

Springer-Lehrbuch

Siegfried Bosch

Lineare Algebra

Vierte, überarbeitete Auflage

Prof. Dr. Siegfried Bosch
Mathematisches Institut
Universität Münster
Einsteinstraße 62
48149 Münster
bosch@math.uni-muenster.de

ISBN 978-3-540-76437-3

e-ISBN 978-3-540-76438-0

DOI 10.1007/978-3-540-76438-0

Springer-Lehrbuch ISSN 0937-7433

Bibliografische Information der Deutschen Nationalbibliothek

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

Mathematics Subject Classification (2000): 15-01

© 2008, 2006, 2003, 2001 Springer-Verlag Berlin Heidelberg

Dieses Werk ist urheberrechtlich geschützt. Die dadurch begründeten Rechte, insbesondere die der Übersetzung, des Nachdrucks, des Vortrags, der Entnahme von Abbildungen und Tabellen, der Funk-sendung, der Mikroverfilmung oder der Vervielfältigung auf anderen Wegen und der Speicherung in Datenverarbeitungsanlagen, bleiben, auch bei nur auszugsweiser Verwertung, vorbehalten. Eine Vervielfältigung dieses Werkes oder von Teilen dieses Werkes ist auch im Einzelfall nur in den Grenzen der gesetzlichen Bestimmungen des Urheberrechtsgesetzes der Bundesrepublik Deutschland vom 9. September 1965 in der jeweils geltenden Fassung zulässig. Sie ist grundsätzlich vergütungspflichtig. Zuwider-handlungen unterliegen den Strafbestimmungen des Urheberrechtsgesetzes.

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.

Satz: Datenerstellung durch den Autor unter Verwendung eines Springer T_EX-Makropakets

Herstellung: LE-T_EX Jelonek, Schmidt & Vöckler GbR, Leipzig

Umschlaggestaltung: WMX Design GmbH, Heidelberg

Gedruckt auf säurefreiem Papier

9 8 7 6 5 4 3 2 1

springer.de

Aus dem Vorwort zur ersten Auflage

Die Mathematik ist eine Wissenschaft, die sich heutzutage in einem äußerst vielfältigen und schillernden Gewand präsentiert. Daher stellt sich zwangsläufig die Frage, welche Bereiche für die ersten Schritte im Vordergrund stehen sollten, wenn man ein Studium der Mathematik aufnehmen möchte. Natürlich hat sich die Art der Ausbildung mit der Zeit gewandelt. In kontinuierlicher Weise sind grundlegende Einsichten, die im Rahmen der Erforschung aktueller Probleme zutage getreten sind, mit in die Lehre eingeflossen. Dabei geht es in der Mathematik keineswegs um komplizierte Details, sondern vielmehr um oftmals wiederkehrende tragende Grundmuster, die sich als wichtig erwiesen haben und die ihrerseits bereits auf elementarem Niveau an sinnvollen Beispielen studiert werden können. So hat es sich bewährt, die Mathematikausbildung an Universitäten mit je einer Einführung in die Infinitesimalrechnung und die Lineare Algebra zu beginnen, meist in zwei getrennten Vorlesungen. Beide Gebiete ergänzen sich gegenseitig und beinhalten in idealer Weise eine Vielzahl interessanter mathematischer Grundmuster. Ja, man kann mit Recht sagen, dass die Methoden der Infinitesimalrechnung und der Linearen Algebra grundlegend für so gut wie alle anderen Bereiche der Mathematik sind.

Der Text dieses Bandes repräsentiert das Pensum einer zwei-semesterigen Einführungsvorlesung zur Linearen Algebra, eine Vorlesung, die ich mehrfach an der Universität Münster gehalten habe. Die meisten Studierenden verfügen bereits über gewisse Vorkenntnisse zur Linearen Algebra, wenn sie sich für ein Mathematikstudium entscheiden, etwa was die Vektorrechnung oder das Lösen linearer Gleichungssysteme angeht. Sie sind dagegen aller Erfahrung nach weniger mit den allgemeinen Begriffsbildungen der Linearen Algebra vertraut, die diese Theorie so universell einsatzfähig machen. Man kann sicherlich sagen, dass diese abstrakte Seite der Linearen Algebra für viele Studierende neue und ungewohnte Schwierigkeiten aufwirft. Ich habe mich dafür entschieden, diese Schwierigkeiten nicht zu kaschieren, sondern ihre Überwindung gezielt in den Vordergrund zu stellen. Deshalb wird in diesem Text von Anfang an großer Wert auf eine klare und systematische, aber dennoch behutsame Entwicklung der in der Linearen Algebra üblichen theoretischen Begriffsbildungen gelegt. Ad-hoc-Lösungen, die bei späteren Überlegungen oder Verallgemeinerungen revidiert werden müssten, werden nach Möglichkeit vermieden. Erst wenn die theoretische Seite eines Themenkomplexes geklärt ist, erfolgt die Behandlung

der zugehörigen Rechenverfahren, unter Ausschöpfung des vollen Leistungsumfangs.

Nun ist allerdings eine Theorie wie die Lineare Algebra, die sich in beträchtlichem Maße von ihren ursprünglichen geometrischen Wurzeln entfernt hat, nur schwerlich zu verdauen, wenn nicht gleichzeitig erklärt wird, *warum* man in dieser oder jener Weise vorgeht, *was* die zugehörige Strategie ist, oder an *welche* Hauptanwendungsfälle man mit einer gewissen Definition denkt. Um solche Fragen abzudecken, wird in einer Vorlesung neben der rein stofflichen Seite in erheblichem Maße auch das zugehörige motivierende Umfeld erläutert. In Lehrbüchern ist diese Komponente oftmals nur in geringem Maße realisiert, da ansonsten ein permanenter Wechsel zwischen der logisch-stringenten mathematischen Vorgehensweise und mehr oder weniger heuristisch-anschaulichen Überlegungen erforderlich wäre, was natürlich für die Einheitlichkeit und Übersichtlichkeit der Darstellung nicht förderlich ist. In dem vorliegenden Text wird nun jedes Kapitel mit einer Reihe von "Vorbemerkungen" eingeleitet, deren Ziel es ist, das motivierende Umfeld des jeweiligen Kapitels zu beleuchten. Ausgehend vom momentanen Kenntnisstand eines Lesers werden die zu behandelnden Hauptfragestellungen einschließlich des zugehörigen geometrischen Hintergrunds (soweit gegeben) erläutert und darüber hinaus mögliche Lösungsansätze und Lösungsstrategien, die Art der erhaltenen Lösung, wie auch die hiermit verbundenen Schwierigkeiten diskutiert. Es wird empfohlen, die Vorbemerkungen während des Studiums eines Kapitels je nach Bedarf mehrfach zu konsultieren, um größtmöglichen Nutzen aus ihnen zu ziehen. Ausdrücklich möchte ich aber darauf hinweisen, dass es sich bei diesen Einführungen zu einem großen Teil um Plausibilitätsbetrachtungen handelt. Diese sind daher nicht mit der üblichen mathematischen Prägnanz abgefasst, und sie sind infolgedessen auch nicht Teil des eigentlichen Lehrstoffes.

Der stoffliche Umfang des Buches bietet nur wenig Besonderheiten. Es werden Vektorräume und ihre linearen Abbildungen, Matrizen und lineare Gleichungssysteme, Determinanten, Polynome, Eigenwert- und Normalformentheorie sowie euklidische und unitäre Vektorräume behandelt. Ein Abschnitt über äußere Produkte (mit einem Stern * gekennzeichnet), in dem als Anwendung der allgemeine Laplacesche Entwicklungssatz für Determinanten bewiesen wird, ist optional. Die Herleitung der Normalformen für Endomorphismen von Vektorräumen erfolgt, der Gesamtstrategie des Buches folgend, im Rahmen von Moduln über Hauptidealringen, wobei solche Moduln allerdings erst zu Beginn von Abschnitt 6.3 eingeführt werden. Wer sich hier auf die elementare Seite der Normalformentheorie beschränken möchte, kann im Anschluss an die Abschnitte 6.1 (Eigenwerte und Eigenvektoren) und 6.2 (Minimalpolynom und charakteristisches Polynom) auch gleich zu den euklidischen und unitären Vektorräumen in Kapitel 7 übergehen.

Vorwort zur vierten Auflage

Seit Erscheinen meiner LINEAREN ALGEBRA hat mich eine Vielzahl von Kommentaren und Vorschlägen zu diesem Thema erreicht, für die ich allen Lesern herzlich danken möchte. Insbesondere gilt dies für die Hörer meiner Vorlesungen, die in den vergangenen Semestern die Grundlagen der Linearen Algebra mit Unterstützung dieses Textbuches erlernt haben.

In der vorliegenden Neuauflage gibt es nur wenige Änderungen. Einige Druckfehler wurden beseitigt und einige kleinere Unstimmigkeiten begradigt. Ansonsten liefert der Text nach wie vor eine detaillierte Einführung in die Lineare Algebra, die sich an den Anforderungen einer zweisemestrigen Einführungsvorlesung orientiert. Beibehalten und konsequent genutzt wurde der bewährte Standpunkt, mit $\mathbb{N}$ die natürlichen Zahlen *einschließlich* der 0 zu bezeichnen. Diese Konvention wird in der Literatur leider nicht einheitlich gehandhabt, sie bringt für uns allerdings manche Vereinfachung mit sich, indem sie es an vielen Stellen ermöglicht, auf die Hervorhebung trivialer (und daher eigentlich uninteressanter) Ausnahmefälle zu verzichten.

Münster, im November 2007

Siegfried Bosch

Inhalt

1	Vektorräume	1
1.1	Mengen und Abbildungen	9
1.2	Gruppen	12
1.3	Körper	17
1.4	Vektorräume	26
1.5	Linear unabhängige Systeme und Basen von Vektorräumen . . .	32
1.6	Direkte Summen	44
2	Lineare Abbildungen	51
2.1	Grundbegriffe	57
2.2	Quotientenvektorräume	65
2.3	Der Dualraum	75
3	Matrizen	85
3.1	Lineare Abbildungen und Matrizen	90
3.2	Das Gaußsche Eliminationsverfahren und der Rang einer Matrix	99
3.3	Matrizenringe und invertierbare Matrizen	109
3.4	Basiswechsel	115
3.5	Lineare Gleichungssysteme	119
4	Determinanten	131
4.1	Permutationen	134
4.2	Determinantenfunktionen	139
4.3	Determinanten von Matrizen und Endomorphismen	143
4.4	Die Cramersche Regel	151
4.5	Äußere Produkte*	155
5	Polynome	165
5.1	Ringe	166
5.2	Teilbarkeit in Integritätsringen	176
5.3	Nullstellen von Polynomen	185
6	Normalformentheorie	189
6.1	Eigenwerte und Eigenvektoren	192
6.2	Minimalpolynom und charakteristisches Polynom	198

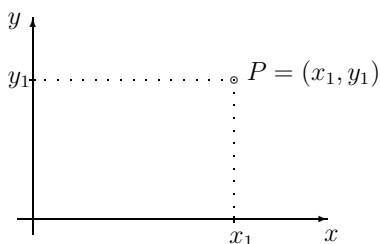
6.3	Der Elementarteilersatz	206
6.4	Endlich erzeugte Moduln über Hauptidealringen	219
6.5	Allgemeine und Jordansche Normalform für Matrizen	224
7	Euklidische und unitäre Vektorräume	243
7.1	Sesquilinearformen	246
7.2	Orthogonalität	251
7.3	Sesquilinearformen und Matrizen	258
7.4	Die adjungierte Abbildung	263
7.5	Isometrien, orthogonale und unitäre Matrizen	269
7.6	Selbstadjungierte Abbildungen	278
	Symbolverzeichnis	285
	Namen- und Sachverzeichnis	291

1. Vektorräume

Vorbemerkungen

Konkrete geometrische Fragestellungen in der Ebene oder im drei-dimensionalen Raum waren vielfach Ausgangspunkt bedeutender mathematischer Entwicklungen. Als Hilfsmittel zur Behandlung solcher Fragen wurden beispielsweise geometrische Konstruktionsverfahren mittels Zirkel und Lineal entwickelt. Eine andere Strategie besteht darin, geometrische Fragen in rechnerische Probleme umzusetzen, um durch “Ausrechnen” zu Lösungen zu gelangen. Dies ist das Vorgehen der *analytischen Geometrie*, die 1637 von René Descartes in seinem berühmten Werk “La Géométrie” begründet wurde. Ein Großteil der rechnerischen Methoden der analytischen Geometrie wiederum wird heute in erweiterter Form unter dem Begriff der *Linearen Algebra* zusammengefasst.

Wir wollen im Folgenden etwas näher auf die grundlegenden Ideen des Descartes’schen Ansatzes eingehen. Hierzu betrachten wir eine Ebene E (etwa in dem uns umgebenden drei-dimensionalen Raum), zeichnen einen Punkt von E als so genannten Nullpunkt 0 aus und wählen dann ein Koordinatensystem mit Koordinatenachsen x und y , die sich im Nullpunkt 0 schneiden. Identifizieren wir die Achsen x und y jeweils noch mit der Menge $\mathbb{R}$ der reellen Zahlen, so lassen sich die Punkte P von E als Paare reeller Zahlen interpretieren:



In der Tat, ist P ein Punkt in E , so konstruiere man die Parallele zu y durch P . Diese schneidet die Achse x in einem Punkt x_1 . Entsprechend schneidet die Parallele zu x durch P die Achse y in einem Punkt y_1 , so dass man aus P das Koordinatenpaar (x_1, y_1) erhält. Umgekehrt lässt sich P aus dem Paar (x_1, y_1) in einfacher Weise zurückgewinnen, und zwar als Schnittpunkt der Parallelen zu y durch x_1 und der Parallelen zu x durch y_1 . Genauer stellt man fest, dass die Zuordnung $P \mapsto (x_1, y_1)$ eine umkehrbar eindeutige Beziehung zwischen

den Punkten von E und den Paaren reeller Zahlen darstellt und man deshalb wie behauptet eine Identifizierung

$$E = \mathbb{R}^2 = \text{Menge aller Paare reeller Zahlen}$$

vornehmen kann. Natürlich hängt diese Identifizierung von der Wahl des Nullpunktes 0 sowie der Koordinatenachsen x und y ab. Wir haben in obiger Abbildung ein rechtwinkliges Koordinatensystem angedeutet. Im Prinzip brauchen wir jedoch an dieser Stelle noch nichts über Winkel zu wissen. Es genügt, wenn wir als Koordinatenachsen zwei verschiedene Geraden x und y durch den Nullpunkt 0 verwenden. Genauer hierzu werden wir noch in den Vorbemerkungen zu Kapitel 2 besprechen.

Es soll nun auch die Identifizierung der beiden Koordinatenachsen x und y mit der Menge $\mathbb{R}$ der reellen Zahlen noch etwas genauer beleuchtet werden. Durch Festlegen des Nullpunktes ist auf x und y jeweils die Streckungsabbildung mit Zentrum 0 und einer reellen Zahl als Streckungsfaktor definiert. Wählen wir etwa einen von 0 verschiedenen Punkt $1_x \in x$ aus und bezeichnen mit $\alpha \cdot 1_x$ das Bild von 1_x unter der Streckung mit Faktor α , so besteht x gerade aus allen Punkten $\alpha \cdot 1_x$, wobei α die reellen Zahlen durchläuft. Genauer können wir sagen, dass die Zuordnung $\alpha \mapsto \alpha \cdot 1_x$ eine umkehrbar eindeutige Beziehung zwischen den reellen Zahlen und den Punkten von x erklärt. Nach Auswahl je eines von 0 verschiedenen Punktes $1_x \in x$ und entsprechend $1_y \in y$ sind daher x und y auf natürliche Weise mit der Menge $\mathbb{R}$ der reellen Zahlen zu identifizieren, wobei die Punkte $0, 1_x \in x$ bzw. $0, 1_y \in y$ den reellen Zahlen 0 und 1 entsprechen. Die Möglichkeit der freien Auswahl der Punkte $1_x \in x$ und $1_y \in y$ wie auch die Verwendung nicht notwendig rechtwinkliger Koordinatensysteme machen allerdings auf ein Problem aufmerksam: Der Abstand von Punkten in E wird unter der Identifizierung $E = \mathbb{R}^2$ nicht notwendig dem auf $\mathbb{R}^2$ üblichen euklidischen Abstand entsprechen, der für Punkte $P_1 = (x_1, y_1)$ und $P_2 = (x_2, y_2)$ durch

$$d(P_1, P_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

gegeben ist. Eine korrekte Reflektierung von Abständen auf E ist jedoch mit Hilfe der später noch zu diskutierenden *Skalarprodukte* möglich.

In der Mathematik ist man stets darum bestrebt, bei der Analyse von Phänomenen und Problemen, für die man sich interessiert, zu gewissen "einfachen Grundstrukturen" zu gelangen, die für das Bild, das sich dem Betrachter bietet, verantwortlich sind. Solchermaßen als wichtig erkannte Grundstrukturen untersucht man dann oftmals losgelöst von der eigentlichen Problematik, um herauszufinden, welche Auswirkungen diese haben; man spricht von einem *Modell*, das man untersucht. Modelle haben den Vorteil, dass sie in der Regel leichter zu überschauen sind, aber manchmal auch den Nachteil, dass sie den eigentlich zu untersuchenden Sachverhalt möglicherweise nur in Teilaspekten beschreiben können.

In unserem Falle liefert der Descartes'sche Ansatz die Erkenntnis, dass Punkte von Geraden, Ebenen oder des drei-dimensionalen Raums mittels Koordinaten zu beschreiben sind. Hierauf gestützt können wir, wie wir gesehen

haben, die Menge $\mathbb{R}^2$ aller Paare reeller Zahlen als Modell einer Ebene ansehen. Entsprechend bildet die Menge $\mathbb{R}^3$ aller Tripel reeller Zahlen ein Modell des drei-dimensionalen Raums, sowie natürlich $\mathbb{R} = \mathbb{R}^1$ ein Modell einer Geraden. Die Untersuchung solcher Modelle führt uns zum zentralen Thema dieses Kapitels, nämlich zu den *Vektorräumen*. Vektorräume beinhalten als fundamentale Struktur zwei Rechenoperationen, zum einen die Multiplikation von Skalaren (in unserem Falle reellen Zahlen) mit Vektoren, was man sich als einen Streckungsprozess vorstellen kann, und zum anderen die Addition von Vektoren. Wir wollen dies mit den zugehörigen geometrischen Konsequenzen einmal am Beispiel einer Ebene E und ihrem Modell $\mathbb{R}^2$ erläutern.

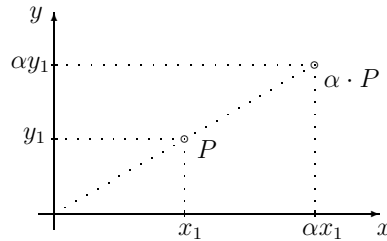
Wir beginnen mit der skalaren Multiplikation. Für

$$\alpha \in \mathbb{R}, \quad P = (x_1, y_1) \in \mathbb{R}^2$$

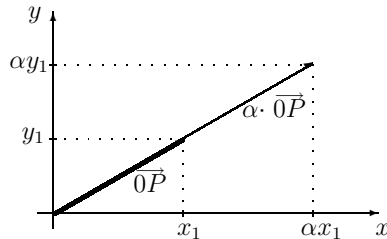
bezeichnet man mit

$$\alpha \cdot P = \alpha \cdot (x_1, y_1) := (\alpha x_1, \alpha y_1)$$

das Produkt von α und P , wobei sich in E folgendes Bild ergibt:



Die Multiplikation von Punkten $P \in E$ mit einem Skalar $\alpha \in \mathbb{R}$ ist folglich zu interpretieren als Streckungsabbildung mit Streckungszentrum 0 und Streckungsfaktor α . Besonders instruktiv lässt sich dies beschreiben, wenn man die Punkte $P \in E$ als "Vektoren" im Sinne gerichteter Strecken $\overrightarrow{OP}$ auffasst. Vektoren sind somit charakterisiert durch ihre Länge und ihre Richtung (außer für den Nullvektor $\overrightarrow{00}$, der keine bestimmte Richtung besitzt). Der Vektor $\alpha \cdot \overrightarrow{OP}$ geht dann aus $\overrightarrow{OP}$ hervor, indem man ihn mit α streckt, d. h. seine Länge mit α (oder, besser, mit dem Betrag $|\alpha|$) multipliziert und ansonsten die Richtung des Vektors beibehält bzw. invertiert, je nachdem ob $\alpha \geq 0$ oder $\alpha < 0$ gilt:



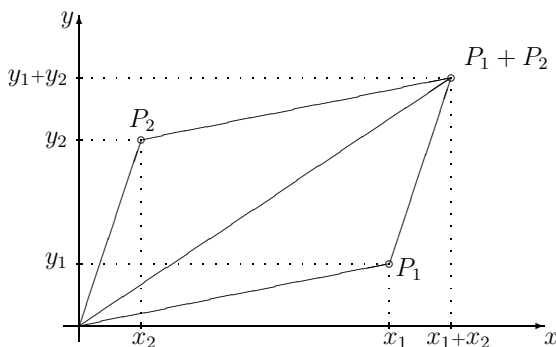
Als weitere Rechenoperation betrachten wir die Addition von Punkten in $\mathbb{R}^2$. Für

$$P_1 = (x_1, y_1), P_2 = (x_2, y_2) \in \mathbb{R}^2$$

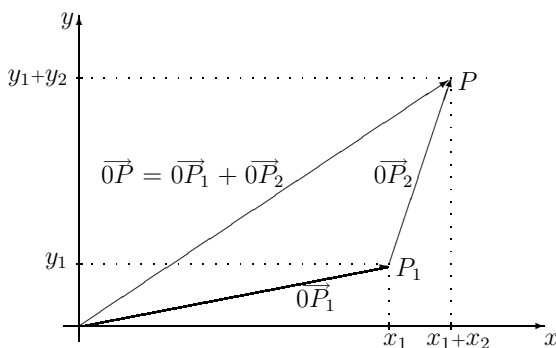
setzt man

$$P_1 + P_2 := (x_1 + x_2, y_1 + y_2),$$

was in E mittels folgender Skizze verdeutlicht werden möge:



Auch die Beschreibung der Addition in E gestaltet sich instruktiver, wenn man den Vektorstandpunkt im Sinne gerichteter Strecken zugrunde legt. Allerdings sollte man dabei zulassen, dass Vektoren als gerichtete Strecken parallel zu sich selbst verschoben und somit vom Koordinatenursprung als ihrem natürlichen Fußpunkt gelöst werden können. Die Summe der Vektoren $\overrightarrow{0P_1}$ und $\overrightarrow{0P_2}$ ergibt sich dann als Vektor $\overrightarrow{0P}$, wobei P derjenige Endpunkt ist, den man erhält, indem man beide Vektoren miteinander kombiniert, also den Vektor $\overrightarrow{0P_1}$ in 0 anlegt und den Vektor $\overrightarrow{0P_2}$ im Endpunkt P_1 von $\overrightarrow{0P_1}$, etwa wie folgt:



Dabei zeigt die obige Parallelogrammkonstruktion, dass sich das Ergebnis der Addition nicht ändert, wenn man alternativ den Vektor $\overrightarrow{0P_2}$ in 0 anlegt und anschließend den Vektor $\overrightarrow{0P_1}$ im Endpunkt von $\overrightarrow{0P_2}$. Die Addition von Vektoren hängt daher nicht von der Reihenfolge der Summanden ab, sie ist *kommutativ*.

Es mag etwas verwirrend wirken, wenn wir die Elemente des $\mathbb{R}^2$ einerseits als Punkte, sowie andererseits auch als (verschiebbare) Vektoren im Sinne gerichteter Strecken interpretieren. Im Prinzip könnte man eine begriffliche Trennung zwischen Punkten und Vektoren vornehmen, indem man den einem Punkt $P \in \mathbb{R}^2$ zugeordneten Vektor $\overrightarrow{0P}$ als *Translation* $Q \mapsto P + Q$ interpretiert, d. h. als Abbildung von $\mathbb{R}^2$ nach $\mathbb{R}^2$, die einem Element $Q \in \mathbb{R}^2$ das Element $P + Q$ als Bild zuordnet. Wir wollen von dieser Möglichkeit allerdings keinen Gebrauch machen, da eine Trennung der Begriffe für unsere Zwecke keine Vorteile bringt und die Dinge lediglich komplizieren würde.

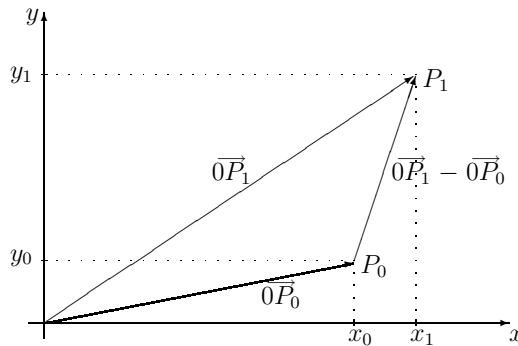
Als Nächstes wollen wir besprechen, dass die Addition von Punkten und Vektoren in $\mathbb{R}^2$ bzw. E auf natürliche Weise auch eine Subtraktion nach sich zieht. Für $P_0 = (x_0, y_0) \in \mathbb{R}^2$ setzt man

$$-P_0 = -(x_0, y_0) := (-1) \cdot (x_0, y_0) = (-x_0, -y_0)$$

und nennt dies das negative oder inverse Element zu P_0 . Dieses ist in eindeutiger Weise charakterisiert als Element $Q \in \mathbb{R}^2$, welches der Gleichung $P_0 + Q = 0$ genügt. Die Subtraktion zweier Elemente $P_1 = (x_1, y_1)$ und $P_0 = (x_0, y_0)$ in $\mathbb{R}^2$ wird dann in nahe liegender Weise auf die Addition zurückgeführt, und zwar durch

$$P_1 - P_0 := P_1 + (-P_0) = (x_1 - x_0, y_1 - y_0).$$

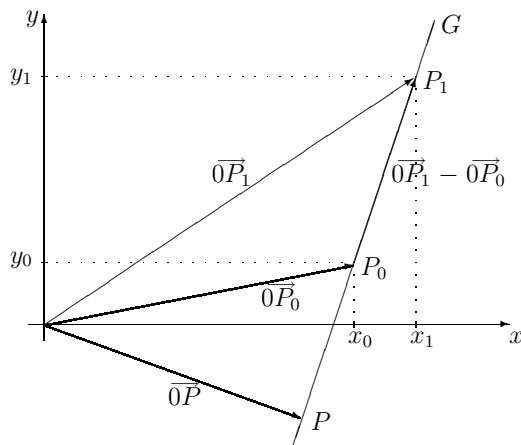
Legen wir wieder den Vektorstandpunkt in E zugrunde, so entsteht also $-\overrightarrow{0P_0}$ aus dem Vektor $\overrightarrow{0P_0}$ durch Invertieren seiner Richtung, wobei die Länge erhalten bleibt. Damit lässt sich die Differenz zweier Vektoren $\overrightarrow{0P_1}$ und $\overrightarrow{0P_0}$ wie folgt illustrieren:



Insbesondere erkennt man, dass die Summe der Vektoren $\overrightarrow{0P_0}$ und $\overrightarrow{0P_1} - \overrightarrow{0P_0}$ gerade den Vektor $\overrightarrow{0P_1}$ ergibt, was eine sinnvoll definierte Addition bzw. Subtraktion natürlich ohnehin leisten sollte. Allgemeiner kann man Summen des Typs

$$\overrightarrow{0P} = \overrightarrow{0P_0} + \alpha \cdot (\overrightarrow{0P_1} - \overrightarrow{0P_0})$$

mit unterschiedlichen Skalaren $\alpha \in \mathbb{R}$ bilden. Der Punkt P liegt dann für $P_0 \neq P_1$ stets auf der Geraden G , die durch P_0 und P_1 festgelegt ist, und zwar durchläuft P ganz G , wenn α ganz $\mathbb{R}$ durchläuft:



Die Gerade in E bzw. $\mathbb{R}^2$, welche die gegebenen Punkte P_0 und P_1 enthält, wird daher durch die Gleichung

$$G = \{P_0 + t \cdot (P_1 - P_0); t \in \mathbb{R}\}$$

beschrieben. Sind zwei solche Geraden

$$G = \{P_0 + t \cdot (P_1 - P_0); t \in \mathbb{R}\}, \quad G' = \{P'_0 + t \cdot (P'_1 - P'_0); t \in \mathbb{R}\}$$

mit $P_0 \neq P_1$ und $P'_0 \neq P'_1$ gegeben, so sind diese genau dann parallel, wenn $P_1 - P_0$ ein skalares Vielfaches von $P'_1 - P'_0$ ist, bzw. umgekehrt, wenn $P'_1 - P'_0$ ein skalares Vielfaches von $P_1 - P_0$ ist. Ist Letzteres nicht der Fall, so besitzen G und G' genau einen Schnittpunkt, wobei eine Berechnung dieses Schnittpunktes auf die Lösung eines so genannten linearen Gleichungssystems führt, welches aus 2 Gleichungen mit 2 Unbekannten, nämlich den Koordinaten des Schnittpunktes von G und G' besteht. Die Lösung von Gleichungssystemen dieses Typs wird uns noch ausführlich in Kapitel 3 beschäftigen.

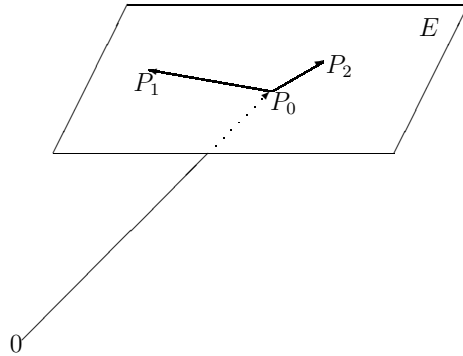
Die vorstehenden Überlegungen lassen sich ohne Probleme auf den dreidimensionalen Raum und sein Modell $\mathbb{R}^3$ verallgemeinern. Beispielsweise ist für zwei Punkte $P_0, P_1 \in \mathbb{R}^3$ wiederum

$$G = \{P_0 + t \cdot (P_1 - P_0); t \in \mathbb{R}\}$$

die durch P_0 und P_1 bestimmte Gerade im $\mathbb{R}^3$. Für Punkte $P_0, P_1, P_2 \in \mathbb{R}^3$ kann man mit $P'_1 := P_1 - P_0$ und $P'_2 := P_2 - P_0$ entsprechend das Gebilde

$$E = \{P_0 + s \cdot P'_1 + t \cdot P'_2; s, t \in \mathbb{R}\}$$

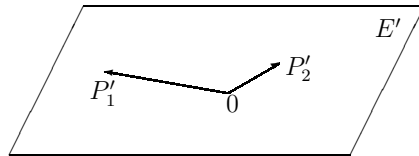
betrachten:



Wenn P'_1 kein Vielfaches von P'_2 und P'_2 kein Vielfaches von P'_1 ist, die Vektoren in 0 angetragen also nicht auf einer Geraden durch 0 liegen, so bezeichnet man P'_1 und P'_2 als *linear unabhängig*. In diesem Falle erkennt man E als Ebene, ansonsten als Gerade oder auch nur als Punkt. Da die Vektoren P'_1 und P'_2 hier eine entscheidende Rolle spielen, sollten wir auch das Gebilde

$$E' = \{s \cdot P'_1 + t \cdot P'_2; s, t \in \mathbb{R}\}$$

betrachten, welches durch Verschieben von E um den Vektor $-\overrightarrow{OP}$ entsteht:



Im Rahmen der Vektorräume nennt man E' den von P'_1 und P'_2 *aufgespannten* oder *erzeugten linearen Unterraum* von $\mathbb{R}^3$. Allgemeiner kann man im $\mathbb{R}^3$ den von beliebigen Vektoren $Q_1, \dots, Q_r$ erzeugten linearen Unterraum

$$U = \{t_1 Q_1 + \dots + t_r Q_r; t_1, \dots, t_r \in \mathbb{R}\}$$

betrachten. Für einen Vektor $Q \in \mathbb{R}^3$ sagt man, dass Q *linear von* $Q_1, \dots, Q_r$ *abhängt*, falls $Q \in U$ gilt. Folgende Fälle sind möglich: Für $Q_1 = \dots = Q_r = 0$ besteht U nur aus dem Nullpunkt 0. Ist aber einer der Vektoren $Q_1, \dots, Q_r$ von 0 verschieden, etwa $Q_1 \neq 0$, so enthält U zumindest die durch Q_1 gegebene Gerade $G = \{t Q_1; t \in \mathbb{R}\}$. Gehören auch $Q_2, \dots, Q_r$ zu G , d. h. sind $Q_2, \dots, Q_r$ linear abhängig von Q_1 , so stimmt U mit G überein. Ist Letzteres nicht der Fall und gilt etwa $Q_2 \notin G$, so spannen Q_1 und Q_2 die Ebene $E = \{t_1 Q_1 + t_2 Q_2; t_1, t_2 \in \mathbb{R}\}$ auf, so dass U zumindest diese Ebene enthält. Im Falle $Q_3, \dots, Q_r \in E$, also wenn $Q_3, \dots, Q_r$ linear von Q_1, Q_2 abhängen, stimmt U mit E überein. Ansonsten gibt es einen dieser Vektoren, etwa Q_3 , der nicht zu E gehört. Die Vektoren Q_1, Q_2, Q_3 bilden dann sozusagen ein Koordinatensystem im $\mathbb{R}^3$, und man sieht dass U mit ganz $\mathbb{R}^3$ übereinstimmt, dass

also alle Vektoren im $\mathbb{R}^3$ linear von Q_1, Q_2, Q_3 abhängen. Insbesondere ergibt sich, dass ein linearer Unterraum im $\mathbb{R}^3$ entweder aus dem Nullpunkt, aus einer Geraden durch 0, aus einer Ebene durch 0 oder aus ganz $\mathbb{R}^3$ besteht.

Das soeben eingeführte Konzept der *linearen Abhängigkeit* von Vektoren ist ein ganz zentraler Punkt, der in diesem Kapitel ausführlich im Rahmen der Vektorräume behandelt werden wird. Dabei nennt man ein System von Vektoren $Q_1, \dots, Q_r$ *linear unabhängig*, wenn keiner dieser Vektoren von den restlichen linear abhängt. Die oben durchgeführte Überlegung zeigt beispielsweise, dass linear unabhängige Systeme im $\mathbb{R}^3$ aus höchstens 3 Elementen bestehen. Insbesondere werden uns linear unabhängige Systeme, so wie wir sie im obigen Beispiel für lineare Unterräume des $\mathbb{R}^3$ konstruiert haben, gestatten, den Begriff des Koordinatensystems oder der Dimension im Kontext der Vektorräume zu präzisieren. Als Verallgemeinerung linear unabhängiger Systeme von Vektoren werden wir schließlich noch so genannte *direkte Summen* von linearen Unterräumen eines Vektorraums studieren.

Wir haben bisher im Hinblick auf Vektorräume lediglich die Modelle $\mathbb{R}^n$ mit $n = 1, 2, 3$ betrachtet, wobei unser geometrisches Vorstellungsvermögen in erheblichem Maße bei unseren Argumentationen mit eingeflossen ist. Bei der Behandlung der Vektorräume in den nachfolgenden Abschnitten werden wir jedoch grundsätzlichlicher vorgehen, indem wir eine Reihe von Verallgemeinerungen zulassen und uns bei der Entwicklung der Theorie lediglich auf gewisse axiomatische Grundlagen stützen. Zunächst beschränken wir uns bei dem zugrunde liegenden Skalarenbereich nicht auf die reellen Zahlen $\mathbb{R}$, sondern lassen beliebige *Körper* zu. Körper sind zu sehen als Zahlssysteme mit gewissen Axiomen für die Addition und Multiplikation, die im Wesentlichen den Regeln für das Rechnen mit den reellen Zahlen entsprechen. So kennt man neben dem Körper $\mathbb{R}$ der reellen Zahlen beispielsweise den Körper $\mathbb{Q}$ der rationalen Zahlen wie auch den Körper $\mathbb{C}$ der komplexen Zahlen. Es gibt aber auch Körper, die nur aus endlich vielen Elementen bestehen.

Die Axiome eines Körpers bauen auf denen einer *Gruppe* auf, denn ein Körper bildet mit seiner Addition insbesondere auch eine Gruppe. So werden wir in diesem Kapitel nach gewissen Vorbereitungen über Mengen zunächst Gruppen studieren, ausgehend von den zugehörigen Gruppenaxiomen. Wir beschäftigen uns dann weiter mit Körpern und deren Rechenregeln und gelangen anschließend zu den Vektorräumen. Vektorräume sind immer in Verbindung mit einem entsprechenden Skalarenbereich zu sehen, dem zugehörigen Körper; man spricht von einem Vektorraum über einem Körper K oder von einem K -Vektorraum. Ein K -Vektorraum V ist ausgerüstet mit einer Addition und einer skalaren Multiplikation, d. h. für $a, b \in V$ und $\alpha \in K$ sind die Summe $a + b$ sowie das skalare Produkt $\alpha \cdot a$ als Elemente von V erklärt. Addition und skalare Multiplikation genügen dabei den so genannten Vektorraumaxiomen, welche bezüglich der Addition insbesondere die Gruppenaxiome enthalten. Prototyp eines K -Vektorraums ist für eine gegebene natürliche Zahl n die Menge

$$K^n = \{(a_1, \dots, a_n); a_1, \dots, a_n \in K\}$$

aller n -Tupel mit Komponenten aus K , wobei Addition und skalare Multiplikation durch

$$\begin{aligned}(a_1, \dots, a_n) + (b_1, \dots, b_n) &:= (a_1 + b_1, \dots, a_n + b_n), \\ \alpha \cdot (a_1, \dots, a_n) &:= (\alpha a_1, \dots, \alpha a_n)\end{aligned}$$

gegeben sind.

Insbesondere wird mit dieser Definition die oben angesprochene Reihe von Modellen $\mathbb{R}^n$ für $n = 1, 2, 3$ auf beliebige *Dimensionen* n verallgemeinert. Dies hat durchaus einen realen Hintergrund, denn um beispielsweise ein Teilchen im drei-dimensionalen Raum in zeitlicher Abhängigkeit zu beschreiben, benötigt man neben den 3 räumlichen Koordinaten noch eine zusätzliche zeitliche Koordinate, so dass man sich im Grunde genommen im Vektorraum $\mathbb{R}^4$ bewegt. In analoger Weise lassen sich Paare von Punkten im drei-dimensionalen Raum als Punkte des $\mathbb{R}^6$ charakterisieren.

1.1 Mengen und Abbildungen

Normalerweise müsste man hier mit einer streng axiomatischen Begründung der Mengenlehre beginnen. Da dies jedoch einen unverhältnismäßig großen Aufwand erfordern würde, wollen wir uns an dieser Stelle mit einem naiven Standpunkt begnügen und unter einer *Menge* lediglich eine Zusammenfassung gewisser Objekte verstehen, der so genannten *Elemente* dieser Menge. Eine Menge X ist somit in eindeutiger Weise durch ihre Elemente festgelegt, wobei wir $x \in X$ schreiben, wenn x ein Element von X ist, bzw. $x \notin X$, wenn dies nicht der Fall ist. Insbesondere werden wir folgende Mengen in natürlicher Weise als gegeben annehmen:

$$\begin{aligned}\emptyset &= \text{leere Menge,} \\ \mathbb{N} &= \{0, 1, 2, \dots\} \text{ natürliche Zahlen,} \\ \mathbb{Z} &= \{0, \pm 1, \pm 2, \dots\} \text{ ganze Zahlen,} \\ \mathbb{Q} &= \{p/q; p, q \in \mathbb{Z}, q \neq 0\} \text{ rationale Zahlen,} \\ \mathbb{R} &= \text{reelle Zahlen.}\end{aligned}$$

Es sei angemerkt, dass bei einer Menge, sofern wir sie in aufzählender Weise angeben, etwa $X = \{x_1, \dots, x_n\}$, die Elemente $x_1, \dots, x_n$ nicht notwendig paarweise verschieden sein müssen. Diese Konvention gilt auch für unendliche Mengen; man vergleiche hierzu etwa die obige Beschreibung von $\mathbb{Q}$.

Um Widersprüche zu vermeiden, sind die Mengenaxiome so ausgelegt, dass die Bildung von Mengen gewissen Restriktionen unterworfen ist. Beispielsweise darf eine Menge niemals sich selbst als Element enthalten, so dass insbesondere die Gesamtheit aller Mengen nicht als Menge angesehen werden kann, da sie sich selbst als Element enthalten würde. Einen Hinweis auf die hiermit verbundene Problematik liefert das folgende Paradoxon von Russel: Wir nehmen einmal in

naiver Weise an, dass man die Gesamtheit aller Mengen, die sich nicht selbst enthalten, also

$$X = \{\text{Mengen } A \text{ mit } A \notin A\},$$

als Menge betrachten kann. Fragt man sich dann, ob $X \in X$ oder $X \notin X$ gilt, so erhält man im Falle $X \in X$ nach Definition von X sofort $X \notin X$ und im Falle $X \notin X$ entsprechend $X \in X$. Es ergibt sich also $X \in X$ und $X \notin X$ zugleich, was keinen Sinn macht.

Wichtig für die Handhabung von Mengen sind die erlaubten Prozesse der Mengenbildung, auf die wir nachfolgend eingehen.

(1) Teilmengen. – Es sei X eine Menge und $P(x)$ eine Aussage, deren Gültigkeit (wahr oder falsch) man für Elemente $x \in X$ testen kann. Dann nennt man $Y = \{x \in X; P(x) \text{ ist wahr}\}$ eine *Teilmenge* von X und schreibt $Y \subset X$. Dabei ist auch $Y = X$ zugelassen. Gilt allerdings $Y \neq X$, so nennt man Y eine *echte* Teilmenge von X . Beispielsweise ist $\mathbb{R}_{>0} := \{x \in \mathbb{R}; x > 0\}$ eine (echte) Teilmenge von $\mathbb{R}$. Für eine gegebene Menge X bilden die Teilmengen von X wiederum eine Menge, die so genannte *Potenzmenge* $\mathfrak{P}(X)$.

(2) Vereinigung und Durchschnitt. – Es sei X eine Menge und I eine Indexmenge, d. h. eine Menge, deren Elemente wir als Indizes verwenden wollen. Ist dann für jedes $i \in I$ eine Teilmenge $X_i \subset X$ gegeben, so nennt man

$$\bigcup_{i \in I} X_i := \{x \in X; \text{ es existiert ein } i \in I \text{ mit } x \in X_i\}$$

die *Vereinigung* der Mengen X_i , $i \in I$, sowie

$$\bigcap_{i \in I} X_i := \{x \in X; x \in X_i \text{ für alle } i \in I\}$$

den *Durchschnitt* dieser Mengen, wobei wir in beiden Fällen wiederum eine Teilmenge von X erhalten. Im Falle einer endlichen Indexmenge $I = \{1, \dots, n\}$ schreibt man auch $X_1 \cup \dots \cup X_n$ statt $\bigcup_{i \in I} X_i$ sowie $X_1 \cap \dots \cap X_n$ statt $\bigcap_{i \in I} X_i$. Zwei Teilmengen $X', X'' \subset X$ werden als *disjunkt* bezeichnet, wenn ihr Durchschnitt leer ist, also $X' \cap X'' = \emptyset$ gilt. Als Variante zur Vereinigung von Mengen X_i , $i \in I$, kann man deren *disjunkte Vereinigung* $\coprod_{i \in I} X_i$ bilden. Hierunter versteht man die Gesamtheit aller Elemente, die in irgendeiner der Mengen X_i enthalten sind, wobei man allerdings für verschiedene Indizes $i, j \in I$ die Elemente von X_i als verschieden von allen Elementen aus X_j ansieht.

(3) Differenz von Mengen. – Sind X_1, X_2 Teilmengen einer Menge X , so heißt

$$X_1 - X_2 := \{x \in X_1; x \notin X_2\}$$

die *Differenz* von X_1 und X_2 . Auch dies ist wieder eine Teilmenge von X , sogar von X_1 .

(4) Kartesisches Produkt von Mengen. – Es seien $X_1, \dots, X_n$ Mengen. Dann heißt

$$\prod_{i=1}^n X_i := \{(x_1, \dots, x_n); x_1 \in X_1, \dots, x_n \in X_n\}$$

das *kartesische Produkt* der Mengen $X_1, \dots, X_n$, wobei man für dieses Produkt auch die Notation $X_1 \times \dots \times X_n$ verwendet bzw. X^n , falls $X_1 = \dots = X_n = X$ gilt. Die Elemente $(x_1, \dots, x_n)$ werden als *n-Tupel* mit Komponenten $x_i \in X_i$, $i = 1, \dots, n$, bezeichnet. Es gilt genau dann $(x_1, \dots, x_n) = (x'_1, \dots, x'_n)$ für zwei *n-Tupel*, wenn man $x_i = x'_i$ für $i = 1, \dots, n$ hat. In ähnlicher Weise lässt sich für eine Indexmenge I das kartesische Produkt $\prod_{i \in I} X_i$ von gegebenen Mengen X_i , $i \in I$, bilden. Man schreibt die Elemente eines solchen Produktes als Familien $(x_i)_{i \in I}$ von Elementen $x_i \in X_i$ und meint damit Tupel, deren Einträge mittels I indiziert werden. Sind die X_i Exemplare ein und derselben Menge X , so verwendet man statt $\prod_{i \in I} X_i$ auch die Notation X^I . Eine Familie $(x_i)_{i \in \emptyset}$, welche durch die leere Indexmenge $I = \emptyset$ indiziert ist, wird als *leer* bezeichnet. Demgemäß bestehen die kartesischen Produkte $\prod_{i \in I} X_i$ und X^I im Falle $I = \emptyset$ aus genau einem Element, nämlich der leeren Familie.

Als Nächstes kommen wir auf den Begriff der Abbildung zwischen Mengen zu sprechen.

Definition 1. Eine Abbildung $f: X \longrightarrow Y$ zwischen zwei Mengen X und Y ist eine Vorschrift, welche jedem $x \in X$ ein wohlbestimmtes Element $y \in Y$ zuordnet, das dann mit $f(x)$ bezeichnet wird; man schreibt hierbei auch $x \longmapsto f(x)$. Dabei heißt X der Definitionsbereich und Y der Bild- oder Wertebereich der Abbildung f .

Zu einer Menge X gibt es stets die *identische Abbildung* $\text{id}_X: X \longrightarrow X$, $x \longmapsto x$. Im Übrigen kann man beispielsweise ein kartesisches Produkt des Typs X^I auch als Menge aller Abbildungen $I \longrightarrow X$ interpretieren.

Im Folgenden sei $f: X \longrightarrow Y$ wieder eine Abbildung zwischen zwei Mengen. Ist $g: Y \longrightarrow Z$ eine weitere Abbildung, so kann man f mit g komponieren; man erhält als Resultat die Abbildung

$$g \circ f: X \longrightarrow Z, \quad x \longmapsto g(f(x)).$$

Für Teilmengen $M \subset X$ und $N \subset Y$ bezeichnet man

$$f(M) := \{y \in Y; \text{ es existiert ein } x \in M \text{ mit } y = f(x)\}$$

als das *Bild* von M unter f sowie

$$f^{-1}(N) := \{x \in X; f(x) \in N\}$$

als das *Urbild* von N unter f ; es handelt sich hierbei um Teilmengen von Y bzw. X . Besteht N aus nur einem einzigen Element y , also $N = \{y\}$, so schreibt man $f^{-1}(y)$ anstelle von $f^{-1}(\{y\})$. Weiter nennt man f *injektiv*, wenn aus $x, x' \in X$ mit $f(x) = f(x')$ stets $x = x'$ folgt, und *surjektiv*, wenn es zu jedem $y \in Y$

ein $x \in X$ mit $f(x) = y$ gibt. Schließlich heißt f *bijektiv*, wenn f injektiv und surjektiv zugleich ist.

Man kann sagen, dass f genau dann injektiv ist, wenn das Urbild $f^{-1}(y)$ eines jeden Punktes $y \in Y$ entweder leer ist oder aus genau einem Punkt $x \in X$ besteht. Weiter ist f genau dann surjektiv, wenn für jedes $y \in Y$ das Urbild $f^{-1}(y)$ nicht leer ist. Somit ist f genau dann bijektiv, wenn für jedes Element $y \in Y$ das Urbild $f^{-1}(y)$ aus genau einem Punkt x besteht. Man kann dann zu f die so genannte *Umkehrabbildung* $g: Y \rightarrow X$ betrachten. Sie ordnet einem Punkt $y \in Y$ das eindeutig bestimmte Element $x \in f^{-1}(y)$ zu, und es gilt $g \circ f = \text{id}_X$ sowie $f \circ g = \text{id}_Y$. Zu einer Abbildung $f: X \rightarrow Y$ bezeichnet man die Umkehrabbildung, sofern diese existiert, meist mit $f^{-1}: Y \rightarrow X$.

Aufgaben

- Es seien A, B, C Teilmengen einer Menge X . Man zeige:
 - $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
 - $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
 - $A - (B \cup C) = (A - B) \cap (A - C)$
 - $A - (B \cap C) = (A - B) \cup (A - C)$
- Es sei $f: X \rightarrow Y$ eine Abbildung zwischen Mengen. Man zeige für Teilmengen $M_1, M_2 \subset X$ und $N_1, N_2 \subset Y$:
 - $f(M_1 \cup M_2) = f(M_1) \cup f(M_2)$
 - $f(M_1 \cap M_2) \subset f(M_1) \cap f(M_2)$
 - $f^{-1}(N_1 \cup N_2) = f^{-1}(N_1) \cup f^{-1}(N_2)$
 - $f^{-1}(N_1 \cap N_2) = f^{-1}(N_1) \cap f^{-1}(N_2)$
 Gilt in (ii) sogar Gleichheit?
- Es seien $X \xrightarrow{f} Y \xrightarrow{g} X$ Abbildungen von Mengen mit $g \circ f = \text{id}$. Man zeige, dass f injektiv und g surjektiv ist.
 - Gibt es eine bijektive Abbildung $\mathbb{N} \rightarrow \mathbb{Z}$?
 - Gibt es für $n \in \mathbb{N}$ eine bijektive Abbildung $\mathbb{N} \rightarrow \mathbb{N} \times \{1, \dots, n\}$?
 - Gibt es eine bijektive Abbildung $\mathbb{N} \rightarrow \mathbb{N} \times \mathbb{N}$?
 - Gibt es eine bijektive Abbildung $\mathbb{N} \rightarrow \mathbb{Q}$?
- Es sei X eine Menge und $f: X \rightarrow \mathfrak{P}(X)$ eine Abbildung von X in die zugehörige Potenzmenge. Man zeige, dass f nicht surjektiv sein kann.

1.2 Gruppen

Unter einer *inneren Verknüpfung* auf einer Menge M versteht man eine Abbildung $f: M \times M \rightarrow M$. Sie ordnet jedem Paar (a, b) von Elementen aus M ein Element $f(a, b) \in M$ zu. Um den Charakter einer Verknüpfung auch in der Notation zum Ausdruck kommen zu lassen, werden wir anstelle von

$f(a, b)$ meist $a \cdot b$ schreiben. Bei kommutativen Verknüpfungen, also solchen, die $f(a, b) = f(b, a)$ für alle $a, b \in M$ erfüllen, verwenden wir auch die additive Schreibweise $a + b$.

Definition 1. Eine Menge G mit einer inneren Verknüpfung $G \times G \longrightarrow G$, $(a, b) \longmapsto a \cdot b$, heißt eine Gruppe, wenn die folgenden Eigenschaften erfüllt sind:

- (i) Die Verknüpfung ist assoziativ, d. h. es gilt $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ für alle $a, b, c \in G$.
- (ii) Es existiert ein neutrales Element e in G , d. h. ein Element $e \in G$ mit $e \cdot a = a \cdot e = a$ für alle $a \in G$.¹
- (iii) Zu jedem $a \in G$ gibt es ein inverses Element, d. h. ein Element $b \in G$ mit $a \cdot b = b \cdot a = e$. Dabei ist e das nach (ii) existierende (eindeutig bestimmte) neutrale Element von G .

Die Gruppe heißt kommutativ oder abelsch, falls die Verknüpfung kommutativ ist, d. h. falls zusätzlich gilt:

- (iv) $a \cdot b = b \cdot a$ für alle $a, b \in G$.

In der obigen Situation sagt man gewöhnlich einfach, G sei eine Gruppe, ohne die Verknüpfung “ $\cdot$ ” explizit zu erwähnen. Beispiele für Gruppen sind:

- (1) $\mathbb{Z}$ mit der Addition “ $+$ ”
- (2) $\mathbb{Q}$ mit der Addition “ $+$ ” und $\mathbb{Q}^* := \mathbb{Q} - \{0\}$ mit der Multiplikation “ $\cdot$ ”
- (3) $\mathbb{R}$ mit der Addition “ $+$ ” und $\mathbb{R}^* := \mathbb{R} - \{0\}$ mit der Multiplikation “ $\cdot$ ”
- (4) Für eine Menge X ist die Menge $\text{Bij}(X, X)$ der bijektiven Selbstabbildungen $X \longrightarrow X$ eine Gruppe unter der Komposition von Abbildungen als Verknüpfung. Man prüft leicht nach, dass diese Gruppe nicht kommutativ ist, sofern X mindestens 3 paarweise verschiedene Elemente enthält.

Wie bereits behauptet, ist in einer Gruppe G das neutrale Element e eindeutig bestimmt. Ist nämlich $e' \in G$ ein weiteres neutrales Element, so folgt $e = e' \cdot e = e'$. Auf ähnliche Weise zeigt man, dass das zu einem Element $a \in G$ gehörige inverse Element $b \in G$ eindeutig bestimmt ist. Hat man nämlich ein weiteres inverses Element $b' \in G$ zu a , so folgt

$$b = e \cdot b = (b' \cdot a) \cdot b = b' \cdot (a \cdot b) = b' \cdot e = b'.$$

Die gerade durchgeführten Schlüsse benötigen (neben den Eigenschaften von e und b) lediglich, dass e' links-neutral ist, d. h. die Eigenschaft $e' \cdot a = a$ für alle $a \in G$ besitzt, sowie dass b' links-invers zu a ist, d. h. die Gleichung $b' \cdot a = e$ erfüllt. Entsprechend kann man für rechts-neutrale bzw. rechts-inverse Elemente schließen. In einer Gruppe stimmt daher jedes links- (bzw. rechts-) neutrale Element mit dem eindeutigen neutralen Element $e \in G$ überein, ist

¹Das neutrale Element e ist, wie wir sogleich sehen werden, durch seine definierende Eigenschaft eindeutig bestimmt.

also insbesondere auch rechts- (bzw. links-) neutral. In ähnlicher Weise sieht man, dass links-inverse Elemente auch rechts-invers bzw. rechts-inverse Elemente auch links-invers sind. Wir können sogar noch einen Schritt weitergehen und die definierenden Bedingungen einer Gruppe in diesem Sinne abschwächen:

Bemerkung 2. *Es genügt, in Definition 1 anstelle von (ii) und (iii) lediglich die Existenz eines Elementes $e \in G$ mit folgenden Eigenschaften zu fordern:*

- (ii') *e ist links-neutral in G , d. h. es gilt $e \cdot a = a$ für alle $a \in G$.*
- (iii') *Zu jedem $a \in G$ existiert ein bezüglich e links-inverses Element in G , d. h. ein Element $b \in G$ mit $b \cdot a = e$.*

Beweis. Es sei G eine Menge mit einer multiplikativ geschriebenen Verknüpfung und einem Element $e \in G$, so dass die Bedingungen (i), (ii') und (iii') erfüllt sind. Um zu sehen, dass G eine Gruppe ist, haben wir zu zeigen, dass die Bedingungen (ii) und (iii) von Definition 1 gelten. Wir zeigen zunächst für Elemente $a \in G$, dass jedes Element $b \in G$, welches links-invers zu a bezüglich e ist, auch rechts-invers zu a bezüglich e ist. Gelte also $b \cdot a = e$, und sei c ein links-inverses Element zu b , so dass also $c \cdot b = e$ gilt. Hieraus folgt

$$\begin{aligned} a \cdot b &= (e \cdot a) \cdot b = ((c \cdot b) \cdot a) \cdot b = (c \cdot (b \cdot a)) \cdot b \\ &= (c \cdot e) \cdot b = c \cdot (e \cdot b) = c \cdot b = e, \end{aligned}$$

so dass b rechts-invers zu a bezüglich e ist. Es bleibt noch zu zeigen, dass das links-neutrale Element e auch rechts-neutral ist. Sei also $a \in G$. Ist dann $b \in G$ links-invers zu a bezüglich e , so ist b , wie wir gesehen haben, auch rechts-invers zu a bezüglich e , und es folgt

$$a \cdot e = a \cdot (b \cdot a) = (a \cdot b) \cdot a = e \cdot a = a,$$

also ist e rechts-neutral. □

Gewöhnlich wird das neutrale Element e einer Gruppe G bei multiplikativer Schreibweise der Verknüpfung als *Einselement* bezeichnet, und man schreibt 1 anstelle von e . Für das inverse Element zu $a \in G$ benutzt man die Schreibweise a^{-1} . Im Übrigen ist es bei multiplikativ geschriebenen Gruppenverknüpfungen üblich, das Verknüpfungszeichen “ $\cdot$ ” zu unterdrücken, sofern dies nicht zu Verwechslungen führt. Für endlich viele Elemente $a_1, \dots, a_n \in G$ definiert man das Produkt dieser Elemente durch

$$\prod_{i=1}^n a_i := a_1 \cdot \dots \cdot a_n.$$

Eine spezielle Klammerung ist hierbei aufgrund des Assoziativgesetzes nicht notwendig; auf einen detaillierten Beweis dieser “offensichtlichen” Tatsache verzichten wir jedoch an dieser Stelle. Wir werden im Folgenden endliche Folgen $a_1, \dots, a_n \in G$ meist für Indizes $n \in \mathbb{N}$ betrachten, so dass hier insbesondere

auch der Fall $n = 0$ zugelassen ist. Es handelt sich dann um die *leere* Folge, und man erklärt das zugehörige leere Produkt durch

$$\prod_{i=1}^0 a_i := 1.$$

Wie schon gesagt, verwendet man bei kommutativen Verknüpfungen meist die additive Schreibweise. Das neutrale Element einer kommutativen Gruppe wird dann als *Nullelement* 0 geschrieben und das Inverse zu einem Element $a \in G$ als $-a$. Statt $a + (-a')$ verwendet man üblicherweise die Notation $a - a'$. Endliche Summen von Elementen $a_i \in G$, $i = 1, \dots, n$, schreibt man in der Form $\sum_{i=1}^n a_i$, wobei die leere Summe durch $\sum_{i=1}^0 a_i := 0$ definiert ist.

Definition 3. *Es sei G eine Gruppe. Eine Teilmenge $H \subset G$ heißt Untergruppe von G , wenn gilt²:*

- (i) $a, b \in H \implies ab \in H$,
- (ii) $1 \in H$,
- (iii) $a \in H \implies a^{-1} \in H$.

Ist also $H \subset G$ eine Untergruppe, so induziert die Gruppenverknüpfung $G \times G \longrightarrow G$ eine Verknüpfung $H \times H \longrightarrow H$, und H ist mit dieser Verknüpfung selbst wieder eine Gruppe. Umgekehrt, ist Letzteres der Fall, so kann man leicht zeigen, dass H eine Untergruppe von G ist. Im Übrigen sieht man sofort ein, dass eine nicht-leere Teilmenge $H \subset G$ bereits dann eine Untergruppe von G ist, wenn die Bedingung $a, b \in H \implies ab^{-1} \in H$ erfüllt ist. Eine Gruppe G enthält stets die trivialen Untergruppen $\{1\}$ und G .

Als Nächstes wollen wir einige elementare Rechenregeln für das Rechnen in Gruppen behandeln. Für Elemente $a, b, c \in G$ gilt:

- (1) $ab = ac \implies b = c$ (Kürzungsregeln)
 $ac = bc \implies a = b$
- (2) $(a^{-1})^{-1} = a$
- (3) $(ab)^{-1} = b^{-1}a^{-1}$

Zum Nachweis von (1) multipliziert man von links mit a^{-1} bzw. von rechts mit c^{-1} . Im Falle (2) schließt man wie folgt. $(a^{-1})^{-1}$ ist, wie wir gesehen haben, dasjenige eindeutig bestimmte Element in G , welches (von links oder rechts) mit a^{-1} multipliziert 1 ergibt. Wegen $a^{-1}a = 1$ ergibt sich $(a^{-1})^{-1} = a$. Entsprechend erhält man $(ab)^{-1} = b^{-1}a^{-1}$, da $(b^{-1}a^{-1})(ab) = b^{-1}(a^{-1}a)b = b^{-1}b = 1$ gilt.

Abschließend wollen wir noch eine spezielle Charakterisierung von Gruppen geben.

² Nachfolgend steht $\implies$ für die so genannte *Implikation*. Für Aussagen A und B schreibt man $A \implies B$ oder $B \Leftarrow A$, wenn B aus A folgt. Entsprechend bedeutet $A \iff B$, dass A und B äquivalent sind.

Satz 4. Eine nicht-leere Menge G mit einer Verknüpfung $(a, b) \mapsto a \cdot b$ ist genau dann eine Gruppe, wenn gilt:

(i) Die Verknüpfung ist assoziativ.

(ii) Zu $a, b \in G$ gibt es stets Elemente $x, y \in G$ mit $x \cdot a = b$ und $a \cdot y = b$.

Sind diese Bedingungen erfüllt, so sind die Elemente x, y in (ii) eindeutig durch a, b bestimmt.

Beweis. Ist G eine Gruppe, so multipliziert man die Gleichungen in (ii) von links bzw. rechts mit a^{-1} . Es folgt, dass $x = ba^{-1}$ bzw. $y = a^{-1}b$ die eindeutig bestimmten Lösungen sind. Seien nun umgekehrt die Bedingungen des Satzes erfüllt, und sei $a \in G$. Dann existiert nach (ii) ein Element $e \in G$ mit $ea = a$. Zu $b \in G$ existiert weiter ein $y \in G$ mit $ay = b$, und es folgt

$$eb = eay = ay = b,$$

also ist e links-neutral. Weiter folgt die Existenz links-inverser Elemente nach (ii). Somit ist G eine Gruppe nach Bemerkung 2. $\square$

Aufgaben

1. Für eine Menge X betrachte man die Menge $\text{Bij}(X, X)$ der bijektiven Selbstabbildungen. Man prüfe nach, dass $\text{Bij}(X, X)$ unter der Komposition von Abbildungen eine Gruppe bildet und zeige, dass diese nicht kommutativ ist, sofern X mindestens 3 verschiedene Elemente besitzt.
2. Es sei G eine Gruppe und $H \subset G$ eine Teilmenge. Man zeige, dass H genau dann eine Untergruppe von G ist, wenn die Gruppenverknüpfung von G eine Verknüpfung auf H induziert (d. h. wenn für $a, b \in H$ stets $ab \in H$ gilt) und wenn H mit dieser Verknüpfung selbst wieder eine Gruppe ist.
3. Es sei G eine Gruppe und $H \subset G$ eine Teilmenge. Man zeige, dass H genau dann eine Untergruppe von G ist, wenn gilt:
 - (i) $H \neq \emptyset$
 - (ii) $a, b \in H \implies ab^{-1} \in H$
4. Es sei G eine Gruppe mit Untergruppen $H_1, H_2 \subset G$. Man zeige, dass $H_1 \cup H_2$ genau dann eine Untergruppe von G ist, wenn $H_1 \subset H_2$ oder $H_2 \subset H_1$ gilt.
5. Für eine Gruppe G betrachte man die Abbildung $i: G \rightarrow G, g \mapsto g^{-1}$. Man zeige:
 - (i) i ist bijektiv.
 - (ii) Ist $A \subset G$ eine Teilmenge mit $i(A) \subset A$, so gilt bereits $i(A) = A$; man nennt A dann *symmetrisch*.
 - (iii) Für jede Teilmenge $A \subset G$ sind $A \cup i(A)$ und $A \cap i(A)$ symmetrisch.
6. Es sei G eine Gruppe mit $a^2 = 1$ für alle $a \in G$. Man zeige, dass G abelsch ist.
7. Es sei G eine endliche abelsche Gruppe. Dann gilt $\prod_{g \in G} g^2 = 1$.

8. Für ein $n \in \mathbb{N} - \{0\}$ betrachte man die Teilmenge $R_n = \{0, 1, \dots, n-1\} \subset \mathbb{N}$. Es sei $\pi: \mathbb{Z} \rightarrow R_n$ die Abbildung, welche einer ganzen Zahl aus $\mathbb{Z}$ jeweils deren nicht-negativen Rest bei Division durch n zuordnet. Man zeige:
- (i) Es existiert eine eindeutig bestimmte Verknüpfung $(a, b) \mapsto a + b$ auf R_n , so dass für $x, y \in \mathbb{Z}$ stets $\pi(x + y) = \pi(x) + \pi(y)$ gilt.
 - (ii) R_n ist mit dieser Verknüpfung eine abelsche Gruppe.
9. Es sei G eine Gruppe. Auf der Potenzmenge $\mathfrak{P}(G)$ betrachte man die durch

$$(A, B) \mapsto A \cdot B = \{a \cdot b \in G; a \in A, b \in B\}$$

gegebene Verknüpfung. Man zeige, dass diese Verknüpfung assoziativ ist und ein neutrales Element besitzt. Ist $\mathfrak{P}(G)$ mit dieser Verknüpfung sogar eine Gruppe? Falls nein, zu welchen Elementen $A \in \mathfrak{P}(G)$ gibt es inverse Elemente?

1.3 Körper

Ein Körper ist eine additiv geschriebene abelsche Gruppe, auf der zusätzlich eine Multiplikation mit gewissen Eigenschaften definiert ist, nach dem Vorbild der rationalen oder der reellen Zahlen. Genauer:

Definition 1. *Ein Körper ist eine Menge K mit zwei inneren Verknüpfungen, geschrieben als Addition “+” und Multiplikation “·”, so dass folgende Bedingungen erfüllt sind:*

- (i) $(a + b) + c = a + (b + c)$ für $a, b, c \in K$ (Assoziativgesetz der Addition).
- (ii) *Es existiert ein Element $0 \in K$ mit $0 + a = a$ für alle $a \in K$ (neutrales Element der Addition).*
- (iii) *Zu $a \in K$ existiert ein Element $b \in K$ mit $b + a = 0$ (inverses Element der Addition).*
- (iv) $a + b = b + a$ für $a, b \in K$ (Kommutativgesetz der Addition).
- (v) $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ für $a, b, c \in K$ (Assoziativgesetz der Multiplikation).
- (vi) *Es existiert ein Element $1 \in K$ mit $1 \cdot a = a$ für alle $a \in K$ (neutrales Element der Multiplikation).*
- (vii) *Zu $a \in K - \{0\}$ existiert ein Element $b \in K$ mit $b \cdot a = 1$ (inverses Element der Multiplikation).*
- (viii) $a \cdot b = b \cdot a$ für $a, b \in K$ (Kommutativgesetz der Multiplikation).
- (ix) $a \cdot (b + c) = a \cdot b + a \cdot c$ und $(a + b) \cdot c = a \cdot c + b \cdot c$ für $a, b, c \in K$ (Distributivgesetze).
- (x) $1 \neq 0$.

Bei den Distributivgesetzen (ix) hätten wir eigentlich auf der rechten Seite die Terme $a \cdot b$, $a \cdot c$, $b \cdot c$ jeweils in Klammern setzen müssen. Man vereinbart jedoch, dass die Multiplikation “·” Vorrang vor der Addition “+” hat, so dass Klammerungen dann entbehrlich sind. Auch sei darauf hingewiesen, dass das Multiplikationszeichen “·”, ähnlich wie im Falle von Gruppen, vielfach nicht

ausgeschrieben wird. Schließlich nennt man 0 das *Nullelement* und 1 das *Eins-element* von K .

Als Nächstes wollen wir einige simple Rechenregeln für das Rechnen in Körpern K behandeln.

(1) $0a = a0 = 0$ für $a \in K$, denn es gilt

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

(2) $(-1)a = -a$ für $a \in K$, denn

$$a + (-1)a = 1a + (-1)a = (1 - 1)a = 0a = 0.$$

(3) $(-a)b = a(-b) = -ab$, $(-a)(-b) = ab$ für $a, b \in K$; dies ergibt sich unter Benutzung von (2).

(4) Für $a, b \in K$ folgt aus $ab = 0$ bereits $a = 0$ oder $b = 0$. Denn aus $ab = 0$ mit $a \neq 0 \neq b$ würde sich sonst als Widerspruch

$$1 = abb^{-1}a^{-1} = 0b^{-1}a^{-1} = 0$$

ergeben.

Man kann also in Körpern in etwa so rechnen, wie man dies von den rationalen oder reellen Zahlen her gewohnt ist. Doch sei schon an dieser Stelle auf Unterschiede zum Vorbild vertrauter Zahlbereiche hingewiesen. Für eine natürliche Zahl $n \in \mathbb{N}$ und ein Element $a \in K$ ist es üblich, die n -fache Summe von a mit sich selbst als $n \cdot a$ zu bezeichnen, wobei dann insbesondere $n \cdot a = 0$ für $n = 0$ oder $a = 0$ gilt. Weiter setzt man $n \cdot a = (-n) \cdot (-a)$ für negative ganze Zahlen n . Es folgt jedoch aus $n \cdot a = 0$ nicht notwendig $n = 0$ oder $a = 0$, wie wir an konkreten Beispielen noch feststellen werden.

Unter Verwendung des Gruppenbegriffs lassen sich Körper in übersichtlicher Weise wie folgt charakterisieren:

Bemerkung 2. Die Bedingungen (i) - (x) in Definition 1 sind äquivalent zu den folgenden Bedingungen:

- (i) K ist eine abelsche Gruppe bezüglich der Addition.
- (ii) $K^* = K - \{0\}$ ist eine abelsche Gruppe bezüglich der Multiplikation.
- (iii) Es gelten die Distributivgesetze (ix) aus Definition 1.

Beweis. Zunächst ist klar, dass die Bedingungen (i) - (iv) aus Definition 1 diejenigen einer kommutativen additiven Gruppe sind. Weiter folgt aus obiger Regel (4), dass für einen Körper K die Teilmenge $K^* = K - \{0\}$ abgeschlossen unter der Multiplikation ist und dass mit einem Element $a \in K^*$ wegen $a \cdot a^{-1} = 1$ auch dessen inverses a^{-1} zu K^* gehört. Somit sieht man, dass K^* eine abelsche Gruppe bezüglich der Multiplikation ist, und es implizieren die Bedingungen aus Definition 1 die Bedingungen von Bemerkung 2.

Seien nun umgekehrt die Bedingungen aus Bemerkung 2 erfüllt. Um hieraus die Bedingungen von Definition 1 abzuleiten, braucht man lediglich zu wissen,

dass in der Situation von Bemerkung 2 die Beziehung $0a = 0 = a0$ für alle $a \in K$ gilt. Diese kann man jedoch mit Hilfe der Distributivgesetze auf gleiche Weise herleiten, wie wir dies bereits oben bei den Rechenregeln getan haben. $\square$

Ähnlich wie bei Gruppen hat man auch bei Körpern den Begriff des Unter- oder Teilkörpers.

Definition 3. *Es sei K ein Körper. Eine Teilmenge $L \subset K$ heißt ein Teilkörper von K , wenn gilt:*

- (i) $a, b \in L \implies a + b, a \cdot b \in L$.
- (ii) $0, 1 \in L$.
- (iii) $a \in L \implies -a \in L$.
- (iv) $a \in L, a \neq 0 \implies a^{-1} \in L$.

Es ist klar, dass eine Teilmenge $L \subset K$ genau dann ein Teilkörper von K ist, wenn Addition und Multiplikation auf K sich zu Verknüpfungen $L \times L \longrightarrow L$ einschränken und wenn L unter diesen Verknüpfungen selbst ein Körper ist. Bekannte Beispiele für Körper sind die rationalen Zahlen $\mathbb{Q}$ und die reellen Zahlen $\mathbb{R}$, wobei $\mathbb{Q}$ ein Teilkörper von $\mathbb{R}$ ist. Ein Körper enthält mindestens 2 verschiedene Elemente, nämlich das neutrale Element der Addition und das neutrale Element der Multiplikation, also 0 und 1. Andererseits gibt es aber auch einen Körper K , der aus genau 2 Elementen besteht. Man betrachte nämlich die Teilmenge $\{0, 1\} \subset \mathbb{Z}$ und setze:

$$\begin{array}{lll} 0 + 0 = 0, & 0 + 1 = 1 + 0 = 1, & 1 + 1 = 0, \\ 0 \cdot 0 = 0, & 0 \cdot 1 = 1 \cdot 0 = 0, & 1 \cdot 1 = 1. \end{array}$$

Eine Verifikation der Körperaxiome zeigt, dass diese Verknüpfungen auf $\{0, 1\}$ in der Tat die Struktur eines Körpers definieren; man bezeichnet diesen meist mit $\mathbb{F}_2$. Natürlich ist $\mathbb{F}_2$ kein Teilkörper von $\mathbb{Q}$ oder $\mathbb{R}$, denn es gilt $2 \cdot 1 = 1 + 1 = 0$, wobei 2 als natürliche Zahl, nicht aber als Element von $\mathbb{F}_2$ aufzufassen ist.

Als Nächstes wollen wir den kleinsten Teilkörper von $\mathbb{R}$ konstruieren, der $\sqrt{2}$ enthält, also diejenige positive reelle Zahl, die mit sich selbst multipliziert 2 ergibt. Dieser Körper wird üblicherweise mit $\mathbb{Q}(\sqrt{2})$ bezeichnet. Zunächst zeigen wir:

Lemma 4. $\sqrt{2} \notin \mathbb{Q}$.

Beweis. Wir führen den Beweis indirekt, also durch Widerspruch, und nehmen $\sqrt{2} \in \mathbb{Q}$ an, etwa $\sqrt{2} = p/q$ mit $p, q \in \mathbb{Z} - \{0\}$. Den Bruch p/q können wir als gekürzt annehmen. Insbesondere sind dann p und q nicht beide durch 2 teilbar. Aus der Gleichung $p^2/q^2 = 2$ ergibt sich $p^2 = 2q^2$ und damit, dass p^2 gerade ist. Da das Quadrat einer ungeraden Zahl stets ungerade ist, muss auch p gerade sein, etwa $p = 2\tilde{p}$ mit einem Element $\tilde{p} \in \mathbb{Z}$. Es folgt $2q^2 = 4\tilde{p}^2$ bzw. $q^2 = 2\tilde{p}^2$ und damit wie soeben, dass 2 ein Teiler von q ist. Damit ist 2 sowohl

ein Teiler von p wie auch von q . Dies hatten wir jedoch zuvor ausgeschlossen. Die Annahme $\sqrt{2} \in \mathbb{Q}$ führt daher zu einem Widerspruch, ist folglich nicht haltbar, und es gilt $\sqrt{2} \notin \mathbb{Q}$. $\square$

Als Folgerung erhalten wir:

Lemma 5. Für $a, b \in \mathbb{Q}$ gilt

$$a + b\sqrt{2} \neq 0 \iff a \neq 0 \text{ oder } b \neq 0.$$

Beweis. Die Implikation “ $\implies$ ” ist trivial. Um die Umkehrung “ $\impliedby$ ” zu zeigen, gehen wir wieder indirekt vor und nehmen an, es gäbe Zahlen $a, b \in \mathbb{Q}$ mit $a + b\sqrt{2} = 0$, wobei a und b nicht beide verschwinden mögen. Dann folgt notwendig $a \neq 0 \neq b$ und somit $\sqrt{2} = -ab^{-1} \in \mathbb{Q}$ im Widerspruch zu Lemma 4. $\square$

Wir definieren nun $\mathbb{Q}(\sqrt{2})$ als Teilmenge von $\mathbb{R}$ durch

$$\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2}; a, b \in \mathbb{Q}\}.$$

Satz 6. $\mathbb{Q}(\sqrt{2})$ ist ein echter Teilkörper von $\mathbb{R}$, der wiederum $\mathbb{Q}$ als echten Teilkörper enthält. Es ist $\mathbb{Q}(\sqrt{2})$ der kleinste Teilkörper von $\mathbb{R}$, der $\sqrt{2}$ enthält.

Beweis. Zunächst soll gezeigt werden, dass $\mathbb{Q}(\sqrt{2})$ ein Teilkörper von $\mathbb{R}$ ist. Um die Abgeschlossenheit von $\mathbb{Q}(\sqrt{2})$ unter der Addition und Multiplikation zu zeigen, betrachte man Elemente $a + b\sqrt{2}, a' + b'\sqrt{2} \in \mathbb{Q}(\sqrt{2})$ mit $a, b, a', b' \in \mathbb{Z}$. Dann folgt

$$\begin{aligned} (a + b\sqrt{2}) + (a' + b'\sqrt{2}) &= (a + a') + (b + b')\sqrt{2} \in \mathbb{Q}(\sqrt{2}), \\ (a + b\sqrt{2}) \cdot (a' + b'\sqrt{2}) &= (aa' + 2bb') + (ab' + a'b)\sqrt{2} \in \mathbb{Q}(\sqrt{2}), \end{aligned}$$

d. h. Bedingung (i) aus Definition 3 ist erfüllt. Dasselbe gilt für Bedingung (ii), denn $0 = 0 + 0\sqrt{2} \in \mathbb{Q}(\sqrt{2})$ und $1 = 1 + 0\sqrt{2} \in \mathbb{Q}(\sqrt{2})$. Weiter ist mit $a + b\sqrt{2}$ auch $-(a + b\sqrt{2}) = (-a) + (-b)\sqrt{2}$ als inverses Element bezüglich der Addition in $\mathbb{Q}(\sqrt{2})$ enthalten, so dass auch Bedingung (iii) aus Definition 3 erfüllt ist.

Etwas schwieriger ist Bedingung (iv) aus Definition 3 nachzuweisen. Sei $a + b\sqrt{2} \in \mathbb{Q}(\sqrt{2})$ von Null verschieden, also $a \neq 0$ oder $b \neq 0$ nach Lemma 5. Dann gilt $a - b\sqrt{2} \neq 0$, ebenfalls nach Lemma 5, und wir können schreiben:

$$\frac{1}{a + b\sqrt{2}} = \frac{a - b\sqrt{2}}{a^2 - 2b^2} = \frac{a}{a^2 - 2b^2} - \frac{b}{a^2 - 2b^2}\sqrt{2} \in \mathbb{Q}(\sqrt{2}).$$

Insgesamt ergibt sich, dass $\mathbb{Q}(\sqrt{2})$ ein Teilkörper von $\mathbb{R}$ ist, und zwar ein echter Teilkörper, da beispielsweise $\sqrt{3}$ nicht zu $\mathbb{Q}(\sqrt{2})$ gehört. Letzteres zeigt man, indem man ähnlich argumentiert wie im Beweis zu Lemma 4. Im Übrigen enthält $\mathbb{Q}(\sqrt{2})$ den Körper der rationalen Zahlen als echten Teilkörper wegen $\sqrt{2} \notin \mathbb{Q}$.

Es bleibt noch zu zeigen, dass $\mathbb{Q}(\sqrt{2})$ der kleinste Teilkörper von $\mathbb{R}$ ist, der $\sqrt{2}$ enthält. Ist zunächst K ein beliebiger Teilkörper von $\mathbb{R}$, so enthält K notwendig alle Elemente der Form $n \cdot 1$ mit $n \in \mathbb{Z}$, es gilt also $\mathbb{Z} \subset K$. Dann muss K aber auch alle Brüche der Form p/q mit $p, q \in \mathbb{Z}$, $q \neq 0$, und damit $\mathbb{Q}$ enthalten. Folglich ist $\mathbb{Q}$ der kleinste Teilkörper von $\mathbb{R}$. Gilt nun $\sqrt{2} \in K$, so enthält K notwendig auch alle Ausdrücke der Form $a + b\sqrt{2}$ mit $a, b \in \mathbb{Q}$ und damit $\mathbb{Q}(\sqrt{2})$. Also ist $\mathbb{Q}(\sqrt{2})$ der (eindeutig bestimmte) kleinste Teilkörper von $\mathbb{R}$, der $\sqrt{2}$ enthält. $\square$

Als Nächstes wollen wir von dem Körper $\mathbb{R}$ der reellen Zahlen ausgehen und diesen zum Körper $\mathbb{C}$ der komplexen Zahlen erweitern. Man setze

$$\mathbb{C} := \mathbb{R} \times \mathbb{R} = \{(a, a') ; a, a' \in \mathbb{R}\}$$

und definiere Addition bzw. Multiplikation auf $\mathbb{C}$ durch

$$\begin{aligned}(a, a') + (b, b') &:= (a + b, a' + b'), \\ (a, a') \cdot (b, b') &:= (ab - a'b', ab' + a'b).\end{aligned}$$

Man prüft leicht nach, dass $\mathbb{C}$ mit diesen Verknüpfungen einen Körper bildet. Dabei ist $0_{\mathbb{C}} = (0, 0)$ das Nullelement sowie $-(a, a') = (-a, -a')$ das inverse Element bezüglich der Addition zu $(a, a') \in \mathbb{C}$. Weiter ist $1_{\mathbb{C}} = (1, 0)$ das Einselement von $\mathbb{C}$, und das inverse Element bezüglich der Multiplikation zu einem Element $(a, a') \neq 0_{\mathbb{C}}$ wird gegeben durch

$$(a, a')^{-1} = \left(\frac{a}{a^2 + a'^2}, -\frac{a'}{a^2 + a'^2} \right).$$

Exemplarisch wollen wir das Assoziativgesetz der Multiplikation nachweisen. Für $(a, a'), (b, b'), (c, c') \in \mathbb{C}$ rechnet man

$$\begin{aligned}((a, a')(b, b'))(c, c') &= (ab - a'b', ab' + a'b)(c, c') \\ &= (abc - a'b'c - ab'c' - a'bc', abc' - a'b'c' + ab'c + a'bc)\end{aligned}$$

sowie

$$\begin{aligned}(a, a')((b, b')(c, c')) &= (a, a')(bc - b'c', bc' + b'c) \\ &= (abc - ab'c' - a'bc' - a'b'c, abc' + ab'c + a'bc - a'b'c'),\end{aligned}$$

d. h. es gilt

$$((a, a')(b, b'))(c, c') = (a, a')((b, b')(c, c')).$$

Man stellt weiter fest, dass die Elemente der Form $(a, 0)$ einen Teilkörper $K \subset \mathbb{C}$ bilden. Es gilt nämlich $0_{\mathbb{C}}, 1_{\mathbb{C}} \in K$ sowie für $(a, 0), (b, 0) \in K$

$$\begin{aligned}(a, 0) + (b, 0) &= (a + b, 0) \in K, \\ (a, 0) \cdot (b, 0) &= (a \cdot b, 0) \in K, \\ -(a, 0) &= (-a, 0) \in K, \\ (a, 0)^{-1} &= (a^{-1}, 0) \in K, \text{ falls } a \neq 0.\end{aligned}$$

Man kann nun durch $a \mapsto (a, 0)$ eine natürliche Identifikation zwischen den Elementen von $\mathbb{R}$ und denen von K erklären. Da diese Identifikation auch die Körperstrukturen von $\mathbb{R}$ bzw. K respektiert, lässt sich $\mathbb{R}$ sogar als Körper mit dem Teilkörper $K \subset \mathbb{C}$ identifizieren. Somit können wir nun $\mathbb{R}$ als Teilkörper von $\mathbb{C}$ auffassen und brauchen nicht mehr zwischen dem Null- bzw. Einselement in $\mathbb{R}$ und $\mathbb{C}$ zu unterscheiden.

Üblicherweise bezeichnet man das Element $(0, 1) \in \mathbb{C}$ als komplexe Zahl i ; diese besitzt die Eigenschaft $i^2 = -1$, ist also zu interpretieren als Quadratwurzel aus -1 . Komplexe Zahlen $z = (a, a')$ lassen sich sodann in der Form

$$z = (a, 0) + (0, a') = (a, 0) + (a', 0) \cdot (0, 1) = a + a'i$$

schreiben. Dabei wird a als *Realteil* und a' als *Imaginärteil* von z bezeichnet. Es gelten die Formeln

$$\begin{aligned}(a + a'i) + (b + b'i) &= (a + b) + (a' + b')i, \\ (a + a'i) \cdot (b + b'i) &= (ab - a'b') + (ab' + a'b)i, \\ -(a + a'i) &= -a - a'i, \\ (a + a'i)^{-1} &= \frac{a}{a^2 + a'^2} - \frac{a'}{a^2 + a'^2}i,\end{aligned}$$

letztere unter der Voraussetzung $a + a'i \neq 0$, also $a \neq 0$ oder $a' \neq 0$.

Als Beispiel für das Rechnen in Körpern wollen wir schließlich noch die binomische Formel herleiten. Sei also K ein beliebiger Körper. Für $a \in K$ und $n \in \mathbb{N}$ definiert man üblicherweise a^n als das n -fache Produkt von a mit sich selbst. Dabei ist a^0 das leere Produkt, also $a^0 = 1$. Außerdem kann man a^{-n} für $a \neq 0$ durch $(a^{-1})^n$ erklären, so dass dann a^n für ganzzahlige Exponenten n definiert ist. Für das Rechnen mit solchen Potenzen gelten die gewöhnlichen Potenzgesetze.

Seien $a, b \in K$, und sei $n \in \mathbb{N}$ eine natürliche Zahl. Zur Berechnung von $(a + b)^n$ wählen wir zunächst eine kombinatorische Methode. Hierzu stellen wir uns $(a + b)^n$ als n -faches Produkt vor:

$$(a + b)^n = (a + b) \cdot \dots \cdot (a + b)$$

Die rechte Seite kann man unter sukzessiver Benutzung der Distributivgesetze ausrechnen, indem man aus jeder Klammer einen Summanden auswählt (also jeweils a oder b), das Produkt über die ausgewählten Elemente bildet und schließlich alle Produkte dieses Typs zu verschiedenen Wahlen summiert. Somit folgt

$$(a + b)^n = \sum_{i=0}^n \alpha(i) a^{n-i} b^i,$$

wobei $\alpha(i)$ gleich der Anzahl der Möglichkeiten ist, den Summanden b genau i -mal aus den n Klammern $(a + b)$ auszuwählen, mit anderen Worten, gleich der Anzahl der i -elementigen Teilmengen in $\{1, \dots, n\}$. Will man i Elemente in $\{1, \dots, n\}$ auswählen, so gibt es für das erste Element n Wahlmöglichkeiten, für

das zweite $n - 1$ und so weiter, schließlich für das i -te Element noch $n - i + 1$ Möglichkeiten. Insgesamt haben wir daher

$$n(n-1)\dots(n-i+1)$$

Möglichkeiten für diesen Auswahlprozess. Nun ist aber zu berücksichtigen, dass eine i -elementige Teilmenge $\{t_1, \dots, t_i\}$ von $\{1, \dots, n\}$, die in einem solchen Prozess konstruiert wird, nicht davon abhängt, in welcher Reihenfolge die Elemente $t_1, \dots, t_i$ ausgewählt werden. Wir müssen daher die obige Anzahl noch durch die Anzahl der Möglichkeiten dividieren, die Elemente $t_1, \dots, t_i$ in ihrer Reihenfolge zu vertauschen, also durch die Anzahl der bijektiven Selbstabbildungen $\pi: \{1, \dots, i\} \rightarrow \{1, \dots, i\}$. Will man eine solche Abbildung π definieren, so hat man zur Festsetzung von $\pi(1)$ zunächst i Möglichkeiten, für $\pi(2)$ noch $i - 1$ Möglichkeiten usw. Die Anzahl der bijektiven Selbstabbildungen von $\{1, \dots, i\}$ ist deshalb $i! = 1 \cdot \dots \cdot i$, und es ergibt sich

$$\alpha(i) = \frac{n(n-1)\dots(n-i+1)}{1 \cdot 2 \cdot \dots \cdot i},$$

wobei man hierfür auch $\binom{n}{i}$ schreibt, also

$$\binom{n}{i} = \frac{n(n-1)\dots(n-i+1)}{1 \cdot 2 \cdot \dots \cdot i} = \frac{n!}{i!(n-i)!}, \quad 0 \leq i \leq n.$$

In den Extremfällen $i = 0$ bzw. $i = n$ erweist sich unsere Konvention bezüglich leerer Produkte als sinnvoll, es gilt $0! = 1$ sowie $\binom{n}{0} = 1 = \binom{n}{n}$ und insbesondere $\binom{0}{0} = 1$. Insgesamt folgt die bekannte *binomische Formel*

$$(a+b)^n = \sum_{i=0}^n \binom{n}{i} a^{n-i} b^i.$$

Die Koeffizienten $\binom{n}{i} \in \mathbb{N}$ werden als *Binomialkoeffizienten* bezeichnet.

Wir wollen noch einen präziseren Beweis für diese Formel geben, wobei wir die Gelegenheit nutzen, um das Prinzip der *vollständigen Induktion* zu erklären. Wenn man zeigen will, dass eine Aussage $A(n)$ für alle natürlichen Zahlen $n \in \mathbb{N}$ gültig ist, so genügt es nach diesem Prinzip, Folgendes zu zeigen:

(1) Es gilt $A(0)$ (Induktionsanfang).

(2) Für beliebiges $n \in \mathbb{N}$ kann man aus der Gültigkeit von $A(n)$ (Induktionsvoraussetzung) auf die Gültigkeit von $A(n+1)$ schließen (Induktionsschluss).

Natürlich kann man die vollständige Induktion statt bei $n = 0$ auch bei einer anderen Zahl $n = n_0 \in \mathbb{N}$ oder sogar bei einer Zahl $n = n_0 \in \mathbb{Z}$ beginnen. Führt man den Induktionsschluss dann für ganze Zahlen $n \geq n_0$ durch, so ergibt sich die Gültigkeit von $A(n)$ für alle ganzen Zahlen $n \geq n_0$. Als Variante dieses Prinzips darf man beim Induktionsschluss zum Nachweis von $A(n+1)$ zusätzlich benutzen, dass die Aussage $A(m)$ bereits für alle m mit $n_0 \leq m \leq n$ gilt, wobei

der Induktionsanfang wiederum bei $n = n_0$ liegen möge. In unserem Fall soll die Aussage $A(n)$ aus zwei Teilen bestehen und für $n \in \mathbb{N}$ wie folgt lauten:

$$\binom{n}{i} \in \mathbb{N} \quad \text{für } 0 \leq i \leq n,$$

$$(a+b)^n = \sum_{i=0}^n \binom{n}{i} a^{n-i} b^i;$$

die Binomialkoeffizienten $\binom{n}{i}$ sind dabei wie oben durch

$$\binom{n}{i} = \frac{n(n-1) \dots (n-i+1)}{1 \cdot 2 \cdot \dots \cdot i} = \frac{n!}{i!(n-i)!}$$

gegeben. Der Induktionsanfang bei $n = 0$ ist leicht durchzuführen; denn man hat $\binom{0}{0} = 1 \in \mathbb{N}$ und $(a+b)^0 = 1 = \binom{0}{0} a^0 b^0$, d. h. $A(0)$ ist richtig. Zum Induktionsschluss betrachten wir ein beliebiges $n \in \mathbb{N}$ und nehmen an, dass $A(n)$ richtig ist. Dann können wir wie folgt rechnen:

$$\begin{aligned} (a+b)^{n+1} &= (a+b)(a+b)^n = (a+b) \sum_{i=0}^n \binom{n}{i} a^{n-i} b^i \\ &= \sum_{i=0}^n \binom{n}{i} a^{n+1-i} b^i + \sum_{i=0}^n \binom{n}{i} a^{n-i} b^{i+1} \\ &= a^{n+1} + \sum_{i=1}^n \binom{n}{i} a^{n+1-i} b^i + \sum_{i=0}^{n-1} \binom{n}{i} a^{n-i} b^{i+1} + b^{n+1} \\ &= a^{n+1} + \sum_{i=1}^n \binom{n}{i} a^{n+1-i} b^i + \sum_{i=1}^n \binom{n}{i-1} a^{n+1-i} b^i + b^{n+1} \\ &= a^{n+1} + \sum_{i=1}^n \left[\binom{n}{i} + \binom{n}{i-1} \right] a^{n+1-i} b^i + b^{n+1} \end{aligned}$$

Nun hat man aber für $1 \leq i \leq n$

$$\begin{aligned} \binom{n}{i} + \binom{n}{i-1} &= \frac{n!}{i!(n-i)!} + \frac{n!}{(i-1)!(n-i+1)!} \\ &= \frac{n!(n-i+1) + n!i}{i!(n-i+1)!} = \frac{n!(n+1)}{i!(n-i+1)!} \\ &= \frac{(n+1)!}{i!(n+1-i)!} = \binom{n+1}{i}, \end{aligned}$$

so dass sich wie gewünscht

$$(a+b)^{n+1} = \sum_{i=0}^{n+1} \binom{n+1}{i} a^{n+1-i} b^i$$

ergibt. Außerdem folgt aus $\binom{n}{i}, \binom{n}{i-1} \in \mathbb{N}$, dass auch $\binom{n+1}{i}$ eine natürliche Zahl ist. Die binomische Formel ist daher per Induktion bewiesen.

Aufgaben

1. Es sei K eine endliche Menge mit zwei Verknüpfungen “+” und “·”, welche den Bedingungen (i) – (x) von Definition 1 genügen, wobei jedoch die Bedingung (vii) ersetzt sei durch

(vii') Für $a, b \in K - \{0\}$ gilt $ab \in K - \{0\}$.

Man zeige, dass K ein Körper ist.

2. Es sei K ein endlicher Körper. Für $n \in \mathbb{N}$ und $a \in K$ bezeichne $na = a + \dots + a$ die n -fache Summe von a mit sich selber.

