur die magnetische Anziehungseigenschaft zu zeigen. Für jedes Polynom g aus $K[x]$ und jedes Polynom f aus U ist zu zeigen, dass gf in U liegt. Aus der Voraussetzung folgt, dass für jede Potenz x^k das Polynom $x^k \cdot f$ in U liegt. Da U ein Unterraum ist, ist dann auch $g = b_0 + b_1x + b_2x^2 + \dots + b_mx^m$ in U enthalten.
20. (a) $(r + I) + (r' + I) := r + r' + I$, $s \cdot (r + I) := s \cdot r + I$.
21. Sei f ein Polynom aus $K[x]$. Zeigen Sie:
 (a) folgt direkt aus der vorigen Aufgabe.
 (b) Wenn f irreduzibel ist, dann gibt es zu jedem Polynom g mit $0 \leq \text{Grad}(g) < \text{Grad}(f)$ Polynome f' und g' mit $1 = ff' + gg'$. Dann ist $g' + I(f)$ eine multiplikative Inverse von $g + I(f)$ in $K[x]/I(f)$. Also ist $K[x]/I(f)$ ein Körper.
 (c) Wenn f ein über $\text{GF}(p)$ irreduzibles Polynom vom Grad n ist, ist $\text{GF}(p)[x]/I(f)$ ein Körper, dessen Elemente durch alle Polynome über $\text{GF}(p)$ vom Grad $< n$ repräsentiert werden. Also hat dieser Körper genau p^n Elemente.
22. (a) Nachrechnen der definierenden Eigenschaften eines Ideals.
 (b) $x^2 + 1, x^2 - 2x + 2$.
23. Nachrechnen der definierenden Eigenschaften eines Ideals.
24. Man wendet schrittweise Polynomdivision an. Sei $f = qg + r$ mit $\text{Grad}(r) < \text{Grad}(g)$. Dann ist die Menge der Teiler von f und g gleich der Menge der Teiler von g und r . Somit ist jeder ggT von f und g auch ein ggT von g und r und umgekehrt. Damit kann man die Berechnung der ggT auf Polynome immer kleineren Grades zurückführen.
25. $I(f, g) \subseteq I(h)$, da jedes Polynom der Form $rf + sg$ ($r, s \in K[x]$) ist auch ein Vielfaches von h , da f und g Vielfache von h sind.
 $I(h) \subseteq I(f, g)$, denn h kann in der Form $rf + sg$ geschrieben werden (Vielfachsummen-darstellung). Also ist jedes Vielfache von h ein Element der Form $r'f + s'g$.
26. $x^7 + x^5 + x^3 + 1 = (x^3 + x + 1)(x^4 + x) + x + 1$. Da der ggT von $x^3 + x + 1$ und $x + 1$ gleich 1, erzeugen die Polynome $x^7 + x^5 + x^3 + 1$ und $x^3 + x + 1$ nach Aufgabe 25 ganz $\text{GF}(2)[x]$.
27. Das Polynom $x + 1$ ist ein größter gemeinsamer Teiler von $x^2 - 1$, $x^3 + 2x^2 + 2x + 1$ und $x^4 + x^3 + x + 1$.
28. Die Aussage folgt direkt aus Aufgabe 25.

Kapitel 7

1. $(1\ 2\ 3), (1\ 3\ 2), (1\ 2\ 3\ 4), (1\ 2\ 4\ 3), (1\ 3\ 2\ 4), (1\ 3\ 4\ 2), (1\ 4\ 2\ 3), (1\ 4\ 3\ 2)$.
Allgemein: Man kann o. B. d. A. das Element 1 an die erste Stelle setzen und die restlichen Elemente beliebig permutieren.
2. Wenn (i, j) ein Fehlstand von π ist, dann ist $(\pi(j), \pi(i))$ ein Fehlstand von π^{-1} .
3. 0, 21, 108.
4. 28, -80. Die Determinante der dritten Matrix kann man rekursiv berechnen: Sei d_n der Wert der Determinante bei einer $n \times n$ Matrix. Dann ist $d_1 = 1$, $d_2 = 1 - a^2$ und $d_n = d_{n-1} - a^2 \cdot d_{n-2}$ für $n > 2$.
5. Wenn man zunächst die erste Zeile von allen anderen Zeilen subtrahiert und dann die 2-te, 3-te, ..., n -te Spalte zur ersten addiert, kommt man leicht auf die Determinante $(a-b)^{n-1}(a + (n-1)b)$. [Vergleichen Sie auch Aufgabe 11 aus Kap. 4.]
6. Die Determinante einer Permutationsmatrix ist 1 oder -1. Beweis durch Induktion nach n .
7. Tun Sie das!
8. Die Homogenität ergibt sich genau so wie die Additivität, ist aber einfacher nachzuweisen.
9. Sie können sich dabei Schritt für Schritt am Beweis für die Entwicklung nach der ersten Zeile entlanghangeln.
10. Der Eintrag an der Stelle (i, j) von $(AB)^T$ ist $\sum_{k=1}^n a_{jk} b_{ki} = \sum_{k=1}^n b_{ki} a_{jk}$. Dieser ist gleich dem Eintrag an der Stelle (i, j) von $B^T A^T$.
11. Es ist klar, dass die Transponierte der Summe von Matrizen gleich der Summe der Transponierten ist (Additivität). Da auch die Homogenität gilt, ist das Transponieren von Matrizen aus $K^n \times^n$ eine lineare Abbildung.
12. Multiplikationssatz.
13. Man rechnet nach, dass für diesen Kandidaten für A^{-1} gilt: $AA^{-1} = E$ und $A^{-1}A = E$.

Kapitel 8

1. Sei $B = \{v_1, v_2, \dots, v_n\}$. Dann gilt $\text{Bild}(f) = \{f(v_1), f(v_2), \dots, f(v_n)\}$. Sei o. B. d. A. $w_1 = f(v_1), \dots, w_s = f(v_s)$ eine Basis von $\text{Bild}(f)$. Man ergänze diese zu einer Basis $B' = \{w_1, \dots, w_s, w_{s+1}, \dots, w_m\}$ von W . Dann hat ${}_B M(f)_{B'}$ die gewünschte Gestalt.
2. $Mx = kx \Leftrightarrow Mx - kx = 0 \Leftrightarrow (M - kE)x = 0$. Umgekehrt ist k genau dann ein Eigenwert von M , wenn $\det(M - kE) = 0$ gilt, d. h. wenn die Spaltenvektoren von $M - kE$ linear abhängig sind, wenn es also einen Vektor $x \neq 0$ gibt mit $(M - kE)x = 0$.
3. Die Eigenwerte sind $-1, 3 + \sqrt{15}, 3 - \sqrt{15}$.
4. Viele reelle Matrizen haben keine reellen Eigenwerte, z. B. die Matrix $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.
5. (a) Sei x ein Eigenvektor zum Eigenwert k von M . Sei $y := A(x)$. Da A regulär ist, ist $y \neq 0$. Damit folgt: $AMA^{-1}(y) = AM(x) = A(kx) = k \cdot A(x) = ky$. Also ist y ein Eigenvektor zum Eigenwert k von AMA^{-1} .

- (b) Basiswechsel entspricht Übergang zu einer ähnlichen Matrix. Ähnliche Matrizen haben nach (a) die gleichen Eigenwerte.
6. Das charakteristische Polynom ist $-(x-1)(x-2)^2$. Also hat M die Eigenwerte 1 und 2. Der Eigenraum zum Eigenwert 1 ist $\langle (1, 1, 2) \rangle$; der Eigenraum zum Eigenwert 2 ist $\langle (1, 2, 3) \rangle$.
7. $\chi_M = -(x-1)(x-2)(x+3)$, $\chi_M = -(x-1)^2(x-2) \dots$
8. (a) $\chi_M = -x^3 + cx^2 + bx + a$.
 (b) Wählen Sie die Werte für a, b, c so, dass χ_M keine Nullstelle in $\text{GF}(3)$ hat.
9. $\dots k^{-1}$.
10. (b) Mit Induktion folgt, dass M^t den Eigenwert k^t hat: Sei v ein Eigenvektor zum Eigenwert k . Dann ist $M^t(v) = M^{t-1}M(v) = M^{t-1}(k \cdot v) = k \cdot (M^{t-1}(v)) = k \cdot (k^{t-1} \cdot v) = \dots$
 (c) Sei $M = E_n$. Dann ist auch $M^2 = E_n$. Zwar hat M^2 den Eigenwert $1 = (-1)^2$, aber M hat nicht den Eigenwert -1 .
11. Sei v ein Eigenvektor zum Eigenwert k . Mit der vorigen Aufgabe folgt: $k^2 \cdot v = f^2(v) = f(v) = k \cdot v$. Also ist $k(k-1) \cdot v = 0$, und somit \dots
12. 1. Lösungsmöglichkeit: Sie „sehen“, dass die Vektoren $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ und $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ Eigenvektoren zu den verschiedenen Eigenwerten $a+b$ und $a-b$ sind.
 2. Lösungsmöglichkeit: Sie berechnen das charakteristische Polynom: $x^2 - 2ax + a^2 - b^2 = (x - (a+b))(x - (a-b))$, das in verschiedene Nullstellen zerfällt.
 Achtung! Bei Körpern der Charakteristik 2 sind die Eigenwerte bzw. die Linearfaktoren nicht verschieden!
13. $\chi_M = -(x-1)(x-2)^2$, $\mu_M = (x-1)(x-2)$. Also ist M diagonalisierbar.
14. Suchen Sie (zum Beispiel) eine 2×2 -Matrix, deren charakteristisches Polynom $x^2 - 2$ ist.
15. Der gesuchte Eigenwert ist die konstante Zeilensumme, ein zugehöriger Eigenvektor der Vektor, der nur aus Einsen besteht.
16. Eigenvektoren zu den angegebenen Eigenwerten sind:

$$\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ i \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -i \\ -1 \\ i \end{pmatrix}.$$

17. Wegen $\dim(K^n \times^n) = n^2 =: m$ gilt $k_0 E + k_1 M + k_2 M^2 + \dots + k_m M^m = 0$ (nicht alle $k_i = 0$). Dann ist $f = k_0 + k_1 x + k_2 x^2 + \dots + k_m x^m$ ein Polynom, das nicht das Nullpolynom ist, die Matrix M als Nullstelle hat und einen Grad $\leq m = n^2$ hat.
18. Aus $f^2 = 0$ folgt $\chi_f = x^2$. Also ist $\mu_f = x$ oder $\mu_f = x^2$. Im ersten Fall ist $f = 0$, im zweiten Fall zeigt man, dass es Darstellungsmatrix in der angegebenen Form gibt.
19. Entwicklung nach der ersten Spalte und Induktion.

$$20. \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

21. $x^2 - k_1x - k_0, -x^3 + k_2x^2 + k_1x + k_1, x^4 - k_3x^3 - k_2x^2 + k_1x + k_0, \dots$
22. $\mu_f = (x-1)(x-2)^2; \mu_M = (x-1)(x-2)(x+3), \mu_M = -(x-1)(x-2).$
23. Da das charakteristische Polynom gleich $\pm x^n$ ist, gilt $\mu_f = x^k$ mit $k \leq n$. Wenn man die Potenzen von M bildet, rutscht die Diagonale aus Einsen jeweils um eine Stelle weiter nach links unten. Also ist $\mu_f = x^n$.
24. Nachrechnen und Anwenden des Satzes über Ideale eines Polynomrings.

Kapitel 9

- Ein Ideal ist eine additive Untergruppe eines Rings, die die magnetische Anziehungseigenschaft hat.
- Sei U eine Untergruppe von $(\mathbf{Z}, +)$. Wenn $U = \{0\}$ ist, dann ist $U = 0\mathbf{Z}$. Sei also $U \neq \{0\}$. Dann liegen in U auch positive Zahlen. Sei n die kleinste positive Zahl in U . Dann ist natürlich $n\mathbf{Z} \subseteq U$. Umgekehrt kann man leicht zeigen, dass jedes Element aus U ein ganzzahliges Vielfaches von n ist. D.h. $U \subseteq n\mathbf{Z}$.
- Untergruppenkriterium: Offenbar ist \mathbf{Z}_n^* nicht leer, da zumindest die Zahl 1 in \mathbf{Z}_n^* liegt. Dass mit jedem Element auch das Inverse in \mathbf{Z}_n^* liegt, folgt aus der Definition von \mathbf{Z}_n^* . Seien schließlich $a, b \in \mathbf{Z}_n^*$. Da ab das Inverse $b^{-1}a^{-1} \in \mathbf{Z}_n^*$ hat, liegt auch das Produkt in \mathbf{Z}_n^* .
- Die Symmetrien eines Rechtecks, das kein Quadrat ist, bestehen neben der Identität aus zwei Spiegelungen an den Symmetrieachsen und der Punktspiegelung am Mittelpunkt des Rechtecks. Die Symmetrien bilden eine Gruppe, da jedes Element sein eigenes Inverses ist und die Operation abgeschlossen ist: Das Produkt der beiden Spiegelungen ist die Punktspiegelung usw.
- Neben der Identität sind dies (a) die Drehungen um $90^\circ, 180^\circ, 270^\circ$ um die Gerade, die die Mittelpunkte der quadratischen Seitenflächen verbindet, (b) Drehungen um 180° um die Geraden, die die Mittelpunkte der nichtquadratischen Seiten verbinden, (c) Drehungen um 180° um die Gerade, die die Mittelpunkte gegenüberliegender Kanten der Länge b verbindet. Wenn man auch Spiegelungen zulässt, kommen weitere acht Abbildungen hinzu.
- Neben der Identität besteht die Symmetriegruppe (a) aus den Drehungen um $\pm 120^\circ$ um eine Achse, die eine Ecke mit dem Mittelpunkt der gegenüberliegenden Seiten verbindet, (b) um Drehungen um 180° um eine Achse, die die Mittelpunkte gegenüberliegender Kanten verbindet. Wenn man auch Spiegelungen zulässt (die man an keinem Tetraeder real ausführen kann), dann ergeben sich (c) noch weitere 12 Symmetrien.
- 4, 8 (16), 12 (24).
- (a) Man bezeichnet zunächst die Flächen irgendwie, z. B. mit a, b, c, d, e, f (a unten, f oben, b, c, d, e im Uhrzeigersinn).