(i) Es existiert ein $n \in \mathbb{N} - \{0\}$, so dass $na = 0$ für alle $a \in K$ gilt.

(ii) Wählt man n wie vorstehend minimal, so ist n eine Primzahl, die so genannte *Charakteristik* von K .

3. Man betrachte für $n \in \mathbb{N} - \{0\}$ die Menge R_n aus Abschnitt 1.2, Aufgabe 8 mit der dort erklärten Addition, welche auf R_n die Struktur einer additiven abelschen Gruppe definiert. Man zeige:

(i) Auf R_n lässt sich in eindeutiger Weise eine Multiplikation erklären, so dass alle Bedingungen von Definition 1, mit eventueller Ausnahme von (vii) erfüllt sind.

(ii) Ist p eine Primzahl, so ist R_p sogar ein Körper; dieser wird auch mit $\mathbb{F}_p$ bezeichnet.

4. Man konstruiere einen Körper mit 4 Elementen.

5. Man weise nach, dass $\sqrt{3}$ nicht zu $\mathbb{Q}(\sqrt{2})$ gehört.

6. Man bestimme den kleinsten Teilkörper von $\mathbb{C}$, welcher die komplexe Zahl i enthält.

7. Für eine Aussage $A(n)$, die für $n \in \mathbb{N}$ definiert ist, betrachte man folgende Bedingungen:

(i) $A(0)$ ist wahr.

(ii) Für alle $n \in \mathbb{N}$ gilt: Ist $A(n)$ wahr, so auch $A(n+1)$.

(iii) Für alle $n \in \mathbb{N}$ gilt: Ist $A(i)$ für alle $i \in \mathbb{N}$ mit $i \leq n$ wahr, so auch $A(n+1)$.

Man zeige mittels eines formalen Schlusses, dass das Induktionsprinzip, welches die Bedingungen (i) und (ii) umfasst, äquivalent zu demjenigen ist, das die Bedingungen (i) und (iii) umfasst.

8. Es sei $A(m, n)$ eine Aussage, die für $m, n \in \mathbb{N}$ erklärt sei. Die folgenden Aussagen seien wahr:

(i) $A(0, 0)$

(ii) $A(i, j) \implies A(i+1, j)$ für $i, j \in \mathbb{N}$.

(iii) $A(i, j) \implies A(i, j+1)$ für $i, j \in \mathbb{N}$.

Man zeige, dass dann $A(i, j)$ für alle $i, j \in \mathbb{N}$ wahr ist (*Prinzip der Doppelinduktion*). Lassen sich die Bedingungen (ii) bzw. (iii) noch abschwächen?

9. Für $n \in \mathbb{N}$ und Elemente $q \neq 1$ eines Körpers K leite man die Formel für die *geometrische Reihe* her:

$$\sum_{i=0}^n q^i = \frac{1 - q^{n+1}}{1 - q}$$

10. Man beweise für $k, n \in \mathbb{N}$ mit $n \geq k \geq 1$:

$$\sum_{i=k-1}^{n-1} \binom{i}{k-1} = \binom{n}{k}$$

11. Für $k, n \in \mathbb{N}$ zeige man, dass die Menge $\{(a_1, \dots, a_n) \in \mathbb{N}^n; a_1 + \dots + a_n = k\}$ genau

$$\binom{k+n-1}{n-1}$$

Elemente besitzt.

1.4 Vektorräume

Wir wollen nun die eingangs angedeutete Vektorrechnung auf eine axiomatische Grundlage stellen, indem wir Vektorräume über Körpern betrachten. Vektoren werden wir im Folgenden stets mit lateinischen Buchstaben $a, b, c, \dots$ bezeichnen, Skalare aus dem zugehörigen Körper dagegen mit griechischen Buchstaben $\alpha, \beta, \gamma, \dots$

Definition 1. *Es sei K ein Körper. Ein K -Vektorraum ist eine Menge V mit einer inneren Verknüpfung $V \times V \longrightarrow V$, $(a, b) \longmapsto a + b$, genannt Addition, und einer äußeren Verknüpfung $K \times V \longrightarrow V$, genannt skalare Multiplikation, so dass gilt:*

- (i) V ist eine abelsche Gruppe bezüglich der Addition “+”.
- (ii) $(\alpha + \beta) \cdot a = \alpha \cdot a + \beta \cdot a$, $\alpha \cdot (a + b) = \alpha \cdot a + \alpha \cdot b$ für alle $\alpha, \beta \in K$, $a, b \in V$, d. h. Addition und Multiplikation verhalten sich distributiv.
- (iii) $(\alpha \cdot \beta) \cdot a = \alpha \cdot (\beta \cdot a)$ für alle $\alpha, \beta \in K$, $a \in V$, d. h. die skalare Multiplikation ist assoziativ.
- (iv) $1 \cdot a = a$ für das Einselement $1 \in K$ und alle $a \in V$.

Elemente eines Vektorraums werden auch als Vektoren bezeichnet. Wie jede Gruppe enthält ein K -Vektorraum mindestens ein Element, nämlich den Nullvektor 0 als neutrales Element. Andererseits kann man eine einelementige Menge $V = \{0\}$ stets zu einem K -Vektorraum machen, indem man $0 + 0 = 0$ und $\alpha \cdot 0 = 0$ für $\alpha \in K$ definiert. Man nennt V dann den *Nullraum* und schreibt in suggestiver Weise $V = 0$, wobei man streng genommen zwischen 0 als Nullelement und 0 als Nullraum zu unterscheiden hat. Ist L ein Körper

und K ein Teilkörper, so kann man L stets als K -Vektorraum auffassen. Als Vektorraumaddition auf L nehme man die gegebene Körperaddition und als skalare Multiplikation $K \times L \rightarrow L$ die Einschränkung der Körpermultiplikation $L \times L \rightarrow L$. Insbesondere ist $\mathbb{C}$ auf diese Weise ein Vektorraum über $\mathbb{Q}$, $\mathbb{Q}(\sqrt{2})$ oder $\mathbb{R}$. Im Übrigen ist jeder Körper K ein Vektorraum über sich selbst.

Für das Rechnen mit Vektoren gelten die gewöhnlichen Rechenregeln, die wir im Folgenden auflisten. Dabei haben wir an dieser Stelle der Deutlichkeit halber 0_K für das Nullelement von K und 0_V für den Nullvektor in V geschrieben, eine Unterscheidung, die wir im Weiteren allerdings nicht mehr machen werden.

- (1) $\alpha \cdot 0_V = 0_V$ für alle $\alpha \in K$.
- (2) $0_K \cdot a = 0_V$ für alle $a \in V$.
- (3) $(-\alpha) \cdot a = \alpha \cdot (-a) = -\alpha \cdot a$ für alle $\alpha \in K, a \in V$.
- (4) Aus $\alpha \cdot a = 0_V$ für $\alpha \in K$ und $a \in V$ folgt bereits $\alpha = 0_K$ oder $a = 0_V$.

Die Regeln (1) - (3) beweist man genauso wie die entsprechenden Regeln für das Rechnen in Körpern. Gleiches gilt für (4), wobei wir hier die Argumentation noch einmal ausführen wollen. Gilt nämlich $\alpha \cdot a = 0$ mit $\alpha \neq 0$, so ergibt sich

$$a = (\alpha^{-1} \cdot \alpha) \cdot a = \alpha^{-1} \cdot (\alpha \cdot a) = \alpha^{-1} \cdot 0_V = 0_V.$$

Als weitere Regeln führen wir noch die allgemeinen Distributivgesetze auf; es seien $\alpha, \alpha_i, \beta_i \in K$ sowie $a, a_i \in V$ für $i = 1, \dots, n$.

$$\begin{aligned} \alpha \cdot \sum_{i=1}^n a_i &= \sum_{i=1}^n \alpha a_i \\ \left(\sum_{i=1}^n \alpha_i \right) \cdot a &= \sum_{i=1}^n \alpha_i a \\ \sum_{i=1}^n \alpha_i a_i + \sum_{i=1}^n \beta_i a_i &= \sum_{i=1}^n (\alpha_i + \beta_i) a_i \end{aligned}$$

Definition 2. Es sei V ein K -Vektorraum. Eine Teilmenge $U \subset V$ heißt ein K -Untervektorraum oder linearer Unterraum von V , wenn gilt:

- (i) $U \neq \emptyset$
- (ii) $a, b \in U \implies a + b \in U$
- (iii) $\alpha \in K, a \in U \implies \alpha a \in U$

Für einen Vektor $a \in V$ ist

$$K \cdot a := \{\alpha a; \alpha \in K\}$$

stets ein linearer Unterraum von V . In Falle $a \neq 0$ kann man hier von einer "Geraden" sprechen, für $a = 0$ ist $K \cdot a$ der Nullraum. Jeder Vektorraum enthält folglich den Nullraum und sich selbst als lineare Unterräume. Fassen wir weiter

etwa $\mathbb{C}$ als $\mathbb{Q}$ -Vektorraum auf, so erkennt man $\mathbb{R}$ und $\mathbb{Q}(\sqrt{2})$ als lineare Unterräume. Im Übrigen ist die Bezeichnung K -Untervektorraum in Definition 2 gerechtfertigt, denn es gilt:

Bemerkung 3. *Eine Teilmenge U eines K -Vektorraumes V ist genau dann ein K -Untervektorraum, wenn U abgeschlossen unter der Addition und der skalaren Multiplikation mit Elementen aus K ist, und wenn U mit diesen Verknüpfungen selbst ein K -Vektorraum ist.*

Beweis. Die behauptete Äquivalenz ist in einfacher Weise zu verifizieren. Wir wollen hier nur zeigen, dass jeder lineare Unterraum $U \subset V$ die in Bemerkung 3 genannten Bedingungen erfüllt. Sei also $U \subset V$ wie in Definition 2. Zunächst besagen die Bedingungen (ii) und (iii), dass U abgeschlossen unter der Addition und der skalaren Multiplikation ist. Weiter übertragen sich allgemeine Eigenschaften der Verknüpfungen wie Assoziativität, Kommutativität, Distributivität usw. in direkter Weise von V auf U . Nach Voraussetzung gilt $U \neq \emptyset$. Es enthält U daher ein Element a . Dann gehört auch $-a = (-1)a$ zu U und damit der Nullvektor $0 = a - a$. Also ist klar, dass U eine additive Untergruppe von V und insgesamt mit den von V induzierten Verknüpfungen ein K -Vektorraum ist. $\square$

Als wichtigstes Beispiel eines Vektorraums über einem Körper K wollen wir das n -fache kartesische Produkt

$$K^n = \{(\alpha_1, \dots, \alpha_n) ; \alpha_i \in K \text{ für } i = 1, \dots, n\}$$

betrachten, wobei $n \in \mathbb{N}$ sei. Die Addition $K^n \times K^n \longrightarrow K^n$ werde erklärt durch

$$(\alpha_1, \dots, \alpha_n) + (\beta_1, \dots, \beta_n) = (\alpha_1 + \beta_1, \dots, \alpha_n + \beta_n),$$

sowie die skalare Multiplikation $K \times K^n \longrightarrow K^n$ durch

$$\alpha \cdot (\alpha_1, \dots, \alpha_n) = (\alpha \cdot \alpha_1, \dots, \alpha \cdot \alpha_n).$$

Das n -Tupel $(0, \dots, 0) \in K^n$ definiert dann den Nullvektor in K^n , den wir üblicherweise wieder mit 0 bezeichnen, und es ist $(-\alpha_1, \dots, -\alpha_n)$ das inverse Element bezüglich der Addition zu einem Element $(\alpha_1, \dots, \alpha_n) \in K^n$. Im Falle $n = 0$ ist K^n als einelementige Menge anzusehen, welche nur aus dem leeren Tupel besteht; K^0 ist somit der Nullraum. Weiter lässt sich K^m für $m \leq n$ in kanonischer³ Weise als linearer Unterraum von K^n auffassen, indem man die Elemente $(\alpha_1, \dots, \alpha_m) \in K^m$ mit denen des Typs $(\alpha_1, \dots, \alpha_m, 0, \dots, 0) \in K^n$ identifiziert.

Anschaulich können wir den Vektorraum K^n für $K = \mathbb{R}$ und $n = 2$ als Modell einer Ebene und für $n = 3$ als Modell des gewöhnlichen dreidimensionalen

³ Die Bezeichnung "kanonisch" werden wir im Folgenden noch häufiger verwenden. Wir meinen hiermit eine Möglichkeit, die sich in nahe liegender Weise als die einfachste Lösung anbietet.

Raumes ansehen. Als Untervektorräume der Ebene $\mathbb{R}^2$ gibt es, wie wir noch sehen werden, außer den trivialen linearen Unterräumen 0 und $\mathbb{R}^2$ lediglich die Geraden des Typs $\mathbb{R}a$ zu von Null verschiedenen Vektoren $a \in \mathbb{R}^2$.

Die obige Konstruktion des Vektorraums K^n lässt sich allgemeiner für einen K -Vektorraum W anstelle von K durchführen. Man erhält dann das n -fache kartesische Produkt W^n von W als K -Vektorraum mit komponentenweiser Addition und skalarer Multiplikation. Darüber hinaus kann man für eine beliebige Familie von K -Vektorräumen $(V_i)_{i \in I}$ das kartesische Produkt $V = \prod_{i \in I} V_i$ als K -Vektorraum auffassen, wiederum mit komponentenweisen Verknüpfungen, indem man also für $\alpha \in K$ und $(v_i)_{i \in I}, (v'_i)_{i \in I} \in V$ setzt:

$$(v_i)_{i \in I} + (v'_i)_{i \in I} = (v_i + v'_i)_{i \in I}, \quad \alpha \cdot (v_i)_{i \in I} = (\alpha \cdot v_i)_{i \in I}.$$

Viele interessante Vektorräume sind als Räume von Abbildungen oder Funktionen zu sehen. Sei etwa K ein Körper und X eine Menge. Dann bildet die Menge $V = \text{Abb}(X, K)$ aller Abbildungen von X nach K auf natürliche Weise einen K -Vektorraum. Man erkläre nämlich die Summe zweier Elemente $f, g \in V$ als Abbildung

$$f + g: X \longrightarrow K, \quad x \longmapsto f(x) + g(x),$$

sowie das skalare Produkt eines Elementes $\alpha \in K$ mit einem Element $f \in V$ durch

$$\alpha f: X \longrightarrow K, \quad x \longmapsto \alpha f(x).$$

Es ist leicht nachzurechnen, dass V mit diesen Verknüpfungen einen K -Vektorraum bildet, den so genannten *Vektorraum der K -wertigen Funktionen auf X* (der im Übrigen mit dem kartesischen Produkt K^X übereinstimmt, dessen Faktoren K durch die Elemente der Menge X parametrisiert werden). Die Nullabbildung

$$0: X \longrightarrow K, \quad x \longmapsto 0$$

ist das Nullelement, und das negative Element zu einem $f \in V$ wird gegeben durch

$$-f: X \longrightarrow K, \quad x \longmapsto -(f(x)).$$

Setzt man beispielsweise $K = \mathbb{R}$ und $X = \{\alpha \in \mathbb{R}; 0 \leq \alpha \leq 1\}$, so ist $V = \text{Abb}(X, \mathbb{R})$ der $\mathbb{R}$ -Vektorraum aller reellwertigen Funktionen auf dem Einheitsintervall in $\mathbb{R}$. Lineare Unterräume werden gebildet von den stetigen Funktionen, den differenzierbaren Funktionen bzw. von den Polynomen.

Im Folgenden sei K stets ein Körper. Wir wollen uns etwas genauer mit dem Problem der Konstruktion von linearen Unterräumen in einem K -Vektorraum V beschäftigen.

Lemma 4. *Es sei V ein K -Vektorraum und $(U_i)_{i \in I}$ eine Familie von linearen Unterräumen. Dann ist $U = \bigcap_{i \in I} U_i$ ebenfalls ein linearer Unterraum von V .*

Beweis. Um zu sehen, dass U ein linearer Unterraum von V ist, verifizieren wir die Bedingungen von Definition 2. Aus $0 \in U_i$ für alle i folgt $0 \in U$. Seien nun

$\alpha \in K$ und $a, b \in U$. Dann ergibt sich $a, b \in U_i$ für alle i , also $a + b, \alpha a \in U_i$ und somit $a + b, \alpha a \in U$. Folglich erfüllt U die definierenden Eigenschaften eines linearen Unterraums von V . $\square$

Satz und Definition 5. *Es sei V ein K -Vektorraum und $A \subset V$ eine Teilmenge. Dann ist*

$$\langle A \rangle := \left\{ \sum_{i=1}^r \alpha_i a_i ; r \in \mathbb{N}, \alpha_i \in K, a_i \in A \text{ für } i = 1, \dots, r \right\}$$

ein linearer Unterraum von V , und dieser stimmt überein mit dem linearen Unterraum

$$\bigcap_{A \subset U} U \subset V,$$

den man gemäß Lemma 4 erhält, wenn man den Durchschnitt über alle linearen Unterräume U in V bildet, die A enthalten.

Folglich ist $\langle A \rangle$ der kleinste lineare Unterraum in V , der A enthält, was bedeutet, dass jeder lineare Unterraum $U \subset V$, der A enthält, auch bereits $\langle A \rangle$ enthalten muss. Man nennt $\langle A \rangle$ den von A in V erzeugten linearen Unterraum oder auch die lineare Hülle von A in V .

In ähnlicher Weise definiert man für eine Familie $\mathfrak{A} = (a_i)_{i \in I}$ von Elementen aus V den von $\mathfrak{A}$ erzeugten linearen Unterraum $\langle \mathfrak{A} \rangle \subset V$ durch $\langle \mathfrak{A} \rangle = \langle A \rangle$ mit $A = \{a_i; i \in I\}$. Aus der Definition und obigem Satz ergeben sich in direkter Weise die folgenden elementaren Eigenschaften für erzeugte lineare Unterräume in einem Vektorraum V :

- (1) $\langle \emptyset \rangle = 0$
- (2) $A \subset \langle A \rangle$ für eine Teilmenge $A \subset V$.
- (3) $\langle U \rangle = U$ für einen linearen Unterraum $U \subset V$.
- (4) $A \subset B \implies \langle A \rangle \subset \langle B \rangle$ und $A \subset \langle B \rangle \implies \langle A \rangle \subset \langle B \rangle$ für Teilmengen $A, B \subset V$.

Nun zum *Beweis* von Satz 5. Wir zeigen zunächst, dass $\langle A \rangle$ ein linearer Unterraum von V ist. Es gilt $\langle A \rangle \neq \emptyset$, denn der Nullvektor 0 lässt sich als leere Summe $\sum_{i=1}^0 \alpha_i a_i$ schreiben (oder für $A \neq \emptyset$ auch als entsprechende echte Summe mit Koeffizienten $\alpha_i = 0$), gehört also zu $\langle A \rangle$. Seien weiter $\alpha \in K$ sowie

$$a = \sum_{i=1}^r \alpha_i a_i, \quad b = \sum_{j=1}^s \beta_j b_j$$

Elemente von $\langle A \rangle$. Dann folgt

$$\alpha a = \sum_{i=1}^r (\alpha \alpha_i) a_i \in \langle A \rangle$$

sowie

$$a + b = \sum_{i=1}^r \alpha_i a_i + \sum_{j=1}^s \beta_j b_j = \sum_{i=1}^{r+s} \alpha_i a_i \in \langle A \rangle,$$

wenn wir $\alpha_{r+j} = \beta_j$ und $a_{r+j} = b_j$ für $j = 1, \dots, s$ setzen. Somit ist $\langle A \rangle$ ein linearer Unterraum von V .

Ist U ein beliebiger linearer Unterraum von V , der A enthält, so muss U aufgrund der definierenden Eigenschaften eines linearen Unterraums auch alle Linearkombinationen $\sum_{i=1}^r \alpha_i a_i$ mit Elementen $a_1, \dots, a_r \in A$ und Koeffizienten $\alpha_1, \dots, \alpha_r \in K$ enthalten. Somit ergibt sich $\langle A \rangle \subset U$ und damit $\langle A \rangle \subset \bigcap_{A \subset U} U$. Andererseits schließt man aus der Gleichung $a = 1 \cdot a$ für $a \in A$ natürlich $A \subset \langle A \rangle$, so dass auch $\langle A \rangle$ zu der Menge aller linearen Unterräume $U \subset V$ gehört, die A enthalten. Insbesondere ergibt sich $\langle A \rangle = \bigcap_{A \subset U} U$, und man erkennt $\langle A \rangle$ als kleinsten linearen Unterraum von V , der A enthält. $\square$

Definition 6. Es sei V ein K -Vektorraum. Eine Familie $\mathfrak{A} = (a_i)_{i \in I}$ von Elementen aus V heißt ein Erzeugendensystem von V , wenn jedes $a \in V$ eine Darstellung $a = \sum_{i \in I} \alpha_i a_i$ mit Koeffizienten $\alpha_i \in K$ besitzt, wobei $\alpha_i = 0$ für fast alle $i \in I$ gilt, d. h. für alle $i \in I$, bis auf endlich viele Ausnahmen. Mit anderen Worten, $\mathfrak{A}$ ist ein Erzeugendensystem von V , wenn $V = \langle \mathfrak{A} \rangle$ gilt. Weiter nennt man V endlich erzeugt, wenn V ein endliches Erzeugendensystem $a_1, \dots, a_n$ besitzt.

Jeder K -Vektorraum V besitzt ein Erzeugendensystem, denn es gilt beispielsweise $\langle V \rangle = V$. Weiter gilt:

$V = \langle 1 \rangle$ für $V = \mathbb{Q}$ als $\mathbb{Q}$ -Vektorraum,

$V = \langle 1, \sqrt{2} \rangle$ für $V = \mathbb{Q}(\sqrt{2})$ als $\mathbb{Q}$ -Vektorraum,

$V = \langle 1, i \rangle$ für $V = \mathbb{C}$ als $\mathbb{R}$ -Vektorraum,

$V = \langle e_1, \dots, e_n \rangle$ für $V = K^n$ als K -Vektorraum.

Dabei sei $e_i \in K^n$ für $i = 1, \dots, n$ der i -te Einheitsvektor, also

$$e_i = (0, \dots, 0, 1, 0, \dots, 0),$$

wobei die 1 genau an der i -ten Stelle steht. Auf präzisere Weise können wir

$$e_i = (\delta_{1i}, \dots, \delta_{ni})$$

schreiben mit

$$\delta_{hi} = \begin{cases} 1 & \text{für } h = i \\ 0 & \text{sonst} \end{cases} ;$$

δ_{hi} ist das so genannte *Kronecker-Symbol*.

Aufgaben

K sei stets ein Körper.

1. Es sei V ein K -Vektorraum und $U \subset V$ ein linearer Unterraum. Für welche Elemente $a \in V$ ist $a + U := \{a + u; u \in U\}$ wiederum ein linearer Unterraum von V ?
2. Es sei V ein K -Vektorraum und $\mathfrak{A} = (A_i)_{i \in I}$ eine Familie von Teilmengen von V . Die Familie $\mathfrak{A}$ möge folgende Bedingung erfüllen: Zu je zwei Indizes $i, j \in I$ existiert stets ein Index $k \in I$ mit $A_i \cup A_j \subset A_k$. Man zeige

$$\langle \bigcup_{i \in I} A_i \rangle = \bigcup_{i \in I} \langle A_i \rangle.$$

Gilt diese Beziehung auch ohne die Voraussetzung an die Familie $\mathfrak{A}$?

3. Es sei V ein endlich erzeugter K -Vektorraum. Dann lässt sich jedes beliebige Erzeugendensystem von V zu einem endlichen Erzeugendensystem verkleinern.
4. Es sei K Teilkörper eines Körpers L und V ein L -Vektorraum. Ist dann $x_1, \dots, x_n$ ein Erzeugendensystem von V als L -Vektorraum und $\alpha_1, \dots, \alpha_m$ ein Erzeugendensystem von L , aufgefasst als K -Vektorraum, so bilden die Produkte $\alpha_i x_j$ mit $i = 1, \dots, m$ und $j = 1, \dots, n$ ein Erzeugendensystem von V als K -Vektorraum.
5. Es seien $x, y \in \mathbb{R}^2$ Punkte, die nicht gemeinsam auf einer Geraden durch den Nullpunkt $0 \in \mathbb{R}^2$ liegen, d. h. es gelte $x \neq 0 \neq y$ sowie $\alpha x \neq \beta y$ für alle $\alpha, \beta \in \mathbb{R}^*$. Man zeige, dass x, y bereits ein Erzeugendensystem von $\mathbb{R}^2$ bilden. Gilt eine entsprechende Aussage auch, wenn man $\mathbb{R}$ durch einen beliebigen Körper K ersetzt?
6. Man betrachte das kartesische Produkt $\mathbb{Q}^{\mathbb{N}} = \prod_{i \in \mathbb{N}} \mathbb{Q}$ als $\mathbb{Q}$ -Vektorraum. Kann dieser Vektorraum ein abzählbares Erzeugendensystem besitzen, d. h. ein Erzeugendensystem des Typs $(x_i)_{i \in \mathbb{N}}$?

1.5 Linear unabhängige Systeme und Basen von Vektorräumen

Sind $a_1, \dots, a_n$ Vektoren eines K -Vektorraums V , so sagt man, wie bereits in den Vorbemerkungen erwähnt, a_n *hänge linear von* $a_1, \dots, a_{n-1}$ *ab*, wenn es Koeffizienten $\alpha_1, \dots, \alpha_{n-1} \in K$ mit $a_n = \sum_{i=1}^{n-1} \alpha_i a_i$ gibt, wenn also $a_n \in \langle a_1, \dots, a_{n-1} \rangle$ gilt. Man sagt in diesem Falle auch, a_n lasse sich aus den Vektoren $a_1, \dots, a_{n-1}$ *linear kombinieren* oder a_n sei eine *Linearkombination* von $a_1, \dots, a_{n-1}$. Wenn man für ein System von Vektoren $a_1, \dots, a_n$ weiß, dass irgendeiner dieser Vektoren von den übrigen linear abhängt, so bezeichnet man das System gemeinhin als *linear abhängig*. (System ist hier im Sinne von Familie gemeint; das System der $a_1, \dots, a_n$ wäre präziser als Familie $(a_i)_{i=1 \dots n}$ zu notieren.) Andererseits heißt das System der $a_1, \dots, a_n$ *linear unabhängig*, wenn keiner dieser Vektoren von den übrigen linear abhängt. Der Begriff der linearen Abhängigkeit bzw. Unabhängigkeit von Vektoren ist in der Linearen Algebra von fundamentaler

Wichtigkeit. Für eine formelmäßige Handhabung dieses Begriffes ist folgende (äquivalente) Definition besonders geeignet, auf die wir uns im Weiteren stets stützen werden.

Definition 1. Ein System von Vektoren $a_1, \dots, a_n$ eines K -Vektorraums V heißt linear unabhängig, wenn aus einer Gleichung $\sum_{i=1}^n \alpha_i a_i = 0$ mit Koeffizienten $\alpha_1, \dots, \alpha_n \in K$ notwendig $\alpha_1 = \dots = \alpha_n = 0$ folgt, wenn sich also der Nullvektor $0 \in V$ nur in trivialer Weise als Linearkombination der Vektoren $a_1, \dots, a_n$ darstellen lässt. Ist diese Bedingung nicht gegeben, so bezeichnet man das System $a_1, \dots, a_n$ als linear abhängig.

Ein System von Vektoren $a_1, \dots, a_n$ ist also genau dann linear abhängig, wenn es Koeffizienten $\alpha_1, \dots, \alpha_n \in K$ mit $\sum_{i=1}^n \alpha_i a_i = 0$ gibt, wobei die α_i nicht sämtlich verschwinden. Dies ist äquivalent zu der bereits oben erwähnten Bedingung, dass einer der Vektoren $a_1, \dots, a_n$ eine Linearkombination der restlichen ist, denn die Gleichung $\sum_{i=1}^n \alpha_i a_i = 0$ ist für $\alpha_{i_0} \neq 0$ äquivalent zu $a_{i_0} = -\sum_{i \neq i_0} \alpha_i^{-1} \alpha_i a_i$. Beispielsweise bildet der Nullvektor $0 \in V$ ein linear abhängiges System, aber auch jedes System von Vektoren, in dem einer der Vektoren mehrfach vorkommt, ist linear abhängig. Dagegen ist ein System, welches aus genau einem Vektor $a \neq 0$ besteht, stets linear unabhängig. Ähnlich wie bei der Konvention der leeren Summe betrachtet man Systeme von Vektoren $a_1, \dots, a_n$ auch im Falle $n = 0$ und meint damit dann das leere System. Auch das leere System erkennt man in nahe liegender Weise als linear unabhängig.

Um die Sprache zu vereinfachen, erwähnt man in der Situation von Definition 1 meist nur die zu betrachtenden Vektoren $a_1, \dots, a_n$, ohne besonders darauf hinzuweisen, dass das System dieser Vektoren gemeint ist. So sagt man etwa in unpräziser Ausdrucksweise, die Vektoren $a_1, \dots, a_n$ seien linear unabhängig, womit man natürlich nicht meint, dass jeder der Vektoren a_i für sich genommen ein linear unabhängiges System bildet (was lediglich $a_i \neq 0$ bedeuten würde), sondern dass das System $(a_i)_{i=1 \dots n}$ linear unabhängig ist.

Mit 1.3/5 sehen wir beispielsweise, dass die Elemente $1, \sqrt{2} \in \mathbb{Q}(\sqrt{2})$ ein linear unabhängiges System bilden, wenn wir $\mathbb{Q}(\sqrt{2})$ als $\mathbb{Q}$ -Vektorraum auffassen. Entsprechendes gilt für die Elemente $1, i$ in $\mathbb{C}$ als $\mathbb{R}$ -Vektorraum. Wichtig ist auch, dass für $n \in \mathbb{N}$ die "Einheitsvektoren" $e_1, \dots, e_n \in K^n$ ein linear unabhängiges System bilden. Denn für $\alpha_1, \dots, \alpha_n \in K$ gilt

$$\sum_{i=1}^n \alpha_i e_i = (\alpha_1, \dots, \alpha_n),$$

also verschwindet diese Summe genau dann, wenn das Element $(\alpha_1, \dots, \alpha_n)$ verschwindet, d. h. wenn $\alpha_i = 0$ für $i = 1, \dots, n$ gilt.

Wir haben die lineare Abhängigkeit bzw. Unabhängigkeit in Definition 1 der Einfachheit halber nur für endliche Systeme von Vektoren formuliert. Die Begriffe übertragen sich aber in nahe liegender Weise auf beliebige Systeme $(a_i)_{i \in I}$, wenn man vereinbart, dass eine Linearkombination der a_i ein Ausdruck der Form $\sum_{i \in I} \alpha_i a_i$ mit Koeffizienten $\alpha_i \in K$ ist, wobei die α_i für fast alle

$i \in I$ verschwinden, d. h. für alle $i \in I$ bis auf endlich viele Ausnahmen. Eine solche Linearkombination ist daher in Wahrheit eine *endliche* Linearkombination, stellt also ein Element in V dar. Man bezeichnet ein System $(a_i)_{i \in I}$ von Vektoren aus V als *linear unabhängig*, wenn aus dem Verschwinden einer Linearkombination der a_i , also einer Gleichung $\sum_{i \in I} \alpha_i a_i = 0$, notwendig $\alpha_i = 0$ für alle $i \in I$ folgt. Das System $(a_i)_{i \in I}$ ist daher genau dann linear unabhängig, wenn jedes endliche Teilsystem von $(a_i)_{i \in I}$ linear unabhängig im Sinne von Definition 1 ist. Entsprechend ist $(a_i)_{i \in I}$ genau dann linear abhängig, wenn es ein endliches Teilsystem gibt, welches linear abhängig im Sinne von Definition 1 ist.

Satz 2. *Es seien $a_1, \dots, a_n$ Vektoren eines K -Vektorraums V . Dann ist äquivalent:*

- (i) *Die Vektoren $a_1, \dots, a_n$ sind linear unabhängig.*
- (ii) *Ist $a = \sum_{i=1}^n \alpha_i a_i$ eine Darstellung eines Elementes $a \in \langle a_1, \dots, a_n \rangle$ mit Koeffizienten $\alpha_1, \dots, \alpha_n \in K$, so sind diese eindeutig durch a bestimmt.*

Beweis. Wir nehmen zunächst Bedingung (i) als gegeben an. Sind dann

$$a = \sum_{i=1}^n \alpha_i a_i = \sum_{i=1}^n \alpha'_i a_i$$

zwei Darstellungen von a als Linearkombination der a_i , so ist $\sum_{i=1}^n (\alpha_i - \alpha'_i) a_i$ eine Linearkombination, die den Nullvektor 0 darstellt. Mit (i) folgt $\alpha_i - \alpha'_i = 0$, also $\alpha_i = \alpha'_i$ für alle i , d. h. die Darstellung von a als Linearkombination der a_i ist eindeutig.

Sei nun umgekehrt Bedingung (ii) gegeben. Um die lineare Unabhängigkeit des Systems der a_i zu zeigen, betrachten wir eine Gleichung $\sum_{i=1}^n \alpha_i a_i = 0$ mit Koeffizienten $\alpha_1, \dots, \alpha_n \in K$. Da trivialerweise $\sum_{i=1}^n 0 \cdot a_i = 0$ gilt, ergibt sich $\alpha_i = 0$ für alle i , wenn man (ii) benutzt. $\square$

Sind die Bedingungen des Satzes erfüllt, so nennt man das System der a_i eine *Basis* des linearen Unterraumes $\langle a_1, \dots, a_n \rangle$ von V . Man vereinbart nämlich:

Definition 3. *Ein System von Vektoren $a_1, \dots, a_n$ eines K -Vektorraums V wird als (endliche) Basis von V bezeichnet, wenn gilt:*

- (i) *Die Vektoren $a_1, \dots, a_n$ bilden ein Erzeugendensystem von V ; d. h. man hat $V = \langle a_1, \dots, a_n \rangle$.*
- (ii) *Das System der Vektoren $a_1, \dots, a_n$ ist linear unabhängig.*

Allgemeiner heißt ein (nicht notwendig endliches) System von Vektoren eines Vektorraums V eine Basis, wenn es sich um ein Erzeugendensystem handelt, welches linear unabhängig ist.

Mit Satz 2 ergibt sich sofort:

Bemerkung 4. Vektoren $a_1, \dots, a_n$ eines K -Vektorraumes V bilden genau dann eine Basis, wenn gilt: Jedes $a \in V$ besitzt eine Darstellung $a = \sum_{i=1}^n \alpha_i a_i$ mit eindeutig bestimmten Koeffizienten $\alpha_1, \dots, \alpha_n \in K$.

Fassen wir die bisher betrachteten Beispiele von Erzeugendensystemen und linear unabhängigen Systemen zusammen, so ergibt sich:

- (1) Das leere System bildet eine Basis des Nullraums über einem gegebenen Körper K , also des K -Vektorraums $V = 0$.
- (2) Die Elemente $1, \sqrt{2}$ bilden eine Basis von $\mathbb{Q}(\sqrt{2})$ als $\mathbb{Q}$ -Vektorraum.
- (3) Die Elemente $1, i$ bilden eine Basis von $\mathbb{C}$ als $\mathbb{R}$ -Vektorraum.
- (4) Für einen Körper K und $n \in \mathbb{N}$ bilden die Einheitsvektoren $e_1, \dots, e_n$ eine Basis des K -Vektorraums K^n .

Die Kenntnis von Basen in Vektorräumen ist verantwortlich dafür, dass man etwa Fragen zur linearen Unabhängigkeit von Vektoren auf das Lösen linearer Gleichungssysteme zurückführen kann. Wir wollen dies am Beispiel des K -Vektorraums K^n und der Basis $e_1, \dots, e_n$ einmal demonstrieren. Gegeben seien Vektoren $a_1, \dots, a_r \in K^n$, etwa

$$a_j = (\alpha_{1j}, \dots, \alpha_{nj}) = \sum_{i=1}^n \alpha_{ij} e_i, \quad j = 1, \dots, r.$$

Die Frage, ob $a_1, \dots, a_r$ linear abhängig sind oder nicht, ist dann äquivalent zu der Frage, ob es ein nicht-triviales r -Tupel $(\xi_1, \dots, \xi_r) \in K^r$ gibt mit $\sum_{j=1}^r \xi_j a_j = 0$, d. h. ob das lineare Gleichungssystem

$$\begin{aligned} \xi_1 \alpha_{11} + \dots + \xi_r \alpha_{1r} &= 0 \\ &\dots \\ \xi_1 \alpha_{n1} + \dots + \xi_r \alpha_{nr} &= 0 \end{aligned}$$

eine nicht-triviale Lösung $(\xi_1, \dots, \xi_r) \in K^r$ besitzt. Techniken zur Lösung solcher Gleichungssysteme werden wir im Abschnitt 3.5 kennen lernen.

Als Nächstes wollen wir ein technisches Lemma beweisen, welches insbesondere für die Handhabung und Charakterisierung von Vektorraumbasen von großem Nutzen ist.

Lemma 5. Für Vektoren $a_1, \dots, a_n$ eines K -Vektorraums V ist äquivalent:

- (i) $a_1, \dots, a_n$ sind linear abhängig.
- (ii) Einer der Vektoren $a_1, \dots, a_n$ ist eine Linearkombination der restlichen, d. h. es existiert ein $p \in \{1, \dots, n\}$ mit $a_p \in \langle a_1, \dots, a_{p-1}, a_{p+1}, \dots, a_n \rangle$.
- (iii) Es existiert ein $p \in \{1, \dots, n\}$ mit

$$\langle a_1, \dots, a_n \rangle = \langle a_1, \dots, a_{p-1}, a_{p+1}, \dots, a_n \rangle.$$

Sind die Vektoren $a_1, \dots, a_r$ für ein $r < n$ linear unabhängig, so folgen aus (i) die Bedingungen (ii) und (iii) bereits für ein $p \in \{r+1, \dots, n\}$.

Beweis. Wir beginnen mit der Implikation von (i) nach (ii). Seien also $a_1, \dots, a_n$ linear abhängig. Man wähle dann $r \in \{0, \dots, n\}$ maximal mit der Eigenschaft, dass das System der Vektoren $a_1, \dots, a_r$ linear unabhängig ist; im Falle $r = 0$ sei hiermit das leere System gemeint, welches stets linear unabhängig ist. Insbesondere gilt $r < n$ aufgrund der Voraussetzung in (i), und $a_1, \dots, a_{r+1}$ sind linear abhängig. Es existiert folglich eine Gleichung $\sum_{i=1}^{r+1} \alpha_i a_i = 0$ mit Koeffizienten $\alpha_i \in K$, die nicht sämtlich verschwinden. Dabei gilt notwendigerweise $\alpha_{r+1} \neq 0$, denn anderenfalls hätte man die Gleichung $\sum_{i=1}^r \alpha_i a_i = 0$, wobei die Koeffizienten nicht sämtlich verschwinden würden, die $a_1, \dots, a_r$ also linear abhängig wären. Die erstere Gleichung lässt sich daher nach a_{r+1} auflösen, man erhält $a_{r+1} = -\sum_{i=1}^r \alpha_i^{-1} \alpha_i a_i$ und damit $a_{r+1} \in \langle a_1, \dots, a_r \rangle$, wie in (ii) und der Zusatzaussage behauptet.

Sei nun Bedingung (ii) erfüllt, d. h. es gelte für ein $p \in \{1, \dots, n\}$ die Beziehung $a_p \in \langle a_1, \dots, a_{p-1}, a_{p+1}, \dots, a_n \rangle$. Man hat dann

$$a_1, \dots, a_n \in \langle a_1, \dots, a_{p-1}, a_{p+1}, \dots, a_n \rangle$$

und somit

$$(*) \quad \langle a_1, \dots, a_n \rangle \subset \langle a_1, \dots, a_{p-1}, a_{p+1}, \dots, a_n \rangle,$$

denn $\langle a_1, \dots, a_n \rangle$ ist der kleinste lineare Unterraum von V , der $a_1, \dots, a_n$ enthält. Da die umgekehrte Inklusion trivialerweise erfüllt ist, ergibt sich Bedingung (iii).

Der Vollständigkeit halber wollen wir hier auch noch darauf hinweisen, dass sich die Inklusion $(*)$ leicht durch direktes Nachrechnen herleiten lässt. Es gelte etwa $a_p = \sum_{i \neq p} \alpha_i a_i$ mit Koeffizienten $\alpha_i \in K$. Für jedes $b \in \langle a_1, \dots, a_n \rangle$ mit einer Darstellung $b = \sum_{i=1}^n \beta_i a_i$ und Koeffizienten $\beta_i \in K$ ergibt sich dann

$$b = \sum_{i \neq p} \beta_i a_i + \beta_p \sum_{i \neq p} \alpha_i a_i = \sum_{i \neq p} (\beta_i + \beta_p \alpha_i) a_i,$$

also $b \in \langle a_1, \dots, a_{p-1}, a_{p+1}, \dots, a_n \rangle$, und somit

$$\langle a_1, \dots, a_n \rangle \subset \langle a_1, \dots, a_{p-1}, a_{p+1}, \dots, a_n \rangle.$$

Sei schließlich Bedingung (iii) gegeben, für ein $p \in \{1, \dots, n\}$ gelte also

$$\langle a_1, \dots, a_n \rangle = \langle a_1, \dots, a_{p-1}, a_{p+1}, \dots, a_n \rangle$$

Dann folgt insbesondere $a_p \in \langle a_1, \dots, a_{p-1}, a_{p+1}, \dots, a_n \rangle$, etwa $a_p = \sum_{i \neq p} \alpha_i a_i$ mit gewissen Koeffizienten $\alpha_i \in K$, und die Gleichung $(-1)a_p + \sum_{i \neq p} \alpha_i a_i = 0$ zeigt, dass $a_1, \dots, a_n$ linear abhängig sind. Damit ist gezeigt, dass die Bedingungen (i), (ii) und (iii) äquivalent sind.

Sind nun die Vektoren $a_1, \dots, a_r$ für ein gegebenes $r < n$ linear unabhängig, $a_1, \dots, a_n$ aber insgesamt linear abhängig, so gilt, wie wir gesehen haben, Bedingung (ii) für ein $p \in \{r+1, \dots, n\}$. Für dieses p ist dann auch Bedingung (iii) erfüllt, so dass die zusätzliche Behauptung ebenfalls bewiesen ist. $\square$

Das gerade bewiesene Lemma lässt einige interessante Schlussfolgerungen zu.

Satz 6. *Jeder endlich erzeugte K -Vektorraum besitzt eine Basis, und jede solche Basis ist endlich.*

Beweis. Es sei $a_1, \dots, a_n$ ein Erzeugendensystem des betrachteten K -Vektorraums V , d. h. es gelte $V = \langle a_1, \dots, a_n \rangle$. Indem wir dieses System verkleinern, können wir $a_1, \dots, a_n$ als minimales Erzeugendensystem voraussetzen. Die Äquivalenz der Bedingungen (i) und (iii) in Lemma 5 zeigt dann, dass die Vektoren $a_1, \dots, a_n$ linear unabhängig sind, also eine Basis bilden.

Ist nun $(b_j)_{j \in J}$ eine weitere Basis von V , so lässt sich jeder der Vektoren $a_1, \dots, a_n$ als Linearkombination von endlich vielen der Vektoren b_j , $j \in J$, darstellen. Es existiert deshalb eine endliche Teilmenge $J' \subset J$ mit

$$V = \langle a_1, \dots, a_n \rangle \subset \langle b_j ; j \in J' \rangle \subset V.$$

Das System $(b_j)_{j \in J'}$ bildet somit ein Erzeugendensystem von V . Dieses ist als Teilsystem von $(b_j)_{j \in J}$ sogar linear unabhängig und stellt deshalb, ebenso wie $(b_j)_{j \in J}$, eine Basis dar. Dann folgt aber notwendig $J = J'$, und man erkennt J als endlich. $\square$

Satz 7. *Es sei V ein K -Vektorraum und $a_1, \dots, a_n$ ein System von Vektoren aus V . Dann ist äquivalent:*

- (i) $a_1, \dots, a_n$ bilden eine Basis von V .
- (ii) $a_1, \dots, a_n$ ist ein maximales linear unabhängiges System in V .
- (iii) $a_1, \dots, a_n$ ist ein minimales Erzeugendensystem von V .

Beweis. Sei zunächst Bedingung (i) als gegeben angenommen, sei also $a_1, \dots, a_n$ eine Basis von V . Für beliebiges $a \in V$ gilt dann

$$V = \langle a_1, \dots, a_n \rangle = \langle a, a_1, \dots, a_n \rangle,$$

und man schließt aus der Äquivalenz (i) $\iff$ (iii) von Lemma 5, dass das System $a, a_1, \dots, a_n$ linear abhängig ist. Also ist $a_1, \dots, a_n$ ein maximales linear unabhängiges System in V .

Als Nächstes gehen wir von Bedingung (ii) aus, sei also $a_1, \dots, a_n$ ein maximales linear unabhängiges System in V . Ist dann $a \in V$ beliebig, so ist das System $a_1, \dots, a_n, a$ linear abhängig, und es existiert eine nicht-triviale Linearkombination mit Koeffizienten aus K

$$\alpha a + \sum_{i=1}^n \alpha_i a_i = 0,$$

welche die Null darstellt. Aus der linearen Unabhängigkeit der $a_1, \dots, a_n$ ergibt sich mittels Lemma 5 (man vergleiche den Beweis der Implikation (i) $\implies$ (ii) in Lemma 5), dass zumindest der Koeffizient α nicht verschwindet. Folglich lässt

sich vorstehende Gleichung nach a auflösen, und man erhält $a \in \langle a_1, \dots, a_n \rangle$, d. h. $a_1, \dots, a_n$ ist ein Erzeugendensystem von V . Weiter folgt aus der linearen Unabhängigkeit der $a_1, \dots, a_n$, indem man die Äquivalenz (i) $\iff$ (iii) aus Lemma 5 benutzt, dass $a_1, \dots, a_n$ ein minimales Erzeugendensystem von V ist.

Nehmen wir schließlich $a_1, \dots, a_n$ wie in Bedingung (iii) als minimales Erzeugendensystem an, so zeigt die Äquivalenz (i) $\iff$ (iii) aus Lemma 5, dass $a_1, \dots, a_n$ dann notwendig ein linear unabhängiges System ist, also eine Basis, da es bereits ein Erzeugendensystem ist. $\square$

Satz 8 (Basisergänzungssatz). *In einem K -Vektorraum V betrachte man ein linear unabhängiges System $a_1, \dots, a_r$ sowie ein Erzeugendensystem $b_1, \dots, b_m$. Dann lässt sich das System der a_i durch Elemente des Systems der b_j zu einer Basis von V ergänzen, d. h. es existieren paarweise verschiedene Indizes $i(r+1), \dots, i(n) \in \{1, \dots, m\}$ mit der Eigenschaft, dass die Vektoren*

$$a_1, \dots, a_r, b_{i(r+1)}, \dots, b_{i(n)}$$

eine Basis von V bilden.

Beweis. Für $n \geq r$ betrachte man paarweise verschiedene Indizes

$$i(r+1), \dots, i(n) \in \{1, \dots, m\},$$

so dass

$$(*) \quad V = \langle a_1, \dots, a_r, b_{i(r+1)}, \dots, b_{i(n)} \rangle$$

gilt. Die Gleichung ist beispielsweise für $n = r + m$ erfüllt, wenn man $i_{r+j} = j$ für $j = 1, \dots, m$ setzt. Man betrachte nun eine Gleichung (*), wobei $n \geq r$ minimal gewählt sei. Dann ist $a_1, \dots, a_r, b_{i(r+1)}, \dots, b_{i(n)}$ ein linear unabhängiges Erzeugendensystem, stellt also eine Basis von V dar. Anderenfalls wäre dieses System nämlich linear abhängig, und man könnte es aufgrund der Äquivalenz (i) $\iff$ (iii) aus Lemma 5 zu einem echt kleineren Erzeugendensystem verkürzen. Da die Vektoren $a_1, \dots, a_r$ jedoch linear unabhängig sind, ergibt sich mit Lemma 5 in dieser Situation, dass man einen der Vektoren $b_{i(r+1)}, \dots, b_{i(n)}$ fortlassen kann, was aber wegen der Minimalität von n ausgeschlossen ist. Das Erzeugendensystem $a_1, \dots, a_r, b_{i(r+1)}, \dots, b_{i(n)}$ ist daher linear unabhängig und folglich eine Basis.

Wir wollen noch auf einen zweiten Beweis eingehen, der den Vorteil hat, dass er im Hinblick auf nicht-endliche Basen verallgemeinerungsfähig ist. Hierzu betrachten wir Indizes

$$i(r+1), \dots, i(n) \in \{1, \dots, m\},$$

nunmehr aber mit der Bedingung, dass die Vektoren

$$a_1, \dots, a_r, b_{i(r+1)}, \dots, b_{i(n)}$$

linear unabhängig sind. Wir dürfen n als maximal gewählt annehmen. Mit Lemma 5 ergibt sich dann

$$b_1, \dots, b_m \in \langle a_1, \dots, a_r, b_{i(r+1)}, \dots, b_{i(n)} \rangle$$

und folglich

$$V = \langle b_1, \dots, b_m \rangle \subset \langle a_1, \dots, a_r, b_{i(r+1)}, \dots, b_{i(n)} \rangle,$$

so dass $a_1, \dots, a_r, b_{i(r+1)}, \dots, b_{i(n)}$ ein Erzeugendensystem und damit eine Basis von V bilden. $\square$

Theorem 9. *In einem K -Vektorraum V mögen die Elemente $a_1, \dots, a_n$ eine Basis sowie $b_1, \dots, b_m$ ein Erzeugendensystem bilden. Dann gilt $n \leq m$. Weiter ist $b_1, \dots, b_m$ genau dann eine Basis, wenn $n = m$ gilt. Je zwei Basen eines endlich erzeugten K -Vektorraums V bestehen folglich aus gleichviel Elementen.*

Beweis. Aufgrund des Basisergänzungssatzes 8 lässt sich das System $a_2, \dots, a_n$ durch Elemente des Systems $b_1, \dots, b_m$ zu einer Basis $b_{i(1)}, \dots, b_{i(r_1)}, a_2, \dots, a_n$ ergänzen, wobei natürlich $r_1 \geq 1$ gelten muss; vgl. Lemma 5. Lässt man bei dieser Basis das Element a_2 fort, so kann man das entstehende System wiederum durch Elemente des Systems $b_1, \dots, b_m$ zu einer Basis von V ergänzen, etwa zu

$$b_{i(1)}, \dots, b_{i(r_1)}, b_{i(r_1+1)}, \dots, b_{i(r_1+r_2)}, a_3, \dots, a_n.$$

Fährt man auf diese Weise fort, so gelangt man nach n Schritten zu einer Basis $b_{i(1)}, \dots, b_{i(r_1+\dots+r_n)}$, wobei die Indizes $i(1), \dots, i(r_1+\dots+r_n) \in \{1, \dots, m\}$ notwendig paarweise verschieden sind. Es folgt $r_1 + \dots + r_n \leq m$ und wegen $r_i \geq 1$ insbesondere $n \leq m$, wie behauptet.

Ist nun $b_1, \dots, b_m$ bereits eine Basis, so kann man die Rolle der a_i und b_j vertauschen und erhält auf diese Weise $m \leq n$, also insbesondere $m = n$. Bildet andererseits $b_1, \dots, b_m$ mit $m = n$ ein Erzeugendensystem von V , so kann man dieses System zu einem minimalen Erzeugendensystem von V verkleinern, also zu einer Basis; vgl. Satz 7. Da wir aber schon wissen, dass Basen in V aus genau n Elementen bestehen, folgt, dass $b_1, \dots, b_m$ notwendig eine Basis von V ist.

Da endlich erzeugte K -Vektorräume gemäß Satz 6 lediglich endliche Basen besitzen, ergibt sich insbesondere, dass je zwei Basen eines solchen Vektorraums aus gleichviel Elementen bestehen. $\square$

Für ein System $a_1, \dots, a_n$ von Elementen bezeichnet man die natürliche Zahl n als die *Länge* dieses Systems. Gelegentlich werden wir auch unendlichen Systemen $(a_i)_{i \in I}$, also Systemen mit unendlicher Indexmenge I , eine Länge zuordnen, nämlich die Länge ∞ . Wir werden dabei nicht zwischen verschiedenen Graden der Unendlichkeit unterscheiden, etwa abzählbar unendlich (z. B. $I = \mathbb{N}$) oder überabzählbar unendlich (z. B. $I = \mathbb{R}$).

Definition 10. *Es sei V ein K -Vektorraum. Besitzt dann V eine Basis endlicher Länge n , so bezeichnet man n als die Dimension von V , in Zeichen*

$\dim_K V = n$. Gibt es andererseits in V keine Basis endlicher Länge, also kein endliches maximales linear unabhängiges System, so sagen wir, die Dimension von V sei unendlich, $\dim_K V = \infty$.

Aufgrund von Theorem 9 ist die Dimension eines Vektorraums wohldefiniert. Der Nullraum $V = 0$ hat die Dimension 0, jeder K -Vektorraum $V \neq 0$ eine Dimension > 0 . Wir wollen noch einige weitere Eigenschaften der Dimension eines Vektorraums zusammenstellen, die sich auf einfache Weise aus den bisher gewonnenen Ergebnissen folgern lassen.

Korollar 11. *Es sei V ein K -Vektorraum und $n \in \mathbb{N}$. Dann ist äquivalent:*

- (i) $\dim_K V = n$.
- (ii) *Es existiert in V ein linear unabhängiges System von n Vektoren, und jeweils $n + 1$ Vektoren sind linear abhängig.*

Beweis. Sei zunächst Bedingung (i) gegeben. Jede Basis von V bildet dann ein linear unabhängiges System bestehend aus n Vektoren. Ist andererseits $y_1, \dots, y_{n+1}$ ein System von $n + 1$ Vektoren aus V und nehmen wir an, dass dieses linear unabhängig ist, so können wir das System gemäß Satz 8 zu einer Basis von V ergänzen. Man hätte dann $\dim_K V \geq n + 1$ im Widerspruch zu unserer Voraussetzung. Aus (i) ergibt sich folglich (ii).

Ist umgekehrt Bedingung (ii) gegeben, so gibt es in V ein maximales linear unabhängiges System bestehend aus n Vektoren. Dieses bildet eine Basis, und es folgt $\dim_K V = n$. $\square$

Korollar 12. *Es sei V ein K -Vektorraum und $n \in \mathbb{N}$. Dann ist äquivalent:*

- (i) $\dim_K V \geq n$.
- (ii) *Es existiert in V ein linear unabhängiges System von n Vektoren.*

Beweis. Bedingung (i) impliziert trivialerweise Bedingung (ii), auch im Falle unendlicher Dimension, da dann keine endlichen Basen, also keine endlichen maximalen linear unabhängigen Systeme in V existieren können. Gehen wir umgekehrt von (ii) aus, so ist nur im Falle $\dim_K V < \infty$ etwas zu zeigen. Jedes linear unabhängige System von Vektoren $a_1, \dots, a_n \in V$ lässt sich dann gemäß Satz 8 zu einer Basis von V ergänzen, und es folgt wie gewünscht $\dim_K V \geq n$. $\square$

Korollar 13. *Für einen K -Vektorraum V ist äquivalent:*

- (i) $\dim_K V = \infty$.
- (ii) *Es existiert eine Folge von Vektoren $a_1, a_2, \dots \in V$, so dass für jedes $n \in \mathbb{N}$ das System $a_1, \dots, a_n$ linear unabhängig ist.*
- (iii) *Es existiert eine Folge von Vektoren $a_1, a_2, \dots \in V$, so dass das System $(a_i)_{i \in \mathbb{N}}$ linear unabhängig ist.*
- (iv) *Zu jedem $n \in \mathbb{N}$ gibt es ein linear unabhängiges System, bestehend aus n Vektoren von V .*

Beweis. Wir gehen aus von Bedingung (i). Sei also $\dim_K V = \infty$. Dann gibt es in V keine endlichen Basen und somit keine endlichen maximalen linear unabhängigen Systeme. Als Konsequenz ist es möglich, eine Folge von Vektoren $a_1, a_2, \dots \in V$ wie in (ii) gewünscht zu konstruieren. Weiter folgt aus (ii) unmittelbar Bedingung (iii), da zu jeder endlichen Teilmenge $I \subset \mathbb{N}$ ein $n \in \mathbb{N}$ existiert mit $I \subset \{1, \dots, n\}$. Die Implikation (iii) $\implies$ (iv) ist trivial, und (iv) $\implies$ (i) schließlich ergibt sich mit Korollar 12. $\square$

Korollar 14. *Es sei V ein K -Vektorraum und $U \subset V$ ein Teilraum. Dann gilt:*

- (i) $\dim_K U \leq \dim_K V$
- (ii) *Aus $\dim_K U = \dim_K V < \infty$ folgt bereits $U = V$.*

Beweis. Die erste Behauptung folgt mittels Korollar 12 aus der Tatsache, dass ein linear unabhängiges System von Vektoren aus U auch in V linear unabhängig ist. Die zweite Behauptung gilt, da man in einem endlich-dimensionalen K -Vektorraum V ein linear unabhängiges System, beispielsweise eine Basis von U , stets zu einer Basis von V ergänzen kann. $\square$

Wir wollen nun noch einige Beispiele betrachten.

- (1) Ist K ein Körper, $n \in \mathbb{N}$, so folgt $\dim_K K^n = n$.
- (2) $\dim_{\mathbb{R}} \mathbb{C} = 2$
- (3) $\dim_{\mathbb{Q}} \mathbb{Q}(\sqrt{2}) = 2$
- (4) $\dim_{\mathbb{Q}} \mathbb{R} = \infty$. Dies zeigt man am einfachsten mit Hilfe eines Abzählbarkeitsarguments. Jeder endlich-dimensionale $\mathbb{Q}$ -Vektorraum wäre, ebenso wie $\mathbb{Q}$, abzählbar, jedoch ist $\mathbb{R}$ nicht abzählbar.

(5) Sei K ein Körper, X eine Menge und $V = \text{Abb}(X, K)$ der K -Vektorraum der K -wertigen Funktionen auf X . Besteht X dann aus $n < \infty$ Elementen, so gilt $\dim_K V = n$, wohingegen man für unendliches X die Gleichung $\dim_K V = \infty$ hat. Wir wollen dies im Folgenden begründen. Für $x \in X$ bezeichne $f_x: X \rightarrow K$ diejenige Funktion, die durch $f_x(x) = 1$ und $f_x(y) = 0$ für $y \neq x$ gegeben ist. Dann ist für jeweils endlich viele paarweise verschiedene Elemente $x_1, \dots, x_n \in X$ das System $f_{x_1}, \dots, f_{x_n}$ linear unabhängig in V , denn aus einer Gleichung $\sum_{i=1}^n \alpha_i f_{x_i} = 0$ mit Koeffizienten $\alpha_1, \dots, \alpha_n \in K$ folgt

$$0 = \left(\sum_{i=1}^n \alpha_i f_{x_i} \right)(x_j) = \alpha_j$$

für $j = 1, \dots, n$. Hieraus ergibt sich bereits $\dim_K V = \infty$, wenn X unendlich viele Elemente besitzt. Da wir andererseits für endliches X jedes $f \in V$ in der Form

$$f = \sum_{x \in X} f(x) f_x$$

schreiben können, ist das System $(f_x)_{x \in X}$ in diesem Falle ein Erzeugendensystem und somit eine Basis von V , so dass man $\dim_K V = n$ hat, wenn X aus $n < \infty$ Elementen besteht.

Abschließend soll noch angedeutet werden, wie die Theorie dieses Abschnitts aussieht, wenn man sich nicht auf endlich erzeugte K -Vektorräume beschränkt. Man muss dann auch unendliche Basen zulassen, wie sie in Definition 3 mit eingeschlossen sind. Man prüft leicht nach, dass die in Satz 2 und Lemma 5 gegebenen Charakterisierungen linearer Abhängigkeit bzw. Unabhängigkeit sinngemäß auch für beliebige Systeme von Vektoren gelten. Als Folgerung übertragen sich die Resultate von Bemerkung 4 und Satz 7 auf den Fall nicht notwendig endlicher Basen.

Etwas problematischer ist der Beweis des Analogons zu Satz 6, dass nämlich jeder K -Vektorraum V eine Basis oder, in äquivalenter Sprechweise, ein maximales linear unabhängiges System besitzt. Die Existenz eines solchen Systems zeigt man am einfachsten mit Hilfe des so genannten Zornschen Lemmas, welches dem Gebiet der Mengenlehre zuzuordnen ist. Das Lemma geht von einer *teilweise geordneten* Menge M aus, wobei teilweise geordnet bedeutet, dass zwischen gewissen Elementen von M eine Relation " $\leq$ " besteht, und zwar mit den folgenden Eigenschaften:

$$\begin{aligned} x &\leq x \text{ für alle } x \in M \\ x \leq y, y \leq z &\implies x \leq z \\ x \leq y, y \leq x &\implies x = y \end{aligned}$$

Man nennt eine Teilmenge $N \subset M$ *streng geordnet*, wenn für je zwei Elemente $x, y \in N$ stets $x \leq y$ oder $y \leq x$ gilt. Weiter heißt ein Element $z \in M$ eine *obere Schranke* von N , wenn $x \leq z$ für alle $x \in N$ gilt. Das Lemma von Zorn lautet nun wie folgt:

Ist M eine teilweise geordnete Menge und besitzt jede streng geordnete Teilmenge von M eine obere Schranke in M , so existiert in M ein maximales Element.

Dabei heißt ein Element $z \in M$ *maximal*, wenn aus $z \leq x$ mit $x \in M$ stets $x = z$ folgt. In unserer konkreten Situation definiere man M als die Menge aller Teilmengen von V , deren Elemente ein linear unabhängiges System von Vektoren in V bilden. Für zwei solche Mengen $A, B \subset V$ setze man $A \leq B$, falls $A \subset B$ gilt. Die Voraussetzungen des Lemmas von Zorn sind dann für M erfüllt, als obere Schranke einer streng geordneten Teilmenge $N \subset M$ dient beispielsweise die Vereinigung aller Teilmengen $A \in N$, also

$$\bigcup_{A \in N} A \subset V.$$

Man erhält somit aus dem Zornschen Lemma die Existenz eines maximalen Elementes in V , d. h. eines maximalen linear unabhängigen Systems von Vektoren in V und damit gemäß Satz 7 einer Basis von V .

Wie wir gesehen haben, lässt sich die Existenz maximaler linear unabhängiger Systeme problemlos mit Hilfe des Zornschen Lemmas beweisen. Ähnliches kann man für minimale Erzeugendensysteme nicht behaupten, und dies ist der Grund dafür, dass die im Beweis zu Satz 6 benutzte Idee, Basen durch Minimieren von Erzeugendensystemen zu konstruieren, im Allgemeinfall nicht zum Ziel führt. Auch der Basisergänzungssatz 8 lässt sich mit Hilfe des Zornschen Lemmas auf den Fall unendlicher Systeme verallgemeinern, wenn man die im Beweis zu Satz 8 gegebene Argumentation im Sinne maximaler linear unabhängiger Systeme mit dem Zornschen Lemma kombiniert. Man kann sogar die Aussage von Theorem 9, dass nämlich je zwei Basen $(a_i)_{i \in I}$ und $(b_j)_{j \in J}$ eines K -Vektorraums V aus “gleichvielen” Elementen bestehen, auf unendlich-dimensionale Vektorräume verallgemeinern. Dabei ist “gleichviel” in dem Sinne zu präzisieren, dass es eine bijektive Abbildung $I \longrightarrow J$ gibt. Man nennt I und J bzw. die Basen $(a_i)_{i \in I}$ und $(b_j)_{j \in J}$ dann auch *gleichmächtig*⁴. Die Mächtigkeitsklasse einer solchen Basis könnten wir als Dimension von V bezeichnen, jedoch wollen wir im Sinne von Definition 10 nicht zwischen verschiedenen unendlichen Dimensionen unterscheiden.

Schließlich sei noch angemerkt, dass man in Korollar 14 (ii) nicht auf die Bedingung $\dim_K V < \infty$ verzichten kann. Um dies einzusehen, betrachte man einen K -Vektorraum V von unendlicher Dimension und ein abzählbar unendliches linear unabhängiges System $(a_i)_{i \in \mathbb{N}}$ von Vektoren in V . Dann ist einerseits $(a_{2 \cdot i})_{i \in \mathbb{N}}$ gleichmächtig zu $(a_i)_{i \in \mathbb{N}}$, andererseits aber $\langle a_0, a_2, a_4, \dots \rangle$ ein echter linearer Unterraum von $\langle a_0, a_1, a_2, \dots \rangle$.

Aufgaben

- Man betrachte $\mathbb{R}^3$ als $\mathbb{R}$ -Vektorraum und überprüfe folgende Systeme von Vektoren auf lineare Abhängigkeit bzw. lineare Unabhängigkeit:
 - $(1, 0, -1), (1, 2, 1), (0, -3, 2)$
 - $(1, 1, 1), (1, 1, 0), (1, 0, 0)$
 - $(9, 1, 5), (17, 11, 14), (9, 1, 5)$
 - $(1, 2, 3), (4, 5, 6), (6, 9, 12)$
 - $(1, 9, 7), (2, 3, 4), (9, 7, 6), (6, 6, 6)$
 - $(1, \alpha, 0), (\alpha, 1, 0), (0, \alpha, 1)$, wobei α eine reelle Zahl sei.
- Es seien U, U' lineare Unterräume eines K -Vektorraums V mit $U \cap U' = 0$. Bilden $x_1, \dots, x_r \in U$ und $y_1, \dots, y_s \in U'$ linear unabhängige Systeme, so auch die Vektoren $x_1, \dots, x_r, y_1, \dots, y_s$ in V .
- Für welche natürlichen Zahlen $n \in \mathbb{N}$ gibt es in $\mathbb{R}^n$ eine unendliche Folge von Vektoren $a_1, a_2, \dots$ mit der Eigenschaft, dass je zwei Vektoren dieser Folge linear unabhängig über $\mathbb{R}$ sind, also ein linear unabhängiges System im $\mathbb{R}$ -Vektorraum $\mathbb{R}^n$ bilden?

⁴ Dass je zwei Basen eines K -Vektorraums gleichmächtig sind, beweist man wie in “Bosch, Algebra”, Abschnitt 7.1. Die dortige Argumentation im Sinne von Transzendenzbasen und algebraischer Unabhängigkeit überträgt sich in direkter Weise auf den Fall von Vektorraum-basen und linearer Unabhängigkeit.

4. Man betrachte den $\mathbb{R}$ -Vektorraum aller Funktionen $p: \mathbb{R} \rightarrow \mathbb{R}$, die durch polynomiale Ausdrücke der Form $p(x) = \sum_{i=1}^r \alpha_i x^i$ mit Koeffizienten $\alpha_i \in \mathbb{R}$ und variablem $r \in \mathbb{N}$ gegeben sind. Man gebe eine Basis dieses Vektorraums an. (Hinweis: Man darf benutzen, dass nicht-triviale reelle Polynome höchstens endlich viele Nullstellen haben.)
5. Es sei $x_1, \dots, x_n$ eine Basis eines K -Vektorraums V . Für gegebene Koeffizienten $\alpha_{ij} \in K$, $1 \leq i < j \leq n$ setze man $y_j = x_j + \sum_{i < j} \alpha_{ij} x_i$, $j = 1, \dots, n$, und zeige, dass dann auch $y_1, \dots, y_n$ eine Basis von V bilden.
6. Es sei $\mathbb{F}$ ein endlicher Körper mit q Elementen und V ein $\mathbb{F}$ -Vektorraum der Dimension n .
 - (i) Man bestimme die Anzahl der Elemente von V .
 - (ii) Man bestimme die Anzahl der Teilmengen in V , deren Elemente jeweils eine Basis von V bilden.
7. Es sei $X = X_1 \amalg \dots \amalg X_n$ eine Menge mit einer endlichen Zerlegung in nicht-leere paarweise disjunkte Teilmengen. Für einen Körper K betrachte man den K -Vektorraum V aller K -wertigen Funktionen $X \rightarrow K$, sowie den linearen Unterraum U derjenigen Funktionen, die auf jedem der X_i konstant sind. Man berechne $\dim_K U$.

Welche Dimension erhält man, wenn die X_i nicht notwendig paarweise disjunkt sind und lediglich $X = \bigcup_{i=1}^n X_i$ gilt?

8. Es sei K ein Körper. Für gegebene Elemente $\gamma_1, \dots, \gamma_n \in K$ betrachte man die Teilmenge

$$U = \{(\alpha_1, \dots, \alpha_n) \in K^n; \sum_{i=1}^n \alpha_i \gamma_i = 0\}.$$

Man zeige, dass U ein linearer Unterraum von K^n ist und berechne $\dim_K U$.

1.6 Direkte Summen

Zu Vektoren $a_1, \dots, a_r$ eines K -Vektorraums V kann man die linearen Unterräume $Ka_i = \langle a_i \rangle$, $i = 1, \dots, r$, betrachten. Bilden die a_i ein Erzeugendensystem von V , so lässt sich jedes Element $b \in V$ in der Form $b = \sum_{i=1}^r b_i$ schreiben mit Vektoren $b_i \in Ka_i$; wir werden sagen, dass V die *Summe* der linearen Unterräume Ka_i ist. Bilden $a_1, \dots, a_r$ sogar eine Basis von V , so überlegt man leicht mit 1.5/4, dass in einer solchen Darstellung die Vektoren $b_i \in Ka_i$ eindeutig durch b bestimmt sind. Wir werden sagen, dass V die *direkte Summe* der Ka_i ist. Im Folgenden sollen Summe und direkte Summe für den Fall beliebiger linearer Unterräume von V erklärt werden.

Definition 1. Es seien $U_1, \dots, U_r$ lineare Unterräume eines K -Vektorraums V . Dann wird die Summe dieser Unterräume erklärt durch

$$\sum_{i=1}^r U_i = \left\{ \sum_{i=1}^r b_i; b_i \in U_i \text{ für } i = 1, \dots, r \right\}.$$

Es ist unmittelbar klar, dass $\sum_{i=1}^r U_i$ wieder ein linearer Unterraum von V ist, nämlich der von $U_1, \dots, U_r$ erzeugte lineare Unterraum $\langle U_1 \cup \dots \cup U_r \rangle \subset V$. Insbesondere sieht man, dass die Summe von linearen Unterräumen in V assoziativ ist.

Satz 2. Für eine Summe $U = \sum_{i=1}^r U_i$ von linearen Unterräumen $U_1, \dots, U_r$ eines K -Vektorraums V sind folgende Aussagen äquivalent:

- (i) Jedes $b \in U$ hat eine Darstellung $b = \sum_{i=1}^r b_i$ mit eindeutig bestimmten Vektoren $b_i \in U_i$, $i = 1, \dots, r$.
- (ii) Aus einer Gleichung $\sum_{i=1}^r b_i = 0$ mit Vektoren $b_i \in U_i$ folgt $b_i = 0$ für $i = 1, \dots, r$.
- (iii) Für $p = 1, \dots, r$ gilt $U_p \cap \sum_{i \neq p} U_i = 0$.

Beweis. Bedingung (i) impliziert trivialerweise (ii). Um (iii) aus (ii) herzuleiten, betrachte man einen Index $p \in \{1, \dots, r\}$ und einen Vektor $b \in U_p \cap \sum_{i \neq p} U_i$, etwa $b = \sum_{i \neq p} b_i$ mit Summanden $b_i \in U_i$. Dann gilt $-b + \sum_{i \neq p} b_i = 0$, und es folgt aus (ii) insbesondere $b = 0$. Somit hat man $U_p \cap \sum_{i \neq p} U_i = 0$.

Sei nun Bedingung (iii) gegeben, und sei $b \in V$ auf zwei Weisen als Summe von Vektoren $b_i, b'_i \in U_i$ dargestellt, also

$$b = \sum_{i=1}^r b_i = \sum_{i=1}^r b'_i.$$

Dann folgt $\sum_{i=1}^r b_i - b'_i = 0$ und somit

$$b_p - b'_p = - \sum_{i \neq p} b_i + b'_i \in U_p \cap \sum_{i \neq p} U_i = 0, \quad p = 1, \dots, r.$$

Insbesondere ergibt sich $b_i = b'_i$ für $i = 1, \dots, r$, d. h. (i) ist erfüllt. □

Definition 3. Es sei $U = \sum_{i=1}^r U_i$ eine Summe von linearen Unterräumen $U_1, \dots, U_r$ eines K -Vektorraums V . Dann heißt U die direkte Summe der U_i , in Zeichen $U = \bigoplus_{i=1}^r U_i$, wenn die äquivalenten Bedingungen von Satz 2 erfüllt sind.

Die Summe $U = \sum_{i=1}^r U_i$ von linearen Unterräumen $U_1, \dots, U_r \subset V$ wird auch mit $U_1 + \dots + U_r$ bezeichnet, und man schreibt $U_1 \oplus \dots \oplus U_r$, falls diese Summe direkt ist. Eine Summe $U_1 + U_2$ zweier linearer Unterräume $U_1, U_2 \subset V$ ist genau dann direkt, wenn $U_1 \cap U_2 = 0$ gilt. Wie schon angedeutet, ist die Direktheit einer Summe von linearen Unterräumen anzusehen als Verallgemeinerung der linearen Unabhängigkeit eines Systems von Vektoren. Für Vektoren $a_1, \dots, a_r \in V$ ist nämlich

$$\langle a_1, \dots, a_r \rangle = \bigoplus_{i=1}^r K a_i, \quad a_i \neq 0 \text{ für } i = 1, \dots, r$$

äquivalent zur linearen Unabhängigkeit von $a_1, \dots, a_r$.

Eine leichte Variante der gerade definierten direkten Summe stellt die so genannte *konstruierte direkte Summe* dar. Man geht hierbei von gegebenen K -Vektorräumen $V_1, \dots, V_r$ aus und konstruiert einen K -Vektorraum V , in dem sich die V_i als lineare Unterräume auffassen lassen, und zwar derart, dass V die direkte Summe der V_i ist, also $V = \bigoplus_{i=1}^r V_i$ im Sinne von Definition 3 gilt. Dabei wird V als das kartesische Produkt der V_i definiert, $V = \prod_{i=1}^r V_i$, und man fasst dieses Produkt als K -Vektorraum mit komponentenweiser Addition und skalarer Multiplikation auf. In V bilden dann die Teilmengen

$$V'_i = \{(v_1, \dots, v_r) \in V; v_j = 0 \text{ für } j \neq i\}, \quad i = 1, \dots, r,$$

lineare Unterräume, und man stellt unschwer fest, dass V die direkte Summe der V'_i ist. Da die natürliche Abbildung

$$\iota_i: V_i \longrightarrow V'_i, \quad v \longmapsto (v_1, \dots, v_r) \text{ mit } v_j = \begin{cases} v & \text{für } j = i \\ 0 & \text{für } j \neq i \end{cases},$$

für $i = 1, \dots, r$ jeweils bijektiv ist und zudem die Vektorraumstrukturen auf V_i und V'_i respektiert, können wir jeweils V_i mit V'_i unter ι_i identifizieren und auf diese Weise V als direkte Summe der linearen Unterräume $V_1, \dots, V_r$ auffassen, in Zeichen $V = \bigoplus_{i=1}^r V_i$. Wir nennen V die konstruierte direkte Summe der V_i . Hierbei ist jedoch ein wenig Vorsicht geboten. Ist beispielsweise ein K -Vektorraum V die (nicht notwendig direkte) Summe gewisser linearer Unterräume $U_1, \dots, U_r \subset V$, gilt also $V = \sum_{i=1}^r U_i$, so kann man zusätzlich die direkte Summe $U = \bigoplus_{i=1}^r U_i$ konstruieren. Mit Hilfe der nachfolgenden Dimensionsformeln kann man dann $\dim_K U \geq \dim_K V$ zeigen, wobei Gleichheit nur dann gilt, wenn V bereits die direkte Summe der U_i ist. Im Allgemeinen wird daher U wesentlich verschieden von V sein.

Satz 4. *Es sei V ein K -Vektorraum und $U \subset V$ ein linearer Unterraum. Dann existiert ein linearer Unterraum $U' \subset V$ mit $V = U \oplus U'$. Für jedes solche U' gilt*

$$\dim_K V = \dim_K U + \dim_K U'.$$

Man nennt U' in dieser Situation ein *Komplement* zu U . Komplemente von linearen Unterräumen sind abgesehen von den trivialen Fällen $U = 0$ und $U = V$ nicht eindeutig bestimmt.

Beweis zu Satz 4. Ist V von endlicher Dimension, so wähle man eine Basis $a_1, \dots, a_r$ von U und ergänze diese durch Vektoren $a_{r+1}, \dots, a_n$ gemäß 1.5/8 zu einer Basis von V . Dann gilt

$$V = \bigoplus_{i=1}^n K a_i, \quad U = \bigoplus_{i=1}^r K a_i,$$

und es ergibt sich $V = U \oplus U'$ mit $U' = \bigoplus_{i=r+1}^n K a_i$.

Ist V nicht notwendig von endlicher Dimension, so können wir im Prinzip genauso schließen, indem wir nicht-endliche Basen zulassen und die Ausführungen am Schluss von Abschnitt 1.5 beachten. Wir können dann eine Basis $(a_i)_{i \in I}$ von U wählen und diese mittels des Basisergänzungssatzes durch ein System $(a'_j)_{j \in J}$ von Vektoren zu einer Basis von V ergänzen. Es folgt $V = U \oplus U'$ mit $U' = \langle (a'_j)_{j \in J} \rangle$.

Zum Beweis der Dimensionsformel betrachte man ein Komplement U' zu U . Sind U und U' von endlicher Dimension, so wähle man eine Basis $b_1, \dots, b_r$ von U bzw. $b_{r+1}, \dots, b_n$ von U' . Dann bilden $b_1, \dots, b_n$ eine Basis von V , und es folgt

$$\dim_K V = n = r + (n - r) = \dim_K U + \dim_K U',$$

wie gewünscht. Ist mindestens einer der beiden Räume U und U' von unendlicher Dimension, so gilt dies erst recht für V , und die Dimensionsformel ist trivialerweise erfüllt. $\square$

Als Anwendung wollen wir noch eine Dimensionsformel für Untervektorräume beweisen.

Satz 5. *Es seien U, U' lineare Unterräume eines K -Vektorraumes V . Dann gilt*

$$\dim_K U + \dim_K U' = \dim_K (U + U') + \dim_K (U \cap U').$$

Beweis. Zunächst seien U und U' von endlicher Dimension. Dann ist $U + U'$ endlich erzeugt und damit nach 1.5/6 ebenfalls von endlicher Dimension. Man wähle nun gemäß Satz 4 ein Komplement W von $U \cap U'$ in U , sowie ein Komplement W' von $U \cap U'$ in U' , also mit

$$U = (U \cap U') \oplus W, \quad U' = (U \cap U') \oplus W'.$$

Dann gilt

$$U + U' = (U \cap U') + W + W',$$

und wir behaupten, dass diese Summe direkt ist. In der Tat, gilt $a + b + b' = 0$ für gewisse Elemente $a \in U \cap U'$, $b \in W$, $b' \in W'$, so ergibt sich

$$b = -(a + b') \in (U \cap U') + W' = U'$$

und wegen $b \in W \subset U$ sogar

$$b \in (U \cap U') \cap W = 0,$$

da U die direkte Summe der linearen Unterräume $U \cap U'$ und W ist. Man erhält also $b = 0$ und auf entsprechende Weise auch $b' = 0$. Damit folgt aber auch $a = 0$, was bedeutet, dass $U + U'$ die direkte Summe der linearen Unterräume $U \cap U'$, W und W' ist. Die behauptete Dimensionsformel ergibt sich dann mittels Satz 4 aus der Rechnung

$$\begin{aligned}
\dim(U + U') &= \dim(U \cap U') + \dim W + \dim W' \\
&= [\dim(U \cap U') + \dim W] + [\dim(U \cap U') + \dim W'] - \dim(U \cap U') \\
&= \dim U + \dim U' - \dim(U \cap U').
\end{aligned}$$

Abschließend bleibt noch der Fall zu betrachten, wo (mindestens) einer der linearen Unterräume U, U' nicht von endlicher Dimension ist. Gilt etwa $\dim_K U = \infty$, so folgt hieraus insbesondere $\dim_K(U + U') = \infty$, und wir haben $\dim_K U + \dim_K U'$ wie auch $\dim_K(U + U') + \dim_K(U \cap U')$ als ∞ zu interpretieren. Die behauptete Dimensionsformel ist also auch in diesem Fall gültig. $\square$

Korollar 6. *Es seien U, U' endlich-dimensionale lineare Unterräume eines K -Vektorraums V . Dann ist äquivalent:*

- (i) $U + U' = U \oplus U'$.
- (ii) $\dim_K(U + U') = \dim_K U + \dim_K U'$.

Beweis. Bedingung (ii) ist aufgrund von Satz 5 äquivalent zu $\dim_K(U \cap U') = 0$, also zu $U \cap U' = 0$ und damit zu Bedingung (i). $\square$

Aufgaben

- Man bestimme Komplemente zu folgenden linearen Unterräumen des $\mathbb{R}^3$ bzw. $\mathbb{R}^4$:
 - (i) $U = \langle (1, 2, 3), (-2, 3, 1), (4, 1, 5) \rangle$
 - (ii) $U = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4; 3x_1 - 2x_2 + x_3 + 2x_4 = 0\}$
- Man betrachte eine Zerlegung $V = \sum_{i=1}^n U_i$ eines endlich-dimensionalen K -Vektorraums V in Untervektorräume $U_i \subset V$ und zeige, dass die Summe der U_i genau dann direkt ist, wenn $\dim_K V = \sum_{i=1}^n \dim_K U_i$ gilt.
- Man betrachte eine direkte Summenzerlegung $V = \bigoplus_{i=1}^n U_i$ eines K -Vektorraums V in lineare Unterräume $U_i \subset V$, sowie für jedes $i = 1, \dots, n$ eine Familie $(x_{ij})_{j \in J_i}$ von Elementen $x_{ij} \in U_i$. Man zeige:
 - (i) Die Elemente x_{ij} , $i = 1, \dots, n$, $j \in J_i$, bilden genau dann ein Erzeugendensystem von V , wenn für jedes $i = 1, \dots, n$ die Elemente x_{ij} , $j \in J_i$, ein Erzeugendensystem von U_i bilden.
 - (ii) Die Elemente x_{ij} , $i = 1, \dots, n$, $j \in J_i$, sind genau dann linear unabhängig in V , wenn für jedes $i = 1, \dots, n$ die Elemente x_{ij} , $j \in J_i$, linear unabhängig in U_i sind.
 - (iii) Die Elemente x_{ij} , $i = 1, \dots, n$, $j \in J_i$, bilden genau dann eine Basis von V , wenn für jedes $i = 1, \dots, n$ die Elemente x_{ij} , $j \in J_i$, eine Basis von U_i bilden.
- Es seien U_1, U_2, U_3 lineare Unterräume eines K -Vektorraums V . Man zeige:

$$\begin{aligned}
&\dim_K U_1 + \dim_K U_2 + \dim_K U_3 \\
&= \dim_K(U_1 + U_2 + U_3) + \dim_K((U_1 + U_2) \cap U_3) + \dim_K(U_1 \cap U_2)
\end{aligned}$$

5. Es sei $U_1 \subset U_2 \subset \dots \subset U_n$ eine Folge linearer Unterräume eines K -Vektorraums V mit $\dim_K V < \infty$. Man zeige, dass es jeweils ein Komplement U'_i zu U_i gibt, $i = 1, \dots, n$, mit $U'_1 \supset U'_2 \supset \dots \supset U'_n$.
6. Es sei U ein linearer Unterraum eines endlich-dimensionalen K -Vektorraums V . Unter welcher Dimensionsbedingung gibt es lineare Unterräume U_1, U_2 in V mit $U \subsetneq U_i \subsetneq V$, $i = 1, 2$, und $U = U_1 \cap U_2$?
7. Es seien U, U' zwei lineare Unterräume eines endlich-dimensionalen K -Vektorraums V . Unter welcher Dimensionsbedingung besitzen U und U' ein gemeinsames Komplement in V ?
8. Es sei U ein linearer Unterraum eines endlich-dimensionalen K -Vektorraums V . Unter welcher Dimensionsbedingung kann man Komplemente $U_1, \dots, U_r$ zu U in V finden, derart dass $\sum_{i=1}^r U_i = \bigoplus_{i=1}^r U_i$ gilt?

2. Lineare Abbildungen

Vorbemerkungen

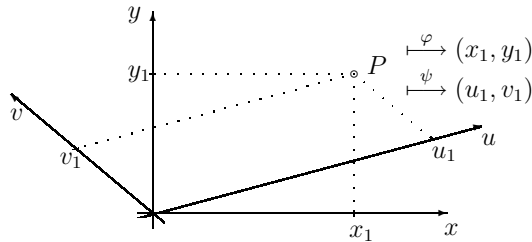
In Kapitel 1 haben wir Vektorräume über einem Körper K als Mengen eingeführt, die mit einer gewissen Struktur ausgestattet sind, nämlich einer Addition und einer skalaren Multiplikation mit Elementen aus K , wobei die Gültigkeit gewisser Axiome (Rechenregeln) gefordert wird. Unberührt blieb dabei zunächst die Frage, wann zwei K -Vektorräume V und V' als “gleich” oder “im Wesentlichen gleich” anzusehen sind. Es gibt eine natürliche Methode, um dies festzustellen: Man versuche, die Elemente von V mittels einer bijektiven Abbildung $f: V \longrightarrow V'$ mit denjenigen von V' zu identifizieren, und zwar in der Weise, dass dabei die Addition und die skalare Multiplikation von V und V' in Übereinstimmung gebracht werden. Für die Abbildung f erfordert dies die Gleichungen

$$f(a + b) = f(a) + f(b), \quad f(\alpha a) = \alpha f(a),$$

für alle $a, b \in V$ und alle $\alpha \in K$. Lässt sich eine solche bijektive Abbildung f finden, so kann man die Vektorräume V und V' als “im Wesentlichen gleich” ansehen, und man nennt f in diesem Falle einen *Isomorphismus* zwischen V und V' . Allgemeiner betrachtet man auch nicht notwendig bijektive Abbildungen $f: V \longrightarrow V'$ mit den oben genannten Verträglichkeitseigenschaften und bezeichnet diese als *Homomorphismen* oder *lineare Abbildungen* von V nach V' . Die Untersuchung von Abbildungen dieses Typs ist das zentrale Thema des vorliegenden Kapitels.

In den Vorbemerkungen zu Kapitel 1 hatten wir erläutert, wie man $\mathbb{R}^2$, die Menge aller Paare reeller Zahlen, als Modell einer gegebenen anschaulichen Ebene E auffassen kann, und zwar indem man in E ein Koordinatensystem auszeichnet. Wir wollen hier noch etwas genauer untersuchen, in welcher Weise dieses Modell von der Wahl des Koordinatensystems abhängt, wobei wir $\mathbb{R}^2$ unter der komponentenweisen Addition und skalaren Multiplikation als $\mathbb{R}$ -Vektorraum auffassen. Wir wählen also einen Nullpunkt $0 \in E$ und ein Koordinatensystem in 0 mit den Achsen x und y ; in den nachfolgenden Skizzen ist dieses System der besseren Übersicht halber als rechtwinkliges Koordinatensystem gezeichnet, was aber keinesfalls erforderlich ist. Indem wir einem Punkt $P \in E$ das Paar $(x_1, y_1) \in \mathbb{R}^2$ bestehend aus den Koordinaten von P bezüglich x und y zuordnen, erhalten wir eine bijektive Abbildung $\varphi: E \longrightarrow \mathbb{R}^2$, welche wir in den Vorbemerkungen zu Kapitel 1 jeweils als Identifizierung angesehen hatten. Für

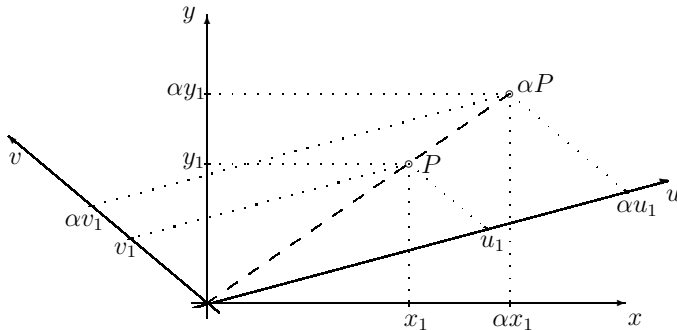
ein zweites Koordinatensystem in 0 mit den Koordinatenachsen u und v ergibt sich entsprechend eine bijektive Abbildung $\psi: E \rightarrow \mathbb{R}^2$, was mittels folgender Skizze verdeutlicht werden möge:



Der Übergang vom ersten zum zweiten Modell wird somit durch die Abbildung

$$f = \psi \circ \varphi^{-1}: \mathbb{R}^2 \rightarrow \mathbb{R}^2$$

beschrieben, und wir wollen plausibel machen, dass es sich hierbei um einen Isomorphismus von $\mathbb{R}$ -Vektorräumen in dem oben beschriebenen Sinne handelt. Zunächst einmal ist die Abbildung f bijektiv, da φ und ψ beide bijektiv sind. Es bleibt also lediglich noch zu zeigen, dass f mit der Vektorraumstruktur von $\mathbb{R}^2$ verträglich ist. Wir betrachten daher ein Paar $(x_1, y_1) \in \mathbb{R}^2$, den zugehörigen Punkt $P \in E$ mit $\varphi(P) = (x_1, y_1)$, sowie das Paar $(u_1, v_1) := \psi(P)$. Folglich besitzt P die Koordinaten x_1, y_1 bezüglich der Achsen x, y und entsprechend die Koordinaten u_1, v_1 bezüglich u, v . Nach Konstruktion gilt dann $f(x_1, y_1) = (u_1, v_1)$, und es ergibt sich für Skalare $\alpha \in \mathbb{R}$ folgendes Bild:



Dabei sei $\alpha P \in E$ derjenige Punkt, der aus P durch Streckung mit Zentrum 0 und Faktor α entsteht. Mittels des Strahlensatzes folgt dann, dass die Koordinaten von αP bezüglich der Achsen x, y bzw. u, v ebenfalls durch Streckung mit Faktor α aus den entsprechenden Koordinaten von P hervorgehen. Dies bedeutet

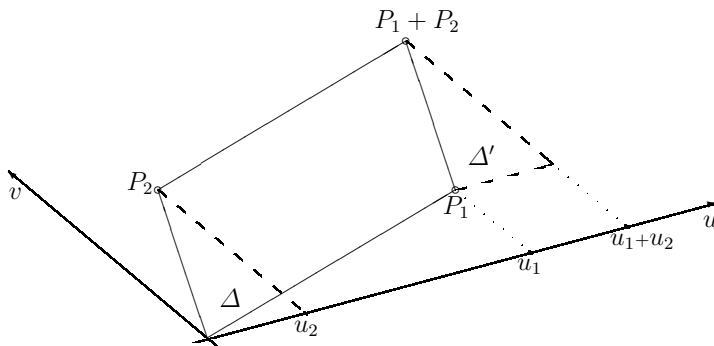
$$\varphi(\alpha P) = (\alpha x_1, \alpha y_1) = \alpha(x_1, y_1), \quad \psi(\alpha P) = (\alpha u_1, \alpha v_1) = \alpha(u_1, v_1),$$

und damit

$$f(\alpha(x_1, y_1)) = \psi(\varphi^{-1}(\alpha(x_1, y_1))) = \psi(\alpha P) = \alpha(u_1, v_1) = \alpha f((x_1, y_1)),$$

d. h. f ist verträglich mit der skalaren Multiplikation von $\mathbb{R}^2$.

Um auch die Verträglichkeit mit der Addition nachzuweisen, betrachten wir zunächst folgende Skizze:



Der Übersichtlichkeit halber haben wir uns hier auf ein einziges (schiefwinkliges) Koordinatensystem mit den Achsen u und v beschränkt. Ausgehend von den Punkten $P_1, P_2 \in E$ sei der Punkt $P_1 + P_2$ mittels der üblichen Parallelogrammkonstruktion definiert. Die Dreiecke Δ und Δ' , die jeweils die gestrichelten Strecken enthalten, erkennt man dann als kongruent, und es folgt, dass sich die u -Koordinate von $P_1 + P_2$ als Summe der u -Koordinaten von P_1 und P_2 ergibt. Entsprechendes gilt für die v -Koordinaten und in gleicher Weise für die Koordinaten bezüglich anderer Koordinatensysteme.

Gehen wir nun von zwei Punkten $(x_1, y_1), (x_2, y_2) \in \mathbb{R}^2$ aus und betrachten die zugehörigen Punkte $P_1, P_2 \in E$ mit $\varphi(P_i) = (x_i, y_i)$ für $i = 1, 2$, so ergibt sich mit $\psi(P_i) = (u_i, v_i)$ aufgrund vorstehender Beobachtung

$$\begin{aligned}\varphi(P_1 + P_2) &= (x_1 + x_2, y_1 + y_2) = (x_1, y_1) + (x_2, y_2), \\ \psi(P_1 + P_2) &= (u_1 + u_2, v_1 + v_2) = (u_1, v_1) + (u_2, v_2).\end{aligned}$$

Die gewünschte Verträglichkeit von f mit der Addition auf $\mathbb{R}^2$ folgt dann aus der Gleichung

$$\begin{aligned}f((x_1, y_1) + (x_2, y_2)) &= \psi(\varphi^{-1}((x_1, y_1) + (x_2, y_2))) = \psi(P_1 + P_2) \\ &= (u_1, v_1) + (u_2, v_2) = f((x_1, y_1)) + f((x_2, y_2)),\end{aligned}$$

und man erkennt $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, wie behauptet, als Isomorphismus von $\mathbb{R}$ -Vektorräumen.

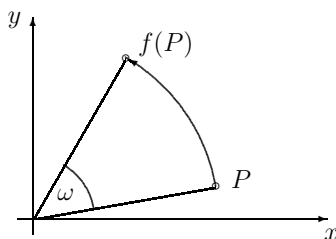
Die vorstehende Überlegung hat eine wichtige Konsequenz. Zeichnet man in der Ebene E einen Punkt 0 als Nullpunkt aus, so lässt sich E auf ganz natürliche Weise mit der Struktur eines $\mathbb{R}$ -Vektorraums versehen. Man wähle nämlich ein Koordinatensystem in 0 und fasse die zugehörige Bijektion $\varphi: E \rightarrow \mathbb{R}^2$, welche einem Punkt $P \in E$ die zugehörigen Koordinaten zuordnet, als Identifizierung auf. Die so von $\mathbb{R}^2$ auf E übertragene Vektorraumstruktur ist unabhängig von der Wahl des Koordinatensystems von E in 0 . In der Tat, ist $\psi: E \rightarrow \mathbb{R}^2$

eine Bijektion, die zu einem weiteren Koordinatensystem von E in 0 korrespondiert, so haben wir gerade gezeigt, dass die Komposition $\psi \circ \varphi^{-1}: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ ein Isomorphismus ist. Dies besagt aber, dass die mittels φ von $\mathbb{R}^2$ auf E übertragene Addition und skalare Multiplikation jeweils mit derjenigen übereinstimmt, die mittels ψ von $\mathbb{R}^2$ auf E übertragen werden kann. Wir können daher mit gutem Recht sagen, dass eine punktierte Ebene, also eine Ebene mit einem als Nullpunkt ausgezeichneten Punkt 0 , auf natürliche Weise die Struktur eines $\mathbb{R}$ -Vektorraums besitzt und dass dieser Vektorraum im Grunde genommen nichts anderes als das wohlbekannte Modell $\mathbb{R}^2$ ist, genauer, dass E als $\mathbb{R}$ -Vektorraum zu $\mathbb{R}^2$ isomorph ist. Entsprechendes gilt natürlich auch für eine Gerade und $\mathbb{R}^1$ als Modell, sowie für den drei-dimensionalen anschaulichen Raum und $\mathbb{R}^3$ als Modell. Es ist allgemeiner auch möglich, die Abhängigkeit der Modelle von der Wahl des Nullpunktes zu vermeiden. Anstelle linearer Abbildungen hat man dann so genannte *affine* Abbildungen zu betrachten, die sich als Komposition linearer Abbildungen mit Translationen darstellen.

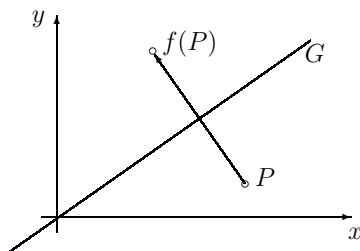
Beispiele für Isomorphismen $f: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ gibt es zur Genüge, etwa die Streckung

$$f: \mathbb{R}^2 \longrightarrow \mathbb{R}^2, \quad (x_1, y_1) \longmapsto (\alpha_1 x_1, \beta_1 y_1),$$

mit fest vorgegebenen Konstanten $\alpha_1, \beta_1 \in \mathbb{R}^*$, die Drehung um $0 \in \mathbb{R}^2$ mit einem gewissen Winkel ω



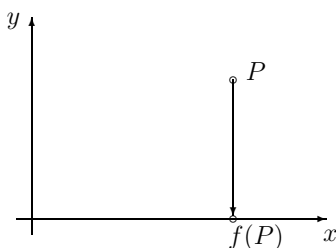
oder die Spiegelung an einer Geraden G , die den Nullpunkt $0 \in \mathbb{R}$ enthält:



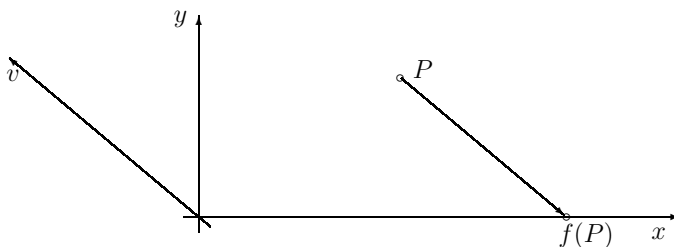
Ein einfaches Beispiel einer linearen Abbildung, die nicht bijektiv und damit kein Isomorphismus ist, stellt die so genannte *Projektion*

$$f: \mathbb{R}^2 \longrightarrow \mathbb{R}, \quad (x_1, y_1) \longmapsto x_1,$$

dar. Diese Abbildung beschreibt sich in der Tat als (Parallel-) Projektion auf die x -Achse, indem man einen Punkt $P \in \mathbb{R}^2$ auf den Schnittpunkt der x -Achse mit der Parallelen zur y -Achse durch P abbildet:



Man kann auch parallel zu einer anderen Achse v projizieren, etwa in der folgenden Weise:



Im Prinzip erhält man hier keine wesentlich andere lineare Abbildung, denn diese wird bezüglich der Koordinatenachsen x, v wiederum durch

$$f: \mathbb{R}^2 \longrightarrow \mathbb{R}, \quad (x_1, v_1) \longmapsto x_1,$$

beschrieben. Ähnliche Beispiele für Isomorphismen oder, allgemeiner, lineare Abbildungen, lassen sich auch für höher-dimensionale Vektorräume $\mathbb{R}^n$ angeben.

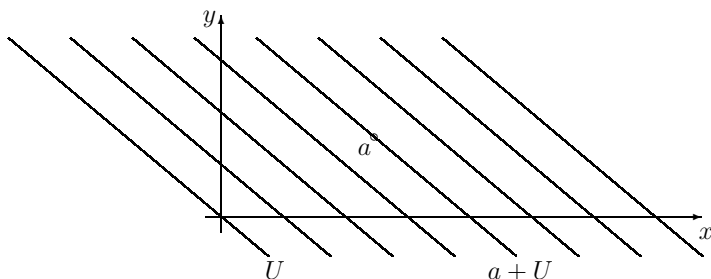
Im vorliegenden Kapitel wollen wir zunächst einmal einige allgemeine Eigenschaften linearer Abbildungen untersuchen, wobei wir wiederum Vektorräume über einem beliebigen Körper K betrachten. Zum Beispiel werden wir sehen, dass die Dimension des Bildes V' einer surjektiven linearen Abbildung $f: V \longrightarrow V'$, man spricht hier vom *Rang* von f , höchstens gleich der Dimension von V sein kann. Somit ist die Existenz beispielsweise einer surjektiven linearen Abbildung $\mathbb{R}^2 \longrightarrow \mathbb{R}^3$ ausgeschlossen. Allgemeiner werden Phänomene dieser Art durch die so genannte *Dimensionsformel* für lineare Abbildungen geregelt. Auch wird sich zeigen, dass eine lineare Abbildung $f: V \longrightarrow V'$ bereits eindeutig durch die Bilder $f(x_i) \in V'$ einer vorgegebenen Basis $(x_i)_{i \in I}$ von V festgelegt ist, ja dass man sogar eine lineare Abbildung $f: V \longrightarrow V'$ eindeutig definieren kann, indem man lediglich die Bilder $f(x_i)$ in V' (in beliebiger Weise) vorgibt. Diese Eigenschaft ist wichtig für die Beschreibung linearer Abbildungen mittels Matrizen (das sind rechteckige Koeffizientenschemata mit Einträgen aus dem Grundkörper K) und damit für die rechnerische Handhabung linearer Abbildungen. Eine tiefergehende Betrachtung von Matrizen wird allerdings erst in Kapitel 3 erfolgen.

Ein weiteres Problem, das wir in diesem Kapitel lösen werden, beschäftigt sich mit der Konstruktion linearer Abbildungen $f: V \longrightarrow V'$ mit vorgegebenen Fasern $f^{-1}(f(a)) \subset V$, wobei a in V variiere. Man sieht leicht ein, dass für eine lineare Abbildung $f: V \longrightarrow V'$ die Faser zu $a = 0$, also $f^{-1}(f(0)) = f^{-1}(0)$

einen linearen Unterraum $U \subset V$ definiert, den so genannten *Kern* von f , der mit $\ker f$ bezeichnet wird, und dass für $a \in V$

$$f^{-1}(f(a)) = a + U = \{a + u; u \in U\}$$

gilt. Man erhält daher die Fasern von f , indem man den linearen Unterraum $U = \ker f$ mit allen Vektoren $a \in V$ parallel verschiebt, wobei man jeden solchen parallel verschobenen Unterraum $A = a + U$ als einen zu U *gehörigen affinen Unterraum* von V bezeichnet. Im Falle $V = \mathbb{R}^2$ und für eine Gerade $U \subset \mathbb{R}^2$ ergibt sich beispielsweise folgendes Bild:



Umgekehrt werden wir zu einem linearen Unterraum $U \subset V$ die Menge V/U aller zu U gehörigen affinen Unterräume betrachten und diese auf natürliche Weise zu einem K -Vektorraum machen, indem wir für $a, b \in V$ und $\alpha \in K$

$$(a + U) + (b + U) := (a + b) + U, \quad \alpha(a + U) := \alpha a + U$$

setzen. Dass man auf diese Weise tatsächlich einen Vektorraum erhält, den so genannten *Quotienten-* oder *Restklassenvektorraum* V/U , bedarf einiger Verifizierungen, die wir im Einzelnen durchführen werden. Insbesondere werden wir sehen, dass dann die kanonische Abbildung

$$\pi: V \longrightarrow V/U, \quad a \longmapsto a + U,$$

eine lineare Abbildung mit $\ker \pi = U$ ergibt, deren Fasern also gerade die zu U gehörigen affinen Unterräume von V sind. Die Abbildung π erfüllt eine wichtige, so genannte *universelle* Eigenschaft: Ist $f: V \longrightarrow V'$ eine lineare Abbildung mit einem Kern, der $U = \ker \pi \subset \ker f$ erfüllt, so zerlegt sich f in eine Komposition

$$f: V \xrightarrow{\pi} V/U \xrightarrow{\bar{f}} V'$$

mit einer eindeutig bestimmten linearen Abbildung $\bar{f}: V/U \longrightarrow V'$; dies ist im Wesentlichen die Aussage des *Homomorphiesatzes* für lineare Abbildungen.

Die Definition eines Vektorraums als Menge mit einer Addition und skalarer Multiplikation macht keinerlei Vorschriften über die Art der Elemente dieser Menge. Wir nutzen dies insbesondere bei der Definition des Restklassenvektorraums V/U aus, dessen Elemente affine Unterräume von V und damit

Teilmengen von V sind. Ein weiteres Beispiel eines Vektorraums, der zu einem gegebenen K -Vektorraum V konstruiert werden kann, ist der *Dualraum* V^* . Seine Elemente sind lineare Abbildungen, und zwar die linearen Abbildungen $V \rightarrow K$. Zwischen V und V^* bestehen enge symmetrische Beziehungen; man spricht von einer Dualität zwischen V und V^* , daher auch der Name Dualraum. Beispielsweise lässt sich V im Falle endlicher Dimension selbst wieder als Dualraum von V^* interpretieren. Mit dem Dualraum eines Vektorraums V decken wir in einem gewissen Sinne die linearen Eigenschaften von V auf einem höheren Niveau auf. Die zu beweisenden Ergebnisse werden es uns in gewissen Fällen ermöglichen, ansonsten erforderliche konventionelle Rechnungen durch konzeptionelle Argumente zu ersetzen.

2.1 Grundbegriffe

Zu Vektoren $a_1, \dots, a_n$ eines K -Vektorraums V kann man stets die Abbildung

$$f: K^n \rightarrow V, \quad (\alpha_1, \dots, \alpha_n) \mapsto \sum_{i=1}^n \alpha_i a_i,$$

betrachten. Für die Einheitsvektoren $e_i = (\delta_{1i}, \dots, \delta_{ni}) \in K^n$, $i = 1, \dots, n$, welche die kanonische Basis von K^n bilden, gilt dann $f(e_i) = a_i$, und wir können in äquivalenter Weise sagen, dass f durch die Vorschrift

$$f\left(\sum_{i=1}^n \alpha_i e_i\right) = \sum_{i=1}^n \alpha_i a_i$$

beschrieben wird. Bilden nun $a_1, \dots, a_n$ sogar eine Basis von V , so hat jedes Element $a \in V$ eine eindeutig bestimmte Darstellung $a = \sum_{i=1}^n \alpha_i a_i$ mit Koeffizienten $\alpha_i \in K$, und man sieht, dass in diesem Falle die Abbildung f bijektiv ist. Da weiter f verträglich mit den Vektorraumstrukturen auf K^n und V ist — es gilt nämlich $f(a + b) = f(a) + f(b)$ für $a, b \in K^n$ sowie $f(\alpha a) = \alpha f(a)$ für $\alpha \in K$, $a \in K^n$ —, können wir V nach Auswahl der Basis $a_1, \dots, a_n$ unter Verwendung der Abbildung f mit K^n identifizieren. Dies zeigt insbesondere, dass Abbildungen zwischen Vektorräumen, welche die Vektorraumstrukturen respektieren, von Interesse sind.

Definition 1. Eine Abbildung $f: V \rightarrow V'$ zwischen K -Vektorräumen V, V' heißt K -Homomorphismus oder K -lineare Abbildung, falls gilt:

- (i) $f(a + b) = f(a) + f(b)$ für $a, b \in V$.
- (ii) $f(\alpha a) = \alpha f(a)$ für $\alpha \in K$, $a \in V$.

Die Bedingungen (i) und (ii) lassen sich zusammenfassen, indem man für $\alpha, \beta \in K$, $a, b \in V$ in äquivalenter Weise fordert:

$$f(\alpha a + \beta b) = \alpha f(a) + \beta f(b)$$

Als einfache Rechenregeln prüft man leicht nach:

$$\begin{aligned} f(0) &= 0 \\ f(-a) &= -f(a) \quad \text{für } a \in V \end{aligned}$$

Im Übrigen ist die Komposition linearer Abbildungen wieder linear. Wir wollen einige einfache Beispiele linearer Abbildungen anschauen.

(1) Die Abbildung $\mathbb{C} \longrightarrow \mathbb{C}$, $z \longmapsto \bar{z} = \operatorname{Re}(z) - i\operatorname{Im}(z)$, ist $\mathbb{R}$ -linear, nicht aber $\mathbb{C}$ -linear.

(2) Für einen K -Vektorraum V ist die identische Abbildung $\operatorname{id}: V \longrightarrow V$, $a \longmapsto a$, ein Beispiel einer K -linearen Abbildung, ebenso die Nullabbildung $0: V \longrightarrow 0$, welche jedes Element $a \in V$ auf 0 abbildet. Weiter ist für einen linearen Unterraum $U \subset V$ die Inklusionsabbildung $U \hookrightarrow V$ ein Beispiel einer K -linearen Abbildung.

(3) Es sei K ein Körper. Zu $m, n \in \mathbb{N}$ betrachte man ein System

$$(\lambda_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} = \begin{pmatrix} \lambda_{11} & \dots & \lambda_{1n} \\ \cdot & \dots & \cdot \\ \lambda_{m1} & \dots & \lambda_{mn} \end{pmatrix}$$

von Elementen aus K ; man spricht von einer *Matrix*. Dann wird durch

$$K^n \longrightarrow K^m, \quad (\alpha_1, \dots, \alpha_n) \longmapsto \left(\sum_{j=1}^n \lambda_{1j} \alpha_j, \dots, \sum_{j=1}^n \lambda_{mj} \alpha_j \right),$$

eine K -lineare Abbildung gegeben. Wir werden sogar im Weiteren sehen, dass jede lineare Abbildung $K^n \longrightarrow K^m$ von dieser Gestalt ist, also durch eine Matrix von Elementen aus K beschrieben werden kann. Allerdings werden wir dann Vektoren in K^n bzw. K^m in konsequenter Weise als *Spalten*-Vektoren und nicht mehr wie bisher als Zeilen-Vektoren schreiben, da dies besser mit dem dann einzuführenden Matrizenprodukt harmonisiert.

Man nennt eine K -lineare Abbildung $f: V \longrightarrow V'$ zwischen Vektorräumen einen *Monomorphismus*, falls f injektiv ist, einen *Epimorphismus*, falls f surjektiv ist, und einen *Isomorphismus*, falls f bijektiv ist.

Bemerkung 2. Ist $f: V \longrightarrow V'$ ein Isomorphismus zwischen K -Vektorräumen, so existiert die Umkehrabbildung $f^{-1}: V' \longrightarrow V$, und diese ist K -linear, also wiederum ein Isomorphismus.

Beweis. Es ist nur die K -Linearität der Umkehrabbildung f^{-1} zu zeigen. Seien also $a', b' \in V'$ gegeben. Für $a = f^{-1}(a')$ und $b = f^{-1}(b')$ hat man dann $a' = f(a)$ und $b' = f(b)$, sowie $a' + b' = f(a + b)$ aufgrund der Linearität von f . Folglich gilt

$$f^{-1}(a' + b') = a + b = f^{-1}(a') + f^{-1}(b').$$

Für $\alpha \in K$ gilt weiter $\alpha a' = f(\alpha a)$ und daher

$$f^{-1}(\alpha a') = \alpha a = \alpha f^{-1}(a'),$$

d. h. f^{-1} ist K -linear. □

Isomorphismen schreiben wir häufig in der Form $V \xrightarrow{\sim} V'$. Im Falle $V = V'$ bezeichnet man eine K -lineare Abbildung $f: V \rightarrow V$ auch als einen *Endomorphismus* von V , und man versteht unter einem *Automorphismus* von V einen bijektiven Endomorphismus.

Bemerkung 3. *Es sei $f: V \rightarrow V'$ eine K -lineare Abbildung zwischen Vektorräumen. Dann sind*

$$\begin{aligned}\ker f &= f^{-1}(0) = \{a \in V; f(a) = 0\} \quad \text{und} \\ \operatorname{im} f &= f(V) = \{f(a); a \in V\}\end{aligned}$$

lineare Unterräume von V bzw. V' . Diese werden als Kern bzw. Bild von f bezeichnet.

Der *Beweis* ist einfach zu führen. Für $a, b \in \ker f$ hat man $f(a + b) = f(a) + f(b) = 0 + 0 = 0$, also $a + b \in \ker f$. Weiter gilt $f(\alpha a) = \alpha f(a) = \alpha 0 = 0$ und damit $\alpha a \in \ker f$ für $\alpha \in K$ und $a \in \ker f$. Da $\ker f \neq \emptyset$ wegen $0 \in \ker f$, erkennt man $\ker f$ als linearen Unterraum von V .

Ähnlich sieht man, dass $\operatorname{im} f$ ein linearer Unterraum von V' ist. Wegen $0 \in \operatorname{im} f$ ist jedenfalls $\operatorname{im} f$ nicht leer. Seien weiter $\alpha \in K$, $a', b' \in \operatorname{im} f$, etwa $a' = f(a)$, $b' = f(b)$ mit $a, b \in V$. Dann folgt $\alpha a' = f(\alpha a)$ sowie $a' + b' = f(a + b)$, d. h. $\alpha a', a' + b' \in \operatorname{im} f$. Also ist auch $\operatorname{im} f$ ein linearer Unterraum von V' . □

Bemerkung 4. *Eine K -lineare Abbildung $f: V \rightarrow V'$ zwischen Vektorräumen ist genau dann injektiv, wenn $\ker f = 0$ gilt.*

Beweis. Sei zunächst f injektiv. Dann besteht insbesondere $\ker f = f^{-1}(0)$ aus höchstens einem Element, und wegen $0 \in f^{-1}(0)$ folgt $\ker f = 0$. Sei nun umgekehrt die Beziehung $\ker f = 0$ gegeben, und seien $a, b \in V$ mit $f(a) = f(b)$. Dann ergibt sich

$$f(a - b) = f(a) - f(b) = 0,$$

also $a - b \in \ker f = 0$ und damit $a = b$, d. h. f ist injektiv. □

Wir wollen weiter untersuchen, wie sich Erzeugendensysteme sowie linear abhängige bzw. unabhängige Systeme unter Anwendung linearer Abbildungen verhalten.

Bemerkung 5. *Es sei $f: V \rightarrow V'$ eine K -lineare Abbildung zwischen Vektorräumen.*

(i) *Für Teilmengen $A \subset V$ gilt $f(\langle A \rangle) = \langle f(A) \rangle$.*

(ii) Es seien die Vektoren $a_1, \dots, a_n \in V$ linear abhängig. Dann sind auch deren Bilder $f(a_1), \dots, f(a_n) \in V'$ linear abhängig. Die Umkehrung hierzu gilt, wenn f injektiv ist.

(iii) Es seien $a_1, \dots, a_n \in V$ Vektoren, deren Bilder $f(a_1), \dots, f(a_n) \in V'$ linear unabhängig sind. Dann sind auch $a_1, \dots, a_n$ linear unabhängig. Die Umkehrung hierzu gilt, wenn f injektiv ist.

Der Beweis kann durch einfache Verifikation der Definitionen geführt werden. Die entsprechenden Rechnungen seien jedoch dem Leser überlassen. Als Konsequenz von Aussage (iii) vermerken wir noch:

Bemerkung 6. Für eine K -lineare Abbildung $f: V \longrightarrow V'$ zwischen Vektorräumen gilt $\dim_K f(V) \leq \dim_K V$.

Eine wichtige Eigenschaft linearer Abbildungen besteht darin, dass sie bereits durch die Werte, die sie auf einer Basis des Urbildraumes annehmen, festgelegt sind.

Satz 7. Es sei V ein K -Vektorraum mit Erzeugendensystem $a_1, \dots, a_n$. Sind dann $a'_1, \dots, a'_n$ beliebige Vektoren eines weiteren K -Vektorraums V' , so gilt:

(i) Es gibt höchstens eine K -lineare Abbildung $f: V \longrightarrow V'$ mit $f(a_i) = a'_i$, $i = 1, \dots, n$.

(ii) Ist $a_1, \dots, a_n$ sogar eine Basis von V , so existiert genau eine K -lineare Abbildung $f: V \longrightarrow V'$ mit $f(a_i) = a'_i$ für $i = 1, \dots, n$.

Beweis. Wir beginnen mit der Eindeutigkeitsaussage (i). Sei also $f: V \longrightarrow V'$ K -linear mit $f(a_i) = a'_i$, $i = 1, \dots, n$, und sei $a \in V$. Dann besitzt a eine Darstellung der Form $a = \sum_{i=1}^n \alpha_i a_i$ mit Koeffizienten $\alpha_i \in K$. Aufgrund der Linearität von f folgt

$$f(a) = f\left(\sum_{i=1}^n \alpha_i a_i\right) = \sum_{i=1}^n \alpha_i f(a_i),$$

was bedeutet, dass $f(a)$ durch die Werte $f(a_i)$, $i = 1, \dots, n$, eindeutig bestimmt ist.

Bilden nun $a_1, \dots, a_n$ im Falle (ii) eine Basis von V , so sind für $a \in V$ die Koeffizienten α_i in der Darstellung $a = \sum_{i=1}^n \alpha_i a_i$ eindeutig bestimmt, und man kann eine Abbildung $f: V \longrightarrow V'$ erklären, indem man setzt:

$$f(a) = \sum_{i=1}^n \alpha_i f(a_i)$$

Diese Abbildung ist K -linear. Sind nämlich $a = \sum_{i=1}^n \alpha_i a_i$ und $b = \sum_{i=1}^n \beta_i a_i$ zwei Elemente von V , so gilt

$$a + b = \sum_{i=1}^n (\alpha_i + \beta_i) a_i.$$

Da dies die (eindeutig bestimmte) Darstellung von $a + b$ als Linearkombination von $a_1, \dots, a_n$ ist, ergibt sich

$$f(a + b) = \sum_{i=1}^n (\alpha_i + \beta_i) a'_i = \sum_{i=1}^n \alpha_i a'_i + \sum_{i=1}^n \beta_i a'_i = f(a) + f(b).$$

Entsprechend rechnet man für $\alpha \in K$

$$f(\alpha a) = f\left(\sum_{i=1}^n \alpha \alpha_i a_i\right) = \sum_{i=1}^n \alpha \alpha_i a'_i = \alpha \sum_{i=1}^n \alpha_i a'_i = \alpha f(a),$$

und man sieht, dass $f: V \rightarrow V'$ K -linear ist. Es existiert also eine K -lineare Abbildung $f: V \rightarrow V'$ mit $f(a_i) = a'_i$, und diese ist eindeutig bestimmt, wie wir in (i) gesehen haben. $\square$

Für zwei K -Vektorräume V und V' bilden die K -linearen Abbildungen $f: V \rightarrow V'$ einen K -Vektorraum, sozusagen den *Vektorraum aller K -linearen V' -wertigen Funktionen auf V* ; dieser wird meist mit $\text{Hom}_K(V, V')$ bezeichnet, also

$$\text{Hom}_K(V, V') = \{f: V \rightarrow V'; f \text{ } K\text{-linear}\}.$$

Dabei ist für $f, g \in \text{Hom}_K(V, V')$ die Summe $f + g$ durch

$$f + g: V \rightarrow V', \quad x \mapsto f(x) + g(x),$$

erklärt und entsprechend für $\alpha \in K$ das Produkt $\alpha \cdot f$ durch

$$\alpha \cdot f: V \rightarrow V', \quad x \mapsto \alpha \cdot (f(x)).$$

Dass $f + g$ und $\alpha \cdot f$ wiederum K -lineare Abbildungen von V nach V' darstellen, ist mit leichter Rechnung nachzuprüfen. Der Sachverhalt von Aussage (ii) in Satz 7 ist dann präziser so zu formulieren, dass für eine Basis $a_1, \dots, a_n$ von V die Zuordnung $f \mapsto (f(a_1), \dots, f(a_n))$ Anlass zu einem Isomorphismus von K -Vektorräumen $\text{Hom}_K(V, V') \xrightarrow{\sim} (V')^n$ gibt.

Diese Korrespondenz kann man noch konkreter beschreiben, wenn man neben der Basis $a_1, \dots, a_n$ von V auch in V' eine Basis fixiert. Sei also $b_1, \dots, b_m$ eine Basis von V' , wobei wir V' als endlich-dimensional annehmen. Dann besitzt jeder Vektor $a' \in V'$ eine eindeutige Darstellung $a' = \sum_{i=1}^m \lambda_i b_i$, ist also durch das Koeffiziententupel $(\lambda_1, \dots, \lambda_m)$ auf umkehrbar eindeutige Weise bestimmt. Eine lineare Abbildung $f: V \rightarrow V'$ korrespondiert daher zu n solchen Koeffiziententupeln $(\lambda_{1j}, \dots, \lambda_{mj})$, $j = 1, \dots, n$, die wir als Spalten interpretieren und zu einer Matrix zusammenfügen:

$$(\lambda_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} = \begin{pmatrix} \lambda_{11} & \dots & \lambda_{1n} \\ \cdot & \dots & \cdot \\ \lambda_{m1} & \dots & \lambda_{mn} \end{pmatrix}$$

Man nennt dies die zu der linearen Abbildung f gehörige Matrix, wobei diese Bezeichnung natürlich relativ zu den in V und V' fixierten Basen zu verstehen ist. Die lineare Abbildung f lässt sich aus der Matrix (λ_{ij}) mit Hilfe der Gleichungen

$$f(a_j) = \sum_{i=1}^m \lambda_{ij} b_i, \quad j = 1, \dots, n.$$

bzw.

$$f\left(\sum_{j=1}^n \alpha_j a_j\right) = \sum_{i=1}^m \left(\sum_{j=1}^n \lambda_{ij} \alpha_j\right) b_i$$

rekonstruieren.

Satz 8. *Es seien V, V' K -Vektorräume.*

- (i) *Falls es einen Isomorphismus $f: V \xrightarrow{\sim} V'$ gibt, so stimmen die Dimensionen von V und V' überein, also $\dim_K V = \dim_K V'$.*
- (ii) *Umgekehrt existiert im Falle $\dim_K V = \dim_K V' < \infty$ stets ein Isomorphismus $f: V \xrightarrow{\sim} V'$.*

Beweis. Ist $f: V \xrightarrow{\sim} V'$ ein Isomorphismus, so sieht man mit Bemerkung 5, dass ein System von Vektoren $a_1, \dots, a_n \in V$ genau dann linear unabhängig ist, wenn die Bilder $f(a_1), \dots, f(a_n) \in V'$ linear unabhängig sind. Dies bedeutet aber $\dim_K V = \dim_K V'$.

Gilt umgekehrt $\dim_K V = \dim_K V' = n < \infty$, so wähle man zwei Basen $a_1, \dots, a_n \in V$ und $a'_1, \dots, a'_n \in V'$. Nach Satz 7 (ii) existiert dann eine eindeutig bestimmte K -lineare Abbildung $f: V \rightarrow V'$ mit $f(a_i) = a'_i$ für $i = 1, \dots, n$. Ist nun $a \in \ker f$, etwa $a = \sum_{i=1}^n \alpha_i a_i$ mit Koeffizienten $\alpha_i \in K$, so folgt $\sum_{i=1}^n \alpha_i a'_i = f(a) = 0$ und wegen der linearen Unabhängigkeit von $a'_1, \dots, a'_n$ bereits $\alpha_i = 0$ für alle i , also $a = 0$. Damit ist f injektiv. Aber f ist auch surjektiv. Denn Elemente $a' \in V'$ kann man stets in der Form $a' = \sum_{i=1}^n \alpha_i a'_i$ mit Koeffizienten $\alpha_i \in K$ darstellen, und es gilt dann $f(\sum_{i=1}^n \alpha_i a_i) = a'$. $\square$

Korollar 9. *Ist V ein K -Vektorraum der Dimension $n < \infty$, so gibt es einen Isomorphismus $V \xrightarrow{\sim} K^n$.*

Dies bedeutet, dass es bis auf Isomorphie als endlich-dimensionale K -Vektorräume nur die Vektorräume K^n gibt. Wir wollen abschließend noch die so genannte Dimensionsformel für K -lineare Abbildungen beweisen.

Satz 10. *Es sei $f: V \rightarrow V'$ eine K -lineare Abbildung zwischen Vektorräumen. Dann gilt*

$$\dim_K V = \dim_K(\ker f) + \dim_K(\operatorname{im} f).$$

Statt $\dim_K(\operatorname{im} f)$, also der Dimension des Bildraumes $f(V)$, schreibt man häufig auch $\operatorname{rg} f$ und nennt dies den *Rang* von f .

Beweis zu Satz 10. Ist einer der Vektorräume $\ker f$ oder $\operatorname{im} f$ von unendlicher Dimension, so auch V ; falls $\dim_K(\operatorname{im} f) = \infty$, so folgt dies aus Bemerkung 5.

Wir dürfen daher $\ker f$ und $\operatorname{im} f$ als endlich-dimensional ansehen. Man wähle dann Vektoren $a_1, \dots, a_m$ in V , so dass deren Bilder $f(a_1), \dots, f(a_m)$ eine Basis von $\operatorname{im} f$ bilden. Weiter wähle man eine Basis $a_{m+1}, \dots, a_n$ von $\ker f$. Es genügt dann nachzuweisen, dass $a_1, \dots, a_n$ eine Basis von V bilden. Hierzu betrachte man einen Vektor $a \in V$. Es existiert eine Darstellung $f(a) = \sum_{i=1}^m \alpha_i f(a_i)$ mit Koeffizienten $\alpha_i \in K$, da $f(a_1), \dots, f(a_m)$ den Vektorraum $\operatorname{im} f$ erzeugen. Weiter liegt der Vektor $a - \sum_{i=1}^m \alpha_i a_i$ im Kern von f , und, da dieser von $a_{m+1}, \dots, a_n$ erzeugt wird, gibt es $\alpha_{m+1}, \dots, \alpha_n \in K$ mit $a - \sum_{i=1}^m \alpha_i a_i = \sum_{i=m+1}^n \alpha_i a_i$, also mit $a = \sum_{i=1}^n \alpha_i a_i$. Dies zeigt, dass $a_1, \dots, a_n$ ein Erzeugendensystem von V bilden.

Um zu sehen, dass $a_1, \dots, a_n$ auch linear unabhängig sind, betrachte man eine Relation $\sum_{i=1}^n \alpha_i a_i = 0$ mit Koeffizienten $\alpha_i \in K$. Hieraus ergibt sich

$$0 = f\left(\sum_{i=1}^n \alpha_i a_i\right) = \sum_{i=1}^n \alpha_i f(a_i) = \sum_{i=1}^m \alpha_i f(a_i)$$

und damit $\alpha_1 = \dots = \alpha_m = 0$, da $f(a_1), \dots, f(a_m)$ linear unabhängig sind. Dann folgt aber $\sum_{i=m+1}^n \alpha_i a_i = 0$ und weiter $\alpha_{m+1} = \dots = \alpha_n = 0$, da $a_{m+1}, \dots, a_n$ linear unabhängig sind. Die Koeffizienten $\alpha_1, \dots, \alpha_n$ verschwinden daher sämtlich, und wir sehen, dass $a_1, \dots, a_n$ linear unabhängig sind. $\square$

Korollar 11. *Es sei $f: V \longrightarrow V'$ eine K -lineare Abbildung zwischen Vektorräumen mit $\dim_K V = \dim_K V' < \infty$. Dann ist äquivalent:*

- (i) *f ist ein Monomorphismus.*
- (ii) *f ist ein Epimorphismus.*
- (iii) *f ist ein Isomorphismus.*

Beweis. Sei zunächst Bedingung (i) erfüllt, also f ein Monomorphismus. Dann folgt unter Benutzung von Satz 10

$$\dim_K V' = \dim_K V = \dim_K(\ker f) + \dim_K(\operatorname{im} f).$$

Wegen $\dim_K(\ker f) = 0$ besitzt daher der lineare Unterraum $\operatorname{im} f \subset V'$ dieselbe (endliche) Dimension wie V' . Mit 1.5/14 (ii) ergibt sich $\operatorname{im} f = V'$, d. h. f ist ein Epimorphismus, und (ii) ist erfüllt.

Sei nun f mit Bedingung (ii) als Epimorphismus vorausgesetzt. Die Dimensionsformel in Satz 10 ergibt dann wegen $\operatorname{im} f = V'$ und $\dim_K V' = \dim_K V$ notwendig $\ker f = 0$, d. h. f ist injektiv und erfüllt damit Bedingung (iii). Dass schließlich (iii) die Bedingung (i) impliziert, ist trivial. $\square$

Auch hier sei darauf hingewiesen, dass Korollar 11 ähnlich wie 1.5/14 (ii) nicht auf den Fall unendlich-dimensionaler Vektorräume zu verallgemeinern ist, da ein unendlich-dimensionaler K -Vektorraum stets echte lineare Unterräume unendlicher Dimension besitzt.

Aufgaben

- Für einen Körper K und ein $n \in \mathbb{N}$ betrachte man K^n als K -Vektorraum. Es bezeichne $p_i: K^n \rightarrow K$ für $i = 1, \dots, n$ jeweils die Projektion auf die i -te Komponente. Man zeige:
 - Die Abbildungen p_i sind K -linear.
 - Eine Abbildung $f: V \rightarrow K^n$ von einem K -Vektorraum V nach K^n ist genau dann K -linear, wenn alle Kompositionen $p_i \circ f$ K -linear sind.
- Gibt es $\mathbb{R}$ -lineare Abbildungen $\mathbb{R}^4 \rightarrow \mathbb{R}^3$ die die folgenden Vektoren $a_i \in \mathbb{R}^4$ jeweils auf die angegebenen Vektoren $b_i \in \mathbb{R}^3$ abbilden?
 - $a_1 = (1, 1, 0, 0), \quad a_2 = (1, 1, 1, 0), \quad a_3 = (0, 1, 1, 1), \quad a_4 = (0, 0, 1, 1)$
 $b_1 = (1, 2, 3), \quad b_2 = (2, 3, 1), \quad b_3 = (3, 1, 2), \quad b_4 = (2, 0, 4)$
 - $a_1 = (0, 1, 1, 1), \quad a_2 = (1, 0, 1, 1), \quad a_3 = (1, 1, 0, 1)$
 b_1, b_2, b_3 wie in (i)
 - $a_1 = (0, 1, 1, 1), \quad a_2 = (1, 0, 1, 1), \quad a_3 = (1, 1, 0, 1), \quad a_4 = (-1, 1, 0, 0)$
 b_1, b_2, b_3, b_4 wie in (i)
 - $a_1 = (0, 1, 1, 1), \quad a_2 = (1, 0, 1, 1), \quad a_3 = (1, 1, 0, 1), \quad a_4 = (0, 2, 0, 1)$
 b_1, b_2, b_3, b_4 wie in (i)
- Man bestimme alle $\mathbb{R}$ -linearen Abbildungen $\mathbb{R} \rightarrow \mathbb{R}$.
- Man bestimme alle Körperhomomorphismen $f: \mathbb{Q}(\sqrt{2}) \rightarrow \mathbb{Q}(\sqrt{2})$, d. h. alle $\mathbb{Q}$ -linearen Abbildungen, die zusätzlich $f(\alpha\beta) = f(\alpha)f(\beta)$ erfüllen.
- Es sei V ein K -Vektorraum und $f: V \rightarrow V$ ein Endomorphismus mit $f^2 = f$. Man zeige $V = \ker f \oplus \operatorname{im} f$.
- Für lineare Unterräume U, U' eines K -Vektorraums V betrachte man die Abbildung

$$\varphi: U \times U' \rightarrow V, \quad (a, b) \mapsto a - b.$$

- Man zeige, dass φ K -linear ist, wenn man $U \times U'$ mit komponentenweiser Addition und skalarer Multiplikation als K -Vektorraum auffasst.
- Man berechne $\dim_K(U \times U')$.
- Man wende die Dimensionsformel für lineare Abbildungen auf φ an und folgere die Dimensionsformel für lineare Unterräume von V

$$\dim_K U + \dim_K U' = \dim_K(U + U') + \dim_K(U \cap U').$$

- Für K -Vektorräume V, V' bestimme man $\dim_K \operatorname{Hom}_K(V, V')$, also die Dimension des Vektorraums aller K -linearen Abbildungen von V nach V' .
- Es seien $V_1 \xrightarrow{f} V_2 \xrightarrow{g} V_3$ K -lineare Abbildungen zwischen endlich-dimensionalen K -Vektorräumen. Man zeige:

$$\operatorname{rg} f + \operatorname{rg} g \leq \operatorname{rg}(g \circ f) + \dim V_2$$

2.2 Quotientenvektorräume

Neben den *linearen Unterräumen* von Vektorräumen V , also den Untervektorräumen im Sinne von 1.4/2, wollen wir in diesem Abschnitt auch so genannte *affine Unterräume* von V betrachten. Diese entstehen im Wesentlichen aus den linearen Unterräumen durch Translation (oder Parallelverschiebung) um Vektoren $a \in V$.

Definition 1. *Es sei V ein K -Vektorraum und $A \subset V$ eine Teilmenge. Man bezeichnet A als affinen Unterraum von V , wenn A leer ist, oder wenn es ein Element $a \in V$ und einen linearen Unterraum $U \subset V$ gibt mit*

$$A = a + U := \{a + u; u \in U\}.$$

In der Situation der Definition ergibt sich für $a \notin U$ insbesondere $0 \notin a + U$, und man sieht, dass $A = a + U$ in diesem Fall kein linearer Unterraum von V sein kann. Ziel dieses Abschnittes ist es zu zeigen, dass die Menge aller affinen Unterräume des Typs $a + U$, $a \in V$, in nahe liegender Weise einen K -Vektorraum bildet. Wir werden diesen Vektorraum mit V/U (man lese V modulo U) bezeichnen und zeigen, dass die Zuordnung $a \mapsto a + U$ eine surjektive K -lineare Abbildung $V \rightarrow V/U$ definiert, welche U als Kern besitzt.

Um das Problem etwas weiter zu verdeutlichen, wollen wir zunächst einmal die Existenz einer surjektiven K -linearen Abbildung $f: V \rightarrow V'$ mit $\ker f = U$ als gegeben annehmen. Für $a' \in V'$ nennt man $f^{-1}(a')$ die *Faser von f über a'* .

Bemerkung 2. *Es sei $f: V \rightarrow V'$ eine K -lineare Abbildung zwischen Vektorräumen mit $\ker f = U$. Dann erhält man für $a \in V$ als Faser von f über $f(a)$ gerade den affinen Unterraum $a + U$, d. h. $f^{-1}(f(a)) = a + U$.*

Beweis. Für $u \in U$ gilt $f(a+u) = f(a)$ und damit $a+U \subset f^{-1}(f(a))$. Umgekehrt folgt aus $a' \in f^{-1}(f(a))$ die Beziehung $f(a') = f(a)$ bzw. $f(a' - a) = 0$. Dies impliziert $a' - a \in \ker f = U$ bzw. $a' \in a + U$, was $f^{-1}(f(a)) \subset a + U$ und folglich die Gleichheit beider Mengen zeigt. $\square$

Wir können also vermerken, dass die Fasern einer surjektiven K -linearen Abbildung $f: V \rightarrow V'$ mit $\ker f = U$ gerade aus den affinen Unterräumen $a + U$, $a \in V$, bestehen. Da über jedem $a' \in f(V) = V'$ genau eine Faser liegt, nämlich $f^{-1}(a')$, können wir die Menge der affinen Unterräume des Typs $a + U \subset V$ mit V' identifizieren und zu einem K -Vektorraum V/U machen, indem wir die Vektorraumstruktur von V' übernehmen. Man addiert dann zwei Fasern $f^{-1}(f(a)) = a + U$, $f^{-1}(f(b)) = b + U$, indem man die Faser über $f(a) + f(b) = f(a + b)$ bildet, also

$$(a + U) + (b + U) = (a + b) + U.$$

Entsprechend multipliziert man eine Faser $f^{-1}(f(a)) = a + U$ mit einem Skalar $\alpha \in K$ indem man die Faser über $\alpha f(a) = f(\alpha a)$ bildet, also $\alpha(a+U) = \alpha a + U$. Schließlich ist klar, dass die Abbildung $V \longrightarrow V/U$, $a \longmapsto a + U$, ebenso wie $f: V \longrightarrow V'$, eine surjektive K -lineare Abbildung mit Kern U ist.

Die vorstehende Überlegung zeigt, dass das Problem, die Menge der affinen Unterräume des Typs $a + U \subset V$ zu einem K -Vektorraum zu machen, eine natürliche Lösung besitzt, wenn man über eine surjektive lineare Abbildung $f: V \longrightarrow V'$ mit $\ker f = U$ verfügt. Eine solche Abbildung kann man sich aber leicht verschaffen. Man wähle nämlich gemäß 1.6/4 ein Komplement U' zu U , also einen linearen Unterraum $U' \subset V$ mit $V = U \oplus U'$. Dann lässt sich jedes $a \in V$ auf eindeutige Weise in der Form $a = u + u'$ mit $u \in U$ und $u' \in U'$ schreiben. Indem wir a jeweils den Summanden u' zuordnen, erhalten wir wie gewünscht eine surjektive K -lineare Abbildung $p: V \longrightarrow U'$ mit Kern U . Man nennt p die Projektion von $U \oplus U'$ auf den zweiten Summanden.

Wir wollen im Folgenden jedoch den Quotientenvektorraum V/U auf eine andere Art konstruieren. Die verwendete Methode hat den Vorteil, dass sie nicht auf speziellen Eigenschaften von Vektorräumen (Existenz eines Komplements zu einem linearen Unterraum) beruht, sondern auch noch in anderen Situationen anwendbar ist.

Definition 3. Eine Relation auf einer Menge M besteht aus einer Teilmenge $R \subset M \times M$; man schreibt $a \sim b$ für $(a, b) \in R$ und spricht von der Relation “ $\sim$ ”. Eine Relation “ $\sim$ ” auf M heißt eine Äquivalenzrelation, wenn sie folgende Eigenschaften besitzt:

- (i) Reflexivität: $a \sim a$ für alle $a \in M$.
- (ii) Symmetrie: $a \sim b \implies b \sim a$.
- (iii) Transitivität: $a \sim b, b \sim c \implies a \sim c$.

Ist “ $\sim$ ” eine Äquivalenzrelation auf einer Menge M , so nennt man zwei Elemente $a, b \in M$ äquivalent, wenn für diese die Relation $a \sim b$ gilt. Weiter bezeichnet man

$$[a] = \{b \in M; b \sim a\}$$

als die Äquivalenzklasse des Elementes a in M . Um ein einfaches Beispiel zu geben, betrachte man eine K -lineare Abbildung $f: V \longrightarrow V'$ zwischen Vektorräumen und setze für $a, b \in V$

$$a \sim b \iff f(a) = f(b).$$

Es ist dann unmittelbar klar, dass f auf diese Weise eine Äquivalenzrelation auf V induziert und dass $[a] = f^{-1}(f(a))$ die Äquivalenzklasse eines Elementes $a \in V$ ist.

Im Falle einer Äquivalenzrelation “ $\sim$ ” auf einer Menge M heißt jedes Element b einer Äquivalenzklasse $[a]$ ein Repräsentant dieser Klasse. Aufgrund von Definition 3 (i) ist ein Element $a \in M$ stets Repräsentant der Klasse $[a]$. Äquivalenzklassen sind daher stets nicht-leer. Wir wollen zeigen, dass zwei Elemente $a, b \in M$ genau dann dieselbe Äquivalenzklasse induzieren, dass also

$[a] = [b]$ gilt, wenn a, b zueinander äquivalent sind oder, alternativ, wenn sie beide zu einem dritten Element $c \in M$ äquivalent sind. Dies impliziert insbesondere, dass die Äquivalenzrelation auf M eine Unterteilung von M in disjunkte Äquivalenzklassen induziert.

Satz 4. *Es sei “ $\sim$ ” eine Äquivalenzrelation auf einer Menge M . Für Elemente $a, b \in M$ ist dann gleichbedeutend:*

- (i) $a \sim b$
- (ii) $[a] = [b]$
- (iii) $[a] \cap [b] \neq \emptyset$

Insbesondere ist M die disjunkte Vereinigung der zu “ $\sim$ ” gehörigen Äquivalenzklassen.

Beweis. Gelte zunächst (i), also $a \sim b$. Für $c \in M$ mit $c \sim a$ folgt aus $a \sim b$ bereits $c \sim b$ und somit $[a] \subset [b]$. Die umgekehrte Inklusion verifiziert man entsprechend, so dass sich insgesamt die Gleichung $[a] = [b]$ in (ii) ergibt. Da Äquivalenzklassen stets nicht-leer sind, ist weiter (iii) eine Konsequenz aus (ii). Ist schließlich Bedingung (iii) gegeben, so wähle man ein Element $c \in [a] \cap [b]$. Dann gilt $c \sim a$, bzw. $a \sim c$ aufgrund der Symmetrie, sowie $c \sim b$ und damit aufgrund der Transitivität auch $a \sim b$, d. h. (i). $\square$

Definiert man nun $M/\sim$ als Menge der Äquivalenzklassen in M , so kann man die kanonische Abbildung

$$M \longrightarrow M/\sim, \quad a \longmapsto [a],$$

betrachten, welche ein Element $a \in M$ auf die zugehörige Äquivalenzklasse $[a] \subset M$ abbildet. Dabei ist $[a]$ nach Satz 4 die eindeutig bestimmte Äquivalenzklasse in M , die a enthält. Im Folgenden sollen nun Äquivalenzrelationen auf Vektorräumen studiert werden.

Bemerkung 5. *Es sei V ein K -Vektorraum und $U \subset V$ ein linearer Unterraum. Dann wird durch*

$$a \sim b :\Longleftrightarrow a - b \in U$$

eine Äquivalenzrelation auf V erklärt, auch als Kongruenz modulo U bezeichnet. Es gilt $[a] = a + U$ für die Äquivalenzklasse $[a]$ zu einem Element $a \in V$. Zwei Äquivalenzklassen $[a] = a + U$ und $[b] = b + U$ stimmen genau dann überein, wenn $a \sim b$, also $a - b \in U$ gilt.

Beweis. Für alle $a \in V$ gilt $a - a = 0 \in U$ und damit $a \sim a$; die Relation ist also reflexiv. Als Nächstes zeigen wir die Symmetrie. Gelte $a \sim b$, d. h. $a - b \in U$. Dann folgt $b - a = -(a - b) \in U$, was aber $b \sim a$ bedeutet. Schließlich überprüfen wir die Transitivität. Gelte $a \sim b$ und $b \sim c$, also $a - b, b - c \in U$. Dann ergibt sich

$$a - c = (a - b) + (b - c) \in U,$$

also $a \sim c$.

Im Übrigen berechnet sich die Äquivalenzklasse $[a]$ eines Vektors $a \in V$ zu

$$[a] = \{b \in V; b \sim a\} = \{b \in V; b - a \in U\} = a + U,$$

wie behauptet. Schließlich stimmen zwei Äquivalenzklassen $[a]$ und $[b]$ gemäß Bemerkung 5 genau dann überein, wenn $a \sim b$, also $a - b \in U$ gilt. $\square$

Für $a \in V$ bezeichnet man die Äquivalenzklasse $[a] = a + U$ genauer als die *Restklasse von a modulo U* oder auch als die *Nebenklasse von a bezüglich U* . Zur Definition des Quotientenvektorraums V/U haben wir auf der Menge dieser Restklassen eine Struktur als K -Vektorraum zu definieren. Um dies zu bewerkstelligen, leiten wir einige Regeln für das Rechnen mit Restklassen her, wobei wir aus Gründen der Übersichtlichkeit Restklassen $a + U$ noch einmal in der Form von Äquivalenzklassen $[a]$ schreiben.

Bemerkung 6. Auf einem K -Vektorraum V betrachte man die zu einem linearen Unterraum $U \subset V$ gehörige Äquivalenzrelation der Kongruenz modulo U . Es gelte $[a] = [a']$ und $[b] = [b']$ für Vektoren $a, a', b, b' \in V$. Dann folgt $[a + b] = [a' + b']$ sowie $[\alpha a] = [\alpha a']$ für $\alpha \in K$.

Beweis. Aus $[a] = [a']$, $[b] = [b']$ ergibt sich $a - a', b - b' \in U$ und damit $(a + b) - (a' + b') = (a - a') + (b - b') \in U$, also $[a + b] = [a' + b']$. Weiter folgt $\alpha a - \alpha a' = \alpha(a - a') \in U$, also $[\alpha a] = [\alpha a']$. $\square$

Die Bildung der Äquivalenzklassen $[a]$ zu Vektoren $a \in V$ ist also mit der Vektorraumstruktur von V verträglich. Bezeichnen wir mit V/U die Menge der Äquivalenzklassen $[a] = a + U$, wobei a in V variiert, so kann man unter Benutzung von Bemerkung 6 eine Addition

$$V/U \times V/U \longrightarrow V/U, \quad ([a], [b]) \longmapsto [a + b],$$

sowie eine skalare Multiplikation

$$K \times V/U \longrightarrow V/U, \quad (\alpha, [a]) \longmapsto [\alpha a],$$

auf V/U definieren, wobei in beiden Fällen nachzuprüfen ist, dass die Verknüpfung wohldefiniert ist. Genau genommen besagt etwa die Abbildungsvorschrift im Falle der Addition: Man betrachte zwei Äquivalenzklassen $A, B \subset V$, wähle Repräsentanten $a \in A$, $b \in B$, so dass man $A = [a]$, $B = [b]$ schreiben kann, und bilde dann das Paar (A, B) auf die Äquivalenzklasse $[a + b]$ ab. Man muss dabei wissen, dass das Resultat $[a + b]$ nicht von der Wahl der Repräsentanten $a \in A$, $b \in B$ abhängt; Letzteres war aber in Bemerkung 6 gezeigt worden. Entsprechend schließt man im Falle der skalaren Multiplikation.

Man prüft nun leicht unter Benutzung der Vektorraumstruktur von V nach, dass V/U mit den genannten Verknüpfungen ein K -Vektorraum ist und dass die kanonische Abbildung

$$\pi: V \longrightarrow V/U, \quad a \longmapsto [a],$$

K -linear und damit ein Epimorphismus ist. Somit können wir formulieren:

Satz 7. *Es sei V ein K -Vektorraum und $U \subset V$ ein linearer Unterraum. Dann ist*

$$V/U = \{a + U; a \in V\},$$

also die Menge der Restklassen modulo U , ein K -Vektorraum unter der Addition

$$V/U \times V/U \longrightarrow V/U, \quad (a + U, b + U) \longmapsto (a + b) + U,$$

sowie der skalaren Multiplikation

$$K \times V/U \longrightarrow V/U, \quad (\alpha, a + U) \longmapsto \alpha a + U.$$

Die kanonische Abbildung $\pi: V \longrightarrow V/U$ ist ein Epimorphismus, dessen Fasern gerade die Restklassen modulo U in V sind. Insbesondere gilt $\ker \pi = U$ und damit

$$\dim_K V = \dim_K U + \dim_K(V/U)$$

aufgrund der Dimensionsformel 2.1/10.

Man nennt V/U den *Quotienten-* oder *Restklassenvektorraum von V modulo U* . Die Konstruktion dieses Vektorraums zeigt insbesondere, dass es zu einem linearen Unterraum $U \subset V$ stets einen Epimorphismus $p: V \longrightarrow V'$ mit $\ker p = U$ gibt. Umgekehrt hatten wir zu Beginn gesehen, dass jede solche Abbildung benutzt werden kann, um die Menge V/U aller Restklassen modulo U in V mit der Struktur eines K -Vektorraums zu versehen, derart dass man auf natürliche Weise einen Epimorphismus $V \longrightarrow V/U$ mit Kern U erhält. Wir wollen diesen Sachverhalt hier weiter präzisieren, indem wir den Homomorphiesatz beweisen, welcher die so genannte universelle Abbildungseigenschaft für Quotientenvektorräume beinhaltet.

Satz 8 (Homomorphiesatz). *Es sei U ein linearer Unterraum eines K -Vektorraums V und $\pi: V \longrightarrow \bar{V}$ der kanonische Epimorphismus auf den Restklassenvektorraum $\bar{V} = V/U$ oder, allgemeiner, ein Epimorphismus mit $\ker \pi = U$. Dann existiert zu jeder K -linearen Abbildung $f: V \longrightarrow V'$ mit $U \subset \ker f$ genau eine K -lineare Abbildung $\bar{f}: \bar{V} \longrightarrow V'$, so dass das Diagramm*

$$\begin{array}{ccc} V & \xrightarrow{f} & V' \\ & \searrow \pi & \nearrow \bar{f} \\ & \bar{V} & \end{array}$$

kommutiert, d. h. so dass $f = \bar{f} \circ \pi$ gilt. Weiter ist $\bar{f}$ genau dann injektiv, wenn $U = \ker f$ gilt und genau dann surjektiv, wenn f surjektiv ist.

Beweis. Zunächst zeigen wir die Eindeutigkeit von $\bar{f}$. Sei also $\bar{f}: \bar{V} \rightarrow V'$ K -linear mit $f = \bar{f} \circ \pi$. Ist dann $\bar{a} \in \bar{V}$ ein Vektor und $a \in \pi^{-1}(\bar{a})$ ein (beliebiges) Urbild, so folgt $\pi(a) = \bar{a}$ und daher $\bar{f}(\bar{a}) = \bar{f}(\pi(a)) = f(a)$. Damit ist $\bar{f}(\bar{a})$ eindeutig durch $\bar{a}$ bestimmt.

Zum Nachweis der Existenz einer K -linearen Abbildung $\bar{f}$ mit den geforderten Eigenschaften betrachte man einen Vektor $\bar{a} \in \bar{V}$ sowie zwei Urbilder $a, b \in \pi^{-1}(\bar{a})$. Es folgt dann

$$a - b \in \ker \pi = U \subset \ker f$$

und damit $f(a) = f(b)$. Der Wert $f(a)$ ist also unabhängig von der speziellen Wahl eines Urbildes $a \in \pi^{-1}(\bar{a})$, und wir können eine Abbildung $\bar{f}: \bar{V} \rightarrow V'$ durch $\bar{a} \mapsto f(a)$ erklären, wobei wir mit a jeweils (irgend)ein π -Urbild zu $\bar{a}$ meinen. Insbesondere gilt dann $f(a) = \bar{f}(\pi(a))$ für $a \in V$, d. h. die geforderte Beziehung $f = \bar{f} \circ \pi$ ist erfüllt. Schließlich ist $\bar{f}$ K -linear. Zu $\bar{a}, \bar{b} \in \bar{V}$ betrachte man nämlich π -Urbilder $a, b \in V$. Dann ist $a + b$ ein π -Urbild zu $\bar{a} + \bar{b}$ sowie αa für $\alpha \in K$ ein π -Urbild zu $\alpha \bar{a}$, und es gilt

$$\begin{aligned}\bar{f}(\bar{a} + \bar{b}) &= f(a + b) = f(a) + f(b) = \bar{f}(\bar{a}) + \bar{f}(\bar{b}), \\ \bar{f}(\alpha \bar{a}) &= f(\alpha a) = \alpha f(a) = \alpha \bar{f}(\bar{a}).\end{aligned}$$

Es bleiben noch die zusätzlich behaupteten Eigenschaften von $\bar{f}$ nachzuweisen. Sei zunächst $\bar{f}$ injektiv. Für $a \in \ker f$ folgt dann $\bar{f}(\pi(a)) = f(a) = 0$, also $\pi(a) = 0$ bzw. $a \in \ker \pi$ aufgrund der Injektivität von $\bar{f}$. Dies bedeutet aber $\ker f \subset \ker \pi = U$ und, da die entgegengesetzte Inklusion ohnehin gilt, bereits $U = \ker f$. Ist umgekehrt diese Gleichung gegeben, so betrachte man ein Element $\bar{a} \in \ker \bar{f}$ sowie ein Urbild $a \in \pi^{-1}(\bar{a})$. Man hat dann $f(a) = \bar{f}(\pi(a)) = \bar{f}(\bar{a}) = 0$ und damit $a \in \ker f = U$. Wegen $U = \ker \pi$ ergibt sich dann aber $\bar{a} = \pi(a) = 0$, also insgesamt $\ker \bar{f} = 0$ und damit die Injektivität von $\bar{f}$.

Ist $\bar{f}$ surjektiv, so ist auch f als Komposition der surjektiven Abbildungen π und $\bar{f}$ surjektiv. Umgekehrt folgt aus der Surjektivität von $f = \bar{f} \circ \pi$, dass auch $\bar{f}$ surjektiv sein muss. $\square$

Als Folgerung wollen wir noch den so genannten Isomorphiesatz für Vektorräume herleiten.

Korollar 9 (Isomorphiesatz). *Es sei $f: V \rightarrow V'$ eine K -lineare Abbildung zwischen Vektorräumen. Dann induziert f in kanonischer Weise einen Isomorphismus $V/\ker f \xrightarrow{\sim} \text{im } f$, also einen kanonischen Isomorphismus $V/\ker f \xrightarrow{\sim} V'$, falls f surjektiv ist.*

Beweis. Gemäß Satz 8 existiert ein kommutatives Diagramm K -linearer Abbildungen

$$\begin{array}{ccc}
 V & \xrightarrow{f} & V' \\
 \searrow \pi & & \nearrow \bar{f} \\
 & V/\ker f &
 \end{array}$$

wobei $\bar{f}$ wegen $\ker \pi = \ker f$ injektiv ist. Es definiert also $\bar{f}$ einen Isomorphismus von $V/\ker f$ auf den linearen Unterraum $\text{im } \bar{f} \subset V'$, und letzterer stimmt überein mit $\text{im } f$. $\square$

Wir wollen schließlich noch die Ergebnisse dieses Abschnitts zur Charakterisierung affiner Unterräume von Vektorräumen verwenden. Dabei wollen wir zunächst bemerken, dass jeder nicht-leere affine Unterraum eines K -Vektorraums V in eindeutiger Weise einen zugehörigen linearen Unterraum bestimmt.

Bemerkung 10. *Es sei V ein K -Vektorraum und $A \subset V$ ein nicht-leerer affiner Unterraum, d. h. eine Teilmenge der Form $a + U$ mit einem Vektor $a \in V$ und einem linearen Unterraum $U \subset V$. Dann ist U eindeutig durch A bestimmt.*

Beweis. Gilt etwa $a + U = A = a' + U'$ für Elemente $a, a' \in V$ und lineare Unterräume $U, U' \subset V$, so folgt

$$U = (a' - a) + U'.$$

Wegen $0 \in U$ bedeutet dies, dass $a' - a$ ein inverses Element in U' besitzt und daher selbst zu U' gehört. Dies aber impliziert $U = U'$, wie behauptet. $\square$

Insbesondere kann man daher nicht-leeren affinen Unterräumen eines Vektorraums V eine Dimension zuordnen, nämlich die Dimension des zugehörigen linearen Unterrums von V . Beispielsweise nennt man einen affinen Unterraum von V der Dimension 1 eine *Gerade*, der Dimension 2 eine *Ebene* oder der Dimension $n - 1$ eine *Hyperebene*, wenn $n = \dim_K V < \infty$.

Satz 11. *Es sei V ein K -Vektorraum und $A \subset V$ eine Teilmenge. Dann ist äquivalent:*

- (i) *A ist ein affiner Unterraum von V , d. h. A ist leer, oder es existieren ein Vektor $a \in V$ sowie ein linearer Unterraum $U \subset V$ mit $A = a + U$.*
- (ii) *Es existiert eine K -lineare Abbildung $f: V \rightarrow V'$, so dass A eine Faser von f ist, d. h. es existiert ein $a' \in V'$ mit $A = f^{-1}(a')$.*
- (iii) *Für jeweils endlich viele Elemente $a_0, \dots, a_r \in A$ sowie Koeffizienten $\alpha_0, \dots, \alpha_r \in K$ mit $\sum_{i=0}^r \alpha_i = 1$ folgt $\sum_{i=0}^r \alpha_i a_i \in A$.*

Beweis. Der Fall $A = \emptyset$ ist trivial, da sich die leere Menge stets als Faser einer linearen Abbildung $f: V \rightarrow V'$ realisieren lässt. Man betrachte etwa $V' := V \times K$ als K -Vektorraum mit komponentenweiser Addition und skalarer Multiplikation. Dann ist

$$f: V \longrightarrow V', \quad v \longmapsto (v, 0),$$

eine K -lineare Abbildung mit $f^{-1}(0, 1) = \emptyset$.

Wir dürfen also im Folgenden $A \neq \emptyset$ voraussetzen. Sei zunächst Bedingung (i) gegeben, also A ein affiner Unterraum von V , etwa $A = a + U$. Betrachtet man dann den kanonischen Epimorphismus $\pi: V \longrightarrow V/U$, so gilt $A = \pi^{-1}(\pi(a))$ nach Bemerkung 2, d. h. A ist Faser einer K -linearen Abbildung und erfüllt damit Bedingung (ii).

Sei nun (ii) gegeben, also $A = f^{-1}(a')$ mit einem Element $a' \in V'$, und seien endlich viele Elemente $a_0, \dots, a_r \in A$ fixiert. Für Koeffizienten $\alpha_0, \dots, \alpha_r \in K$ mit $\sum_{i=0}^r \alpha_i = 1$ gilt dann

$$f\left(\sum_{i=0}^r \alpha_i a_i\right) = \sum_{i=0}^r \alpha_i f(a_i) = \left(\sum_{i=0}^r \alpha_i\right) \cdot a' = a'$$

und damit $\sum_{i=0}^r \alpha_i a_i \in f^{-1}(a') = A$. Bedingung (iii) ist also erfüllt.

Sei schließlich (iii) gegeben. Wir wählen dann ein Element $a_0 \in A$ aus und behaupten, dass $\Delta A = \{a - a_0; a \in A\}$ ein linearer Unterraum von V ist. In der Tat, wir haben $0 \in \Delta A$ und damit $\Delta A \neq \emptyset$. Sind weiter $a, b \in \Delta A$, also $a = a_1 - a_0, b = b_1 - a_0$ mit $a_1, b_1 \in A$, so folgt $a_1 + (-1)a_0 + b_1 \in A$ gemäß (iii) und damit

$$a + b = (a_1 - a_0 + b_1) - a_0 \in \Delta A.$$

Für $\alpha \in K$ gilt weiter $\alpha a_1 + (1 - \alpha)a_0 \in A$, ebenfalls unter Benutzung von (iii), also

$$\alpha a = (\alpha a_1 + (1 - \alpha)a_0) - a_0 \in \Delta A.$$

Folglich ist ΔA ein linearer Unterraum in V , und wir sehen, dass $A = a_0 + \Delta A$ ein affiner Unterraum von V ist. $\square$

Wir wollen noch zeigen, wie man affine Unterräume von Vektorräumen in konkreter Weise erzeugen kann. Hierzu betrachte man einen K -Vektorraum V sowie Vektoren $a_0, \dots, a_r \in V, r \geq 0$. Dann existiert ein kleinster affiner Unterraum $A \subset V$, der diese Vektoren enthält, nämlich

$$A = \left\{a_0 + \sum_{i=1}^r \alpha_i (a_i - a_0); \alpha_1, \dots, \alpha_r \in K\right\}.$$

Um dies einzusehen, stellen wir zunächst fest, dass A ein affiner Unterraum in V ist, der die Vektoren $a_0, \dots, a_r$ enthält; der zugehörige lineare Unterraum $U \subset V$ mit $A = a_0 + U$ wird von den Vektoren $a_i - a_0, i = 1, \dots, r$, erzeugt. Da man eine Summe $a_0 + \sum_{i=1}^r \alpha_i (a_i - a_0)$ auch in der Form $(1 - \sum_{i=1}^r \alpha_i) \cdot a_0 + \sum_{i=1}^r \alpha_i a_i$ schreiben kann, erhält man weiter

$$A = \left\{\sum_{i=0}^r \alpha_i a_i; \alpha_0, \dots, \alpha_r \in K \text{ mit } \sum_{i=0}^r \alpha_i = 1\right\}.$$

Hieraus folgt unter Benutzung von Satz 11 (iii), dass A wie behauptet der kleinste affine Unterraum von V ist, der $a_0, \dots, a_r$ enthält. Beispielsweise kann man die durch zwei verschiedene Punkte a, b eines K -Vektorraums bestimmte affine Gerade betrachten. In gewohnter Weise ergibt sich

$$G = \{a + \alpha(b - a); \alpha \in K\} = \{\alpha a + \beta b; \alpha, \beta \in K \text{ mit } \alpha + \beta = 1\}.$$

Aufgaben

- Man betrachte folgende Mengen M mit der jeweils angegebenen Relation $\sim$ und entscheide, ob es sich um eine Äquivalenzrelation handelt. Falls möglich gebe man die Äquivalenzklassen an.
 - $M = \mathbb{R}, a \sim b : \Longleftrightarrow |a| = |b|$
 - $M = \mathbb{R}, a \sim b : \Longleftrightarrow |a - b| < 1$
 - $M = \mathbb{Z}, a \sim b : \Longleftrightarrow a - b$ wird von p geteilt (d. h. es existiert eine Zerlegung $a - b = c \cdot p$ mit einem $c \in \mathbb{Z}$), wobei p eine fest vorgegebene ganze Zahl ist.
- Es seien V, V' Vektorräume über einem Körper K und $A \subset V$ sowie $A' \subset V'$ affine Unterräume. Man zeige:
 - Für eine lineare Abbildung $f: V \longrightarrow V'$ ist $f(A)$ ein affiner Unterraum von V' und $f^{-1}(A')$ ein affiner Unterraum von V .
 - Für $V = V'$ sind $A + A' = \{a + a'; a \in A, a' \in A'\}$ und $A \cap A'$ affine Unterräume von V .
 - Es ist $A \times A'$ ein affiner Unterraum von $V \times V'$, wobei man $V \times V'$ mit komponentenweiser Addition und skalarer Multiplikation als K -Vektorraum auffasse.
- Es seien A, A' affine Unterräume eines K -Vektorraums V . Man zeige, $A \cup A'$ ist genau dann ein affiner Unterraum von V , wenn $A \subset A'$ oder $A' \subset A$ gilt.
- Es sei V ein K -Vektorraum und $A \subset V$ eine Teilmenge. Man zeige, dass es eine kleinste Teilmenge $A' \subset V$ mit $A \subset A'$ gibt, so dass die durch

$$a \sim b : \Longleftrightarrow a - b \in A'$$

definierte Relation eine Äquivalenzrelation auf V definiert. Man bestimme A' unter der Voraussetzung, dass A abgeschlossen unter der skalaren Multiplikation ist, dass also aus $\alpha \in K, a \in A$ stets $\alpha a \in A$ folgt.

- Für eine Matrix $(\lambda_{ij})_{\substack{i=1, \dots, m \\ j=1, \dots, n}}$ mit Koeffizienten aus einem Körper K und für Elemente $\delta_1, \dots, \delta_m \in K$ zeige man, dass durch

$$A = \{(\alpha_1, \dots, \alpha_n) \in K^n; \sum_{j=1}^n \lambda_{ij} \alpha_j = \delta_i, i = 1, \dots, m\}$$

ein affiner Unterraum im K^n erklärt wird. Ist jeder affine Unterraum des K^n von dieser Bauart? (Hinweis: Man vermeide "unnötiges Rechnen", indem man geeignete lineare Abbildungen betrachtet. Man beginne mit dem Fall $m = 1$.)

6. Für Vektoren $a_0, \dots, a_r$ eines K -Vektorraums V betrachte man den von den a_i erzeugten affinen Unterraum $A \subset V$ sowie den von den a_i erzeugten linearen Unterraum $U \subset V$. Man zeige $A \subset U$ und weiter:

$$\dim_K A = \begin{cases} \dim_K U & \text{falls } 0 \in A \\ \dim_K U - 1 & \text{falls } 0 \notin A \end{cases}$$

7. Eine Abbildung $f: V \longrightarrow V'$ zwischen K -Vektorräumen heie *affin*, wenn es ein Element $x_0 \in V'$ gibt, so dass die Abbildung

$$V \longrightarrow V', \quad a \longmapsto f(a) - x_0,$$

K -linear ist. Man zeige: Eine Abbildung $f: V \longrightarrow V'$ ist genau dann affin, wenn für jeweils endlich viele Vektoren $a_0, \dots, a_r \in V$ und Elemente $\alpha_0, \dots, \alpha_r \in K$ mit $\sum_{i=0}^r \alpha_i = 1$ stets

$$f\left(\sum_{i=0}^r \alpha_i a_i\right) = \sum_{i=0}^r \alpha_i f(a_i)$$

gilt.

8. (1. *Isomorphiesatz*) Es seien U, U' lineare Unterräume eines K -Vektorraums V . Man zeige: Die kanonische Abbildung $U \hookrightarrow U + U' \longrightarrow (U + U')/U'$ besitzt $U \cap U'$ als Kern und induziert einen Isomorphismus

$$U/(U \cap U') \xrightarrow{\sim} (U + U')/U'.$$

9. (2. *Isomorphiesatz*) Es seien V ein K -Vektorraum und $U \subset U' \subset V$ lineare Unterräume. Man zeige:

- (i) Die kanonische Abbildung $U' \hookrightarrow V \longrightarrow V/U$ besitzt U als Kern und induziert einen Monomorphismus $U'/U \hookrightarrow V/U$. Folglich lässt sich U'/U mit seinem Bild in V/U identifizieren und somit als linearer Unterraum von V/U auffassen.
- (ii) Die Projektion $V \longrightarrow V/U'$ faktorisiert über V/U , d. h. lässt sich als Komposition $V \xrightarrow{\pi} V/U \xrightarrow{f} V/U'$ schreiben, mit einer linearen Abbildung f und der kanonischen Projektion π .
- (iii) f besitzt U'/U als Kern und induziert einen Isomorphismus

$$(V/U)/(U'/U) \xrightarrow{\sim} V/U'.$$

10. Für lineare Unterräume U_1, U_2 eines K -Vektorraums V betrachte man die kanonischen Projektionen $\pi_i: V \longrightarrow V/U_i$, $i = 1, 2$, sowie die Abbildung

$$(\pi_1, \pi_2): V \longrightarrow V/U_1 \times V/U_2, \quad x \longmapsto (\pi_1(x), \pi_2(x)).$$

Man zeige:

- (i) (π_1, π_2) ist K -linear, wenn man $V/U_1 \times V/U_2$ mit komponentenweiser Addition und skalarer Multiplikation als K -Vektorraum auffasst.
- (ii) (π_1, π_2) ist genau dann injektiv, wenn $U_1 \cap U_2 = 0$ gilt.
- (iii) (π_1, π_2) ist genau dann surjektiv, wenn $V = U_1 + U_2$ gilt.
- (iv) (π_1, π_2) ist genau dann bijektiv, wenn $V = U_1 \oplus U_2$ gilt.

11. Man konstruiere ein Beispiel eines K -Vektorraums V mit einem linearen Unterraum $0 \subsetneq U \subset V$, derart dass es einen Isomorphismus $V/U \xrightarrow{\sim} V$ gibt. Gibt es ein solches Beispiel im Falle $\dim_K V < \infty$?
12. Es seien $V_1 \xrightarrow{f} V_2 \xrightarrow{g} V_3 \xrightarrow{h} V_4$ K -lineare Abbildungen zwischen endlich-dimensionalen K -Vektorräumen. Man zeige:

$$\operatorname{rg}(g \circ f) + \operatorname{rg}(h \circ g) \leq \operatorname{rg}(h \circ g \circ f) + \operatorname{rg} g$$

2.3 Der Dualraum

Für zwei K -lineare Abbildungen $f, g: V \longrightarrow V'$ zwischen K -Vektorräumen hatten wir in Abschnitt 2.1 die Summe durch

$$f + g: V \longrightarrow V', \quad a \longmapsto f(a) + g(a),$$

und für eine Konstante $\alpha \in K$ das Produkt αf durch

$$\alpha f: V \longrightarrow V', \quad a \longmapsto \alpha(f(a)),$$

definiert. Man rechnet ohne Schwierigkeiten nach, dass $f + g$ und αf wieder K -linear sind und dass die K -linearen Abbildungen $V \longrightarrow V'$ mit diesen Verknüpfungen einen K -Vektorraum bilden; dieser wird mit $\operatorname{Hom}_K(V, V')$ bezeichnet. Speziell werden wir in diesem Abschnitt für V' als K -Vektorraum den Körper K selbst betrachten. K -lineare Abbildungen $V \longrightarrow K$ werden auch als *Linearformen* bezeichnet.

Definition 1. *Es sei V ein K -Vektorraum. Dann heißt der K -Vektorraum $\operatorname{Hom}_K(V, K)$ aller Linearformen auf V der Dualraum zu V . Dieser wird mit V^* bezeichnet.*

Beispielsweise ist für $n \in \mathbb{N} - \{0\}$ die Abbildung

$$K^n \longrightarrow K, \quad (\alpha_1, \dots, \alpha_n) \longmapsto \alpha_1,$$

eine Linearform auf K^n . Betrachten wir allgemeiner einen K -Vektorraum V mit Basis $a_1, \dots, a_n$, so können wir aus 2.1/7 ablesen, dass es zu gegebenen Elementen $\lambda_1, \dots, \lambda_n \in K$ stets eine eindeutig bestimmte Linearform $\varphi \in V^*$ mit $\varphi(a_i) = \lambda_i$, $i = 1, \dots, n$, gibt. Die Linearformen auf V korrespondieren folglich in bijektiver Weise zu den n -Tupeln von Konstanten aus K , wobei diese Zuordnung genauer einen Isomorphismus $V^* \xrightarrow{\sim} K^n$ darstellt. Legen wir etwa im Falle $V = K^n$ die kanonische Basis $e_1, \dots, e_n$ zugrunde, so werden dementsprechend die Linearformen auf K^n gerade durch die Abbildungen

$$K^n \longrightarrow K, \quad (\alpha_1, \dots, \alpha_n) \longmapsto \sum_{i=1}^n \lambda_i \alpha_i,$$

zu n -Tupeln $(\lambda_1, \dots, \lambda_n) \in K^n$ gegeben. Insbesondere ist der Dualraum des Nullraums wieder der Nullraum.

Satz 2. *Es sei $f: U \longrightarrow V$ eine K -lineare Abbildung zwischen Vektorräumen. Dann ist die Abbildung*

$$f^*: V^* \longrightarrow U^*, \quad \varphi \longmapsto \varphi \circ f,$$

K -linear; man nennt f^ die duale Abbildung zu f . Ist $g: V \longrightarrow W$ eine weitere K -lineare Abbildung zwischen Vektorräumen, so gilt $(g \circ f)^* = f^* \circ g^*$.*

Man kann sich die Definition des Bildes $f^*(\varphi) = \varphi \circ f$ anhand des folgenden kommutativen Diagramms verdeutlichen:

$$\begin{array}{ccc} U & \xrightarrow{f} & V \\ & \searrow f^*(\varphi) & \downarrow \varphi \\ & & K \end{array}$$

Es wird die Linearform φ sozusagen unter der Abbildung f^* zu einer Linearform auf U zurückgezogen, indem man mit f komponiert.

Nun zum *Beweis von Satz 2*. Für $\varphi, \psi \in V^*$ sowie für $\alpha \in K$ ergibt sich in nahe liegender Weise

$$\begin{aligned} f^*(\varphi + \psi) &= (\varphi + \psi) \circ f = (\varphi \circ f) + (\psi \circ f) = f^*(\varphi) + f^*(\psi), \\ f^*(\alpha\varphi) &= (\alpha\varphi) \circ f = \alpha(\varphi \circ f) = \alpha f^*(\varphi) \end{aligned}$$

und damit die K -Linearität von f^* . Um die Formel für die Komposition dualer Abbildungen nachzuweisen, betrachte man für eine Linearform $\varphi \in W^*$ das kommutative Diagramm:

$$\begin{array}{ccccc} U & \xrightarrow{f} & V & \xrightarrow{g} & W \\ & \searrow (g \circ f)^*(\varphi) & \searrow g^*(\varphi) & & \downarrow \varphi \\ & & & & K \end{array}$$

Es ergibt sich

$$\begin{aligned} (g \circ f)^*(\varphi) &= \varphi \circ (g \circ f) = (\varphi \circ g) \circ f \\ &= (g^*(\varphi)) \circ f = f^*(g^*(\varphi)) = (f^* \circ g^*)(\varphi) \end{aligned}$$

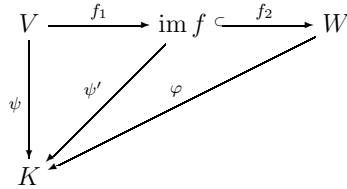
und damit die gewünschte Relation $(g \circ f)^* = f^* \circ g^*$, indem man φ in W^* variieren lässt. $\square$

Im Folgenden wollen wir einige Zusammenhänge zwischen linearen Abbildungen und den zugehörigen dualen Abbildungen untersuchen. Wir schließen dabei auch den Fall unendlich-dimensionaler Vektorräume mit ein, beschränken uns bei den Beweisen aus technischen Gründen aber meist auf den Fall endlich-dimensionaler Vektorräume.

Satz 3. *Es sei $f: V \rightarrow W$ eine K -lineare Abbildung zwischen Vektorräumen sowie $f^*: W^* \rightarrow V^*$ die zugehörige duale Abbildung.*

- (i) *Ist f ein Monomorphismus, so ist f^* ein Epimorphismus.*
- (ii) *Ist f ein Epimorphismus, so ist f^* ein Monomorphismus.*

Beweis. Wir beginnen mit Aussage (i). Sei also f injektiv. Um zu sehen, dass f^* surjektiv ist, zeigen wir, dass jede Linearform $\psi \in V^*$ ein Urbild $\varphi \in W^*$ besitzt, dass also zu ψ eine Linearform $\varphi \in W^*$ mit $\psi = \varphi \circ f$ existiert. Hierzu schreiben wir f als Komposition zweier K -linearer Abbildungen, nämlich der von f induzierten Abbildung $f_1: V \rightarrow \text{im } f$, sowie der Inklusionsabbildung $f_2: \text{im } f \hookrightarrow W$. Dabei ist f_1 aufgrund der Injektivität von f ein Isomorphismus. Sodann betrachte man das Diagramm



wobei die Linearformen $\psi': \text{im } f \rightarrow K$ und $\varphi: W \rightarrow K$ noch zu konstruieren sind, und zwar so, dass das Diagramm kommutiert. Da f_1 ein Isomorphismus ist, können wir $\psi' = \psi \circ f_1^{-1}$ setzen. Dann bleibt noch das Problem, die Linearform $\psi': \text{im } f \rightarrow K$ zu einer Linearform $\varphi: W \rightarrow K$ fortzusetzen.

Um dies zu bewerkstelligen, wählen wir gemäß 1.6/4 ein Komplement W' zu $\text{im } f$ in W , also einen Untervektorraum $W' \subset W$ mit $W = (\text{im } f) \oplus W'$. Jedes $w \in W$ hat dann eine eindeutige Zerlegung $w = w_1 \oplus w_2$ mit $w_1 \in \text{im } f$ und $w_2 \in W'$. Setzt man jeweils $\varphi(w) = \psi'(w_1)$, so erhält man wie gewünscht eine Linearform $\varphi: W \rightarrow K$, die ψ' fortsetzt.

Zum Nachweis von (ii) setze man $f: V \rightarrow W$ als surjektiv voraus. Weiter seien $\varphi, \varphi' \in W^*$ mit $f^*(\varphi) = f^*(\varphi')$, also mit $\varphi \circ f = \varphi' \circ f$. Ist dann $w \in W$ beliebig gewählt und $v \in V$ ein Urbild unter f , so gilt

$$\varphi(w) = \varphi(f(v)) = (\varphi \circ f)(v) = (\varphi' \circ f)(v) = \varphi'(f(v)) = \varphi'(w).$$

Hieraus folgt $\varphi = \varphi'$, wenn man w in W variieren lässt. □

Die Aussage von Satz 3 lässt sich mittels so genannter exakter Sequenzen allgemeiner formulieren. Unter einer *Sequenz* wollen wir eine Kette

$$\dots \xrightarrow{f_{n-2}} V_{n-1} \xrightarrow{f_{n-1}} V_n \xrightarrow{f_n} V_{n+1} \xrightarrow{f_{n+1}} \dots$$

von K -linearen Abbildungen verstehen, wobei der Index n einen endlichen oder unendlichen Abschnitt in $\mathbb{Z}$ durchlaufe. Man sagt, die Sequenz besitze bei V_n die Eigenschaft eines *Komplexes*, wenn $f_n \circ f_{n-1} = 0$ oder, äquivalent hierzu, im $f_{n-1} \subset \ker f_n$ gilt. Weiter heißt die Sequenz *exakt bei V_n* , wenn sogar im $f_{n-1} = \ker f_n$ gilt. Erfüllt die Sequenz die Komplex-Eigenschaft an allen Stellen V_n (natürlich außer bei solchen V_n , die die Sequenz möglicherweise terminieren), so spricht man von einem Komplex. Entsprechend heißt die Sequenz *exakt*, wenn sie an allen Stellen exakt ist. Beispielsweise ist für eine K -lineare Abbildung $f: V' \rightarrow V$ die Sequenz

$$0 \rightarrow V' \xrightarrow{f} V$$

genau dann exakt, wenn f injektiv ist; dabei ist 0 der Nullvektorraum sowie $0 \rightarrow V'$ (als einzig mögliche K -lineare Abbildung von 0 nach V') die Nullabbildung. Entsprechend ist die Sequenz

$$V' \xrightarrow{f} V \rightarrow 0$$

genau dann exakt, wenn f surjektiv ist; auch hierbei ist $V \rightarrow 0$ (zwangsläufig) die Nullabbildung. Exakte Sequenzen des Typs

$$0 \rightarrow V' \xrightarrow{f} V \xrightarrow{g} V'' \rightarrow 0$$

werden auch als *kurze exakte Sequenzen* bezeichnet. Die Exaktheit einer solchen Sequenz ist äquivalent zu den folgenden Bedingungen:

- (1) f ist injektiv,
- (2) $\operatorname{im} f = \ker g$,
- (3) g ist surjektiv.

Bei einer kurzen exakten Sequenz können wir daher V' unter f als Untervektorraum von V auffassen, und es folgt mit dem Isomorphiesatz 2.2/9, dass g einen kanonischen Isomorphismus $V/V' \xrightarrow{\sim} V''$ induziert. Umgekehrt kann man zu jedem Untervektorraum $U \subset V$ in kanonischer Weise die exakte Sequenz

$$0 \rightarrow U \rightarrow V \rightarrow V/U \rightarrow 0$$

betrachten.

Satz 4. *Es sei*

$$\dots \rightarrow V_{n-1} \xrightarrow{f_{n-1}} V_n \xrightarrow{f_n} V_{n+1} \rightarrow \dots$$

eine exakte Sequenz von K -Vektorräumen. Dann ist auch die zugehörige duale Sequenz

$$\dots \leftarrow V_{n-1}^* \xleftarrow{f_{n-1}^*} V_n^* \xleftarrow{f_n^*} V_{n+1}^* \leftarrow \dots$$

exakt.

Formulieren wir Satz 3 im Sinne exakter Sequenzen um, so sieht man, dass für eine exakte Sequenz $0 \rightarrow U \rightarrow V$ auch die zugehörige duale Sequenz $V^* \rightarrow U^* \rightarrow 0$ exakt ist; vgl. Satz 3 (i). Entsprechend besagt Satz 3 (ii), dass für eine exakte Sequenz $U \rightarrow V \rightarrow 0$ die zugehörige duale Sequenz $0 \rightarrow V^* \rightarrow U^*$ exakt ist. Satz 4 ist daher eine Verallgemeinerung von Satz 3.

Beweis zu Satz 4. Es genügt zu zeigen, dass für eine exakte Sequenz

$$U \xrightarrow{f} V \xrightarrow{g} W$$

die zugehörige duale Sequenz

$$W^* \xrightarrow{g^*} V^* \xrightarrow{f^*} U^*$$

exakt ist. Zunächst folgt aus $g \circ f = 0$ für beliebiges $\varphi \in W^*$ die Beziehung $(f^* \circ g^*)(\varphi) = \varphi \circ g \circ f = 0$, d. h. es gilt $f^* \circ g^* = 0$ und damit $\text{im } g^* \subset \ker f^*$. Um auch die umgekehrte Inklusion zeigen zu können, betrachte man ein Element $\varphi \in \ker f^*$, also eine Linearform $\varphi: V \rightarrow K$ mit $\varphi \circ f = 0$. Dann haben wir $\varphi(\text{im } f) = 0$ und wegen der Exaktheit von $U \rightarrow V \rightarrow W$ auch $\varphi(\ker g) = 0$; Letzteres bedeutet $\ker g \subset \ker \varphi$. Um nun $\varphi \in \text{im } g^*$ zu zeigen, haben wir eine Linearform $\psi \in W^*$ mit $\varphi = \psi \circ g$ zu konstruieren. Hierzu zerlegen wir $g: V \rightarrow W$ ähnlich wie im Beweis zu Satz 3 in die von g induzierte surjektive Abbildung $g_1: V \rightarrow \text{im } g$ sowie die Inklusion $g_2: \text{im } g \hookrightarrow W$. Aufgrund des Homomorphiesatzes 2.2/8 existiert dann eine Linearform $\varphi': \text{im } g \rightarrow K$ mit $\varphi = \varphi' \circ g_1$. Weiter lässt sich φ' , wie im Beweis zu Satz 3 gezeigt oder unter Benutzung von Satz 3 (i), zu einer Linearform $\psi: W \rightarrow K$ ausdehnen, d. h. mit $\psi \circ g_2 = \varphi'$. Dann gilt

$$g^*(\psi) = \psi \circ g = \psi \circ g_2 \circ g_1 = \varphi' \circ g_1 = \varphi,$$

also ist ψ ein Urbild zu φ unter g^* . Hieraus ergibt sich $\ker f^* \subset \text{im } g^*$ bzw. $\ker f^* = \text{im } g^*$ und damit die Exaktheit der Sequenz $W^* \rightarrow V^* \rightarrow U^*$. $\square$

Als Nächstes wollen wir für endlich-dimensionale K -Vektorräume und ihre Dualräume Basen zueinander in Beziehung setzen. Insbesondere wollen wir zeigen, dass ein Vektorraum und sein Dualraum dieselbe Dimension besitzen. Wir erinnern dabei nochmals an das Kronecker-Symbol δ_{ij} , welches für Indizes i, j aus einer vorgegebenen Menge erklärt ist. Und zwar gilt $\delta_{ij} = 1$ für $i = j$ und $\delta_{ij} = 0$ sonst.

Satz 5. *Es sei V ein K -Vektorraum mit Basis $a_1, \dots, a_n$. Erklärt man dann Linearformen $\varphi_1, \dots, \varphi_n \in V^*$ gemäß 2.1/7 durch $\varphi_i(a_j) = \delta_{ij}$, so bilden diese eine Basis von V^* , die so genannte duale Basis zu $a_1, \dots, a_n$.*

Beweis. Wir überprüfen zunächst, dass $\varphi_1, \dots, \varphi_n$ linear unabhängig sind. Hierzu betrachte man eine Gleichung $\sum_{i=1}^n \alpha_i \varphi_i = 0$ mit Koeffizienten $\alpha_i \in K$. Einsetzen von a_j liefert

$$\alpha_j = \sum_{i=1}^n \alpha_i \varphi_i(a_j) = 0, \quad j = 1, \dots, n,$$

d. h. $\varphi_1, \dots, \varphi_n$ sind linear unabhängig. Ist andererseits $\varphi \in V^*$ gegeben, so gilt

$$\varphi(a_j) = \sum_{i=1}^n \varphi(a_i) \varphi_i(a_j), \quad j = 1, \dots, n,$$

was $\varphi = \sum_{i=1}^n \varphi(a_i) \varphi_i$ gemäß 2.1/7 impliziert. $\varphi_1, \dots, \varphi_n$ erzeugen also V^* und bilden damit sogar eine Basis von V^* . $\square$

Korollar 6. *Es sei V ein endlich-dimensionaler K -Vektorraum. Dann gilt $\dim_K V = \dim_K V^*$. Insbesondere existiert ein Isomorphismus $V \xrightarrow{\sim} V^*$.*

Man beachte, dass sich die Aussage von Satz 5 nicht auf unendlich-dimensionale K -Vektorräume verallgemeinern lässt. Ist nämlich $(a_i)_{i \in I}$ eine unendliche Basis, so erkläre man wiederum durch $\varphi_i(a_j) = \delta_{ij}$ ein System $(\varphi_i)_{i \in I}$ von Linearformen auf V . Wie im Beweis zu Satz 5 ergibt sich, dass dieses linear unabhängig ist und dass der von den φ_i erzeugte lineare Unterraum von V^* gerade aus denjenigen Linearformen $\varphi: V \rightarrow K$ besteht, für die $\varphi(a_i) = 0$ für fast alle $i \in I$ gilt. Andererseits kann man aber Linearformen $V \rightarrow K$ definieren, indem man die Bilder der Basisvektoren a_i beliebig vorgibt. Insbesondere gibt es Linearformen $\varphi \in V^*$ mit $\varphi(a_i) \neq 0$ für unendlich viele Indizes $i \in I$. Somit kann $(\varphi_i)_{i \in I}$ im Falle einer unendlichen Indexmenge I kein Erzeugendensystem und damit auch keine Basis von V^* sein.

Als Nächstes wollen wir für lineare Abbildungen $f: V \rightarrow W$ den Rang $\operatorname{rg} f$, also die Dimension des Bildes von f , mit dem Rang der zugehörigen dualen Abbildung $f^*: W^* \rightarrow V^*$ vergleichen.

Satz 7. *Ist $f: V \rightarrow W$ eine K -lineare Abbildung endlich-dimensionaler Vektorräume und $f^*: W^* \rightarrow V^*$ die zugehörige duale Abbildung, so gilt $\operatorname{rg} f = \operatorname{rg} f^*$.*

Beweis. Wir spalten $f: V \rightarrow W$ auf in die induzierte surjektive lineare Abbildung $f_1: V \rightarrow \operatorname{im} f$, sowie die Inklusion $f_2: \operatorname{im} f \hookrightarrow W$. Nach Satz 2 ist dann $f^*: W^* \rightarrow V^*$ die Komposition von $f_2^*: W^* \rightarrow (\operatorname{im} f)^*$ und $f_1^*: (\operatorname{im} f)^* \rightarrow V^*$. Nun ist f_2^* surjektiv nach Satz 3 (i) und f_1^* injektiv nach Satz 3 (ii). Es wird daher $(\operatorname{im} f)^*$ unter f_1^* isomorph auf das Bild $\operatorname{im} f^*$ abgebildet, und man erhält mit Korollar 6

$$\operatorname{rg} f = \dim_K(\operatorname{im} f) = \dim_K(\operatorname{im} f)^* = \dim_K(\operatorname{im} f^*) = \operatorname{rg} f^*.$$

$\square$

Abschließend wollen wir noch den doppelt dualen Vektorraum $V^{**} = (V^*)^*$ zu einem K -Vektorraum V betrachten. Man hat stets eine kanonische Abbildung

$$\Phi: V \longrightarrow V^{**}, \quad a \longmapsto a^*,$$

wobei a^* für $a \in V$ gegeben ist durch

$$a^*: V^* \longrightarrow K, \quad \varphi \longmapsto \varphi(a).$$

Satz 8. *Ist V ein endlich-dimensionaler K -Vektorraum, so ist die vorstehend beschriebene Abbildung $\Phi: V \longrightarrow V^{**}$ ein Isomorphismus von K -Vektorräumen.*

Beweis. Wir zeigen zunächst, dass Φ K -linear ist. Seien $a, b \in V$ und $\alpha \in K$. Dann stimmen die Abbildungen

$$\begin{aligned} (a+b)^*: V^* &\longrightarrow K, & \varphi &\longmapsto \varphi(a+b) = \varphi(a) + \varphi(b), \\ a^* + b^*: V^* &\longrightarrow K, & \varphi &\longmapsto \varphi(a) + \varphi(b), \end{aligned}$$

überein, d. h. es gilt $\Phi(a+b) = \Phi(a) + \Phi(b)$. Entsprechend hat man

$$\begin{aligned} (\alpha a)^*: V^* &\longrightarrow K, & \varphi &\longmapsto \varphi(\alpha a) = \alpha \varphi(a), \\ \alpha a^*: V^* &\longrightarrow K, & \varphi &\longmapsto \alpha \varphi(a), \end{aligned}$$

d. h. es folgt $\Phi(\alpha a) = \alpha \Phi(a)$. Insgesamt ergibt sich die K -Linearität von Φ .