Eine Drehung um 90° um eine Gerade, die zwei gegenüberliegende Seitenmitten verbindet, ist z. B. die Permutation $(a)(f)(b\ c\ d\ e)$.

Eine Drehung um 180° um eine Gerade, die zwei gegenüberliegende Kantenmitten verbindet, ist z. B. die Permutation $(a\ f)(b\ e)(c\ d)$.

Eine Drehung um 120° um eine Raumdiagonale ist z. B. die Permutation $(a\ b\ e)(c\ f\ d)$.

9. (a) Satz von Lagrange.
- (b) Die Untergruppen einer Ordnung $k \leq 4$ erhält man als Erzeugnis eines k -Zyklus. Die Menge der Elemente, die 4 festlassen, ist eine Untergruppe der Ordnung 6 (isomorph zu S_3). Die A_4 ist eine Untergruppe der Ordnung 12, und S_4 ist die Untergruppe der Ordnung 24. Eine Untergruppe der Ordnung 8 besteht neben der Identität aus den Elementen $(1\ 2\ 3\ 4)$, $(1\ 3)(2\ 4)$, $(1\ 4\ 3\ 2)$, $(1\ 3)$, $(2\ 4)$, $(1\ 2)(3\ 4)$, $(1\ 4)(3\ 2)$.
10. 1 und -1 sind die einzigen Elemente, die \mathbf{Z} erzeugen.
11. \mathbf{Z}_7^* wird von den Elementen 3 und 5 erzeugt; \mathbf{Z}_8^* ist eine Gruppe der Ordnung 7, sie wird also von allen Elementen $\neq 1$ erzeugt.
12. Sei a ein Element von \mathbf{Z}_n , das ein multiplikatives Inverses a' hat. Das bedeutet $aa' \equiv 1 \pmod{n}$, also $aa' = 1 + kn$. Daraus sieht man, dass jeder gemeinsame Teiler von a und n auch 1 teilt. Also ist $\text{ggT}(a, n) = 1$. Sei umgekehrt $\text{ggT}(a, n) = 1$. Dann gibt es nach dem Lemma von Bézout ganze Zahlen a' und n' mit $1 = aa' + nn'$, also $aa' = 1 - nn'$. Es folgt, dass a' das multiplikativ Inverse von a ist.
13. $|\mathbf{Z}_p^*| = p - 1$, $|\mathbf{Z}_{p^2}^*| = p(p - 1)$
14. Beim Erzeugnis von Elementen einer Gruppe werden diese Elemente in beliebiger Weise mit Hilfe der Gruppenoperation verknüpft. Beim Erzeugnis von Vektoren werden diese in beliebiger Weise mit Hilfe der Vektoraddition verknüpft – sie können aber zusätzlich mit Skalaren multipliziert werden.
15. Sei G eine Gruppe gerader Ordnung. Es ist zu zeigen, dass es ein Element $g \neq e$ gibt, das zu sich selbst invers ist. (Denn aus $g^{-1} = g$ folgt $e = gg^{-1} = gg = g^2$.) Angenommen, es gäbe kein solches Element. Dann könnte man die Elemente $\neq e$ von G in Paare $\{g, g^{-1}\}$ disjunkt aufteilen. Das geht aber höchstens dann, wenn $|G| - 1$ gerade ist. Dies würde aber bedeuten, dass $|G|$ ungerade ist: ein Widerspruch!
16. Sei G eine abelsche Gruppe.
 - (a) $(gh)^2 = (gh)(gh) = (gh)(hg) = g(hh)g = geg = gg = e$.
 - (b) Seien g und h zwei verschiedene Elemente der Ordnung 2. Dann ist die Menge $\{e, g, h, gh\}$ aufgrund des Untergruppenkriteriums eine Untergruppe von G , also gleich dem Erzeugnis von g und h .
 - (c) Sei x das erzeugende Element von G , d. h. $G = \{x, x^2, x^3, \dots, x^{2n} = e\}$ für eine natürliche Zahl n . Dann ist g^n das einzige Element der Ordnung 2.
17. Sei $g \in G$ beliebig. Wenn g in U liegt, dann ist natürlich $g^{-1}Ug = U$. Sei also $g \notin U$. Dann ist $gU \cup U = G$, da U den Index 2 hat. Wegen $Ug \cap U = \emptyset$ muss $Ug \subseteq gU$, also $g^{-1}Ug \subseteq U$ gelten. Daraus folgt, dass U ein Normalteiler ist.
18. Die Untergruppen von S_3 haben die Ordnungen 1, 2, 3, 6. Die nichttrivialen Untergruppen haben die Ordnungen 2 oder 3. Die Untergruppen der Ordnung 2 werden