Bilden nun $a_1, \dots, a_n$ eine Basis von V und ist $\varphi_1, \dots, \varphi_n$ die hierzu duale Basis, so ist $a_1^*, \dots, a_n^*$ die duale Basis zu $\varphi_1, \dots, \varphi_n$; denn es gilt

$$a_i^*(\varphi_j) = \varphi_j(a_i) = \delta_{ij} \quad \text{für } i, j = 1, \dots, n.$$

Unter Φ wird also eine Basis von V auf eine Basis des Bildraumes V^{**} abgebildet. Dies bedeutet dann, dass Φ notwendig ein Isomorphismus ist. $\square$

Aufgaben

- Man prüfe, ob die folgenden Linearformen $\varphi_i: \mathbb{R}^5 \longrightarrow \mathbb{R}$ ein linear unabhängiges System in $(\mathbb{R}^5)^*$ bilden:
 - $$\begin{aligned} \varphi_1: (\alpha_1, \dots, \alpha_5) &\longmapsto \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 \\ \varphi_2: (\alpha_1, \dots, \alpha_5) &\longmapsto \alpha_1 + 2\alpha_2 + 3\alpha_3 + 4\alpha_4 + 5\alpha_5 \\ \varphi_3: (\alpha_1, \dots, \alpha_5) &\longmapsto \alpha_1 - \alpha_2 \end{aligned}$$
 - $$\begin{aligned} \varphi_1: (\alpha_1, \dots, \alpha_5) &\longmapsto \alpha_1 + 7\alpha_2 + 7\alpha_3 + 7\alpha_4 + \alpha_5 \\ \varphi_2: (\alpha_1, \dots, \alpha_5) &\longmapsto \alpha_1 + 2\alpha_2 + 3\alpha_3 + 4\alpha_4 + 5\alpha_5 \\ \varphi_3: (\alpha_1, \dots, \alpha_5) &\longmapsto 12\alpha_1 + 7\alpha_2 + 13\alpha_3 + 8\alpha_4 + 9\alpha_5 \\ \varphi_4: (\alpha_1, \dots, \alpha_5) &\longmapsto 10\alpha_1 - 2\alpha_2 + 3\alpha_3 - 3\alpha_4 + 3\alpha_5 \end{aligned}$$
- Für einen Körper K und ein $n \in \mathbb{N}$ betrachte man K^n als K -Vektorraum. Es sei $p_i: K^n \longrightarrow K$, $i = 1, \dots, n$, jeweils die Projektion auf die i -te Koordinate. Man zeige, dass $p_1, \dots, p_n$ eine Basis des Dualraums $(K^n)^*$ bilden, und zwar die duale Basis zur kanonischen Basis $e_1, \dots, e_n \in K^n$, bestehend aus den Einheitsvektoren $e_i = (\delta_{1i}, \dots, \delta_{ni})$.

Im Falle $K = \mathbb{R}$ und $n = 4$ fixiere man

$$(1, 0, 0, 0), \quad (1, 1, 0, 0), \quad (1, 1, 1, 0), \quad (1, 1, 1, 1)$$

als Basis von $\mathbb{R}^4$ und bestimme die hierzu duale Basis von $(\mathbb{R}^4)^*$, indem man deren Elemente als Linearkombinationen der p_i angibt.

3. Es seien V, W K -Vektorräume. Man zeige, dass die Abbildung

$$\text{Hom}_K(V, W) \longrightarrow \text{Hom}_K(W^*, V^*), \quad f \longmapsto f^*,$$

K -linear ist, was bedeutet, dass für Elemente $\alpha \in K$ und K -lineare Abbildungen $f, g: V \longrightarrow W$ stets $(f + g)^* = f^* + g^*$ sowie $(\alpha f)^* = \alpha \cdot f^*$ gilt.

4. Man betrachte einen K -Vektorraum V und zu einem Element $a \in V$ das induzierte Element $a^* \in V^{**}$, welches definiert ist durch

$$a^*: V^* \longrightarrow K, \quad \varphi \longmapsto \varphi(a).$$

Man zeige:

- (i) Die Abbildung $\iota: V \longrightarrow \text{Hom}_K(K, V)$, welche einem Element $a \in V$ die Abbildung $\alpha \longmapsto \alpha \cdot a$ zuordnet, ist ein Isomorphismus von K -Vektorräumen.
- (ii) Man identifiziere V unter ι mit $\text{Hom}_K(K, V)$ und entsprechend K als K -Vektorraum mit seinem Dualraum $\text{Hom}_K(K, K)$. Dann ist für $a \in V$ das Element $a^* \in V^{**}$ zu interpretieren als die duale Abbildung zu $a = \iota(a)$.
5. Es sei V ein endlich-dimensionaler K -Vektorraum und V^* sein Dualraum. Für lineare Unterräume $U_1 \subset V$ und $U_2 \subset V^*$ setze man

$$U_1^\perp = \{\varphi \in V^* ; \varphi(U_1) = 0\},$$

$$U_2^\perp = \{a \in V ; \varphi(a) = 0 \text{ für alle } \varphi \in U_2\}.$$

Man zeige für lineare Unterräume $U, U' \subset V$ bzw. $U, U' \subset V^*$:

- (i) $(U + U')^\perp = U^\perp \cap U'^\perp$
- (ii) $(U^\perp)^\perp = U$
- (iii) $(U \cap U')^\perp = U^\perp + U'^\perp$
- (iv) $\dim_K U + \dim_K U^\perp = \dim_K V$
6. Es sei $f: V \longrightarrow V'$ eine K -lineare Abbildung zwischen K -Vektorräumen und $f^*: V'^* \longrightarrow V^*$ die zugehörige duale Abbildung. Man konstruiere auf kanonische Weise einen Isomorphismus

$$\text{coker } f^* \xrightarrow{\sim} (\ker f)^*,$$

wobei man mit $\text{coker } f^*$ den *Kokern* von f^* , d. h. den Quotienten $V^*/\text{im } f^*$ bezeichnet.

7. Es sei $0 \longrightarrow V' \xrightarrow{f} V \xrightarrow{g} V'' \longrightarrow 0$ eine kurze exakte Sequenz von K -Vektorräumen. Man zeige, dass eine solche Sequenz stets *spaltet*, d. h. dass es eine K -lineare Abbildung $\tilde{g}: V'' \longrightarrow V$ mit $g \circ \tilde{g} = \text{id}_{V''}$ gibt. Die Wahl einer solchen Abbildung $\tilde{g}$ hat folgende Konsequenzen:

- (i) $\tilde{g}$ ist injektiv, und es gilt $V = \ker g \oplus \text{im } \tilde{g}$ bzw. $V = V' \oplus V''$, wenn wir V' unter f mit $\ker g = \text{im } f$ identifizieren und entsprechend V'' unter $\tilde{g}$ mit $\text{im } \tilde{g}$.

- (ii) Es existiert eine eindeutig bestimmte K -lineare Abbildung $\tilde{f}: V \longrightarrow V'$ mit $\tilde{f} \circ f = \text{id}_{V'}$ und $\tilde{f} \circ \tilde{g} = 0$.
- (iii) Es ist $0 \longleftarrow V' \xleftarrow{\tilde{f}} V \xleftarrow{\tilde{g}} V'' \longleftarrow 0$ eine kurze exakte Sequenz.
8. Man betrachte eine exakte Sequenz von endlich-dimensionalen K -Vektorräumen

$$0 \longrightarrow V_1 \xrightarrow{f_1} \dots \xrightarrow{f_{n-1}} V_n \longrightarrow 0$$

und zeige

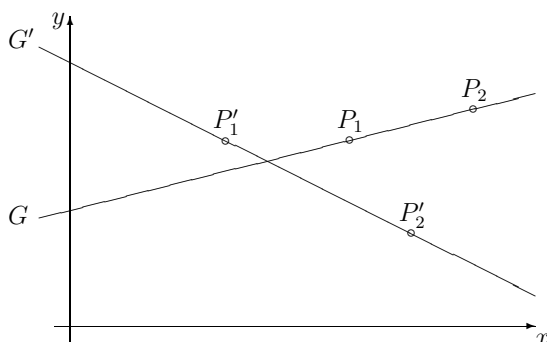
$$\sum_{i=1}^n (-1)^i \dim_K V_i = 0.$$

3. Matrizen

Vorbemerkungen

Bisher haben wir uns im Wesentlichen mit der Untersuchung von grundsätzlichen Eigenschaften bei Vektorräumen und linearen Abbildungen sowie der hiermit verbundenen Konstruktionen beschäftigt. Wir wollen nun verstärkt auf den rechnerischen Standpunkt eingehen und in diesem Kapitel zeigen, wie man konkret gegebene Probleme der Linearen Algebra in effektiver Weise rechnerisch lösen kann. Im Zentrum stehen hier *Matrizen* und das so genannte *Gaußsche Eliminationsverfahren*, insbesondere in der Version zur Lösung linearer Gleichungssysteme.

Als Beispiel gehen wir von der folgenden geometrischen Situation aus:



Gegeben seien zwei Geraden $G, G' \subset \mathbb{R}^2$, wobei G durch die (verschiedenen) Punkte $P_1, P_2 \in \mathbb{R}^2$ festgelegt sei und entsprechend G' durch die (verschiedenen) Punkte $P'_1, P'_2 \in \mathbb{R}^2$. Man untersuche dann in Abhängigkeit von P_1, P_2, P'_1, P'_2 , ob G und G' Schnittpunkte, also gemeinsame Punkte, besitzen und, wenn ja, wie sich diese aus den Punkten P_1, P_2, P'_1, P'_2 berechnen lassen. Wie wir bereits in den Vorbemerkungen zu Kapitel 1 gesehen haben, gelten für G und G' die Parametrisierungen

$$G = \{P_1 + t(P_2 - P_1) ; t \in \mathbb{R}\}, \quad G' = \{P'_1 + t'(P'_2 - P'_1) ; t' \in \mathbb{R}\},$$

und ein Schnittpunkt $S \in G \cap G'$ ist offenbar dadurch charakterisiert, dass es Parameter $t, t' \in \mathbb{R}$ mit

$$P_1 + t(P_2 - P_1) = S = P'_1 + t'(P'_2 - P'_1)$$

gibt. Mit anderen Worten, zur Ermittlung der Schnittpunkte von G und G' haben wir alle Paare $(t, t') \in \mathbb{R}^2$ zu bestimmen, die der Gleichung

$$(*) \quad t(P_2 - P_1) + t'(P'_1 - P'_2) = P'_1 - P_1$$

genügen, oder, in Komponentenschreibweise mit

$$P_2 - P_1 = (\alpha_{11}, \alpha_{21}), \quad P'_1 - P'_2 = (\alpha_{12}, \alpha_{22}), \quad P'_1 - P_1 = (\beta_1, \beta_2),$$

die den Gleichungen

$$(**) \quad \begin{aligned} t\alpha_{11} + t'\alpha_{12} &= \beta_1 \\ t\alpha_{21} + t'\alpha_{22} &= \beta_2 \end{aligned}$$

genügen. Es ist also ein Gleichungssystem bestehend aus 2 Gleichungen mit den beiden Unbekannten t und t' und den gegebenen Konstanten $\alpha_{ij}, \beta_i \in \mathbb{R}$ zu lösen. Dabei handelt es sich um ein *lineares* Gleichungssystem, womit wir meinen, dass die Unbekannten t, t' separiert in höchstens erster Potenz auftreten.

Um die Gleichung $(*)$ zu lösen, nehmen wir zunächst einmal an, dass die beiden Vektoren $P_2 - P_1$ und $P'_1 - P'_2$ *linear abhängig* sind, was in diesem Fall zur Folge hat, dass jeder der beiden Vektoren ein skalares Vielfaches des anderen ist; geometrisch bedeutet dies, dass G und G' parallel sind. Es ist also $P'_1 - P'_2$ ein skalares Vielfaches von $P_2 - P_1$, etwa $P'_1 - P'_2 = \gamma(P_2 - P_1)$, und wir können die Gleichung $(*)$ zu

$$P'_1 = P_1 + (t + \gamma t')(P_2 - P_1)$$

umschreiben. Somit ist die Existenz von Lösungen t, t' äquivalent zu $P'_1 \in G$, was aber dann bereits $G = G'$ impliziert. Es können daher lediglich folgende Fälle auftreten, die sich gegenseitig ausschließen:

1. $P_2 - P_1$ und $P'_1 - P'_2$ sind linear abhängig, $G = G'$.
2. $P_2 - P_1$ und $P'_1 - P'_2$ sind linear abhängig, $G \neq G'$; dann gilt $G \cap G' = \emptyset$.
3. $P_2 - P_1$ und $P'_1 - P'_2$ sind linear unabhängig; dann besitzen G und G' genau einen Schnittpunkt (was anschaulich klar ist), den wir sogleich berechnen wollen.

Wir nehmen nun also gemäß 3. an, dass die Vektoren $P_2 - P_1 = (\alpha_{11}, \alpha_{21})$ und $P'_1 - P'_2 = (\alpha_{12}, \alpha_{22})$ *linear unabhängig* sind. Insbesondere gilt $P_2 - P_1 \neq 0$, und es sei etwa $\alpha_{11} \neq 0$. Dann können wir das System $(**)$ in äquivalenter Weise umformen, indem wir das $\frac{\alpha_{21}}{\alpha_{11}}$ -fache der ersten Gleichung von der zweiten subtrahieren, so dass wir in der zweiten Gleichung die Unbekannte t "eliminieren":

$$\begin{aligned} t\alpha_{11} + t'\alpha_{12} &= \beta_1 \\ t'(\alpha_{22} - \frac{\alpha_{21}}{\alpha_{11}}\alpha_{12}) &= \beta_2 - \frac{\alpha_{21}}{\alpha_{11}}\beta_1 \end{aligned}$$

Nun gilt aber $\alpha_{22} - \frac{\alpha_{21}}{\alpha_{11}}\alpha_{12} \neq 0$, denn anderenfalls würde sich

$$\alpha_{12} = \frac{\alpha_{12}}{\alpha_{11}}\alpha_{11}, \quad \alpha_{22} = \frac{\alpha_{12}}{\alpha_{11}}\alpha_{21}$$

ergeben, und $P'_1 - P'_2$ wäre linear abhängig von $P_2 - P_1$, was aber unserer Annahme widerspricht. Somit können wir die zweite Gleichung weiter äquivalent umformen zu

$$t' = \frac{\beta_2 - \frac{\alpha_{21}}{\alpha_{11}}\beta_1}{\alpha_{22} - \frac{\alpha_{21}}{\alpha_{11}}\alpha_{12}} = \frac{\alpha_{11}\beta_2 - \alpha_{21}\beta_1}{\alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12}},$$

und wir erhalten aus der ersten Gleichung

$$t = \frac{1}{\alpha_{11}} \left(\beta_1 - \alpha_{12} \frac{\alpha_{11}\beta_2 - \alpha_{21}\beta_1}{\alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12}} \right) = \frac{\beta_1\alpha_{22} - \beta_2\alpha_{12}}{\alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12}}.$$

Dabei haben wir unter der Annahme von $\alpha_{11} \neq 0$ die Gleichungen in äquivalenter Weise umgeformt, so dass t, t' auch tatsächlich Lösungen von $(**)$ sind. Es bleibt nun aber noch der Fall $\alpha_{11} = 0$ zu betrachten. In diesem Fall ist allerdings $\alpha_{21} \neq 0$, und wir können eine entsprechende Rechnung unter dieser Annahme machen. Das Ergebnis für t, t' lässt sich formal aus dem vorstehenden ableiten, indem wir in den Vektoren $(\alpha_{11}, \alpha_{21}), (\alpha_{12}, \alpha_{22}), (\beta_1, \beta_2)$ jeweils die Reihenfolge der Komponenten vertauschen, was einer Vertauschung der beiden Gleichungen des Systems $(**)$ entspricht. Da sich aber das vorstehende Ergebnis bei dieser Vertauschung nicht ändert, wie man leicht feststellt, erhalten wir

$$t = \frac{\beta_1\alpha_{22} - \beta_2\alpha_{12}}{\alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12}}, \quad t' = \frac{\alpha_{11}\beta_2 - \alpha_{21}\beta_1}{\alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12}}$$

als Lösung der Gleichung $(*)$ bzw. des linearen Gleichungssystems $(**)$ im Falle nicht-paralleler Geraden G, G' . Insbesondere besteht $G \cap G'$ aus genau einem Schnittpunkt, und dieser berechnet sich zu

$$S = P_1 + \frac{\beta_1\alpha_{22} - \beta_2\alpha_{12}}{\alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12}}(P_2 - P_1) = P'_1 + \frac{\alpha_{11}\beta_2 - \alpha_{21}\beta_1}{\alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12}}(P'_2 - P'_1),$$

jeweils als Punkt von G bzw. G' .

In ähnlicher Weise führt das Problem, den Schnitt zweier Ebenen im $\mathbb{R}^3$, oder allgemeiner, den Schnitt zweier affiner Unterräume im K^n zu bestimmen, auf ein lineares Gleichungssystem. Im ersten Falle handelt es sich beispielsweise um ein System von 3 Gleichungen mit 4 Unbekannten, von denen im Allgemeinfall eine frei wählbar ist, so dass der Schnitt dann als Gerade parametrisiert wird. Im zweiten Fall hat man ein System von n Gleichungen mit $r + s$ Unbekannten zu betrachten, wobei r und s die Dimensionen der betrachteten affinen Unterräume im K^n sind.

Aber auch direkte Fragen über Vektoren in Vektorräumen können auf lineare Gleichungssysteme führen. Über einem Körper K betrachte man beispielsweise Vektoren $a_i = (\alpha_{1i}, \dots, \alpha_{ni}) \in K^n$, $i = 1, \dots, r$, und teste, ob diese linear abhängig sind. Wir müssen dann also feststellen, ob es Elemente $t_1, \dots, t_r \in K$ gibt, die nicht alle verschwinden, so dass

$$t_1 a_1 + \dots + t_r a_r = 0$$

gilt, oder, nach Übergang zu den einzelnen Komponenten, ob das lineare Gleichungssystem

$$\begin{aligned} t_1 \alpha_{11} + \dots + t_r \alpha_{1r} &= 0 \\ t_1 \alpha_{21} + \dots + t_r \alpha_{2r} &= 0 \\ &\dots \\ t_1 \alpha_{n1} + \dots + t_r \alpha_{nr} &= 0 \end{aligned}$$

eine nicht-triviale Lösung $t_1, \dots, t_r \in K$ besitzt. Alternativ können wir eine lineare Abbildung $f: K^r \longrightarrow K^n$ betrachten, welche die kanonische Basis bestehend aus den Einheitsvektoren $e_1, \dots, e_r \in K^r$ auf Vektoren $a_1, \dots, a_r \in K^n$ abbildet, und danach fragen, ob ein gegebener Vektor $b = (\beta_1, \dots, \beta_n)$ zum Bild von f gehört, also zu dem Unterraum, der von den Vektoren $a_i = f(e_i)$, $i = 1, \dots, r$ erzeugt wird. Zu testen ist daher, ob die Gleichung

$$t_1 a_1 + \dots + t_r a_r = b,$$

bzw. für $a_i = (\alpha_{1i}, \dots, \alpha_{ni})$, $i = 1, \dots, r$, ob das lineare Gleichungssystem

$$\begin{aligned} t_1 \alpha_{11} + \dots + t_r \alpha_{1r} &= \beta_1 \\ t_1 \alpha_{21} + \dots + t_r \alpha_{2r} &= \beta_2 \\ &\dots \\ t_1 \alpha_{n1} + \dots + t_r \alpha_{nr} &= \beta_n \end{aligned}$$

eine Lösung zulässt.

Die Anordnung der Koeffizienten in obigem Gleichungssystem legt es nahe, die Vektoren $a_1, \dots, a_r, b$ nicht mehr als *Zeilen*-Vektoren zu schreiben, sondern als *Spalten*-Vektoren, also

$$a_1 = \begin{pmatrix} \alpha_{11} \\ \vdots \\ \alpha_{n1} \end{pmatrix}, \quad \dots, \quad a_r = \begin{pmatrix} \alpha_{1r} \\ \vdots \\ \alpha_{nr} \end{pmatrix}, \quad b = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}.$$

Im Übrigen werden wir das rechteckige Koeffizientenschema

$$A = (a_1, \dots, a_r) = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1r} \\ \alpha_{21} & \dots & \alpha_{2r} \\ \cdot & \dots & \cdot \\ \alpha_{n1} & \dots & \alpha_{nr} \end{pmatrix}$$

betrachten und dies als *Koeffizientenmatrix* des Gleichungssystems bezeichnen. Wir entschließen uns auch, die Unbekannten $t_1, \dots, t_r$ zu einem Spalten-Vektor zusammenzufassen, und können dann das obige lineare Gleichungssystem, ähnlich wie im Falle einer Gleichung mit einer Unbekannten, in der übersichtlichen Form

$$\begin{pmatrix} \alpha_{11} & \dots & \alpha_{1r} \\ \alpha_{21} & \dots & \alpha_{2r} \\ \vdots & \dots & \vdots \\ \alpha_{n1} & \dots & \alpha_{nr} \end{pmatrix} \cdot \begin{pmatrix} t_1 \\ \vdots \\ t_r \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}$$

schreiben, wobei das Produkt der Koeffizientenmatrix mit dem Vektor der Unbekannten durch

$$\begin{pmatrix} \alpha_{11} & \dots & \alpha_{1r} \\ \alpha_{21} & \dots & \alpha_{2r} \\ \vdots & \dots & \vdots \\ \alpha_{n1} & \dots & \alpha_{nr} \end{pmatrix} \cdot \begin{pmatrix} t_1 \\ \vdots \\ t_r \end{pmatrix} := \begin{pmatrix} \sum_{j=1}^r \alpha_{1j} t_j \\ \vdots \\ \sum_{j=1}^r \alpha_{nj} t_j \end{pmatrix}$$

definiert wird. Nach wie vor lautet das Gleichungssystem dann

$$\sum_{j=1}^r \alpha_{ij} t_j = \beta_i, \quad i = 1, \dots, n,$$

wobei wir, wie hier geschehen, die Koeffizienten α_{ij} in Zukunft immer links von den Unbekannten t_j schreiben werden.

Die Bedeutung von Matrizen als rechteckige Koeffizientenschemata ist nicht auf lineare Gleichungssysteme beschränkt. Matrizen treten in der Linearen Algebra überall dort auf, wo es darum geht, Skalare zu katalogisieren, die von zwei Index-Parametern abhängen. So betrachte man etwa eine lineare Abbildung $f: V \rightarrow W$ zwischen K -Vektorräumen V, W mit Basen $X = (x_1, \dots, x_r)$ und $Y = (y_1, \dots, y_n)$. Es ist f dann vollständig charakterisiert durch die Bilder $f(x_1), \dots, f(x_r)$ der Basis X . Jedes dieser Bilder $f(x_j)$ wiederum ist vollständig durch die Angabe seiner Koordinaten bezüglich Y charakterisiert, d. h. der Koeffizienten $\alpha_{1j}, \dots, \alpha_{nj} \in K$, die man benötigt, um $f(x_j)$ als Linearkombination der Basis Y darzustellen, etwa

$$f(x_j) = \sum_{i=1}^n \alpha_{ij} y_i, \quad j = 1, \dots, r.$$

Somit ist f eindeutig charakterisiert durch die Angabe der Matrix

$$A_{f,X,Y} = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1r} \\ \alpha_{21} & \dots & \alpha_{2r} \\ \vdots & \dots & \vdots \\ \alpha_{n1} & \dots & \alpha_{nr} \end{pmatrix},$$

und man nennt dies die Matrix, die f bezüglich der Basen X von V und Y von W beschreibt. Die Korrespondenz $f \mapsto A_{f,X,Y}$ besitzt sehr gute Eigenschaften, wie wir noch sehen werden. Beispielsweise lässt sich der Übergang von einem Vektor $x \in V$ zu seinem Bild $f(x) \in W$ auf dem Niveau der Koordinaten bezüglich X bzw. Y als Multiplikation von $A_{f,X,Y}$ mit dem so genannten Koordinatenspaltenvektor von x bezüglich X interpretieren. Weiter entspricht die

Komposition linearer Abbildungen bei Verwendung geeigneter Basen dem noch zu definierenden Produkt der beschreibenden Matrizen.

Im Spezialfall $V = W$ und $f = \text{id}$ stellt die Matrix $A_{\text{id}, X, Y}$ gerade diejenigen Koeffizienten bereit, die man benötigt, um die Elemente der Basis X als Linearkombinationen der Basis Y darzustellen. Insbesondere sehen wir, dass der Wechsel zwischen verschiedenen Basen eines Vektorraums ebenfalls mittels Matrizen beschrieben wird. Ein weiteres Anwendungsfeld für Matrizen stellen die in Kapitel 7 zu behandelnden Skalarprodukte dar.

Wir studieren in diesem Kapitel zunächst die Korrespondenz zwischen linearen Abbildungen f und den zugehörigen beschreibenden Matrizen $A_{f, X, Y}$ genauer. Sodann besprechen wir das Gaußsche Eliminationsverfahren, und zwar zuerst als Verfahren, welches mittels so genannter elementarer Zeilenumformungen den Zeilenrang einer Matrix bestimmt, d. h. die Dimension des von den Zeilen der Matrix erzeugten Vektorraums. In analoger Weise lässt sich der Spaltenrang einer Matrix betrachten. Es ist leicht einzusehen, dass dieser im Falle einer Matrix $A_{f, X, Y}$, die zu einer linearen Abbildung f gehört, mit dem Rang von f , also der Dimension des Bildes von f , übereinstimmt. Dass der Spaltenrang einer Matrix stets auch mit deren Zeilenrang übereinstimmt, ist hingegen ein nicht-triviales Resultat.

Als Nächstes behandeln wir die Invertierbarkeit von (quadratischen) Matrizen. Eine quadratische Matrix A heißt *invertierbar*, wenn es eine quadratische Matrix B mit $A \cdot B = B \cdot A = E$ gibt, wobei E die so genannte *Einheitsmatrix* bezeichnet, deren Diagonaleinträge alle 1 sind und die ansonsten nur aus Nullen besteht. Wir werden insbesondere sehen, dass die Eigenschaft einer linearen Abbildung, ein Isomorphismus zu sein, auf der Seite der Matrizen der Invertierbarkeit entspricht. Im Übrigen verwenden wir das Gaußsche Eliminationsverfahren, um konkret gegebene Matrizen auf Invertierbarkeit zu überprüfen und, falls möglich, zu invertieren. Es folgt die Interpretation invertierbarer Matrizen im Rahmen von Basiswechselmatrizen.

Schließlich kommen wir dann noch ausführlich auf lineare Gleichungssysteme zu sprechen. Wir ordnen diese in den Zusammenhang linearer Abbildungen ein und zeigen insbesondere, wie man die Lösung konkreter Gleichungssysteme mittels des Gaußschen Eliminationsverfahrens formalisieren kann.

3.1 Lineare Abbildungen und Matrizen

Wir hatten bereits in Abschnitt 2.1 gesehen, dass man mit Hilfe von Matrizen lineare Abbildungen definieren kann und dass man umgekehrt lineare Abbildungen bezüglich gewählter Basen durch Matrizen beschreiben kann. Dieser Sachverhalt soll in diesem Abschnitt genauer studiert werden. Wie immer sei K ein Körper.

Definition 1. *Es seien $m, n \in \mathbb{N}$. Eine $(m \times n)$ -Matrix mit Koeffizienten aus K ist ein System von Elementen $\alpha_{ij} \in K$, welches durch die Bedingung*

$(i, j) \in \{1, \dots, m\} \times \{1, \dots, n\}$ parametrisiert wird, also

$$A = (\alpha_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \cdot & \dots & \cdot \\ \alpha_{m1} & \dots & \alpha_{mn} \end{pmatrix}.$$

Dabei heißt i der Zeilenindex und j der Spaltenindex von A ; man sagt, A bestehe aus m Zeilen und n Spalten. Für $m = 0$ oder $n = 0$ ist A eine so genannte leere Matrix¹.

Wir bezeichnen die Menge aller $(m \times n)$ -Matrizen mit $K^{m \times n}$. Dabei lässt sich $K^{m \times n}$ unter der Zuordnung

$$\begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \cdot & \dots & \cdot \\ \alpha_{m1} & \dots & \alpha_{mn} \end{pmatrix} \mapsto (\alpha_{11}, \dots, \alpha_{1n}, \alpha_{21}, \dots, \alpha_{2n}, \dots, \alpha_{m1}, \dots, \alpha_{mn})$$

mit $K^{m \cdot n}$ identifizieren. Insbesondere können wir $K^{m \times n}$, ebenso wie $K^{m \cdot n}$, als K -Vektorraum auffassen. Damit ist für Matrizen

$$A = (\alpha_{ij})_{i,j}, \quad B = (\beta_{ij})_{i,j} \in K^{m \times n}$$

deren Summe gegeben durch

$$A + B = (\alpha_{ij} + \beta_{ij})_{i,j},$$

sowie das Produkt von A mit einem Skalar $\lambda \in K$ durch

$$\lambda A = (\lambda \alpha_{ij})_{i,j}.$$

Das Nullelement $0 \in K^{m \times n}$ ist die so genannte *Nullmatrix*, deren Koeffizienten α_{ij} sämtlich verschwinden. Insbesondere ergibt sich als negatives Element zu einer Matrix $A = (\alpha_{ij})_{i,j} \in K^{m \times n}$ die Matrix $-A = (-\alpha_{ij})_{i,j}$. Die kanonische Basis von $K^{m \times n}$ wird gegeben durch die Matrizen

$$E_{ij} = (\delta_{i\mu} \delta_{j\nu})_{\substack{\mu=1,\dots,m, \\ \nu=1,\dots,n}}, \quad i = 1, \dots, m, \quad j = 1, \dots, n,$$

wobei E_{ij} genau am Schnittpunkt der i -ten Zeile mit der j -ten Spalte eine 1 stehen hat und ansonsten aus lauter Nullen besteht. Beispielsweise gilt für eine Matrix $(\alpha_{ij})_{i,j} \in K^{m \times n}$

$$A = \sum_{\substack{i=1,\dots,m \\ j=1,\dots,n}} \alpha_{ij} E_{ij}.$$

¹ Zur Vermeidung von Sonderfällen im Zusammenhang mit linearen Abbildungen ist es praktisch, leere Matrizen nicht von den Betrachtungen auszuschließen. Wir werden insbesondere für $m = 0$ leere Matrizen unterschiedlicher Spaltenanzahlen n betrachten, sowie im Falle $n = 0$ leere Matrizen unterschiedlicher Zeilenanzahlen m .

Im Zusammenhang mit Matrizen ist es üblich, Vektoren $a \in K^m$ nicht wie bisher als *Zeilen*-Vektoren $a = (\alpha_1, \dots, \alpha_m)$ zu schreiben, sondern als *Spalten*-Vektoren:

$$a = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix}$$

Insbesondere kann man aus n solchen Vektoren $a_1, \dots, a_n \in K^m$ eine Matrix $A = (a_1, \dots, a_n) \in K^{m \times n}$ aufbauen. Ist W ein endlich-dimensionaler K -Vektorraum mit Basis $Y = (y_1, \dots, y_m)$, so besitzt jedes $a \in W$ eine Darstellung $a = \sum_{i=1}^m \alpha_i y_i$ mit eindeutig bestimmten Koeffizienten $\alpha_i \in K$. Wir bezeichnen

$$a_Y = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix} \in K^m$$

als den zu $a \in W$ gehörigen *Koordinatenspaltenvektor* bezüglich der Basis Y von W .

Bemerkung 2. Es sei W ein K -Vektorraum mit Basis $Y = (y_1, \dots, y_m)$. Dann ist die Abbildung

$$\kappa_Y: W \longrightarrow K^m, \quad a \longmapsto a_Y,$$

ein Isomorphismus von K -Vektorräumen.

Beweis. Man prüft sofort nach, dass κ_Y linear ist und die Basis Y von V auf die kanonische Basis von K^m abbildet. Somit ist κ_Y notwendigerweise ein Isomorphismus. $\square$

Unter Verwendung von Koordinatenspaltenvektoren können wir die bereits in Abschnitt 2.1 angedeutete Zuordnung einer Matrix zu einer linearen Abbildung in einfacher Weise charakterisieren:

Definition 3. Es seien V, W K -Vektorräume mit Basen $X = (x_1, \dots, x_n)$ von V sowie $Y = (y_1, \dots, y_m)$ von W . Ist dann $f: V \longrightarrow W$ eine K -lineare Abbildung, so heißt

$$A = A_{f,X,Y} = (f(x_1)_Y, \dots, f(x_n)_Y) \in K^{m \times n}$$

die zu f gehörige Matrix bezüglich der Basen X und Y .

Die zu f gehörige Matrix $A_{f,X,Y}$ besteht also gerade aus den Koordinatenspaltenvektoren bezüglich Y , die sich aus den Bildern $f(x_1), \dots, f(x_n)$ der Basis $x_1, \dots, x_n$ von V ergeben. Um $A_{f,X,Y}$ in expliziter Weise aufzuschreiben, stelle man also die Bilder von $x_1, \dots, x_n$ mit Hilfe der Basis Y dar, etwa

$$f(x_j) = \sum_{i=1}^m \alpha_{ij} y_i, \quad j = 1, \dots, n,$$

und es folgt dann:

$$A_{f,X,Y} = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn} \end{pmatrix}$$

Beispielsweise ist die zur Nullabbildung $0: V \longrightarrow W$ gehörige Matrix $A_{0,X,Y}$ gerade die Nullmatrix $0 \in K^{m \times n}$. Betrachtet man andererseits für $V = W$ und $X = Y$ die identische Abbildung $\text{id}: V \longrightarrow V$, so ist die zugehörige Matrix $A_{\text{id},X,X}$ die so genannte $(m \times m)$ -*Einheitsmatrix*:

$$E := E_m := (\delta_{ij})_{i,j=1,\dots,m} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \in K^{m \times m}$$

Wir wollen nun die Zuordnung $f \longmapsto A_{f,X,Y}$ genauer untersuchen und erinnern daran, dass die Menge aller K -linearen Abbildungen eines K -Vektorraums V in einen K -Vektorraum W wiederum einen K -Vektorraum bildet, der mit $\text{Hom}_K(V, W)$ bezeichnet wird. Die Summe zweier Elemente $f, g \in \text{Hom}_K(V, W)$ ist definiert durch $(f+g)(a) = f(a) + g(a)$ und das skalare Vielfache von f mit einem Element $\alpha \in K$ durch $(\alpha f)(a) = \alpha f(a)$, jeweils für $a \in V$.

Satz 4. *Es seien V, W endlich-dimensionale K -Vektorräume mit gegebenen Basen $X = (x_1, \dots, x_n)$ und $Y = (y_1, \dots, y_m)$. Dann ist die Abbildung*

$$\Psi: \text{Hom}_K(V, W) \longrightarrow K^{m \times n}, \quad f \longmapsto A_{f,X,Y},$$

ein Isomorphismus von K -Vektorräumen.

Beweis. Wir prüfen zunächst nach, dass Ψ K -linear ist. Dabei verwenden wir den Isomorphismus $W \longrightarrow K^m$, $a \longmapsto a_Y$, und benutzen insbesondere, dass diese Abbildung K -linear ist. Für $f, g \in \text{Hom}_K(V, W)$, $\alpha \in K$ gilt

$$\begin{aligned} A_{f+g,X,Y} &= ((f+g)(x_1)_Y, \dots, (f+g)(x_n)_Y) \\ &= ((f(x_1) + g(x_1))_Y, \dots, (f(x_n) + g(x_n))_Y) \\ &= (f(x_1)_Y + g(x_1)_Y, \dots, f(x_n)_Y + g(x_n)_Y) \\ &= (f(x_1)_Y, \dots, f(x_n)_Y) + (g(x_1)_Y, \dots, g(x_n)_Y) \\ &= A_{f,X,Y} + A_{g,X,Y}, \end{aligned}$$

sowie

$$\begin{aligned} A_{(\alpha f),X,Y} &= ((\alpha f)(x_1)_Y, \dots, (\alpha f)(x_n)_Y) \\ &= ((\alpha(f(x_1)))_Y, \dots, (\alpha(f(x_n)))_Y) \\ &= (\alpha(f(x_1))_Y, \dots, \alpha(f(x_n))_Y) \\ &= \alpha(f(x_1)_Y, \dots, f(x_n)_Y) \\ &= \alpha \cdot A_{f,X,Y}, \end{aligned}$$

also $\Psi(f + g) = \Psi(f) + \Psi(g)$, $\Psi(\alpha f) = \alpha\Psi(f)$, d. h. Ψ ist K -linear. Weiter ist Ψ injektiv aufgrund von 2.1/7 (i) und surjektiv aufgrund von 2.1/7 (ii). $\square$

Korollar 5. Für endlich-dimensionale K -Vektorräume V, W gilt

$$\dim_K \operatorname{Hom}_K(V, W) = \dim_K V \cdot \dim_K W.$$

Als Nächstes wollen wir das Produkt von Matrizen definieren und zeigen, dass dieses der Komposition linearer Abbildungen entspricht. Für natürliche Zahlen $\ell, m, n \in \mathbb{N}$ kann man Matrizen $A \in K^{\ell \times m}$ mit Matrizen $B \in K^{m \times n}$ multiplizieren und erhält dabei Matrizen aus $K^{\ell \times n}$, und zwar ist für

$$A = (\alpha_{ij})_{\substack{i=1,\dots,\ell \\ j=1,\dots,m}} \in K^{\ell \times m}, \quad B = (\beta_{jk})_{\substack{j=1,\dots,m \\ k=1,\dots,n}} \in K^{m \times n}$$

das Produkt $A \cdot B$ erklärt durch:

$$A \cdot B = \left(\sum_{j=1}^m \alpha_{ij} \cdot \beta_{jk} \right)_{\substack{i=1,\dots,\ell \\ k=1,\dots,n}}$$

Das Produkt kann also nur dann gebildet werden, wenn die Spaltenanzahl von A gleich der Zeilenanzahl von B ist. Ein einfacher Fall eines Matrizenprodukts liegt vor, wenn man eine Zeile aus $K^{1 \times m}$ mit einer Spalte aus $K^{m \times 1}$ multipliziert:

$$(\alpha_1, \dots, \alpha_m) \cdot \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} = (\alpha_1\beta_1 + \dots + \alpha_m\beta_m)$$

Es entsteht eine (1×1) -Matrix, wobei wir diese unter Fortlassen der Klammern mit dem entsprechenden Element von K identifizieren wollen. So kann man sagen, dass für Matrizen $A \in K^{\ell \times m}$, $B \in K^{m \times n}$ das Produkt $A \cdot B = (\gamma_{ik})_{i,k}$ aus allen Produkten von Zeilen von A mit Spalten von B besteht, und zwar ist das Element γ_{ik} in der i -ten Zeile und k -ten Spalte von $A \cdot B$ gerade das Produkt der i -ten Zeile von A mit der k -ten Spalte von B . Andererseits beachte man, dass das Produkt einer Spalte aus $K^{\ell \times 1}$ mit einer Zeile aus $K^{1 \times n}$ eine Matrix aus $K^{\ell \times n}$ ergibt:

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_\ell \end{pmatrix} \cdot (\beta_1, \dots, \beta_n) = \begin{pmatrix} \alpha_1\beta_1 & \dots & \alpha_1\beta_n \\ \cdot & \dots & \cdot \\ \alpha_\ell\beta_1 & \dots & \alpha_\ell\beta_n \end{pmatrix}$$

Bemerkung 6. Das Matrizenprodukt ist assoziativ, d. h. für $m, n, p, q \in \mathbb{N}$ und Matrizen $A \in K^{m \times n}$, $B \in K^{n \times p}$, $C \in K^{p \times q}$ gilt stets

$$(A \cdot B) \cdot C = A \cdot (B \cdot C).$$

Beweis. Mit

$$A = (\alpha_{ij})_{\substack{i=1,\dots,m, \\ j=1,\dots,n}}, \quad B = (\beta_{jk})_{\substack{j=1,\dots,n, \\ k=1,\dots,p}}, \quad C = (\gamma_{kl})_{\substack{k=1,\dots,p, \\ \ell=1,\dots,q}}$$

erhält man

$$(A \cdot B) \cdot C = \left(\sum_{j=1}^n \alpha_{ij} \beta_{jk} \right)_{i,k} \cdot C = \left(\sum_{k=1}^p \sum_{j=1}^n \alpha_{ij} \beta_{jk} \gamma_{kl} \right)_{i,l},$$

$$A \cdot (B \cdot C) = A \cdot \left(\sum_{k=1}^p \beta_{jk} \gamma_{kl} \right)_{j,l} = \left(\sum_{j=1}^n \sum_{k=1}^p \alpha_{ij} \beta_{jk} \gamma_{kl} \right)_{i,l},$$

also wie gewünscht $(A \cdot B) \cdot C = A \cdot (B \cdot C)$. $\square$

Wir wollen nun Matrizen im Sinne linearer Abbildungen interpretieren. Zunächst ist festzustellen, dass jede Matrix $A = (\alpha_{ij})_{i,j} \in K^{m \times n}$ Anlass zu einer Abbildung

$$f: K^n \longrightarrow K^m, \quad x \longmapsto A \cdot x,$$

gibt, wobei wir K^n mit $K^{n \times 1}$ und K^m mit $K^{m \times 1}$ identifizieren. In ausführlicher Schreibweise lautet die Abbildungsvorschrift

$$\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \longmapsto \begin{pmatrix} \sum_{j=1}^n \alpha_{1j} \xi_j \\ \vdots \\ \sum_{j=1}^n \alpha_{mj} \xi_j \end{pmatrix}$$

und man sieht unmittelbar, dass f K -linear ist. Es ist A offenbar gerade die im Sinne von Definition 3 zu f gehörige Matrix, wenn man in K^n und K^m jeweils die aus den Einheitsvektoren bestehende kanonische Basis zugrunde legt. Wir wollen zeigen, dass dieses Beispiel in gewissem Sinne schon den Allgemeinfall beschreibt:

Satz 7. *Es sei $f: V \longrightarrow W$ eine K -lineare Abbildung zwischen Vektorräumen V und W mit Basen $X = (x_1, \dots, x_n)$ und $Y = (y_1, \dots, y_m)$. Dann gilt für $a \in V$*

$$f(a)_Y = A_{f,X,Y} \cdot a_X.$$

Beweis. Hat man $A_{f,X,Y} = (\alpha_{ij})_{i,j}$ und $a_X = (\xi_j)$, so gilt

$$f(x_j) = \sum_{i=1}^m \alpha_{ij} y_i, \quad j = 1, \dots, n,$$

sowie $a = \sum_{j=1}^n \xi_j x_j$. Dies ergibt

$$f(a) = \sum_{j=1}^n \xi_j f(x_j) = \sum_{j=1}^n \xi_j \sum_{i=1}^m \alpha_{ij} y_i = \sum_{i=1}^m \left(\sum_{j=1}^n \alpha_{ij} \xi_j \right) y_i,$$

also wie gewünscht $f(a)_Y = A_{f,X,Y} \cdot a_X$. $\square$

Die Formel aus Satz 7 lässt sich insbesondere auf den Fall $V = W$ und die identische Abbildung $\text{id}: V \longrightarrow V$ anwenden. Sie beschreibt dann einen Basiswechsel und zeigt, wie sich Koordinatenvektoren bezüglich der Basis X in solche bezüglich der Basis Y umrechnen lassen:

$$a_Y = A_{\text{id}, X, Y} \cdot a_X$$

Um die Aussage von Satz 7 noch etwas genauer zu interpretieren, führen wir die Abbildungen $\kappa_X: V \longrightarrow K^n$, $a \longmapsto a_X$, und $\kappa_Y: W \longrightarrow K^m$, $b \longmapsto b_Y$, ein, welche einem Vektor jeweils den zugehörigen Koordinatenspaltenvektor zuordnen; κ_X, κ_Y sind nach Bemerkung 2 Isomorphismen von K -Vektorräumen. Weiter sei $\tilde{f}: K^n \longrightarrow K^m$ gegeben durch $x \longmapsto A_{f, X, Y} \cdot x$. Dann besagt die Formel in Satz 7 gerade:

Korollar 8. *In der vorstehenden Situation ist das Diagramm*

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \downarrow \kappa_X & & \downarrow \kappa_Y \\ K^n & \xrightarrow{\tilde{f}} & K^m \end{array}$$

kommutativ.

Fassen wir daher die Abbildungen κ_X, κ_Y als Identifizierungen auf, so geht f in $\tilde{f}$ über und ist nichts anderes als die Multiplikation mit der Matrix $A_{f, X, Y}$.

Als Nächstes wollen wir zeigen, dass das Matrizenprodukt als Komposition linearer Abbildungen interpretiert werden kann.

Satz 9. *Es seien U, V, W K -Vektorräume mit Basen X, Y, Z . Für lineare Abbildungen $f: U \longrightarrow V$ und $g: V \longrightarrow W$ gilt dann*

$$A_{g \circ f, X, Z} = A_{g, Y, Z} \cdot A_{f, X, Y}.$$

Beweis. Auch diese Formel ist leicht nachzurechnen. Gelte

$$X = (x_1, \dots, x_n), \quad Y = (y_1, \dots, y_m), \quad Z = (z_1, \dots, z_\ell)$$

sowie

$$A_{f, X, Y} = (\alpha_{jk})_{\substack{j=1, \dots, m \\ k=1, \dots, n}}, \quad A_{g, Y, Z} = (\beta_{ij})_{\substack{i=1, \dots, \ell \\ j=1, \dots, m}}.$$

Dann folgt $f(x_k) = \sum_{j=1}^m \alpha_{jk} y_j$ für $k = 1, \dots, n$ und somit

$$\begin{aligned}
g \circ f(x_k) &= \sum_{j=1}^m \alpha_{jk} g(y_j) \\
&= \sum_{j=1}^m \alpha_{jk} \sum_{i=1}^{\ell} \beta_{ij} z_i \\
&= \sum_{i=1}^{\ell} \left(\sum_{j=1}^m \beta_{ij} \alpha_{jk} \right) \cdot z_i,
\end{aligned}$$

also

$$A_{g \circ f, X, Z} = \left(\sum_{j=1}^m \beta_{ij} \alpha_{jk} \right)_{i,k} = A_{g, Y, Z} \cdot A_{f, X, Y},$$

wie gewünscht.

Wir können den Beweis aber auch anders führen und uns dabei jegliche Rechnung ersparen. Man hat nämlich nach Korollar 8 ein kommutatives Diagramm

$$\begin{array}{ccccc}
U & \xrightarrow{f} & V & \xrightarrow{g} & W \\
\downarrow \kappa_X & & \downarrow \kappa_Y & & \downarrow \kappa_Z \\
K^n & \xrightarrow{\tilde{f}} & K^m & \xrightarrow{\tilde{g}} & K^\ell
\end{array}$$

wobei $\tilde{f}$ durch $x \mapsto A_{f, X, Y} \cdot x$ und $\tilde{g}$ durch $y \mapsto A_{g, Y, Z} \cdot y$ erklärt ist. Mit Bemerkung 6 sieht man dann, dass $A_{g \circ f, X, Z}$ und $A_{g, Y, Z} \cdot A_{f, X, Y}$ zwei Matrizen sind, die die K -lineare Abbildung $g \circ f$ bezüglich der Basen X und Z beschreiben. Mit Satz 4 ergibt sich daraus $A_{g \circ f, X, Z} = A_{g, Y, Z} \cdot A_{f, X, Y}$. $\square$

Wir wollen abschließend noch einige Regeln für das Rechnen mit Matrizen auflisten, die sich leicht durch direkte Verifikation nachrechnen lassen; vgl. auch Bemerkung 6. Sei $\alpha \in K$ und seien A, B, C Matrizen mit Koeffizienten aus K . Weiter bezeichne 0 eine Nullmatrix und E eine (quadratische) Einheitsmatrix. Dann gelten die Formeln

$$\begin{aligned}
A + B &= B + A, \\
A + 0 &= A, \\
EA &= A, \quad BE = B, \\
(\alpha A)B &= A(\alpha B) = \alpha(AB), \\
A(B + C) &= AB + AC, \\
(A + B)C &= AC + BC, \\
(AB)C &= A(BC),
\end{aligned}$$

wobei wir in den einzelnen Gleichungen verlangen, dass die Zeilen- bzw. Spaltenanzahlen so gewählt sind, dass die aufgeführten Summen und Produkte auch gebildet werden können. Mit Matrizen kann man daher, was die Addition und

Multiplikation angeht, (fast) wie gewohnt rechnen. Allerdings darf man in einem Produkt $A \cdot B$ die Faktoren nicht vertauschen, selbst dann nicht, wenn für Matrizen A, B die Produkte $A \cdot B$ und $B \cdot A$ beide erklärt sind; man betrachte etwa den Fall einer Zeile A und einer Spalte B , jeweils mit n Komponenten. Sogar für quadratische Matrizen, also solche, bei denen die Zeilenzahl mit der Spaltenzahl übereinstimmt, ist das Produkt $A \cdot B$ im Allgemeinen verschieden von $B \cdot A$, wie man an einfachen Beispielen nachprüfen kann.

Aufgaben

1. Man berechne die Produkte AB und BA für die Matrizen

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

2. Man wähle in den angegebenen K -Vektorräumen V eine kanonische Basis X und bestimme zu den linearen Abbildungen $f: V \rightarrow V$ jeweils die zugehörige Matrix $A_{f,X,X}$.

- (i) $V = \mathbb{R}^2$, $K = \mathbb{R}$, f Drehung um 90° im mathematisch positiven Sinne.
- (ii) $V = \mathbb{R}^2$, $K = \mathbb{R}$, f Spiegelung an der Geraden $y = x$.
- (iii) $V = \mathbb{Q}(\sqrt{2})$, $K = \mathbb{Q}$, f Multiplikation mit $\alpha + \beta\sqrt{2}$, wobei $\alpha, \beta \in \mathbb{Q}$.

3. Es sei K ein Körper. Man zeige, dass die Abbildung

$$f: K \rightarrow K^{2 \times 2}, \quad x \mapsto \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}$$

injektiv ist und für $x, y \in K$ die Beziehung $f(x + y) = f(x) \cdot f(y)$ erfüllt.

4. Es sei $f: V \rightarrow W$ eine K -lineare Abbildung zwischen endlich-dimensionalen K -Vektorräumen mit $\text{rg } f = r$. Man zeige: Es existieren Basen X von V und Y von W , so dass die zu f gehörige Matrix $A_{f,X,Y}$ die folgende Gestalt besitzt:

$$A_{f,X,Y} = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

Dabei bezeichnet E_r die $(r \times r)$ -Einheitsmatrix, sowie 0 jeweils geeignete Nullmatrizen.

5. Für $m_1, m_2, n_1, n_2 \in \mathbb{N} - \{0\}$ betrachte man Matrizen

$$A \in K^{m_1 \times n_1}, \quad B \in K^{m_1 \times n_2}, \quad C \in K^{m_2 \times n_1}, \quad D \in K^{m_2 \times n_2}, \\ E \in K^{n_1 \times r}, \quad F \in K^{n_2 \times r}$$

und zeige

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} E \\ F \end{pmatrix} = \begin{pmatrix} AE + BF \\ CE + DF \end{pmatrix},$$

wobei man diese Gleichung in nahe liegender Weise als Gleichung zwischen Matrizen in $K^{(m_1+m_2) \times r}$ auffasse.

6. Es sei V ein endlich-dimensionaler K -Vektorraum und $f: V \longrightarrow V$ ein Endomorphismus. Man zeige:

- (i) Es existiert genau dann ein nicht-trivialer linearer Unterraum $U \subset V$ mit $f(U) \subset U$, wenn es in V eine Basis X gibt mit $A_{f,X,X} = \begin{pmatrix} \square & * \\ 0 & \square \end{pmatrix}$.
- (ii) Es existieren genau dann nicht-triviale lineare Unterräume $U_1, U_2 \subset V$ mit $V = U_1 \oplus U_2$ und $f(U_i) \subset U_i$ für $i = 1, 2$, wenn es in V eine Basis X gibt mit $A_{f,X,X} = \begin{pmatrix} \square & 0 \\ 0 & \square \end{pmatrix}$.

Dabei bezeichnen die Symbole 0 jeweils eine Nullmatrix geeigneten Typs, $\square$ nicht-leere quadratische Matrizen, sowie $*$ eine geeignete rechteckige Matrix, jeweils mit (unspezifizierten) Koeffizienten aus K .

7. Es sei V ein nicht-trivialer K -Vektorraum mit Basis X und $f: V \longrightarrow V$ ein Endomorphismus, so dass gilt:

$$A_{f,X,X} = \begin{pmatrix} 0 & 1 & 1 & 1 & \dots & 1 \\ 0 & 0 & 1 & 1 & \dots & 1 \\ 0 & 0 & 0 & 1 & \dots & 1 \\ \cdot & \cdot & \cdot & \cdot & \dots & \cdot \\ 0 & 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

Man berechne $\dim_K(\ker f)$.

8. Es sei $E \in K^{m \times m}$, $m \geq 1$, die Einheitsmatrix und $N = (\alpha_{ij})_{i,j=1,\dots,m} \in K^{m \times m}$ eine Matrix mit $\alpha_{ij} = 0$ für $i \geq j$. Man zeige:

- (i) $N^m = 0$.
- (ii) Die Matrix $B = E + N$ ist invertierbar, d. h. es existiert eine Matrix C in $K^{m \times m}$ mit $BC = CB = E$. (Hinweis: Man versuche, die Formel für die geometrische Reihe zu verwenden.)

3.2 Das Gaußsche Eliminationsverfahren und der Rang einer Matrix

Es sei $A = (\alpha_{ij})_{i,j} \in K^{m \times n}$ eine $(m \times n)$ -Matrix und

$$f: K^n \longrightarrow K^m, \quad x \longmapsto Ax,$$

die durch A definierte lineare Abbildung. Wie wir in Abschnitt 3.1 bemerkt haben, ist A die zu f gehörige Matrix, wenn man in K^n und K^m jeweils die aus den Einheitsvektoren bestehende kanonische Basis zugrunde legt. In Abschnitt 2.1 hatten wir den Rang einer linearen Abbildung f als die Dimension des Bildes im f erklärt. In unserer konkreten Situation wird das Bild im f von den Bildern $f(e_1), \dots, f(e_n)$ der kanonischen Basis $e_1, \dots, e_n$ von K^n erzeugt, also von den Spaltenvektoren der Matrix A . Der Rang von f ist daher gleich der

Dimension des von den Spalten von A in K^n erzeugten linearen Unterraums. Man bezeichnet diese Dimension auch als den Spaltenrang von A . In ähnlicher Weise besitzt A einen Zeilenrang.

Definition 1. Es sei $A = (\alpha_{ij})_{i,j} \in K^{m \times n}$ eine Matrix. Dann nennt man die Dimension des von den Spalten

$$\begin{pmatrix} \alpha_{11} \\ \vdots \\ \alpha_{m1} \end{pmatrix}, \quad \dots, \quad \begin{pmatrix} \alpha_{1n} \\ \vdots \\ \alpha_{mn} \end{pmatrix}$$

von A erzeugten Unterraums in K^m den Spaltenrang von A ; dieser wird mit $\text{rg}_s A$ bezeichnet. Entsprechend heißt die Dimension des von den Zeilen

$$(\alpha_{11}, \dots, \alpha_{1n}), \quad \dots, \quad (\alpha_{m1}, \dots, \alpha_{mn})$$

von A erzeugten Unterraums in K^n (wobei wir die Elemente von K^n hier als Zeilenvektoren auffassen) der Zeilenrang von A ; dieser wird mit $\text{rg}_z A$ bezeichnet.

Als wichtiges Resultat werden wir zeigen, dass Spaltenrang und Zeilenrang einer Matrix stets übereinstimmen. Man schreibt dann einfach $\text{rg } A$ anstelle von $\text{rg}_s A$ oder $\text{rg}_z A$ und nennt dies den *Rang* von A . Zunächst wollen wir jedoch die oben erwähnte Beziehung zwischen dem Rang einer linearen Abbildung $K^n \rightarrow K^m$ und dem Rang der zugehörigen Matrix auf beliebige lineare Abbildungen ausdehnen.

Bemerkung 2. Es sei $f: V \rightarrow W$ eine K -lineare Abbildung zwischen endlich-dimensionalen Vektorräumen mit Basen X bzw. Y und mit beschreibender Matrix $A_{f,X,Y}$. Dann gilt $\text{rg } f = \text{rg}_s A_{f,X,Y}$. Insbesondere hängt der Spaltenrang der f beschreibenden Matrix $A_{f,X,Y}$ nicht von der Wahl der Basen X und Y ab.

Beweis. Man betrachte das kommutative Diagramm

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \downarrow \kappa_X & & \downarrow \kappa_Y \\ K^n & \xrightarrow{\tilde{f}} & K^m \end{array}$$

aus 3.1/8, wobei κ_X (bzw. κ_Y) gerade derjenige Isomorphismus ist, der einem Vektor aus V (bzw. W) den zugehörigen Koordinatenspaltenvektor bezüglich X (bzw. Y) zuordnet. Weiter ist $\tilde{f}$ die Multiplikation mit $A_{f,X,Y}$. Es wird dann im f unter κ_Y isomorph auf im $\tilde{f}$ abgebildet, und es folgt

$$\text{rg } f = \dim_K(\text{im } f) = \dim_K(\text{im } \tilde{f}) = \text{rg } \tilde{f},$$

wobei, wie eingangs festgestellt, $\text{rg } \tilde{f}$ gerade der Spaltenrang der Matrix $A_{f,X,Y}$ ist. $\square$

Wir wollen nun das *Gaußsche Eliminationsverfahren* besprechen, welches es erlaubt, den Zeilen- bzw. Spaltenrang von Matrizen mittels elementarer Rechnung zu bestimmen, und zwar werden wir uns hier speziell mit der Bestimmung des Zeilenrangs beschäftigen. Zu diesem Zwecke interpretieren wir die Vektoren in K^n als Zeilenvektoren und bauen dementsprechend Matrizen aus $K^{m \times n}$ aus m solcher Zeilenvektoren auf. Wir schreiben also Matrizen $A \in K^{m \times n}$ in der Form

$$A = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & \dots & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn} \end{pmatrix}$$

mit den Zeilenvektoren $a_i = (\alpha_{i1}, \dots, \alpha_{in}) \in K^n$, $i = 1, \dots, m$. Auf solche Matrizen wollen wir so genannte *elementare Zeilenumformungen* anwenden, die, wie wir noch sehen werden, insbesondere den Zeilenrang nicht ändern. Diese Umformungen lassen sich auch durch Multiplikation von links mittels Matrizen, die man als *Elementarmatrizen* bezeichnet, realisieren. Wir wollen daher einige quadratische Matrizen einführen, die wir zur Beschreibung der Elementarmatrizen benötigen. Für $i, j = 1, \dots, m$ sei $E_{ij} \in K^{m \times m}$ diejenige $(m \times m)$ -Matrix, die am Schnittpunkt der i -ten Zeile und j -ten Spalte eine 1 stehen hat und ansonsten aus lauter Nullen besteht. Weiter bezeichne $E = (\delta_{ij})_{i,j} \in K^{m \times m}$ die *Einheitsmatrix*, wobei δ_{ij} das Kronecker-Symbol ist. Für diese Matrix gilt also $E = \sum_{i=1}^m E_{ii}$, sie besteht auf der Diagonalen aus Einsen und ansonsten aus Nullen. Die Matrix E heißt Einheitsmatrix, da $E \cdot A = A$ für jede Matrix $A \in K^{m \times n}$ und $B \cdot E = B$ für jede Matrix $B \in K^{n \times m}$ gilt. Folgende Typen elementarer Zeilenumformungen werden wir verwenden:

Typ I: Man multipliziere eine Zeile von A , etwa die i -te, mit einem Skalar $\alpha \in K^*$, also:

$$\begin{pmatrix} \vdots \\ a_i \\ \vdots \end{pmatrix} \mapsto \begin{pmatrix} \vdots \\ \alpha \cdot a_i \\ \vdots \end{pmatrix}$$

Diese Umformung von A kann man auch erreichen, indem man A von links mit der Elementarmatrix $E + (\alpha - 1)E_{ii}$ multipliziert. Letztere Matrix unterscheidet sich von der Einheitsmatrix E dadurch, dass auf der Diagonalen an der Stelle (i, i) statt einer 1 der Faktor α steht.

Typ II: Man addiere zu einer Zeile von A , etwa der i -ten, eine weitere Zeile von A , etwa die j -te, wobei $i \neq j$ gelte, also:

$$\begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} \mapsto \begin{pmatrix} \vdots \\ a_i + a_j \\ \vdots \\ a_j \\ \vdots \end{pmatrix}$$

Diese Umformung kann man auch erreichen, indem man A von links mit der Elementarmatrix $E + E_{ij}$ multipliziert.

Typ III: Man addiere zu einer Zeile von A , etwa der i -ten, ein Vielfaches einer weiteren Zeile von A , etwa der j -ten, wobei $i \neq j$ gelte, also:

$$\begin{pmatrix} \dots \\ a_i \\ \dots \\ a_j \\ \dots \end{pmatrix} \mapsto \begin{pmatrix} \dots \\ a_i + \alpha a_j \\ \dots \\ a_j \\ \dots \end{pmatrix}$$

mit $\alpha \in K$. Diese Umformung kann man auch erreichen, indem man A von links mit der Elementarmatrix $E + \alpha E_{ij}$ multipliziert.

Typ IV: Man vertausche zwei Zeilen von A , etwa die i -te und die j -te, also:

$$\begin{pmatrix} \dots \\ a_i \\ \dots \\ a_j \\ \dots \end{pmatrix} \mapsto \begin{pmatrix} \dots \\ a_j \\ \dots \\ a_i \\ \dots \end{pmatrix}$$

Diese Umformung kann man auch erreichen, indem man A von links mit der Elementarmatrix $E - E_{ii} - E_{jj} + E_{ij} + E_{ji}$ multipliziert. Letztere Matrix erhält man aus der Einheitsmatrix, indem man die i -te und j -te Zeile (oder, alternativ, Spalte) miteinander vertauscht.

Wie man leicht sehen kann, sind die Typen III und IV Kombinationen der Typen I und II.

Satz 3. *Es sei $A \in K^{m \times n}$ eine Matrix und $B \in K^{m \times n}$ eine weitere, die mittels elementarer Zeilentransformationen aus A hervorgeht. Dann erzeugen die Zeilenvektoren von A den gleichen linearen Unterraum in K^n wie die Zeilenvektoren von B . Insbesondere ist der Zeilenrang einer Matrix invariant unter elementaren Zeilenumformungen.*

Beweis. Es seien $a_1, \dots, a_m$ Vektoren eines K -Vektorraums V . Wählt man dann $\alpha \in K^*$ und $i, j \in \{1, \dots, m\}$, $i \neq j$, so gilt offenbar für den von $a_1, \dots, a_m$ erzeugten Unterraum

$$\begin{aligned} \langle a_1, \dots, a_i, \dots, a_j, \dots, a_m \rangle &= \langle a_1, \dots, \alpha a_i, \dots, a_j, \dots, a_m \rangle, \\ \langle a_1, \dots, a_i, \dots, a_j, \dots, a_m \rangle &= \langle a_1, \dots, a_i + a_j, \dots, a_j, \dots, a_m \rangle, \\ \langle a_1, \dots, a_i, \dots, a_j, \dots, a_m \rangle &= \langle a_1, \dots, a_i + \alpha a_j, \dots, a_j, \dots, a_m \rangle, \\ \langle a_1, \dots, a_i, \dots, a_j, \dots, a_m \rangle &= \langle a_1, \dots, a_j, \dots, a_i, \dots, a_m \rangle. \end{aligned}$$

Indem wir vorstehende Überlegung iterativ auf die Situation $V = K^n$ anwenden, folgt die Behauptung. $\square$

Theorem 4 (Gaußsches Eliminationsverfahren). *Jede Matrix $A \in K^{m \times n}$ lässt sich durch elementare Zeilenumformungen auf Zeilenstufenform bringen, d. h. auf die Form*

$$B = \begin{pmatrix} 0 & \dots & 0 & \boxed{\beta_1} & \dots & * & * & \dots & * & \dots & * & \dots & * & \dots & * \\ 0 & \dots & 0 & 0 & \dots & 0 & \boxed{\beta_2} & \dots & * & \dots & * & \dots & * & \dots & * \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & \dots & * & \dots & * & \dots & * \\ \dots & & \dots & & \dots & & \dots & & \dots & & \dots & & \dots & & \dots \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & \dots & \overline{0 \dots 0} & \boxed{\beta_r} & \dots & * \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ \dots & & \dots & & \dots & & \dots & & \dots & & \dots & & \dots & & \dots \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 \end{pmatrix}$$

mit Koeffizienten $\beta_1, \dots, \beta_r \in K^*$. Die ersten r Zeilen von B bilden dann eine Basis des von den Zeilenvektoren von A in K^n erzeugten linearen Unterraums. Insbesondere gilt

$$\operatorname{rg}_z A = \operatorname{rg}_z B = r.$$

Beweis. Es seien $b_1, \dots, b_m$ die Zeilen der obigen Matrix B , das Element β_i gehöre dabei zur Spalte mit Index j_i , $i = 1, \dots, r$. Wir wollen zunächst zeigen, dass $b_1, \dots, b_r$ linear unabhängig sind und damit eine Basis des linearen Unterraums

$$\langle b_1, \dots, b_r \rangle = \langle b_1, \dots, b_m \rangle \subset K^n$$

bilden. Wenn wir A mittels elementarer Zeilentransformationen in B überführen können (was wir weiter unten zeigen werden), so handelt es sich gemäß Satz 3 bei diesem Unterraum gerade um den von den Zeilenvektoren von A erzeugten linearen Unterraum von K^n . Insbesondere ergibt sich $\operatorname{rg}_z A = \operatorname{rg}_z B = r$.

Gelte etwa $\sum_{i=1}^r \alpha_i b_i = 0$ für gewisse Koeffizienten $\alpha_i \in K$. Hat man dann $B = (\beta_{ij})_{i,j}$, also mit $\beta_{i,j_i} = \beta_i$ für $i = 1, \dots, r$, so erhält man die Gleichungen

$$\begin{aligned} \alpha_1 \beta_1 &= 0, \\ \alpha_1 \beta_{1,j_2} + \alpha_2 \beta_2 &= 0, \\ \alpha_1 \beta_{1,j_3} + \alpha_2 \beta_{2,j_3} + \alpha_3 \beta_3 &= 0, \\ &\dots \\ \alpha_1 \beta_{1,j_r} + \alpha_2 \beta_{2,j_r} + \alpha_3 \beta_{3,j_r} + \dots + \alpha_r \beta_r &= 0, \end{aligned}$$

woraus sich zunächst $\alpha_1 = 0$, dann $\alpha_2 = 0$ usw. bis schließlich $\alpha_r = 0$ ergibt. Es folgt die lineare Unabhängigkeit von $b_1, \dots, b_r$.

Es bleibt noch zu zeigen, dass sich jede Matrix $A = (\alpha_{ij})_{i,j} \in K^{m \times n}$ auf Zeilenstufenform bringen lässt. Hierzu gehe man wie folgt vor. Da die Nullmatrix bereits Zeilenstufenform besitzt, darf man $A \neq 0$ annehmen. Man wähle einen minimalen Index $j = j_1$, so dass es ein i mit $\alpha_{ij} \neq 0$ gibt, vertausche die erste mit der i -ten Zeile und setze $\beta_1 = \alpha_{i,j_1}$. Durch $m - 1$ Umformungen des Typs III, welche die erste Zeile invariant lassen, kann man dann erreichen, dass alle Elemente in der j_1 -ten Spalte unterhalb β_1 zu Null werden. Die resultierende Matrix hat folgende Gestalt:

$$\left(\begin{array}{cccc|ccc} 0 & \dots & 0 & \beta_1 & * & \dots & * \\ 0 & \dots & 0 & 0 & & & \\ \cdot & \dots & \cdot & \cdot & & & \\ 0 & \dots & 0 & 0 & & & \end{array} \middle| \begin{array}{c} A^{(1)} \end{array} \right)$$

Man kann nun dasselbe Verfahren in einem zweiten Schritt auf $A^{(1)}$ anstelle von A anwenden. Indem man die für $A^{(1)}$ benötigten Transformationen als Zeilentransformationen der Gesamtmatrix interpretiert und das Verfahren genügend oft wiederholt, lässt sich in rekursiver Weise die behauptete Zeilenstufenform realisieren. Erhält man dabei in einem r -ten Schritt erstmalig $A^{(r)}$ als Null- oder leere Matrix, so ist die Konstruktion beendet. $\square$

Wir wollen ein Beispiel zur Transformation einer Matrix auf Zeilenstufenform betrachten:

$$\begin{pmatrix} 0 & 0 & 2 \\ 0 & 2 & 4 \\ 0 & 1 & 1 \end{pmatrix} \mapsto \begin{pmatrix} 0 & 1 & 1 \\ 0 & 2 & 4 \\ 0 & 0 & 2 \end{pmatrix} \mapsto \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 2 \end{pmatrix} \mapsto \begin{pmatrix} 0 & \boxed{1} & 1 \\ 0 & 0 & \boxed{2} \\ 0 & 0 & 0 \end{pmatrix}$$

Das Gaußsche Eliminationsverfahren kann man insbesondere verwenden, um die Dimension eines von gewissen gegebenen Vektoren erzeugten linearen Unterraums $U = \langle a_1, \dots, a_m \rangle \subset K^n$ zu berechnen bzw. um eine Basis von U anzugeben. In Verbindung mit der Dimensionsformel 1.6/5 ist es dann möglich, die Dimension des Schnittes $U \cap U'$ zweier linearer Unterräume $U, U' \subset K^n$ zu ermitteln. Weiter kann man natürlich zu einer linearen Abbildung $f: K^n \rightarrow K^m$ eine Basis des Bildes und damit insbesondere den Rang $\text{rg } f$ bestimmen (wobei man Vektoren in K^m am besten als Zeilenvektoren interpretiert). Aus der Dimensionsformel 2.1/10 ergibt sich dann auch die Dimension des Kerns von f . Dass das Gaußsche Eliminationsverfahren überdies dazu geeignet ist, eine Basis von $\ker f$ anzugeben, werden wir genauer noch im Abschnitt 3.5 über lineare Gleichungssysteme sehen. Bei der Lösung dieser Gleichungssysteme führt das Gaußsche Verfahren zu einer sukzessiven Reduzierung des Systems der unbekannten Größen, bis man schließlich ein Restsystem von Größen erhält, deren Werte frei gewählt werden können. Einige der unbekannten Größen werden also entfernt (eliminiert), so dass die Bezeichnung *Eliminationsverfahren* plausibel wird.

In Analogie zu den elementaren Zeilenumformungen kann man natürlich auch elementare Spaltenumformungen von Matrizen erklären. Solche Umformungen lassen sich als Multiplikation von rechts mit geeigneten Elementarmatrizen interpretieren, und man kann ähnlich wie in Theorem 4 zeigen, dass sich jede Matrix auf Spaltenstufenform transformieren lässt. Man braucht sich dies aber nicht in allen Details zu überlegen, denn wir wollen als Nächstes das Transponieren von Matrizen behandeln. Unter diesem Prozess gehen elementare Zeilenumformungen über in elementare Spaltenumformungen und umgekehrt. Ähnliches gilt für Matrizen in Zeilenstufenform bzw. Spaltenstufenform.

Definition 5. Es sei

$$A = (\alpha_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} \in K^{m \times n}.$$

Dann heißt

$$A^t = (\alpha_{ij})_{\substack{j=1,\dots,n \\ i=1,\dots,m}} \in K^{n \times m}$$

die zu A transponierte Matrix.

Es geht also A^t aus A hervor, indem man Spalten- und Zeilenindex miteinander vertauscht. Mit anderen Worten, man spiegelt A an der Hauptdiagonalen, welche durch die Positionen (i, i) , $i = 1, \dots, \min(m, n)$, charakterisiert ist. Dabei werden die Zeilenvektoren von A zu den Spaltenvektoren von A^t und entsprechend die Spaltenvektoren von A zu den Zeilenvektoren von A^t . Unmittelbar ersichtlich ist:

Bemerkung 6. Die Abbildung $K^{m \times n} \longrightarrow K^{n \times m}$, $A \longmapsto A^t$, ist ein Isomorphismus von K -Vektorräumen. Insbesondere gilt $(A + B)^t = A^t + B^t$ und $(\alpha A)^t = \alpha A^t$ für $A, B \in K^{m \times n}$, $\alpha \in K$.

Bemerkung 7. Für $A \in K^{m \times n}$ gilt $(A^t)^t = A$ sowie $\text{rg}_z A = \text{rg}_s A^t$ und $\text{rg}_s A = \text{rg}_z A^t$.

Bemerkung 8. Für komпонierbare Matrizen A, B gilt $(A \cdot B)^t = B^t \cdot A^t$.

Wir wollen nur den Beweis zu Bemerkung 8 angeben. Für

$$A = (\alpha_{ij})_{i,j} \in K^{\ell \times m}, \quad B = (\beta_{jk})_{j,k} \in K^{m \times n}$$

ergibt sich

$$\begin{aligned} (A \cdot B)^t &= \left(\left(\sum_j \alpha_{ij} \beta_{jk} \right)_{i,k} \right)^t = \left(\sum_j \alpha_{ij} \beta_{jk} \right)_{k,i} \\ &= \left(\sum_j \beta_{jk} \alpha_{ij} \right)_{k,i} = B^t \cdot A^t, \end{aligned}$$

wie behauptet. □

Als Nächstes zeigen wir, dass das Transponieren von Matrizen dem Dualisieren linearer Abbildungen entspricht.

Satz 9. Es sei $f: V \longrightarrow W$ eine lineare Abbildung zwischen endlich-dimensionalen K -Vektorräumen mit Basen X von V und Y von W . Ist dann X^* (bzw. Y^*) die duale Basis zu X (bzw. Y), und bezeichnet $f^*: W^* \longrightarrow V^*$ die zu f duale lineare Abbildung, so gilt

$$A_{f^*, Y^*, X^*} = (A_{f, X, Y})^t.$$

Beweis. Es sei

$$\begin{aligned} X &= (x_1, \dots, x_n), & Y &= (y_1, \dots, y_m), \\ X^* &= (x_1^*, \dots, x_n^*), & Y^* &= (y_1^*, \dots, y_m^*), \end{aligned}$$

wobei man also $x_j^*(x_\nu) = \delta_{j\nu}$ und $y_i^*(y_\mu) = \delta_{i\mu}$ hat. Insbesondere wird dann für $\varphi \in V^*$ die eindeutig bestimmte Darstellung als Linearkombination der Basis X^* durch $\varphi = \sum_{j=1}^n \varphi(x_j) x_j^*$ gegeben. Dies ist eine Konsequenz von 2.1/7, wie wir bereits im Beweis zu 2.3/5 gesehen hatten. Somit lässt sich die zu f^* gehörige Matrix wie folgt beschreiben:

$$A_{f^*, Y^*, X^*} = (f^*(y_1^*)_{X^*}, \dots, f^*(y_m^*)_{X^*}) = (f^*(y_i^*)(x_j))_{\substack{j=1, \dots, n \\ i=1, \dots, m}}$$

Ist nun die Matrix zu f gegeben durch

$$A_{f, X, Y} = (\alpha_{ij})_{\substack{i=1, \dots, m, \\ j=1, \dots, n}}, \quad \text{d. h. durch} \quad f(x_j) = \sum_{\mu=1}^m \alpha_{\mu j} y_\mu,$$

so erhält man für $i = 1, \dots, m$

$$\begin{aligned} f^*(y_i^*) &= \sum_{j=1}^n f^*(y_i^*)(x_j) \cdot x_j^* = \sum_{j=1}^n (y_i^* \circ f)(x_j) \cdot x_j^* \\ &= \sum_{j=1}^n y_i^*(f(x_j)) \cdot x_j^* = \sum_{j=1}^n y_i^* \left(\sum_{\mu=1}^m \alpha_{\mu j} y_\mu \right) \cdot x_j^* \\ &= \sum_{j=1}^n \alpha_{ij} \cdot x_j^*, \end{aligned}$$

und dies bedeutet

$$A_{f^*, Y^*, X^*} = (\alpha_{ij})_{\substack{j=1, \dots, n \\ i=1, \dots, m}} = (A_{f, X, Y})^t,$$

wie behauptet. □

Benutzen wir nun 2.3/7, dass nämlich in der Situation von Satz 9 die Abbildungen f und f^* gleichen Rang haben, so ergibt sich mit Bemerkung 2 und Bemerkung 7

$$\begin{aligned} \text{rg}_s A_{f, X, Y} &= \text{rg } f = \text{rg } f^* = \text{rg}_s A_{f^*, Y^*, X^*} \\ &= \text{rg}_s (A_{f, X, Y})^t = \text{rg}_z A_{f, X, Y}. \end{aligned}$$

Da man Matrizen stets als lineare Abbildungen realisieren kann, erhält man als Folgerung zu Satz 9:

Korollar 10. Für Matrizen $A \in K^{m \times n}$ gilt $\text{rg}_s A = \text{rg}_z A$, d. h. Spalten- und Zeilenrang stimmen überein.

Wie bereits angedeutet, werden wir von nun an $\text{rg } A$ anstelle von $\text{rg}_s A$ bzw. $\text{rg}_z A$ schreiben und diese Zahl als den *Rang* der Matrix A bezeichnen. Die Übereinstimmung von Spalten- und Zeilenrang bei Matrizen ist ein nicht-triviales Resultat, das wir hier auf einfache Weise als Anwendung der Theorie dualer Abbildungen gewonnen haben. Man kann dieses Resultat aber auch anders herleiten, indem man benutzt, dass sich der Zeilen- bzw. Spaltenrang einer Matrix A nicht ändert, wenn man A von links oder rechts mit so genannten invertierbaren Matrizen multipliziert; vgl. Abschnitt 3.4.

Schließlich wollen wir unter Benutzung von Korollar 10 noch zeigen, dass sich die in Theorem 4 beschriebene Transformation einer Matrix auf Zeilenstufenform auch im Sinne von Spaltenvektoren interpretieren lässt. Dies ist insbesondere dann von Vorteil, wenn man aus einem gegebenen System von Vektoren ein maximales linear unabhängiges Teilsystem auswählen möchte.

Satz 11. Es sei $A = (a_1, \dots, a_n) \in K^{m \times n}$ eine Matrix mit den Spalten $a_1, \dots, a_n \in K^m$, welche sich mittels elementarer Zeilenumformungen auf Zeilenstufenform

$$B = \begin{pmatrix} 0 \dots 0 \mid \beta_1 \dots * & * \dots * & \dots & * \dots * & * \dots * \\ 0 \dots 0 \mid 0 \dots 0 \mid \beta_2 \dots * & \dots & * \dots * & * \dots * \\ 0 \dots 0 \mid 0 \dots 0 \mid 0 \dots 0 & \dots & * \dots * & * \dots * \\ \dots & \dots & \dots & \dots \\ 0 \dots 0 \mid 0 \dots 0 \mid 0 \dots 0 & \dots & 0 \dots 0 \mid \beta_r \dots * \\ 0 \dots 0 \mid 0 \dots 0 \mid 0 \dots 0 & \dots & 0 \dots 0 \mid 0 \dots 0 \\ \dots & \dots & \dots & \dots \\ 0 \dots 0 \mid 0 \dots 0 \mid 0 \dots 0 & \dots & 0 \dots 0 \mid 0 \dots 0 \end{pmatrix}$$

bringen lässt. Dabei sei für $i = 1, \dots, r$ das Element $\beta_i \in K^*$ jeweils in der Spalte mit Index j_i positioniert, mit $1 \leq j_1 < \dots < j_r \leq n$. Dann sind die Vektoren $a_{j_1}, \dots, a_{j_r}$ linear unabhängig, und es gilt

$$\langle a_1, \dots, a_j \rangle = \langle a_{j_1}, \dots, a_{j_i} \rangle, \quad j = 1, \dots, n,$$

wenn $i \leq r$ jeweils maximal mit $j_i \leq j$ gewählt ist.

Beweis. Verkleinert man die Matrix A zu einer Matrix A' , indem man gewisse Spalten streicht, so erhält man eine zugehörige Zeilenstufenform von A' , indem man B durch Streichen der entsprechenden Spalten zu einer Matrix B' verkleinert. Es folgt dann $\text{rg}_z B' = \text{rg}_z A' = \text{rg}_s A'$, so dass die Dimension des von den Spalten von A' erzeugten linearen Unterraums gerade gleich dem Rang von B' ist. Insbesondere folgt mit diesem Argument die lineare Unabhängigkeit des Systems $a_{j_1}, \dots, a_{j_r} \in K^m$ und mittels 1.5/14(ii) auch die behauptete Gleichheit linearer Unterräume. $\square$

Aufgaben

1. Man bringe die Matrix

$$A = \begin{pmatrix} 1 & 2 & 1 & 2 & 1 & 2 \\ 2 & 5 & 4 & 5 & 4 & 5 \\ 1 & 4 & 6 & 6 & 6 & 6 \\ 2 & 5 & 6 & 9 & 7 & 11 \end{pmatrix} \in \mathbb{R}^{4 \times 6}$$

auf Zeilenstufenform.

2. Man berechne den Zeilenrang der Matrix

$$A = \begin{pmatrix} 1 & 1 & 3 & 2 \\ 1 & 1 & 2 & 3 \\ 1 & 1 & 0 & 0 \end{pmatrix} \in K^{3 \times 4},$$

jeweils für die Körper $K = \mathbb{Q}$ und $K = \mathbb{F}_5$; dabei ist $\mathbb{F}_5$ der Körper mit 5 Elementen aus Abschnitt 1.3, Aufgabe 3.

3. Man prüfe, ob das System der folgenden Vektoren in
- $\mathbb{R}^5$
- linear unabhängig ist:

$$(5, 4, 3, 2, 1), \quad (1, 2, 3, 4, 5), \quad (2, 2, 2, 2, 1), \quad (1, 0, 0, 1, 1), \quad (1, 0, 1, 0, 1)$$

4. Es sei
- $U \subset \mathbb{R}^4$
- der lineare Unterraum, der von den Vektoren

$$(1, 2, 1, 2), \quad (2, 5, 4, 5), \quad (1, 4, 6, 6), \quad (2, 5, 6, 9), \quad (2, 6, 7, 8), \quad (1, 1, 0, 3)$$

erzeugt wird. Man berechne $\dim_{\mathbb{R}} U$ und gebe eine Basis von U an.

5. Es seien in
- $\mathbb{R}^5$
- die folgenden linearen Unterräume gegeben:

$$U = \langle (1, 0, 1, 0, 1), (2, 3, 4, 1, 2), (0, 3, 2, 1, 0) \rangle$$

$$U' = \langle (1, -1, 1, 0, 2), (1, 3, 3, 1, 1), (1, 2, 3, 1, 2) \rangle$$

Man berechne $\dim_{\mathbb{R}} U$ und $\dim_{\mathbb{R}} U'$ und zeige $U \subset U'$.

6. Man betrachte in
- $\mathbb{R}^6$
- die linearen Unterräume

$$U = \langle (1, 1, 1, 0, 1, 1), (2, 3, 4, 0, 2, 1), (0, 3, 2, 1, 1, 1) \rangle,$$

$$U' = \langle (2, 6, 6, 2, 2, 1), (0, 9, 4, 3, 0, 1), (1, 2, 3, 1, 2, 1) \rangle$$

und berechne $\dim_{\mathbb{R}}(U \cap U')$.

7. Man betrachte die Linearformen

$$f_1: \mathbb{R}^5 \longrightarrow \mathbb{R}, \quad (\alpha_1, \dots, \alpha_5) \longmapsto (\alpha_2 + \alpha_3 + \alpha_4 + \alpha_5)$$

$$f_2: \mathbb{R}^5 \longrightarrow \mathbb{R}, \quad (\alpha_1, \dots, \alpha_5) \longmapsto (\alpha_1 + 2\alpha_2 - \alpha_3 - \alpha_4)$$

$$f_3: \mathbb{R}^5 \longrightarrow \mathbb{R}, \quad (\alpha_1, \dots, \alpha_5) \longmapsto (5\alpha_1 + 2\alpha_2 - \alpha_3 + 2\alpha_4 - 2\alpha_5)$$

$$f_4: \mathbb{R}^5 \longrightarrow \mathbb{R}, \quad (\alpha_1, \dots, \alpha_5) \longmapsto (\alpha_1 - \alpha_2 + \alpha_4 - \alpha_5)$$

und prüfe, ob diese ein linear unabhängiges System im Dualraum $(\mathbb{R}^5)^*$ definieren.

8. Die Linearformen f_1, f_2, f_3, f_4 seien wie in Aufgabe 7. Für die $\mathbb{R}$ -lineare Abbildung

$$f: \mathbb{R}^5 \longrightarrow \mathbb{R}^4, \quad x \longmapsto (f_1(x), f_2(x), f_3(x), f_4(x)),$$

bestimme man $\dim_{\mathbb{R}} \ker f$ und $\dim_{\mathbb{R}} \operatorname{im} f$. Weiter gebe man eine Basis von $\operatorname{im} f$ an.

9. Es sei $\mathbb{F}_2$ der Körper mit 2 Elementen. Im $\mathbb{F}_2$ -Vektorraum $\mathbb{F}_2^{2000}$ sei eine Familie von Vektoren $(v_i)_{i=1, \dots, 2000}$ definiert durch

$$v_i = \begin{cases} 0 & \text{falls } i \text{ nicht durch 4 teilbar ist,} \\ (1 + \delta_{ij})_{j=1, \dots, 2000} & \text{falls } i \text{ durch 4, nicht durch 8 teilbar ist,} \\ (1 + \delta_{i-1, j} + \delta_{i+1, j})_{j=1, \dots, 2000} & \text{falls } i \text{ durch 8 teilbar ist.} \end{cases}$$

Man betrachte den von den v_i erzeugten linearen Unterraum $U \subset \mathbb{F}_2^{2000}$, bestimme $\dim_{\mathbb{F}_2} U$ und gebe eine Basis von U an.

3.3 Matrizenringe und invertierbare Matrizen

Für $n \in \mathbb{N}$ ist auf der Menge $K^{n \times n}$ aller $(n \times n)$ -Matrizen ähnlich wie bei einem Körper neben der Addition eine Multiplikation gegeben, die für $n \geq 2$ allerdings nicht kommutativ ist. Dabei zeigen Gleichungen des Typs

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

dass es für $n \geq 2$ nicht-triviale Nullteiler in $K^{n \times n}$ gibt. Eine Matrix $A \in K^{n \times n}$ heißt ein *Nullteiler*, wenn es eine von Null verschiedene Matrix $B \in K^{n \times n}$ mit $A \cdot B = 0$ oder $B \cdot A = 0$ gibt.

Definition 1. Eine Menge R mit zwei Verknüpfungen “+” (Addition) und “ $\cdot$ ” (Multiplikation) heißt ein Ring, wenn folgende Bedingungen erfüllt sind:

- (i) R ist eine abelsche Gruppe bezüglich der Addition.
- (ii) Die Multiplikation ist assoziativ, d. h. für $a, b, c \in R$ gilt

$$(a \cdot b) \cdot c = a \cdot (b \cdot c).$$

(iii) Addition und Multiplikation verhalten sich distributiv, d. h. für $a, b, c \in R$ gilt

$$a \cdot (b + c) = a \cdot b + a \cdot c, \quad (a + b) \cdot c = a \cdot c + b \cdot c.$$

Das neutrale Element bezüglich der Addition 0 wird als Nullelement von R bezeichnet. Der Ring R heißt kommutativ, wenn die Multiplikation kommutativ ist. Weiter nennt man ein Element $e \in R$ ein Einselement, wenn $e \cdot a = a = a \cdot e$ für alle $a \in R$ gilt; man schreibt dann auch 1 anstelle von e .

Natürlich ist ein Einselement eines Ringes, sofern es existiert, eindeutig bestimmt. Das einfachste Beispiel eines Ringes ist der *Nullring* 0. Dieser besteht

nur aus einem Element 0 mit den Verknüpfungen $0+0=0$ und $0\cdot 0=0$; hier ist also 0 sowohl das Nullelement wie auch das Einselement. Wir wollen aber noch einige interessantere Beispiele anführen.

(1) Der Ring $\mathbb{Z}$ der ganzen Zahlen ist ein kommutativer Ring mit Eins unter der gewöhnlichen Addition und Multiplikation.

(2) Jeder Körper, etwa $\mathbb{Q}$, $\mathbb{R}$ oder $\mathbb{C}$, ist ein kommutativer Ring mit Eins.

(3) Es sei V ein K -Vektorraum und $R = \text{Hom}_K(V, V)$ der K -Vektorraum der Endomorphismen von V . Dann ist R ein Ring, der so genannte *Endomorphismenring* von V , wenn man in R als Addition die gewöhnliche Addition linearer Abbildungen sowie als Multiplikation die Komposition linearer Abbildungen betrachtet. Dieser Ring wird mit $\text{End}_K(V)$ bezeichnet. Er enthält die Nullabbildung $0: V \rightarrow V$ als Nullelement, sowie die identische Abbildung $\text{id}: V \rightarrow V$ als Einselement. Man kann sich überlegen, dass $\text{End}_K(V)$ für $\dim_K V \geq 2$ nicht kommutativ ist.

(4) Für $n \in \mathbb{N}$ ist der K -Vektorraum $K^{n \times n}$ der quadratischen n -reihigen Matrizen ein Ring unter der in Abschnitt 3.1 eingeführten Addition und Multiplikation von Matrizen. Es ist $K^{0 \times 0}$ der Nullring und $K^{1 \times 1}$ ein Ring, den wir in kanonischer Weise mit dem Körper K identifizieren können. Weiter ist die Multiplikation in $K^{n \times n}$ für $n \geq 2$ nicht mehr kommutativ. Die Matrix $E = (\delta_{ij})_{i,j} \in K^{n \times n}$, oftmals auch mit 1 bezeichnet, ist ein Einselement in $K^{n \times n}$ und wird als n -reihige *Einheitsmatrix* bezeichnet. Wie wir bereits zu Beginn gesehen haben, enthält $K^{n \times n}$ für $n \geq 2$ nicht-triviale Nullteiler.

Die Beispiele (3) und (4) sind, wie man mit Hilfe der Resultate aus Abschnitt 3.1 einsieht, verwandt. Genauer gilt:

Satz 2. *Es sei V ein K -Vektorraum und $X = (x_1, \dots, x_n)$ eine Basis von V . Dann ist die Abbildung*

$$\Psi: \text{End}_K(V) \longrightarrow K^{n \times n}, \quad f \longmapsto A_{f,X,X},$$

ein Isomorphismus von Ringen, d. h. eine bijektive Abbildung mit

$$\Psi(f+g) = \Psi(f) + \Psi(g), \quad \Psi(f \circ g) = \Psi(f) \cdot \Psi(g)$$

für alle Elemente $f, g \in \text{End}_K(V)$.

Beweis. Aufgrund von 3.1/4 ist Ψ eine bijektive Abbildung, die mit der Addition verträglich ist. Die Verträglichkeit mit der Multiplikation folgt aus 3.1/9. $\square$

Ist a ein Element eines Ringes R mit Eins, so heißt ein Element $b \in R$ invers zu a , wenn $a \cdot b = 1 = b \cdot a$ gilt. Das inverse zu einem Element $a \in R$ ist, falls es existiert, stets eindeutig bestimmt. Sind nämlich zwei Elemente $b, b' \in R$ invers zu a , oder genauer, ist b invers zu a und gilt lediglich $a \cdot b' = 1$ (bzw. alternativ $b' \cdot a = 1$), so folgt

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

bzw.

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

Man schreibt a^{-1} für das inverse Element zu a und nennt a in diesem Fall *invertierbar* oder eine *Einheit*. In einem Körper ist jedes Element $a \neq 0$ eine Einheit. Andererseits gibt es in Matrizenringen $K^{n \times n}$ für $n \geq 2$ stets nicht-triviale Nullteiler, und solche Matrizen können keine Einheiten sein. Ist nämlich $A \in K^{n \times n}$ eine Einheit und $B \in K^{n \times n}$ eine Matrix mit $A \cdot B = 0$, so folgt notwendig $B = A^{-1} \cdot A \cdot B = 0$ mit der inversen Matrix A^{-1} zu A .

Bemerkung 3. *In jedem Ring R mit Eins ist die Menge R^* aller Einheiten eine Gruppe bezüglich der Multiplikation.*

Diese Aussage ergibt sich unmittelbar aus den definierenden Eigenschaften eines Ringes bzw. einer Einheit.

Bemerkung 4. *Es sei V ein K -Vektorraum und $f \in \text{End}_K(V)$. Dann ist äquivalent:*

(i) *Es existiert ein Element $g \in \text{End}_K(V)$ mit $f \circ g = \text{id}_V = g \circ f$, d. h. f ist eine Einheit in $\text{End}_K(V)$.*

(ii) *f ist ein Automorphismus von V .*

Beweis. Die Beziehung $f \circ g = \text{id}_V$ impliziert, dass f surjektiv ist, und entsprechend $g \circ f = \text{id}_V$, dass f injektiv ist. Aus Bedingung (i) ergibt sich somit (ii), und umgekehrt folgert man aus (ii) auch (i), wenn man g als Umkehrabbildung zu f erklärt. $\square$

Die Einheitengruppe des Endomorphismenrings $\text{End}_K(V)$ besteht also gerade aus allen Automorphismen von V . Wir bezeichnen diese Gruppe mit $\text{Aut}_K(V)$ und nennen dies die *Automorphismengruppe* von V . Die Einheitengruppe des Matrizenrings $K^{n \times n}$ wird mit $\text{Gl}(n, K)$ bezeichnet; man spricht hier von der *allgemeinen linearen Gruppe* (general linear group). Die Elemente von $\text{Gl}(n, K)$ heißen *invertierbare* (oder umkehrbare, bzw. reguläre, bzw. nicht-singuläre, bzw. nicht ausgeartete) Matrizen.

Satz 5. *Es sei V ein K -Vektorraum und $X = (x_1, \dots, x_n)$ eine Basis von V . Dann beschränkt sich der Isomorphismus*

$$\Psi: \text{End}_K(V) \xrightarrow{\sim} K^{n \times n}, \quad f \longmapsto A_{f, X, X},$$

aus Satz 2 zu einem Isomorphismus

$$\text{Aut}_K(V) \xrightarrow{\sim} \text{Gl}(n, K)$$

der zugehörigen Einheitengruppen, d. h. zu einer bijektiven Abbildung, welche die jeweiligen Gruppenstrukturen respektiert.

Beweis. Ψ bildet die identische Abbildung $\text{id}_V \in \text{End}_K(V)$ auf die Einheitsmatrix $E \in K^{n \times n}$ ab, also das Einselement von $\text{End}_K(V)$ auf das Einselement von $K^{n \times n}$. Hieraus ergibt sich unmittelbar, dass Ψ Einheiten in Einheiten überführt. Da die Umkehrabbildung Ψ^{-1} entsprechende Eigenschaften besitzt, folgt die Behauptung. $\square$

Wir wollen den Inhalt des Satzes noch in einem etwas allgemeineren Rahmen formulieren.

Satz 6. *Es sei $f: V \longrightarrow W$ eine K -lineare Abbildung zwischen Vektorräumen mit Basen $X = (x_1, \dots, x_n)$ und $Y = (y_1, \dots, y_n)$. Die beschreibende Matrix $A = A_{f,X,Y}$ ist genau dann invertierbar, wenn f ein Isomorphismus ist.*

Beweis. Man überlegt sich wie im Beweis zu Bemerkung 4, dass f genau dann bijektiv ist, wenn es eine K -lineare Abbildung $g: W \longrightarrow V$ mit $f \circ g = \text{id}_W$ und $g \circ f = \text{id}_V$ gibt. Letzteres ist aber aufgrund von 3.1/4 und 3.1/9 äquivalent zu der Existenz einer Matrix $B \in K^{n \times n}$ mit $A_{f,X,Y} \cdot B = 1$ und $B \cdot A_{f,X,Y} = 1$. $\square$

Korollar 7. *Für eine Matrix $A \in K^{n \times n}$ ist äquivalent:*

- (i) *A ist invertierbar.*
- (ii) *Es existiert eine Matrix $B \in K^{n \times n}$ mit $B \cdot A = 1$.*
- (iii) *Es existiert eine Matrix $B \in K^{n \times n}$ mit $A \cdot B = 1$.*
- (iv) *$\text{rg } A = n$.*

Beweis. Wir realisieren A als lineare Abbildung, etwa als Abbildung

$$f: K^n \longrightarrow K^n, \quad x \longmapsto A \cdot x.$$

Dann ist A die zu f gehörige Matrix bezüglich der kanonischen Basis auf K^n . Indem wir Satz 5 sowie 3.2/2 benutzen, genügt es zu zeigen, dass die folgenden Bedingungen äquivalent sind:

- (i') *f ist ein Automorphismus.*
- (ii') *Es existiert eine K -lineare Abbildung $g: K^n \longrightarrow K^n$ mit $g \circ f = \text{id}$.*
- (iii') *Es existiert eine K -lineare Abbildung $g: K^n \longrightarrow K^n$ mit $f \circ g = \text{id}$.*
- (iv') *$\text{rg } f = n$.*

Ist f ein Automorphismus, so folgen natürlich die Bedingungen (ii'), (iii') und (iv'). Umgekehrt, hat man (ii'), so ist f injektiv und nach 2.1/11 ein Automorphismus. Unter den Bedingungen (iii') bzw. (iv') schließlich ist f surjektiv und damit ebenfalls ein Automorphismus aufgrund von 2.1/11. $\square$

Bedingung (iv) aus Korollar 7 gibt uns ein nützliches Kriterium für die Invertierbarkeit einer Matrix $A \in K^{n \times n}$, insbesondere deshalb, weil wir den Rang einer Matrix mittels elementarer Zeilenumformungen in expliziter Weise bestimmen können; vgl. 3.2/4. So sieht man etwa, dass für $i, j \in \{1, \dots, n\}$, $i \neq j$, und $\alpha \in K^*$ die Elementarmatrizen

$$E + (\alpha - 1)E_{ii}, \quad E + \alpha E_{ij}, \quad E - E_{ii} - E_{jj} + E_{ij} + E_{ji}$$

aus $K^{n \times n}$ invertierbar sind; dabei sei E wie in Abschnitt 3.2 die Einheitsmatrix in $K^{n \times n}$ sowie $E_{ij} = (\delta_{i\mu} \delta_{j\nu})_{\mu, \nu}$. Die Elementarmatrizen gehen nämlich durch elementare Zeilenumformungen aus der Einheitsmatrix E hervor, haben also denselben Rang wie diese, also n ; vgl. 3.2/3. Weiter unten werden wir die Inversen zu den aufgeführten Elementarmatrizen (die sich im Übrigen leicht “erraten” lassen) auch noch explizit bestimmen.

Wir können die gerade beschriebene Argumentation dazu nutzen, um das in Abschnitt 3.2 behandelte Gaußsche Eliminationsverfahren auf die Invertierung von Matrizen auszudehnen. Man gehe etwa aus von einer Matrix $A \in K^{n \times n}$. Dann kann man A gemäß 3.2/4 mittels elementarer Zeilenumformungen auf Zeilenstufenform B bringen. Nehmen wir nun A als invertierbar an, so gilt $\text{rg } A = \text{rg } B = n$ gemäß Korollar 7, und es stehen in der Situation von 3.2/4 die Elemente $\beta_1, \dots, \beta_n$ gerade auf der Hauptdiagonalen, also:

$$B = \begin{pmatrix} \beta_1 & * & * & \dots & * \\ 0 & \beta_2 & * & \dots & * \\ . & . & . & . & . \\ . & . & . & . & * \\ 0 & 0 & \dots & 0 & \beta_n \end{pmatrix}$$

Man kann nun durch Multiplikation der i -ten Zeile mit β_i^{-1} , $i = 1, \dots, n$, annehmen, dass alle β_i den Wert 1 haben. Sodann kann man Vielfache der n -ten Zeile von den übrigen Zeilen subtrahieren und auf diese Weise erreichen, dass alle Elemente in der n -ten Spalte oberhalb von β_n zu Null werden. Entsprechend kann man die übrigen Spalten behandeln, und es folgt, dass sich A im Falle der Invertierbarkeit mittels elementarer Zeilenumformungen in die Einheitsmatrix E überführen lässt. Jede solche Zeilenumformung lässt sich interpretieren als Multiplikation mit einer Elementarmatrix von links, wie wir in Abschnitt 3.2 gesehen haben. Wir finden daher Elementarmatrizen $S_1, \dots, S_r \in K^{n \times n}$ mit

$$S_r \cdot \dots \cdot S_1 \cdot A = E,$$

und Multiplikation mit A^{-1} von rechts ergibt $A^{-1} = S_r \cdot \dots \cdot S_1$. Indem wir

$$A^{-1} = S_r \cdot \dots \cdot S_1 \cdot E$$

schreiben, sehen wir folgendes: Diejenigen elementaren Zeilenumformungen, die A in die Einheitsmatrix E überführen, führen E selbst in die Matrix A^{-1} über!

Das Verfahren zur Invertierung von Matrizen kann man daher wie folgt beschreiben: Man bringe A mittels einer Folge elementarer Zeilenumformungen auf Zeilenstufenform und führe bei jedem Schritt die entsprechende Zeilenumformung auch an der Einheitsmatrix durch. Wenn man dann an der Zeilenstufenform $\text{rg } A = n$ ablesen kann, ist A invertierbar, und man fahre fort, bis man A in die Einheitsmatrix überführt hat. Die aus der Einheitsmatrix durch die entsprechenden Umformungen gewonnene Matrix ist dann die zu A inverse Matrix A^{-1} . Beispielsweise erkennt man auf diese Weise für $\alpha \in K$, $i \neq j$:

$$\begin{aligned}
(E + (\alpha - 1)E_{ii})^{-1} &= E + (\alpha^{-1} - 1)E_{ii} \\
(E + \alpha E_{ij})^{-1} &= E - \alpha E_{ij} \\
(E - E_{ii} - E_{jj} + E_{ij} + E_{ji})^{-1} &= E - E_{ii} - E_{jj} + E_{ij} + E_{ji}
\end{aligned}$$

Als Nebenprodukt dieses Verfahrens können wir noch vermerken:

Satz 8. *Jede invertierbare Matrix $A \in \text{Gl}(n, K)$ ist ein Produkt von Elementarmatrizen des Typs*

$$\begin{aligned}
E + (\alpha - 1)E_{ii}, & \quad \alpha \in K^*, \\
E + E_{ij}, & \quad i \neq j,
\end{aligned}$$

mit $1 \leq i, j \leq n$. Dabei ist $E \in K^{n \times n}$ die Einheitsmatrix und $E_{ij} = (\delta_{\mu i} \delta_{\nu j})_{\mu, \nu}$.

Beweis. Wir haben gerade gesehen, dass A mittels elementarer Zeilenumformungen des Typs I - IV aus Abschnitt 3.2 in die Einheitsmatrix überführt werden kann und dass A^{-1} das Produkt der entsprechenden Elementarmatrizen ist. Nun kann man aber eine elementare Zeilenumformung des Typs III, also Addition des α -fachen der j -ten Zeile zur i -ten Zeile für $\alpha \in K^*$ und gewisse Indizes $i \neq j$, auch als eine Folge von elementaren Zeilenumformungen der Typen I und II interpretieren: Man multipliziere die j -te Zeile mit α , addiere sie zur i -ten Zeile und multipliziere die j -te Zeile anschließend wieder mit α^{-1} . Entsprechend kann man auch eine elementare Zeilenumformung des Typs IV, also das Vertauschen von i -ter und j -ter Zeile für gewisse Indizes $i \neq j$, als eine Folge von Umformungen der Typen I und II interpretieren: Man addiere die j -te Zeile zur i -ten Zeile, multipliziere die j -te Zeile mit -1 , addiere die i -te Zeile zur j -ten Zeile und subtrahiere schließlich die j -te Zeile von der i -ten Zeile. Dabei ist der letzte Schritt eine Umformung vom Typ III, also zerlegbar in eine Folge von Umformungen der Typen I und II.

Indem wir vorstehende Überlegung auf A^{-1} anstelle von A anwenden, sehen wir, dass A wie behauptet ein Produkt von Elementarmatrizen ist, die zu den elementaren Zeilenumformungen der Typen I und II korrespondieren. $\square$

Möchte man zu einer gegebenen Matrix $A \in \text{Gl}(n, K)$ deren Zerlegung in Elementarmatrizen konkret bestimmen, so verfährt man am besten wie folgt: Man überführt A mittels elementarer Zeilenumformungen in die Einheitsmatrix und erhält auf diese Weise eine Zerlegung der Form $A^{-1} = S_r \cdot \dots \cdot S_1$ mit Elementarmatrizen $S_1, \dots, S_r$. Inversenbildung liefert dann $A = S_1^{-1} \cdot \dots \cdot S_r^{-1}$, wobei es sich bei den S_i^{-1} , wie oben beschrieben, wiederum um Elementarmatrizen handelt. Will man sich auf die in Satz 8 genannten Elementarmatrizen beschränken, so sind die S_i^{-1} gegebenenfalls noch weiter zu zerlegen.

Aufgaben

1. Es sei X eine Menge und R ein Ring mit 1. Man zeige, dass die Menge $\text{Abb}(X, R)$ aller Abbildungen $X \rightarrow R$ unter der gewöhnlichen Addition bzw. Multiplikation

R -wertiger Funktionen einen Ring bildet. Man beschreibe die Einheitengruppe dieses Ringes.

2. Man gebe einen K -Vektorraum V an, zu dem es Nichteinheiten $f, g \in \text{End}_K(V)$ mit $f \circ g = \text{id}_V$ gibt.
3. Es sei $f: U \rightarrow U$ ein Endomorphismus eines endlich-dimensionalen K -Vektorraums. Dann ist äquivalent:
 - (i) $\text{rg } f < \dim_K U$.
 - (ii) f ist ein *Links-Nullteiler* in $\text{End}_K(V)$, d. h. es existiert $g \in \text{End}_K(V)$, $g \neq 0$, mit $f \circ g = 0$.
 - (iii) f ist ein *Rechts-Nullteiler* in $\text{End}_K(V)$, d. h. es existiert $g \in \text{End}_K(V)$, $g \neq 0$, mit $g \circ f = 0$.

Insbesondere ist ein Links-Nullteiler in $\text{End}_K(V)$ auch Rechts-Nullteiler und umgekehrt. Entsprechendes gilt im Matrizenring $K^{n \times n}$.

4. Man prüfe folgende Matrizen auf Invertierbarkeit und gebe gegebenenfalls die inverse Matrix an:

$$\begin{pmatrix} 1 & 2 & 3 & 2 & 4 \\ 1 & 1 & 2 & 2 & 4 \\ 1 & 3 & 4 & 3 & 7 \\ 2 & 3 & 5 & 3 & 8 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 1 & 2 & 0 \\ 1 & 0 & 2 & 1 & 0 \\ 1 & 1 & 1 & 2 & 1 \\ 0 & 1 & 2 & 1 & 0 \\ 1 & 1 & 1 & 2 & 0 \end{pmatrix} \in \mathbb{R}^{5 \times 5}$$

5. Falls möglich schreibe man die folgenden Matrizen als Produkt von Elementarmatrizen:

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 3 \\ 1 & 1 & 1 & 4 \\ 1 & 0 & 1 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \in \mathbb{R}^{4 \times 4}$$

6. Für $A \in \text{GL}(n, K)$ zeige man $(A^{-1})^t = (A^t)^{-1}$.
7. Für $m, r, s \in \mathbb{N} - \{0\}$ betrachte man Matrizen

$$A \in K^{m \times m}, \quad B \in K^{m \times s}, \quad C \in K^{r \times m}, \quad D \in K^{r \times s}$$

mit

$$\text{rg} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = m,$$

wobei A invertierbar sei. Man zeige $D = C \cdot A^{-1} \cdot B$.

3.4 Basiswechsel

Wir haben gesehen, dass man K -lineare Abbildungen zwischen Vektorräumen bei Fixierung von Basen durch Matrizen beschreiben kann. In diesem Abschnitt soll insbesondere untersucht werden, wie sich die beschreibende Matrix ändert, wenn man zu anderen Basen übergeht. Wir beginnen mit einer Charakterisierung von Basen.

Bemerkung 1. Es sei V ein K -Vektorraum mit einer Basis $X = (x_1, \dots, x_n)$. Zu n Elementen $y_j = \sum_{i=1}^n \alpha_{ij} x_i \in V$, $j = 1, \dots, n$, mit Koeffizienten $\alpha_{ij} \in K$ betrachte man die Matrix $A = (\alpha_{ij})_{i,j} \in K^{n \times n}$, also die Matrix $A_{f,X,X}$ der durch $x_j \mapsto y_j$ erklärten K -linearen Abbildung $f: V \rightarrow V$. Dann ist äquivalent:

- (i) $Y = (y_1, \dots, y_n)$ ist eine Basis von V .
- (ii) $\operatorname{rg} A = n$.
- (iii) $A \in \operatorname{Gl}(n, K)$, d. h. A ist invertierbar.

Beweis. Ist Y eine Basis von V , so ist f ein Isomorphismus, und es folgt mit 3.2/2 $\operatorname{rg} A = \operatorname{rg} f = \dim_K V = n$. Gilt $\operatorname{rg} A = n$, so hat man $A \in \operatorname{Gl}(n, K)$ nach 3.3/7. Aus letzter Bedingung wiederum ergibt sich mit 3.3/6, dass f ein Isomorphismus ist, und folglich, dass Y eine Basis ist. $\square$

Ist Y in der Situation von Bemerkung 1 eine Basis von V , so bezeichnen wir $A = (\alpha_{ij})_{i,j}$ als *Matrix eines Basiswechsels*. Man kann dann, und dies werden wir zur Präzisierung der Art des Basiswechsels stets so handhaben, A auch als Matrix zur identischen Abbildung $\operatorname{id}: V \rightarrow V$ bezüglich der Basen Y und X interpretieren, also $A = A_{\operatorname{id},Y,X}$. Insbesondere ergibt sich aus Bemerkung 1, dass sich bei fixierter Basis X von V jede Matrix $A \in \operatorname{Gl}(n, K)$ als Basiswechselmatrix der Form $A_{\operatorname{id},Y,X}$ mit einer geeigneten Basis Y von V interpretieren lässt. Aus der Gleichung

$$A_{\operatorname{id},Y,X} \cdot A_{\operatorname{id},X,Y} = A_{\operatorname{id},X,X} = 1,$$

vgl. 3.1/9, lesen wir ab:

Bemerkung 2. Es sei V ein K -Vektorraum mit endlichen Basen X und Y . Dann ist die Basiswechselmatrix $A_{\operatorname{id},Y,X}$ invertierbar, und ihr Inverses wird gegeben durch die Basiswechselmatrix $A_{\operatorname{id},X,Y}$.

Weiter folgt aus 3.1/7:

Bemerkung 3. Es sei V ein K -Vektorraum mit endlichen Basen X und Y . Für $a \in V$ bezeichne a_X bzw. a_Y den Koordinatenspaltenvektor von a bezüglich X bzw. Y . Dann gilt

$$a_X = A_{\operatorname{id},Y,X} \cdot a_Y, \quad a_Y = A_{\operatorname{id},X,Y} \cdot a_X.$$

Als Nächstes soll untersucht werden, wie sich die beschreibende Matrix einer linearen Abbildung unter Basiswechsel verhält.

Satz 4. Es sei $f: V \rightarrow W$ eine K -lineare Abbildung zwischen endlich-dimensionalen K -Vektorräumen. Sind dann X, X' Basen von V und Y, Y' Basen von W , so gilt

$$A_{f,X',Y'} = (A_{\operatorname{id},Y',Y})^{-1} \cdot A_{f,X,Y} \cdot A_{\operatorname{id},X',X}.$$

Beweis. Indem man 3.1/9 auf $f = \text{id}_W \circ f \circ \text{id}_V$ anwendet, erhält man

$$A_{f,X',Y'} = A_{\text{id},Y,Y'} \cdot A_{f,X,Y} \cdot A_{\text{id},X',X},$$

mit $A_{\text{id},Y,Y'} = (A_{\text{id},Y',Y})^{-1}$. □

Korollar 5. *Es sei $f: V \longrightarrow V$ ein Endomorphismus eines endlich-dimensionalen K -Vektorraums V mit Basen X, X' . Dann gilt*

$$A_{f,X',X'} = (A_{\text{id},X',X})^{-1} \cdot A_{f,X,X} \cdot A_{\text{id},X',X}.$$

Wir wollen nun noch einige Anwendungen zum Rang einer Matrix geben.

Satz 6. *Für Matrizen $S \in \text{Gl}(m, K)$, $A \in K^{m \times n}$ und $T \in \text{Gl}(n, K)$ gilt*

$$\text{rg}(S \cdot A \cdot T) = \text{rg } A.$$

Beweis. Man betrachte die lineare Abbildung $f: K^n \longrightarrow K^m$, $a \longmapsto A \cdot a$. Dann ist A die zu f gehörige Matrix, wenn man in K^n und K^m jeweils die kanonische Basis zugrunde legt. Interpretiert man dann S^{-1} als Matrix eines Basiswechsels in K^m und T als Matrix eines Basiswechsels in K^n , so ergibt sich

$$\text{rg}(S \cdot A \cdot T) = \text{rg } f = \text{rg } A$$

mit 3.2/2 und Satz 4. □

Lemma 7. *Es sei $f: V \longrightarrow W$ eine K -lineare Abbildung zwischen Vektorräumen mit $\dim_K V = n$ und $\dim_K W = m$. Sei $r = \text{rg } f$. Dann existiert eine Basis $X = (x_1, \dots, x_n)$ von V sowie eine Basis $Y = (y_1, \dots, y_m)$ von W mit*

$$A_{f,X,Y} = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix},$$

wobei E_r die $(r \times r)$ -Einheitsmatrix ist und 0 jeweils geeignete (möglicherweise auch leere) Bereiche mit Koeffizienteneinträgen 0 bezeichnet.

Beweis. Man wähle eine Basis $y_1, \dots, y_r$ von $\text{im } f$ und ergänze diese zu einer Basis $Y = (y_1, \dots, y_m)$ von W . Seien weiter $x_1, \dots, x_r$ f -Urbilder zu $y_1, \dots, y_r$. Diese sind linear unabhängig, vgl. 2.1/5, und lassen sich, wie im Beweis zu 2.1/10 gezeigt, durch Elemente $x_{r+1}, \dots, x_n \in \ker f$ zu einer Basis X von V ergänzen. Sodann gilt

$$f(x_j) = \begin{cases} y_j & \text{für } j = 1, \dots, r, \\ 0 & \text{für } j = r+1, \dots, n, \end{cases}$$

d. h. $A_{f,X,Y}$ ist von der gewünschten Form. □

Satz 8. *Es sei $A \in K^{m \times n}$ mit $\operatorname{rg} A = r$. Dann existieren invertierbare Matrizen $S \in \operatorname{Gl}(m, K)$ und $T \in \operatorname{Gl}(n, K)$ mit*

$$S \cdot A \cdot T = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix},$$

wobei E_r die $(r \times r)$ -Einheitsmatrix ist und 0 jeweils geeignete (möglicherweise auch leere) Bereiche mit Koeffizienteneinträgen 0 bezeichnet.

Beweis. Wir betrachten die lineare Abbildung

$$f: K^n \longrightarrow K^m, \quad a \longmapsto A \cdot a,$$

deren Matrix bezüglich der kanonischen Basen in K^n und K^m durch A gegeben wird. Nach Lemma 7 gibt es dann Basen X von K^n und Y von K^m , so dass $A_{f,X,Y}$ von der behaupteten Gestalt ist. Nach Satz 4 gilt $A_{f,X,Y} = S \cdot A \cdot T$ mit Basiswechselmatrizen, also invertierbaren Matrizen S und T . $\square$

Abschließend wollen wir noch andeuten, wie man die Überlegungen dieses Abschnitts dazu benutzen kann, um zu zeigen, dass bei einer Matrix stets der Spalten- mit dem Zeilenrang übereinstimmt. Wir hatten dieses Resultat bereits in 3.2/10 hergeleitet, und zwar unter Verwendung der Theorie dualer Abbildungen. Sei also $A \in K^{m \times n}$. Dann können wir A als Matrix $A_{f,X,Y}$ zu einer K -linearen Abbildung $f: V \longrightarrow W$ bezüglich einer Basis X von V bzw. Y von W interpretieren, wobei sich der Spaltenrang von A nach 3.2/2 zu $\operatorname{rg}_s A = \operatorname{rg} f$ berechnet. Invertierbare Matrizen $S \in \operatorname{Gl}(m, K)$, $T \in \operatorname{Gl}(n, K)$ kann man als Basiswechselmatrizen auffassen, so dass man $\operatorname{rg}_s(S \cdot A \cdot T) = \operatorname{rg} f$ unter Benutzung von Satz 4 erhält. Mit dieser Überlegung ergibt sich

$$\operatorname{rg}_s(S \cdot A \cdot T) = \operatorname{rg}_s A$$

für beliebige Matrizen $S \in \operatorname{Gl}(m, K)$, $T \in \operatorname{Gl}(n, K)$. Diese Gleichung ist aber auch für den Zeilenrang richtig, denn es gilt

$$\operatorname{rg}_z(S \cdot A \cdot T) = \operatorname{rg}_s(S \cdot A \cdot T)^t = \operatorname{rg}_s(T^t \cdot A^t \cdot S^t) = \operatorname{rg}_s A^t = \operatorname{rg}_z A.$$

Man benötigt hierfür außer den Resultaten 3.2/7 und 3.2/8 lediglich, dass mit $S \in \operatorname{Gl}(m, K)$ auch S^t invertierbar ist, entsprechend für T , was aber unmittelbar klar ist. Wählt man nun S und T so, dass $S \cdot A \cdot T$ die in Satz 8 angegebene einfache Gestalt besitzt, so gilt natürlich $\operatorname{rg}_s(S \cdot A \cdot T) = \operatorname{rg}_z(S \cdot A \cdot T)$, und es folgt die gewünschte Beziehung

$$\operatorname{rg}_s A = \operatorname{rg}_s(S \cdot A \cdot T) = \operatorname{rg}_z(S \cdot A \cdot T) = \operatorname{rg}_z A.$$

Aufgaben

1. Man zeige, dass die folgenden Systeme von Vektoren

$$X = \{(1, 1, 1, 2, 0), (1, 0, 2, 1, 0), (1, 1, 1, 2, 1), (1, 1, 2, 1, 0), (0, 1, 1, 2, 0)\},$$

$$Y = \{(1, 2, 3, 4, 4), (1, 1, 2, 3, 4), (1, 3, 4, 6, 7), (2, 3, 5, 6, 8), (0, 0, 1, 1, 1)\}$$

jeweils eine Basis von $\mathbb{R}^5$ bilden. Wie lauten die Basiswechselmatrizen $A_{\text{id}, X, Y}$ und $A_{\text{id}, Y, X}$?

2. Es sei V ein K -Vektorraum endlicher Dimension $n > 0$. Man beschreibe alle Basiswechselmatrizen $A \in K^{n \times n}$, welche eine gegebene Basis X von V , abgesehen von der Reihenfolge der Basisvektoren, wieder in sich selbst überführen.
3. Es sei V ein K -Vektorraum endlicher Dimension $n > 0$. Für gegebene Matrizen $A, B \in K^{n \times n}$ beweise man die Äquivalenz folgender Bedingungen:
- (i) Es existiert eine Matrix $S \in \text{Gl}(n, K)$ mit $B = S^{-1}AS$.
 - (ii) Es existieren $f \in \text{End}_K(V)$ und Basen X, Y von V mit $A_{f, X, X} = A$ und $A_{f, Y, Y} = B$.
4. Für Matrizen $A, B \in K^{m \times n}$ schreibe man $A \sim B$, falls es $S \in \text{Gl}(m, K)$ und $T \in \text{Gl}(n, K)$ gibt mit $B = SAT$. Man zeige, dass die Relation “ $\sim$ ” eine Äquivalenzrelation ist, beschreibe die zugehörigen Äquivalenzklassen und gebe insbesondere deren Anzahl an.
5. Es sei $V \neq 0$ ein endlich-dimensionaler K -Vektorraum mit den Basen X, Y und V^* sein Dualraum mit den dualen Basen X^*, Y^* . Für die Basiswechselmatrizen $A = A_{\text{id}, X, Y}$ und A_{id, X^*, Y^*} gilt dann $A_{\text{id}, X^*, Y^*} = (A^{-1})^t$.

3.5 Lineare Gleichungssysteme

Für eine Matrix $A = (\alpha_{ij})_{i,j} \in K^{m \times n}$ und einen Vektor $b = (b_1, \dots, b_m)^t \in K^m$, den wir als Spaltenvektor auffassen wollen, nennt man

$$\begin{array}{ccccccc} \alpha_{11}x_1 & + & \dots & + & \alpha_{1n}x_n & = & b_1 \\ \alpha_{21}x_1 & + & \dots & + & \alpha_{2n}x_n & = & b_2 \\ & & \dots & & & & \\ \alpha_{m1}x_1 & + & \dots & + & \alpha_{mn}x_n & = & b_m \end{array}$$

oder in Matrizenschreibweise

$$A \cdot x = b$$

ein *lineares Gleichungssystem* mit Koeffizienten $\alpha_{ij} \in K$ und den “Unbekannten” $x_1, \dots, x_n$, bzw. $x = (x_1, \dots, x_n)^t$. Genauer versteht man hierunter das Problem, alle $x \in K^n$ zu bestimmen, die die Gleichung $A \cdot x = b$ erfüllen. Im Falle $b = 0$ heißt das Gleichungssystem *homogen*, ansonsten *inhomogen*. Wir wollen eine spezielle Bezeichnung für den Raum der Lösungen eines linearen Gleichungssystems einführen.

Definition 1. Für eine Matrix $A \in K^{m \times n}$ und einen Spaltenvektor $b \in K^m$ bezeichnen wir die Menge

$$M_{A,b} = \{x \in K^n; A \cdot x = b\}$$

als den Lösungsraum des linearen Gleichungssystems $A \cdot x = b$.

Wir können sofort eine triviale, aber sehr wichtige Feststellung treffen, die Informationen über die Struktur solcher Lösungsräume liefert:

Bemerkung 2. Zu einem linearen Gleichungssystem $A \cdot x = b$ mit $A \in K^{m \times n}$, $b \in K^m$ betrachte man die K -lineare Abbildung $f: K^n \longrightarrow K^m$, $a \longmapsto A \cdot a$. Dann gilt

$$M_{A,b} = f^{-1}(b).$$

Der Lösungsraum des Gleichungssystems ist daher ein affiner Unterraum von K^n ; vgl. 2.2/11.

Für $b = 0$ folgt insbesondere $M_{A,0} = \ker f$. In diesem Falle ist der Lösungsraum sogar ein linearer Unterraum von K^n .

Wir wollen zunächst homogene lineare Gleichungssysteme, also lineare Gleichungssysteme des Typs $A \cdot x = 0$ genauer studieren. Der Lösungsraum $M_{A,0}$ ist dann ein linearer Unterraum von K^n , enthält stets die triviale Lösung $0 \in K^n$ und ist folglich nicht leer.

Satz 3. Für $A \in K^{m \times n}$ ist der Lösungsraum $M_{A,0}$ des homogenen linearen Gleichungssystems $A \cdot x = 0$ ein linearer Unterraum von K^n mit Dimension $\dim_K(M_{A,0}) = n - \operatorname{rg} A$.

Beweis. Die lineare Abbildung $f: K^n \longrightarrow K^m$, $a \longmapsto A \cdot a$, besitzt $M_{A,0}$ als Kern, und die Dimensionsformel 2.1/10 liefert $n = \dim_K(\ker f) + \operatorname{rg} f$, also $\dim_K(M_{A,0}) = n - \operatorname{rg} A$. $\square$

Lemma 4. Für $A \in K^{m \times n}$ und $S \in \operatorname{Gl}(m, K)$ haben die linearen Gleichungssysteme $A \cdot x = 0$ und $(S \cdot A) \cdot x = 0$ dieselben Lösungen, d. h. $M_{A,0} = M_{SA,0}$.

Beweis. Gilt $x \in M_{A,0}$, also $A \cdot x = 0$, so folgt mittels Multiplikation mit S von links $S \cdot A \cdot x = 0$, also $x \in M_{SA,0}$. Umgekehrt, hat man $x \in M_{SA,0}$, also $S \cdot A \cdot x = 0$, so ergibt sich durch Multiplikation mit S^{-1} von links $A \cdot x = 0$, also $x \in M_{A,0}$. $\square$

Die Aussage des Lemmas ist von besonderem Nutzen, wenn man ein konkret gegebenes homogenes lineares Gleichungssystem der Form $A \cdot x = 0$ explizit lösen möchte. Man kann nämlich das Gaußsche Eliminationsverfahren anwenden

und A gemäß 3.2/4 mittels elementarer Zeilenumformungen auf Zeilenstufenform bringen. Da solche Umformungen auch als Multiplikation von links mit Elementarmatrizen, also invertierbaren Matrizen, interpretiert werden können, ändert sich der Lösungsraum des betrachteten homogenen linearen Gleichungssystems dabei nicht. Man darf daher ohne Beschränkung der Allgemeinheit annehmen, dass A Zeilenstufenform besitzt. Um unsere Bezeichnungen übersichtlich zu gestalten, wollen wir weiter annehmen, dass die Elemente von A , bei denen sozusagen die einzelnen Stufen beginnen, auf der Hauptdiagonalen von A stehen, dass also A von der Gestalt

$$A = \begin{pmatrix} \alpha_{11} & \dots & \dots & \dots & \alpha_{1n} \\ & \alpha_{22} & \dots & \dots & \alpha_{2n} \\ & & \dots & & \\ & & & \dots & \\ & 0 & & \alpha_{rr} & \dots & \alpha_{rn} \\ & & & 0 & \dots & 0 \\ & & & & \dots & \\ & & & 0 & \dots & 0 \end{pmatrix}$$

mit Koeffizienten $\alpha_{11}, \dots, \alpha_{rr} \in K^*$ ist, wobei $r = \text{rg } A$ gilt. Eine solche Zeilenstufenform kann man aus einer Zeilenstufenform allgemeinen Typs durch Vertauschen von Spalten herstellen, ein Prozess, der sich beispielsweise durch Umnummerieren der Unbekannten $x_1, \dots, x_n$ realisieren lässt. Im Übrigen kann man durch Ausführen weiterer elementarer Zeilenumformungen stets $\alpha_{ii} = 1$ für $i = 1, \dots, r$ erreichen und außerdem, dass alle Elemente in der i -ten Spalte oberhalb von α_{ii} verschwinden. Die Matrix A kann also von der Form

$$\begin{pmatrix} 1 & & & \alpha_{1,r+1} & \dots & \alpha_{1,n} \\ & 1 & & \alpha_{2,r+1} & \dots & \alpha_{2,n} \\ & & \ddots & \ddots & & \ddots \\ & & & \ddots & & \ddots \\ & & & & 1 & \alpha_{r,r+1} & \dots & \alpha_{r,n} \\ 0 & \dots & \dots & \dots & \dots & \dots & 0 \\ 0 & \dots & \dots & \dots & \dots & \dots & 0 \\ \dots & & & & & & \dots \\ 0 & \dots & \dots & \dots & \dots & \dots & 0 \end{pmatrix}$$

angenommen werden, wobei sich in der linken oberen Ecke die $(r \times r)$ -Einheitsmatrix befindet. Somit ist folgendes Gleichungssystem zu lösen:

$$x_i + \sum_{j=r+1}^n \alpha_{ij} x_j = 0, \quad i = 1, \dots, r$$

Dies ist ohne Aufwand möglich, da man die Werte $x_{r+1}, \dots, x_n \in K$ beliebig vorgeben darf und sich die Werte von $x_1, \dots, x_r$ hieraus zu

$$x_i = - \sum_{j=r+1}^n \alpha_{ij} x_j, \quad i = 1, \dots, r,$$

bestimmen.

Die Arbeit beim Lösen des linearen Gleichungssystems $A \cdot x = 0$ reduziert sich damit auf das Herstellen der oben angegebenen speziellen Zeilenstufenform von A . Insbesondere wird deutlich, warum dieses nach Gauß benannte Verfahren als *Eliminationsverfahren* bezeichnet wird. Aus der ersten Gleichung ergibt sich x_1 in Abhängigkeit von $x_{r+1}, \dots, x_n$, aus der zweiten x_2 in Abhängigkeit von $x_{r+1}, \dots, x_n$ usw. Es werden also nach und nach unbekannte Größen eliminiert, bis man zu einem Restsystem von Größen gelangt, deren Werte frei wählbar sind. Dies äußert sich darin, dass die Projektion

$$K^n \longrightarrow K^{n-r}, \quad (a_1, \dots, a_n)^t \longmapsto (a_{r+1}, \dots, a_n)^t,$$

einen Isomorphismus $M_{A,0} \xrightarrow{\sim} K^{n-r}$ induziert. Insbesondere sehen wir nochmals $\dim_K M_{A,0} = n - r$ ein, und es wird klar, dass man durch Liften einer Basis von K^{n-r} , etwa der kanonischen, eine Basis von $M_{A,0}$ erhält. In obiger Notation besteht diese dann aus den Vektoren

$$v_j = (-\alpha_{1j}, \dots, -\alpha_{rj}, \delta_{r+1j}, \dots, \delta_{nj})^t, \quad j = r+1, \dots, n.$$

Das Verfahren zur Bestimmung einer Basis des Lösungsraums $M_{A,0}$ eines homogenen linearen Gleichungssystems $A \cdot x = 0$ gestaltet sich daher wie folgt:

Man transformiere A auf die spezielle oben beschriebene Zeilenstufenform. Für $j = r+1, \dots, n$ sei $v_j \in K^n$ derjenige Vektor, dessen Komponenten mit Index $i = 1, \dots, r$ jeweils aus dem Negativen der entsprechenden Komponenten der j -ten Spalte der Zeilenstufenform von A bestehen und dessen Komponenten mit Index $i = r+1, \dots, n$ gerade diejenigen des $(j-r)$ -ten Einheitsvektors aus K^{n-r} seien. Dann bilden $v_{r+1}, \dots, v_n$ eine Basis von $M_{A,0}$.

Im Prinzip behält diese Regel auch dann ihre Gültigkeit, wenn wir bei der Herstellung der Zeilenstufenform von A auf Spaltenvertauschungen und damit auf ein Ummummerieren der Unbekannten verzichten. Die Rolle der Indizes $i = 1, \dots, r$ (bzw. $i = r+1, \dots, n$) wird, was die Spaltenindizes der Zeilenstufenform von A wie auch die Indizes der Komponenten der v_j angeht, in diesem Falle von denjenigen Spaltenindizes übernommen, bei denen die Zeilenstufenform “springt” (bzw. nicht “springt”). Hiermit meinen wir diejenigen Spalten, die im Sinne der Notation von 3.2/4 eines (bzw. keines) der dort positionierten Elemente $\beta_1, \dots, \beta_r \in K^*$ enthalten. Um die Lösungen auch in diesem Falle formelmäßig zu beschreiben, gehen wir von der entsprechenden speziellen Zeilenstufenform von A aus, die nunmehr die Gestalt

$$\begin{array}{ccccccc}
1 & j_1 & j_2 & \longleftarrow \text{Spaltenindex} \longrightarrow & j_r & n \\
\left(\begin{array}{cccc|cccc}
0 & \dots & 0 & 1 & * & \dots & * & 0 & * & \dots & * \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & * & \dots & * \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\
\dots & & & & & & & & & & \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\
\dots & & & & & & & & & & \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0
\end{array} \right. & \dots & & \left. \begin{array}{cccc}
* & \dots & * & 0 & * & \dots & * \\
* & \dots & * & 0 & * & \dots & * \\
* & \dots & * & 0 & * & \dots & * \\
\dots & & & & & & \\
0 & \dots & 0 & 1 & * & \dots & * \\
0 & \dots & 0 & 0 & 0 & \dots & 0
\end{array} \right)
\end{array}$$

besitzt. Bezeichnet dann J' das Komplement der Menge der Spaltenindizes, bei denen die Zeilenstufenform A “springt”, also $J' = \{1, \dots, n\} - \{j_1, \dots, j_r\}$, so ist das lineare Gleichungssystem

$$x_{j_i} + \sum_{j' \in J'} \alpha_{ij'} x_{j'} = 0, \quad i = 1, \dots, r,$$

zu lösen. Wiederum kann man die Werte der $x_{j'} \in K$ für $j' \in J'$ beliebig vorgeben und die Werte der x_{j_i} für $i = 1, \dots, r$ daraus berechnen. Die Projektion

$$K^n \longrightarrow K^{n-r}, \quad (a_1, \dots, a_n)^t \longmapsto (a_{j'})_{j' \in J'}^t,$$

liefert daher einen Isomorphismus $M_{A,0} \xrightarrow{\sim} K^{n-r}$, und durch Liften der kanonischen Basis von K^{n-r} erhält man eine Basis von $M_{A,0}$ bestehend aus den Vektoren

$$v_{j'} = (\xi_{1j'}, \dots, \xi_{nj'})^t, \quad j' \in J',$$

mit

$$\xi_{i'j'} = \begin{cases} -\alpha_{ij'} & \text{für } i' = j_i \text{ mit } i \in \{1, \dots, r\}, \\ \delta_{i'j'} & \text{für } i' \in J' \end{cases}.$$

Wir wollen ein Beispiel betrachten. Sei $K = \mathbb{R}$, $m = 3$, $n = 4$ und

$$A = \begin{pmatrix} 0 & 0 & 1 & 2 \\ 1 & 2 & 1 & 3 \\ 1 & 2 & 2 & 5 \end{pmatrix}.$$

Das lineare Gleichungssystem $A \cdot x = 0$ schreibt sich dann ausführlich in der Form:

$$\begin{aligned}
x_3 + 2x_4 &= 0 \\
x_1 + 2x_2 + x_3 + 3x_4 &= 0 \\
x_1 + 2x_2 + 2x_3 + 5x_4 &= 0
\end{aligned}$$

Bringen wir nun A mittels elementarer Zeilenumformungen auf die spezielle Zeilenstufenform, also

$$A \mapsto \begin{pmatrix} 1 & 2 & 1 & 3 \\ 0 & 0 & 1 & 2 \\ 1 & 2 & 2 & 5 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 2 & 1 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 1 & 2 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

so “springt” diese Stufenform genau bei den Spaltenindizes 1 und 3. Wir lesen daher nach der oben beschriebenen Regel als Basis des Lösungsraumes $M_{A,0}$ ab:

$$v_2 = (-2, 1, 0, 0)^t, \quad v_4 = (-1, 0, -2, 1)^t$$

Argumentieren wir etwas ausführlicher, so bleibt das lineare Gleichungssystem

$$\begin{aligned} x_1 + 2x_2 + x_4 &= 0, \\ x_3 + 2x_4 &= 0 \end{aligned}$$

zu lösen. Die Projektion

$$K^4 \longrightarrow K^2, \quad (a_1, \dots, a_4)^t \longmapsto (a_2, a_4)^t,$$

liefert einen Isomorphismus $M_{A,0} \xrightarrow{\sim} K^2$, und wir liften die kanonische Basis von K^2 zu einer Basis von $M_{A,0}$:

$$\begin{aligned} x_2 = 1, \quad x_4 = 0 &\implies x_1 = -2, \quad x_3 = 0 \\ x_2 = 0, \quad x_4 = 1 &\implies x_1 = -1, \quad x_3 = -2 \end{aligned}$$

Insbesondere gilt $\dim_K M_{A,0} = 2$, und wir erkennen, wie bereits oben angegeben,

$$M_{A,0} = \langle (-2, 1, 0, 0)^t, (-1, 0, -2, 1)^t \rangle.$$

Als Nächstes wollen wir den Allgemeinfall behandeln, also inhomogene lineare Gleichungssysteme des Typs $A \cdot x = b$, wobei der Fall $b = 0$ nicht explizit ausgeschlossen werden soll. Im Folgenden bezeichnen wir mit $(A, b) \in K^{m \times (n+1)}$ diejenige Matrix, die aus A durch Hinzufügen von b als $(n+1)$ -ter Spalte entsteht.

Satz 5. Zu $A \in K^{m \times n}$ und $b \in K^m$ betrachte man das lineare Gleichungssystem $A \cdot x = b$.

(i) $A \cdot x = b$ ist genau dann lösbar (d. h. besitzt mindestens eine Lösung), wenn $\operatorname{rg} A = \operatorname{rg}(A, b)$ gilt.

(ii) $A \cdot x = b$ ist genau dann universell lösbar (d. h. besitzt für jedes $b \in K^m$ mindestens eine Lösung), wenn $\operatorname{rg} A = m$ gilt.

(iii) $A \cdot x = b$ besitzt genau dann für alle $b \in K^m$ höchstens eine Lösung, wenn $\operatorname{rg} A = n$ gilt.

Beweis. Man betrachte die lineare Abbildung $f: K^n \longrightarrow K^m$, $a \longmapsto A \cdot a$. Es gilt $M_{A,b} = f^{-1}(b)$ für den Lösungsraum zu $A \cdot x = b$; vgl. Bemerkung 2. Somit ist $M_{A,b}$ genau dann nicht leer, wenn b zum Bild von f gehört. Sind $a_1, \dots, a_n \in K^m$ die Spalten von A , so gilt im $f = \langle a_1, \dots, a_n \rangle$, und es ist $b \in \operatorname{im} f$ äquivalent zu $\langle a_1, \dots, a_n \rangle = \langle a_1, \dots, a_n, b \rangle$. Da aber $\langle a_1, \dots, a_n \rangle$ stets ein linearer Teilraum

von $\langle a_1, \dots, a_n, b \rangle$ ist, kann man mit 1.5/14 (ii) bereits dann auf die Gleichheit beider Räume schließen, wenn ihre Dimensionen übereinstimmen. Die Dimensionen sind aber gerade die Ränge der Matrizen A bzw. (A, b) . Somit sehen wir, dass $M_{A,b} \neq \emptyset$ äquivalent zu $\operatorname{rg} A = \operatorname{rg}(A, b)$ ist, wie in (i) behauptet.

Wegen $M_{A,b} = f^{-1}(b)$ ist $A \cdot x = b$ genau dann universell lösbar, wenn f surjektiv ist, also im $f = K^m$ gilt. Letzteres ist äquivalent zu $\operatorname{rg} f = m$ und somit zu $\operatorname{rg} A = m$, wie in (ii) behauptet.

Die eindeutige Lösbarkeit von $A \cdot x = b$ schließlich, wie in (iii) betrachtet, ist äquivalent zur Injektivität von f . Indem man die Dimensionsformel 2.1/10 für f benutzt, also

$$n = \dim_K(\ker f) + \operatorname{rg} f,$$

sieht man, dass die Injektivität von f äquivalent zu $\operatorname{rg} f = n$ bzw. $\operatorname{rg} A = n$ ist. $\square$

Gilt $m = n$ in der Situation von Satz 5, so ist die universelle Lösbarkeit in Bedingung (ii) äquivalent zu der höchstens eindeutigen Lösbarkeit in Bedingung (iii). Mit 3.3/7 folgt daher:

Korollar 6. *Für eine Matrix $A \in K^{n \times n}$ ist äquivalent:*

- (i) *Das lineare Gleichungssystem $A \cdot x = b$ ist universell für $b \in K^n$ lösbar.*
- (ii) *Das lineare Gleichungssystem $A \cdot x = 0$ besitzt nur die triviale Lösung.*
- (iii) *A ist invertierbar.*

Wir wollen noch etwas genauer auf die Struktur von Lösungsräumen inhomogener linearer Gleichungssysteme eingehen. Das nachfolgende Resultat zeigt dabei nochmals, dass es sich bei solchen Lösungsräumen um affine Unterräume handelt.

Satz 7. *Für $A \in K^{m \times n}$ und $b \in K^m$ habe man eine Lösung $v_0 \in M_{A,b}$ des linearen Gleichungssystems $A \cdot x = b$. Dann gilt $M_{A,b} = v_0 + M_{A,0}$. Mit anderen Worten, die Gesamtheit aller Lösungen des inhomogenen Systems $A \cdot x = b$ erhält man in der Form $v_0 + v$, wobei v_0 eine beliebige, so genannte partikuläre Lösung dieses Systems ist und v alle Lösungen des zugehörigen homogenen Systems $A \cdot x = 0$ durchläuft.*

Beweis. Wir betrachten wieder die durch A gegebene lineare Abbildung

$$f: K^n \longrightarrow K^m, \quad x \longmapsto A \cdot x,$$

wobei $M_{A,b} = f^{-1}(b)$ gilt. Für $v_0 \in M_{A,b}$ folgt dann mit 2.2/2

$$M_{A,b} = f^{-1}(f(v_0)) = v_0 + \ker f = v_0 + M_{A,0},$$

wie behauptet. $\square$

Wir wollen nun noch zeigen, wie man mit Hilfe des Gaußschen Eliminationsverfahrens auch inhomogene lineare Gleichungssysteme lösen kann. Zunächst

eine nützliche Beobachtung, die als Verallgemeinerung von Lemma 4 zu sehen ist:

Lemma 8. Für $A \in K^{m \times n}$, $b \in K^m$ und $S \in \text{Gl}(m, K)$ haben die linearen Gleichungssysteme $A \cdot x = b$ und $(S \cdot A) \cdot x = S \cdot b$ dieselben Lösungen, d. h. $M_{A,b} = M_{SA, Sb}$.

Beweis. Indem man mit S bzw. S^{-1} von links multipliziert, sieht man, dass $A \cdot x = b$ für $x \in K^n$ äquivalent zu $S \cdot A \cdot x = S \cdot b$ ist. $\square$

Man darf also zur Lösung eines linearen Gleichungssystems $A \cdot x = b$ die Matrix A mittels elementarer Zeilenumformungen abändern, wenn man gleichzeitig diese Umformungen auch bei b , aufgefasst als $(m \times 1)$ -Matrix, durchführt; solche Umformungen lassen sich nämlich als Multiplikation von links mit Elementarmatrizen, also invertierbaren Matrizen auffassen. Am einfachsten ist es dann, A und b zu der Matrix (A, b) zusammenzufügen und diese Matrix mittels elementarer Zeilenumformungen zu verändern. Solche Umformungen wirken separat auf die einzelnen Spalten, insbesondere also separat auf die Matrix A sowie die Spalte b , d. h. es gilt $S \cdot (A, b) = (SA, Sb)$ für $S \in \text{Gl}(m, K)$.

Wie im homogenen Fall transformiere man nun (A, b) zunächst auf Zeilenstufenform. Nach eventueller Umnummerierung der Unbekannten $x_1, \dots, x_n$ können wir annehmen, dass die transformierte Matrix die Gestalt

$$\left(\begin{array}{cccccc} 1 & & & \alpha_{1,r+1} & \dots & \alpha_{1,n} & \beta_1 \\ & 1 & & \alpha_{2,r+1} & \dots & \alpha_{2,n} & \beta_2 \\ & & \ddots & \ddots & & \ddots & \ddots \\ & & & \ddots & & \ddots & \ddots \\ & & & & 1 & \alpha_{r,r+1} & \dots & \alpha_{r,n} & \beta_r \\ 0 & \dots & \dots & \dots & \dots & \dots & 0 & \beta_{r+1} \\ 0 & \dots & \dots & \dots & \dots & \dots & 0 & 0 \\ \ddots & & & & & & \ddots & \ddots \\ 0 & \dots & \dots & \dots & \dots & \dots & 0 & 0 \end{array} \right)$$

SA

Sb

besitzt, wobei sich in der linken oberen Ecke die $(r \times r)$ -Einheitsmatrix befindet und S das Produkt der benötigten Elementarmatrizen andeutet. Wie im homogenen Fall folgt $r = \text{rg } A$, und man sieht, dass die Bedingung $\text{rg } A = \text{rg}(A, b)$ bzw. $\text{rg}(SA) = \text{rg}(SA, Sb)$ äquivalent zu $\beta_{r+1} = 0$ ist. Das System $A \cdot x = b$ ist also genau dann lösbar, wenn in obiger Zeilenstufenform $\beta_{r+1} = 0$ gilt. Ist Letzteres der Fall, so ergibt sich

$$x_i + \sum_{j=r+1}^n \alpha_{ij} x_j = \beta_i, \quad i = 1, \dots, r,$$

als transformiertes Gleichungssystem. Indem man $x_{r+1}, \dots, x_n = 0$ setzt, gelangt man zu einer partikulären Lösung v_0 mit den Komponenten

$$\xi_j = \begin{cases} \beta_j & \text{für } j = 1, \dots, r \\ 0 & \text{für } j = r+1, \dots, n \end{cases}.$$

Bestimmt man nun auch noch, wie bereits vorgeführt, den Lösungsraum $M_{A,0}$ des homogenen Systems $A \cdot x = 0$, so ergibt sich $M_{A,b} = v_0 + M_{A,0}$ gemäß Satz 7. Insgesamt lässt sich feststellen, dass sich die Lösung v_0 , wie auch eine Basis $v_{r+1}, \dots, v_n$ des Lösungsraumes $M_{A,0}$ in direkter Weise aus der hergeleiteten speziellen Zeilenstufenform der Ausgangsmatrix (A, b) ablesen lassen. Ähnlich wie im Falle homogener linearer Gleichungssysteme gilt dies auch dann, wenn man keine Umnummerierung der Unbekannten zulässt und stattdessen die Indizes $i = 1, \dots, r$ durch diejenigen Spaltenindizes $j_1, \dots, j_r$ ersetzt, bei denen die Zeilenstufenform von A “springt”. Es ist dann das Gleichungssystem

$$x_{j_i} + \sum_{j' \in J'} \alpha_{ij'} x_{j'} = \beta_i, \quad i = 1, \dots, r,$$

mit $J' = \{1, \dots, n\} - \{j_1, \dots, j_r\}$ zu lösen. Eine partikuläre Lösung $v_0 \in M_{A,b}$ wird in diesem Falle durch den Vektor $v_0 = (\xi_1, \dots, \xi_n)^t \in K^n$ mit den Komponenten

$$\xi_j = \begin{cases} \beta_i & \text{für } j = j_i \text{ mit } i \in \{1, \dots, r\} \\ 0 & \text{sonst} \end{cases}$$

gegeben.

Als Beispiel wollen wir für $K = \mathbb{R}$ das System $A \cdot x = b$ lösen mit

$$A = \begin{pmatrix} 0 & 0 & 1 & 2 \\ 1 & 2 & 1 & 3 \\ 1 & 2 & 2 & 5 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}.$$

Den Lösungsraum $M_{A,0}$ des zugehörigen homogenen Systems hatten wir bereits bestimmt. Wir bringen zunächst die Matrix (A, b) auf spezielle Zeilenstufenform, also

$$(A, b) \mapsto \begin{pmatrix} 1 & 2 & 1 & 3 & 1 \\ 0 & 0 & 1 & 2 & 1 \\ 1 & 2 & 2 & 5 & 3 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 2 & 1 & 3 & 1 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 2 & 2 \end{pmatrix} \mapsto \begin{pmatrix} \boxed{1} & 2 & 0 & 1 & 0 \\ 0 & 0 & \boxed{1} & 2 & 1 \\ 0 & 0 & 0 & 0 & \boxed{1} \end{pmatrix},$$

woraus man $\text{rg}(A, b) = 3 > 2 = \text{rg } A$ entnimmt. Das System $A \cdot x = b$ besitzt daher keine Lösung, d. h. es gilt $M_{A,b} = \emptyset$.

Alternativ wollen wir das Gleichungssystem auch noch für $b = (1, 1, 2)^t$ betrachten. Die Transformation von (A, b) auf spezielle Zeilenstufenform liefert dann

$$(A, b) \mapsto \begin{pmatrix} 1 & 2 & 1 & 3 & 1 \\ 0 & 0 & 1 & 2 & 1 \\ 1 & 2 & 2 & 5 & 2 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 2 & 1 & 3 & 1 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 2 & 1 \end{pmatrix} \mapsto \begin{pmatrix} \boxed{1} & 2 & 0 & 1 & 0 \\ 0 & 0 & \boxed{1} & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

In diesem Fall gilt $\text{rg } A = \text{rg}(A, b) = 2$, und man liest $v_0 = (0, 0, 1, 0)$ als partikuläre Lösung ab.

In ausführlicherer Argumentation ist das System

$$\begin{aligned}x_1 + 2x_2 + x_4 &= 0, \\x_3 + 2x_4 &= 1\end{aligned}$$

zu betrachten. Um eine partikuläre Lösung zu berechnen, setzen wir $x_2 = x_4 = 0$ und erhalten $x_1 = 0$, $x_3 = 1$, also $(0, 0, 1, 0)^t \in M_{A,b}$. Da wir bereits gezeigt haben, dass die Vektoren $(-2, 1, 0, 0)^t$ und $(-1, 0, -2, 1)^t$ eine Basis von $M_{A,0}$ bilden, ergibt sich

$$M_{A,b} = (0, 0, 1, 0)^t + \langle (-2, 1, 0, 0)^t, (-1, 0, -2, 1)^t \rangle.$$

Aufgaben

1. Für eine Matrix und Vektoren

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 2 & 3 & 4 & 5 & 6 \\ 0 & 2 & 2 & 4 & 4 \end{pmatrix} \in \mathbb{R}^{4 \times 5}, \quad b = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \in \mathbb{R}^4, \quad b' = \begin{pmatrix} 1 \\ 1 \\ 1 \\ -1 \end{pmatrix} \in \mathbb{R}^4$$

bestimme man alle Lösungen

- (i) des homogenen linearen Gleichungssystems $A \cdot x = 0$,
- (ii) des inhomogenen linearen Gleichungssystems $A \cdot x = b$,
- (iii) des inhomogenen linearen Gleichungssystems $A \cdot x = b'$.

2. Für reelle Matrizen

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 4 & 9 \\ 1 & 8 & 27 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 4 & 9 & 16 \\ 1 & 8 & 27 & 64 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 4 & 9 & 16 & 25 \\ 1 & 8 & 27 & 64 & 125 \end{pmatrix}$$

und Vektoren $b \in \mathbb{R}^4$ untersuche man die linearen Gleichungssysteme $A \cdot x = b$, $B \cdot x = b$ und $C \cdot x = b$ auf universelle bzw. höchstens eindeutige Lösbarkeit in $x \in \mathbb{R}^3$, bzw. $x \in \mathbb{R}^4$, bzw. $x \in \mathbb{R}^5$. Sind diese Systeme speziell für $b = (1, 0, 0, 0)^t$ lösbar?

- 3. Man zeige, dass jeder affine Unterraum in K^n Lösungsraum eines geeigneten linearen Gleichungssystems mit Koeffizienten aus K ist.
- 4. Es sei $f: V \rightarrow W$ eine K -lineare Abbildung, die bezüglich geeigneter Basen X von V und Y von W durch die Matrix

$$A_{f,X,Y} = \begin{pmatrix} 1 & 1 & 2 & 4 & 8 \\ 1 & 2 & 1 & 2 & 1 \\ 2 & 2 & 2 & 2 & 1 \\ 1 & 2 & 2 & 2 & 2 \end{pmatrix}$$

gegeben sei. Man bestimme eine Basis von $\ker f$.

5. Man betrachte in $\mathbb{R}^6$ die linearen Unterräume

$$U = \langle (1, 1, 1, 0, 1, 1), (2, 3, 4, 0, 2, 1), (0, 3, 2, 1, 1, 1) \rangle,$$

$$U' = \langle (2, 6, 6, 2, 2, 1), (0, 9, 4, 3, 0, 1), (1, 2, 3, 1, 2, 1) \rangle$$

und bestimme $U \cap U'$ durch Angabe einer Basis.

6. Es sei $f: V \rightarrow W$ eine K -lineare Abbildung zwischen (endlich-dimensionalen) K -Vektorräumen und $f^*: W^* \rightarrow V^*$ die zugehörige duale Abbildung. Man zeige für $b \in W$, dass die "lineare Gleichung" $f(x) = b$ genau dann in $x \in V$ lösbar ist, wenn $\varphi(b) = 0$ für alle $\varphi \in \ker f^*$ gilt.
7. Zu einem K -Vektorraum V endlicher Dimension n betrachte man Linearformen $\varphi_1, \dots, \varphi_m \in V^*$ und zeige:
- (i) Es gilt $\dim_K \langle \varphi_1, \dots, \varphi_m \rangle = \operatorname{rg} f$, wobei $f: V \rightarrow K^m$ definiert sei durch $x \mapsto (\varphi_1(x), \dots, \varphi_m(x))$.
 - (ii) Das "lineare Gleichungssystem" $\varphi_i(x) = b_i$, $i = 1, \dots, m$, hat genau dann für alle Wahlen von Elementen $b_1, \dots, b_m \in K$ eine Lösung $x \in V$, wenn die Linearformen $\varphi_1, \dots, \varphi_m$ linear unabhängig sind.
8. Es seien V ein endlich-dimensionaler K -Vektorraum, $\varphi_1, \dots, \varphi_m \in V^*$ Linearformen und $b_1, \dots, b_m \in K$. Man zeige: Das "lineare Gleichungssystem"

$$\varphi_i(x) = b_i, \quad i = 1, \dots, m,$$

ist genau dann in $x \in V$ lösbar, wenn folgende Bedingung erfüllt ist: Sind $\alpha_1, \dots, \alpha_m \in K$ Koeffizienten mit $\sum_{i=1}^m \alpha_i \varphi_i = 0$, so folgt auch $\sum_{i=1}^m \alpha_i b_i = 0$.

4. Determinanten

Vorbemerkungen

Jeder quadratischen Matrix A mit Koeffizienten aus einem Körper K kann man mittels einer gewissen Rechenvorschrift eine Invariante zuordnen, die so genannte *Determinante*. Diese ist genau dann von Null verschieden, wenn die Spalten oder, alternativ, die Zeilen von A linear unabhängig sind, d. h. genau dann, wenn A invertierbar ist.

Um einen Eindruck davon zu bekommen, wie man in mehr oder weniger zwangsläufiger Weise auf die Bildung der Determinante geführt wird, betrachten wir eine Matrix

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} \in K^{2 \times 2}$$

und untersuchen, welche Anforderungen die Koeffizienten α_{ij} erfüllen müssen, damit die Spalten von A linear abhängig werden. Im Falle $A = 0$ sind die Spalten von A natürlich trivialerweise linear abhängig. Es sei deshalb $A \neq 0$, etwa $\alpha_{11} \neq 0$. Dann ist der erste Spaltenvektor von A nicht Null, und die beiden Spalten von A sind genau dann linear abhängig, wenn es ein $c \in K$ mit

$$\alpha_{12} = c\alpha_{11}, \quad \alpha_{22} = c\alpha_{21}$$

gibt. Hieraus folgt

$$c = \frac{\alpha_{12}}{\alpha_{11}}, \quad \alpha_{22} = \frac{\alpha_{12}}{\alpha_{11}}\alpha_{21}$$

und damit die Beziehung

$$(*) \quad \alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12} = 0.$$

Ein Zurückverfolgen der Rechnung zeigt, dass die Spalten von A unter der Bedingung $\alpha_{11} \neq 0$ genau dann linear abhängig sind, wenn der Ausdruck in $(*)$ verschwindet. Man kann nun exakt die gleiche Rechnung durchführen

- für $\alpha_{21} \neq 0$, wenn man die beiden Zeilen von A vertauscht,
- für $\alpha_{12} \neq 0$, wenn man die beiden Spalten von A vertauscht,
- für $\alpha_{22} \neq 0$, wenn man die beiden Spalten und Zeilen von A vertauscht.

Da bei allen diesen Vertauschungen der Ausdruck in $(*)$ bis auf das Vorzeichen unverändert bleibt und da dieser natürlich auch im Falle $A = 0$ verschwindet, können wir aus unserer Rechnung ablesen:

Eine (beliebige) Matrix $A = (\alpha_{ij}) \in K^{2 \times 2}$ hat genau dann maximalen Rang, wenn der Ausdruck $\det(A) := \alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12}$, den man als Determinante von A bezeichnet, nicht verschwindet.

Gegenüber konventionellen Betrachtungen bietet der Determinanten-Ausdruck $\det(A)$ den ganz wesentlichen Vorteil, dass mit seiner Hilfe die Maximalität des Rangs von A in direkter Weise von den Koeffizienten α_{ij} der Matrix A abgelesen werden kann, ohne dass Fallunterscheidungen zu beachten sind. Dieses Phänomen haben wir schon bei den Vorbemerkungen zu Kapitel 3 beobachten können. Dort hatten wir für

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} \in K^{2 \times 2}, \quad b = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \in K^2,$$

wobei A maximalen Rang habe, und einen Vektor von Unbekannten $t = \begin{pmatrix} t_1 \\ t_2 \end{pmatrix}$ das Gleichungssystem $A \cdot t = b$ studiert und eine Lösung erhalten, die sich unter Verwendung des oben definierten Determinantenausdrucks in der Form

$$t_1 = \frac{\det \begin{pmatrix} \beta_1 & \alpha_{12} \\ \beta_2 & \alpha_{22} \end{pmatrix}}{\det \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix}}, \quad t_2 = \frac{\det \begin{pmatrix} \alpha_{11} & \beta_1 \\ \alpha_{21} & \beta_2 \end{pmatrix}}{\det \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix}}$$

schreiben lässt. Auch hier waren bei der Herleitung Fallunterscheidungen notwendig, obwohl sich das Ergebnis am Ende als fallunabhängig herausstellte. Es handelt sich um den einfachsten Fall der *Cramerschen Regel*, die wir in allgemeiner Form in diesem Kapitel beweisen werden.

Historisch ist die Einführung von Determinanten in der Tat im Zusammenhang mit der Lösung linearer Gleichungssysteme zu sehen, und zwar mit der Entdeckung der nach Cramer benannten Regel. Wir wollen dies am Beispiel eines Systems bestehend aus 3 linearen Gleichungen mit 3 Unbekannten einmal erläutern:

$$\begin{aligned} \alpha_{11}t_1 + \alpha_{12}t_2 + \alpha_{13}t_3 &= \beta_1 \\ \alpha_{21}t_1 + \alpha_{22}t_2 + \alpha_{23}t_3 &= \beta_2 \\ \alpha_{31}t_1 + \alpha_{32}t_2 + \alpha_{33}t_3 &= \beta_3 \end{aligned}$$

Die Koeffizientenmatrix $A = (\alpha_{ij})_{i,j=1,2,3}$ habe den Rang 3. Dann hat die Matrix $A' = (\alpha_{ij})_{i=1,2,3;j=2,3}$ noch Rang 2, wobei wir mittels Vertauschungen der Reihenfolge der Gleichungen annehmen können, dass $(\alpha_{ij})_{i=1,2;j=2,3}$ Rang 2 hat. Wir bemühen uns nun, Koeffizienten $c_1, c_2 \in K$ zu finden, derart dass bei Addition des c_1 -fachen der ersten Zeile und des c_2 -fachen der zweiten Zeile zur dritten Zeile die Unbekannten t_2 und t_3 eliminiert werden. Dies ist möglich, da nach unserer Annahme die beiden ersten Zeilen von A' linear unabhängig sind. Wir müssen hierzu das Gleichungssystem

$$\begin{aligned} \alpha_{12}c_1 + \alpha_{22}c_2 &= -\alpha_{32} \\ \alpha_{13}c_1 + \alpha_{23}c_2 &= -\alpha_{33} \end{aligned}$$

lösen und erhalten gemäß der Cramerschen Regel für 2×2 -Matrizen

$$c_1 = \frac{\det \begin{pmatrix} -\alpha_{32} & \alpha_{22} \\ -\alpha_{33} & \alpha_{23} \end{pmatrix}}{\det \begin{pmatrix} \alpha_{12} & \alpha_{22} \\ \alpha_{13} & \alpha_{23} \end{pmatrix}} = \frac{\det \begin{pmatrix} \alpha_{22} & \alpha_{23} \\ \alpha_{32} & \alpha_{33} \end{pmatrix}}{\det \begin{pmatrix} \alpha_{12} & \alpha_{13} \\ \alpha_{22} & \alpha_{23} \end{pmatrix}},$$

$$c_2 = \frac{\det \begin{pmatrix} \alpha_{12} & -\alpha_{32} \\ \alpha_{13} & -\alpha_{33} \end{pmatrix}}{\det \begin{pmatrix} \alpha_{12} & \alpha_{22} \\ \alpha_{13} & \alpha_{23} \end{pmatrix}} = - \frac{\det \begin{pmatrix} \alpha_{12} & \alpha_{13} \\ \alpha_{32} & \alpha_{33} \end{pmatrix}}{\det \begin{pmatrix} \alpha_{12} & \alpha_{13} \\ \alpha_{22} & \alpha_{23} \end{pmatrix}},$$

wobei $\det \begin{pmatrix} \alpha_{12} & \alpha_{13} \\ \alpha_{22} & \alpha_{23} \end{pmatrix}$ nicht verschwindet. Somit ergibt sich als dritte Gleichung unseres ursprünglichen Systems

$$(\alpha_{11}c_1 + \alpha_{21}c_2 + \alpha_{31})t_1 = \beta_1c_1 + \beta_2c_2 + \beta_3$$

mit einem Faktor $\alpha_{11}c_1 + \alpha_{21}c_2 + \alpha_{31}$, der nicht verschwindet, da sich der Rang der Koeffizientenmatrix des ursprünglichen Gleichungssystems nicht geändert hat. Dann aber können wir durch diesen Koeffizienten dividieren, und es folgt

$$t_1 = \frac{\beta_1 \det \begin{pmatrix} \alpha_{22} & \alpha_{23} \\ \alpha_{32} & \alpha_{33} \end{pmatrix} - \beta_2 \det \begin{pmatrix} \alpha_{12} & \alpha_{13} \\ \alpha_{32} & \alpha_{33} \end{pmatrix} + \beta_3 \det \begin{pmatrix} \alpha_{12} & \alpha_{13} \\ \alpha_{22} & \alpha_{23} \end{pmatrix}}{\alpha_{11} \det \begin{pmatrix} \alpha_{22} & \alpha_{23} \\ \alpha_{32} & \alpha_{33} \end{pmatrix} - \alpha_{21} \det \begin{pmatrix} \alpha_{12} & \alpha_{13} \\ \alpha_{32} & \alpha_{33} \end{pmatrix} + \alpha_{31} \det \begin{pmatrix} \alpha_{12} & \alpha_{13} \\ \alpha_{22} & \alpha_{23} \end{pmatrix}}.$$

Wir haben diese Formel für t_1 unter der Annahme hergeleitet, dass die beiden ersten Zeilen der Matrix $A' = (\alpha_{ij})_{i=1,2,3;j=2,3}$ linear unabhängig sind, ein Fall, den wir durch Vertauschen der Reihenfolge der Gleichungen unseres Systems stets herstellen können. Um zu sehen, dass die Formel für t_1 auch im Allgemeinen gültig ist, bleibt noch zu zeigen, dass diese invariant gegenüber solchen Vertauschungen ist. Dies ist aber leicht nachzuprüfen, wenn man benutzt, dass die Determinante (einer 2×2 -Matrix) invariant unter Transponieren ist und das Vorzeichen wechselt, wenn man zwei Spalten oder Zeilen der Matrix vertauscht.

In der Formel für t_1 erkennt man, dass Zähler und Nenner nach dem gleichen Bildungsgesetz aus den Matrizen

$$\begin{pmatrix} \beta_1 & \alpha_{12} & \alpha_{13} \\ \beta_2 & \alpha_{22} & \alpha_{23} \\ \beta_3 & \alpha_{32} & \alpha_{33} \end{pmatrix}, \quad \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{pmatrix}$$

hervorgehen. In Kenntnis der Determinantentheorie würde man sagen, dass im Zähler und Nenner von t_1 jeweils die Determinanten dieser Matrizen stehen, und zwar in Form ihrer Entwicklung nach der ersten Spalte. Man kann nun den Ausdruck im Zähler bzw. Nenner von t_1 als *Definition* der Determinante einer 3×3 -Matrix nehmen, wobei diese Definition dann Bezug nimmt auf die Kenntnis von Determinanten von 2×2 -Matrizen.

Das gerade vorgestellte Verfahren funktioniert in induktiver Weise allgemeiner für Systeme von n linearen Gleichungen mit n Unbekannten, sofern die Koeffizientenmatrix maximalen Rang hat. Es ergibt sich eine Formel für t_1 ähnlich wie oben, wobei im Zähler und Nenner jeweils eine Summe von Vielfachen von Determinanten von $(n-1) \times (n-1)$ -Matrizen steht, in Form der Entwicklung der Determinante einer $n \times n$ -Matrix nach ihrer ersten Spalte. Die gewonnenen Ausdrücke kann man zur Definition der Determinante einer $n \times n$ -Matrix erheben, wodurch man dann insgesamt eine rekursive Definition der Determinante einer Matrix erhält.

Man beachte allerdings, dass das obige Verfahren zwar für t_1 auf die Entwicklung der Determinante einer $n \times n$ -Matrix nach ihrer ersten Spalte führt, für beliebiges t_j mit $1 \leq j \leq n$ jedoch auf die Entwicklung nach der j -ten Spalte. Dies sieht man etwa, indem man die Unbekannten t_1 und t_j vertauscht. Will man also die Cramersche Regel beweisen und gleichzeitig dabei Determinanten einführen, so hat man bei jedem Rekursionsschritt auch einige allgemeine Eigenschaften von Determinanten mitzubeweisen, etwa die Möglichkeit, Determinanten nach beliebigen Spalten (oder Zeilen) zu entwickeln. Wir werden natürlich nicht in dieser Weise vorgehen. Stattdessen beginnen wir mit dem Studium so genannter Determinantenfunktionen, in deren Kontext sich automatisch eine allgemeine nicht-rekursive Definition für Determinanten ergeben wird. Dabei werden einige Fakten über Permutationen benötigt, die wir zu Beginn des Kapitels zusammenstellen. Erst nach Klärung der Definition und der Eigenschaften von Determinanten beweisen wir schließlich die Cramersche Regel und im Zusammenhang hiermit auch die Möglichkeit, Determinanten nach Spalten oder Zeilen zu entwickeln. In einem optionalen Abschnitt geht es schließlich noch um äußere Produkte, in deren Zusammenhang wir auch den so genannten allgemeinen Laplaceschen Entwicklungssatz für Determinanten herleiten.

4.1 Permutationen

Wir hatten in Abschnitt 1.2 gesehen, dass für eine Menge X die bijektiven Selbstabbildungen $X \rightarrow X$ eine Gruppe G bilden, wenn man die Komposition von Abbildungen als Verknüpfung nimmt. Die identische Abbildung $\text{id} \in G$ ist das Einselement, und das inverse Element zu einem Element $f: X \rightarrow X$ von G wird gegeben durch die inverse Abbildung $f^{-1}: X \rightarrow X$.

Wir wollen hier für eine natürliche Zahl n speziell die Menge $X = \{1, \dots, n\}$ betrachten, wobei wir X im Falle $n = 0$ als die leere Menge annehmen. Für die zugehörige Gruppe der bijektiven Selbstabbildungen von X schreibt man dann $\mathfrak{S}_n$ und nennt dies die *symmetrische Gruppe* oder die *Permutationsgruppe* zum Index n . Die Elemente von $\mathfrak{S}_n$ heißen *Permutationen* der Zahlen $1, \dots, n$, da sie sozusagen diese Elemente in ihrer Reihenfolge vertauschen. Gilt für eine Permutation $\sigma \in \mathfrak{S}_n$ etwa $\sigma(i) = a_i$ für $i = 1, \dots, n$, so schreibt man auch

$$\sigma = \begin{pmatrix} 1 & \dots & n \\ a_1 & \dots & a_n \end{pmatrix}.$$

Die Permutationsgruppen $\mathfrak{S}_0$ und $\mathfrak{S}_1$ bestehen jeweils nur aus dem Einselement, welches die identische Abbildung auf $X = \emptyset$ bzw. $X = \{1\}$ repräsentiert. Weiter ist $\mathfrak{S}_2$ eine Gruppe von 2 Elementen und damit insbesondere abelsch, wohingegen $\mathfrak{S}_n$ für $n \geq 3$ nicht mehr abelsch ist. Für

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad \tau = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$$

gilt beispielsweise

$$\sigma \circ \tau = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \quad \tau \circ \sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}.$$

Wir wollen die Anzahl der Elemente von $\mathfrak{S}_n$ bestimmen. Hierzu bezeichnen wir für $n \in \mathbb{N}$ die natürliche Zahl $\prod_{i=1}^n i$ mit $n!$, wobei das Rufzeichen “!” als *Fakultät* gelesen wird. Insbesondere gilt $0! = 1! = 1$ sowie $2! = 2$.

Bemerkung 1. *Es besteht $\mathfrak{S}_n$ aus genau $n!$ Elementen.*

Beweis. Will man eine bijektive Selbstabbildung

$$\sigma: \{1, \dots, n\} \longrightarrow \{1, \dots, n\}$$

erklären, so kann man schrittweise vorgehen und zunächst $\sigma(1)$ festlegen, dann $\sigma(2)$ und so weiter, bis man schließlich $\sigma(n)$ festlegt. Dabei genügt es, eine injektive Abbildung zu konstruieren; diese ist wegen der Endlichkeit der Menge $\{1, \dots, n\}$ automatisch bijektiv. Für die Wahl von $\sigma(1)$ hat man zunächst n Möglichkeiten, denn für $\sigma(1)$ kann man jedes Element aus $\{1, \dots, n\}$ nehmen. Für die Wahl von $\sigma(2)$ verbleiben dann noch $n - 1$ Möglichkeiten, denn man muss $\sigma(2) \in \{1, \dots, n\}$ verschieden von $\sigma(1)$ wählen, da man eine injektive Abbildung konstruieren möchte. Bei der Festlegung von $\sigma(3)$ hat man noch $n - 2$ Möglichkeiten und so weiter, bis man schließlich bei der Wahl von $\sigma(n)$ lediglich noch eine Möglichkeit hat. Indem man das Produkt über die einzelnen Anzahlen bildet, erhält man insgesamt $n! = \prod_{i=1}^n i$ als Anzahl der Möglichkeiten, eine injektive Selbstabbildung von $\{1, \dots, n\}$ zu erklären. $\square$

Definition 2. *Eine Permutation $\tau \in \mathfrak{S}_n$ heißt Transposition, wenn es zwei verschiedene Zahlen $i, j \in \{1, \dots, n\}$ mit $\tau(i) = j$, $\tau(j) = i$ sowie $\tau(k) = k$ für alle restlichen Zahlen $k \in \{1, \dots, n\}$ gibt. Man schreibt dann auch $\tau = (i, j)$.*

Die Transposition $\tau = (i, j)$ vertauscht also gerade die Zahlen i und j und lässt ansonsten alle Elemente von $\{1, \dots, n\}$ fest. Insbesondere gilt $\tau^2 = \text{id}$, d. h. τ ist zu sich selbst invers. Es besteht $\mathfrak{S}_2$ gerade aus der identischen Abbildung und der Transposition $(1, 2)$, wohingegen die Gruppen $\mathfrak{S}_0$ und $\mathfrak{S}_1$ keine Transpositionen enthalten.

Satz 3. *Jedes $\pi \in \mathfrak{S}_n$ ist ein Produkt von Transpositionen.*

Beweis. Wir schließen mit fallender Induktion nach $r(\pi)$, wobei $r(\pi)$ maximal in $\{0, 1, \dots, n\}$ gewählt sei mit der Eigenschaft $\pi(i) = i$ für $i = 1, \dots, r(\pi)$. Für $r(\pi) = n$ gilt $\pi = \text{id}$, und dies ist ein leeres Produkt von Transpositionen. Für $r = r(\pi) < n$ folgt notwendig $\pi(r+1) > r+1$. Setzen wir $\tau_1 = (r+1, \pi(r+1))$, so bildet das Produkt $\tau_1\pi$ die Elemente $1, \dots, r+1$ identisch auf sich selbst ab, erfüllt also $r(\tau_1\pi) > r(\pi)$, und ist somit nach Induktionsvoraussetzung ein Produkt von Transpositionen, etwa $\tau_1\pi = \tau_2 \dots \tau_s$. Multiplikation mit $\tau_1 = \tau_1^{-1}$ von links liefert dann wie gewünscht $\pi = \tau_1 \dots \tau_s$. $\square$

Man nennt eine Permutation $\pi \in \mathfrak{S}_n$ *gerade* oder *ungerade*, je nachdem ob sich π als ein Produkt einer geraden oder ungeraden Anzahl von Transpositionen darstellen lässt. Dabei wollen wir zeigen, dass eine Permutation π nicht gerade und ungerade zugleich sein kann. Als Hilfsmittel führen wir das so genannte *Signum* von π ein.

Definition 4. Sei $\pi \in \mathfrak{S}_n$. Dann heißt

$$\text{sgn } \pi = \prod_{1 \leq i < j \leq n} \frac{\pi(j) - \pi(i)}{j - i}$$

das Signum der Permutation π .

Im Grunde genommen erstreckt sich die Produktbildung bei der Definition des Signums einer Permutation $\pi \in \mathfrak{S}_n$ über alle zwei-elementigen Teilmengen von $\{1, \dots, n\}$. Genauer meinen wir hiermit:

Bemerkung 5. Es sei M eine Menge von Indexpaaren (i, j) mit $1 \leq i, j \leq n$, so dass die Zuordnung $(i, j) \mapsto \{i, j\}$ eine Bijektion von M auf die Menge der zwei-elementigen Teilmengen von $\{1, \dots, n\}$ erklärt. Dann gilt

$$\text{sgn } \pi = \prod_{(i,j) \in M} \frac{\pi(j) - \pi(i)}{j - i}$$

Beweis. Zunächst ist $\{\{i, j\}; 1 \leq i < j \leq n\}$ gerade die Menge aller zwei-elementigen Teilmengen von $\{1, \dots, n\}$. Berücksichtigt man dann noch die Gleichung

$$\frac{\pi(j) - \pi(i)}{j - i} = \frac{\pi(i) - \pi(j)}{i - j}$$

so folgt die Behauptung. $\square$

Wir wollen einige Folgerungen aus dieser Bemerkung ziehen.

Satz 6. Sei $\pi \in \mathfrak{S}_n$. Dann gilt $\text{sgn } \pi = (-1)^s$, wobei s die Anzahl aller Fehlstände in der Folge $\pi(1), \dots, \pi(n)$ ist, d. h. die Anzahl aller Paare

$(\pi(i), \pi(j))$, $1 \leq i < j \leq n$, mit $\pi(i) > \pi(j)$. Insbesondere folgt $\operatorname{sgn} \pi = \pm 1$. Für eine Transposition $\pi \in \mathfrak{S}_n$ berechnet sich das Signum zu $\operatorname{sgn} \pi = -1$.

Beweis. Eine Permutation $\pi \in \mathfrak{S}_n$ bildet die zwei-elementigen Teilmengen von $\{1, \dots, n\}$ bijektiv auf sich selbst ab. Daher gilt

$$\prod_{1 \leq i < j \leq n} (j - i) = \pm \prod_{1 \leq i < j \leq n} (\pi(j) - \pi(i))$$

und damit $\operatorname{sgn} \pi = \pm 1$. Genauer erhält man $\operatorname{sgn} \pi = (-1)^s$, wobei s gleich der Anzahl der Faktoren auf der rechten Seite ist, die negativ sind, also gleich der Anzahl der Fehlstände in der Folge $\pi(1), \dots, \pi(n)$.

Für eine Transposition $\pi = (i, j)$ mit $i < j$, also

$$\pi = \begin{pmatrix} 1 & \dots & i-1 & i & i+1 & \dots & j-1 & j & j+1 & \dots & n \\ 1 & \dots & i-1 & j & i+1 & \dots & j-1 & i & j+1 & \dots & n \end{pmatrix},$$

gibt es in der Folge $\pi(1), \dots, \pi(n)$ genau $j-i$ Fehlstände der Form $(j, *)$ sowie $j-i$ Fehlstände der Form $(*, i)$. Da wir dabei (j, i) als Fehlstand in zweifacher Weise berücksichtigt haben, ergibt sich

$$\operatorname{sgn} \pi = (-1)^{2(j-i)-1} = -1,$$

wie behauptet. □

Satz 7. Die Abbildung $\operatorname{sgn}: \mathfrak{S}_n \longrightarrow \{1, -1\}$ ist ein Gruppenhomomorphismus, d. h. es gilt $\operatorname{sgn}(\sigma \circ \tau) = \operatorname{sgn} \sigma \cdot \operatorname{sgn} \tau$ für $\sigma, \tau \in \mathfrak{S}_n$.

Beweis. Mit Bemerkung 5 können wir wie folgt rechnen:

$$\begin{aligned} \operatorname{sgn}(\sigma \circ \tau) &= \prod_{i < j} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{j - i} \\ &= \prod_{i < j} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)} \cdot \prod_{i < j} \frac{\tau(j) - \tau(i)}{j - i} \\ &= \operatorname{sgn} \sigma \cdot \operatorname{sgn} \tau \end{aligned}$$

□

Korollar 8. Ist $\pi \in \mathfrak{S}_n$ darstellbar als ein Produkt von s Transpositionen, so gilt $\operatorname{sgn} \pi = (-1)^s$.

Wir sehen damit, dass eine Permutation $\pi \in \mathfrak{S}_n$ genau dann gerade, also ein Produkt einer geraden Anzahl von Transpositionen ist, wenn $\operatorname{sgn} \pi = 1$ gilt. Entsprechend ist π genau dann ungerade, also ein Produkt einer ungeraden Anzahl von Transpositionen, wenn $\operatorname{sgn} \pi = -1$ gilt.

In Satz 7 haben wir von einem *Gruppenhomomorphismus* gesprochen. Allgemein nennt man eine Abbildung $f: G \longrightarrow G'$ zwischen zwei Gruppen G und

G' einen Gruppenhomomorphismus, wenn f mit der Produktbildung in G und G' verträglich ist, wenn also für $g, h \in G$ stets $f(g \cdot h) = f(g) \cdot f(h)$ folgt. Dann gilt $f(1) = 1$ für die Einselemente in G und G' , sowie $f(g^{-1}) = f(g)^{-1}$ für $g \in G$. Man kann sich leicht überlegen, dass der *Kern* von f , also $\ker f = f^{-1}(1)$, eine Untergruppe von G ist, sogar ein *Normalteiler*. Hierunter versteht man eine Untergruppe $N \subset G$ mit $g^{-1} \cdot h \cdot g \in N$ für alle $g \in G$ und $h \in N$. Letzteres ist natürlich nur für nicht-kommutative Gruppen eine echte Bedingung; in einer kommutativen Gruppe ist jede Untergruppe ein Normalteiler.

In der Situation von Satz 7 ist also sgn ein Gruppenhomomorphismus von $\mathfrak{S}_n$ in die multiplikative Gruppe $\{1, -1\}$. Der Kern dieses Homomorphismus besteht aus allen geraden Permutationen aus $\mathfrak{S}_n$. Diese bilden einen Normalteiler und insbesondere eine Untergruppe $\mathfrak{A}_n \subset \mathfrak{S}_n$. Man nennt $\mathfrak{A}_n$ die *alternierende Gruppe* zum Index n . Zusammenfassend können wir sagen:

Bemerkung 9. Die geraden Permutationen in $\mathfrak{S}_n$ bilden einen Normalteiler

$$\mathfrak{A}_n = \{\pi \in \mathfrak{S}_n; \operatorname{sgn} \pi = 1\} \subset \mathfrak{S}_n.$$

Weiter besteht

$$\mathfrak{S}_n - \mathfrak{A}_n = \{\pi \in \mathfrak{S}_n; \operatorname{sgn} \pi = -1\}$$

aus allen ungeraden Permutationen in $\mathfrak{S}_n$, und es gilt

$$\mathfrak{S}_n - \mathfrak{A}_n = \tau \mathfrak{A}_n = \mathfrak{A}_n \tau$$

für jede ungerade Permutation $\tau \in \mathfrak{S}_n$. Dabei bezeichnet $\tau \mathfrak{A}_n$ die Menge aller Produkte $\tau \circ \pi$ mit $\pi \in \mathfrak{A}_n$, entsprechend für $\mathfrak{A}_n \tau$.

Beweis. Es ist nur zeigen, dass sich die ungeraden Permutationen aus $\mathfrak{S}_n$ in der Form $\tau \mathfrak{A}_n$ bzw. $\mathfrak{A}_n \tau$ beschreiben lassen. Ist τ ungerade, so sieht man mittels Satz 7, dass $\tau \mathfrak{A}_n$ und $\mathfrak{A}_n \tau$ in $\mathfrak{S}_n - \mathfrak{A}_n$ enthalten sind. Ist umgekehrt $\pi \in \mathfrak{S}_n - \mathfrak{A}_n$, so sind $\tau^{-1} \circ \pi$ und $\pi \circ \tau^{-1}$ gerade, woraus $\pi \in \tau \mathfrak{A}_n$ bzw. $\pi \in \mathfrak{A}_n \tau$ folgt. $\square$

Aufgaben

1. Man bestimme das Signum der folgenden Permutationen:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 2 & 1 \end{pmatrix} \in \mathfrak{S}_5, \quad \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 3 & 1 & 5 & 6 & 8 & 4 & 7 \end{pmatrix} \in \mathfrak{S}_8$$

2. Eine Permutation $\pi \in \mathfrak{S}_n$ heißt ein *r-Zyklus*, wenn es paarweise verschiedene Elemente $a_1, \dots, a_r \in \{1, \dots, n\}$ gibt mit

$$\begin{aligned} \pi(a_i) &= a_{i+1} \text{ für } i = 1, \dots, r-1, \\ \pi(a_r) &= a_1, \end{aligned}$$

und π alle übrigen Elemente von $\{1, \dots, n\}$ fest lässt. Man bestimme $\operatorname{sgn} \pi$ für einen *r-Zyklus* $\pi \in \mathfrak{S}_n$.

3. Man zeige, dass jedes Element $\pi \in \mathfrak{S}_n$ endliche Ordnung besitzt, d. h. dass es ein $r \in \mathbb{N} - \{0\}$ gibt mit $\pi^r = 1$.
4. Man zeige: Ist $N \subset \mathfrak{S}_n$ ein Normalteiler, der eine Transposition enthält, so gilt bereits $N = \mathfrak{S}_n$.
5. Man bestimme die Anzahl der Elemente der alternierenden Gruppe $\mathfrak{A}_n$.
6. Man zeige für eine Gruppe G und eine Untergruppe $N \subset G$, dass N bereits dann ein Normalteiler in G ist, wenn $gN \subset Ng$ für alle $g \in G$ gilt.
7. Für eine Gruppe G bezeichne $S(G)$ die Menge aller bijektiven Abbildungen von G nach G . Man zeige:
 - (i) $S(G)$ ist eine Gruppe unter der Komposition von Abbildungen.
 - (ii) Es existiert ein injektiver Gruppenhomomorphismus $G \longrightarrow S(G)$.

4.2 Determinantenfunktionen

In diesem Abschnitt fixieren wir einen K -Vektorraum V endlicher Dimension n und betrachten Abbildungen $\Delta: V^n \longrightarrow K$, sozusagen K -wertige Funktionen von n Argumenten aus V , wobei V^n im Falle $n = 0$ nur aus dem leeren Tupel besteht. Eine solche Abbildung heißt *multilinear*, wenn sie linear in jedem Argument ist. Dabei nennt man Δ linear im i -ten Argument, wenn für $a_1, \dots, a_n, a'_i \in V$ und $\alpha, \alpha' \in K$ stets

$$\begin{aligned} & \Delta(a_1, \dots, a_{i-1}, \alpha a_i + \alpha' a'_i, a_{i+1}, \dots, a_n) \\ &= \alpha \cdot \Delta(a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_n) \\ &+ \alpha' \cdot \Delta(a_1, \dots, a_{i-1}, a'_i, a_{i+1}, \dots, a_n) \end{aligned}$$

gilt. Weiter heißt Δ *alternierend*, wenn $\Delta(a_1, \dots, a_n)$ verschwindet, sobald zwei der Vektoren $a_1, \dots, a_n$ übereinstimmen.

Definition 1. *Es sei V ein K -Vektorraum endlicher Dimension n . Eine Determinantenfunktion auf V ist eine alternierende multilineare Abbildung $\Delta: V^n \longrightarrow K$. Man bezeichnet Δ als nicht-trivial, wenn Δ nicht die Nullabbildung ist.*

Ähnlich wie K -lineare Abbildungen kann man Determinantenfunktionen addieren oder mit Skalaren aus K multiplizieren, um neue solche Funktionen zu erhalten. Die Determinantenfunktionen auf V bilden folglich einen K -Vektorraum, und wir werden im Folgenden sehen, dass dessen Dimension 1 ist, unabhängig von der Dimension von V . Doch zunächst wollen wir einige elementare Eigenschaften von Determinantenfunktionen angeben.

Bemerkung 2. *Zu einem K -Vektorraum V der Dimension n betrachte man eine multilineare Abbildung $\Delta: V^n \longrightarrow K$. Dann ist äquivalent:*

- (i) Δ ist alternierend und damit eine Determinantenfunktion.

(ii) Für jedes linear abhängige System von Vektoren $a_1, \dots, a_n \in V$ gilt $\Delta(a_1, \dots, a_n) = 0$.

Beweis. Es ist nur die Implikation (i) $\implies$ (ii) zu zeigen. Seien also $a_1, \dots, a_n$ linear abhängig. Dann lässt sich einer dieser Vektoren, etwa a_1 , als Linearkombination der restlichen darstellen, also $a_1 = \sum_{i=2}^n \beta_i a_i$, und es folgt aus (i) unter Benutzung der Multilinearität von Δ

$$\Delta(a_1, \dots, a_n) = \sum_{i=2}^n \beta_i \Delta(a_i, a_2, \dots, a_n) = 0.$$

□

Bemerkung 3. Sei Δ eine Determinantenfunktion auf V . Für $a_1, \dots, a_n \in V$ gilt dann:

(i) $\Delta(\dots a_i \dots a_j \dots) = -\Delta(\dots a_j \dots a_i \dots)$, d. h. Δ ändert bei Vertauschung zweier verschiedener Argumente das Vorzeichen.

(ii) $\Delta(a_{\pi(1)}, \dots, a_{\pi(n)}) = \text{sgn } \pi \cdot \Delta(a_1, \dots, a_n)$ für $\pi \in \mathfrak{S}_n$.

(iii) $\Delta(\dots a_i + \alpha a_j \dots a_j \dots) = \Delta(\dots a_i \dots a_j \dots)$ für $\alpha \in K$, d. h. Δ bleibt invariant, wenn man zum i -ten Argument ein skalares Vielfaches eines weiteren Arguments addiert.