- jeweils von einer Transposition erzeugt, während die Untergruppe der Ordnung 3 von einem (der beiden) 3-Zykel erzeugt wird.
19. Seien g und h beliebige Elemente von G . Aus $(hg)(hg) = e$ ergibt sich $hg = (hg)^{-1} = g^{-1}h^{-1}$. Mit $(gh)(gh) = e$ folgt also: $e = (gh)(gh) = g(hg)h = g(g^{-1}h^{-1})h = e$.
 20. Wir addieren einen von Null verschiedenen Vektor k mal: $v+v+\dots+v = k \cdot v$. Diese Summe ist nur dann Null, wenn $k = 0$ ist. Dies ist über $\text{GF}(p)$ zum ersten Mal bei $k = p$ Summanden der Fall.
 21. Wir fixieren eine Basis v_1, v_2, \dots, v_n von V . Dann liefert die Abbildung, die jeder linearen Abbildung f ihre Darstellungsmatrix ${}_B M_B(f)$ zuordnet, den gewünschten Isomorphismus.
 22. $|\text{GL}(n, q)| = (q^n - 1) \cdot \dots \cdot (q^n - q^{n-1})$, $|\text{GL}(2, q)| = (q^2 - 1) \cdot (q^2 - q) = q(q^2 - 1)(q - 1)$, $|\text{GL}(3, q)| = (q^3 - 1)(q^3 - q)(q^3 - q^2)$.
 23. Untergruppenkriterium. Eine Nebenklasse besteht genau aus den Matrizen mit fester Determinante $\neq 0$. Also ist die Anzahl der Nebenklassen gleich ...
 24. $\{k \cdot E_n \mid k \in \text{GF}(q)^*\}$.
 25. Untergruppenkriterium.
 26. $f(g^{-1})f(g) = f(g^{-1}g) = f(e) = e$. Also ist $f(g^{-1})$ invers zu $f(g)$, d. h. $f(g^{-1}) = f(g)^{-1}$.
 27. Der Isomorphismus ist die Abbildung $f: x \mapsto x+1$. Damit folgt $f(x^*y) = f(xy+x+y) = xy+x+y+1 = (x+1)(y+1) = f(x) \cdot f(y)$.
 28. Eine Richtung ist klar. Sei umgekehrt $g^{-1}Ng \subseteq N$ für alle $g \in G$. Zu zeigen: $N \subseteq g^{-1}Ng$ für alle $g \in G$. Wir betrachten ein beliebiges $n \in N$. Nach Voraussetzung ist $n' := gng^{-1} = (g^{-1})^{-1}ng^{-1} \in N$. Wieder nach Voraussetzung gilt dann $n = g^{-1}n'g \in N$.
 29. Führen Sie die Eigenschaften von G/N auf die entsprechenden Eigenschaften von G zurück!
 30. Offenbar gibt es nur eine Gruppe der Ordnung 1. Die Gruppen von Primzahlordnung sind zyklisch, also isomorph zu \mathbb{Z}_p . Es bleiben also die Ordnungen 4, 6, 8 zu betrachten. In einer Gruppe der Ordnung 4 gibt es nur Elemente der Ordnungen 1, 2, oder 4. Wenn es ein Element der Ordnung 4 gibt, ist die Gruppe zyklisch, also isomorph zu \mathbb{Z}_4 . Andernfalls hat jedes Element $\neq e$ die Ordnung 2. Damit ist die Gruppe isomorph zur Kleinschen Vierergruppe. Jede Gruppe der Ordnung 6 ist entweder zyklisch, also isomorph zu \mathbb{Z}_6 oder isomorph zu S_3 . Es gibt fünf Gruppen der Ordnung 8: die zyklische Gruppe (also \mathbb{Z}_8), die Gruppe $\mathbb{Z}_2 \times \mathbb{Z}_4$, die „Diedergruppe“ der Ordnung 8 (= Symmetriegruppe des Quadrats), die „Quaternionengruppe“ und die Gruppe $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$.
 31. Untergruppenkriterium.
 32. Die Abbildung f ist ein Gruppenhomomorphismus von K^* auf Q (siehe vorige Aufgabe). $\text{Bild}(f) = Q$, $\text{Kern}(f) = \{1, -1\}$. f ist nur dann ein Körperhomomorphismus, wenn gilt $f(a+b) = f(a) + f(b)$, d. h. $(a+b)^2 = a^2 + b^2$. Das ist nur der Fall in Körpern der Charakteristik 2.
 33. Eine Spiegelung hat die Ordnung 2, ebenso eine Punktspiegelung. Eine Drehung um den Winkel $360^\circ/n$ hat die Ordnung n .

Kapitel 10

1. Nachrechnen.
2. Eigenschaften durch Ausrechnen nachweisen.
3. Das liegt letztlich daran, dass die linearen Abbildungen f_v und g_w den Nullvektor auf den Nullvektor abbilden.
4. Nachrechnen.
5. Eine symmetrische Matrix ist genau dann positiv definit, wenn die Matrix aus den ersten k Elementen in den ersten k Zeilen eine positive Determinante hat ($1 \leq k \leq 4$). Ein Beispiel dafür ist die Matrix, die auf der Hauptdiagonale die Einträge 2 und sonst die Einträge 1 hat.
6. $\langle v, w+w' \rangle$

$$\begin{aligned} &= \left\langle \sum_{i=1}^n k_i v_i, \sum_{i=1}^n h_i v_i + \sum_{i=1}^n h'_i v_i \right\rangle = \sum_{i=1}^n k_i (h_i + h'_i) a_{ii} = \sum_{i=1}^n k_i h_i a_{ii} + \sum_{i=1}^n k_i h'_i a_{ii} \\ &= \langle v, w \rangle + \langle v, w' \rangle. \end{aligned}$$

7. Unterraumkriterium.
8. Jeder Vektor des Radikals muss zumindest auf den Einheitsvektoren senkrecht stehen. Daraus sieht man, dass nur der Nullvektor in diesem Radikal liegen kann.
9. Analog zu Nachweis der Eigenschaften von f_v .
10. Explizites Nachprüfen der Definition.
11. Unterraumkriterium. Bei einer symmetrischen $n \times n$ -Matrix können die Einträge auf der Diagonalen und in der rechten oberen Hälfte frei gewählt werden, alle anderen Einträge sind dadurch bestimmt. Also ist die Dimension dieses Vektorraums gleich $1+2+\dots+n = n(n+1)/2$.
12. Eine Richtung ist klar. Sei umgekehrt $\|f(v)\| = \|v\|$ für alle $v \in V$. Dann ist auch $\langle f(v+w), f(v+w) \rangle = \langle v+w, v+w \rangle$. Es folgt $\langle f(v), f(v) \rangle + \langle f(v), f(w) \rangle + \langle f(w), f(v) \rangle + \langle f(w), f(w) \rangle = \langle v, v \rangle + \langle v, w \rangle + \langle w, v \rangle + \langle w, w \rangle$, also $\langle v, v \rangle + 2\langle f(v), f(w) \rangle + \langle w, w \rangle = \langle v, v \rangle + 2\langle v, w \rangle + \langle w, w \rangle$, und somit $\langle f(v), f(w) \rangle = \langle v, w \rangle$.
13. Man muss zu s ($s \leq n$) orthonormalen Vektoren w_1, \dots, w_s einen zu diesen orthonormalen Vektor w_{s+1} finden. Dazu wählt man einen zu w_1, \dots, w_s linear unabhängigen Vektor v_{s+1} und setzt $w := v_{s+1} + \sum_{j=1}^s k_j w_j$. Man bestimmt die k_j durch Ausnutzen der Orthonormalitäten von w mit w_i und durch die Normiertheit.
14. Nein.
15. Die Menge $\{(\frac{1}{2}, \frac{1}{2}, \frac{1}{\sqrt{2}}), (\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0), (\frac{1}{2}, \frac{1}{2}, -\frac{1}{\sqrt{2}})\}$ ist eine Orthonormalbasis von \mathbf{R}^3 .
16. Die Menge $\{(1/2, 1/2, 1/2, 1/2), (-1/2, -1/2, 1/2, 1/2), (-1/2, 1/2, -1/2, 1/2), (-1/2, 1/2, 1/2, -1/2)\}$ ist eine Orthonormalbasis von \mathbf{R}^4 .
17. Untergruppenkriterium.
18. Benutzen Sie die Linearität in der ersten und zweiten Komponente.
19. Benutzen Sie die Linearität in der ersten und zweiten Komponente.
20. Nachrechnen.