Beweis. Unter Ausnutzung der Eigenschaften einer Determinantenfunktion können wir wie folgt rechnen:

$$\begin{aligned} 0 &= \Delta(\dots a_i + a_j \dots a_i + a_j \dots) \\ &= \Delta(\dots a_i \dots a_i + a_j \dots) + \Delta(\dots a_j \dots a_i + a_j \dots) \\ &= \Delta(\dots a_i \dots a_i \dots) + \Delta(\dots a_i \dots a_j \dots) \\ &\quad + \Delta(\dots a_j \dots a_i \dots) + \Delta(\dots a_j \dots a_j \dots) \\ &= \Delta(\dots a_i \dots a_j \dots) + \Delta(\dots a_j \dots a_i \dots) \end{aligned}$$

Dabei mögen die explizit aufgeschriebenen Argumente immer an den Stellen i und j stehen, wobei an den übrigen Stellen stets $a_1, \dots, a_{i-1}$ bzw. $a_{i+1}, \dots, a_{j-1}$ bzw. $a_{j+1}, \dots, a_n$ als Argumente einzusetzen sind. Aus der Rechnung folgt dann Aussage (i). Weiter ist (ii) eine Folgerung von (i), da jede Permutation nach 4.1/3 ein Produkt von Transpositionen ist und da man nach 4.1/8 $\text{sgn } \pi = (-1)^s$ hat, wenn π ein Produkt von s Transpositionen ist. Aussage (iii) schließlich ergibt sich aus folgender Rechnung:

$$\begin{aligned} &\Delta(\dots a_i + \alpha a_j \dots a_j \dots) \\ &= \Delta(\dots a_i \dots a_j \dots) + \alpha \Delta(\dots a_j \dots a_j \dots) \\ &= \Delta(\dots a_i \dots a_j \dots) \end{aligned}$$

□

Eine multilineare Abbildung $\Delta: V^n \longrightarrow K$, die nach unserer Definition alternierend ist, ist also insbesondere alternierend in dem Sinne, dass sie das

Vorzeichen bei Vertauschung zweier ihrer Argumente ändert; daher die Bezeichnung alternierend. Man beachte jedoch, dass die Umkehrung hierzu nur dann richtig ist, wenn $1 \neq -1$ in K gilt.

Als Nächstes wollen wir ein nicht-triviales Beispiel einer Determinantenfunktion auf dem Vektorraum $V = K^n$ konstruieren.

Lemma 4. Für Matrizen $A = (\alpha_{ij})_{i,j} \in K^{n \times n}$ definiere man die Determinante von A durch

$$\det(A) = \sum_{\pi \in \mathfrak{S}_n} \operatorname{sgn} \pi \cdot \alpha_{\pi(1),1} \cdot \dots \cdot \alpha_{\pi(n),n}.$$

Dann ist $\det$, aufgefasst als Funktion in den n Spalten der Matrix A , eine Determinantenfunktion auf $V = K^n$, und es gilt $\det(E) = 1$ für die Einheitsmatrix $E \in K^{n \times n}$. Insbesondere ist die Determinantenfunktion $\det$ nicht-trivial.

Beweis. Zunächst ist die Funktion $\det$ multilinear in den Spalten der Matrix A , da ein Produkt $c_1 \cdot \dots \cdot c_n$ linear in jedem seiner Faktoren ist. Weiter ist zu zeigen, dass $\det(A) = 0$ gilt, sofern zwei Spalten in A identisch sind. Seien also $a_1, \dots, a_n$ die Spalten von A . Gilt dann $a_i = a_j$ für zwei Indizes $i \neq j$, so kann man die Transposition $\tau = (i, j) \in \mathfrak{S}_n$ betrachten. Durchläuft π die geraden Permutationen in $\mathfrak{S}_n$, so durchläuft $\pi \circ \tau$ nach 4.1/9 alle ungeraden Permutationen in $\mathfrak{S}_n$. Man erhält daher

$$\begin{aligned} & \det(a_1, \dots, a_i, \dots, a_j, \dots, a_n) \\ &= \sum_{\pi \in \mathfrak{A}_n} \alpha_{\pi(1),1} \cdot \dots \cdot \alpha_{\pi(i),i} \cdot \dots \cdot \alpha_{\pi(j),j} \cdot \dots \cdot \alpha_{\pi(n),n} \\ &- \sum_{\pi \in \mathfrak{A}_n} \alpha_{\pi \circ \tau(1),1} \cdot \dots \cdot \alpha_{\pi \circ \tau(i),i} \cdot \dots \cdot \alpha_{\pi \circ \tau(j),j} \cdot \dots \cdot \alpha_{\pi \circ \tau(n),n} \\ &= 0, \end{aligned}$$

denn es gilt

$$\begin{aligned} \alpha_{\pi \circ \tau(i),i} &= \alpha_{\pi(j),i} = \alpha_{\pi(j),j} \cdot \\ \alpha_{\pi \circ \tau(j),j} &= \alpha_{\pi(i),j} = \alpha_{\pi(i),i}, \end{aligned}$$

sowie $\alpha_{\pi \circ \tau(k),k} = \alpha_{\pi(k),k}$ für $k \neq i, j$.

Schließlich verifiziert man leicht $\det(E) = 1$. In diesem Falle verschwinden nämlich in der definierenden Summe für $\det(E)$ alle Terme bis auf denjenigen zur Identität $\operatorname{id} \in \mathfrak{S}_n$. $\square$

Im Übrigen werden wir in 7.2/9 und 7.2/10 sehen, dass der Betrag der Determinante $\det(A)$ geometrisch als das Volumen des von den Spaltenvektoren von A in K^n aufgespannten Parallelotops interpretiert werden kann.

Als Nächstes wollen wir einsehen, dass das soeben gegebene Beispiel einer Determinantenfunktion auf ganz natürliche Weise entsteht.

Lemma 5. *Es sei Δ eine Determinantenfunktion auf V und $X = (x_1, \dots, x_n)$ eine Basis von V . Für beliebige Vektoren $a_1, \dots, a_n \in V$ gilt dann:*

$$\Delta(a_1, \dots, a_n) = \det(a_{1,X}, \dots, a_{n,X}) \cdot \Delta(x_1, \dots, x_n).$$

Dabei bezeichnet $a_{j,X} \in K^n$ für $j = 1, \dots, n$ den Koordinatenspaltenvektor von a_j bezüglich der Basis X , also $a_{j,X} = (\alpha_{1j}, \dots, \alpha_{nj})^t$, wenn $a_j = \sum_{i=1}^n \alpha_{ij} x_i$.

Beweis. Unter Benutzung der Eigenschaften einer Determinantenfunktion ergibt sich:

$$\begin{aligned} & \Delta(a_1, \dots, a_n) \\ &= \sum_{i_1, \dots, i_n=1}^n \alpha_{i_1,1} \dots \alpha_{i_n,n} \cdot \Delta(x_{i_1}, \dots, x_{i_n}) \\ &= \sum_{\pi \in \mathfrak{S}_n} \alpha_{\pi(1),1} \dots \alpha_{\pi(n),n} \cdot \Delta(x_{\pi(1)}, \dots, x_{\pi(n)}) \\ &= \sum_{\pi \in \mathfrak{S}_n} \operatorname{sgn} \pi \cdot \alpha_{\pi(1),1} \dots \alpha_{\pi(n),n} \cdot \Delta(x_1, \dots, x_n) \\ &= \det(a_{1,X}, \dots, a_{n,X}) \cdot \Delta(x_1, \dots, x_n) \end{aligned}$$

□

Wir können nun in einfacher Weise einige Folgerungen aus den beiden Lemmata ziehen.

Satz 6. *Es sei $X = (x_1, \dots, x_n)$ eine Basis eines K -Vektorraums V und weiter Δ eine Determinantenfunktion auf V . Dann ist äquivalent:*

- (i) Δ ist trivial.
- (ii) $\Delta(x_1, \dots, x_n) = 0$.

Beweis. Es ist nur die Implikation (ii) $\implies$ (i) zu begründen, und diese ergibt sich aus der Formel in Lemma 5. □

Satz 7. *Es sei Δ eine nicht-triviale Determinantenfunktion auf einem n -dimensionalen K -Vektorraum V . Für n Vektoren $x_1, \dots, x_n \in V$ ist dann äquivalent:*

- (i) $x_1, \dots, x_n$ sind linear abhängig.
- (ii) $\Delta(x_1, \dots, x_n) = 0$.

Beweis. Die Implikation (i) $\implies$ (ii) ergibt sich aus Bemerkung 2. Um die umgekehrte Implikation zu zeigen, gehen wir indirekt vor und nehmen an, dass $x_1, \dots, x_n$ linear unabhängig sind, also eine Basis X von V bilden. Nach Satz 6 ist Δ dann trivial, im Widerspruch zu unserer Voraussetzung. □

Satz 8. *Es sei V ein K -Vektorraum mit einer Basis $X = (x_1, \dots, x_n)$. Bezeichnet a_X für Vektoren $a \in V$ jeweils den Koordinatenspaltenvektor von a zur Basis X , so wird durch*

$$\det_X(a_1, \dots, a_n) = \det(a_{1,X}, \dots, a_{n,X})$$

eine Determinantenfunktion mit $\det_X(x_1, \dots, x_n) = 1$ auf V erklärt. Der K -Vektorraum der Determinantenfunktionen auf V wird von $\det_X$ erzeugt und besitzt folglich die Dimension 1.

Beweis. Indem man den Isomorphismus $V \xrightarrow{\sim} K^n$, $a \mapsto a_X$, verwendet, kann man sehen, dass $\det_X$ die Eigenschaften einer nicht-trivialen Determinantenfunktion mit $\det_X(x_1, \dots, x_n) = 1$ besitzt, da $\det$ die entsprechenden Eigenschaften hat; vgl. Lemma 4. Im Übrigen zeigt Lemma 5, dass jede weitere Determinantenfunktion Δ auf V ein skalares Vielfaches von $\det_X$ ist. Der Vektorraum der Determinantenfunktionen auf V besitzt also die Dimension 1. $\square$

Korollar 9. *Es seien Δ, Δ' Determinantenfunktionen auf dem K -Vektorraum V . Ist dann Δ nicht-trivial, so existiert ein eindeutig bestimmtes Element $\alpha \in K$ mit $\Delta' = \alpha \cdot \Delta$.*

Aufgaben

1. Es seien V ein n -dimensionaler K -Vektorraum, $U \subset V$ ein r -dimensionaler linearer Unterraum und $x_{r+1}, \dots, x_n$ ein fest gegebenes System von Vektoren in V . Für eine Determinantenfunktion $\Delta: V^n \rightarrow K$ zeige man, dass sich durch

$$\Delta_U(a_1, \dots, a_r) := \Delta(a_1, \dots, a_r, x_{r+1}, \dots, x_n)$$

eine Determinantenfunktion $\Delta: U^r \rightarrow K$ definieren lässt. Wann ist Δ_U nicht-trivial?

2. Es sei $\Delta: V^n \rightarrow K$ eine nicht-triviale Determinantenfunktion auf einem n -dimensionalen K -Vektorraum V und $x_1, \dots, x_{n-1}$ ein linear unabhängiges System von Vektoren in V . Man zeige: Es existiert ein $x_n \in V$ mit $\Delta(x_1, \dots, x_n) = 1$, wobei die Restklasse von x_n in $V/\langle x_1, \dots, x_{n-1} \rangle$ eindeutig bestimmt ist.
3. Für einen n -dimensionalen K -Vektorraum V mit einer Basis $X = (x_1, \dots, x_n)$ betrachte man den K -Vektorraum W aller multilinearen Abbildungen $V^n \rightarrow K$. Man bestimme $\dim_K W$ und gebe eine Basis von W an.

4.3 Determinanten von Matrizen und Endomorphismen

Für eine Matrix $A = (\alpha_{ij})_{i,j} \in K^{n \times n}$ haben wir bereits in 4.2/4 die Determinante durch die Formel

$$\det(A) = \sum_{\pi \in \mathfrak{S}_n} \operatorname{sgn} \pi \cdot \alpha_{\pi(1),1} \cdot \dots \cdot \alpha_{\pi(n),n}.$$

erklärt. Diese Formel ergibt für $n = 0$ als Determinante der leeren Matrix den Wert 1. Im Folgenden wollen wir nun allgemeiner die Determinante von Endomorphismen endlich-dimensionaler K -Vektorräume einführen.

Bemerkung 1. Es sei $f: V \longrightarrow V$ ein Endomorphismus eines K -Vektorraums V endlicher Dimension n . Weiter sei Δ eine nicht-triviale Determinantenfunktion auf V ; eine solche existiert stets nach 4.2/8. Dann gilt:

(i) Auf V wird durch

$$\Delta_f(a_1, \dots, a_n) = \Delta(f(a_1), \dots, f(a_n)), \quad a_1, \dots, a_n \in V,$$

eine Determinantenfunktion Δ_f definiert.

(ii) Es existiert ein eindeutig bestimmtes Element $\alpha_f \in K$ mit $\Delta_f = \alpha_f \cdot \Delta$.

(iii) Ist $X = (x_1, \dots, x_n)$ eine Basis von V , so folgt

$$\alpha_f = \det(f(x_1)_X, \dots, f(x_n)_X) = \det(A_{f,X,X}),$$

wobei $f(x_i)_X$ der Koordinatenspaltenvektor von $f(x_i)$ bezüglich der Basis X und $A_{f,X,X}$ die Matrix von f bezüglich der Basis X sei. Insbesondere erkennt man, dass α_f unabhängig von der Wahl der Determinantenfunktion Δ ist.

Beweis. Die Funktion Δ_f ist offenbar multilinear, da Δ diese Eigenschaft hat und f linear ist. Sind weiter $a_1, \dots, a_n$ linear abhängig, so gilt Gleiches für $f(a_1), \dots, f(a_n)$, und es folgt $\Delta_f(a_1, \dots, a_n) = 0$. Folglich ist Δ_f eine Determinantenfunktion. Nach 4.2/9 gibt es dann eine eindeutig bestimmte Konstante $\alpha_f \in K$ mit $\Delta_f = \alpha_f \Delta$. Betrachten wir schließlich eine Basis $X = (x_1, \dots, x_n)$ von V , so ergibt sich mit 4.2/5

$$\begin{aligned} \alpha_f \cdot \Delta(x_1, \dots, x_n) &= \Delta_f(x_1, \dots, x_n) = \Delta(f(x_1), \dots, f(x_n)) \\ &= \det(f(x_1)_X, \dots, f(x_n)_X) \cdot \Delta(x_1, \dots, x_n), \end{aligned}$$

wobei $\Delta(x_1, \dots, x_n) \neq 0$ wegen 4.2/7 und folglich

$$\alpha_f = \det(f(x_1)_X, \dots, f(x_n)_X)$$

gilt. Insbesondere erkennt man, dass α_f unabhängig von der Wahl von Δ ist. $\square$

Definition 2. Es sei $f: V \longrightarrow V$ ein Endomorphismus eines endlich-dimensionalen K -Vektorraums V . Dann nennt man $\det f := \alpha_f$, wobei α_f wie in Bemerkung 1 bestimmt ist, die Determinante des Endomorphismus f . Es gilt

$$\det f = \det(A_{f,X,X})$$

für jede Matrix $A_{f,X,X}$, die f bezüglich einer Basis X von V beschreibt. Insbesondere erhält man $\det f = 1$ im Falle $V = 0$.

Wir wollen zunächst einige grundlegende Eigenschaften von Determinanten herleiten.

Satz 3. Es sei V ein K -Vektorraum endlicher Dimension n und $\text{End}_K(V)$ der Endomorphismenring von V . Dann gilt für $f, g \in \text{End}_K(V)$ und $\alpha \in K$:

- (i) $\det(\text{id}) = 1$ für die Identität $\text{id} \in \text{End}_K(V)$.
- (ii) $\det(\alpha f) = \alpha^n \cdot \det(f)$.
- (iii) $\det(f \circ g) = \det(f) \cdot \det(g)$.
- (iv) $\det(f) \neq 0 \iff f$ invertierbar.
- (v) $\det(f^{-1}) = (\det f)^{-1}$, falls f invertierbar ist.
- (vi) $\det(f^*) = \det(f)$ für die duale Abbildung $f^*: V^* \longrightarrow V^*$ zu f .

Bevor wir zum Beweis kommen, formulieren wir die entsprechenden Aussagen auch noch für Determinanten von Matrizen.

Satz 4. Für $A, B \in K^{n \times n}$ und $\alpha \in K$ gilt:

- (i) $\det(E) = 1$ für die Einheitsmatrix $E \in K^{n \times n}$.
- (ii) $\det(\alpha A) = \alpha^n \cdot \det(A)$.
- (iii) $\det(A \cdot B) = \det(A) \cdot \det(B)$.
- (iv) $\det(A) \neq 0 \iff A$ invertierbar.
- (v) $\det(A^{-1}) = (\det A)^{-1}$, falls A invertierbar ist.
- (vi) $\det(A^t) = \det(A)$ für die transponierte Matrix A^t zu A .

Beweis zu den Sätzen 3 und 4. Indem wir zwischen Endomorphismen und zugehörigen Matrizen wechseln, vgl. Abschnitt 3.1 sowie Satz 3.2/9, genügt es, alternativ entweder die Aussagen aus Satz 3 oder Satz 4 zu beweisen. Zudem ist nur der Fall $n = \dim V > 0$ von Interesse.

(i) $\det E = 1$ haben wir bereits in 4.2/4 gezeigt.

(ii) Man benutze die Multilinearität von Determinantenfunktionen: Sind $a_1, \dots, a_n$ die Spalten einer Matrix $A \in K^{n \times n}$, so folgt

$$\det(\alpha A) = \det(\alpha a_1, \dots, \alpha a_n) = \alpha^n \det(a_1, \dots, a_n) = \alpha^n \det(A).$$

(iii) Für diese Aussage ist es günstiger, im Sinne von Determinanten von Endomorphismen zu argumentieren. Man wähle eine nicht-triviale Determinantenfunktion Δ auf V sowie eine Basis $X = (x_1, \dots, x_n)$ von V . Dann gilt

$$\begin{aligned} \det(f \circ g) \cdot \Delta(x_1, \dots, x_n) &= \Delta(f \circ g(x_1), \dots, f \circ g(x_n)) \\ &= \det f \cdot \Delta(g(x_1), \dots, g(x_n)) \\ &= \det f \cdot \det g \cdot \Delta(x_1, \dots, x_n) \end{aligned}$$

und damit $\det(f \circ g) = \det f \cdot \det g$ wegen $\Delta(x_1, \dots, x_n) \neq 0$.

(iv) Eine Matrix $A \in K^{n \times n}$ ist genau dann invertierbar, wenn ihr Rang n ist, vgl. 3.3/7, d. h. genau dann, wenn die Spalten von A linear unabhängig sind. Damit ergibt sich die Aussage als Konsequenz von 4.2/4 und 4.2/7.

(v) Für $A \in \text{Gl}(n, K)$ gilt nach (i) und (iii)

$$\det(A) \cdot \det(A^{-1}) = \det(A \cdot A^{-1}) = \det(E) = 1.$$

(vi) Für $A = (\alpha_{ij})_{i,j} \in K^{n \times n}$ gilt

$$\begin{aligned}
 \det(A) &= \sum_{\pi \in \mathfrak{S}_n} \operatorname{sgn} \pi \cdot \alpha_{\pi(1),1} \cdot \dots \cdot \alpha_{\pi(n),n} \\
 &= \sum_{\pi \in \mathfrak{S}_n} \operatorname{sgn} \pi \cdot \alpha_{1,\pi^{-1}(1)} \cdot \dots \cdot \alpha_{n,\pi^{-1}(n)} \\
 &= \sum_{\pi \in \mathfrak{S}_n} \operatorname{sgn} \pi^{-1} \cdot \alpha_{1,\pi^{-1}(1)} \cdot \dots \cdot \alpha_{n,\pi^{-1}(n)} \\
 &= \sum_{\pi \in \mathfrak{S}_n} \operatorname{sgn} \pi \cdot \alpha_{1,\pi(1)} \cdot \dots \cdot \alpha_{n,\pi(n)} \\
 &= \det(A^t).
 \end{aligned}$$

□

Korollar 5. *Es sei $A \in K^{m \times n}$ eine Matrix vom Rang r . Dann lässt sich aus A durch Streichen von Spalten und Zeilen eine quadratische Untermatrix $A' \in K^{r \times r}$ mit $\det(A') \neq 0$ konstruieren, und es ist r maximal mit dieser Eigenschaft.*

Beweis. Gemäß 3.2/10 dürfen wir den Rang von Matrizen wahlweise als Spalten- oder Zeilenrang interpretieren. Insbesondere können wir in A ein linear unabhängiges System von r Spalten auswählen. Durch Streichen der restlichen Spalten entsteht eine Untermatrix $A_1 \in K^{m \times r}$ von A mit Rang r . Entsprechend können wir in A_1 ein linear unabhängiges System von r Zeilen auswählen. Indem wir die restlichen Zeilen von A_1 streichen, entsteht eine quadratische Untermatrix $A' \in K^{r \times r}$ von A_1 bzw. A mit Rang r . Diese ist nach 3.3/7 invertierbar und erfüllt $\det(A') \neq 0$ gemäß Satz 4.

Ist andererseits $A' \in K^{s \times s}$ eine quadratische Untermatrix von A mit einer Spaltenzahl $s > r = \operatorname{rg} A$, so sind die Spalten von A' notwendigerweise linear abhängig, denn jeweils s Spalten von A sind linear abhängig. Es folgt dann $\det(A') = 0$, ebenfalls mit Satz 4. □

Wir gehen schließlich noch auf die Berechnung der Determinanten spezieller Matrizen $A \in K^{n \times n}$ ein und untersuchen zunächst, wie sich die Determinante ändert, wenn man auf A elementare Zeilen- oder Spaltentransformationen anwendet.

Satz 6. *Es seien $\alpha \in K$, $A \in K^{n \times n}$, sowie i, j zwei verschiedene Indizes mit $1 \leq i, j \leq n$.*

(i) *$\det A$ ändert sich nicht, wenn man zu der i -ten Zeile (bzw. Spalte) von A das α -fache der j -ten Zeile (bzw. Spalte) addiert.*

(ii) *$\det A$ ändert das Vorzeichen, wenn man in A die i -te mit der j -ten Zeile (bzw. Spalte) vertauscht.*

(iii) *$\det A$ multipliziert sich mit α , wenn man die i -te Zeile (bzw. Spalte) von A mit α multipliziert.*

Beweis. Nach Satz 4 gilt $\det A = \det A^t$. Wir können uns daher auf die Betrachtung elementarer Spaltenumformungen beschränken. Dann ergeben sich die behaupteten Aussagen jedoch unmittelbar aus der Tatsache, dass $\det A$ eine Determinantenfunktion in den Spalten von A ist; vgl. 4.2/3 und 4.2/4. $\square$

Der obige Satz bietet eine natürliche Möglichkeit, Determinanten von Matrizen $A \in K^{n \times n}$ explizit zu berechnen: Man bestimme zunächst mittels des Gaußschen Eliminationsverfahrens den Rang $r = \operatorname{rg} A$. Für $r < n$ folgt dann $\det A = 0$ gemäß 4.2/2. Im Falle $r = n$ jedoch lässt sich A , wie wir in Abschnitt 3.3 gesehen haben, mittels elementarer Zeilenumformungen in die Einheitsmatrix E überführen, wobei $\det E = 1$ gilt. Benötigt man s solcher Umformungen (wobei man elementare Zeilen- und Spaltenumformungen gemischt verwenden darf), so ändert sich der Wert der Determinante bei jedem Schritt um einen gewissen Faktor α_σ , $\sigma = 1, \dots, s$, und es ergibt sich $\det A = \alpha_1^{-1} \dots \alpha_s^{-1}$.

Wir wollen einige Beispiele zur Berechnung spezieller Determinanten geben und dabei auch zeigen, wie das gerade beschriebene Verfahren in konkreten Situationen angewendet werden kann.

(1) Sei $A = (\alpha_{ij})_{i,j} \in K^{n \times n}$ mit $\alpha_{ij} = 0$ für $i > j$, also

$$A = \begin{pmatrix} \alpha_{11} & \dots & \dots & * \\ 0 & \alpha_{22} & \dots & * \\ \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \alpha_{nn} \end{pmatrix}.$$

Dann gilt $\det(A) = \prod_{i=1}^n \alpha_{ii}$. In der Tat, betrachten wir einen Term der Summe

$$\det(A) = \sum_{\pi \in \mathfrak{S}_n} \operatorname{sgn} \pi \cdot \alpha_{\pi(1),1} \cdot \dots \cdot \alpha_{\pi(n),n}$$

zu einer Permutation $\pi \in \mathfrak{S}_n$, so verschwindet dieser aufgrund unserer Voraussetzung über die α_{ij} , sofern es einen Index $i \in \{1, \dots, n\}$ mit $\pi(i) > i$ gibt. Da andererseits aber $\pi(i) \leq i$, $i = 1, \dots, n$, nur für die identische Permutation gilt, reduziert sich obige Summe auf einen Term, und zwar auf $\alpha_{11} \cdot \dots \cdot \alpha_{nn}$.

Alternativ können wir im Sinne elementarer Zeilenumformungen argumentieren. Gilt $\alpha_{ii} = 0$ für einen Index $i \in \{1, \dots, n\}$, so sind die Spalten mit den Indizes $1, \dots, i$ offenbar linear abhängig, und es gilt $\det A = 0$. Ansonsten multipliziere man für $i = 1, \dots, n$ die i -te Zeile von A mit α_{ii}^{-1} , wobei sich die Determinante jeweils um den Faktor α_{ii}^{-1} ändert. Die resultierende Matrix lässt sich dann in die Einheitsmatrix überführen, indem man für $i = 2, \dots, n$ geeignete Vielfache der i -ten Zeile zu den vorhergehenden Zeilen addiert. Der Wert der Determinante bleibt dabei unverändert, so dass man wie gewünscht $\det A = \alpha_{11} \dots \alpha_{nn}$ erhält.

(2) Zu $m, n \in \mathbb{N}$ betrachte man Matrizen $A_{11} \in K^{m \times m}$, $A_{12} \in K^{m \times n}$, $A_{22} \in K^{n \times n}$ sowie die Nullmatrix $0 \in K^{n \times m}$. Für die zusammengesetzte Matrix

$$A = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}$$

gilt dann $\det(A) = \det(A_{11}) \cdot \det(A_{22})$. Um dies nachzuweisen, nehmen wir A von der Form $(\alpha_{ij})_{i,j=1,\dots,m+n}$ an. Es gilt dann $\alpha_{ij} = 0$ für $m+1 \leq i \leq m+n$, $1 \leq j \leq m$. In der Summe

$$\det(A) = \sum_{\pi \in \mathfrak{S}_{m+n}} \operatorname{sgn} \pi \cdot \alpha_{\pi(1),1} \cdot \dots \cdot \alpha_{\pi(m+n),m+n}.$$

haben wir daher nur Terme zu solchen Permutationen $\pi \in \mathfrak{S}_n$ zu berücksichtigen, welche $\pi(i) \leq m$ für $i = 1, \dots, m$ erfüllen. Dann gilt automatisch $m+1 \leq \pi(i) \leq m+n$ für $i = m+1, \dots, m+n$, und es “zerfällt” π in Permutationen σ von $\{1, \dots, m\}$ und τ von $\{m+1, \dots, m+n\}$, wobei $\operatorname{sgn} \pi = \operatorname{sgn}(\sigma) \cdot \operatorname{sgn}(\tau)$ gilt. Da auf diese Weise die zu betrachtenden Permutationen $\pi \in \mathfrak{S}_{m+n}$ bijektiv den Paaren $(\sigma, \tau) \in \mathfrak{S}_m \times \mathfrak{S}_n$ entsprechen, ergibt sich mit $A_{22} = (\beta_{ij})_{i,j=1,\dots,n}$

$$\begin{aligned} \det(A) &= \sum_{\sigma \in \mathfrak{S}_m, \tau \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) \operatorname{sgn}(\tau) \alpha_{\sigma(1),1} \dots \alpha_{\sigma(m),m} \cdot \beta_{\tau(1),1} \dots \beta_{\tau(n),n} \\ &= \sum_{\sigma \in \mathfrak{S}_m} \operatorname{sgn}(\sigma) \alpha_{\sigma(1),1} \dots \alpha_{\sigma(m),m} \cdot \sum_{\tau \in \mathfrak{S}_n} \operatorname{sgn}(\tau) \beta_{\tau(1),1} \dots \beta_{\tau(n),n} \\ &= \det(A_{11}) \cdot \det(A_{22}), \end{aligned}$$

wie behauptet.

Alternativ kann man dieses Resultat allerdings auch ohne größere Rechnung erhalten, indem man die Theorie der Determinantenfunktionen verwendet. Seien $E_m \in K^{m \times m}$ und $E_n \in K^{n \times n}$ die Einheitsmatrizen. Wir benutzen, dass

$$\det(A) = \det \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}$$

einerseits eine Determinantenfunktion in den Spalten von A_{11} ist, sowie andererseits auch in den Zeilen von A_{22} . Mit 4.2/5 ergibt sich daher die Gleichung

$$\det \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix} = \det(A_{11}) \cdot \det \begin{pmatrix} E_m & A_{12} \\ 0 & A_{22} \end{pmatrix}$$

und entsprechend für $A_{11} = E_m$ die Gleichung

$$\det \begin{pmatrix} E_m & A_{12} \\ 0 & A_{22} \end{pmatrix} = \det(A_{22}) \cdot \det \begin{pmatrix} E_m & A_{12} \\ 0 & E_n \end{pmatrix}.$$

Da die Matrix

$$\begin{pmatrix} E_m & A_{12} \\ 0 & E_n \end{pmatrix}$$

nach Beispiel (i) die Determinante 1 besitzt, folgt wie gewünscht

$$\det(A) = \det(A_{11}) \cdot \det(A_{22}).$$

Schließlich wollen wir noch zeigen, wie man mittels elementarer Umformungen argumentieren kann. Gilt $\operatorname{rg} A_{11} < m$, so kann A nicht maximalen Rang

besitzen, da dann die ersten m Spalten von A linear abhängig sind. Entsprechendes gilt für $\text{rg } A_{22} < n$, da in diesem Falle die letzten n Zeilen von A linear abhängig sind. Insgesamt ergibt sich $\det A = 0 = \det A_{11} \cdot \det A_{22}$, falls mindestens eine der Matrizen A_{11} und A_{22} nicht invertierbar ist. Sind jedoch A_{11} und A_{22} beide invertierbar, so wende man auf die ersten m Zeilen von A elementare Zeilenumformungen an, welche die Untermatrix A_{11} in die Einheitsmatrix überführen. Die Determinante von A ändert sich hierbei offenbar um den Faktor $(\det A_{11})^{-1}$. Entsprechend wende man auf die letzten n Zeilen von A elementare Zeilenumformungen an, die die Untermatrix A_{22} in die Einheitsmatrix überführen. Dies bewirkt eine Änderung der Determinante um den Faktor $(\det A_{22})^{-1}$. Die resultierende Matrix besitzt dann die in Beispiel (1) betrachtete Form, wobei die Diagonalelemente alle den Wert 1 haben. Somit ist die Determinante dieser Matrix 1, und es folgt $\det A = \det A_{11} \cdot \det A_{22}$.

(3) Ein berühmtes Beispiel einer konkret gegebenen Determinante ist die so genannte *Vandermondesche Determinante*

$$V(\alpha_1, \alpha_2, \dots, \alpha_n) = \det \begin{pmatrix} 1 & 1 & \dots & 1 \\ \alpha_1 & \alpha_2 & \dots & \alpha_n \\ \alpha_1^2 & \alpha_2^2 & \dots & \alpha_n^2 \\ \vdots & \vdots & \dots & \vdots \\ \alpha_1^{n-1} & \alpha_2^{n-1} & \dots & \alpha_n^{n-1} \end{pmatrix},$$

welche zu Elementen $\alpha_1, \alpha_2, \dots, \alpha_n \in K$ gebildet wird. Wir wollen induktiv

$$V(\alpha_1, \alpha_2, \dots, \alpha_n) = \prod_{i < j} (\alpha_j - \alpha_i)$$

zeigen. Für $n = 1$ ist die Behauptung trivial. Für $n > 1$ subtrahiere man in obiger Matrix von der n -ten Zeile das α_1 -fache der $(n-1)$ -ten Zeile usw. bis schließlich von der 2. Zeile das α_1 -fache der 1. Zeile. Dann gilt

$$V(\alpha_1, \alpha_2, \dots, \alpha_n) = \det \begin{pmatrix} 1 & 1 & \dots & 1 \\ 0 & \alpha_2 - \alpha_1 & \dots & \alpha_n - \alpha_1 \\ 0 & \alpha_2(\alpha_2 - \alpha_1) & \dots & \alpha_n(\alpha_n - \alpha_1) \\ \vdots & \vdots & \dots & \vdots \\ 0 & \alpha_2^{n-2}(\alpha_2 - \alpha_1) & \dots & \alpha_n^{n-2}(\alpha_n - \alpha_1) \end{pmatrix}$$

und es folgt mit obigem Beispiel (2)

$$V(\alpha_1, \alpha_2, \dots, \alpha_n) = \left(\prod_{i > 1} (\alpha_i - \alpha_1) \right) V(\alpha_2, \dots, \alpha_n).$$

Per Induktion ergibt sich daraus die Behauptung.

Schließlich wollen für Matrizen $A = (\alpha_{ij})_{i,j} \in K^{n \times n}$, $n \leq 3$, deren Determinanten noch explizit angeben.

$$\begin{aligned}
n = 1 : \quad \det(A) &= \alpha_{11} \\
n = 2 : \quad \det(A) &= \alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12} \\
n = 3 : \quad \det(A) &= \alpha_{11}\alpha_{22}\alpha_{33} + \alpha_{21}\alpha_{32}\alpha_{13} + \alpha_{31}\alpha_{12}\alpha_{23} \\
&\quad - \alpha_{11}\alpha_{32}\alpha_{23} - \alpha_{31}\alpha_{22}\alpha_{13} - \alpha_{21}\alpha_{12}\alpha_{33}
\end{aligned}$$

Letztere Formel entspricht der Auflistung

$$\mathfrak{S}_3 = \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \right. \\
\left. \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \right\}$$

Die Art der Summanden kann man sich dabei an folgendem Schema (bezeichnet als Regel von Sarrus) in Erinnerung bringen:

$$\begin{array}{ccccccc}
\alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{11} & \alpha_{12} & & \\
\alpha_{21} & \alpha_{22} & \alpha_{23} & \alpha_{21} & \alpha_{22} & & \\
\alpha_{31} & \alpha_{32} & \alpha_{33} & \alpha_{31} & \alpha_{32} & &
\end{array}$$

(Note: In the original image, the last two columns are connected by dashed lines to the first two columns, representing the cyclic nature of the Sarrus rule.)

Man hat also zur Berechnung von $\det(\alpha_{ij})_{i,j=1,2,3}$ die Produkte über die Elemente der 3 von links oben nach rechts unten verlaufenden Diagonalen zu addieren und davon die Produkte über die Elemente der 3 von links unten nach rechts oben verlaufenden Diagonalen zu subtrahieren.

Aufgaben

1. Für $A \in K^{n \times n}$ mit $A^t = -A$ und ungeradem n zeige man: Es gilt $\det(A) = 0$ oder $1 + 1 = 0$ in K .
2. Für Matrizen $A \in K^{n \times n}$ lasse man folgende Zeilenoperationen zu:
 - (i) Vertauschen zweier verschiedener Zeilen und gleichzeitiges Multiplizieren einer Zeile mit -1 .
 - (ii) Multiplizieren einer Zeile mit einer Konstanten $\alpha \in K^*$ und gleichzeitiges Multiplizieren einer weiteren Zeile mit α^{-1} .
 - (iii) Addieren eines Vielfachen einer Zeile zu einer anderen.

Man zeige, dass sich invertierbare Matrizen $A \in K^{n \times n}$ mittels solcher Zeilenumformungen in Diagonalmatrizen des Typs

$$\begin{pmatrix} 1 & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & 1 & 0 \\ 0 & \dots & 0 & d \end{pmatrix}$$

überführen lassen, wobei $d = \det A$ gilt.

3. Sei $A = (\alpha_{ij}) \in K^{n \times n}$ mit

$$\alpha_{ij} = \begin{cases} 1 & \text{für } i < j \\ 0 & \text{für } i = j \\ -1 & \text{für } i > j \end{cases}$$

und n gerade. Man zeige $\det A = 1$.

4. Man berechne die Determinante der Matrix

$$\begin{pmatrix} 2 & 1 & 1 & \dots & 1 \\ 1 & 2 & 1 & \dots & 1 \\ 1 & 1 & 2 & \dots & 1 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & 1 & \dots & 2 \end{pmatrix} \in \mathbb{R}^{n \times n}.$$

5. Es sei V ein n -dimensionaler K -Vektorraum und $f: V \rightarrow V$ ein Endomorphismus. Für ein $x \in V$ gelte $V = \langle x, f(x), \dots, f^{n-1}(x) \rangle$ sowie

$$f^n(x) = a_{n-1}f^{n-1}(x) + \dots + a_1f(x) + a_0x.$$

Man zeige $\det f = (-1)^{n+1}a_0$.

6. Es seien $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n \in K$ Konstanten mit $\alpha_i + \beta_j \neq 0$ für alle i, j . Man zeige (*Cauchysche Determinante*):

$$\det \begin{pmatrix} (\alpha_1 + \beta_1)^{-1} & \dots & (\alpha_1 + \beta_n)^{-1} \\ \dots & & \dots \\ (\alpha_n + \beta_1)^{-1} & \dots & (\alpha_n + \beta_n)^{-1} \end{pmatrix} = \frac{\prod_{i < j} (\alpha_j - \alpha_i)(\beta_j - \beta_i)}{\prod_{i,j} (\alpha_i + \beta_j)}$$

4.4 Die Cramersche Regel

Wir wollen zunächst zu einer Matrix $A = (\alpha_{ij})_{i,j} \in K^{n \times n}$ die so genannte *adjungierte* Matrix A^{ad} erklären. Hierzu konstruieren wir für beliebige Indizes $i, j \in \{1, \dots, n\}$ eine $(n \times n)$ -Matrix A_{ij} aus A , indem wir alle Elemente $\alpha_{i1}, \dots, \alpha_{in}$ der i -ten Zeile und alle Elemente $\alpha_{1j}, \dots, \alpha_{nj}$ der j -ten Spalte von A durch 0 ersetzen, bis auf das Element α_{ij} im Schnittpunkt von i -ter Zeile und j -ter Spalte, welches wir zu 1 abändern, also:

$$A_{ij} = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1,j-1} & 0 & \alpha_{1,j+1} & \dots & \alpha_{1n} \\ \dots & & \dots & \dots & \dots & & \dots \\ \alpha_{i-1,1} & \dots & \alpha_{i-1,j-1} & 0 & \alpha_{i-1,j+1} & \dots & \alpha_{i-1,n} \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ \alpha_{i+1,1} & \dots & \alpha_{i+1,j-1} & 0 & \alpha_{i+1,j+1} & \dots & \alpha_{i+1,n} \\ \dots & \dots & \dots & \dots & \dots & & \dots \\ \alpha_{n1} & \dots & \alpha_{n,j-1} & 0 & \alpha_{n,j+1} & \dots & \alpha_{nn} \end{pmatrix}$$

Weiter werde die Matrix $A'_{ij} \in K^{(n-1) \times (n-1)}$ durch Streichen der i -ten Zeile und der j -ten Spalte von A_{ij} erklärt.

Bemerkung 1. Es sei $A \in K^{n \times n}$ eine Matrix mit den Spalten $a_1, \dots, a_n \in K^n$. Dann gilt für Indizes $1 \leq i, j \leq n$:

- (i) $\det(A_{ij}) = (-1)^{i+j} \cdot \det(A'_{ij})$.
- (ii) $\det(A_{ij}) = \det(a_1, \dots, a_{j-1}, e_i, a_{j+1}, \dots, a_n)$, wobei $e_i = (\delta_{1i}, \dots, \delta_{ni})^t$ der i -te Einheitsvektor in K^n ist.

Beweis. Für $i = j = 1$ ist A_{ij} von der Gestalt

$$A_{11} = \begin{pmatrix} 1 & 0 \\ 0 & A'_{11} \end{pmatrix},$$

und wir erhalten, wie in Abschnitt 4.3 berechnet,

$$\det A_{11} = 1 \cdot \det A'_{11}.$$

Den Allgemeinfall von (i) führt man leicht hierauf zurück, indem man in A_{ij} die i -te Zeile mit allen vorhergehenden Zeilen sowie die j -te Spalte mit allen vorhergehenden Spalten vertauscht. Dies sind insgesamt $(i-1) + (j-1)$ Vertauschungen, so dass sich die Determinante der neuen Matrix dabei von $\det A_{ij}$ um den Faktor $(-1)^{i+j}$ unterscheidet.

Weiter ergibt sich (ii) aus 4.3/6(i), da sich die Matrix mit den Spalten $a_1, \dots, a_{j-1}, e_i, a_{j+1}, \dots, a_n$ in die Matrix A_{ij} überführen lässt, indem man geeignete Vielfache der j -ten Spalte e_i von den restlichen subtrahiert. $\square$

Definition 2. Für eine Matrix $A = (\alpha_{ij})_{i,j} \in K^{n \times n}$ und für Indizes $1 \leq i, j \leq n$ bilde man die Matrizen A_{ij}, A'_{ij} wie vorstehend beschrieben. Dann heißt

$$\tilde{\alpha}_{ij} = \det A_{ij} = (-1)^{i+j} \cdot \det A'_{ij}$$

der Cofaktor von A zum Indexpaar (i, j) oder, in nicht ganz korrekter Sprechweise, der Cofaktor von A zum Koeffizienten α_{ij} . Die Matrix $\tilde{A} = (\tilde{\alpha}_{ij})_{i,j}$ heißt die Matrix der Cofaktoren von A , und

$$A^{\text{ad}} := \tilde{A}^t = (\tilde{\alpha}_{ji})_{i,j}$$

die zu A adjungierte Matrix (oder die zu A gehörige Komplementärmatrix).

Als Nächstes wollen wir eine fundamentale Beziehung zwischen einer Matrix A und ihrer adjungierten A^{ad} beweisen, welche im weiteren Sinne als so genannte *Cramersche Regel* bezeichnet wird.

Satz 3. Sei $A = (\alpha_{ij})_{i,j} \in K^{n \times n}$, und sei $E \in K^{n \times n}$ die Einheitsmatrix. Dann gilt

$$A^{\text{ad}} \cdot A = A \cdot A^{\text{ad}} = (\det A) \cdot E.$$

In ausführlicher Schreibweise bedeutet diese Identität:

$$\sum_{j=1}^n \tilde{\alpha}_{ji} \alpha_{jk} = \delta_{ik} \cdot \det A,$$

$$\sum_{j=1}^n \alpha_{ij} \tilde{\alpha}_{kj} = \delta_{ik} \cdot \det A,$$

für $1 \leq i, k \leq n$.

Beweis. Es seien $a_1, \dots, a_n$ die n Spalten von A . Dann rechnet man unter Ausnutzung von Bemerkung 1

$$\begin{aligned} \sum_{j=1}^n \tilde{\alpha}_{ji} \alpha_{jk} &= \sum_{j=1}^n \alpha_{jk} \cdot \det(A_{ji}) \\ &= \sum_{j=1}^n \alpha_{jk} \cdot \det(a_1, \dots, a_{i-1}, e_j, a_{i+1}, \dots, a_n) \\ &= \det(a_1, \dots, a_{i-1}, a_k, a_{i+1}, \dots, a_n) \\ &= \delta_{ik} \cdot \det(A), \end{aligned}$$

d. h. es gilt $\tilde{A}^t \cdot A = (\det A) \cdot E$. Um zu sehen, dass dies auch gleich $A \cdot \tilde{A}^t$ ist, benutze man die gerade bewiesene Beziehung für A^t anstelle von A , also

$$(\tilde{A}^t)^t \cdot A^t = (\det A^t) \cdot E.$$

Da man offenbar $\tilde{A}^t = \tilde{A}^t$ hat, ergibt sich

$$\tilde{A} \cdot A^t = (\det A) \cdot E$$

unter Benutzung von $\det(A^t) = \det(A)$, vgl. 4.3/4. Transponieren liefert dann wie gewünscht

$$A \cdot \tilde{A}^t = (\det A) \cdot E.$$

□

Wir wollen einige Folgerungen aus diesem Satz ziehen. Zunächst erhält man als Spezialfall der Gleichungen in Satz 3 den nach Laplace benannten Satz zum Entwickeln einer Determinante nach einer Zeile oder Spalte.

Korollar 4. Sei $A = (\alpha_{ij})_{i,j} \in K^{n \times n}$. Dann gilt:

$$\det A = \sum_{i=1}^n \alpha_{ij} \cdot \det(A_{ij}) = \sum_{i=1}^n (-1)^{i+j} \cdot \alpha_{ij} \cdot \det(A'_{ij}), \quad j = 1, \dots, n,$$

$$\det A = \sum_{j=1}^n \alpha_{ij} \cdot \det(A_{ij}) = \sum_{j=1}^n (-1)^{i+j} \cdot \alpha_{ij} \cdot \det(A'_{ij}), \quad i = 1, \dots, n$$

Man bezeichnet dies als die Formeln zur Entwicklung von $\det A$ nach der j -ten Spalte bzw. i -ten Zeile.

Für invertierbare Matrizen kann man die Formeln in Satz 3 zur Berechnung der inversen Matrix benutzen.

Korollar 5. Für $A \in \text{Gl}(n, K)$ gilt $A^{-1} = (\det A)^{-1} \cdot A^{\text{ad}}$.

Des Weiteren wollen wir noch die Cramersche Regel im engeren Sinne herleiten, die sich auf die Lösung linearer Gleichungssysteme bezieht.

Korollar 6. Für eine Matrix $A \in \text{Gl}(n, K)$ und einen Spaltenvektor $b \in K^n$ betrachte man das lineare Gleichungssystem $A \cdot x = b$. Dann ist $x = A^{-1} \cdot b$ die eindeutige Lösung dieses Systems. Sind $a_1, \dots, a_n$ die Spalten von A , so wird die i -te Komponente von $A^{-1} \cdot b$ gegeben durch

$$(\det A)^{-1} \cdot \det(a_1, \dots, a_{i-1}, b, a_{i+1}, \dots, a_n).$$

Beweis. Wir benutzen die Formel $A^{-1} = (\det A)^{-1} \cdot \tilde{A}^t$ aus Korollar 5. Dann berechnet sich die i -te Komponente von $A^{-1} \cdot b$ für $b = (\beta_1, \dots, \beta_n)^t$ mittels Bemerkung 1 zu

$$\begin{aligned} (\det A)^{-1} \cdot \sum_{j=1}^n \tilde{\alpha}_{ji} \beta_j &= (\det A)^{-1} \cdot \sum_{j=1}^n (\det A_{ji}) \beta_j \\ &= (\det A)^{-1} \cdot \sum_{j=1}^n \det(a_1, \dots, a_{i-1}, \beta_j e_j, a_{i+1}, \dots, a_n) \\ &= (\det A)^{-1} \cdot \det(a_1, \dots, a_{i-1}, b, a_{i+1}, \dots, a_n). \end{aligned}$$

□

Aufgaben

1. Man löse das folgende lineare Gleichungssystem mittels Cramerscher Regel:

$$2x_1 + 2x_2 + 4x_3 = 1$$

$$2x_1 - 3x_2 + 2x_3 = 0$$

$$5x_1 + 4x_2 + 3x_3 = 1$$

2. Es sei $A = (\alpha_{ij}) \in \text{Gl}(n, \mathbb{R})$ eine invertierbare Matrix mit Koeffizienten $\alpha_{ij} \in \mathbb{Z}$. Man zeige:

(i) A^{-1} hat Koeffizienten in $\mathbb{Q}$.

(ii) A^{-1} hat genau dann Koeffizienten in $\mathbb{Z}$, wenn $\det A = \pm 1$ gilt.

3. Man zeige für invertierbare Matrizen $A, B \in \text{Gl}(n, K)$ und $\alpha \in K$:

(i) $(\alpha A)^{\text{ad}} = \alpha^{n-1} A^{\text{ad}}$

(ii) $(A^t)^{\text{ad}} = (A^{\text{ad}})^t$

(iii) $(AB)^{\text{ad}} = B^{\text{ad}} A^{\text{ad}}$

- (iv) $\det(A^{\text{ad}}) = (\det A)^{n-1}$
- (v) $(A^{\text{ad}})^{\text{ad}} = (\det A)^{n-2} A$
- (vi) $(A^{\text{ad}})^{-1} = (A^{-1})^{\text{ad}}$

Bemerkung: Alle vorstehenden Beziehungen, mit Ausnahme der letzten, bleiben auch für beliebige Matrizen aus $K^{n \times n}$ gültig. Um dies aus dem Fall invertierbarer Matrizen zu folgern, sind allerdings Konstruktionen erforderlich, die im Moment noch nicht zur Verfügung stehen.

4. Für Spaltenvektoren $a_1, \dots, a_n, b \in K^n$ und Koeffizienten $\alpha_1, \dots, \alpha_n \in K$ gelte $\sum_{i=1}^n \alpha_i a_i = b$. Man zeige

$$\det(a_1, \dots, a_{i-1}, b, a_{i+1}, \dots, a_n) = \alpha_i \cdot \det(a_1, \dots, a_n), \quad i = 1, \dots, n,$$

und folgere hieraus die Aussage von Korollar 6.

4.5 Äußere Produkte*

In Abschnitt 4.2 haben wir für einen K -Vektorraum V endlicher Dimension n alternierende Multilinearformen in n Variablen studiert. Wir wollen nun allgemeiner für beliebige K -Vektorräume V alternierende multilineare Abbildungen in einer gewissen Anzahl r von Variablen betrachten. Dabei heißt eine Abbildung $\Phi: V^r \rightarrow W, (a_1, \dots, a_r) \mapsto \Phi(a_1, \dots, a_r)$, in einen K -Vektorraum W *multilinear*, wenn Φ linear in jedem Argument ist, sowie *alternierend*, wenn $\Phi(a_1, \dots, a_r) = 0$ gilt, sofern zwei der Vektoren $a_1, \dots, a_r \in V$ übereinstimmen. Entsprechend wie in 4.2/2 bzw. 4.2/3 zeigt man:

Bemerkung 1. Für K -Vektorräume V, W und $r \in \mathbb{N}$ betrachte man eine alternierende multilineare Abbildung $\Phi: V^r \rightarrow W$. Dann gilt für Vektoren $a_1, \dots, a_r \in V$:

- (i) $\Phi(a_1, \dots, a_r) = 0$, sofern $a_1, \dots, a_r$ linear abhängig sind.
- (ii) $\Phi(a_{\pi(1)}, \dots, a_{\pi(r)}) = \text{sgn } \pi \cdot \Phi(a_1, \dots, a_r)$ für jede Permutation $\pi \in \mathfrak{S}_r$.

Insbesondere folgt aus (i), dass es im Falle $r > \dim_K V$ nur triviale multilineare alternierende Abbildungen $V^r \rightarrow W$ gibt. Wir wollen im Folgenden zu V und $r \in \mathbb{N}$ einen K -Vektorraum $\bigwedge^r V$ konstruieren, derart dass die alternierenden multilinearen Abbildungen $V^r \rightarrow W$ in einen beliebigen K -Vektorraum W hinein in bijektiver Weise den K -linearen Abbildungen $\bigwedge^r V \rightarrow W$ entsprechen.

Satz 2. Zu einem K -Vektorraum V und einer natürlichen Zahl $r \in \mathbb{N}$ gibt es stets einen K -Vektorraum D mit einer multilinearen alternierenden Abbildung $\sigma: V^r \rightarrow D$, welche folgende universelle Eigenschaft besitzt:

Ist $\Phi: V^r \rightarrow W$ eine alternierende multilineare Abbildung in einen K -Vektorraum W , so existiert eindeutig eine K -lineare Abbildung $\varphi: D \rightarrow W$ mit $\Phi = \varphi \circ \sigma$, so dass also das Diagramm

$$\begin{array}{ccc}
 V^r & \xrightarrow{\sigma} & D \\
 \searrow \Phi & & \swarrow \varphi \\
 & W &
 \end{array}$$

kommutiert.

Das Paar (D, σ) ist durch diese Abbildungseigenschaft bis auf kanonische Isomorphie eindeutig bestimmt.

Dass in der Situation des Satzes zwei Paare (D, σ) und (D', σ') , welche die genannte universelle Abbildungseigenschaft besitzen, in kanonischer Weise isomorph sind, kann man leicht einsehen. Es gibt dann nämlich eine eindeutig bestimmte K -lineare Abbildung $\iota: D \rightarrow D'$ mit $\sigma' = \iota \circ \sigma$, sowie eine eindeutig bestimmte K -lineare Abbildung $\iota': D' \rightarrow D$ mit $\sigma = \iota' \circ \sigma'$. Also hat man

$$\sigma = \iota' \circ \sigma' = \iota' \circ \iota \circ \sigma, \quad \sigma' = \iota \circ \sigma = \iota \circ \iota' \circ \sigma',$$

und damit

$$\text{id}_D \circ \sigma = (\iota' \circ \iota) \circ \sigma, \quad \text{id}_{D'} \circ \sigma' = (\iota \circ \iota') \circ \sigma'.$$

Es sind deshalb $\text{id}_D, \iota' \circ \iota: D \rightarrow D$ zwei K -lineare Abbildungen, die durch Komposition mit $\sigma: V^r \rightarrow D$ dieselbe alternierende multilineare Abbildung $V^r \rightarrow D$ ergeben. Die Eindeutigkeitsaussage in der Abbildungseigenschaft für (D, σ) ergibt dann $\iota' \circ \iota = \text{id}_D$. Entsprechend erhält man $\iota \circ \iota' = \text{id}_{D'}$, und es folgt, dass ι und ι' zueinander inverse Isomorphismen sind, eben die kanonischen Isomorphismen, deren Existenz im Satz behauptet wird.

In der Situation von Satz 2, dessen Beweis wir weiter unten fortführen werden, nennt man den K -Vektorraum D das r -fache äußere Produkt oder die r -fache äußere Potenz von V und benutzt hierfür die Notation $\bigwedge^r V$. Das Bild eines Tupels $(a_1, \dots, a_r) \in V^r$ unter der multilinearen alternierenden Abbildung $\sigma: V^r \rightarrow \bigwedge^r V$ wird mit

$$a_1 \wedge \dots \wedge a_r$$

bezeichnet, wobei man hier auch von dem äußeren oder Dachprodukt der Elemente $a_1, \dots, a_r$ spricht. Diese Produktbildung ist linear und alternierend in den einzelnen Faktoren, d. h. es gilt für $\alpha, \beta \in K, a_1, \dots, a_r, b_i \in V$

$$\begin{aligned}
 & a_1 \wedge \dots \wedge (\alpha a_i + \beta b_i) \wedge \dots \wedge a_r \\
 &= \alpha \cdot (a_1 \wedge \dots \wedge a_i \wedge \dots \wedge a_r) + \beta \cdot (a_1 \wedge \dots \wedge b_i \wedge \dots \wedge a_r),
 \end{aligned}$$

sowie $a_1 \wedge \dots \wedge a_r = 0$, falls die Vektoren $a_1, \dots, a_r$ nicht paarweise verschieden sind. Wir können sogar aus der universellen Abbildungseigenschaft von $\bigwedge^r V$ folgende Information ableiten:

Bemerkung 3. Der K -Vektorraum $\bigwedge^r V$ wird von den Elementen $a_1 \wedge \dots \wedge a_r$ mit $a_1, \dots, a_r \in V$ erzeugt.

Beweis. Es sei $D' \subset \bigwedge^r V$ der lineare Unterraum, der von allen Elementen des Typs $a_1 \wedge \dots \wedge a_r$ erzeugt wird. Da das Bild von $\sigma: V^r \longrightarrow \bigwedge^r V$ in D' liegt, führt jede Zerlegung

$$\Phi: V^r \xrightarrow{\sigma} \bigwedge^r V \xrightarrow{\varphi} W$$

einer multilinearen alternierenden Abbildung $\Phi: V^r \longrightarrow W$ automatisch zu einer Zerlegung

$$\Phi: V^r \xrightarrow{\sigma} D' \xrightarrow{\varphi'} W$$

mit $\varphi' = \varphi|_{D'}$. Nun ist φ' aufgrund der Gleichung $\Phi = \varphi' \circ \sigma$ notwendigerweise eindeutig bestimmt auf dem Bild im σ , also auch auf dem von im σ erzeugten linearen Unterraum von D' , d. h. auf D' selbst. Somit erfüllt D' zusammen mit σ die universelle Eigenschaft eines äußeren Produkts, und man sieht wie oben, dass die Abbildung $D' \hookrightarrow \bigwedge^r V$ ein Isomorphismus ist, als Inklusion also die Identität darstellt. Somit bilden die Elemente des Typs $a_1 \wedge \dots \wedge a_r$ in der Tat ein Erzeugendensystem von $D' = \bigwedge^r V$. $\square$

Wir erleben hier zum ersten Mal, dass mathematische Objekte, nämlich der K -Vektorraum $\bigwedge^r V$ und die alternierende multilineare Abbildung

$$\sigma: V^r \longrightarrow \bigwedge^r V, \quad (a_1, \dots, a_r) \longmapsto a_1 \wedge \dots \wedge a_r,$$

nicht in konkreter Weise, sondern nur bis auf kanonische Isomorphie definiert werden. Natürlich muss noch gezeigt werden, dass Objekte mit den spezifizierten Eigenschaften auch wirklich existieren. Wir werden hierfür ein explizites Konstruktionsverfahren verwenden, das jedoch wegen seiner allgemeinen Natur nur von untergeordnetem Interesse sein kann. Stattdessen ist die charakterisierende universelle Eigenschaft das entscheidende Hilfsmittel zur Handhabung des Vektorraums $\bigwedge^r V$.

Um nun den *Beweis zu Satz 2* abzuschließen, bleibt noch die Existenz eines K -Vektorraums D zusammen mit einer alternierenden multilinearen Abbildung $\sigma: V^r \longrightarrow D$ nachzuweisen, welche die behauptete universelle Abbildungseigenschaft besitzt. Hierzu betrachten wir, motiviert durch die Aussage von Bemerkung 3, einen K -Vektorraum $\hat{D}$, der V^r als erzeugende Teilmenge enthält. Am einfachsten ist es, den freien von allen Symbolen $v \in V^r$ erzeugten K -Vektorraum zu betrachten, also

$$\hat{D} = K^{(V^r)} = \{(\alpha_v)_{v \in V^r}; \alpha_v \in K, \alpha_v = 0 \text{ für fast alle } v\}.$$

Hier bildet das System der Elemente

$$e_v = (\delta_{v,v'})_{v' \in V^r}, \quad v \in V^r,$$

eine Basis, wobei wir für Tupel $v = (a_1, \dots, a_r) \in V^r$ anstelle von e_v auch ausführlicher $e_{(a_1, \dots, a_r)}$ schreiben werden.

Sei $R \subset \hat{D}$ der lineare Unterraum, der von allen Elementen der Gestalt

$$e_{(\dots, \alpha a + \beta b, \dots)} - \alpha \cdot e_{(\dots, a, \dots)} - \beta \cdot e_{(\dots, b, \dots)}, \\ e_{(\dots, a, \dots, a, \dots)}$$

mit $\alpha, \beta \in K$, $a, b \in V$ erzeugt wird (wobei bei den Elementen ersteren Typs die Einträge an einer festen Stelle i stehen und die restlichen Einträge in allen drei Termen jeweils unverändert sind). Man setze dann $D = \hat{D}/R$ und bezeichne die Restklasse eines Basiselementes $e_{(a_1, \dots, a_r)} \in \hat{D}$ mit $a_1 \wedge \dots \wedge a_r$. Dann ist nach Definition von R die Abbildung

$$\sigma: V^r \longrightarrow D, \quad (a_1, \dots, a_r) \longmapsto e_{(a_1, \dots, a_r)} \longmapsto a_1 \wedge \dots \wedge a_r,$$

sozusagen erzwingen ermaßen multilinear und alternierend. Um zu erkennen, dass σ die behauptete universelle Abbildungseigenschaft besitzt, betrachte man eine alternierende multilineare Abbildung $\Phi: V^r \longrightarrow W$ in einen K -Vektorraum W hinein. Dann kann man eine Abbildung $\hat{\varphi}: \hat{D} \longrightarrow W$ erklären, indem man die Bilder der kanonischen Basis $(e_v)_{v \in V^r}$ von $\hat{D}$ vorgibt, und zwar durch

$$\hat{\varphi}(e_{(a_1, \dots, a_r)}) = \Phi(a_1, \dots, a_r), \quad (a_1, \dots, a_r) \in V^r.$$

Man vergleiche hierzu Satz 2.1/7 (ii), den man in einer Version für nicht notwendig endliche Basen benötigt. Es entsteht somit das folgende kommutative Diagramm, wobei die Existenz der Abbildung φ noch zu erklären ist:

$$\begin{array}{ccc} & \hat{D} & \\ & \uparrow \quad \downarrow \hat{\varphi} & \\ V^r & \xrightarrow{\sigma} & D = \hat{D}/R \\ & \downarrow \Phi \quad \uparrow \varphi & \\ & W & \end{array}$$

Aus der Eigenschaft, dass Φ multilinear und alternierend ist, ergibt sich sofort, dass $\ker \hat{\varphi}$ alle erzeugenden Elemente von R , also R selbst enthält. Mittels des Homomorphiesatzes 2.2/8 folgt dann, dass $\hat{\varphi}$ über den Quotienten $D = \hat{D}/R$ faktorisiert, sich also als Komposition der Projektion $\hat{D} \longrightarrow D$ und einer K -linearen Abbildung $\varphi: D \longrightarrow W$ schreiben lässt, wobei φ die Gleichung

$$(*) \quad \varphi(a_1 \wedge \dots \wedge a_r) = \Phi(a_1, \dots, a_r)$$

für $(a_1, \dots, a_r) \in V^r$ erfüllt; dies bedeutet aber $\Phi = \varphi \circ \sigma$.

Es bleibt nun noch einzusehen, dass φ durch die Gleichung $\Phi = \varphi \circ \sigma$ eindeutig bestimmt ist. Gilt diese Gleichung, d. h. gilt $(*)$ für alle Tupel in V^r , so ist φ als K -lineare Abbildung jedenfalls eindeutig bestimmt auf dem linearen Unterraum von D , der von allen Elementen des Typs $a_1 \wedge \dots \wedge a_r$ erzeugt wird; vgl. 2.1/7. Dieser Unterraum ist aber identisch mit D , da die Elemente $e_{(a_1, \dots, a_r)}$

nach unserer Konstruktion eine Basis von $\hat{D}$ bilden, die Elemente $a_1 \wedge \dots \wedge a_r$ also immerhin noch den Quotientenvektorraum D erzeugen. $\square$

Die vorstehende Konstruktion ergibt im Fall $r = 0$ einen eindimensionalen K -Vektorraum $\bigwedge^0 V$, der von dem Bild des leeren Tupels aus V^0 erzeugt wird. Eine solche Situation läßt sich natürlich konkretisieren: Man setzt $\bigwedge^0 V := K$, wobei man vereinbart, dass das Produkt $a_1 \wedge \dots \wedge a_r$ für $r = 0$ leer ist und demgemäß den Wert 1 hat.

Als Nächstes wollen wir überlegen, wie man, ausgehend von einer Basis $X = (x_1, \dots, x_n)$ eines K -Vektorraums V eine Basis von $\bigwedge^r V$ erhalten kann. Hierzu bezeichne Z_r^n für $n, r \in \mathbb{N}$ die Menge aller r -elementigen Teilmengen von $\{1, \dots, n\}$. Weiter werde für Elemente $H \in Z_r^n$ der Vektor $x_H \in \bigwedge^r V$ durch $x_H = x_{h_1} \wedge \dots \wedge x_{h_r}$ erklärt, wobei man $H = \{h_1, \dots, h_r\}$ mit $h_1 < \dots < h_r$ schreibe. Für Elemente $a_1, \dots, a_r \in V$, etwa $a_j = \sum_{i=1}^n \alpha_{ij} x_i$ mit Koeffizienten $\alpha_{ij} \in K$, setzen wir dann noch

$$\det_{X,H}(a_1, \dots, a_r) = \sum_{\pi \in \mathfrak{S}_r} \text{sgn}(\pi) \cdot \alpha_{h_{\pi(1)},1} \dots \alpha_{h_{\pi(r)},r}.$$

Es ist also $\det_{X,H}(a_1, \dots, a_r)$, abgesehen vielleicht von dem trivialen Fall $r = 0$, die Determinante derjenigen $(r \times r)$ -Matrix, die man aus der $(n \times r)$ -Matrix $(a_{1,X}, \dots, a_{r,X})$ erhält, indem man alle Zeilen mit einem Index $i \notin H$ streicht; dabei ist $a_{j,X}$ jeweils der Koordinatenspaltenvektor von a_j bezüglich der Basis X , $j = 1, \dots, r$. Insbesondere ergibt sich, dass $\det_{X,H}: V^r \longrightarrow K$ eine alternierende multilineare Abbildung ist.

Satz 4. *Wie vorstehend beschrieben betrachte man einen K -Vektorraum V mit einer Basis $X = (x_1, \dots, x_n)$, sowie das r -te äußere Produkt $\bigwedge^r V$.*

(i) *Die Elemente x_H , $H \in Z_r^n$, bilden eine Basis von $\bigwedge^r V$. Insbesondere gilt $\dim_K(\bigwedge^r V) = \binom{n}{r}$.*

(ii) *Für beliebige Elemente $a_1, \dots, a_r \in V$ gilt*

$$a_1 \wedge \dots \wedge a_r = \sum_{H \in Z_r^n} \det_{X,H}(a_1, \dots, a_r) x_H.$$

Beweis. Der Fall $r = 0$ ist trivial; sei also $r > 0$. Als alternierende multilineare Abbildung faktorisiert $\det_{X,H}: V^r \longrightarrow K$ über eine K -lineare Abbildung

$$d_{X,H}: \bigwedge^r V \longrightarrow K,$$

und es gilt $d_{X,H}(x_{H'}) = \delta_{H,H'}$ für $H, H' \in Z_r^n$, wie man leicht einsieht. Letztere Gleichungen können aber nur dann bestehen, wenn die Elemente x_H , $H \in Z_r^n$ linear unabhängig sind. Dass weiter die x_H ein Erzeugendensystem und damit insgesamt eine Basis von $\bigwedge^r V$ bilden, folgt dann aus (ii), da die Elemente des Typs $a_1 \wedge \dots \wedge a_r$ ein Erzeugendensystem bilden.

Es bleibt also noch Aussage (ii) nachzuweisen. Schreibt man $a_j = \sum_{i=1}^n \alpha_{ij} x_i$ mit Koeffizienten $\alpha_{ij} \in K$, so kann man wie folgt rechnen:

$$\begin{aligned} a_1 \wedge \dots \wedge a_r &= \sum_{i_1, \dots, i_r=1}^n \alpha_{i_1,1} \dots \alpha_{i_r,r} \cdot x_{i_1} \wedge \dots \wedge x_{i_r} \\ &= \sum_{H \in Z_r^n} \sum_{\pi \in \mathfrak{S}_r} \operatorname{sgn}(\pi) \alpha_{h_{\pi(1)},1} \dots \alpha_{h_{\pi(r)},r} \cdot x_H \\ &= \sum_{H \in Z_r^n} \det_{X,H}(a_1, \dots, a_r) \cdot x_H. \end{aligned}$$

□

Der gerade gegebene Beweis zeigt im Übrigen, dass die von den alternierenden multilinearen Abbildungen $\det_{X,H}: V^r \longrightarrow K$ induzierten K -linearen Abbildungen $d_{X,H}: \bigwedge^r V \longrightarrow K$ eine Basis des Dualraums $(\bigwedge^r V)^*$ bilden, nämlich gerade die duale Basis zu der Basis von $\bigwedge^r V$, die von den Elementen x_H gebildet wird.

Ist $f: V \longrightarrow W$ eine lineare Abbildung zwischen K -Vektorräumen, so gibt es für $r \in \mathbb{N}$ eine eindeutig bestimmte K -lineare Abbildung

$$\bigwedge^r f: \bigwedge^r V \longrightarrow \bigwedge^r W$$

mit

$$\bigwedge^r f(a_1 \wedge \dots \wedge a_r) = f(a_1) \wedge \dots \wedge f(a_r).$$

Die Eindeutigkeit von $\bigwedge^r f$ ist klar, da die Elemente des Typs $a_1 \wedge \dots \wedge a_r$ den Vektorraum $\bigwedge^r V$ erzeugen. Um die Existenz zu zeigen, betrachte man im Falle $r > 0$ die alternierende multilineare Abbildung

$$V^r \longrightarrow \bigwedge^r W, \quad (a_1, \dots, a_r) \longmapsto f(a_1) \wedge \dots \wedge f(a_r),$$

und nutze aus, dass diese über $\bigwedge^r V$ faktorisiert.

Korollar 5. *Es sei $f: V \longrightarrow V$ ein Endomorphismus eines K -Vektorraums V endlicher Dimension n . Dann ist der zugehörige Endomorphismus*

$$\bigwedge^n f: \bigwedge^n V \longrightarrow \bigwedge^n V$$

gerade die Multiplikation mit $\det f \in K$.

Beweis. Man wähle eine Basis $X = (x_1, \dots, x_n)$ von V . Dann gilt

$$(\bigwedge^n f)(x_1 \wedge \dots \wedge x_n) = f(x_1) \wedge \dots \wedge f(x_n),$$

sowie nach Satz 4

$$f(x_1) \wedge \dots \wedge f(x_n) = \det(f(x_1)_X, \dots, f(x_n)_X) \cdot x_1 \wedge \dots \wedge x_n.$$

Da $x_1 \wedge \dots \wedge x_n$ eine Basis von $\bigwedge^n V$ bildet, folgt die Behauptung. □

Wir wollen nun noch zeigen, dass man das Dachprodukt “ $\wedge$ ” als Produkt in einem geeigneten Ring auffassen kann.

Lemma 6. *Es sei V ein K -Vektorraum. Zu $r, s \in \mathbb{N}$ existiert dann eine K -bilineare Abbildung*

$$\begin{aligned} \wedge: \bigwedge^r V \times \bigwedge^s V &\longrightarrow \bigwedge^{r+s} V, \\ (a_1 \wedge \dots \wedge a_r, b_1 \wedge \dots \wedge b_s) &\longmapsto a_1 \wedge \dots \wedge a_r \wedge b_1 \wedge \dots \wedge b_s, \end{aligned}$$

welche durch die angegebene Abbildungsvorschrift eindeutig charakterisiert ist.

Beweis. Man betrachte die alternierende multilineare Abbildung

$$\begin{aligned} \Phi: V^{r+s} &\longrightarrow \bigwedge^{r+s} V, \\ (a_1, \dots, a_r, b_1, \dots, b_s) &\longmapsto a_1 \wedge \dots \wedge a_r \wedge b_1 \wedge \dots \wedge b_s, \end{aligned}$$

wobei wir $r, s > 0$ annehmen. Die Fälle $r = 0$ oder $s = 0$ sind in ähnlicher Weise zu behandeln, unter Verwendung der Konvention, dass leere Produkte den Wert 1 haben. Für festgewählte Elemente $a_1, \dots, a_r \in V$ ergibt sich eine alternierende multilineare Abbildung

$$\begin{aligned} V^s &\longrightarrow \bigwedge^{r+s} V, \\ (b_1, \dots, b_s) &\longmapsto a_1 \wedge \dots \wedge a_r \wedge b_1 \wedge \dots \wedge b_s, \end{aligned}$$

welche durch $\bigwedge^s V$ faktorisiert. Es induziert Φ daher eine Abbildung

$$\begin{aligned} \Phi': V^r \times \bigwedge^s V &\longrightarrow \bigwedge^{r+s} V, \\ (a_1, \dots, a_r, \sum_j b_{j1} \wedge \dots \wedge b_{js}) &\longmapsto \sum_j a_1 \wedge \dots \wedge a_r \wedge b_{j1} \wedge \dots \wedge b_{js}, \end{aligned}$$

und man sieht, da Dachprodukte der Form $a_1 \wedge \dots \wedge a_r \wedge b_1 \wedge \dots \wedge b_s$ multilinear und alternierend in den einzelnen Faktoren sind, dass $\Phi'(\cdot, b)$ für festes $b \in \bigwedge^s V$ multilinear und alternierend auf V^r ist. Folglich induziert Φ' eine Abbildung

$$\begin{aligned} \Phi'': \bigwedge^r V \times \bigwedge^s V &\longrightarrow \bigwedge^{r+s} V, \\ (\sum_i a_{i1} \wedge \dots \wedge a_{ir}, \sum_j b_{j1} \wedge \dots \wedge b_{js}) &\longmapsto \sum_{ij} a_{i1} \wedge \dots \wedge a_{ir} \wedge b_{j1} \wedge \dots \wedge b_{js}, \end{aligned}$$

derart, dass $\Phi''(\cdot, b)$ für alle $b \in \bigwedge^s V$ K -linear auf $\bigwedge^r V$ ist. Die Rechenregeln für Dachprodukte zeigen dann, dass Φ'' sogar, wie gewünscht, K -bilinear ist. Dass Φ'' durch die angegebene Abbildungsvorschrift eindeutig charakterisiert ist, folgt daraus, dass die Elemente des Typs $a_1 \wedge \dots \wedge a_r$ ein Erzeugendensystem von $\bigwedge^r V$ bilden und Entsprechendes für $\bigwedge^s V$ gilt. $\square$

Man kann nun zu einem K -Vektorraum V die (konstruierte) direkte Summe

$$\bigwedge V = \bigoplus_{r \in \mathbb{N}} \bigwedge^r V$$

aller äußeren Potenzen bilden, womit man ähnlich wie in 1.6 denjenigen Teil des kartesischen Produktes $\prod_{r \in \mathbb{N}} \bigwedge^r V$ meint, der aus allen Familien $(\lambda_r)_{r \in \mathbb{N}}$ mit $\lambda_r \in \bigwedge^r V$ sowie $\lambda_r = 0$ für fast alle $r \in \mathbb{N}$ besteht. Es ist $\bigwedge V$ in natürlicher Weise ein K -Vektorraum, und man kann zeigen, dass die in Lemma 6 betrachteten Abbildungen des Typs $\bigwedge^r V \times \bigwedge^s V \rightarrow \bigwedge^{r+s} V$ auf $\bigwedge V$ eine Multiplikation definieren, derart dass $\bigwedge V$ ein Ring wird. Dabei erkennt man $K = \bigwedge^0 V$ in kanonischer Weise als Unterring von $\bigwedge V$ und spricht von einer K -Algebra. Genauer bezeichnet man $\bigwedge V$ als die *äußere Algebra* zu V .

Wir wollen hier die Produktbildung aus Lemma 6 lediglich dazu benutzen, um den so genannten *allgemeinen Laplaceschen Entwicklungssatz* für Determinanten herzuleiten. Wir betrachten dazu wieder einen K -Vektorraum V mit Basis $X = (x_1, \dots, x_n)$ und verwenden eine Notation wie in Satz 4, insbesondere sei an die alternierenden multilinearen Abbildungen $\det_{X,H}: V^r \rightarrow K$ zu Elementen $H \in Z_r^n$ erinnert. Dabei stimmt $\det_{X,\{1, \dots, n\}}$ mit der in 4.2/8 eingeführten Determinantenfunktion $\det_X$ überein. Weiter sei $H^\dagger \in Z_{n-r}^n$ für $H \in Z_r^n$ erklärt als Komplement $\{1, \dots, n\} - H$, und man setze $\rho_H = (-1)^\nu$, wobei ν die Anzahl aller Paare $(h, h^\dagger) \in H \times H^\dagger$ mit $h > h^\dagger$ bezeichnet.

Satz 7. *Es sei V ein K -Vektorraum mit Basis $X = (x_1, \dots, x_n)$. Mit vorstehender Notation gilt dann für $1 \leq r < n$ und Vektoren $a_1, \dots, a_n \in V$:*

$$\det_X(a_1, \dots, a_n) = \sum_{H \in Z_r^n} \rho_H \cdot \det_{X,H}(a_1, \dots, a_r) \cdot \det_{X,H^\dagger}(a_{r+1}, \dots, a_n)$$

Beweis. Unter Verwendung von Satz 4 und Lemma 6 kann man wie folgt rechnen:

$$\begin{aligned} & \det_X(a_1, \dots, a_n) \cdot x_1 \wedge \dots \wedge x_n \\ &= a_1 \wedge \dots \wedge a_n \\ &= (a_1 \wedge \dots \wedge a_r) \wedge (a_{r+1} \wedge \dots \wedge a_n) \\ &= \left(\sum_{H \in Z_r^n} \det_{X,H}(a_1, \dots, a_r) \cdot x_H \right) \wedge \left(\sum_{H \in Z_{n-r}^n} \det_{X,H}(a_{r+1}, \dots, a_n) \cdot x_H \right) \\ &= \sum_{H \in Z_r^n} \det_{X,H}(a_1, \dots, a_r) \cdot \det_{X,H^\dagger}(a_{r+1}, \dots, a_n) \cdot x_H \wedge x_{H^\dagger} \\ &= \sum_{H \in Z_r^n} \rho_H \cdot \det_{X,H}(a_1, \dots, a_r) \cdot \det_{X,H^\dagger}(a_{r+1}, \dots, a_n) \cdot x_1 \wedge \dots \wedge x_n \end{aligned}$$

Da $x_1 \wedge \dots \wedge x_n$ eine Basis von $\bigwedge^n V$ bildet, insbesondere also von Null verschieden ist, ergibt sich die gewünschte Beziehung. $\square$

Wenden wir Satz 7 auf $V = K^n$ und die kanonische Basis an, so beschreibt die hergeleitete Formel die Entwicklung der Determinante einer $(n \times n)$ -Matrix nach den ersten r Spalten. Durch Spaltenvertauschung und Berücksichtigung

entsprechender Vorzeichen gewinnt man einen Entwicklungssatz nach r beliebig vorgegebenen Spalten. Weiter kann man durch Transponieren hieraus einen Entwicklungssatz nach vorgegebenen r Zeilen gewinnen.

Aufgaben

1. Es sei V ein K -Vektorraum und $U \subset V$ ein linearer Unterraum. Man formuliere eine universelle Eigenschaft, die den Quotientenvektorraum V/U charakterisiert.
2. Es sei V ein K -Vektorraum. Man zeige, gegebene Vektoren $a_1, \dots, a_r \in V$ sind genau dann linear unabhängig, wenn das Element $a_1 \wedge \dots \wedge a_r \in \bigwedge^r V$ nicht trivial ist.
3. Es sei $f: V \longrightarrow W$ eine lineare Abbildung zwischen K -Vektorräumen. Man zeige für $r \in \mathbb{N}$, dass die Abbildung

$$\bigwedge^r f: \bigwedge^r V \longrightarrow \bigwedge^r W$$

injektiv bzw. surjektiv ist, sofern f diese Eigenschaft besitzt.

4. Es sei V ein K -Vektorraum der Dimension $n < \infty$. Man bestimme

$$\dim \bigoplus_{r \in \mathbb{N}} (\bigwedge^r V).$$

5. Für einen K -Vektorraum V betrachte man $\bigoplus_{r \in \mathbb{N}} \bigwedge^r V$ als Ring unter dem Dachprodukt $\wedge$. Man zeige, dass für ein Element a dieses Ringes genau dann $a \cdot a = 0$ gilt, wenn man $a \in \bigoplus_{i>0} \bigwedge^i V$ hat.
6. *Symmetrische Produkte:* Es sei V ein K -Vektorraum, und $r \in \mathbb{N}$. Eine Abbildung $\Phi: V^r \longrightarrow W$ in einen K -Vektorraum W heißt *symmetrisch*, wenn $\Phi(a_{\pi(1)}, \dots, a_{\pi(r)}) = \Phi(a_1, \dots, a_r)$ für alle r -Tupel $(a_1, \dots, a_r) \in V^r$ und alle Permutationen $\pi \in \mathfrak{S}_r$ gilt. Man zeige: Es existiert ein K -Vektorraum P mit einer symmetrischen multilinearen Abbildung $\sigma: V^r \longrightarrow P$, welche folgende universelle Eigenschaft erfüllt:

Zu jeder symmetrischen multilinearen Abbildung $\Phi: V^r \longrightarrow W$ in einen K -Vektorraum W existiert eindeutig eine K -lineare Abbildung $\varphi: P \longrightarrow W$ mit der Eigenschaft $\Phi = \varphi \circ \sigma$.

Man nennt P die r -te *symmetrische Potenz* von V .

5. Polynome

Vorbemerkungen

Für einen K -Vektorraum V der Dimension $n < \infty$ bilden die Endomorphismen $\tau: V \longrightarrow V$ einen Ring $\text{End}_K(V)$, der gemäß 3.3/2 als K -Vektorraum von der Dimension n^2 ist. Betrachtet man daher zu einem Endomorphismus τ von V dessen Potenzen $\tau^{n^2}, \dots, \tau^0 = \text{id}$, so sind diese linear abhängig. Folglich existiert in $\text{End}_K(V)$ eine Gleichung der Form

$$(*) \quad \tau^r + c_1 \tau^{r-1} + \dots + c_r = 0$$

mit Konstanten $c_i \in K$ und einer natürlichen Zahl $r \leq n^2$, wobei man stets $r \leq n$ wählen kann, wie genauere Überlegungen später zeigen werden. Es handelt sich also um eine Gleichung r -ten Grades mit Koeffizienten aus K , eine so genannte *algebraische Gleichung* von τ über K . Diese kann genau dann linear gewählt werden (d. h. mit $r = 1$), wenn τ ein skalares Vielfaches der Identität ist, so dass im Allgemeinen $r > 1$ gelten wird. Nun ist die Theorie algebraischer Gleichungen allerdings nicht mehr der *Linearen* Algebra zuzurechnen, sie gehört thematisch eher zu dem Bereich, den man heute meist als “Algebra” bezeichnet.

Dennoch sind Gleichungen des Typs $(*)$ für unsere Zwecke sehr wichtig, da man aus ihnen bereits wertvolle Informationen über die Struktur des zugehörigen Endomorphismus τ ablesen kann. All dies werden wir im Kapitel 6 über die Normalformentheorie von Endomorphismen genauestens studieren. Dabei sind jedoch gewisse Grundkenntnisse über solche Gleichungen und insbesondere über die zugehörigen *Polynome*

$$(**) \quad t^r + c_1 t^{r-1} + \dots + c_r$$

erforderlich, wobei das zu $(*)$ gehörige Polynom $(**)$ einen Ausdruck darstellt, in dem man τ durch eine so genannte *Variable* t ersetzt hat, die bei Bedarf unterschiedliche Werte annehmen kann.

Wir werden in diesem Kapitel insbesondere den Ring aller Polynome mit Koeffizienten aus einem gegebenen Körper K betrachten und zeigen, dass dieser in Bezug auf Teilbarkeitseigenschaften sehr große Ähnlichkeiten mit dem Ring $\mathbb{Z}$ der ganzen Zahlen aufweist. Beispielsweise werden wir für Polynomringe den Satz von der eindeutigen Primfaktorzerlegung herleiten.

5.1 Ringe

Bereits in Abschnitt 3.3 hatten wir Ringe betrachtet, und zwar zu einer natürlichen Zahl $n \geq 1$ den Matrizenring $K^{n \times n}$ über einem Körper K , sowie zu einem K -Vektorraum V den Endomorphismenring $\text{End}_K(V)$. Um bestimmte Eigenschaften von Elementen in $K^{n \times n}$ oder $\text{End}_K(V)$ genauer zu beschreiben, ist es zweckmäßig, so genannte *Polynomringe* zu benutzen. Wir wollen daher zunächst einige allgemeine Dinge über Ringe zusammenstellen, wobei wir uns grundsätzlich auf Ringe *mit Eins* beschränken werden. Zunächst sei nochmals an die in 3.3/1 gegebene Definition eines Ringes erinnert.

Definition 1. Eine Menge R mit zwei Verknüpfungen “+” (Addition) und “ $\cdot$ ” (Multiplikation) heißt ein Ring (mit Eins), wenn folgende Bedingungen erfüllt sind:

- (i) R ist eine abelsche Gruppe bezüglich der Addition.
- (ii) Die Multiplikation ist assoziativ, d. h. es gilt

$$(a \cdot b) \cdot c = a \cdot (b \cdot c) \quad \text{für } a, b, c \in R.$$

- (iii) Es existiert ein Einselement in R , also ein Element $1 \in R$, so dass $1 \cdot a = a = a \cdot 1$ für alle $a \in R$ gilt.

- (iv) Addition und Multiplikation verhalten sich distributiv, d. h. für $a, b, c \in R$ gilt

$$a \cdot (b + c) = a \cdot b + a \cdot c, \quad (a + b) \cdot c = a \cdot c + b \cdot c.$$

Der Ring R heißt kommutativ, wenn die Multiplikation kommutativ ist.

Es ist klar, dass das Einselement 1 eines Ringes durch seine definierende Eigenschaft eindeutig bestimmt ist. Als nahe liegendes Beispiel eines kommutativen Rings kann man den Ring $\mathbb{Z}$ der ganzen Zahlen betrachten. Im Übrigen ist jeder Körper, also insbesondere $\mathbb{Q}$, $\mathbb{R}$ oder $\mathbb{C}$, ein kommutativer Ring. Wie schon in Abschnitt 3.3 definiert, heißt ein Element a eines Ringes R eine *Einheit*, wenn es ein Element $b \in R$ mit $ab = ba = 1$ gibt. Die Menge R^* aller Einheiten von R bildet eine Gruppe bezüglich der Multiplikation. Für $1 \neq 0$ gilt $R^* \subset R - \{0\}$, wobei dies im Allgemeinen eine echte Inklusion ist. Genauer ist die Gleichung $R^* = R - \{0\}$ äquivalent zu der Bedingung, dass R ein Körper ist. Als triviales Beispiel eines Rings hat man den so genannten *Nullring* 0. Dieser besteht nur aus einem Element 0 mit den Verknüpfungen $0 + 0 = 0$ und $0 \cdot 0 = 0$. Hier ist das Element 0 ein Null- und Einselement zugleich, so dass wir $1 = 0$ schreiben können. Der Nullring ist der einzige Ring, in dem diese Gleichung gilt.

Bei dem Matrizenring $K^{n \times n}$ über einem Körper K bzw. dem Endomorphismenring $\text{End}_K(V)$ eines K -Vektorraums V handelt es sich für $n \geq 2$ bzw. $\dim_K(V) \geq 2$ um nicht-kommutative Ringe; vgl. Abschnitt 3.3. Als weiteres Beispiel kann man für einen Ring R dessen n -faches kartesisches Produkt R^n als Ring betrachten, indem man Addition und Multiplikation komponentenweise erklärt:

$$\begin{aligned}
(\alpha_1, \dots, \alpha_n) + (\beta_1, \dots, \beta_n) &= (\alpha_1 + \beta_1, \dots, \alpha_n + \beta_n) \\
(\alpha_1, \dots, \alpha_n) \cdot (\beta_1, \dots, \beta_n) &= (\alpha_1 \cdot \beta_1, \dots, \alpha_n \cdot \beta_n)
\end{aligned}$$

Es ist dann $0 = (0, \dots, 0)$ das Nullelement und $1 = (1, \dots, 1)$ das Einselement. Im Falle $R \neq 0$ und $n \geq 2$ zeigt die Gleichung

$$(1, 0, \dots, 0) \cdot (0, \dots, 0, 1) = (0, \dots, 0),$$

dass dieser Ring nicht-triviale Nullteiler besitzt. Dabei heißt ein Element a eines Rings ein *Nullteiler*, wenn es ein Element $b \neq 0$ dieses Rings mit $a \cdot b = 0$ oder $b \cdot a = 0$ gibt. Man nennt einen kommutativen Ring R mit $1 \neq 0$ einen *Integritätsring*, wenn R keine nicht-trivialen Nullteiler besitzt, wenn also für $a, b \in R - \{0\}$ stets $a \cdot b \neq 0$ gilt.

Als wichtiges Beispiel wollen wir nunmehr den *Polynomring* $R[T]$ über einem *kommutativen* Ring R in einer Variablen T konstruieren. Um unsere Intentionen klarzulegen, gehen wir dabei zunächst in naiver Weise vor und erklären $R[T]$ als Menge aller formal gebildeten Summen des Typs $\sum_{i=0}^m a_i T^i$, mit Koeffizienten $a_i \in R$ und variabler oberer Grenze $m \in \mathbb{N}$. Addiert und multipliziert man solche Ausdrücke “wie gewöhnlich”, so erkennt man $R[T]$ als kommutativen Ring. Dabei stelle man sich T als eine “variable” bzw. “allgemeine” Größe vor, für die man nach Bedarf Elemente z. B. aus R einsetzen darf. Wichtig ist, dass Addition und Multiplikation in $R[T]$ bei einem solchen Ersetzungsprozess in die entsprechenden Verknüpfungen von R übergehen. Man rechnet daher mit T in gleicher Weise wie mit einer “konkreten” Größe, etwa aus R .

Der Polynomring $R[T]$ soll nun aber auch noch auf präzisere Weise konstruiert werden. Wir setzen $R[T] = R^{(\mathbb{N})}$ und verstehen hierunter die Menge aller Folgen $(a_i)_{i \in \mathbb{N}}$ von Elementen $a_i \in R$, für die $a_i = 0$ für fast alle $i \in \mathbb{N}$ gilt, also für alle $i \in \mathbb{N}$, bis auf endlich viele Ausnahmen. Addition und Multiplikation solcher Folgen seien erklärt durch die Formeln

$$\begin{aligned}
(a_i)_{i \in \mathbb{N}} + (b_i)_{i \in \mathbb{N}} &:= (a_i + b_i)_{i \in \mathbb{N}}, \\
(a_i)_{i \in \mathbb{N}} \cdot (b_i)_{i \in \mathbb{N}} &:= (c_i)_{i \in \mathbb{N}},
\end{aligned}$$

mit

$$c_i := \sum_{\mu+\nu=i} a_\mu b_\nu = \sum_{\mu=0}^i a_\mu b_{i-\mu}.$$

Indem man die Ringeigenschaften von R benutzt, kann man leicht nachrechnen, dass $R^{(\mathbb{N})}$ mit diesen Verknüpfungen einen kommutativen Ring (mit Eins) bildet. Das Nullelement wird gegeben durch die Folge $0 = (0, 0, \dots)$, das Einselement durch die Folge $1 = (1, 0, 0, \dots)$.

Wir wollen exemplarisch nur das Assoziativgesetz der Multiplikation nachrechnen. Für $(a_i)_i, (b_i)_i, (c_i)_i \in R^{(\mathbb{N})}$ gilt

$$\begin{aligned}
[(a_i)_i \cdot (b_i)_i] \cdot (c_i)_i &= \left(\sum_{\lambda+\mu=i} a_\lambda b_\mu \right)_i \cdot (c_i)_i \\
&= \left(\sum_{\kappa+\nu=i} \left(\sum_{\lambda+\mu=\kappa} a_\lambda b_\mu \right) \cdot c_\nu \right)_i = \left(\sum_{\lambda+\mu+\nu=i} a_\lambda b_\mu c_\nu \right)_i,
\end{aligned}$$

sowie in entsprechender Weise

$$\begin{aligned}(a_i)_i \cdot [(b_i)_i \cdot (c_i)_i] &= (a_i)_i \cdot \left(\sum_{\mu+\nu=i} b_\mu c_\nu \right)_i \\ &= \left(\sum_{\lambda+\kappa=i} a_\lambda \cdot \left(\sum_{\mu+\nu=\kappa} b_\mu c_\nu \right) \right)_i = \left(\sum_{\lambda+\mu+\nu=i} a_\lambda b_\mu c_\nu \right)_i.\end{aligned}$$

Um schließlich die Elemente von $R^{(\mathbb{N})}$ wie gewohnt als Polynome, also Ausdrücke der Form $\sum_{i=0}^{\infty} a_i T^i$, zu schreiben, betrachte man das Element

$$T = (0, 1, 0, \dots) \in R^{(\mathbb{N})}.$$

Für $n \in \mathbb{N}$ gilt $T^n = (\delta_{in})_{i \in \mathbb{N}}$, wobei dies insbesondere für $n = 0$ aufgrund unserer Konvention über das leere Produkt richtig ist. Identifizieren wir nun die Elemente $a \in R$ mit den entsprechenden Elementen $(a, 0, 0, \dots) \in R^{(\mathbb{N})}$, so ist diese Identifizierung mit den Verknüpfungen in R und $R^{(\mathbb{N})}$ verträglich. Wir können also R sozusagen als “Unterring” von $R^{(\mathbb{N})}$ auffassen. Elemente $f = (a_i)_{i \in \mathbb{N}} \in R^{(\mathbb{N})}$ lassen sich dann als Polynome in T mit Koeffizienten aus R schreiben, nämlich in der Form

$$f = (a_i)_{i \in \mathbb{N}} = \sum_{i \in \mathbb{N}} (a_i \delta_{ij})_{j \in \mathbb{N}} = \sum_{i \in \mathbb{N}} a_i (\delta_{ij})_{j \in \mathbb{N}} = \sum_{i \in \mathbb{N}} a_i T^i,$$

wobei die Koeffizienten $a_i \in R$ in der Darstellung $f = \sum_{i \in \mathbb{N}} a_i T^i$ jeweils eindeutig bestimmt sind; man nennt a_i den *Koeffizienten von f zum Grad i* .

Wir werden von nun an für den gerade konstruierten Polynomring stets die Notation $R[T]$ verwenden und dessen Elemente als Polynome der Form $\sum_{i \in \mathbb{N}} a_i T^i$ mit Koeffizienten $a_i \in R$ schreiben (wobei anstelle von T natürlich auch ein anderer Buchstabe zur Bezeichnung der Variablen zugelassen ist). Dabei sei immer stillschweigend vorausgesetzt, dass die Koeffizienten a_i für fast alle Indizes $i \in \mathbb{N}$ verschwinden, die vorstehende Summe also jeweils als *endliche* Summe Sinn macht. Addition und Multiplikation in $R[T]$ beschreiben sich dann durch die bekannten Formeln

$$\begin{aligned}\left(\sum_{i \in \mathbb{N}} a_i T^i \right) + \left(\sum_{i \in \mathbb{N}} b_i T^i \right) &= \sum_{i \in \mathbb{N}} (a_i + b_i) T^i, \\ \left(\sum_{i \in \mathbb{N}} a_i T^i \right) \cdot \left(\sum_{i \in \mathbb{N}} b_i T^i \right) &= \sum_{i \in \mathbb{N}} \left(\sum_{\mu+\nu=i} a_\mu b_\nu \right) \cdot T^i.\end{aligned}$$

Ist für ein Polynom $f = \sum_{i \in \mathbb{N}} a_i T^i \in R[T]$ ein $n \in \mathbb{N}$ bekannt mit $a_i = 0$ für $i > n$, so können wir auch $f = \sum_{i=0}^n a_i T^i$ schreiben. Gilt zudem $a_n \neq 0$, so nennt man a_n den *höchsten Koeffizienten* von f , und es wird n als *Grad* von f bezeichnet, $n = \text{grad } f$. Jedes nicht-triviale Polynom $f \in R[T]$ besitzt einen wohlbestimmten Grad. Darüber hinaus trifft man die Konvention, dem Nullpolynom 0 den Grad $-\infty$ zuzuordnen.

Satz 2. *Es sei R ein kommutativer Ring und $R[T]$ der Polynomring einer Variablen über R . Für $f, g \in R[T]$ gilt dann*

$$\begin{aligned}\operatorname{grad}(f + g) &\leq \max\{\operatorname{grad} f, \operatorname{grad} g\}, \\ \operatorname{grad}(f \cdot g) &\leq \operatorname{grad} f + \operatorname{grad} g.\end{aligned}$$

Ist R ein Integritätsring, so gilt sogar

$$\operatorname{grad}(f \cdot g) = \operatorname{grad} f + \operatorname{grad} g.$$

Beweis. Die Behauptung ist problemlos zu verifizieren, falls f oder g das Nullpolynom ist. Wir dürfen daher f und g als nicht-trivial annehmen, also $m = \operatorname{grad} f \geq 0$ sowie $n = \operatorname{grad} g \geq 0$. Sei etwa $f = \sum a_i T^i$, $g = \sum b_i T^i$. Dann folgt $a_i = 0$ für $i > m$ und $b_i = 0$ für $i > n$ und damit $a_i + b_i = 0$ für $i > \max\{m, n\}$, also $\operatorname{grad}(f + g) \leq \max\{m, n\}$. Auf ähnliche Weise sieht man $\sum_{\mu+\nu=i} a_\mu b_\nu = 0$ für $i > m + n$, und dies bedeutet $\operatorname{grad}(f \cdot g) \leq m + n$. Insbesondere ist

$$a_m b_n = \sum_{\mu+\nu=m+n} a_\mu b_\nu$$

der Koeffizient in $f \cdot g$ vom Grad $m + n$ und damit der höchste Koeffizient, falls er nicht verschwindet. Ist nun R ein Integritätsring, so gilt $a_m b_n \neq 0$ wegen $a_m \neq 0 \neq b_n$, und man erkennt $\operatorname{grad}(f \cdot g) = m + n$. $\square$

Korollar 3. Ist R ein Integritätsring, so auch der Polynomring $R[T]$.

Beweis. Seien $f, g \in R[T]$ zwei nicht-triviale Polynome. Dann folgt $\operatorname{grad} f \geq 0$, $\operatorname{grad} g \geq 0$ und somit gemäß Satz 2 $\operatorname{grad}(f \cdot g) = \operatorname{grad} f + \operatorname{grad} g \geq 0$. Dies zeigt $f \cdot g \neq 0$, d. h. $R[T]$ ist ein Integritätsring. $\square$

Wir wollen im Weiteren spezielle Werte für die Variable T eines Polynomrings $R[T]$ einsetzen. Hierzu ist es von Nutzen, neben Homomorphismen von Ringen auch so genannte Algebren und deren Homomorphismen zu betrachten.

Definition 4. Eine Abbildung $\varphi: R \longrightarrow R'$ zwischen Ringen R, R' heißt ein Homomorphismus, genauer ein Ringhomomorphismus, wenn gilt:

- (i) $\varphi(a + b) = \varphi(a) + \varphi(b)$ für $a, b \in R$.
- (ii) $\varphi(a \cdot b) = \varphi(a) \cdot \varphi(b)$ für $a, b \in R$.
- (iii) $\varphi(1_R) = 1_{R'}$, d. h. φ bildet das Einselement $1_R \in R$ auf das Einselement $1_{R'} \in R'$ ab.

Wie bei Vektorraumhomomorphismen spricht man von einem *Mono-*, *Epi-* bzw. *Isomorphismus*, wenn φ injektiv, surjektiv bzw. bijektiv ist. Weiter wird ein Homomorphismus $\varphi: R \longrightarrow R$ auch als *Endomorphismus* von R bezeichnet, bzw. als *Automorphismus*, wenn dieser ein Isomorphismus ist.

Definition 5. Es sei R ein kommutativer Ring. Eine R -Algebra besteht aus einem (nicht notwendig kommutativen) Ring A und einem Ringhomomorphismus $\varphi: R \longrightarrow A$, derart dass alle Elemente aus $\varphi(R)$ mit den Elementen aus A vertauschbar sind, also $\varphi(r)a = a\varphi(r)$ für alle $r \in R$, $a \in A$ gilt.

Häufig spricht man einfach von A als R -Algebra, ohne den Homomorphismus $\varphi: R \rightarrow A$ explizit zu erwähnen. Entsprechend schreibt man $r \cdot a$ anstelle von $\varphi(r) \cdot a$ für Elemente $r \in R$, $a \in A$, wobei der definierende Homomorphismus $R \rightarrow A$ dann durch $r \mapsto r \cdot 1_A$ gegeben ist. Wenn dieser nicht injektiv ist, so darf man allerdings statt $r \cdot 1_A$ keinesfalls wieder r schreiben, denn es ist dann nicht möglich, die Elemente $r \in R$ mit ihren Bildern $r \cdot 1_A \in A$ zu identifizieren. Beispielsweise sind R und der Polynomring $R[T]$ auf kanonische Weisen R -Algebren, indem man die identische Abbildung $R \rightarrow R$ bzw. die Inklusionsabbildung $R \hookrightarrow R[T]$ betrachtet. Weiter definiert für einen Vektorraum V über einem Körper K die Abbildung

$$K \rightarrow \text{End}_K(V), \quad \alpha \mapsto \alpha \cdot \text{id}_V,$$

den Endomorphismenring $\text{End}_K(V)$ als K -Algebra. Entsprechend ist der Matrizenring $K^{n \times n}$ für $n \in \mathbb{N} - \{0\}$ unter der Abbildung

$$K \rightarrow K^{n \times n}, \quad \alpha \mapsto \alpha \cdot E_n,$$

eine K -Algebra, wobei E_n die Einheitsmatrix in $K^{n \times n}$ bezeichne.

Ist A eine R -Algebra, so kann man neben der Addition und Multiplikation des Ringes A auch noch die äußere Multiplikation

$$R \times A \rightarrow A, \quad (r, a) \mapsto r \cdot a,$$

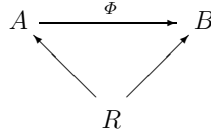
mit Elementen aus R betrachten. Diese Multiplikation erfüllt Eigenschaften, wie sie etwa in 1.4/1 für die skalare Multiplikation eines Vektorraums gefordert werden. In der Tat wird A im Falle eines Körpers $K = R$ unter der äußeren Multiplikation zu einem K -Vektorraum, wobei wir für $\text{End}_K(V)$ und $K^{n \times n}$ jeweils die auch früher schon betrachteten K -Vektorraumstrukturen erhalten. Im Übrigen sei darauf hingewiesen, dass die äußere Multiplikation mit der inneren Multiplikation auf A verträglich ist, d. h. es gilt

$$r \cdot (a \cdot b) = (r \cdot a) \cdot b = a \cdot (r \cdot b) \quad \text{für } r \in R, a, b \in A.$$

Homomorphismen zwischen R -Algebren werden in natürlicher Weise als Ringhomomorphismen erklärt, die zusätzlich die äußere Multiplikation mit R respektieren:

Definition 6. *Es seien A, B Algebren über einem kommutativen Ring R . Ein Homomorphismus von R -Algebren $\Phi: A \rightarrow B$ ist ein Ringhomomorphismus, so dass $\Phi(ra) = r\Phi(a)$ für alle $r \in R$, $a \in A$ gilt.*

Sind $R \rightarrow A$ und $R \rightarrow B$ die definierenden Homomorphismen der betrachteten R -Algebren, so ist ein Ringhomomorphismus $\Phi: A \rightarrow B$ genau dann ein Homomorphismus von R -Algebren, wenn das Diagramm



kommutiert.

Satz 7. *Es sei R ein kommutativer Ring, $R[T]$ der Polynomring einer Variablen über R , sowie A eine beliebige R -Algebra. Zu jedem $t \in A$ gibt es dann einen eindeutig bestimmten R -Algebrahomomorphismus $\Phi: R[T] \rightarrow A$ mit $\Phi(T) = t$. Dieser wird beschrieben durch*

$$\sum_{i \in \mathbb{N}} a_i T^i \mapsto \sum_{i \in \mathbb{N}} a_i t^i$$

oder, in suggestiver Schreibweise, durch

$$f \mapsto f(t),$$

wobei dann insbesondere

$$(f + g)(t) = f(t) + g(t), \quad (f \cdot g)(t) = f(t) \cdot g(t)$$

für $f, g \in R[T]$ gilt. Man nennt Φ auch den Einsetzungshomomorphismus, der t anstelle von T einsetzt.

Beweis. Ist Φ ein R -Algebrahomomorphismus der geforderten Art, so gilt für $\sum_{i \in \mathbb{N}} a_i T^i \in R[T]$

$$\Phi\left(\sum_{i \in \mathbb{N}} a_i T^i\right) = \sum_{i \in \mathbb{N}} \Phi(a_i T^i) = \sum_{i \in \mathbb{N}} a_i \Phi(T)^i = \sum_{i \in \mathbb{N}} a_i t^i.$$

Dies zeigt die Eindeutigkeit von Φ . Um auch die Existenz nachzuweisen, erkläre man Φ durch

$$\Phi\left(\sum_{i \in \mathbb{N}} a_i T^i\right) = \sum_{i \in \mathbb{N}} a_i t^i.$$

Man sieht dann unmittelbar, dass Φ ein R -Algebrahomomorphismus ist. Beispielsweise ergibt sich die Verträglichkeit von Φ mit der Multiplikation aufgrund folgender Rechnung:

$$\begin{aligned}
 \Phi\left(\sum_{i \in \mathbb{N}} a_i T^i \cdot \sum_{i \in \mathbb{N}} b_i T^i\right) &= \Phi\left(\sum_{i \in \mathbb{N}} \left(\sum_{\mu+\nu=i} a_\mu b_\nu\right) \cdot T^i\right) = \sum_{i \in \mathbb{N}} \left(\sum_{\mu+\nu=i} a_\mu b_\nu\right) \cdot t^i \\
 &= \sum_{i \in \mathbb{N}} \sum_{\mu+\nu=i} (a_\mu t^\mu) \cdot (b_\nu t^\nu) = \left(\sum_{i \in \mathbb{N}} a_i t^i\right) \cdot \left(\sum_{i \in \mathbb{N}} b_i t^i\right) \\
 &= \Phi\left(\sum_{i \in \mathbb{N}} a_i T^i\right) \cdot \Phi\left(\sum_{i \in \mathbb{N}} b_i T^i\right)
 \end{aligned}$$

Man beachte dabei, dass wir die Vertauschbarkeit von t mit den (Bildern der) Koeffizienten b_i in A benutzt haben. $\square$

Betrachten wir beispielsweise den Endomorphismenring $A = \text{End}_K(V)$ eines K -Vektorraums V als K -Algebra, so können wir zu jedem Endomorphismus $\tau: V \rightarrow V$ den K -Algebrahomomorphismus

$$\Phi_\tau: K[T] \rightarrow \text{End}_K(V), \quad \sum_{i \in \mathbb{N}} a_i T^i \mapsto \sum_{i \in \mathbb{N}} a_i \tau^i,$$

betrachten, der den Endomorphismus τ anstelle von T einsetzt. Dabei ist τ^i natürlich erklärt als i -fache Komposition $\tau \circ \dots \circ \tau$ von τ mit sich selber, und das Bild eines Polynoms $\sum_{i \in \mathbb{N}} a_i T^i \in K[T]$ unter Φ_τ ergibt sich als Endomorphismus

$$\sum_{i \in \mathbb{N}} a_i \tau^i: V \rightarrow V, \quad v \mapsto \sum_{i \in \mathbb{N}} a_i \tau^i(v).$$

Wir werden den Homomorphismus Φ_τ insbesondere zur Definition des so genannten *Minimalpolynoms* von τ verwenden; vgl. 6.2/9 und 6.2/10. Anstelle des Endomorphismenrings eines Vektorraums kann man natürlich auch den Matrizenring $K^{n \times n}$ betrachten. Zu jeder Matrix $C \in K^{n \times n}$ erhält man dann einen K -Algebrahomomorphismus

$$\Phi_C: K[T] \rightarrow K^{n \times n}, \quad \sum_{i \in \mathbb{N}} a_i T^i \mapsto \sum_{i \in \mathbb{N}} a_i C^i,$$

der C anstelle von T einsetzt.

Wir wollen ein weiteres Beispiel betrachten. Es sei K ein Körper und $\text{Abb}(K, K)$ die Menge aller Abbildungen $K \rightarrow K$ oder, mit anderen Worten, die Menge aller K -wertigen Funktionen auf K . Zu jedem Polynom $f \in K[T]$ lässt sich dann die zugehörige Polynomfunktion $t \mapsto f(t)$ als Element von $\text{Abb}(K, K)$ betrachten. Diese ordnet einem Element $t \in K$ den Wert $f(t) \in K$ zu, den man erhält, indem man t anstelle von T in f einsetzt. Auf diese Weise ergibt sich eine Abbildung

$$\Psi: K[T] \rightarrow \text{Abb}(K, K), \quad f \mapsto (t \mapsto f(t)),$$

die wir im Folgenden als Homomorphismus von K -Algebren deuten wollen. Um $A = \text{Abb}(K, K)$ als K -Algebra zu erklären, betrachten wir die gewöhnliche Addition und Multiplikation K -wertiger Funktionen, gegeben durch

$$\begin{aligned} p + q: K &\rightarrow K, & t &\mapsto p(t) + q(t), \\ p \cdot q: K &\rightarrow K, & t &\mapsto p(t) \cdot q(t). \end{aligned}$$

Mit diesen Verknüpfungen ist A ein kommutativer Ring, in dem die Nullfunktion 0 das Nullelement und die Funktion 1_A , die identisch 1 ist, das Einselement bilden. Für $\alpha \in K$ kann man weiter das α -fache eines Elementes $p \in A$ durch

$$\alpha \cdot p: K \rightarrow K, \quad t \mapsto \alpha \cdot (p(t)),$$

erklären. In diesem Sinne ist dann

$$K \longrightarrow A, \quad \alpha \longmapsto \alpha \cdot 1_A,$$

ein Ringhomomorphismus, der A zu einer K -Algebra macht. Mit Satz 7 folgt leicht, dass die Abbildung Ψ , die einem Polynom $f \in K[T]$ die zugehörige Polynomfunktion $t \mapsto f(t)$ zuordnet, tatsächlich ein Homomorphismus von K -Algebren ist. Gleiches gilt, wenn man allgemeiner anstelle von K einen kommutativen Ring R zugrunde legt. Die Homomorphie-Eigenschaft für Ψ besagt insbesondere, dass man mit dem Element T in gleicher Weise rechnet, wie man es mit konkreten Werten t anstelle von T tun würde. Dies rechtfertigt die Bezeichnung von T als “Variable”.

Enthält der Körper K unendlich viele Elemente, so ist die Abbildung Ψ injektiv, wie wir später aus 5.3/2 ablesen können. Man könnte dann Polynome aus $K[T]$ mit ihren zugehörigen Polynomfunktionen $K \longrightarrow K$ identifizieren. Für einen endlichen Körper $K = \{\alpha_1, \dots, \alpha_q\}$ jedoch ist beispielsweise

$$f = (T - \alpha_1) \dots (T - \alpha_q) \in K[T]$$

ein nicht-triviales Polynom, dessen zugehörige Polynomfunktion $K \longrightarrow K$ identisch verschwindet, so dass also $\Psi(f) = 0$ gilt. Die Abbildung Ψ ist daher im Allgemeinen nicht injektiv, und dies ist einer der Gründe dafür, dass wir Polynome nicht als konkrete Funktionen, sondern als formale Ausdrücke unter Zuhilfenahme einer “Variablen” erklärt haben.

Wir wollen nun noch Ideale und Unterringe von Ringen einführen, Begriffe, die im Zusammenhang mit Ringhomomorphismen von Bedeutung sind.

Definition 8. *Es sei R ein Ring (mit Eins). Ein Unterring von R besteht aus einer Teilmenge $S \subset R$, derart dass gilt:*

- (i) $a, b \in S \implies a - b, a \cdot b \in S$.
- (ii) $1 \in S$.

S ist dann mit den von R induzierten Verknüpfungen selbst wieder ein Ring, und die Inklusionsabbildung $S \hookrightarrow R$ ist ein Ringhomomorphismus. Man überlegt sich leicht, dass das Bild $\varphi(R)$ eines Ringhomomorphismus $\varphi: R \longrightarrow R'$ stets ein Unterring von R' ist.

Definition 9. *Es sei R ein Ring. Eine Teilmenge $\mathfrak{a} \subset R$ heißt Ideal, falls gilt:*

- (i) $\mathfrak{a}$ ist additive Untergruppe von R , d. h. man hat $\mathfrak{a} \neq \emptyset$, und aus $a, b \in \mathfrak{a}$ folgt $a - b \in \mathfrak{a}$.
- (ii) Aus $a \in \mathfrak{a}$, $r \in R$ folgt $ra, ar \in \mathfrak{a}$.

Man rechnet leicht nach, dass für jeden Ringhomomorphismus $\varphi: R \longrightarrow R'$ der Kern

$$\ker \varphi = \{a \in R; \varphi(a) = 0\}$$

ein Ideal ist, wobei φ genau dann injektiv ist, wenn $\ker \varphi$ das Nullideal ist, d. h. nur aus dem Nullelement besteht. Ist R ein beliebiger Ring, so kann man zu

einem Element $a \in R$, welches mit allen übrigen Elementen von R vertauschbar ist, das hiervon erzeugte *Hauptideal*

$$(a) = Ra = \{ra; r \in R\}$$

betrachten. Speziell besteht für $R = \mathbb{Z}$ das Ideal (2) aus allen geraden Zahlen, das Ideal (12) aus allen durch 12 teilbaren Zahlen. In jedem Ring gibt es die so genannten trivialen Ideale, nämlich das *Einheitsideal* $(1) = R$ sowie das *Nullideal* (0) , welches man meist mit 0 bezeichnet. Ein Körper besitzt außer den trivialen keine weiteren Ideale. Umgekehrt kann man zeigen, dass ein kommutativer Ring, der genau zwei Ideale hat, ein Körper ist.

Ist $\mathfrak{a}$ ein Ideal eines Ringes R , so kann man ähnlich wie in Abschnitt 2.2 den *Quotienten-* oder *Restklassenring* $R/\mathfrak{a}$ konstruieren. Da man auf diese Weise interessante Ringe und sogar Körper erhalten kann, wollen wir die Konstruktion hier kurz besprechen. Man erklärt eine Relation “ $\sim$ ” auf R , indem man für $a, b \in R$ setzt:

$$a \sim b \iff a - b \in \mathfrak{a}$$

Man sieht dann unschwer, dass “ $\sim$ ” eine Äquivalenzrelation ist. Für $a \in R$ hat man $a \sim a$ wegen $a - a = 0 \in \mathfrak{a}$. Die Relation ist daher reflexiv. Gilt weiter $a \sim b$, so folgt $a - b \in \mathfrak{a}$ und damit auch $b - a = -(a - b) \in \mathfrak{a}$. Letzteres bedeutet $b \sim a$, und es folgt die Symmetrie von “ $\sim$ ”. Zum Nachweis der Transitivität schließlich nehme man $a \sim b$ und $b \sim c$ für Elemente $a, b, c \in R$ an. Dann gilt $a - b, b - c \in \mathfrak{a}$, und man erhält

$$a - c = (a - b) + (b - c) \in \mathfrak{a},$$

d. h. $a \sim c$. Es ist “ $\sim$ ” also eine Äquivalenzrelation, und wir können zu einem Element $a \in R$ die zugehörige Äquivalenzklasse

$$[a] := \{b \in R; b \sim a\} = a + \mathfrak{a}$$

bilden, wobei $a + \mathfrak{a}$ als *Nebenklasse von a bezüglich $\mathfrak{a}$* , also als Menge aller Summen des Typs $a + a'$ mit $a' \in \mathfrak{a}$ zu interpretieren ist. Gemäß 2.2/4 sind zwei Äquivalenzklassen $[a], [b]$ entweder disjunkt oder aber identisch (im Falle $a \sim b$). Insbesondere ist R die disjunkte Vereinigung aller Äquivalenzklassen.

Man bezeichne nun die Menge aller Nebenklassen bezüglich $\mathfrak{a}$ mit $R/\mathfrak{a}$. Um $R/\mathfrak{a}$ zu einem Ring zu machen, werden folgende Verknüpfungen erklärt:

$$[a] + [b] = [a + b], \quad [a] \cdot [b] = [a \cdot b]$$

Dabei ist zu verifizieren, dass die Klasse $[a + b]$ bzw. $[a \cdot b]$ nicht von der Wahl der Repräsentanten a und b der betrachteten Klassen $[a], [b]$ abhängt. Gelte etwa $[a] = [a']$ und $[b] = [b']$. Dann folgt $a - a', b - b' \in \mathfrak{a}$ und damit

$$(a + b) - (a' + b') \in \mathfrak{a}, \quad ab - a'b' = a(b - b') + (a - a')b' \in \mathfrak{a},$$

also $a + b \sim a' + b'$ und $ab \sim a'b'$. Addition und Multiplikation in $R/\mathfrak{a}$ sind daher wohldefiniert, und man sieht unmittelbar aufgrund der Ringeigenschaften von R , dass auch $R/\mathfrak{a}$ ein Ring ist, und zwar derart, dass die Abbildung

$$\pi: R \longrightarrow R/\mathfrak{a}, \quad a \longmapsto [a],$$

ein surjektiver Ringhomomorphismus mit Kern $\mathfrak{a}$ ist. Man bezeichnet $R/\mathfrak{a}$ als *Restklassen- oder Quotientenring von R modulo $\mathfrak{a}$* . Ähnlich wie in 2.2/8 gilt folgender Homomorphiesatz:

Satz 10. (Homomorphiesatz für Ringe). *Es sei $\varphi: R \longrightarrow R'$ ein Ringhomomorphismus, $\mathfrak{a} \subset R$ ein Ideal mit $\mathfrak{a} \subset \ker \varphi$ und $\pi: R \longrightarrow R/\mathfrak{a}$ die natürliche Projektion. Dann existiert ein eindeutig bestimmter Ringhomomorphismus $\overline{\varphi}: R/\mathfrak{a} \longrightarrow R'$, so dass das Diagramm*

$$\begin{array}{ccc} R & \xrightarrow{\varphi} & R' \\ & \searrow \pi \quad \nearrow \overline{\varphi} & \\ & R/\mathfrak{a} & \end{array}$$

kommutiert. Dabei ist $\overline{\varphi}$ genau dann injektiv, wenn $\mathfrak{a} = \ker \varphi$ gilt und genau dann surjektiv, wenn φ surjektiv ist.

Beweis. Da man die Argumentation aus 2.2/8 mutatis mutandis übernehmen kann, wollen wir uns hier kurz fassen und nur auf die Definition von $\overline{\varphi}$ eingehen. Um $\overline{\varphi}(\overline{a})$ für ein Element $\overline{a} \in R/\mathfrak{a}$ zu erklären, wähle man ein π -Urbild $a \in R$ und setze $\overline{\varphi}(\overline{a}) = \varphi(a)$. Dabei ist natürlich zu verifizieren, dass das Element $\varphi(a)$ unabhängig von der Wahl des π -Urbildes a zu $\overline{a}$ ist. Letzteres folgt aus der Bedingung $\mathfrak{a} \subset \ker \varphi$. $\square$

Aufgaben

R sei stets ein kommutativer Ring.

1. Für ein Element $t \in R$ betrachte man den Homomorphismus $\Phi: R[T] \longrightarrow R$, $f \longmapsto f(t)$, der t anstelle der Variablen T einsetzt. Man zeige:

$$\ker \Phi = R[T] \cdot (T - t)$$

(Hinweis: Man reduziere auf den Fall $t = 0$, indem man den Einsetzungshomomorphismus $R[T] \longrightarrow R[T]$ betrachtet, der T durch $T + t$ ersetzt.)

2. Es sei V ein nicht-trivialer Vektorraum über einem Körper K . Zu einem Endomorphismus $\varphi \in \text{End}_K(V)$ betrachte man den Einsetzungshomomorphismus $\Phi: K[T] \longrightarrow \text{End}_K(V)$, der φ anstelle von T einsetzt. Man bestimme $\ker \Phi$ in den Fällen $\varphi = \text{id}$ und $\varphi = 0$.
3. Es bezeichne $R^{\mathbb{N}}$ die Menge aller Folgen $(a_i)_{i \in \mathbb{N}}$ von Elementen $a_i \in R$.

- (i) Man verwende die gleichen Formeln wie im Falle des Polynomrings einer Variablen über R , um anstelle von $R^{(\mathbb{N})}$ auf $R^{\mathbb{N}}$ die Struktur eines Ringes zu erklären. Dieser Ring wird mit $R[[T]]$ bezeichnet und heißt Ring der *formalen Potenzreihen* einer Variablen über R . Seine Elemente lassen sich als *unendliche* Reihen $\sum_{i=0}^{\infty} a_i T^i$ darstellen.

- (ii) Es sei $q \in R[[T]] \cdot T$. Man zeige, dass $\sum_{n=0}^{\infty} q^n$ zu einem wohldefinierten Element $f \in R[[T]]$ Anlass gibt und dass $f \cdot (1 - q) = 1$ gilt.
 - (iii) Man bestimme die Gruppe aller Einheiten in $R[[T]]$.
4. Es seien $\mathfrak{a}, \mathfrak{b} \subset R$ Ideale. Man zeige, dass die folgenden Teilmengen von R wiederum Ideale bilden:
- (i) $\mathfrak{a} + \mathfrak{b} = \{a + b; a \in \mathfrak{a}, b \in \mathfrak{b}\}$
 - (ii) $\mathfrak{a} \cdot \mathfrak{b} =$ Menge aller endlichen Summen von Produkten $a \cdot b$ mit $a \in \mathfrak{a}$ und $b \in \mathfrak{b}$
 - (iii) $\mathfrak{a} \cap \mathfrak{b}$
5. Man zeige, dass R genau dann ein Körper ist, wenn R genau zwei verschiedene Ideale besitzt.
6. Für ein Ideal $\mathfrak{a} \subset R$ zeige man, dass die Menge aller Elemente $a \in R$, zu denen es ein $n \in \mathbb{N}$ mit $a^n \in \mathfrak{a}$ gibt, ebenfalls wieder ein Ideal in R bildet. Dieses wird mit $\text{rad } \mathfrak{a}$ bezeichnet und heißt das *Nilradikal* von $\mathfrak{a}$.
7. Für eine Familie $(a_i)_{i \in I}$ von Elementen in R zeige man:
- (i) Es existiert ein kleinstes Ideal $\mathfrak{a} \subset R$ mit $a_i \in \mathfrak{a}$ für alle $i \in I$.
 - (ii) Es gilt $\mathfrak{a} = \{\sum_{i \in I} r_i a_i; r_i \in R, r_i = 0 \text{ für fast alle } i \in I\}$.
8. (*Isomorphiesatz*) Es seien $\mathfrak{a} \subset \mathfrak{b} \subset R$ Ideale. Man zeige:
- (i) Die kanonische Abbildung $\mathfrak{b} \hookrightarrow R \longrightarrow R/\mathfrak{a}$ besitzt $\mathfrak{a}$ als Kern und induziert eine Injektion $\mathfrak{b}/\mathfrak{a} \hookrightarrow R/\mathfrak{a}$, wobei man $\mathfrak{b}/\mathfrak{a}$ zunächst als Menge der Restklassen $b + \mathfrak{a}$ mit $b \in \mathfrak{b}$ erkläre. Weiter lässt sich $\mathfrak{b}/\mathfrak{a}$ mit seinem Bild in $R/\mathfrak{a}$ identifizieren und als Ideal in $R/\mathfrak{a}$ auffassen.
 - (ii) Die Projektion $R \longrightarrow R/\mathfrak{b}$ faktorisiert über $R/\mathfrak{a}$, d. h. lässt sich als Komposition $R \xrightarrow{\pi} R/\mathfrak{a} \xrightarrow{f} R/\mathfrak{b}$ schreiben, mit einem Ringhomomorphismus f und der kanonischen Projektion π .
 - (iii) f besitzt $\mathfrak{b}/\mathfrak{a}$ als Kern und induziert einen Isomorphismus

$$(R/\mathfrak{a})/(\mathfrak{b}/\mathfrak{a}) \xrightarrow{\sim} R/\mathfrak{b}.$$

5.2 Teilbarkeit in Integritätsringen

In diesem Abschnitt seien alle Ringe als *Integritätsringe* und damit insbesondere als kommutativ vorausgesetzt. Im Wesentlichen interessieren wir uns für den Polynomring $K[T]$ in einer Variablen T über einem Körper K , für den wir Teilbarkeits- und Faktorisierungsaussagen herleiten wollen. Als Ring mit ganz analogen Eigenschaften soll aber auch der Ring $\mathbb{Z}$ der ganzen Zahlen betrachtet werden. Grundlegend ist in beiden Ringen das Verfahren der *Division mit Rest*, welches wir insbesondere benutzen wollen, um den Satz über die eindeutige Primfaktorzerlegung in Polynomringen herzuleiten. Zunächst erinnern wir an dieses Verfahren, wobei wir davon ausgehen, dass die Division mit Rest im Ring $\mathbb{Z}$ der ganzen Zahlen wohlbekannt ist.

Satz 1. Zu $a, b \in \mathbb{Z}$ mit $b > 0$ existieren eindeutig bestimmte Zahlen $q, r \in \mathbb{Z}$ mit

$$a = qb + r, \quad 0 \leq r < b.$$

Satz 2. Zu Polynomen $f, g \in K[T]$, $g \neq 0$, mit Koeffizienten aus einem Körper K existieren eindeutig bestimmte Polynome $q, r \in K[T]$ mit

$$f = qg + r, \quad \text{grad } r < \text{grad } g.$$

Beweis zu Satz 2. Wir beginnen mit der Existenzaussage. Für $\text{grad } f < \text{grad } g$ gilt die gewünschte Zerlegung trivialerweise mit $q = 0$ und $r = f$. Hat man andererseits

$$m = \text{grad } f \geq \text{grad } g = n,$$

so sei a (bzw. b) der höchste Koeffizient von f (bzw. g), und man setze

$$q_1 := \frac{a}{b} \cdot T^{m-n}, \quad f_1 := f - q_1 g.$$

Dann gilt

$$\text{grad}(q_1 g) = \text{grad } q_1 + \text{grad } g = (m - n) + n = m = \text{grad } f$$

nach 5.1/2, und die höchsten Koeffizienten von $q_1 g$ und f stimmen überein. Insbesondere ergibt sich

$$f = q_1 g + f_1, \quad \text{grad } f_1 < \text{grad } f.$$

Im Falle $\text{grad } f_1 < \text{grad } g$ erhält man die gewünschte Zerlegung mit $q = q_1$ und $r = f_1$. Falls aber $\text{grad } f_1 \geq \text{grad } g$ gilt, so kann man nach dem gerade beschriebenen Verfahren fortfahren und eine Zerlegung

$$f_1 = q_2 g + f_2, \quad \text{grad } f_2 < \text{grad } f_1,$$

finden usw. Nach endlich vielen Schritten gelangt man schließlich zu einer Zerlegung

$$f_{k-1} = q_k g + f_k$$

mit $\text{grad } f_k < \text{grad } g$. Dann ist

$$f = (q_1 + \dots + q_k)g + f_k$$

die gewünschte Zerlegung von f .

Um nun auch die Eindeutigkeitsaussage herzuleiten, betrachte man Zerlegungen

$$f = qg + r = q'g + r'$$

mit $\text{grad } r, \text{grad } r' < \text{grad } g$. Sodann hat man

$$0 = (q - q')g + (r - r') \quad \text{bzw.} \quad (q - q')g = r' - r.$$

Gilt nun $q - q' \neq 0$, so folgt $\text{grad}(q - q') \geq 0$ und damit gemäß 5.1/2

$$\text{grad}((q - q')g) = \text{grad}(q - q') + \text{grad } g \geq \text{grad } g.$$

Andererseits hat man aber

$$\text{grad}(r' - r) \leq \max\{\text{grad } r', \text{grad } r\} < \text{grad } g,$$

was der vorhergehenden Abschätzung widerspricht. Also folgt $q = q'$ und damit $r = r'$. $\square$

Ringe, in denen eine Division wie in den Sätzen 1 oder 2 möglich ist, werden als *euklidische Ringe* bezeichnet, genauer:

Definition 3. Ein Integritätsring R heißt ein euklidischer Ring, wenn es eine Abbildung $\delta: R - \{0\} \rightarrow \mathbb{N}$ (eine so genannte Gradfunktion) mit folgender Eigenschaft gibt: Zu $a, b \in R$, $b \neq 0$, existieren jeweils (nicht notwendig eindeutig bestimmte) Elemente $q, r \in R$ mit $a = qb + r$ und $r = 0$ oder $\delta(r) < \delta(b)$.

Wir können daher feststellen:

Korollar 4. Der Ring $\mathbb{Z}$ der ganzen Zahlen und der Polynomring $K[T]$ über einem Körper K sind euklidische Ringe. Als Gradfunktion nehme man im ersten Fall den Absolutbetrag, im zweiten den gewöhnlichen Grad von Polynomen.

Definition 5. Ein Integritätsring R heißt Hauptidealring, wenn jedes Ideal $\mathfrak{a} \subset R$ ein Hauptideal ist, also die Gestalt $\mathfrak{a} = (a)$ mit einem Element $a \in R$ besitzt.

Satz 6. Es sei R ein euklidischer Ring. Dann ist R auch ein Hauptidealring.

Beweis. Es sei $\mathfrak{a} \subset R$ ein Ideal. Um zu zeigen, dass $\mathfrak{a}$ ein Hauptideal ist, dürfen wir $\mathfrak{a} \neq 0$ annehmen, denn anderenfalls gilt $\mathfrak{a} = 0 = (0)$. Sei nun $a \in \mathfrak{a} - \{0\}$ ein Element, für das der "Grad" $\delta(a)$ minimal ist. Wir behaupten, dass dann $\mathfrak{a} = (a)$ gilt. Natürlich gilt $(a) \subset \mathfrak{a}$. Um auch die umgekehrte Inklusion zu zeigen, wählen wir ein Element $a' \in \mathfrak{a}$. Da R euklidisch ist, gibt es Elemente $q, r \in R$ mit $a' = qa + r$, wobei r entweder verschwindet oder $\delta(r) < \delta(a)$ erfüllt. Nun hat man aber $r = a' - qa \in \mathfrak{a}$, so dass aufgrund der Wahl von a notwendig $r = 0$ und damit $a' = qa \in (a)$ folgt. Es gilt daher $\mathfrak{a} \subset (a)$, insgesamt also $\mathfrak{a} = (a)$, und R ist ein Hauptidealring. $\square$

Korollar 7. Der Ring $\mathbb{Z}$ der ganzen Zahlen und der Polynomring $K[T]$ über einem Körper K sind Hauptidealringe.

Wir wollen als Nächstes den Teilbarkeitsbegriff in einem Integritätsring einführen.

Definition 8. Es seien a, b Elemente eines Integritätsrings R .

- (i) Man sagt, a teile b oder a sei ein Teiler von b , in Zeichen $a|b$, wenn es ein $c \in R$ mit $ac = b$ gibt. Ist a kein Teiler von b , so schreibt man $a \nmid b$.
- (ii) a und b heißen assoziiert, wenn es eine Einheit $e \in R^*$ mit $ae = b$ gibt.

Es ist also a genau dann ein Teiler von b , wenn $b \in (a)$ bzw. $(b) \subset (a)$ gilt. Beispielsweise teilt $T + 1$ das Polynom $T^2 - 1$ in $K[T]$, und man hat $a|b$ sowie $b|a$, falls a und b assoziiert sind. Genauer kann man feststellen:

Bemerkung 9. Für Elemente a, b eines Integritätsrings R ist äquivalent:

- (i) $a|b$ und $b|a$.
- (ii) $(a) = (b)$.
- (iii) a und b sind assoziiert.

Beweis. Wir zeigen lediglich, dass aus (ii) Bedingung (iii) folgt, alle anderen Implikationen ergeben sich mittels direkter Verifikation aus Definition 8. Gelte also $(a) = (b)$. Dann gibt es Elemente $c, d \in R$ mit $ac = b$ und $a = bd$. Hieraus folgt $a = bd = acd$, also $a \cdot (1 - cd) = 0$. Gilt nun $a \neq 0$, so folgt $cd = 1$, da R ein Integritätsring ist, und es sind c, d Einheiten in R . Folglich sind a und b assoziiert. Gleiches gilt aber auch im Falle $a = 0$, denn man hat dann insbesondere $b = ac = 0$. $\square$

Definition 10. Es sei R ein Integritätsring und $p \in R$ ein Element, welches keine Einheit und von Null verschieden ist.

- (i) p heißt irreduzibel, wenn aus einer Gleichung $p = ab$ mit $a, b \in R$ stets folgt, dass a oder b eine Einheit in R ist. Anderenfalls nennt man p reduzibel.
- (ii) p heißt Primelement, wenn aus $p|ab$ mit $a, b \in R$ stets $p|a$ oder $p|b$ folgt oder, in äquivalenter Formulierung, wenn aus $ab \in (p)$ stets $a \in (p)$ oder $b \in (p)$ folgt.

Es ist also p genau dann irreduzibel, wenn aus einer Relation $p \in (a)$ mit $a \in R$ entweder $(a) = (p)$ oder $(a) = R$ folgt. Weiter sieht man mit Induktion, dass ein Primelement $p \in R$ genau dann ein Produkt $a_1 \cdot \dots \cdot a_r$ von Elementen aus R teilt, wenn es einen der Faktoren a_i teilt.

Bemerkung 11. Es sei R ein Integritätsring. Dann ist jedes Primelement von R auch irreduzibel.

Beweis. Sei $p \in R$ ein Primelement und seien $a, b \in R$ mit $p = ab$. Dann teilt p das Produkt ab , und es folgt, dass p einen der beiden Faktoren teilt, etwa $p|a$, d. h. es gibt eine Gleichung $a = pc$ mit $c \in R$. Setzt man dies in die Gleichung $p = ab$ ein, so erhält man $p = pcb$ bzw. $p(1 - cb) = 0$. Da R ein Integritätsring und p von Null verschieden ist, folgt hieraus $cb = 1$, d. h. b ist eine Einheit. Mithin ist p irreduzibel. Alternativ hätten wir auch die Beziehungen $a|p$ (wegen

$p = ab$) und $p|a$ verwenden können. Mit Bemerkung 9 folgt hieraus, dass a und p assoziiert sind. $\square$

Wir werden sogleich zeigen, dass in Hauptidealringen auch die Umkehrung von Bemerkung 11 gilt. Insbesondere dürfen wir dann Primzahlen in $\mathbb{Z}$, die ja gemeinhin als *irreduzible* Elemente definiert werden, auch als *Primelemente* bezeichnen.

Satz 12. *Es sei R ein Hauptidealring und $p \in R$ von 0 verschieden und keine Einheit. Dann ist äquivalent:*

- (i) *p ist irreduzibel.*
- (ii) *p ist ein Primelement.*

Beweis. Wir haben nur noch die Implikation (i) $\implies$ (ii) zu zeigen. Sei also $p \in R$ ein irreduzibles Element, und gelte $p|ab$ sowie $p \nmid a$ für zwei Elemente $a, b \in R$. Um $p|b$ zu zeigen, betrachte man das Ideal

$$Ra + Rp := \{ra + sp; r, s \in R\}$$

in R , welches aufgrund unserer Voraussetzung über R ein Hauptideal ist, etwa $Ra + Rp = Rd$. Insbesondere gilt $a, p \in Rd$ und folglich $d|a$, $d|p$. Nun ist p aber irreduzibel. Daher folgt aus $d|p$ bzw. einer Gleichung $p = cd$, dass c oder d eine Einheit ist. Ist nun c eine Einheit, so können wir $d = c^{-1}p$ schreiben, und man erhält $p|a$ aus $d|a$, im Widerspruch zu $p \nmid a$. Somit bleibt nur der Fall übrig, dass d eine Einheit ist, d. h. es gilt $Ra + Rp = R$ und, nach Multiplikation mit b , die Gleichung $Rab + Rpb = Rb$. Es existieren also $r, s \in R$ mit $rab + spb = b$. Wegen $p|ab$ folgt hieraus wie gewünscht $p|b$. $\square$

Korollar 13. *In einem Hauptidealring R lässt sich jedes Element $a \in R - \{0\}$, welches keine Einheit ist, als endliches Produkt von Primelementen schreiben.*

Beweis. Da jedes irreduzible Element von R bereits prim ist, genügt es, eine Faktorisierung in irreduzible Elemente zu konstruieren. Sei also $a \in R - \{0\}$ eine Nichteinheit. Wir gehen indirekt vor und nehmen an, dass sich a nicht als endliches Produkt irreduzibler Elemente schreiben lässt. Dann ist a reduzibel, und man kann a folglich als Produkt $a_1 a'_1$ zweier Nichteinheiten aus R schreiben. Da a keine endliche Faktorisierung in irreduzible Elemente besitzt, gilt dasselbe für mindestens einen der beiden Faktoren a_1, a'_1 , etwa für a_1 , und wir können a_1 wiederum als Produkt $a_2 a'_2$ zweier Nichteinheiten aus R schreiben. Fährt man auf diese Weise fort, so erhält man eine Folge von Elementen

$$a = a_0, a_1, a_2, \dots \in R,$$

so dass a_{i+1} jeweils ein Teiler von a_i , aber nicht assoziiert zu a_i ist. Mit anderen Worten, man erhält eine aufsteigende Folge von Idealen

$$(a) = (a_0) \subsetneq (a_1) \subsetneq (a_2) \subsetneq \dots,$$

wobei es sich hier gemäß Bemerkung 9 jeweils um echte Inklusionen handelt. Man prüft nun leicht nach, dass die Vereinigung einer aufsteigenden Folge von Idealen wiederum ein Ideal ergibt. Wir können also durch $\mathfrak{b} = \bigcup_{i=0}^{\infty} (a_i)$ ein Ideal in R definieren, und zwar ein Hauptideal, da R ein Hauptidealring ist. Folglich existiert ein Element $b \in \mathfrak{b}$ mit $\mathfrak{b} = (b)$. Nach Definition von $\mathfrak{b}$ gibt es dann einen Index $i_0 \in \mathbb{N}$ mit $b \in (a_{i_0})$, und es folgt

$$\mathfrak{b} = (b) \subset (a_{i_0}) \subset (a_i) \subset \mathfrak{b}$$

für $i \geq i_0$, also $(a_{i_0}) = (a_i)$ für $i \geq i_0$, im Widerspruch dazu, dass die Kette der Ideale (a_i) echt aufsteigend ist. $\square$

Als Nächstes wollen wir zeigen, dass Faktorisierungen in Primelemente im Wesentlichen eindeutig sind.

Lemma 14. *In einem Integritätsring R habe man die Gleichung*

$$p_1 \cdot \dots \cdot p_r = q_1 \cdot \dots \cdot q_s$$

für Primelemente $p_1, \dots, p_r \in R$ und irreduzible Elemente $q_1, \dots, q_s \in R$. Dann gilt $r = s$, und nach Unnummerierung der $q_1, \dots, q_s$ existieren Einheiten $\varepsilon_1, \dots, \varepsilon_r \in R^$ mit $q_i = \varepsilon_i p_i$ für $i = 1, \dots, r$, d. h. p_i ist jeweils assoziiert zu q_i .*

Beweis. Aus $p_1 \cdot \dots \cdot p_r = q_1 \cdot \dots \cdot q_s$ folgt insbesondere $p_1 \mid q_1 \cdot \dots \cdot q_s$. Da p_1 ein Primelement ist, gibt es ein i mit $p_1 \mid q_i$, und wir können durch Unnummerierung der $q_1, \dots, q_s$ annehmen, dass $i = 1$ und somit $p_1 \mid q_1$ gilt. Man hat also eine Gleichung $q_1 = \varepsilon_1 p_1$, wobei ε_1 aufgrund der Irreduzibilität von q_1 eine Einheit ist. Da wir uns in einem Integritätsring befinden, ergibt sich hieraus

$$p_2 \cdot \dots \cdot p_r = \varepsilon_1 q_2 \cdot \dots \cdot q_s.$$

In gleicher Weise zeigt man nun, dass p_2 zu einem der Elemente $q_2, \dots, q_s$ assoziiert ist usw. Man kann also $q_1, \dots, q_s$ so unnummerieren, dass p_i für $i = 1, \dots, r$ jeweils zu q_i assoziiert ist. Insbesondere folgt $r \leq s$. Nach "Auskürzen" aller p_i aus der Gleichung $p_1 \cdot \dots \cdot p_r = q_1 \cdot \dots \cdot q_s$ verbleibt eine Gleichung des Typs

$$1 = q_{r+1} \cdot \dots \cdot q_s,$$

welche zeigt, dass das System der $q_{r+1}, \dots, q_s$ aus Einheiten besteht. Da alle q_i zugleich irreduzibel sind, also keine Einheiten sein können, ist das System leer, und es gilt folglich $r = s$. $\square$

Man sagt, in einem Integritätsring R gelte der *Satz von der eindeutigen Primfaktorzerlegung*, oder auch R sei *faktoriell*, wenn sich jede Nichteinheit $a \in R - \{0\}$ als Produkt von Primelementen in R schreiben lässt. Gemäß Lemma 14 ist eine solche Faktorisierung von a (im Wesentlichen) eindeutig. Benutzen wir weiter, dass jedes Primelement aufgrund von Bemerkung 11 irreduzibel

ist, so kann man mit Lemma 14 schließen, dass sich in einem faktoriellen Ring jede von Null verschiedene Nichteinheit auf (im Wesentlichen) eindeutige Weise als Produkt irreduzibler Elemente schreiben lässt. Man kann darüber hinaus zeigen, dass umgekehrt die letztere Eigenschaft in einem Integritätsring dessen Faktorialität impliziert. Wir wollen jedoch auf Beweise nicht weiter eingehen, sondern nur noch Korollar 13 in neuer Sprechweise formulieren.

Satz 15. *Jeder Hauptidealring ist faktoriell, insbesondere also der Ring $\mathbb{Z}$ der ganzen Zahlen sowie der Polynomring $K[T]$ einer Variablen T über einem Körper K .*

Man kann Primfaktorzerlegungen in einem faktoriellen Ring R weiter standardisieren, indem man in jeder Klasse assoziierter Primelemente eines auswählt und damit ein Repräsentantensystem $P \subset R$ aller Primelemente betrachtet. Man kann dann annehmen, dass in Primfaktorzerlegungen, abgesehen von Einheiten, nur die Primelemente $p \in P$ vorkommen, und man kann darüber hinaus gleiche Primfaktoren zu Potenzen zusammenfassen. Es besitzt dann jedes Element $a \in R - \{0\}$ eine Primfaktorzerlegung der Form

$$a = \varepsilon \prod_{p \in P} p^{\mu_p(a)}$$

mit einer Einheit $\varepsilon \in R^*$ und Exponenten $\mu_p(a) \in \mathbb{Z}$, die für fast alle $p \in P$ trivial sind. Die Eindeutigkeit der Primfaktorzerlegung besagt dann, dass ε und die $\mu_p(a)$ jeweils eindeutig durch a bestimmt sind. Im Übrigen sieht man unmittelbar ein, dass für Elemente $a, b \in R - \{0\}$ die Teilbarkeitsbedingung $a|b$ genau dann erfüllt ist, wenn $\mu_p(a) \leq \mu_p(b)$ für alle $p \in P$ gilt.

In $\mathbb{Z}$ gibt es nur die Einheiten 1 und -1 , und es ist üblich, P als das System aller positiven Primelemente zu definieren. Im Polynomring $K[T]$ über einem Körper K dagegen besteht die Einheitengruppe aufgrund von 5.1/2 aus allen nicht-trivialen konstanten Polynomen, stimmt also mit der Einheitengruppe K^* von K überein. Daher gibt es zu jedem Primpolynom genau ein assoziiertes Primpolynom, welches *normiert* ist, d. h. 1 als höchsten Koeffizienten besitzt, und man definiert P als das System aller normierten Primpolynome.

Wie gewöhnlich lässt sich für zwei von Null verschiedene Elemente $a, b \in R$ mit Primfaktorzerlegung

$$a = \varepsilon \prod_{p \in P} p^{\mu_p(a)}, \quad b = \delta \prod_{p \in P} p^{\mu_p(b)}$$

der *größte gemeinsame Teiler*

$$\text{ggT}(a, b) = \prod_{p \in P} p^{\min(\mu_p(a), \mu_p(b))}$$

erklären. Dies ist ein gemeinsamer Teiler d von a und b mit der charakterisierenden Eigenschaft, dass aus $t|a$ und $t|b$ stets $t|d$ folgt.

Entsprechend kann man zu a und b das *kleinste gemeinsame Vielfache*

$$\text{kgV}(a, b) = \prod_{p \in P} p^{\max(\mu_p(a), \mu_p(b))}$$

betrachten. Dies ist ein gemeinsames Vielfaches v von a und b , so dass jedes weitere gemeinsame Vielfache von a und b auch ein Vielfaches von v ist. Man beachte jedoch, dass die Elemente $\text{ggT}(a, b)$ und $\text{kgV}(a, b)$ nur bis auf Einheiten wohldefiniert sind, da sie außer von a und b noch von der speziellen Wahl von P abhängen.

In Hauptidealringen lässt sich der größte gemeinsame Teiler zweier Elemente $a, b \in R$ idealtheoretisch charakterisieren, was vielfach von Nutzen ist. Hierzu betrachtet man

$$Ra + Rb := \{ra + sb; r, s \in R\}$$

als Ideal in R , wobei die definierenden Eigenschaften eines Ideals leicht zu verifizieren sind.

Satz 16. *Es seien a, b zwei von Null verschiedene Elemente eines Hauptidealrings R . Für den größten gemeinsamen Teiler $d = \text{ggT}(a, b)$ gilt dann*

$$Ra + Rb = Rd.$$

Insbesondere gibt es eine Gleichung $ra + sb = d$ mit Elementen $r, s \in R$, die notwendig teilerfremd sind, d. h. $\text{ggT}(r, s) = 1$ erfüllen.

Beweis. Das Ideal $Ra + Rb \subset R$ ist ein Hauptideal, etwa $Ra + Rb = Rd'$. Dann folgt wegen $a, b \in Rd'$, dass d' ein gemeinsamer Teiler von a, b und damit auch von d ist. Andererseits besteht wegen $Ra + Rb = Rd'$ eine Gleichung des Typs $ra + sb = d'$ mit gewissen Elementen $r, s \in R$. Dies zeigt, dass jeder gemeinsame Teiler von a, b auch ein Teiler von d' ist. Insbesondere gilt also $d \mid d'$. Zusammen mit $d' \mid d$ ergibt sich gemäß Bemerkung 9, dass d und d' assoziiert sind. Somit gilt $Ra + Rb = Rd$, wie behauptet, und man hat eine Gleichung des Typs $ra + sb = d$. Letztere besagt, dass jeder gemeinsame Teiler von a, b , multipliziert mit $\text{ggT}(r, s)$, einen Teiler von d ergibt. Dies ist aber nur im Falle $\text{ggT}(r, s) = 1$ möglich. $\square$

Abschließend wollen wir noch aus Satz 16 eine spezielle Eigenschaft von Primelementen in Hauptidealringen folgern.

Korollar 17. *Es sei R ein Hauptidealring und $p \in R - \{0\}$. Dann ist äquivalent:*

- (i) *p ist ein Primelement.*
- (ii) *Der Restklassenring $R/(p)$ ist ein Körper.*

Beweis. Sei zunächst p ein Primelement. Insbesondere ist p dann keine Einheit und somit $R/(p)$ nicht der Nullring. Um einzusehen, dass $R/(p)$ ein Körper ist, wähle man $\bar{a} \in R/(p) - \{0\}$. Es ist zu zeigen, dass es ein $\bar{b} \in R/(p)$ mit $\bar{b} \cdot \bar{a} = 1$

gibt oder, in äquivalenter Formulierung, dass es zu $a \in R - (p)$ eine Gleichung der Form

$$ba + rp = 1$$

mit $b, r \in R$ gibt. Letzteres folgt aber aus Satz 16, da a und p offenbar teilerfremd sind.

Wenn andererseits p kein Primelement ist, so ist p entweder eine Einheit, oder aber es gibt von Null verschiedene Nichteinheiten $a, b \in R$ mit $p \nmid a$, $p \nmid b$, sowie $p \mid ab$. Im ersten Fall ist der Restklassenring $R/(p)$ der Nullring. Im zweiten sind die Restklassen $\bar{a}, \bar{b} \in R/(p)$ zu a, b von Null verschieden, erfüllen aber $\bar{a} \cdot \bar{b} = 0$. Es ist also $R/(p)$ in beiden Fällen kein Integritätsring und damit insbesondere kein Körper. $\square$

Als Anwendung sehen wir, dass für $p \in \mathbb{Z}$ der Restklassenring $\mathbb{Z}/p\mathbb{Z}$ genau dann ein Körper ist, wenn p prim ist. Insbesondere ist $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ für eine Primzahl $p \in \mathbb{N}$ ein Körper mit p Elementen.

Genauso folgt für einen Körper K und Polynome $f \in K[T]$, dass der Restklassenring $K[T]/(f)$ genau dann ein Körper ist, wenn f prim ist. In $\mathbb{R}[T]$ sind beispielsweise die Polynome $T - 1$ und $T^2 + 1$ irreduzibel und damit auch prim. Es gilt

$$\mathbb{R}[T]/(T - 1) \simeq \mathbb{R}, \quad \mathbb{R}[T]/(T^2 + 1) \simeq \mathbb{C},$$

wie man leicht mit Hilfe des Homomorphiesatzes 5.1/10 zeigen kann. Im Übrigen kann man zeigen, dass die primen Polynome in $\mathbb{R}[T]$ gerade aus allen Polynomen vom Grad 1 sowie den nullstellenfreien Polynomen vom Grad 2 gebildet werden.

Schließlich wollen wir noch die so genannte *Charakteristik* eines Körpers definieren. Ist K ein Körper, so gibt es einen eindeutig bestimmten Ringhomomorphismus $\varphi: \mathbb{Z} \rightarrow K$. Dieser bildet eine natürliche Zahl n ab auf die n -fache Summe $n \cdot 1_K$ des Einselementes $1_K \in K$ und entsprechend $-n$ auf $-(n \cdot 1_K)$. Der Kern von φ ist ein Ideal in $\mathbb{Z}$, also ein Hauptideal, und wird damit von einem eindeutig bestimmten Element $p \in \mathbb{N}$ erzeugt. Es ist p entweder 0 oder ansonsten die kleinste positive natürliche Zahl mit $p \cdot 1_K = 0$. Man nennt p die *Charakteristik* von K ; diese ist entweder 0 oder aber prim, wie man ähnlich wie im zweiten Teil des Beweises zu Korollar 17 sehen kann.

Aufgaben

- Man betrachte die folgenden Polynome $f, g \in \mathbb{R}[T]$ und dividiere jeweils f mit Rest durch g :
 - $f = T^6 + 3T^4 + T^3 - 2$, $g = T^2 - 2T + 1$,
 - $f = T^n - 1$, $g = T - 1$, mit $n \in \mathbb{N} - \{0\}$,
 - $f = T^n + T^{n-1} + \dots + 1$, $g = T + 1$, mit $n \in \mathbb{N} - \{0\}$.
- Es seien a, b von Null verschiedene Elemente eines Hauptidealrings R . Man zeige $Ra \cap Rb = Rv$ für $v = \text{kgV}(a, b)$.

3. Man bestimme alle Unterringe von $\mathbb{Q}$.
4. Es sei R ein Integritätsring und $p \in R - \{0\}$. Man zeige, dass p genau dann prim ist, wenn $R/(p)$ ein Integritätsring ist.
5. Man bestimme die Primfaktorzerlegung des Polynoms $T^4 - 1$ im Polynomring $\mathbb{R}[T]$.
6. Man zeige, dass der Polynomring $\mathbb{Z}[T]$ kein Hauptidealring ist.
7. Man zeige, dass $\mathbb{Z} + \mathbb{Z}i = \{x + yi \in \mathbb{C}; x, y \in \mathbb{Z}\}$ einen Unterring des Körpers der komplexen Zahlen bildet und ein Hauptidealring ist.
8. Man zeige, dass es in $\mathbb{Z}$ unendlich viele paarweise nicht-assoziierte Primelemente gibt. Gleiches gilt für den Polynomring $K[T]$ über einem Körper K .
9. Es sei $\mathbb{F}$ ein endlicher Körper der Charakteristik p , wobei p den Kern des kanonischen Ringhomomorphismus $\mathbb{Z} \rightarrow \mathbb{F}$ erzeugt. Man zeige:
 - (i) p ist eine Primzahl.
 - (ii) Es besteht $\mathbb{F}$ aus p^r Elementen, wobei r eine geeignete natürliche Zahl ist.
10. Es sei K ein Körper und A eine K -Algebra mit $\dim_K A < \infty$. Für ein Element $a \in A$ zeige man: Es existiert ein eindeutig bestimmtes normiertes Polynom $f \in K[T]$ kleinsten Grades mit $f(a) = 0$.

5.3 Nullstellen von Polynomen

Es sei K ein Körper (oder allgemeiner ein kommutativer Ring) und A eine K -Algebra. Ein Element $t \in A$ heißt *Nullstelle* eines Polynoms $f \in K[T]$, wenn $f(t) = 0$ gilt, d. h. wenn das Bild von f unter dem Einsetzungshomomorphismus $K[T] \rightarrow A$, der t anstelle der Variablen T einsetzt (vgl. 5.1/7), trivial ist. Um ein Beispiel zu geben, betrachte man den Endomorphismenring A eines K -Vektorraums V ; dieser wird zu einer K -Algebra unter dem Ringhomomorphismus $K \rightarrow A, c \mapsto c \cdot \text{id}_V$. Für $\dim_K V > 1$ ist leicht zu sehen, dass das Polynom $T^2 \in K[T]$ außer dem Nullelement $0 \in A$ noch weitere Nullstellen besitzt, sogar unendlich viele, wenn K unendlich viele Elemente hat. Die Gleichung $\varphi^2 = 0$ für einen Endomorphismus $\varphi: V \rightarrow V$ ist nämlich gleichbedeutend mit $\text{im } \varphi \subset \ker \varphi$.

Wir wollen uns zunächst aber nur für Nullstellen von Polynomen $f \in K[T]$ in K interessieren, wobei wir K als Körper voraussetzen. Aufgrund der Nullteilerfreiheit von K sind dann stärkere Aussagen möglich, beispielsweise ist das Nullelement $0 \in K$ die einzige Nullstelle in K zu $T^2 \in K[T]$.

Satz 1. *Sei $\alpha \in K$ Nullstelle eines Polynoms $f \in K[T]$. Dann existiert ein Polynom $g \in K[T]$ mit*

$$f = (T - \alpha) \cdot g,$$

wobei g durch diese Gleichung eindeutig bestimmt ist.

Beweis. Division mit Rest von f durch $(T - \alpha)$ führt zu einer Gleichung

$$f = (T - \alpha) \cdot g + r$$

mit $r \in K$. Setzt man hierin α anstelle von T ein, so ergibt sich wegen $f(\alpha) = 0$ unmittelbar $r = r(\alpha) = 0$ und damit die gewünschte Gleichung $f = (T - \alpha) \cdot g$. Die Eindeutigkeit von g folgt aus der Nullteilerfreiheit von K oder aus der Eindeutigkeit der Division mit Rest; vgl. 5.2/2. $\square$

Korollar 2. Sei $f \in K[T]$, $f \neq 0$. Dann besitzt f nur endlich viele paarweise verschiedene Nullstellen $\alpha_1, \dots, \alpha_r \in K$, wobei $r \leq \text{grad } f$ gilt. Weiter existieren $n_1, \dots, n_r \in \mathbb{N} - \{0\}$ sowie ein Polynom $g \in K[T]$ ohne Nullstellen in K mit

$$f = \prod_{i=1}^r (T - \alpha_i)^{n_i} \cdot g.$$

Dabei sind die Exponenten n_i sowie das Polynom g eindeutig durch f bestimmt.

Beweis. Man betrachte Zerlegungen der Form

$$f = \prod_{i=1}^r (T - \alpha_i)^{n_i} \cdot g$$

mit paarweise verschiedenen Nullstellen $\alpha_1, \dots, \alpha_r$ von f (wobei r variieren darf), Exponenten $n_i \in \mathbb{N} - \{0\}$ und einem Polynom $g \in K[T]$. Dann gilt stets

$$\text{grad } f = \sum_{i=1}^r n_i + \text{grad } g,$$

und wir können eine solche Zerlegung finden, für die $\text{grad } g$ minimal ist. Dann ist g aber wie gewünscht ohne Nullstellen in K , da man ansonsten gemäß Satz 1 von g einen Linearfaktor der Form $T - \alpha$ abspalten könnte und dies zu einer Zerlegung mit einem echt kleineren Grad von g führen würde. Man erkennt dann, dass $\alpha_1, \dots, \alpha_r$ aufgrund der Nullteilerfreiheit von K die einzigen Nullstellen von f sind und dass $r \leq \text{grad } f$ gilt.

Die Faktoren $(T - \alpha_i)$ sind irreduzibel und damit insbesondere prim. Die Eindeutigkeitsaussage folgt daher leicht aus der Eindeutigkeit der Primfaktorzerlegung in $K[T]$. Man benutze dabei, dass in der Primfaktorzerlegung von g lediglich Faktoren vom Grad ≥ 2 vorkommen können, da g keine Nullstellen in K besitzt. $\square$

In der vorstehenden Situation nennt man n_i die *Vielfachheit* der Nullstelle α_i . Weiter bezeichnet man einen Körper K als *algebraisch abgeschlossen*, wenn jedes nicht-konstante Polynom $f \in K[T]$ (mindestens) eine Nullstelle in K besitzt. In der Zerlegung von Korollar 2 ist g dann konstant. Man kann daher sagen, dass ein Körper K genau dann algebraisch abgeschlossen ist, wenn sich jedes nicht-konstante Polynom $f \in K[T]$ als Produkt von Linearfaktoren, d. h. Polynomen vom Grad 1 schreiben lässt oder, in äquivalenter Weise, wenn die

irreduziblen Polynome in $K[T]$ gerade die Polynome vom Grad 1 sind. Wir wollen in diesem Zusammenhang den so genannten *Fundamentalsatz der Algebra* formulieren.

Theorem 3. *Der Körper $\mathbb{C}$ der komplexen Zahlen ist algebraisch abgeschlossen.*

Der Beweis erfordert Hilfsmittel, die über die lineare Algebra hinausgehen. Er wird üblicherweise im Rahmen der Funktionentheorie- oder Algebra-Vorlesungen geführt. Als Beispiel wollen wir hier nur noch anfügen, dass die Polynome

$$\begin{aligned} T^2 - 2 &\in \mathbb{Q}[T], \\ T^2 + 1 &\in \mathbb{R}[T], \\ T^2 + T + 1 &\in \mathbb{F}_2[T] \end{aligned}$$

keine Nullstellen in den jeweiligen Körpern besitzen. Da es sich um Polynome vom Grad 2 handelt, folgt hieraus, dass diese irreduzibel und damit prim sind.

Aufgaben

1. Es sei K ein Körper. Für ein Polynom $f = \sum_{i \in \mathbb{N}} a_i T^i \in K[T]$ definiere man dessen Ableitung durch $f' := \sum_{i > 0} i a_i T^{i-1}$, wobei $i a_i$ jeweils als i -fache Summe von a_i zu verstehen ist. Man zeige: Ein Element $\alpha \in K$ ist genau dann eine mehrfache Nullstelle (d. h. der Ordnung > 1) von f , wenn α Nullstelle von $\text{ggT}(f, f')$ ist.
2. Es sei K ein Körper und A eine K -Algebra. Für Polynome $f, g \in K[T] - \{0\}$ und $h = \text{ggT}(f, g)$ zeige man: Ist $a \in A$ eine gemeinsame Nullstelle von f und g , so ist a auch Nullstelle von h . (Hinweis: Man benutze die in 5.2/16 beschriebene Charakterisierung des größten gemeinsamen Teilers.)
3. Für eine komplexe Zahl $\alpha = u + iv \in \mathbb{C}$ mit Realteil u und Imaginärteil v sei die zugehörige konjugiert komplexe Zahl definiert durch $\bar{\alpha} = u - iv$. Man zeige, dass ein normiertes Polynom $f \in \mathbb{R}[T]$ genau dann prim ist, wenn es von der Form $f = T - \alpha$ mit $\alpha \in \mathbb{R}$ oder $f = (T - \alpha)(T - \bar{\alpha})$ mit $\alpha \in \mathbb{C} - \mathbb{R}$ ist. (Hinweis: Man betrachte die Abbildung $\mathbb{C} \rightarrow \mathbb{C}$, $\alpha \mapsto \bar{\alpha}$, welche die Eigenschaften eines $\mathbb{R}$ -Algebraisomorphismus besitzt, und setze diese fort zu einem $\mathbb{R}[T]$ -Algebraisomorphismus $\mathbb{C}[T] \rightarrow \mathbb{C}[T]$.)

6. Normalformentheorie

Vorbemerkungen

In diesem Kapitel geht es darum, für endlich-dimensionale K -Vektorräume V die Struktur der Endomorphismen von V zu klären. Was aber hat man unter der *Struktur* eines Endomorphismus $f: V \longrightarrow V$ zu verstehen? Man kann beispielsweise die folgenden Fragen stellen:

1. Gibt es nicht-triviale Untervektorräume $U \subset V$ mit $f(U) \subset U$, auf denen $f|_U$ von besonders einfacher Gestalt ist, z. B. $f|_U = \lambda \operatorname{id}_U$ mit einem Skalar $\lambda \in K$?
2. Um f auf ganz V zu beschreiben: Kann man V in eine direkte Summe nicht-trivialer Untervektorräume $V = \bigoplus_{i=1}^r U_i$ zerlegen mit $f(U_i) \subset U_i$, so dass sich $f|_{U_i}$ in signifikanter Weise charakterisieren lässt? Gibt es eine feinste Zerlegung dieses Typs, und ist diese in irgendeiner Weise eindeutig charakterisiert?

Dies sind bereits die wichtigsten Fragen, die wir untersuchen wollen. Zunächst ist die Beantwortung der ersten Frage relativ einfach. Wir betrachten für $\lambda \in K$ den Endomorphismus $f - \lambda \operatorname{id}$ von V . Sein Kern gibt genau denjenigen (maximalen) Untervektorraum von V an, auf dem sich f wie $\lambda \operatorname{id}$ verhält, und dieser Unterraum ist genau dann nicht-trivial, wenn der Kern von $f - \lambda \operatorname{id}$ nicht-trivial ist, also gemäß 2.1/11 genau dann, wenn $f - \lambda \operatorname{id}$ nicht invertierbar ist, und damit nach 4.3/3 genau dann, wenn $\det(f - \lambda \operatorname{id}) = 0$ gilt. Es ist also die Gleichung $\det(f - \lambda \operatorname{id}) = 0$ für $\lambda \in K$ zu lösen, und wir werden damit automatisch dazu veranlasst, das so genannte *charakteristische Polynom* $\chi_f \in K[T]$ zu f zu betrachten, das entsteht, wenn wir auf $\det(\lambda \operatorname{id} - f)$ die Definition der Determinante einer beschreibenden Matrix anwenden, dabei jedoch anstelle von λ die Variable T vorsehen.

Die Nullstellen von χ_f in K werden als *Eigenwerte* zu f bezeichnet. Für einen solchen Eigenwert λ heißt $V_\lambda = \ker(\lambda \operatorname{id} - f)$ der zu λ gehörige *Eigenraum*, und die Elemente von $V_\lambda - \{0\}$ werden als *Eigenvektoren* zum Eigenwert λ bezeichnet. Wir werden zeigen, dass Eigenvektoren zu verschiedenen Eigenwerten stets linear unabhängig sind, und daraus folgern, dass die Summe der Eigenräume zu den verschiedenen Eigenwerten von f stets direkt ist. Stimmt diese Summe bereits mit V überein, so ist f *diagonalisierbar*, womit wir meinen,

dass V eine Basis bestehend aus Eigenvektoren zu f besitzt und was zur Folge hat, dass die zugehörige beschreibende Matrix von f eine Diagonalmatrix ist. Diese Situation ist beispielsweise gegeben, wenn die Anzahl der verschiedenen Eigenwerte von f gleich der Dimension von V ist. Wir erhalten damit auch eine erste (partielle) Antwort auf die eingangs gestellte Frage 2.

Es ist relativ leicht einzusehen, dass lineare Abbildungen $f: V \rightarrow V$ im Allgemeinen nicht diagonalisierbar sind. Beispielsweise ist eine Drehung um 0 im $\mathbb{R}^2$ nicht diagonalisierbar, es sei denn, der Drehwinkel beträgt 0° oder 180° . So wird man zur Beantwortung der Frage 2 noch nach anderen Möglichkeiten suchen müssen, um Untervektorräume $U \subset V$ mit $f(U) \subset U$, d. h. so genannte *f-invariante* Untervektorräume, zu konstruieren. Folgende Beobachtung ist hierbei grundlegend: Man betrachte zu einem Vektor $x \in V$ den Untervektorraum $U \subset V$, der von den Elementen $x, f(x), f^2(x), \dots$ erzeugt wird. Dann ist U offenbar ein *f-invarianter* Untervektorraum in V , offenbar der kleinste, der x enthält. Wir nennen U den von x erzeugten *f-zyklischen* Untervektorraum von V . Seine Struktur lässt sich leicht beschreiben, wenn man Ergebnisse aus Kapitel 5 über den Polynomring $K[T]$ verwendet. Man kann nämlich die K -lineare Abbildung

$$\varphi: K[T] \rightarrow U, \quad \sum_{i \in \mathbb{N}} c_i T^i \mapsto \sum_{i \in \mathbb{N}} c_i f^i(x),$$

betrachten und stellt dabei fest, dass $\ker \varphi$ ein Ideal in $K[T]$ ist. Denn für

$$p = \sum_{i=0}^r c_i T^i \in K[T], \quad q = \sum_{j=0}^s d_j T^j \in \ker \varphi$$

gilt

$$\begin{aligned} \varphi(pq) &= \varphi\left(\sum_{k=0}^{r+s} \left(\sum_{i+j=k} c_i d_j\right) T^k\right) = \sum_{k=0}^{r+s} \left(\sum_{i+j=k} c_i d_j\right) f^k(x) \\ &= \left(\sum_{i=0}^r c_i f^i\right)\left(\sum_{j=0}^s d_j f^j\right)(x) = \left(\sum_{i=0}^r c_i f^i\right)(0) = 0, \end{aligned}$$

d. h. $\ker \varphi$ ist insbesondere abgeschlossen unter Multiplikation mit Elementen aus $K[T]$. Nun wissen wir aber, dass $K[T]$ ein Hauptidealring ist, dass es folglich ein eindeutig bestimmtes normiertes Polynom kleinsten Grades

$$p(T) = T^r + c_1 T^{r-1} + \dots + c_r \in \ker \varphi$$

gibt und dass dieses das Ideal erzeugt; denn $\ker \varphi$ besteht nicht nur aus dem Nullpolynom, da $K[T]$ als K -Vektorraum von unendlicher, V aber von endlicher Dimension ist. Hieraus gewinnt man die Gleichung

$$p(f)(x) = f^r(x) + c_1 f^{r-1}(x) + \dots + c_r x = 0,$$

und diese zeigt in induktiver Weise, dass U bereits von $x, f^1(x), \dots, f^{r-1}(x)$ erzeugt wird und, da es in $\ker \varphi$ kein nicht-triviales Polynom vom Grade $< r$

gibt, dass diese Elemente sogar eine Basis von U bilden. Bezüglich dieser Basis wird $f|_U$ dann durch die so genannte *Begleitmatrix*

$$\begin{pmatrix} 0 & & & & -c_0 \\ 1 & 0 & & & -c_1 \\ & 1 & . & & -c_2 \\ & & . & . & \dots \\ & & & . & \dots \\ & & & . & 0 & -c_{n-2} \\ & & & & 1 & -c_{n-1} \end{pmatrix}$$

des Polynoms p beschrieben.

Um die eingangs gestellte Frage 2 vollständig zu klären, ist noch zu untersuchen, in wie weit sich V als eine direkte Summe f -zyklischer Untervektorräume darstellen lässt. Dass dies in der Tat stets möglich ist, und zwar mit so genannten *f -unzerlegbaren* (ebenfalls f -zyklischen) Untervektorräumen, die keine weitere Zerlegung in eine direkte Summe f -invarianter Unterräume mehr zulassen, und mit zugehörigen normierten Polynomen aus $K[T]$, die eindeutig durch f bestimmt sind, ist ein tief liegendes Resultat, dessen Beweis einigen Aufwand erfordert. Um die eigentlichen Gründe für das Zustandekommen dieses Resultats aufzudecken, werden wir die so genannte *Elementarteilertheorie* behandeln, und zwar für *Moduln* über Hauptidealringen. Ein Modul über einem Ring ist formal genauso definiert wie ein Vektorraum über einem Körper, nur dass man als Skalarenbereich anstelle eines Körpers einen Ring vorsieht. Dass man beispielsweise einen K -Vektorraum mit einem Endomorphismus $f: V \rightarrow V$ als einen Modul über dem Polynomring $K[T]$ auffassen sollte, wobei man für $x \in V$ das Produkt Tx durch $f(x)$ definiert, wird durch die obige Betrachtung f -zyklischer Untervektorräume nahe gelegt. In diesem Sinne ist der von einem Vektor $x \in V$ erzeugte f -zyklische Untervektorraum $U \subset V$ zu sehen als der von x erzeugte $K[T]$ -Untermodul von V .

Obwohl Moduln als “Vektorräume” über Ringen interpretiert werden können, gibt es dennoch gravierende Unterschiede zu Vektorräumen über Körpern, die durch das Phänomen der so genannten *Torsion* verursacht sind. Für einen Modul M über einem Ring R gibt es nämlich im Allgemeinen von Null verschiedene Elemente $r \in R$ und $m \in M$ mit $rm = 0$, wobei dann r natürlich keine Einheit sein kann, da ansonsten $m = r^{-1}(rm) = 0$ folgen würde. Insbesondere kann ein solcher Modul keine Basis besitzen. Gibt es zu jedem $m \neq 0$ aus M ein $r \neq 0$ in R mit $rm = 0$, so bezeichnet man M als einen *Torsionsmodul*. Beispielsweise ist in der obigen Situation auch V als $K[T]$ -Modul ein Torsionsmodul. Da V von endlicher Dimension ist, existieren nämlich nicht-triviale Polynome $p \in K[T]$ mit $p \cdot V = 0$. Wir werden insbesondere sehen, dass es wiederum ein eindeutig bestimmtes normiertes Polynom kleinsten Grades $p_f \in K[T]$ mit $p_f \cdot V = 0$ gibt. Man nennt p_f das *Minimalpolynom* zu f , und wir werden mit dem Satz von Cayley-Hamilton zeigen, dass p_f stets ein Teiler des charakteristischen Polynoms χ_f ist und damit einen Grad $\leq n$ besitzt.

Die mittels der Elementarteilertheorie gewonnene Zerlegung von V in f -zyklische bzw. f -invariante Unterräume werden wir schließlich dazu verwenden, um f mittels kanonisch zugeordneter Matrizen zu beschreiben. Ist beispielsweise $V = \bigoplus_{i=1}^r U_i$ eine Zerlegung in f -zyklische Untervektorräume, wobei $U_i \simeq K[T]/(p_i)$ mit normierten Polynomen $p_i \in K[T]$ im Sinne von $K[T]$ -Moduln gelte, so kann man für jedes U_i eine K -Basis wählen, derart dass $f|_{U_i}$ bezüglich dieser Basis durch die Begleitmatrix $A(p_i)$ zu p_i dargestellt wird. Setzt man die einzelnen Basen der U_i zu einer Gesamtbasis von V zusammen, so ist die zugehörige f beschreibende Matrix von der Form $\text{Diag}(A(p_1), \dots, A(p_r))$, d. h. eine Art “Diagonalmatrix”, auf deren Diagonalen die Kästchen $A(p_i)$ angeordnet sind. Geht man von irgendeiner Matrix $A \in K^{n \times n}$ aus, die f beschreibt, und sind die Polynome p_i Primpotenzen, so ist $A' = \text{Diag}(A(p_1), \dots, A(p_r))$ bereits die so genannte *allgemeine Normalform* von A . Im Übrigen werden wir auch die *Jordansche Normalform* von A betrachten (sofern das charakteristische Polynom χ_f in lineare Faktoren zerfällt) und abschließend zeigen, wie man Normalformen explizit berechnen kann, indem man die Elementarteiler der Matrix $TE - A \in K[T]^{n \times n}$ bestimmt; dabei sei $E \in K^{n \times n}$ die Einheitsmatrix.

6.1 Eigenwerte und Eigenvektoren

Wir kehren nunmehr zur Theorie der Vektorräume über einem Körper K zurück und betrachten zunächst eine K -lineare Abbildung $f: V \rightarrow W$ zwischen endlich-dimensionalen K -Vektorräumen V und W . Ist dann $X = (x_1, \dots, x_n)$ eine Basis von V und $Y = (y_1, \dots, y_m)$ eine Basis von W , so lässt sich f durch eine zugehörige Matrix $A_{f,X,Y}$ beschreiben; vgl. 3.1/3. Durch geschickte Wahl von X und Y kann man erreichen, dass $A_{f,X,Y}$ von möglichst einfacher Gestalt ist. So hatten wir in 3.4/7 gesehen, dass es Basen X' von V und Y' von W gibt mit

$$A_{f,X',Y'} = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix};$$

dabei bezeichnet $E_r \in K^{r \times r}$ die Einheitsmatrix und r den Rang von f . Weiter besteht die Relation

$$A_{f,X',Y'} = (A_{\text{id},Y',Y})^{-1} \cdot A_{f,X,Y} \cdot A_{\text{id},X',X}$$

mit den Basiswechselmatrizen $A_{\text{id},X',X}$ und $A_{\text{id},Y',Y}$; vgl. 3.4/4. Unter Benutzung der bijektiven Korrespondenz zwischen linearen Abbildungen und Matrizen können wir daher auch sagen, dass es zu einer Matrix $A \in K^{m \times n}$ vom Rang r stets invertierbare Matrizen $S \in \text{Gl}(m, K)$ und $T \in \text{Gl}(n, K)$ mit

$$S^{-1} \cdot A \cdot T = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

gibt; vgl. hierzu auch 3.4/8.

Wir wollen im Weiteren ein entsprechendes Problem für Endomorphismen $f: V \rightarrow V$ studieren. Genauer soll durch geeignete Wahl einer Basis X von

V erreicht werden, dass die Matrix $A_{f,X,X}$ von möglichst einfacher Gestalt ist. Übersetzt in die Sprache der Matrizen bedeutet dies: Ausgehend von einer Matrix $A \in K^{n \times n}$ ist eine invertierbare Matrix $S \in \text{Gl}(n, K)$ zu konstruieren, derart dass die Matrix $S^{-1} \cdot A \cdot S$ von möglichst einfacher Gestalt ist, beispielsweise eine Diagonalmatrix ist:

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \ddots \\ & & & \lambda_n \end{pmatrix}$$

Dabei sei erwähnt, dass eine solche Diagonalgestalt allerdings nicht in allen Fällen zu erreichen ist. Um eine bequeme Sprechweise für unser Problem zu haben, sagen wir:

Definition 1. Zwei Matrizen $A, B \in K^{n \times n}$ heißen *ähnlich*, wenn es eine invertierbare Matrix $S \in \text{Gl}(n, K)$ mit $B = S^{-1} \cdot A \cdot S$ gibt.

Man sieht unmittelbar, dass die Ähnlichkeit von Matrizen aus $K^{n \times n}$ eine Äquivalenzrelation darstellt. Somit zerfällt $K^{n \times n}$ in disjunkte Klassen ähnlicher Matrizen.

Bemerkung 2. Zu einer Matrix $A \in K^{n \times n}$ betrachte man einen n -dimensionalen K -Vektorraum V mit einer Basis X und den (eindeutig bestimmten) Endomorphismus $f: V \rightarrow V$ mit $A_{f,X,X} = A$ vgl. 3.3/2. Für eine weitere Matrix $B \in K^{n \times n}$ ist dann äquivalent:

- (i) A und B sind ähnlich.
- (ii) Es existiert eine Basis X' von V mit $A_{f,X',X'} = B$.

Beweis. Seien zunächst A und B ähnlich, gelte also $B = S^{-1} \cdot A \cdot S$ mit $S \in \text{Gl}(n, K)$. Fassen wir dann die Matrix S gemäß 3.4/1 (und den sich daran anschließenden Erläuterungen) als Basiswechselmatrix auf, so erhalten wir eine Basis X' von V mit $S = A_{\text{id},X',X}$, und es folgt mit 3.4/4

$$A_{f,X',X'} = (A_{\text{id},X',X})^{-1} \cdot A_{f,X,X} \cdot A_{\text{id},X',X} = S^{-1} \cdot A \cdot S = B,$$

d. h. Bedingung (ii) ist erfüllt.

Ist umgekehrt Bedingung (ii) gegeben, so zeigt die Gleichung

$$A_{f,X',X'} = (A_{\text{id},X',X})^{-1} \cdot A_{f,X,X} \cdot A_{\text{id},X',X},$$

dass A und B ähnlich sind. □

Wir wollen uns nun mit der Frage beschäftigen, wann eine gegebene Matrix $A \in K^{n \times n}$ zu einer Diagonalmatrix ähnlich ist. Dazu führen wir folgende Sprechweise ein:

Definition 3. Eine Matrix $A \in K^{n \times n}$ heißt *diagonalisierbar*, wenn sie zu einer Diagonalmatrix ähnlich ist.

Ein Endomorphismus $f: V \longrightarrow V$ eines endlich-dimensionalen K -Vektorraums V heißt diagonalisierbar, wenn die beschreibende Matrix $A_{f,X,X}$ für eine Basis X von V (und damit für alle Basen von V) diagonalisierbar ist.

Aus Bemerkung 2 kann man ablesen, dass ein Endomorphismus $f: V \longrightarrow V$ eines endlich-dimensionalen K -Vektorraums V genau dann diagonalisierbar ist, wenn es eine Basis $X = (x_1, \dots, x_n)$ von V mit

$$A_{f,X,X} = \begin{pmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \ddots \\ & & & \lambda_n \end{pmatrix}$$

gibt, so dass also $f(x_i) = \lambda_i x_i$ für gewisse Konstanten $\lambda_i \in K$ gilt, $i = 1, \dots, n$. Wir werden in diesem Zusammenhang folgende Terminologie verwenden:

Definition 4. Es sei $f: V \longrightarrow V$ ein Endomorphismus eines K -Vektorraums V . Eine Konstante $\lambda \in K$ heißt Eigenwert zu f , wenn es einen Vektor $a \in V - \{0\}$ mit $f(a) = \lambda a$ gibt. Man nennt in diesem Falle a einen Eigenvektor von f zum Eigenwert λ . Für eine Matrix $A \in K^{n \times n}$ seien Eigenwerte und -Vektoren erklärt als Eigenwerte und -Vektoren der zugehörigen linearen Abbildung $K^n \longrightarrow K^n$, $x \longmapsto Ax$.

Eigenvektoren sind definitionsgemäß immer von 0 verschieden, und wir können formulieren:

Bemerkung 5. Ein Endomorphismus $f: V \longrightarrow V$ eines endlich-dimensionalen K -Vektorraums V ist genau dann diagonalisierbar, wenn es in V eine Basis aus Eigenvektoren bezüglich f gibt.

Als Anwendung der Beschreibung linearer Abbildungen mittels Matrizen, vgl. 3.1/7, ergibt sich:

Bemerkung 6. Es sei $f: V \longrightarrow V$ ein Endomorphismus eines endlich-dimensionalen K -Vektorraums V mit Basis X . Für $\lambda \in K$ ist dann äquivalent:

- (i) λ ist Eigenwert von f .
- (ii) λ ist Eigenwert von $A_{f,X,X}$.

Beweis. Sei $\dim_K V = n$. Wir benutzen das kommutative Diagramm

$$\begin{array}{ccc} V & \xrightarrow{f} & V \\ \kappa_X \downarrow & & \downarrow \kappa_X \\ K^n & \xrightarrow{\tilde{f}} & K^n \end{array}$$

aus 3.1/8. Dabei ist κ_X derjenige Isomorphismus, der einem Vektor $v \in V$ den zugehörigen Koordinatenspaltenvektor $v_X \in K^n$ zuordnet, sowie $\tilde{f}: K^n \longrightarrow K^n$

die durch $u \mapsto A_{f,X,X} \cdot u$ erklärte Abbildung. Ist nun $a \in V$ ein Eigenvektor zu f mit Eigenwert $\lambda \in K$, so gilt insbesondere $a \neq 0$ und damit auch $\kappa_X(a) \neq 0$. Weiter folgt aufgrund der Kommutativität des obigen Diagramms

$$\tilde{f}(\kappa_X(a)) = \kappa_X(f(a)) = \kappa_X(\lambda a) = \lambda \kappa_X(a),$$

d. h. $\kappa_X(a)$ ist Eigenvektor zu $\tilde{f}$, ebenfalls zum Eigenwert λ . Ist umgekehrt $b \in K^n$ ein Eigenvektor zu $\tilde{f}$ zum Eigenwert λ , so folgt entsprechend, dass $\kappa_X^{-1}(b) \in V$ ein Eigenvektor zu f zum Eigenwert λ ist. $\square$

Insbesondere sieht man mit Bemerkung 2 oder auch mittels einfacher direkter Rechnung:

Bemerkung 7. *Ähnliche Matrizen besitzen dieselben Eigenwerte.*

Als Beispiel zeigen wir, dass die Matrizen

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathbb{R}^{2 \times 2}, \quad B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \in \mathbb{C}^{2 \times 2}$$

nicht diagonalisierbar sind. Da das Gleichungssystem

$$\alpha_2 = \lambda \alpha_1, \quad -\alpha_1 = \lambda \alpha_2$$

für einen nicht-trivialen Vektor $(\alpha_1, \alpha_2)^t \in \mathbb{R}^2$ stets auf die Gleichung $\lambda^2 = -1$ führt, die in $\mathbb{R}$ nicht lösbar ist, sehen wir, dass A in $\mathbb{R}^{2 \times 2}$ nicht diagonalisierbar sein kann, da die zugehörige lineare Abbildung $\mathbb{R}^2 \rightarrow \mathbb{R}^2$, $x \mapsto Ax$, keinen Eigenwert besitzt. Das Bild ändert sich jedoch, wenn wir A als Matrix in $\mathbb{C}^{2 \times 2}$ auffassen, denn die durch A gegebene $\mathbb{C}$ -lineare Abbildung $\mathbb{C}^2 \rightarrow \mathbb{C}^2$, $x \mapsto Ax$, wird bezüglich der Basis $\begin{pmatrix} 1 \\ i \end{pmatrix}, \begin{pmatrix} 1 \\ -i \end{pmatrix}$ durch eine Diagonalmatrix beschrieben. Weiter zeigt das Gleichungssystem

$$\alpha_1 = \lambda \alpha_1, \quad \alpha_1 + \alpha_2 = \lambda \alpha_2,$$

dass die Matrix B höchstens $\lambda = 1$ als Eigenwert besitzt. Wäre B also diagonalisierbar, so müsste B zur Einheitsmatrix ähnlich sein und dann schon mit dieser übereinstimmen, da die Einheitsmatrix aus trivialen Gründen nur zu sich selbst ähnlich ist.

Satz 8. *Es sei $f: V \rightarrow V$ ein Endomorphismus eines endlich-dimensionalen K -Vektorraums V . Sind $a_1, \dots, a_r \in V$ Eigenvektoren zu paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_r$, so sind $a_1, \dots, a_r$ linear unabhängig.*

Beweis. Wir schließen mit Induktion nach r , wobei wir $r \geq 1$ annehmen dürfen. Der Fall $r = 1$ ist klar, denn ein Eigenvektor ist nach Definition stets von 0 verschieden. Sei also $r > 1$, und gelte

$$\sum_{i=1}^r \alpha_i a_i = 0$$

mit Koeffizienten $\alpha_1, \dots, \alpha_r \in K$. Man hat dann

$$\sum_{i=1}^r \lambda_i \alpha_i a_i = \sum_{i=1}^r \alpha_i f(a_i) = f\left(\sum_{i=1}^r \alpha_i a_i\right) = f(0) = 0,$$

aber auch

$$\sum_{i=1}^r \lambda_1 \alpha_i a_i = 0$$

und folglich

$$\sum_{i=2}^r (\lambda_i - \lambda_1) \alpha_i a_i = 0.$$

Nun sind $a_2, \dots, a_r$ insgesamt $r - 1$ Eigenvektoren zu paarweise verschiedenen Eigenwerten und somit nach Induktionsvoraussetzung linear unabhängig. Es ergibt sich daher $(\lambda_i - \lambda_1) \alpha_i = 0$ und damit $\alpha_i = 0$ für $i = 2, \dots, r$. Dann zeigt die Gleichung $\sum_{i=1}^r \alpha_i a_i = 0$, dass auch der Term $\alpha_1 a_1$ verschwindet und wegen $a_1 \neq 0$ sogar der Koeffizient α_1 . Die Vektoren $a_1, \dots, a_r$ sind also wie behauptet linear unabhängig. $\square$

Korollar 9. *Ein Endomorphismus $f: V \longrightarrow V$ eines endlich-dimensionalen K -Vektorraums V hat höchstens $n = \dim_K V$ verschiedene Eigenwerte.*

Korollar 10. *Besitzt ein Endomorphismus $f: V \longrightarrow V$ eines endlich-dimensionalen K -Vektorraums V genau $n = \dim_K V$ verschiedene Eigenwerte, so ist f diagonalisierbar.*

Beweis. Seien $\lambda_1, \dots, \lambda_n \in K$ paarweise verschiedene Eigenwerte zu f , und seien $a_1, \dots, a_n \in V$ zugehörige Eigenvektoren. Dann sind diese gemäß Satz 8 linear unabhängig, bilden also wegen $n = \dim_K V$ eine Basis X von V . Die zugehörige Matrix $A_{f,X,X}$ ist eine Diagonalmatrix mit $\lambda_1, \dots, \lambda_n$ als Diagonalelementen. $\square$

Wir wollen Satz 8 noch etwas verallgemeinern.

Definition 11. *Es sei $\lambda \in K$ Eigenwert eines Vektorraumendomorphismus $f: V \longrightarrow V$. Dann heißt*

$$V_\lambda := \ker(f - \lambda \text{id}) = \{a \in V; f(a) = \lambda a\}$$

der Eigenraum von f zum Eigenwert λ .

Korollar 12. *Sei $f: V \longrightarrow V$ ein Endomorphismus eines endlich-dimensionalen K -Vektorraums V , und seien $\lambda_1, \dots, \lambda_r$ die sämtlichen Eigenwerte von f . Ist dann $V' = \sum_{i=1}^r V_{\lambda_i}$ der von den zugehörigen Eigenräumen erzeugte Untervektorraum von V , so gilt*

$$V' = \bigoplus_{i=1}^r V_{\lambda_i};$$

die Summe ist also direkt. Im Übrigen ist f genau dann diagonalisierbar, wenn $V = V'$ gilt, wenn also V von den Eigenräumen zu f erzeugt wird.

Beweis. Da Eigenvektoren zu paarweise verschiedenen Eigenwerten linear unabhängig sind, kann eine Summe $\sum_{i=1}^r v_i$ mit $v_i \in V_{\lambda_i}$ nur dann verschwinden, wenn alle v_i verschwinden. Dies bedeutet aber, dass V' die direkte Summe der Eigenräume V_{λ_i} ist. Ist nun f diagonalisierbar, so besitzt V eine Basis aus Eigenvektoren, und es gilt $V = V'$. Umgekehrt, ist V darstellbar als direkte Summe der Eigenräume V_{λ_i} , so wähle man in jedem dieser Eigenräume eine Basis. Das System aller dieser Elemente bildet dann eine Basis von V , die aus lauter Eigenvektoren von f besteht, d. h. f ist diagonalisierbar. $\square$

Korollar 13. Für eine Diagonalmatrix

$$A = \begin{pmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \ddots \\ & & & \lambda_n \end{pmatrix}$$

sind $\lambda_1, \dots, \lambda_n$ die einzigen Eigenwerte von A .