Literatur

- [Art] Artmann, B.: Lineare Algebra. Birkhäuser Verlag, Basel (1991)
- [Beu1] Beutelspacher, A.: In Mathe war ich immer schlecht ... Verlag Vieweg, Braunschweig, Wiesbaden (2009)
- [Beu2] Beutelspacher, A.: „Das ist o.B.d.A. trivial!“ Tips und Tricks zur Formulierung mathematischer Gedanken, 9. Aufl. Verlag Vieweg, Braunschweig, Wiesbaden (2009)
- [BeuRo] Beutelspacher, A., Rosenbaum, U.: Projektive Geometrie, 2. Aufl. Verlag Vieweg, Braunschweig, Wiesbaden (2004)
- [Bries] Brieskorn, E.: Lineare Algebra und Analytische Geometrie I, II. Verlag Vieweg, Braunschweig, Wiesbaden (1983, 1985)
- [Cro] Crowe, M.J.: A History of Vector Analysis. Dover Publications, New York (1985)
- [Cox] Coxeter, H.S.M.: Introduction to Geometry, 2. Aufl. John Wiley & Sons, New York, London, Sydney, Toronto (1969)
- [Ebbi] Ebbinghaus, H.D., Hermes, H., Hirzebruch, F., Koecher, M., Mainzer, K., Neukirch, J., Prestel, A., Remmert, R.: Zahlen. Grundwissen Mathematik, 3. Aufl. Bd. 1. (1992)
- [Fis] Fischer, G.: Lineare Algebra, 9. Aufl. Verlag Vieweg, Braunschweig, Wiesbaden (1986)
- [FrPr] Friedrichsdorf, U., Prestel, A.: Mengenlehre für den Mathematiker. Verlag Vieweg, Braunschweig, Wiesbaden (1985)
- [Hal] Halmos, P.R.: Naive Mengenlehre. Vandenhoeck & Ruprecht, Göttingen (1968)
- [HeQu] Heise, W., Quattrocchi, P.: Informations- und Codierungstheorie, 3. Aufl. Springer-Verlag, Berlin, Heidelberg (1995)
- [Heu] Heuser, H.: Lehrbuch der Analysis, Teil 1, 8. Aufl. Teubner, Stuttgart (1990)
- [Hill] Hill, R.: A first course in coding theory. Clarendon Press, Oxford (1986)
- [Hun] Hungerford, T.: Algebra. Graduate Texts in Mathematics 73. Springer-Verlag, New York (1974)
- [Jän] Jänich, K.: Lineare Algebra, 5. Aufl. Springer-Verlag, Berlin, Heidelberg (1993)
- [Lan] Lang, S.: Linear Algebra, 2. Aufl. Addison-Wesley, New York (1971)
- [Lor] Lorenz, F.: Lineare Algebra I, II, 2. Aufl. B.I.-Wissenschaftsverlag, Mannheim, Wien, Zürich (1988, 1989)
- [Lün] Lüneburg, H.: Vorlesungen über Lineare Algebra. B.I. Wissenschaftsverlag, Mannheim, Wien, Zürich (1993)
- [MWSL] MacWilliams, F.J., Sloane, N.J.A.: The Theory of Error-Correcting Codes. North-Holland, Amsterdam, New York, Oxford (1978)

- [Mäd] Mäder, P.: Mathematik hat Geschichte. Metzler, Hannover (1992)
- [Scho] Scholz, E. (Hrsg.): Geschichte der Algebra. Eine Einführung. B.I.-Wissenschaftsverlag, Mannheim, Wien, Zürich (1990)
- [Schu] Schulz, R.H.: Codierungstheorie – Eine Einführung. Verlag Vieweg, Braunschweig und Wiesbaden (1991)
- [Tama] Tamaschke, O.: Projektive Geometrie II. Bibliographisches Institut, Mannheim, Wien, Zürich (1972)
- [Wey] Weyl, H.: Raum – Zeit – Materie, 6. Aufl. Springer-Verlag, Berlin (1970)
- [WußA] Wußing, H., Arnold, W.: Biographien bedeutender Mathematiker. Aulis Verlag, Köln (1978)
- [Zad] Zaddach, A.: Graßmanns Algebra in der Geometrie. B.I. Wissenschaftsverlag, Mannheim (1994)

Sachverzeichnis

A

Abbildung, 7, 8
abelsche Gruppe, 272
Abgeschlossenheit, 31
Abgeschlossenheit eines affinen Unterraums, 127
Abstand, 130
abzählbaren Mengen, 14
additive Gruppe eines Körpers, 273
Additive Gruppe eines Rings, 273
Additive Gruppe eines Vektorraums, 274
Additivität der Matrizenmultiplikation, 109
affine Linearkombination, 142
affiner Raum, 123
affiner Unterraum, 126
ähnliche Matrizen, 162
Ähnlichkeit von Darstellungsmatrizen, 161, 162
alternierende Gruppe, 216
Anführer, 137
Anordnung eines Körpers, 32
Anzahl aller Permutationen, 207
Anzahl der Nebenklassen, 88
Anzahl der Vektoren einer Nebenklasse, 87
Anzahl der Vektoren eines endlichen Vektorraums, 72
Äquivalenz von Injektivität und Surjektivität, 25
Äquivalenz von Injektivität und Surjektivität von linearen Abbildungen, 166
Äquivalenzklasse, 5
Äquivalenzrelation, 5
auf, 9
Aus orthogonal mach orthonormal, 313
Austauschlemma, 102
Austauschlemma, 75
Austauschsatz, 102
Austauschsatz von Steinitz, 77

Automorphismus, 149

Automorphismus eines Körpers, 47

Automorphismus von Gruppen, 285

B

Basis, 71, 133
Basisauswahlsatz, 73
Basisergänzungssatz, 73
begleitender Unterraum eines affinen Unterraums, 125
Behauptung, 20
Berechnung der Geraden durch zwei Punkte, 125
Beschreibung einer linearen Abbildung durch die Bilder einer Basis, 150
Beschreibung symmetrischer Bilinearformen durch selbstadjungierte Abbildungen, 325
Beweis, 20
Bézout, Étienne (1730–1783), 189
Bidualcode, 135
Bidualraum, 169
bijektive Abbildung, 9
Bild, 7
Bild eines Gruppenhomomorphismus, 286
Bild(f), 148
Bild(f) ist ein Vektorraum, 148
Bildbereich einer Abbildung, 7
bilinear, 298
Bilinearform, 297, 298
Bilinearformen kommen von Linearformen, 301
Bilinearformen und Abbildungen von V nach V^* , 302
binäre Folge, 25
binäre Verknüpfung, 31

Bunjakowski, Viktor Jakowlewitsch
(1804–1889), 312

C

Cantor, Georg (1845–1918), 1
Cardano, Geronimo (1501–1576), 36
cartesisches Produkt, 2
Cauchy, Augustin-Louis (1789–1857), 312
Cayley, Arthur (1821–1895), 259
Charakterisierung der linearen Abhängigkeit,
68
Charakterisierung einer linearen Abbildung
durch die Bilder einer Basis, 152
Charakterisierung invertierbarer Matrizen, 161
Charakterisierung von Isomorphismen, 159
Charakteristik eines endlichen Körpers, 54
Charakteristisches Polynom einer linearen
Abbildung, 250
charakteristisches Polynom einer linearen
Abbildung, 251
charakteristisches Polynom einer Matrix, 250
Code, 131
Codewort, 129
Codewörter, 131
codieren, 129

D

Darstellung einer Permutation als Produkt
disjunkter Zyklen, 209
Darstellung einer Transposition durch eine
ungerade Anzahl von
Nachbartranspositionen, 212
Darstellung von Permutationen durch
Transpositionen, 211
Darstellungsmatrix, 155
Darstellungsmatrix der Identität, 157
Darstellungsmatrix des Produkts von linearen
Abbildungen, 160
Darstellungsmatrix einer Matrix, 158
Darstellungsmatrix einer selbstadjungierten
Abbildung, 326
Darstellungssatz, 156
Daten, 129
decodieren, 129, 132
Dedekind, Richard (1831–1916), 15
Dedekindsche Beschreibung unendlicher
Mengen, 15
Definition einer Menge durch Auflistung ihrer
Elemente, 1

Definition einer Menge durch eine Eigenschaft,
2

Definitionsbereich einer Abbildung, 7
Descartes, Rene (1596–1650), 3
Determinante einer 22-Matrix, 204
Determinante einer Dreiecksmatrix, 222
Determinante einer Matrix, 204
Determinante einer orthogonalen Matrix, 319
Determinante einer regulären Matrix, 207
Determinantenfunktion, 203
diagonalisierbar, 242
Diagonalmatrix, 242
Differenz zweier Mengen, 2
Dimension, 78
Dimension des Kern einer Linearform, 167
Dimension des Lösungsraums eines
homogenen linearen Gleichungssystems,
114, 165
Dimension eines affinen Unterraums, 127
Dimension eines Eigenraums, 254
Dimensionsformel für Bilinearformen, 306
Dimensionsformel für den dualen Code, 135
Dimensionsformel für lineare Abbildungen, 165
Dimensionssatz, 83
Direkter Beweis, 21
disjunkte Mengen, 2
Division mit Rest, 40
Doppelsumme, 19
Drehung, 316, 321
Dreiecksaxiom, 123
Dreiecksmatrix, 222, 223
duale Basis, 168
dualer Code, 134
Dualraum, 167
Dualraum ist ein Vektorraum, 167
Durchschnitt von Mengen, 2

E

Ebene eines affinen Raums, 127
echte Teilmenge, 1
Effekt einer Nachbartransposition, 214
Eigenraum, 246
Eigenvektor einer linearen Abbildung, 243
Eigenwert einer linearen Abbildung, 243
Eigenwert einer Matrix, 244
Eigenwerte einer orthogonalen Abbildung, 317
Eindeutige Lösung eines linearen
Gleichungssystems, 121
Eindeutigkeit der Determinantenfunktion, 206

Eindeutigkeit der inversen Abbildung, 12
 Eindeutigkeit der inversen Elemente einer Gruppe, 272
 Eindeutigkeit der Nebenklassenanhhrer, 137
 Eindeutigkeit der negativen Vektoren, 61
 Eindeutigkeit des neutralen Elements einer Gruppe, 272
 Eindeutigkeit des Nullvektors, 60
 Eindeutigkeitseigenschaft einer Basis, 71
 Einheitsmatrix, 113
 Einselement eines Krpers, 30
 Einsetzen in ein Polynom, 185
 Einsetzungshomomorphismus, 186
 Element einer Menge, 1
 elementare Spaltenumformung, 117
 elementare Zeilenumformung, 116
 Elementarmatrix, 234
 Empfnger, 129
 endlich erzeugbarer Vektorraum, 73
 endlich-dimensionaler Vektorraum, 78
 endliche Gruppe, 280
 endliche Mengen, 3
 endlicher Vektorraum, 72
 Endomorphismus, 150
 Entwicklung nach der ersten Zeile, 225
 Entwicklung nach einer Spalte, 227
 Entwicklung nach einer Zeile, 225
 Ersetzen eines Vektors durch einen anderen, 75
 erweiterte Matrix eines Gleichungssystems, 113
 erweiterter Hamming-Code, 145
 Erweiterung eines Krpers, 65
 Erzeugendensystem, 71, 102
 Erzeugnis, 69
 Erzeugnis von Elementen einer Gruppe, 282
 Erzeugung von affinen Unterrumen, 128
 Euklidische Norm, 312
 euklidische Norm, 311
 euklidischer Algorithmus, 40
 euklidischer Algorithmus fr Polynome, 197
 euklidischer Vektorraum, 308
 Euler, Leonhard (1707–1783), 284
 Existenz einer Basis, 74

F

Faktorgruppe, 287
 Faktorraum, 88
 fehlererkennender Code, 143
 fehlerkorrigierender Code, 131
 Fehlervektor, 130

Fehlstand einer Permutation, 213
 Fehlstnde einer Transposition, 213
 Fermat, Pierre de (1601–1665), 284
 Fixpunkt, 56
 Fixpunkt einer Permutation, 208
 Folge, 2
 Fundamentalsatz fr endlich-dimensionale Vektorrume, 154

G

Galois, Evariste (1811–1832), 46
 Galoisfeld, 46
 ganze Zahlen, 2
 Gau, Carl Friedrich (1777–1855), 55
 Gausche Zahlenebene, 55
 Gauscher Algorithmus, 120, 121
 general linear group, 275
 Generatormatrix, 133
 Geometrie, 123
 gerade Permutation, 214
 Gewicht, 133
 $GF(q)$, 46
 ggT, 199
 Gleichheitsrelation, 5, 340
 gleichmchtige Mengen, 13
 Gleichmchtigkeit endlicher Mengen, 13
 Gleichmchtigkeit von N und Z , 15
 Gleichmchtigkeit von Z und $2Z$, 14
 Grad des Nullpolynoms, 182
 Grad eines Polynoms, 182
 Gradformel fr Polynome, 184
 Gram, Jrgen Perdersen (1850–1916), 299
 Gramsche Matrix, 299
 Gramann, Hermann Gnther (1809–1877), 92
 groter gemeinsamer Teiler zweier ganzer Zahlen, 199
 Gruppe, 271

H

Hamilton, William Rowan (1805–1865), 36
 Hamming, Richard W., 130
 Hamming-Abstand, 130
 Hamming-Code, 143
 Hauptachsentransformation, 327
 Hauptdiagonale, 113
 Hauptideal, 193
 Hauptidealring, 198
 Hauptsatz der Algebra, 251
 Hintereinanderausfhrung von Abbildungen, 10

$\text{Hom}(V, W)$, 175
 homogenes lineares Gleichungssystem, 105
 Homogenität der Matrizenmultiplikation, 110
 Homomorphieprinzip, 42
 Homomorphiesatz für Gruppen, 288
 Homomorphiesatz für Vektorräume, 164
 Homomorphismus, 147
 Homomorphismus eines Körpers, 47
 Homomorphismus von Gruppen, 285
 Hyperebene, 82, 103

I

i, 34
 Ideal eines Ringes, 192
 Ideale eines Polynomrings, 193
 identische Abbildung, 8, 148
 Identität, 8
 imaginäre Einheit, 34
 imaginäre Einheiten des
 Quaternionenschiefkörpers, 37
 Imaginärteil, 34
 Index einer Untergruppe, 281
 indirekter Beweis, 21
 Induktion, 22
 inhomogenes lineares Gleichungssystem, 105
 injektive Abbildung, 9
 inneres Produkt, 134
 Interpolationsformel von Lagrange, 197
 Invarianz der Determinantenfunktion
 gegenüber elementaren
 Spaltenumformungen, 230
 Invarianz der Determinantenfunktion
 gegenüber elementaren
 Zeilenumformungen, 205
 Invarianz der Geraden unter einer linearen
 Abbildung, 150
 Invarianz der negativen Elemente unter einem
 Automorphismus, 48
 Invarianz der neutralen Elemente unter einem
 Automorphismus, 48
 Invarianz der Ordnungsrelation in \mathbb{R} , 50
 Invarianz des Lösungsraums bei elementaren
 Spaltenvertauschungen, 119
 Invarianz des Lösungsraums bei elementaren
 Zeilenumformungen, 119
 Invarianz des Rangs einer Matrix bei
 elementaren Umformungen, 117, 141,
 348
 inverse Abbildung, 12

Inverse einer Elementarmatrix, 234
 invertierbare Abbildung, 11
 invertierbare Matrix, 160
 Invertierbare Polynome, 184
 Invertierbarkeit bijektiver Abbildungen, 10
 Invertierbarkeit modulo eines irreduziblen
 Polynoms, 191
 irreduzibles Polynom, 190
 Isomorphie von Gruppen, 285
 Isomorphie von Körpern, 47
 Isomorphismus, 149
 Isomorphismus von Gruppen, 285
 Isomorphismus von Körpern, 47

K

Kanal, 129
 Kern einer linearen Abbildung, 162
 Kern einer Linearform, 167
 Kern eines Gruppenhomomorphismus, 286
 Kern eines Homomorphismus, 287
 Klassifikation der zyklischen Gruppen, 285
 Kleiner Satz von Fermat, 284
 Kleiner Satz von Fermat (Originalfassung), 284
 Kn, 62
 Koeffizient eines Polynoms, 182
 Koeffizienten des charakteristischen Polynoms,
 252
 Koeffizientenvergleich, 69
 kommutative Gruppe, 272
 kommutativer Körper, 32
 kommutativer Ring, 178
 Komplement, 80
 Komplement einer Menge, 2
 komplementäre Unterräume, 80, 102
 komplexe Zahl, 33
 komplexer Vektorraum, 60
 komponentenweise, 33
 konjugiert-komplexe Zahl, 51
 konstante Zeilensumme, 268
 Konstruktion von symmetrischen
 Bilinearformen, 308
 Kontraposition, 21
 Kontrollmatrix, 135
 Körper, 29
 Körper der komplexen Zahlen, 36
 Körper von Primzahlordnung, 43
 Kriterien zur Diagonalisierbarkeit, 244, 247,
 254, 262

Kriterium für die Gleichheit von Nebenklassen, [86](#), [280](#)

Kriterium für die Lösbarkeit eines linearen Gleichungssystems, [113](#)

Kugel, [131](#)

k-Zyklus, [209](#)

L

Länge einer binären Folge, [25](#)

Länge eines Hamming-Codes, [143](#)

Länge eines Zyklus, [209](#)

leere Menge, [2](#)

Leibniz, Gottfried Wilhelm (1646–1716), [207](#)

Lemma über das Minimalgewicht, [133](#)

Lemma über Hammingkugeln, [131](#)

Lemma über verschiedene Eigenwerte, [245](#)

Lemma über verschiedene Nebenklassen, [87](#)

Lemma von Bézout, [189](#)

Leonhard Euler (1707–1783), [35](#)

linear abhängig, [67](#)

linear unabhängig, [67](#), [102](#)

linearer Code, [133](#)

lineares Gleichungssystem, [105](#)

Linearform, [167](#)

Linearformen induzieren Bilinearformen, [303](#)

Linearkombination, [67](#), [102](#)

Linearkombination einer unendlichen Menge, [67](#)

Lösung eines linearen Gleichungssystems, [105](#)

Lösungsraum eines homogenen Systems, [111](#)

Lösungsraum eines inhomogenen Systems, [112](#)

M

Mächtigkeit der Potenzmenge, [17](#)

MacMath, [4](#)

Magnetische Anziehungseigenschaft eines Ideals, [192](#)

Matrix, [63](#)

maximal linear unabhängige Menge, [73](#)

Menge, [1](#)

Mersennesche Primzahlen, [281](#)

Methode zur Bestimmung von Parallelen, [126](#)

Minimalabstand, [131](#)

Minimalgewicht, [133](#)

Mittelpunkt, [131](#)

modulo, [40](#)

modulo n , [41](#), [179](#)

Multiplikation mit Elementarmatrizen, [234](#)

Multiplikation von Polynomen, [182](#)

Multiplikationssatz für Determinanten, [233](#)

Multiplikative Gruppe eines Körpers, [273](#)

multiplikative Inverse einer komplexen Zahl, [35](#)

multiplikative Inverse eines Quaternions, [39](#)

N

$n!$, [208](#)

Nachbartransposition, [211](#)

Nachricht, [129](#), [130](#)

natürliche Zahlen, [2](#)

„Natürlicher Isomorphismus“ des Bidualraums, [170](#)

Nebendiagonale einer Matrix, [140](#)

Nebenklasse, [85](#)

Nebenklasse nach einer Untergruppe, [279](#)

neutrales Element einer Gruppe, [271](#)

nichtausgeartete Bilinearform, [300](#)

Normalteiler einer Gruppe, [286](#)

normiertes Polynom, [194](#)

n -Tupel, [2](#)

Nullabbildung, [52](#), [148](#)

Nullelement eines Körpers, [29](#), [177](#)

Nullpolynom, [182](#)

Nullstelle eines Polynoms, [188](#)

Nullstellen des Minimalpolynoms, [261](#)

Nullstellen eines Polynoms, [188](#)

nullteilerfrei, [110](#)

Nullteilerfreiheit, [31](#)

O

obere Dreiecksmatrix, [222](#)

Ordnung einer Gruppe, [280](#)

Ordnung eines Elements einer Gruppe, [283](#)

Ordnung eines Elements in \mathbb{Z}_n , [54](#)

orthogonale Gruppe, [319](#)

orthogonale lineare Abbildung, [316](#)

orthogonale Matrix, [317](#)

orthogonale Vektoren, [300](#)

orthogonales Komplement, [310](#)

Orthonormalisierungssatz von E. Schmidt, [314](#)

Ortsvektoren, [62](#)

P

parallel, [123](#)

Parallelenaxiom, [123](#)

Parallelenschar, [123](#)

Partition einer Menge, [6](#)

Permutation, [207](#)

Polynom, [182](#)

Polynomdivision, 187
 Polynomring, 182
 positiv definite Bilinearform, 308
 Potenz einer Matrix, 140
 Potenzmenge, 16
 Primzahl, 43
 Produkt von Abbildungen, 10
 Produkt von Matrizen, 106
 Produkt von Permutationen, 208
 Produkt von Polynomen, 180
 Produktformel für die Mächtigkeit des
 cartesischen Produkts, 4

Q

Quadratgruppe, 277
 quadratische Matrix, 116, 204
 Quadrupel, 36
 Quaternionen, 37
 Quaternionenschiefkörper, 37

R

Radikal einer Bilinearform, 330
 Radius, 131
 Rang einer Gramschen Matrix, 300
 Rang einer linearen Abbildung, 158
 Rang einer Matrix, 112
 rationale Zahlen, 2
 Realteil, 34
 Reduktion der Determinante auf eine
 Diagonalmatrix, 223
 reelle Zahlen, 2
 reeller Vektorraum, 60
 Reflexivität einer Relation, 4
 Regel von Sarrus, 219
 reguläre Matrix, 160
 Relation, 4
 Repräsentant einer Äquivalenzklasse, 7
 Repräsentant einer Nebenklasse, 85
 Rest bei ganzzahliger Division, 40
 Restklassenring, 179
 Ring, 177
 Ringhomomorphismus, 186
 Ringschluß, 21
 Ruffini, Paolo (1765–1822), 189

S

Sarrus, Pierre-Frédéric (1798–1858), 220
 Satz über Äquivalenzklassen, 6
 Satz über das charakteristische Polynom, 251
 Satz über das Minimalpolynom, 260

Satz über das Standardskalarprodukt, 309
 Satz über den Polynomring, 180
 Satz über die alternierende Gruppe, 216
 Satz über die Anzahl der Nullstellen, 191
 Satz über die Bijektivität invertierbarer
 Abbildungen, 25
 Satz über die Faktorgruppe, 287
 Satz über die Kontrollmatrix, 136
 Satz über die Nullstellen eines Polynoms, 191
 Satz über die Ordnung von Elementen einer
 Gruppe, 283
 Satz über die Symmetriegruppe, 277
 Satz über die Untergruppen einer zyklischen
 Gruppe, 285
 Satz über die Vandermondesche Determinante,
 231
 Satz über Eigenwerte einer Matrix, 245
 Satz über gerade Permutationen, 215
 Satz über Gleichheit von Äquivalenzklassen, 6
 Satz über konjugiert-komplexe Zahlen, 51
 Satz über Matrizenringe, 179
 Satz über Nachbartranspositionen, 214
 Satz über orthogonale Unterräume, 305
 Satz über ungerade Permutationen, 216
 Satz vom Dualraum, 167
 Satz vom Faktorraum, 88
 Satz vom orthogonalen Komplement, 310
 Satz von Cayley-Hamilton, 256
 Satz von der Potenzmenge, 16
 Satz von Euler, 284
 Satz von Lagrange, 281
 Satz von Lagrange für Gruppenelemente, 283
 Satz von Ruffini, 189
 Schiefkörper, 32
 Schmidt, Erhard (1876–1959), 314
 Schwarz, Hermann Amandus (1853–1921), 312
 selbstadjungierte lineare Abbildung, 325
 Semi-bilinearform, 331
 Sender, 129
 senkrecht aufeinander stehende Vektoren, 300
 Signum der inversen Permutation, 216
 Signum einer Permutation, 214
 Signumsformel, 218
 Skalar, 59
 skalares Vielfaches einer Funktion, 64
 Skalarprodukt, 308
 Skalarprodukt eines komplexen Vektorraums,
 332
 Spaltenrang, 112

special linear group, 275
Spektralsatz, 327
Spezielle lineare Gruppe, 275
spezielle orthogonale Gruppe, 319
Spiegelung, 321
Spur einer Matrix, 174, 252
Spurabbildung, 174
Standardbilinearform, 300
Standardskalarprodukt, 300
starrer Körper, 49
Starrheit von \mathbb{Q} , 47
Starrheit von \mathbb{R} , 51
Streckung, 56
Strukturmatrix einer Bilinearform, 299
Summationsindex, 19
Summe von Funktionen, 64
Summe von Unterräumen, 99
Summe zweier Punkte, 55
Summenformel für die Mächtigkeit von Mengen, 3
surjektive Abbildung, 9
Sylvester, James Joseph (1814–1897), 328
Sylvesterscher Trägheitssatz, 328
Symmetrie einer Relation, 4
Symmetriegruppe des Würfels, 277
Symmetriegruppe eines geometrischen Objekts, 276
Symmetriegruppe eines Quadrats, 277
symmetrische Bilinearform, 307
symmetrische Gruppe, 208, 276
Syndrom, 136
Syndrom-Decodierung, 138

T
Tamaschke, Olaf, 123
Tamaschke-Axiom, 123
Teilbarkeit ganzer Zahlen, 198
Teilbarkeit von Polynomen, 189
teilerfremde ganze Zahlen, 291
teilerfremde Polynome, 189
Teilkörper, 66
Teilmenge, 1
Transformation der Darstellungsmatrix bei Basiswechsel, 161
Transitivität einer Relation, 4
transponierte Matrix, 228
Transposition, 211
Transposition eines Vektors, 109
triviale Darstellung des Nullvektors, 67

triviale Untergruppen, 279
trivialer Automorphismus, 47
trivialer Unterraum, 71
triviales Ideal, 192

U

überabzählbare Mengen, 14
überflüssige Spalte, 115
umkehrbare Abbildung, 11
umkehrbare Matrix, 160
Unbestimmte, 182
unendlichdimensionale Vektorräume, 101
unendliche Mengen, 14
ungerade Permutation, 214
unitärer Vektorraum, 332
Untergruppe, 278
Untergruppenkriterium, 279
Unterkörper, 66
Untermenge, 1
Unterraum, 70
Unterraumkriterium, 98
Unterring, 185
Urbild, 7

V

van der Waerden, Bartel Leendert (geb. 1903), 94
Vandermonde, A. T. (1735–1796), 230
Vandermondesche Determinante, 230
Vandermondesche Matrix, 230
Vektor, 59, 92
Vektorraum, 59
Verbindungsaxiom, 123
Vereinigung von Mengen, 2
Verknüpfung, 31
Verschiedene Bilinearformen von V („Wer eine kennt, kennt alle“), 304
Vertauschung der Summationsreihenfolge, 20
Vielfachheit einer Nullstelle eines Polynoms, 189
Vieta, Francois (1540–1603), 36
Volle lineare Gruppe, 275
Voraussetzung, 20

W

Wallis, John (1616–1703), 55
Wessel, Caspar (1745–1818), 55
Weyl, Hermann (1885–1955), 94
Widerspruchsbeweis, 21
wohldefiniert, 7, 90

Würfelgruppe, 277

X

x, 182

Z

Zeilenrang, 114

Zeilenrang = Spaltenrang, 114

Zentrum einer Streckung, 56

Zerlegung eines Vektors bezüglich
komplementärer Unterräume, 81

\mathbb{Z}_n , 41

Zornsches Lemma, 73

zweiter Dimensionssatz, 90

zyklische Gruppe, 282

zyklische Permutation, 209

Zyklus, 209