Beweis. Wir betrachten die lineare Abbildung $f: K^n \rightarrow K^n$, $x \mapsto Ax$. Es ist klar, dass es sich bei den $\lambda_1, \dots, \lambda_n$ um Eigenwerte von A bzw. f handelt. Um Wiederholungen zu vermeiden, schreibe man $\{\lambda_1, \dots, \lambda_n\} = \{\lambda'_1, \dots, \lambda'_s\}$, wobei die Elemente $\lambda'_1, \dots, \lambda'_s$ paarweise verschieden sind. Die Diagonalgestalt von A besagt, dass es in K^n eine Basis gibt, nämlich die kanonische Basis $e_1, \dots, e_n$, so dass e_i jeweils Eigenvektor von f zum Eigenwert λ_i ist. Für $j = 1, \dots, s$ sei nun $U_{\lambda'_j} \subset V$ derjenige lineare Unterraum, der erzeugt wird von allen e_i mit Indizes i , für die $\lambda_i = \lambda'_j$ gilt. Es besteht dann die Zerlegung

$$V = \bigoplus_{j=1}^s U_{\lambda'_j}, \quad U_{\lambda'_j} \subset V_{\lambda'_j},$$

wobei $V_{\lambda'_j}$ jeweils der Eigenraum von f zum Eigenwert λ'_j ist. Ein Vergleich mit der Zerlegung aus Korollar 12 ergibt $U_{\lambda'_j} = V_{\lambda'_j}$ für $j = 1, \dots, s$ und zeigt außerdem, dass es neben $\lambda'_1, \dots, \lambda'_s$ keine weiteren Eigenwerte von A geben kann. $\square$

Aufgaben

V sei stets ein Vektorraum endlicher Dimension über einem Körper K .

1. Es seien $a, b \in V$ Eigenvektoren eines Endomorphismus $f: V \rightarrow V$. Man untersuche, in welchen Fällen auch $a - b$ ein Eigenvektor von f ist.

2. Es sei $\lambda \in K$ Eigenwert eines Endomorphismus $f: V \rightarrow V$. Man zeige, dass für Polynome $q \in K[T]$ jeweils $q(\lambda)$ Eigenwert von $q(f)$ ist.
3. Für die Matrix

$$A = \begin{pmatrix} 2 & 1 & 0 & 1 \\ 0 & 2 & 0 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix} \in \mathbb{R}^{4 \times 4}$$

berechne man alle Eigenwerte und die zugehörigen Eigenräume. Ist A diagonalisierbar?

4. Die Matrizen $A, B \in K^{n \times n}$ seien ähnlich. Man zeige in direkter Weise:
 - (i) Ein Element $\lambda \in K$ ist genau dann ein Eigenwert von A , wenn es Eigenwert von B ist.
 - (ii) Für Eigenwerte $\lambda \in K$ von A bzw. B gilt $\dim_K V_{A,\lambda} = \dim_K V_{B,\lambda}$, wobei $V_{A,\lambda}$ den Eigenraum zu λ bezüglich der linearen Abbildung $K^n \rightarrow K^n$, $x \mapsto Ax$ bezeichne; entsprechend für $V_{B,\lambda}$.
5. Es seien $A, B \in K^{n \times n}$ ähnlich. Dann sind für Polynome $q \in K[T]$ auch die Matrizen $q(A)$ und $q(B)$ ähnlich.
6. Zwei Endomorphismen $f, g: V \rightarrow V$ heißen ähnlich, wenn es einen Automorphismus $h: V \rightarrow V$ mit $g = h^{-1} \circ f \circ h$ gibt. Man zeige: f und g sind genau dann ähnlich, wenn für eine gegebene Basis X von V die beschreibenden Matrizen $A_{f,X,X}$ und $A_{g,X,X}$ ähnlich sind.
7. Für ein kommutatives Diagramm linearer Abbildungen zwischen K -Vektorräumen

$$\begin{array}{ccc} V & \xrightarrow{f} & V \\ h \downarrow & & \downarrow h \\ W & \xrightarrow{g} & W \end{array}$$

zeige man:

- (i) Ist h injektiv, so ist jeder Eigenwert von f auch Eigenwert von g .
- (ii) Ist h surjektiv, so ist jeder Eigenwert von g auch Eigenwert von f .

Man konstruiere einfache Beispiele, die zeigen, dass in den vorstehenden Aussagen die Voraussetzungen “injektiv” bzw. “surjektiv” nicht entbehrlich sind.

6.2 Minimalpolynom und charakteristisches Polynom

Wie wir im vorigen Abschnitt gesehen haben, steht das Problem der Diagonalisierbarkeit von Endomorphismen oder Matrizen in engem Zusammenhang mit dem Problem, die zugehörigen Eigenwerte und Eigenvektoren zu bestimmen. Wir beschäftigen uns daher zunächst mit der Berechnung von Eigenwerten. Generell sei V in diesem Abschnitt ein K -Vektorraum *endlicher* Dimension n .

Satz 1. Sei $f: V \longrightarrow V$ ein Endomorphismus. Für $\lambda \in K$ ist dann äquivalent:

- (i) λ ist Eigenwert zu f .
- (ii) $\ker(\lambda \operatorname{id} - f) \neq 0$.
- (iii) $\lambda \operatorname{id} - f$ ist nicht invertierbar.
- (iv) $\det(\lambda \operatorname{id} - f) = 0$.

Beweis. Sei λ ein Eigenwert zu f . Dann existiert ein Eigenvektor zu λ , d. h. ein $a \in V - \{0\}$ mit $f(a) = \lambda a$. Hieraus ergibt sich $a \in \ker(\lambda \operatorname{id} - f)$ und damit insbesondere $\ker(\lambda \operatorname{id} - f) \neq 0$. Umgekehrt ist jeder von Null verschiedene Vektor $a \in \ker(\lambda \operatorname{id} - f)$ ein Eigenvektor zum Eigenwert λ . Bedingungen (i) und (ii) sind also äquivalent.

Weiter ergibt sich die Äquivalenz (ii) $\iff$ (iii) aus 2.1/11 und die Äquivalenz (iii) $\iff$ (iv) aus 4.3/3 (iv). $\square$

Als Beispiel wollen wir alle Eigenwerte der Matrix

$$A = \begin{pmatrix} 1 & 4 \\ 1 & 1 \end{pmatrix} \in \mathbb{R}^{2 \times 2}$$

bestimmen. Bezeichnet E die Einheitsmatrix in $\mathbb{R}^{2 \times 2}$, so gilt für $\lambda \in \mathbb{R}$

$$\det(\lambda E - A) = \det \begin{pmatrix} \lambda - 1 & -4 \\ -1 & \lambda - 1 \end{pmatrix} = \lambda^2 - 2\lambda - 3.$$

Die Gleichung $\det(\lambda E - A) = 0$ ist daher äquivalent zu $\lambda = 3$ oder $\lambda = -1$. Daher sind 3, -1 die Eigenwerte von A , und man sieht mittels 6.1/10, dass A diagonalisierbar ist. Und zwar ist A ähnlich zu der Matrix

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

Man kann nun leicht die zu den Eigenwerten 3, -1 gehörigen Eigenräume bestimmen, indem man die linearen Gleichungssysteme

$$(3E - A)x = 0 \quad \text{bzw.} \quad (-E - A)x = 0$$

löst.

Wir wollen $\det(\lambda E - A)$ für eine Matrix $A = (\alpha_{ij})_{ij} \in K^{n \times n}$ und eine Konstante $\lambda \in K$ genauer auswerten; E sei nunmehr die Einheitsmatrix in $K^{n \times n}$. Wie in 4.2/4 definiert, gilt

$$\begin{aligned} \det(\lambda E - A) &= \det((\lambda \delta_{ij} - \alpha_{ij})_{ij}) \\ &= \sum_{\pi \in \mathfrak{S}_n} \operatorname{sgn} \pi \cdot \prod_{i=1}^n (\lambda \delta_{\pi(i), i} - \alpha_{\pi(i), i}), \end{aligned}$$

und wir können folgende Definition treffen:

Definition 2. Sei $A = (\alpha_{ij})_{ij} \in K^{n \times n}$. Dann heit

$$\chi_A = \sum_{\pi \in \mathfrak{S}_n} \operatorname{sgn} \pi \cdot \prod_{i=1}^n (T\delta_{\pi(i),i} - \alpha_{\pi(i),i}) \in K[T]$$

das charakteristische Polynom von A .

Insbesondere gilt $\chi_A(\lambda) = \det(\lambda E - A)$ fr alle $\lambda \in K$, und die Nullstellen von χ_A in K sind gerade die Eigenwerte von A .

Satz 3. Das charakteristische Polynom $\chi_A \in K[T]$ zu einer Matrix $A \in K^{n \times n}$ ist normiert vom Grad n . Es gilt

$$\chi_A = \sum_{i=0}^n c_{n-i} T^i, \quad c_i \in K,$$

mit $c_0 = 1$, $-c_1 = \operatorname{Spur} A = \sum_{i=1}^n \alpha_{ii}$ und $(-1)^n c_n = \det A$, wobei die Summe der Diagonalelemente α_{ii} als Spur von A bezeichnet wird.

Beweis. Als Summe n -facher Produkte linearer Polynome in T ist χ_A vom Grad $\leq n$, und es gilt

$$c_n = \chi_A(0) = \det(-A) = (-1)^n \det A,$$

also $(-1)^n c_n = \det A$. Weiter besitzt der zweite Term in der Zerlegung

$$\chi_A = \prod_{i=1}^n (T - \alpha_{ii}) + \sum_{\substack{\pi \in \mathfrak{S}_n \\ \pi \neq \operatorname{id}}} \prod_{i=1}^n (T\delta_{\pi(i),i} - \alpha_{\pi(i),i})$$

einen Grad $\leq n-2$, denn es wird nur ber Permutationen $\pi \neq \operatorname{id}$ summiert. Fr $\pi \neq \operatorname{id}$ gibt es nmlich mindestens zwei verschiedene Indizes $i, j \in \{1, \dots, n\}$ mit $\pi(i) \neq i$, $\pi(j) \neq j$, so dass folglich die Ausdrcke $T\delta_{\pi(i),i}$, $T\delta_{\pi(j),j}$ fr diese Indizes verschwinden. Die Koeffizienten vom Grad n und $n-1$ in χ_A , also c_0 und c_1 , stimmen daher berein mit den Koeffizienten vom Grad n und $n-1$ in $\prod_{i=1}^n (T - \alpha_{ii})$, und es folgt $c_0 = 1$, sowie $c_1 = -\sum_{i=1}^n \alpha_{ii}$, wie behauptet. $\square$

Man kann das charakteristische Polynom $\chi_A \in K[T]$ zu einer Matrix $A \in K^{n \times n}$ auch durch die Gleichung $\chi_A = \det(TE - A)$ erklren, wobei man dann allerdings die Determinante einer Matrix mit Eintrgen aus dem Polynomring $K[T]$ zu bilden hat. Da wir bisher nur Determinanten von Matrizen mit Koeffizienten aus einem Krper betrachtet und auch nur fr diese Situation Rechenregeln fr Determinanten bewiesen haben, greifen wir zu einem Trick. hnlich wie man den Krper $\mathbb{Q}$ der rationalen Zahlen als Krper aller Brche ganzer Zahlen bildet, konstruiert man zu $K[T]$ den so genannten *rationalen Funktionenkrper* $K(T)$ aller Brche von Polynomen aus $K[T]$. Es ist dann $K[T]$ ein Unterring des Krpers $K(T)$, und man kann $TE - A$ als Matrix in $K(T)^{n \times n}$ auffassen. Insbesondere ist $\det(TE - A)$ wohldefiniert, und man darf

zur Berechnung dieser Determinante die bekannten Entwicklungssätze oder andere Rechenregeln für Determinanten anwenden. Erwähnt sei aber auch, dass sich alternativ die Determinantentheorie über beliebigen kommutativen Ringen entwickeln lässt, worauf wir hier aber nicht weiter eingehen wollen.

Satz 4. Sind $A, B \in K^{n \times n}$ ähnlich, so folgt $\chi_A = \chi_B$.

Beweis. Sei $S \in \text{Gl}(n, K)$ mit $B = S^{-1}AS$. Dann gilt aufgrund der Multiplikativität der Determinante

$$\begin{aligned}\chi_B &= \det(TE - S^{-1}AS) = \det(S^{-1}(TE - A)S) \\ &= \det S^{-1} \cdot \det(TE - A) \cdot \det S = \det(TE - A) \\ &= \chi_A.\end{aligned}$$

□

Korollar 5. Ähnliche Matrizen besitzen die gleiche Spur.

Da die charakteristischen Polynome ähnlicher Matrizen übereinstimmen, kann man nunmehr, unter Benutzung von 6.1/2, auch das charakteristische Polynom eines Endomorphismus $f: V \rightarrow V$ erklären.

Definition 6. Sei $f: V \rightarrow V$ ein Endomorphismus und X eine Basis eines endlich-dimensionalen K -Vektorraums V . Dann bezeichnet man

$$\chi_f = \chi_{A_{f,X,X}} \in K[T]$$

als das charakteristische Polynom von f und

$$\text{Spur } f = \text{Spur } A_{f,X,X}$$

als die Spur von f . Für den trivialen Fall $V = 0$ gilt $\chi_f = 1$ und $\text{Spur } f = 0$.

Es folgt mit 6.1/2 und Satz 4, dass χ_f und $\text{Spur } f$ unabhängig von der speziellen Wahl von X sind. Weiter können wir mit Satz 1 feststellen:

Satz 7. Ein Element $\lambda \in K$ ist genau dann ein Eigenwert eines Endomorphismus $f: V \rightarrow V$, wenn λ eine Nullstelle des charakteristischen Polynoms χ_f ist.

Als Nächstes wollen wir einsehen, dass die Diagonalisierbarkeit eines Endomorphismus an gewissen Eigenschaften des zugehörigen charakteristischen Polynoms abzulesen ist.

Satz 8. Für einen Endomorphismus $f: V \rightarrow V$ ist äquivalent:

- (i) f ist diagonalisierbar.

(ii) χ_f zerfällt vollständig in Linearfaktoren, etwa

$$\chi_f = \prod_{i=1}^r (T - \lambda_i)^{n_i},$$

und für den Eigenraum V_{λ_i} zum Eigenwert λ_i gilt

$$\dim V_{\lambda_i} = n_i.$$

Beweis. Sei $\dim_K V = n$. Wir beginnen mit der Implikation (i) $\implies$ (ii) und nehmen f als diagonalisierbar an. Dann existiert eine Basis X von V , bestehend aus Eigenvektoren zu f , also mit

$$A_{f,X,X} = \begin{pmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ & & \ddots \\ 0 & & & \lambda_n \end{pmatrix}.$$

Insbesondere folgt $\chi_f = \prod_{i=1}^n (T - \lambda_i)$. Schreiben wir dies durch Zusammenfassen gleicher Faktoren zu Potenzen in der Form $\chi_f = \prod_{i=1}^r (T - \lambda_i)^{n_i}$, so folgt $\dim_K(V_{\lambda_i}) \geq n_i$. Wegen $V = \bigoplus_{i=1}^r V_{\lambda_i}$, vgl. 6.1/12 und 6.1/13, gilt dann

$$n = \sum_{i=1}^r \dim_K(V_{\lambda_i}) \geq \sum_{i=1}^r n_i = n$$

und damit $\dim_K(V_{\lambda_i}) = n_i$ für alle $i = 1, \dots, r$.

Ist umgekehrt Bedingung (ii) gegeben, so folgt

$$\sum_{i=1}^r \dim_K(V_{\lambda_i}) = \sum_{i=1}^r n_i = n.$$

Nach 6.1/12 gilt $V' = \bigoplus_{i=1}^r V_{\lambda_i}$ für den von den Eigenräumen V_{λ_i} erzeugten Unterraum $V' \subset V$, also $\dim_K V' = n = \dim_K V$ und damit $V' = V$. Dann ist f aber diagonalisierbar, wiederum nach 6.1/12. $\square$

Als Beispiel für die Anwendung von Satz 8 wollen wir einen Endomorphismus $f: V \rightarrow V$ betrachten, der bezüglich einer geeigneten Basis durch eine Dreiecksmatrix der Form

$$A = \begin{pmatrix} \lambda & & * \\ & \lambda & \\ & & \ddots \\ 0 & & & \lambda \end{pmatrix} \in K^{n \times n}$$

beschrieben wird. Dann gilt $\chi_f = \chi_A = (T - \lambda)^n$, und λ ist der einzige Eigenwert zu f bzw. A . Ist nun A keine Diagonalmatrix, so ist $\lambda \operatorname{id} - f$ nicht die Nullabbildung und folglich der Eigenraum $V_\lambda = \ker(\lambda \operatorname{id} - f)$ echt in V enthalten. Nach

Satz 8 kann f bzw. A daher nicht diagonalisierbar sein. Wir können dies aber auch in direkter Weise sehen. Wenn A diagonalisierbar ist, so ist A ähnlich zu λE , wobei $E \in K^{n \times n}$ die Einheitsmatrix bezeichne. Da aber E und damit auch λE mit allen Matrizen in $K^{n \times n}$ vertauschbar ist, kann λE nur zu sich selbst ähnlich sein. Somit müsste schon $A = \lambda E$ gelten.

Neben dem charakteristischen Polynom χ_f zu einem Endomorphismus f eines K -Vektorraums V kann man auch noch ein weiteres Polynom zu f betrachten, nämlich das so genannte *Minimalpolynom*. Um dieses zu definieren, betrachten wir den Endomorphismenring $\text{End}_K(V)$ als K -Algebra unter dem Ringhomomorphismus

$$K \longrightarrow \text{End}_K(V), \quad c \longmapsto c \cdot \text{id}_V,$$

und verwenden folgendes Resultat:

Satz 9. *Zu einem Endomorphismus $f: V \longrightarrow V$ eines endlich-dimensionalen K -Vektorraums V bilde man den K -Algebrahomomorphismus*

$$\varphi_f: K[T] \longrightarrow \text{End}_K(V), \quad p \longmapsto p(f),$$

der f anstelle von T einsetzt; vgl. 5.1/7. Dann ist $\ker \varphi_f$ von Null verschieden, und es existiert ein normiertes Polynom $p_f \in K[T]$ mit $\ker \varphi_f = (p_f)$. Es ist p_f das eindeutig bestimmte normierte Polynom kleinsten Grades in $K[T]$ mit $p_f(f) = 0$.

Beweis. Indem wir die Ringmultiplikation vergessen, können wir φ_f auch als Homomorphismus zwischen K -Vektorräumen auffassen. Man hat dann $\dim_K(K[T]) = \infty$ sowie gemäß 3.3/2

$$\dim_K(\text{End}_K(V)) = n^2 < \infty$$

für $n = \dim_K(V)$. Letzteres hat $\ker \varphi_f \neq 0$ zur Folge, bzw. dass φ_f nicht injektiv sein kann. Genauer, die Elemente $\varphi_f(T^0), \dots, \varphi_f(T^{n^2})$ sind aus Dimensionsgründen linear abhängig in $\text{End}_K(V)$, und $\ker \varphi_f$ enthält daher ein nicht-triviales Polynom vom Grad $\leq n^2$. Nun ist aber $\ker \varphi_f$ ein Ideal in $K[T]$ und $K[T]$ ein Hauptidealring; vgl. 5.2/7. Es existiert daher ein nicht-triviales Polynom $p_f \in K[T]$ mit $\ker \varphi_f = p_f \cdot K[T]$. Als erzeugendes Element eines Hauptideals in einem Integritätsring ist p_f nach 5.2/9 eindeutig bestimmt bis auf eine Einheit. Da aber $K[T]^* = K^*$ gilt, ist p_f eindeutig, wenn wir dieses Polynom als normiert voraussetzen. Natürlich ist p_f dann das normierte Polynom kleinsten Grades, welches f annulliert. $\square$

Definition 10. *Ist $f: V \longrightarrow V$ ein Endomorphismus, so heißt das nach Satz 9 eindeutig in $K[T]$ existierende normierte Polynom kleinsten Grades, welches f annulliert, das Minimalpolynom von f ; dieses wird mit p_f bezeichnet. Entsprechend ist das Minimalpolynom p_A einer Matrix $A \in K^{n \times n}$ erklärt als das normierte Polynom kleinsten Grades in $K[T]$, welches A annulliert.*

Im Beweis zu Satz 9 wurde gezeigt, dass der Kern des Homomorphismus $\varphi_f: K[T] \rightarrow \text{End}_K(V)$ nicht-triviale Polynome vom Grad $\leq n^2$ enthält. Als Konsequenz ergibt sich $\text{grad } p_f \leq n^2$. Diese Abschätzung lässt sich aber noch erheblich verbessern.

Satz 11 (Cayley-Hamilton). *Das Minimalpolynom p_f eines Endomorphismus $f: V \rightarrow V$ ist stets ein Teiler des charakteristischen Polynoms χ_f . Insbesondere gilt $\chi_f(f) = 0$ und $\text{grad } p_f \leq \text{grad } \chi_f = n$.*

Beweis. In der Situation von Satz 9 ist nur $\chi_f \in \ker \varphi_f = (p_f)$, d. h. $\chi_f(f) = 0$ zu zeigen. Indem wir dies in ein Matrizenproblem übersetzen, genügt es $\chi_A(A) = 0$ für Matrizen $A \in K^{n \times n}$ zu zeigen. Um bequem rechnen zu können, betrachten wir wieder den rationalen Funktionenkörper $K(T)$, dessen Elemente Brüche von Polynomen aus $K[T]$ sind. Sodann können wir den Unterring $K[T]^{n \times n}$ des Matrizenrings $K(T)^{n \times n}$ betrachten, der aus allen $(n \times n)$ -Matrizen mit Einträgen aus $K[T]$ besteht. Der Homomorphismus $K(T) \rightarrow K(T)^{n \times n}$, der ein Element $q \in K(T)$ auf das q -fache der Einheitsmatrix $E \in K(T)^{n \times n}$ abbildet, beschränkt sich zu einem Ringhomomorphismus

$$K[T] \rightarrow K[T]^{n \times n}, \quad f \mapsto f \cdot E,$$

und definiert auf $K[T]^{n \times n}$ die Struktur einer $K[T]$ -Algebra. Indem wir die Variable $T \in K[T]$ mit ihrem Bild $T \cdot E \in K[T]^{n \times n}$ identifizieren, lässt sich jedes Element $M \in K[T]^{n \times n}$ in der Form $M = \sum_{i \in \mathbb{N}} M_i T^i$ schreiben, wobei die Koeffizienten $M_i \in K^{n \times n}$ eindeutig durch M bestimmt sind und natürlich für fast alle $i \in \mathbb{N}$ verschwinden. Für $M = (\sum_{i \in \mathbb{N}} m_{\mu\nu i} T^i)_{\mu, \nu=1, \dots, n}$ mit Koeffizienten $m_{\mu\nu i} \in K$ setze man nämlich $M_i = (m_{\mu\nu i})_{\mu, \nu=1, \dots, n}$. In dieser Weise können wir $K[T]^{n \times n}$ als “Polynomring” $K^{n \times n}[T]$ auffassen, wobei allerdings zu berücksichtigen ist, dass der Ring $K^{n \times n}$ für $n > 1$ nicht kommutativ ist. Man kann aber wie in Abschnitt 5.1 den Polynomring $R[T]$ auch für einen nicht-kommutativen Ring R erklären, muss sich dann aber darüber im Klaren sein, dass Einsetzungsabbildungen des Typs

$$R[T] \rightarrow R, \quad \sum a_i T^i \mapsto \sum a_i t^i,$$

für Elemente $t \in R$ im Allgemeinen *keine* Homomorphismen mehr darstellen. Denn T ist in $R[T]$ mit allen Elementen aus R vertauschbar, möglicherweise aber nicht $t \in R$ mit allen anderen Elementen aus R .

Nach diesen Vorbereitungen betrachte man nun eine Matrix $A \in K^{n \times n}$ und ihr charakteristisches Polynom $\chi_A \in K[T]$, welches wir für die Zwecke dieses Beweises einmal in der Form $\chi_A = \sum_{i \in \mathbb{N}} c_i T^i$ schreiben wollen. Weiter fassen wir $TE - A$ als Matrix in $K[T]^{n \times n} \subset K(T)^{n \times n}$ auf und bilden deren adjungierte Matrix $(TE - A)^{\text{ad}} \in K(T)^{n \times n}$; vgl. 4.4/2. Aufgrund der Cramerschen Regel 4.4/3 besteht dann die Gleichung

$$(TE - A)^{\text{ad}} \cdot (TE - A) = (\det(TE - A)) \cdot E = \chi_A(T) \cdot E.$$

Dies ist zunächst eine Gleichung in $K(T)^{n \times n}$, sie gilt aber auch in $K[T]^{n \times n}$, da aufgrund der Definition der adjungierten Matrix mit $TE - A$ auch $(TE - A)^{\text{ad}}$ zu $K[T]^{n \times n}$ gehört. Gilt etwa $(TE - A)^{\text{ad}} = \sum_{i \in \mathbb{N}} A_i T^i$ mit Matrizen $A_i \in K^{n \times n}$, so folgt

$$\begin{aligned} \left(\sum_{i \in \mathbb{N}} A_i T^i \right) \cdot (TE - A) &= -A_0 T^0 A + \sum_{i \in \mathbb{N}} (A_i T^i TE - A_{i+1} T^{i+1} A) \\ &= -A_0 A + \sum_{i \in \mathbb{N}} (A_i - A_{i+1} A) T^{i+1} \\ &= \sum_{i \in \mathbb{N}} c_i T^i \cdot E = \chi_A(T) \cdot E \end{aligned}$$

mit

$$-A_0 A = c_0 E, \quad A_i - A_{i+1} A = c_{i+1} E \quad \text{für } i \in \mathbb{N}.$$

Dieselbe Rechnung lässt sich auch mit einer Matrix $B \in K^{n \times n}$ anstelle der Variablen T durchführen, sofern A mit B vertauschbar ist. Insbesondere dürfen wir in vorstehender Gleichung T durch $B := A$ ersetzen und erhalten dann wie gewünscht $\chi_A(A) = 0$, da der Term $(TE - A)$ nach Ersetzen von T durch A verschwindet. $\square$

Wir wollen noch zwei einfache Beispiele betrachten. Für die Einheitsmatrix $E \in K^{n \times n}$, $n > 0$, gilt $\chi_E = (T - 1)^n$, $p_A = T - 1$, und für die Nullmatrix $0 \in K^{n \times n}$ hat man $\chi_0 = T^n$, $p_0 = T$. Insbesondere sieht man, dass das Minimalpolynom im Allgemeinen nicht mit dem charakteristischen Polynom übereinstimmt.

Aufgaben

V sei stets ein Vektorraum endlicher Dimension n über einem Körper K .

1. Man bestimme Eigenwerte und zugehörige Eigenräume der folgenden Matrix:

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

Ist A diagonalisierbar?

2. Man bestimme das Minimalpolynom p_f zu einem Endomorphismus $f: V \rightarrow V$ in folgenden Fällen:
 - (i) $V = 0$
 - (ii) $f = \text{id}$
 - (iii) $f = 0$
 - (iv) Es existieren lineare Unterräume $V_1, V_2 \subset V$ mit $V = V_1 \oplus V_2$, und es gilt $f(v_1 + v_2) = v_1$ für $v_i \in V_i$, $i = 1, 2$.

3. Es seien $U_1, U_2 \subset V$ lineare Unterräume und $f: V \rightarrow V$ ein Endomorphismus, der sich zu Endomorphismen $f_i: U_i \rightarrow U_i$, $i = 1, 2$, einschränkt. Man zeige:
 - (i) Gilt $V = U_1 + U_2$, so folgt $p_f = \text{kgV}(p_{f_1}, p_{f_2})$ für die Minimalpolynome von f, f_1, f_2 .
 - (ii) f schränkt sich zu einem Endomorphismus $f_{12}: U_1 \cap U_2 \rightarrow U_1 \cap U_2$ ein, und es gilt $p_{f_{12}} | \text{ggT}(p_{f_1}, p_{f_2})$ für die Minimalpolynome von f_{12}, f_1, f_2 . Gilt im Allgemeinen auch die Gleichheit $p_{f_{12}} = \text{ggT}(p_{f_1}, p_{f_2})$?
4. K sei algebraisch abgeschlossen. Man zeige, ein Endomorphismus $f: V \rightarrow V$ ist genau dann nilpotent (d. h. erfüllt eine Gleichung der Form $f^r = 0$), wenn f außer 0 keine weiteren Eigenwerte besitzt.
5. Es sei $f: V \rightarrow V$ ein Automorphismus. Man zeige, es existiert ein Polynom $q \in K[T]$ mit $f^{-1} = q(f)$.
6. Die Folge der *Fibonacci-Zahlen* $c_1, c_2, \dots \in \mathbb{N}$ ist definiert durch $c_1 = c_2 = 1$ und $c_{n+2} = c_{n+1} + c_n$ für $n \in \mathbb{N}$. Man gebe für c_n einen geschlossenen Ausdruck an, der nur von n abhängt. Hinweis: Man bestimme eine Matrix $A \in \mathbb{R}^{2 \times 2}$ mit $A \cdot (c_{n+1}, c_n)^t = (c_{n+2}, c_{n+1})^t$ für $n \geq 1$ und eine Basiswechselmatrix $S \in \text{Gl}(2, \mathbb{R})$, derart dass $S^{-1} \cdot A \cdot S$ Diagonalgestalt besitzt.

6.3 Der Elementarteilersatz

Es sei V ein Vektorraum über einem Körper K und $f: V \rightarrow V$ ein Endomorphismus. Dann setzt sich die auf V definierte skalare Multiplikation $K \times V \rightarrow V$, $(\alpha, v) \mapsto \alpha \cdot v$, fort zu einer äußeren Multiplikation

$$K[T] \times V \rightarrow V, \quad (p, v) \mapsto p \cdot v := p(f)(v).$$

Dabei ist für $p \in K[T]$ wie üblich $p(f)$ derjenige Ausdruck in $\text{End}_K(V)$, der aus p entsteht, indem man die Variable T durch f ersetzt. Weiter ist $p(f)(v)$ das Bild von v unter dem Endomorphismus $p(f): V \rightarrow V$. Man prüft leicht nach, dass V als additive abelsche Gruppe zusammen mit der äußeren Multiplikation $K[T] \times V \rightarrow V$ den in 1.4/1 aufgeführten Vektorraumaxiomen genügt, wenn man einmal davon absieht, dass $K[T]$ nur ein Ring und kein Körper ist; wir sagen, V sei ein $K[T]$ -Modul. (Man beachte: Im Unterschied zu anderem sprachlichen Gebrauch heißt es in der Mathematik “der Modul” bzw. “die Moduln”, mit Betonung auf der ersten Silbe.)

Will man Normalformen von Endomorphismen $f: V \rightarrow V$ studieren, so bedeutet dies, dass man die oben erklärte Struktur von V als $K[T]$ -Modul analysieren muss. Wir wollen daher zunächst ein paar Grundlagen über Moduln zusammenstellen.

Definition 1. Es sei R ein (kommutativer) Ring. Ein R -Modul ist eine Menge M mit einer inneren Verknüpfung $M \times M \rightarrow M$, $(a, b) \mapsto a + b$, genannt *Addition*, und einer äußeren Verknüpfung $R \times M \rightarrow M$, $(\alpha, a) \mapsto \alpha \cdot a$, genannt *skalare Multiplikation*, so dass gilt:

- (i) M ist eine abelsche Gruppe bezüglich der Addition “+”.
- (ii) $(\alpha + \beta) \cdot a = \alpha \cdot a + \beta \cdot a$, $\alpha \cdot (a + b) = \alpha \cdot a + \alpha \cdot b$ für alle $\alpha, \beta \in R$, $a, b \in M$, d. h. Addition und skalare Multiplikation verhalten sich distributiv.
- (iii) $(\alpha \cdot \beta) \cdot a = \alpha \cdot (\beta \cdot a)$ für alle $\alpha, \beta \in R$, $a \in M$, d. h. die skalare Multiplikation ist assoziativ.
- (iv) $1 \cdot a = a$ für das Einselement $1 \in R$ und alle $a \in M$.

Wir wollen einige Beispiele betrachten:

- (1) Ein Vektorraum über einem Körper K ist ein K -Modul.
- (2) Ein Ideal $\mathfrak{a} \subset R$ eines Ringes R ist ein R -Modul, insbesondere ist R selbst ein R -Modul. Für $n \in \mathbb{N} - \{0\}$ ist R^n in nahe liegender Weise ein R -Modul.
- (3) Jede abelsche Gruppe G ist ein $\mathbb{Z}$ -Modul; wie gewöhnlich erkläre man $n \cdot g$ für $n \in \mathbb{N}$ und $g \in G$ als n -fache Summe von g , sowie $(-n) \cdot g$ als $-(n \cdot g)$.
- (4) Jeder Endomorphismus $f: V \rightarrow V$ eines K -Vektorraums V induziert, wie oben erklärt, auf V die Struktur eines $K[T]$ -Moduls. Ist V endlich-dimensional und bezeichnet p_f das Minimalpolynom von f , so gilt $p_f \cdot v = 0$ für alle $v \in V$. Im Unterschied zu Vektorräumen kann man daher bei einem R -Modul M aus einer Gleichung $\alpha \cdot m = 0$ mit $\alpha \in R$, $m \in M$ nicht schließen, dass α oder m verschwinden.

Eine ganze Reihe von Begriffen, die bei Vektorräumen eine Rolle spielen, haben auch für Moduln ihre Bedeutung. Sei etwa M ein Modul über einem Ring R . Ein *Untermodul* von M ist eine nicht-leere Teilmenge $N \subset M$, so dass gilt:

$$\begin{aligned} a, b \in N &\implies a + b \in N, \\ \alpha \in R, a \in N &\implies \alpha \cdot a \in N \end{aligned}$$

Es ist N dann wieder ein R -Modul unter den von M ererbten Verknüpfungen. Betrachtet man einen Ring R als Modul über sich selbst, so stimmen die Ideale des Rings R mit den Untermoduln von R überein.

Ein *Homomorphismus* zwischen R -Moduln M und N ist eine Abbildung $\varphi: M \rightarrow N$, für die

$$\varphi(a + b) = \varphi(a) + \varphi(b), \quad \varphi(\alpha \cdot a) = \alpha \cdot \varphi(a),$$

für $a, b \in M$, $\alpha \in R$ gilt. Man spricht dabei auch von einer *R -linearen Abbildung*. Mono-, Epi- bzw. Isomorphismen von R -Moduln sind wie üblich als injektive, surjektive bzw. bijektive Homomorphismen erklärt.

Zu einem System $(a_i)_{i \in I}$ von Elementen aus M kann man den hiervon erzeugten Untermodul $M' \subset M$ betrachten, wobei letzterer durch

$$M' := \sum_{i \in I} Ra_i := \left\{ \sum_{i \in I} \alpha_i a_i ; \alpha_i \in R, \alpha_i = 0 \text{ für fast alle } i \in I \right\}$$

definiert ist. Gilt $M = \sum_{i \in I} Ra_i$, so nennt man $(a_i)_{i \in I}$ ein *Erzeugendensystem* von M . Man bezeichnet M als *endlich erzeugt* oder (in missbräuchlicher Sprechweise) als *endlich*, wenn M ein endliches Erzeugendensystem besitzt. Weiter

heißt ein System $(a_i)_{i \in I}$ von Elementen aus M *frei* (oder *linear unabhängig*), wenn aus einer Gleichung

$$\sum_{i \in I} \alpha_i a_i = 0$$

mit Koeffizienten $\alpha_i \in R$, die für fast alle $i \in I$ verschwinden, bereits $\alpha_i = 0$ für alle i folgt. Freie Erzeugendensysteme werden auch als *Basen* bezeichnet. Man beachte jedoch, dass Moduln im Unterschied zu Vektorräumen im Allgemeinen keine Basen besitzen; vgl. Beispiel (4) oben. Moduln, die eine Basis besitzen, heißen *frei*, bzw. *endlich frei*, wenn sie eine endliche Basis besitzen. Homomorphismen zwischen freien R -Moduln lassen sich wie gewöhnlich bezüglich gewählter Basen durch Matrizen mit Koeffizienten aus R beschreiben.

Sind $M_1, \dots, M_n$ Untermoduln eines R -Moduls M , so kann man deren *Summe*

$$M' = \left\{ \sum_{i=1}^n a_i ; a_i \in M_i, i = 1, \dots, n \right\}$$

betrachten. Dabei ist M' wiederum ein Untermodul von M , und man schreibt $M' = \sum_{i=1}^n M_i$. Weiter sagt man, M' sei die *direkte Summe* der M_i , in Zeichen $M' = \bigoplus_{i=1}^n M_i$, wenn zusätzlich für jedes $a \in M'$ die jeweilige Darstellung $a = \sum_{i=1}^n a_i$ mit Elementen $a_i \in M_i$ eindeutig ist.

Wie im Falle von Vektorräumen lässt sich die direkte Summe $\bigoplus_{i=1}^n N_i$ von R -Moduln $N_1, \dots, N_n$, die nicht notwendig als Untermoduln eines R -Moduls N gegeben sind, auch *konstruieren*. Man setze nämlich

$$\bigoplus_{i=1}^n N_i = N_1 \times \dots \times N_n$$

und betrachte dieses kartesische Produkt wiederum als R -Modul, und zwar unter der komponentenweisen Addition bzw. skalaren Multiplikation. Es lässt sich dann N_i für $i = 1, \dots, n$ jeweils mit dem Untermodul

$$0 \times \dots \times 0 \times N_i \times 0 \times \dots \times 0 \subset N_1 \times \dots \times N_n$$

identifizieren, so dass der Modul $N_1 \times \dots \times N_n$ in der Tat als direkte Summe der Untermoduln $N_1, \dots, N_n$ aufzufassen ist.

Ist N ein Untermodul eines R -Moduls M , so kann man den *Restklassenmodul* M/N bilden. Wie im Falle von Vektorräumen gibt N nämlich Anlass zu einer Äquivalenzrelation auf M :

$$a \sim b \iff a - b \in N$$

Es besteht M/N aus den zugehörigen Äquivalenzklassen, d. h. aus den Nebenklassen $[a] = a + N$ zu Elementen $a \in M$. Die R -Modulstruktur auf M/N wird durch die Formeln

$$\begin{aligned} [a] + [b] &= [a + b], & a, b \in M, \\ \alpha \cdot [a] &= [\alpha \cdot a], & \alpha \in R, a \in M, \end{aligned}$$

gegeben, wobei natürlich die Wohldefiniertheit zu überprüfen ist. Die Homomorphiesätze 2.2/8 und 2.2/9 lassen sich ohne Änderungen übertragen. Insbesondere induziert ein R -Modulhomomorphismus $\varphi: M \rightarrow N$ stets einen injektiven R -Modulhomomorphismus $\overline{\varphi}: M/\ker \varphi \hookrightarrow N$.

Da Moduln im Allgemeinen keine Basen besitzen, lässt sich der Begriff der Dimension nicht ohne weiteres von Vektorräumen auf Moduln übertragen. Gewisse Aspekte des Dimensionsbegriffes werden durch die so genannte *Länge* eines Moduls abgedeckt. Hierunter versteht man für einen R -Modul M das Supremum $\ell_R(M)$ aller Längen ℓ von echt aufsteigenden Ketten von Untermoduln des Typs

$$0 \subsetneq M_1 \subsetneq M_2 \subsetneq \dots \subsetneq M_\ell = M.$$

Beispielsweise ist $\ell_R(M) = 0$ äquivalent zu $M = 0$. Weiter hat $\mathbb{Z}$ als freier Modul über sich selbst die Länge ∞ . Als Hilfsmittel für später benötigen wir zwei Lemmata.

Lemma 2. *Es sei R ein Hauptidealring und $a \in R$ ein Element mit Primfaktorzerlegung $a = p_1 \dots p_r$. Dann besitzt der Restklassenmodul R/aR die Länge $\ell_R(R/aR) = r$.*

Beweis. Es sei $\pi: R \rightarrow R/aR$ die kanonische Projektion. Da die Ideale $\overline{\mathfrak{a}} \subset R/aR$ unter der Zuordnung $\overline{\mathfrak{a}} \mapsto \pi^{-1}(\overline{\mathfrak{a}})$ bijektiv denjenigen Idealen in R entsprechen, die aR enthalten, stimmt die Länge von R/aR überein mit dem Supremum der Längen ℓ echt aufsteigender Idealketten des Typs

$$Ra \subsetneq \mathfrak{a}_1 \subsetneq \mathfrak{a}_2 \subsetneq \dots \subsetneq \mathfrak{a}_\ell = R.$$

Da R ein Hauptidealring ist, wird jedes $\mathfrak{a}_i$ von einem Element a_i erzeugt. Weiter ist eine echte Inklusion $\mathfrak{a}_{i-1} \subsetneq \mathfrak{a}_i$ gleichbedeutend damit, dass a_i ein echter Teiler von a_{i-1} ist. Die Länge von R/aR ist daher gleich dem Supremum aller $\ell \in \mathbb{N}$, so dass es $a_1, \dots, a_\ell \in R$ gibt mit der Eigenschaft, dass a_i jeweils ein echter Teiler von a_{i-1} ist, $i = 1, \dots, \ell$; dabei ist $a_0 = a$ zu setzen. Da in R der Satz von der eindeutigen Primfaktorzerlegung gilt und a_0 ein Produkt von r Primfaktoren ist, berechnet sich dieses Supremum zu r . $\square$

Lemma 3. *Ist ein R -Modul M die direkte Summe zweier Untermoduln M' und M'' , so gilt $\ell_R(M) = \ell_R(M') + \ell_R(M'')$.*

Beweis. Hat man echt aufsteigende Ketten von Untermoduln

$$\begin{aligned} 0 \subsetneq M'_1 \subsetneq M'_2 \subsetneq \dots \subsetneq M'_r = M', \\ 0 \subsetneq M''_1 \subsetneq M''_2 \subsetneq \dots \subsetneq M''_s = M'', \end{aligned}$$

so ist

$$\begin{aligned} 0 \subsetneq M'_1 \oplus 0 \subsetneq M'_2 \oplus 0 \subsetneq \dots \subsetneq M'_r \oplus 0 \\ \subsetneq M'_r \oplus M''_1 \subsetneq M'_r \oplus M''_2 \subsetneq \dots \subsetneq M'_r \oplus M''_s = M \end{aligned}$$

eine echt aufsteigende Kette der Länge $r + s$ in M . Also gilt

$$\ell_R(M) \geq \ell_R(M') + \ell_R(M'').$$

Zum Nachweis der umgekehrten Abschätzung betrachte man eine echt aufsteigende Kette von Untermoduln

$$0 = M_0 \subsetneq M_1 \subsetneq M_2 \subsetneq \dots \subsetneq M_\ell = M.$$

Es sei $\pi'': M' \oplus M'' \longrightarrow M''$ die Projektion auf den zweiten Summanden, so dass also $\ker \pi'' = M'$ gilt. Dann ist für $0 \leq \lambda < \ell$, wie wir sogleich sehen werden, jeweils $M_\lambda \cap M'$ echt enthalten in $M_{\lambda+1} \cap M'$ oder $\pi''(M_\lambda)$ echt enthalten in $\pi''(M_{\lambda+1})$. Hieraus folgt $\ell \leq \ell_R(M') + \ell_R(M'')$ und damit insgesamt wie gewünscht $\ell_R(M) = \ell_R(M') + \ell_R(M'')$.

Um das gerade behauptete Inklusionsverhalten zu rechtfertigen, nehmen wir einmal $M_\lambda \cap M' = M_{\lambda+1} \cap M'$ sowie $\pi''(M_\lambda) = \pi''(M_{\lambda+1})$ an und zeigen, dass dies bereits $M_\lambda = M_{\lambda+1}$ impliziert, im Widerspruch zu unserer Voraussetzung. In der Tat, zu $a \in M_{\lambda+1}$ gibt es wegen $\pi''(M_\lambda) = \pi''(M_{\lambda+1})$ ein $a' \in M_\lambda$ mit $\pi(a) = \pi(a')$, also mit $a - a' \in \ker \pi = M'$. Dann gilt wegen $a, a' \in M_{\lambda+1}$ sogar $a - a' \in M_{\lambda+1} \cap M' = M_\lambda \cap M'$ und damit $a = a' + (a - a') \in M_\lambda$. Es folgt daher $M_{\lambda+1} \subset M_\lambda$ bzw. $M_{\lambda+1} = M_\lambda$, was aber ausgeschlossen war. $\square$

Theorem 4 (Elementarteilersatz). *Es sei R ein Hauptidealring und F ein endlicher freier R -Modul. Weiter sei $M \subset F$ ein Untermodul. Dann existieren Elemente $x_1, \dots, x_s \in F$, die Teil einer Basis von F sind, sowie Koeffizienten $\alpha_1, \dots, \alpha_s \in R - \{0\}$, so dass gilt:*

- (i) $\alpha_1 x_1, \dots, \alpha_s x_s$ bilden eine Basis von M .
- (ii) $\alpha_i \mid \alpha_{i+1}$ für $1 \leq i < s$.

Dabei sind $\alpha_1, \dots, \alpha_s$ bis auf Assoziiertheit (d. h. bis auf Einheiten) eindeutig durch M bestimmt, unabhängig von der Wahl von $x_1, \dots, x_s$. Man nennt $\alpha_1, \dots, \alpha_s$ die Elementarteiler von $M \subset F$. Insbesondere ist deren Anzahl s eindeutig bestimmt.

Beweis der Existenzaussage von Theorem 4. Es sei $Y = (y_1, \dots, y_m)$ eine Basis von F . Wir zeigen zunächst per Induktion nach m , dass der Untermodul $M \subset F$ endlich erzeugt ist. Für $m = 1$ ist dies klar, denn F ist dann als R -Modul isomorph zu R , und M korrespondiert zu einem Ideal in R . Letzteres ist endlich erzeugt, da R ein Hauptidealring ist. Sei also $m > 1$. Man setze $F' = \sum_{i=1}^{m-1} R y_i$ und $F'' = R y_m$. Weiter betrachte man die Projektion $\pi: F \longrightarrow F''$, welche y_i für $i < m$ auf 0 und y_m auf y_m abbildet; es gilt dann $\ker \pi = F'$, und man hat — in der Sprache von Abschnitt 2.3 — eine kurze exakte Sequenz

$$0 \longrightarrow F' \longrightarrow F \longrightarrow F'' \longrightarrow 0.$$

Nun sind die Untermoduln $M \cap F' \subset F'$ und $\pi(M) \subset F''$ nach Induktionsvoraussetzung endlich erzeugt, und man zeigt wie üblich, z. B. wie im Beweis zu 2.1/10, dass ein Erzeugendensystem von $M \cap F'$ zusammen mit der Liftung

eines Erzeugendensystems von $\pi(M)$ insgesamt ein Erzeugendensystem von M bildet. M ist also endlich erzeugt.

Wir behalten $Y = (y_1, \dots, y_m)$ als Basis von F bei und wählen ein endliches Erzeugendensystem $z_1, \dots, z_n$ von M . Bezeichnet dann $e = (e_1, \dots, e_n)$ die kanonische Basis des R -Moduls R^n , so kann man die durch $e_j \mapsto z_j$ erklärte R -lineare Abbildung $f: R^n \rightarrow F$ betrachten, deren Bild M ergibt. Gilt dann

$$z_j = \sum_{i=1}^m \alpha_{ij} y_i, \quad j = 1, \dots, n,$$

so ist $A = (\alpha_{ij})_{i,j} \in R^{m \times n}$ die Matrix zu f bezüglich der Basen e und Y . Wir verwenden nun folgendes Hilfsresultat, das wir weiter unten beweisen werden:

Lemma 5. *Es sei R ein Hauptidealring und $A = (\alpha_{ij}) \in R^{m \times n}$ eine Matrix mit Koeffizienten aus R . Dann gibt es invertierbare Matrizen $S \in R^{m \times m}$ und $T \in R^{n \times n}$ mit*

$$S \cdot A \cdot T = \begin{pmatrix} \alpha_1 & 0 & & 0 & 0 & \dots & 0 \\ 0 & \alpha_2 & & 0 & 0 & \dots & 0 \\ & & \ddots & & & & \\ & & & \ddots & & & \\ 0 & 0 & & \alpha_s & 0 & \dots & 0 \\ 0 & 0 & & 0 & 0 & \dots & 0 \\ 0 & 0 & & 0 & 0 & \dots & 0 \end{pmatrix}$$

und mit Koeffizienten $\alpha_1, \dots, \alpha_s \in R - \{0\}$ (wobei $0 \leq s \leq \min(m, n)$ gilt), die für $1 \leq i < s$ die Bedingung $\alpha_i \mid \alpha_{i+1}$ erfüllen. Dabei sind $\alpha_1, \dots, \alpha_s$ bis auf Assoziiertheit eindeutig bestimmt; man nennt sie die Elementarteiler der Matrix A .

Indem man S und T als Basiswechselmatrizen auffasst, sieht man, dass die Matrix SAT ebenfalls die Abbildung f beschreibt, allerdings bezüglich geeigneter anderer Basen $e'_1, \dots, e'_n$ von R^n und $x_1, \dots, x_m$ von F . Insbesondere folgt, dass M als Bild von f durch $\alpha_1 x_1, \dots, \alpha_s x_s$ erzeugt wird. Da das System der $x_1, \dots, x_m$ frei ist und wir Koeffizienten aus einem Integritätsring R betrachten, bilden $\alpha_1 x_1, \dots, \alpha_s x_s$ sogar eine Basis von M . Damit haben wir die Existenz der Elementarteiler $\alpha_1, \dots, \alpha_s$ von $M \subset F$ auf die Existenzaussage von Lemma 5 zurückgeführt. $\square$

Beweis der Existenzaussage von Lemma 5. Wir nehmen zunächst R als euklidischen Ring an und zeigen anhand eines konstruktiven Verfahrens unter Verwendung der Division mit Rest, dass sich die Matrix $A = (\alpha_{ij})$ durch reversible elementare Zeilen- und Spaltenumformungen in die gewünschte Gestalt bringen lässt, nämlich durch Vertauschen von Zeilen (bzw. Spalten) sowie durch Addieren eines Vielfachen einer Zeile (bzw. Spalte) zu einer weiteren Zeile (bzw. Spalte). Wie im Fall einer Matrix mit Koeffizienten aus einem Körper sind elementare Umformungen dieses Typs als Multiplikation mit einer invertierbaren

Elementarmatrix von links (bzw. rechts) zu interpretieren. Die benötigten Zeilenumformungen korrespondieren daher insgesamt zur Multiplikation mit einer invertierbaren Matrix $S \in R^{m \times m}$ von links, die benötigten Spaltenumformungen entsprechend zur Multiplikation mit einer invertierbaren Matrix $T \in R^{n \times n}$ von rechts. Anschließend verallgemeinern wir das Verfahren, so dass es in modifizierter Version auch für Hauptidealringe anwendbar ist.

Wir betrachten im Folgenden also zunächst einen euklidischen Ring R mit Gradabbildung $\delta: R - \{0\} \rightarrow \mathbb{N}$. Für $A = 0$ ist nichts zu zeigen, so dass wir $A \neq 0$ annehmen dürfen. Es ist unsere Strategie, A mittels elementarer Umformungen so abzuändern, dass sich das Minimum

$$d(A) := \min\{\delta(\alpha); \alpha \text{ ist Koeffizient} \neq 0 \text{ von } A\}$$

schrittweise verringert. Da δ Werte in $\mathbb{N}$ annimmt, muss dieses Verfahren nach endlich vielen Schritten abbrechen. Ist dann $\alpha \neq 0$ ein Koeffizient der transformierten Matrix mit minimalem Grad $\delta(\alpha)$, so zeigen wir mittels Division mit Rest, dass α alle anderen Koeffizienten der Matrix teilt; α ist dann der erste Elementarteiler von A .

Im Einzelnen gehen wir wie folgt vor. Indem wir Zeilen und Spalten in A vertauschen, können wir $d(A) = \delta(\alpha_{11})$ annehmen, dass also $\delta(\alpha_{11})$ minimal ist unter allen $\delta(\alpha_{ij})$ mit $\alpha_{ij} \neq 0$. Diese Situation stellen wir zu Beginn eines jeden Schrittes her. Ist dann eines der Elemente der 1. Spalte, etwa α_{i1} , nicht durch α_{11} teilbar, so teile man α_{i1} mit Rest durch α_{11} , etwa $\alpha_{i1} = q\alpha_{11} + \beta$ mit $\delta(\beta) < \delta(\alpha_{11})$, und ziehe das q -fache der 1. Zeile von der i -ten Zeile ab. Als Resultat entsteht an der Position $(i, 1)$ das Element β . Das Minimum $d(A)$ der Grade von nichtverschwindenden Koeffizienten von A hat sich daher verringert, und man starte das Verfahren erneut mit einem weiteren Schritt. In gleicher Weise können wir die Elemente der 1. Zeile mittels elementarer Spaltenumformungen abändern. Da $d(A)$ Werte in $\mathbb{N}$ annimmt, also nicht beliebig oft verringert werden kann, ist nach endlich vielen Schritten jedes Element der 1. Spalte sowie der 1. Zeile ein Vielfaches von α_{11} , und wir können durch Addition von Vielfachen der 1. Zeile zu den restlichen Zeilen der Matrix annehmen, dass $\alpha_{i1} = 0$ für $i > 1$ gilt. Entsprechend können wir mit der 1. Zeile verfahren und auf diese Weise $\alpha_{i1} = \alpha_{1j} = 0$ für $i, j > 1$ erreichen. Dabei dürfen wir weiter annehmen, dass das Minimum $d(A)$ mit $\delta(\alpha_{11})$ übereinstimmt; ansonsten ist das Verfahren erneut zu beginnen und ein entsprechendes Element an der Stelle $(1, 1)$ neu zu positionieren. Existieren nun $i, j > 1$ mit $\alpha_{11} \nmid \alpha_{ij}$, so addiere man die j -te Spalte zur ersten, ein Prozess, der α_{11} unverändert lässt. Wie gerade beschrieben, lassen sich die Elemente unterhalb α_{11} erneut trivialisieren, und zwar unter Verringerung des Grades $d(A)$. Nach endlich vielen Schritten gelangt man so zu einer Matrix (α_{ij}) mit $\alpha_{i1} = \alpha_{1j} = 0$ für $i, j > 1$ sowie mit der Eigenschaft, dass α_{11} jedes andere Element α_{ij} mit $i, j > 1$ teilt. Man behandle dann in gleicher Weise die Untermatrix $(\alpha_{ij})_{i,j>1}$ von $A = (\alpha_{ij})$, sofern diese nicht bereits Null ist. Die hierfür benötigten Umformungen lassen die erste Zeile und Spalte von A invariant und erhalten insbesondere die Bedingung, dass α_{11} alle restlichen Koeffizienten von A teilt. Führt man dieses Verfahren in

induktiver Weise fort, so gelangt man schließlich nach endlich vielen Schritten zu einer Matrix, auf deren Hauptdiagonalen die gesuchten Elementarteiler mit der behaupteten Teilbarkeitseigenschaft stehen und deren sonstige Einträge alle verschwinden. Damit ist die Existenzaussage von Lemma 5 und insbesondere auch von Theorem 4 bewiesen, zumindest im Falle eines euklidischen Rings R .

Ist nun R lediglich als Hauptidealring bekannt, so benötigen wir elementare Matrizenumformungen eines etwas allgemeineren Typs, die wir zunächst erläutern wollen. Seien $\sigma, \tau, \sigma', \tau' \in R$ mit $\sigma\tau' - \tau\sigma' = 1$ gewählt. Dann sind die Matrizen

$$\begin{pmatrix} \sigma & \tau \\ \sigma' & \tau' \end{pmatrix}, \quad \begin{pmatrix} \tau' & -\tau \\ -\sigma' & \sigma \end{pmatrix} \in R^{2 \times 2}$$

invers zueinander, insbesondere also invertierbar. Entsprechend sieht man, dass für $\sigma\tau' - \tau\sigma' = \pm 1$ und $1 \leq i < j \leq m$ auch die Matrizen des Typs

$$E^{ij}(\sigma, \tau, \sigma', \tau') = \begin{pmatrix} 1 & & & \\ & \sigma & \tau & \\ & & 1 & \\ & \sigma' & \tau' & \\ & & & 1 \end{pmatrix} \in R^{m \times m}$$

invertierbar sind. Hierbei stehen die Elemente $\sigma, \tau, \sigma', \tau'$ jeweils an den Positionen (i, i) , (i, j) , (j, i) , (j, j) , und mit "1" sind Serien von Elementen 1 auf der Diagonalen angedeutet. Im Übrigen ist die Matrix $E^{ij}(\sigma, \tau, \sigma', \tau')$ mit Elementen 0 aufgefüllt, die der Übersichtlichkeit halber aber nicht ausgedrückt sind. Multipliziert man nun A von links mit $E^{ij}(\sigma, \tau, \sigma', \tau')$, so hat dies folgenden Effekt: Als neue i -te Zeile erhält man die Summe des σ -fachen der alten i -ten Zeile und des τ -fachen der alten j -ten Zeile. Entsprechend ist die neue j -te Zeile die Summe des σ' -fachen der alten i -ten Zeile und des τ' -fachen der alten j -ten Zeile. Beispielsweise ergibt sich eine Vertauschung der i -ten und j -ten Zeile mit den Konstanten

$$\sigma = 0, \quad \tau = 1, \quad \sigma' = 1, \quad \tau' = 0,$$

sowie die Addition des ε -fachen der j -ten Zeile zur i -ten Zeile mit

$$\sigma = 1, \quad \tau = \varepsilon, \quad \sigma' = 0, \quad \tau' = 1.$$

Mittels Transponierens sieht man, dass analoge Spaltenumformungen von A durch Multiplikation von rechts mit Matrizen des Typs $E^{ij}(\sigma, \tau, \sigma', \tau') \in R^{n \times n}$ generiert werden können.

Wir bezeichnen nun für Elemente $\alpha \in R - \{0\}$ mit $\delta(\alpha)$ die Anzahl der Primfaktoren von α ; dies ist gemäß Lemma 2 gerade die Länge des Restklassenrings $R/\alpha R$. Weiter setzen wir, ähnlich wie im Falle euklidischer Ringe,

$$d(A) := \min\{\delta(\alpha); \alpha \text{ ist Koeffizient} \neq 0 \text{ von } A\}$$

mit dem Ziel, $d(A)$ schrittweise zu verringern, solange bis es einen Koeffizienten α von A gibt, der alle übrigen Koeffizienten teilt. Durch Vertauschen von Zeilen

und Spalten können wir wiederum $d(A) = \delta(\alpha_{11})$ annehmen. Ist nun eines der Elemente der 1. Spalte, etwa α_{i1} , kein Vielfaches von α_{11} , so bilde man den größten gemeinsamen Teiler β von α_{11} und α_{i1} . Für diesen gilt dann notwendig $\delta(\beta) < \delta(\alpha_{11})$, und es erzeugt β gemäß 5.2/16 das Ideal $R\alpha_{11} + R\alpha_{i1}$, d. h. es existiert eine Gleichung des Typs

$$\beta = \sigma\alpha_{11} + \tau\alpha_{i1},$$

wobei $\sigma, \tau \in R$ notwendig teilerfremd sind und damit eine Gleichung des Typs

$$\sigma\tau' - \tau\sigma' = 1$$

mit gewissen Elementen $\sigma', \tau' \in R$ erfüllen. Multipliziert man nun A von links mit $E^{1i}(\sigma, \tau, \sigma', \tau')$, so etabliert dieser Prozess in A an der Position $(1, 1)$ das Element β und verringert somit das Minimum $d(A)$. Iteriert man das Verfahren wie im Falle euklidischer Ringe, so kann man schließlich erreichen, dass die Elemente $\alpha_{21}, \dots, \alpha_{m1}$ durch α_{11} teilbar sind bzw., indem man geeignete Vielfache der 1. Zeile von den restlichen subtrahiert, dass $\alpha_{21} = \dots = \alpha_{m1} = 0$ gilt. In gleicher Weise kann man mittels entsprechender Spaltenumformungen die Elemente $\alpha_{12}, \dots, \alpha_{1n}$ trivialisieren usw. Wir sehen also, dass sich die Matrix A schrittweise wie im Falle euklidischer Ringe abändern lässt, bis schließlich die gewünschte Gestalt erreicht ist. $\square$

Wir wollen das für euklidische Ringe beschriebene Verfahren an einem einfachen Beispiel demonstrieren und betrachten hierzu die Matrix

$$A = \begin{pmatrix} 6 & 2 & 5 \\ 32 & 2 & 28 \\ 30 & 2 & 26 \end{pmatrix} \in \mathbb{Z}^{3 \times 3}.$$

Der Bequemlichkeit halber lassen wir zur Bestimmung der Elementarteiler von A neben den oben verwendeten elementaren Zeilen- und Spaltenumformungen auch noch die Multiplikation einer Zeile bzw. Spalte mit einer Einheit unseres Ringes $R = \mathbb{Z}$ zu. Dies ist erlaubt, denn auch diese Umformungen lassen sich als Multiplikation von links bzw. rechts mit invertierbaren Elementarmatrizen interpretieren, und zwar mit solchen, die aus der Einheitsmatrix hervorgehen, indem man einen der Diagonaleinträge 1 durch eine Einheit aus R ersetzt. Wir wollen uns ansonsten aber an das für euklidische Ringe geschilderte Verfahren halten, obwohl sich die Bestimmung der Elementarteiler von A durch eine geschicktere Wahl der elementaren Umformungen noch vereinfachen ließe.

$$\begin{aligned}
A = \begin{pmatrix} 6 & 2 & 5 \\ 32 & 2 & 28 \\ 30 & 2 & 26 \end{pmatrix} &\xrightarrow{(1)} \begin{pmatrix} 2 & 6 & 5 \\ 2 & 32 & 28 \\ 2 & 30 & 26 \end{pmatrix} \xrightarrow{(2)} \begin{pmatrix} 2 & 6 & 5 \\ 0 & 26 & 23 \\ 0 & 24 & 21 \end{pmatrix} \xrightarrow{(3)} \begin{pmatrix} 2 & 0 & 1 \\ 0 & 26 & 23 \\ 0 & 24 & 21 \end{pmatrix} \\
&\xrightarrow{(4)} \begin{pmatrix} 1 & 0 & 2 \\ 23 & 26 & 0 \\ 21 & 24 & 0 \end{pmatrix} \xrightarrow{(5)} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 26 & -46 \\ 0 & 24 & -42 \end{pmatrix} \xrightarrow{(6)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 26 & 46 \\ 0 & 24 & 42 \end{pmatrix} \xrightarrow{(7)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 24 & 42 \\ 0 & 26 & 46 \end{pmatrix} \\
&\xrightarrow{(8)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 24 & 42 \\ 0 & 2 & 4 \end{pmatrix} \xrightarrow{(9)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 4 \\ 0 & 24 & 42 \end{pmatrix} \xrightarrow{(10)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 4 \\ 0 & 0 & -6 \end{pmatrix} \xrightarrow{(11)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 6 \end{pmatrix}
\end{aligned}$$

Es ergeben sich also 1, 2, 6 als die Elementarteiler von A , wobei im Einzelnen die folgenden elementaren Umformungen ausgeführt wurden:

- (1) Vertauschen von 1. und 2. Spalte
- (2) Subtrahieren der 1. von der 2. und der 3. Zeile
- (3) Subtrahieren des 3-fachen bzw. 2-fachen der 1. Spalte von der 2. bzw. 3. Spalte
- (4) Vertauschen von 1. und 3. Spalte
- (5) Subtrahieren des 23-fachen bzw. 21-fachen der 1. Zeile von der 2. bzw. 3. Zeile
- (6) Subtrahieren des 2-fachen der 1. Spalte von der 3. Spalte, Multiplikation der 3. Spalte mit -1
- (7) Vertauschen von 2. und 3. Zeile
- (8) Subtrahieren der 2. Zeile von der 3. Zeile
- (9) Vertauschen von 2. und 3. Zeile
- (10) Subtrahieren des 12-fachen der 2. Zeile von der 3. Zeile
- (11) Subtrahieren des 2-fachen der 2. Spalte von der 3. Spalte, Multiplikation der 3. Spalte mit -1

Als Nächstes wenden wir uns nun der Eindeutigkeitsaussage in Lemma 5 bzw. Theorem 4 zu und beginnen mit einem grundlegenden Lemma, welches auch später noch von Bedeutung sein wird.

Lemma 6. *Es sei R ein Hauptidealring und*

$$Q \simeq \bigoplus_{i=1}^s R/\alpha_i R$$

ein Isomorphismus von R -Moduln, wobei $\alpha_1, \dots, \alpha_s \in R - \{0\}$ Nichteinheiten mit $\alpha_i \mid \alpha_{i+1}$ für $1 \leq i < s$ sind und $\bigoplus_{i=1}^s R/\alpha_i R$ die konstruierte direkte Summe der R -Moduln $R/\alpha_i R$ bezeichne. Dann sind $\alpha_1, \dots, \alpha_s$ bis auf Assoziiertheit eindeutig durch Q bestimmt.

Beweis. Aus technischen Gründen invertieren wir die Nummerierung der α_i und betrachten zwei Zerlegungen

$$Q \simeq \bigoplus_{i=1}^s R/\alpha_i R \simeq \bigoplus_{j=1}^t R/\beta_j R$$

mit $\alpha_{i+1} \mid \alpha_i$ für $1 \leq i < s$ sowie $\beta_{j+1} \mid \beta_j$ für $1 \leq j < t$. Falls es einen Index $k \leq \min\{s, t\}$ mit $\alpha_k R \neq \beta_k R$ gibt, so wähle man k minimal mit dieser Eigenschaft. Da $\alpha_i R = \beta_i R$ für $1 \leq i < k$ und da $\alpha_{k+1}, \dots, \alpha_s$ sämtlich Teiler von α_k sind, zerlegt sich $\alpha_k Q$ zu

$$\begin{aligned} \alpha_k Q &\simeq \bigoplus_{i=1}^{k-1} \alpha_k \cdot (R/\alpha_i R) \\ &\simeq \bigoplus_{i=1}^{k-1} \alpha_k \cdot (R/\alpha_i R) \oplus \bigoplus_{j=k}^t \alpha_k \cdot (R/\beta_j R). \end{aligned}$$

Wir benutzen nun die Lemmata 2 und 3. Wegen $\ell_R(\alpha_k Q) \leq \ell_R(Q) < \infty$ ergibt sich durch Vergleich beider Seiten $\ell_R(\alpha_k \cdot (R/\beta_j R)) = 0$ für $j = k, \dots, t$. Letzteres bedeutet aber insbesondere $\alpha_k \cdot (R/\beta_k R) = 0$ bzw. $\alpha_k R \subset \beta_k R$. Entsprechend zeigt man $\beta_k R \subset \alpha_k R$ und somit $\alpha_k R = \beta_k R$, im Widerspruch zu unserer Annahme. Es gilt daher $\alpha_i R = \beta_i R$ für alle Indizes i mit $1 \leq i \leq \min\{s, t\}$. Hat man weiter $s \leq t$, so folgt, wiederum unter Benutzung von Lemma 3, dass $\bigoplus_{j=s+1}^t R/\beta_j R$ von der Länge 0 ist, also $R/\beta_j R$ für $j = s+1, \dots, t$ die Länge 0 hat. Andererseits ist aber β_j für $j = 1, \dots, t$ keine Einheit. Daher besitzt jeder Modul $R/\beta_j R$ gemäß Lemma 2 eine Länge größer als 0, und es folgt $s = t$. $\square$

Beweis der Eindeutigkeitsaussage von Theorem 4 und Lemma 5. Die Eindeutigkeit in Lemma 5 ist eine Konsequenz der Eindeutigkeitsaussage in Theorem 4. Es genügt daher, die Eindeutigkeit der Elementarteiler in Theorem 4 zu zeigen. Seien also $x_1, \dots, x_s$ Teil einer Basis von F , und seien $\alpha_1, \dots, \alpha_s \in R$ mit $\alpha_i \mid \alpha_{i+1}$ für $1 \leq i < s$, so dass $\alpha_1 x_1, \dots, \alpha_s x_s$ eine Basis von M bilden. Betrachtet man dann den Untermodul

$$F' = \bigoplus_{i=1}^s R x_i = \{a \in F; \text{ es existiert ein } \alpha \in R - \{0\} \text{ mit } \alpha a \in M\}$$

von F , so hängt F' nur von M und nicht von der speziellen Wahl der Elemente $x_1, \dots, x_s$ ab. Der kanonische Homomorphismus

$$F' \longrightarrow \bigoplus_{i=1}^s R/\alpha_i R, \quad \sum_{i=1}^s \gamma_i x_i \longmapsto (\overline{\gamma}_1, \dots, \overline{\gamma}_s),$$

wobei $\overline{\gamma}_i$ jeweils die Restklasse von γ_i in $R/\alpha_i R$ bezeichne, ist dann surjektiv und besitzt M als Kern, induziert also aufgrund des Homomorphiesatzes einen Isomorphismus

$$F'/M \xrightarrow{\sim} \bigoplus_{i=1}^s R/\alpha_i R.$$

Aus der Eindeutigkeitsaussage in Lemma 6 ist dann zu folgern, dass jedenfalls die Nichteinheiten unter den $\alpha_1, \dots, \alpha_s$ bis auf Assoziiertheit eindeutig durch M bestimmt sind.

Um nun einzusehen, dass auch die Anzahl der Einheiten unter den $\alpha_1, \dots, \alpha_s$ eindeutig durch M bestimmt ist, wollen wir zeigen, dass s als Anzahl der Elemente einer Basis von M bzw. F' eindeutig bestimmt ist. Da jede Basis $z_1, \dots, z_s$ von M Anlass zu einem Isomorphismus $M \xrightarrow{\sim} R^s$ gibt, genügt es zu zeigen, dass die Existenz eines Isomorphismus $R^s \xrightarrow{\sim} R^{s'}$ bereits $s = s'$ nach sich zieht. Im Falle eines Körpers R ist dies klar aufgrund der Dimensionstheorie für Vektorräume; vgl. z. B. 2.1/8. Ist jedoch R kein Körper, so enthält $R - R^*$ mindestens ein von Null verschiedenes Element, und dieses lässt sich gemäß 5.2/13 als Produkt von Primelementen schreiben. Man findet daher in R mindestens ein Primelement p , und es ist leicht nachzuprüfen, dass jeder Isomorphismus von R -Moduln $R^s \xrightarrow{\sim} R^{s'}$ einen Isomorphismus von R/pR -Moduln

$$(R/pR)^s = R^s/pR^s \xrightarrow{\sim} R^{s'}/pR^{s'} = (R/pR)^{s'}$$

induziert. Da aber R/pR nach 5.2/17 ein Körper ist, können wir wiederum $s = s'$ schließen. Alternativ kann man an dieser Stelle auch die Eindeutigkeitsaussage von Lemma 6 ausnutzen. $\square$

Korollar 7. *Es sei R ein Hauptidealring und F ein endlicher freier R -Modul. Dann besitzen je zwei Basen von F gleiche Länge.¹ Diese Länge wird auch als der Rang von F bezeichnet.*

Beweis. Die Behauptung wurde bereits in obigem Beweis hergeleitet, um nachzuweisen, dass die Anzahl der Elementarteiler eines Untermoduls $M \subset F$ eindeutig bestimmt ist. Im Übrigen folgt die Aussage von Korollar 7 aber auch formal aus der Eindeutigkeitsaussage von Theorem 4, da die Anzahl der Elementarteiler des trivialen Untermoduls $F \subset F$ eindeutig bestimmt ist. $\square$

Für die Berechnung von Elementarteilern in der Situation von Theorem 4 ist es wichtig zu wissen, dass wir diese als Elementarteiler einer Matrix erhalten haben, denn die Elementarteiler einer Matrix können mit Hilfe des im Beweis zu Lemma 5 gegebenen praktischen Verfahrens bestimmt werden. Wir wollen die genauen Bedingungen hier noch einmal gesondert formulieren.

Korollar 8. *Es sei R ein Hauptidealring, F ein endlicher freier R -Modul und $M \subset F$ ein Untermodul. Dann ist M endlich erzeugt.*

Es sei F' ein weiterer endlicher freier R -Modul und $f: F' \rightarrow F$ eine R -lineare Abbildung, welche eine Basis von F' auf ein Erzeugendensystem von M abbildet, also mit $\text{Bild } f = M$. Ist dann A eine Matrix mit Koeffizienten aus R , welche f bezüglich geeigneter Basen in F und F' beschreibt, so stimmen die Elementarteiler von A überein mit denjenigen von $M \subset F$.

¹ Die Aussage gilt allgemeiner für endliche freie Moduln über beliebigen Ringen (kommutativ mit 1); vgl. Aufgabe 6.

Aufgaben

Es sei R ein beliebiger Ring (kommutativ mit 1), sofern nichts anderes verlangt ist.

1. Man bestimme die Elementarteiler der folgenden Matrizen:

$$\begin{pmatrix} 2 & 6 & 8 \\ 3 & 1 & 2 \\ 9 & 5 & 4 \end{pmatrix}, \quad \begin{pmatrix} 4 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 15 \end{pmatrix} \in \mathbb{Z}^{3 \times 3}$$

2. Es sei R ein Hauptidealring und $A = (\alpha_{ij})_{i,j} \in R^{m \times n}$ eine nicht-triviale Matrix mit Koeffizienten aus R . Sind dann $\alpha_1, \dots, \alpha_s$ die Elementarteiler von A , so gilt $s > 0$ und $\alpha_1 = \text{ggT}(\alpha_{ij}; i = 1, \dots, m, j = 1, \dots, n)$.
3. Es seien $a_{11}, \dots, a_{1n}$ teilerfremde Elemente eines Hauptidealrings R , d. h. es gelte $\text{ggT}(a_{11}, \dots, a_{1n}) = 1$. Man zeige, es gibt Elemente $a_{ij} \in R$, $i = 2, \dots, n$, $j = 1, \dots, n$, so dass die Matrix $(a_{ij})_{i,j=1,\dots,n}$ in $R^{n \times n}$ invertierbar ist.
4. Es sei $f: M \rightarrow N$ ein Homomorphismus endlicher freier Moduln über einem Hauptidealring R , d. h. M und N mögen jeweils endliche Basen besitzen. Man verwende die Aussage des Elementarteilersatzes und folgere die Existenz von Basen X von M und Y von N , sowie von Null verschiedener Elemente $\alpha_1, \dots, \alpha_s \in R$ mit $\alpha_i | \alpha_{i+1}$ für $1 \leq i < s$, so dass gilt:

$$A_{f,X,Y} = \begin{pmatrix} \alpha_1 & 0 & & 0 & 0 & \dots & 0 \\ 0 & \alpha_2 & & 0 & 0 & \dots & 0 \\ & & \ddots & & & & \\ & & & \ddots & & & \\ 0 & 0 & & & \alpha_s & 0 & \dots & 0 \\ 0 & 0 & & & 0 & 0 & \dots & 0 \\ 0 & 0 & & & 0 & 0 & \dots & 0 \end{pmatrix}$$

Dabei sind die Elemente $\alpha_1, \dots, \alpha_s$ bis auf Assoziiiertheit eindeutig bestimmt.

5. Man bestimme die Länge von $(\mathbb{Z}/15\mathbb{Z})^4$ als $\mathbb{Z}$ -Modul.
6. Es sei M ein endlich erzeugter R -Modul, der zudem frei ist. Man zeige:
 - (i) M besitzt eine endliche Basis.
 - (ii) Je zwei Basen von M bestehen aus gleichviel Elementen.

(Hinweis: Um die Argumentation zu erleichtern, nehme man für Teil (ii) an, dass R ein Integritätsring ist.)

7. Für den Elementarteilersatz (Theorem 4) hatten wir einen Untermodul M eines endlichen freien Moduls F betrachtet. Man zeige, dass die Aussage des Satzes erhalten bleibt, wenn man alternativ F als frei und M als endlich erzeugt voraussetzt.
8. Es sei V ein Vektorraum über einem Körper K . Zu jedem Endomorphismus $f: V \rightarrow V$ kann man auf V die zugehörige Struktur als $K[T]$ -Modul betrachten, welche charakterisiert ist durch $T \cdot v = f(v)$ für Elemente $v \in V$. Man zeige, dass man auf diese Weise eine Bijektion zwischen der Menge $\text{End}_K(V)$ und der Menge der $K[T]$ -Modul-Strukturen auf V erhält, die verträglich sind mit der Struktur von V als K -Vektorraum.

9. (1. *Isomorphiesatz für Moduln*) Es seien N, N' Untermoduln eines R -Moduls M . Man zeige: Die kanonische Abbildung $N \hookrightarrow N + N' \rightarrow (N + N')/N'$ besitzt $N \cap N'$ als Kern und induziert einen Isomorphismus

$$N/(N \cap N') \xrightarrow{\sim} (N + N')/N'.$$

10. (2. *Isomorphiesatz für Moduln*) Es seien M ein R -Modul und $N \subset N' \subset M$ Untermoduln. Man zeige:

- (i) Die kanonische Abbildung $N' \hookrightarrow M \rightarrow M/N$ besitzt N als Kern und induziert einen Monomorphismus $N'/N \hookrightarrow M/N$. Folglich lässt sich N'/N mit seinem Bild in M/N identifizieren und somit als Untermodul von M/N auffassen.
- (ii) Die Projektion $M \rightarrow M/N'$ faktorisiert über M/N , d. h. lässt sich als Komposition $M \xrightarrow{\pi} M/N \xrightarrow{f} M/N'$ schreiben, mit einem Modulhomomorphismus f und der kanonischen Projektion π .
- (iii) f besitzt N'/N als Kern und induziert einen Isomorphismus

$$(M/N)/(N'/N) \xrightarrow{\sim} M/N'.$$

6.4 Endlich erzeugte Moduln über Hauptidealringen

Wir wollen nun einige Folgerungen zur Struktur endlich erzeugter Moduln über Hauptidealringen aus dem Elementarteilersatz ziehen. Im nächsten Abschnitt sollen die gewonnenen Struktursätze dann in Ergebnisse über Normalformen von Endomorphismen von Vektorräumen umgesetzt werden. Als Hilfsmittel benötigen wir noch den so genannten *Chinesischen Restsatz*, den wir als Erstes beweisen.

Satz 1. *Es sei R ein Hauptidealring und*

$$a = \varepsilon p_1^{n_1} \dots p_r^{n_r}$$

eine Primfaktorzerlegung in R mit einer Einheit ε und paarweise nicht-assozierten Primelementen p_i . Ist dann

$$\pi_i: R \longrightarrow R/(p_i^{n_i}), \quad i = 1, \dots, r,$$

jeweils die kanonische Projektion, so ist der Homomorphismus

$$\varphi: R \longrightarrow R/(p_1^{n_1}) \times \dots \times R/(p_r^{n_r}), \quad a \longmapsto (\pi_1(a), \dots, \pi_r(a)),$$

surjektiv und erfüllt $\ker \varphi = (a)$, induziert also einen Isomorphismus

$$R/(a) \xrightarrow{\sim} R/(p_1^{n_1}) \times \dots \times R/(p_r^{n_r}).$$

Dabei ist $R/(p_1^{n_1}) \times \dots \times R/(p_r^{n_r})$ als Ring unter komponentenweiser Addition und Multiplikation zu verstehen.

Beweis. Zunächst zeigen wir, dass φ surjektiv ist. Hierzu genügt es offenbar nachzuprüfen, dass es Elemente $e_1, \dots, e_r \in R$ gibt mit

$$\pi_i(e_j) = \begin{cases} 1 & \text{für } i = j, \\ 0 & \text{sonst.} \end{cases}$$

Da die Elemente $p_j^{n_j}$ und $\prod_{i \neq j} p_i^{n_i}$ für fest gewähltes $j \in \{1, \dots, r\}$ teilerfremd sind, ist das von ihnen erzeugte Hauptideal das Einheitsideal. Folglich existiert für jedes j eine Gleichung des Typs

$$e_j + d_j = 1 \quad \text{mit } e_j \in \left(\prod_{i \neq j} p_i^{n_i}\right), \quad d_j \in (p_j^{n_j}).$$

Es gilt dann $\pi_i(e_j) = 0$ für $i \neq j$ und $\pi_j(e_j) = 1$, wie gewünscht.

Der Kern von φ besteht aus allen Elementen aus R , die durch die Potenzen $p_i^{n_i}$, $i = 1, \dots, r$, teilbar sind, und damit aus den Elementen, die durch a teilbar sind. Somit ergibt sich $\ker \varphi = (a)$, und der behauptete Isomorphismus folgt aus dem Isomorphiesatz 5.1/10. $\square$

In der Situation von Satz 1 lässt sich der Restklassenring $R/(a)$ auch als R -Modul auffassen. Die Aussage des Chinesischen Restsatzes besagt dann, dass $R/(a)$ isomorph zu der konstruierten direkten Summe der R -Moduln $R/(p_i^{n_i})$ ist, also

$$R/(a) \simeq R/(p_1^{n_1}) \oplus \dots \oplus R/(p_r^{n_r}),$$

wobei wir, wie zu Beginn von Abschnitt 6.3 erläutert, den i -ten Summanden $R/(p_i^{n_i})$ mit dem entsprechenden Untermodul

$$0 \times \dots \times 0 \times R/(p_i^{n_i}) \times 0 \times \dots \times 0 \subset R/(p_1^{n_1}) \times \dots \times R/(p_r^{n_r})$$

zu identifizieren haben.

Für einen Modul M über einem *Integritätsring* R definiert man den so genannten *Torsionsuntermodul* T durch

$$T = \{a \in M; \text{ es existiert ein } \alpha \in R - \{0\} \text{ mit } \alpha a = 0\}.$$

Man prüft leicht nach, dass T in der Tat ein Untermodul von M ist, indem man die Nullteilerfreiheit von R benutzt. Es heißt M ein *Torsionsmodul*, wenn M mit seinem Torsionsuntermodul übereinstimmt. Der Torsionsuntermodul eines freien Moduls ist stets trivial, freie Moduln sind daher sozusagen als das “Gegenstück” zu den Torsionsmoduln anzusehen.

Ist M ein Modul über einem Ring R , so bezeichnet man eine exakte Sequenz von R -linearen Abbildungen

$$R^n \xrightarrow{\psi} R^m \xrightarrow{\varphi} M \longrightarrow 0$$

mit geeigneten endlichen freien R -Moduln R^m und R^n auch als eine *endliche Präsentation* von M . Genauer besteht diese aus einem Epimorphismus

$\varphi: R^m \longrightarrow M$ und einer R -linearen Abbildung $\psi: R^n \longrightarrow R^m$ mit $\text{im } \psi = \ker \varphi$. Eine solche endliche Präsentation existiert stets, wenn M ein endlich erzeugter Modul über einem Hauptidealring R ist. Man wähle nämlich ein endliches Erzeugendensystem $z_1, \dots, z_m$ in M und betrachte die R -lineare Abbildung $\varphi: R^m \longrightarrow M$, welche die kanonische Basis von R^m auf $z_1, \dots, z_m$ abbildet. Dann ist φ ein Epimorphismus, und der Untermodul $\ker \varphi \subset R^m$ ist gemäß 6.3/4 endlich erzeugt. Wir können daher einen weiteren Epimorphismus $\psi: R^n \longrightarrow \ker \varphi \subset R^m$ und folglich eine endliche Präsentation von M finden. Ist in dieser Situation $A \in R^{m \times n}$ eine Matrix, welche die R -lineare Abbildung ψ bezüglich geeigneter Basen in R^m und R^n beschreibt, so bezeichnen wir die Elementarteiler von A auch als die *Elementarteiler der betrachteten endlichen Präsentation* von M . Nach 6.3/8 stimmen diese mit den Elementarteilern von $\ker \varphi \subset R^m$ überein und hängen daher nicht von der Auswahl der Basen in R^m und R^n ab.

Wir wollen nun aus dem Elementarteilersatz 6.3/4 verschiedene Versionen des so genannten Hauptsatzes über endlich erzeugte Moduln über Hauptidealringen herleiten. Wir beginnen mit einer grundlegenden Folgerung aus dem Elementarteilersatz.

Satz 2. *Es sei M ein endlich erzeugter Modul über einem Hauptidealring R und $T \subset M$ sein Torsionsuntermodul. Dann gibt es einen endlichen freien Untermodul $F \subset M$, etwa $F \simeq R^d$, sowie Nichteinheiten $\alpha_1, \dots, \alpha_s \in R - \{0\}$ mit $\alpha_j | \alpha_{j+1}$ für $1 \leq j < s$ und*

$$M = F \oplus T, \quad T \simeq \bigoplus_{j=1}^s R/\alpha_j R.$$

Dabei ist d eindeutig bestimmt; man bezeichnet d als den Rang von M . Weiter sind die Elemente $\alpha_1, \dots, \alpha_s$ eindeutig bestimmt bis auf Assoziiertheit. Sie stimmen überein mit den Nichteinheiten unter den Elementarteilern einer jeden endlichen Präsentation von M .

Beweis. Wir gehen von einer endlichen Präsentation von M aus und betrachten den zugehörigen Epimorphismus $\varphi: R^m \longrightarrow M$. Dann folgt $M \simeq R^m / \ker \varphi$ aufgrund des Homomorphiesatzes. Auf die Situation $\ker \varphi \subset R^m$ können wir nun den Elementarteilersatz 6.3/4 anwenden (wobei wir berücksichtigen, dass jede Basis von R^m gemäß 6.3/7 aus genau m Elementen besteht). Es existieren daher Elemente $x_1, \dots, x_m$, die eine Basis von R^m bilden, sowie Elemente $\alpha_1, \dots, \alpha_s \in R - \{0\}$, $s \leq m$, mit $\alpha_j | \alpha_{j+1}$ für $1 \leq j < s$, so dass $\alpha_1 x_1, \dots, \alpha_s x_s$ eine Basis von $\ker \varphi$ bilden. Indem wir $\alpha_{s+1} = \dots = \alpha_m = 0$ setzen, können wir den Epimorphismus

$$\varphi': R^m = \bigoplus_{j=1}^m R x_j \longrightarrow \bigoplus_{j=1}^m R x_j / R \alpha_j x_j, \quad (\gamma_1, \dots, \gamma_m) \longmapsto (\bar{\gamma}_1, \dots, \bar{\gamma}_m),$$

betrachten, wobei $\bar{\gamma}_j$ jeweils die Restklasse von γ_j in $R x_j / R \alpha_j x_j$ bezeichne. Nach Konstruktion gilt $\ker \varphi' = \ker \varphi$ und folglich aufgrund des Homomorphiesatzes

$$M \simeq R^m / \ker \varphi = \left(\bigoplus_{j=1}^m R x_j \right) / \ker \varphi' \simeq \bigoplus_{j=1}^m R x_j / R \alpha_j x_j \simeq R^{m-s} \oplus \bigoplus_{j=1}^s R / \alpha_j R,$$

also mit $d := m - s$ eine Zerlegung des behaupteten Typs, wenn wir Summanden $R / \alpha_j R = 0$, d. h. mit $\alpha_j \in R^*$ unterdrücken.

In dieser Zerlegung korrespondiert $\bigoplus_{j=1}^s R / \alpha_j R$ zu dem Torsionsuntermodul $T \subset M$ und ist daher eindeutig bestimmt. Indem wir triviale Summanden $R / \alpha_j R$ ignorieren, bzw. annehmen, dass $\alpha_1, \dots, \alpha_s$ keine Einheiten sind, ergibt sich die Eindeutigkeit der α_j mit 6.3/6. Insbesondere sind die α_j unabhängig von der betrachteten Präsentation von M und bestehen gemäß 6.3/8 aus den Nichteinheiten unter den Elementarteilern einer solchen Präsentation.

Um zu sehen, dass auch d eindeutig ist, betrachte man den Epimorphismus

$$M \xrightarrow{\sim} R^d \oplus \bigoplus_{j=1}^s R / \alpha_j R \longrightarrow R^d,$$

der sich aus dem obigen Isomorphismus und der Projektion auf den Summanden R^d zusammensetzt. Der Kern dieser Abbildung ist offenbar gerade der Torsionsuntermodul $T \subset M$, so dass man aufgrund des Homomorphiesatzes einen Isomorphismus $M/T \xrightarrow{\sim} R^d$ erhält. Hieraus folgt die Eindeutigkeit von d mit 6.3/7. $\square$

Speziellere Versionen des Hauptsatzes über endlich erzeugte Moduln über Hauptidealringen lassen sich mittels des Chinesischen Restsatzes aus Satz 2 folgern.

Satz 3. *Es sei M ein endlich erzeugter Modul über einem Hauptidealring R und $T \subset M$ der Torsionsuntermodul. Weiter sei $P \subset R$ ein Vertretersystem der Primelemente von R , und für $p \in P$ bezeichne*

$$M_p = \{x \in M ; p^n x = 0 \text{ für geeignetes } n \in \mathbb{N}\}$$

den so genannten Untermodul der p -Torsion in M . Dann gilt

$$T = \bigoplus_{p \in P} M_p,$$

und es gibt einen endlichen freien Untermodul $F \subset M$, etwa $F \simeq R^d$, mit

$$M = F \oplus T,$$

wobei d eindeutig bestimmt ist und M_p für fast alle $p \in P$ verschwindet. Weiter gibt es zu jedem $p \in P$ natürliche Zahlen $1 \leq n(p, 1) \leq \dots \leq n(p, s_p)$ mit

$$M_p \simeq \bigoplus_{j_p=1 \dots s_p} R / p^{n(p, j_p)} R,$$

wobei die $s_p, n(p, j_p)$ durch die Isomorphie

$$M \simeq F \oplus \bigoplus_{p \in P} \bigoplus_{j_p=1 \dots s_p} R/p^{n(p,j_p)} R$$

eindeutig bestimmt sind und $s_p = 0$ für fast alle p gilt.

Bevor wir zum Beweis kommen, wollen wir diesen Hauptsatz auch noch speziell für endlich erzeugte $\mathbb{Z}$ -Moduln formulieren, als *Hauptsatz über endlich erzeugte abelsche Gruppen*.

Korollar 4. *Es sei G eine endlich erzeugte abelsche Gruppe, P sei die Menge der Primzahlen in $\mathbb{N}$. Dann gestattet G eine Zerlegung in Untergruppen*

$$G = F \oplus \bigoplus_{p \in P} \bigoplus_{j_p=1 \dots s_p} G_{p,j_p},$$

wobei F frei ist, etwa $F \simeq \mathbb{Z}^d$, und G_{p,j_p} zyklisch von p -Potenz-Ordnung, etwa $G_{p,j_p} \simeq \mathbb{Z}/p^{n(p,j_p)}\mathbb{Z}$ mit $1 \leq n(p,1) \leq \dots \leq n(p,s_p)$. Die Zahlen $d, s_p, n(p,j_p)$ sind eindeutig durch G bestimmt, ebenso die Untergruppen $G_p = \bigoplus_{j_p=1 \dots s_p} G_{p,j_p}$, wobei s_p für fast alle $p \in P$ verschwindet.

Wenn G eine endlich erzeugte Torsionsgruppe ist, also ein über $\mathbb{Z}$ endlich erzeugter Torsionsmodul, so besitzt G keinen freien Anteil und besteht daher, wie man insbesondere mit Korollar 4 sieht, nur aus endlich vielen Elementen. Umgekehrt ist jede endliche abelsche Gruppe natürlich eine endlich erzeugte Torsionsgruppe.

Nun zum *Beweis von Satz 3*. Wir beginnen mit der Zerlegung

$$M \simeq R^d \oplus \bigoplus_{j=1}^s R/\alpha_j R$$

aus Satz 2, wobei insbesondere d eindeutig bestimmt ist. Hierbei korrespondiert $\bigoplus_{j=1}^s R/\alpha_j R$ zu dem Torsionsuntermodul $T \subset M$, sowie R^d zu einem freien Modul $F \subset M$, und es gilt $M = F \oplus T$.

Man zerlege die α_j nun in Primfaktoren, etwa $\alpha_j = \varepsilon_j p_1^{n(1,j)} \dots p_r^{n(r,j)}$ mit Einheiten $\varepsilon_j \in R^*$, paarweise verschiedenen Primelementen $p_1, \dots, p_r \in P$, sowie Exponenten $n(i,j) \in \mathbb{N}$, die auch trivial sein dürfen. Dann gilt aufgrund von Satz 1

$$T \simeq \bigoplus_{j=1}^s \bigoplus_{i=1}^r R/p_i^{n(i,j)} R \simeq \bigoplus_{i=1}^r \bigoplus_{j=1}^s R/p_i^{n(i,j)} R.$$

In dieser Zerlegung korrespondiert $\bigoplus_{j=1}^s R/p_i^{n(i,j)} R$ offenbar gerade zu dem Untermodul $M_{p_i} \subset M$ der p_i -Torsion und ist deshalb eindeutig bestimmt; die Restklasse von p_i ist nämlich in jedem Restklassenring der Form $R/p_{i'}^n R$ mit $i' \neq i$ eine Einheit. Somit folgt aus obiger Zerlegung insbesondere $T = \bigoplus_{p \in P} M_p$. Da im Übrigen in der Zerlegung

$$M_{p_i} \simeq \bigoplus_{j=1}^s R/p_i^{n(i,j)} R$$

die Terme zu Exponenten $n(i, j) = 0$ trivial sind und die restlichen Terme gemäß 6.3/6 eindeutig bestimmt sind, ergibt sich insgesamt die Behauptung von Satz 3. $\square$

Aufgaben

Für eine Gruppe G bezeichnet man mit $\text{ord } G$ die Anzahl ihrer Elemente und nennt dies die *Ordnung* von G .

1. Es sei G eine abelsche Gruppe der Ordnung $n < \infty$. Man zeige: Zu jedem Teiler d von n gibt es eine Untergruppe $H \subset G$ der Ordnung d . Andererseits besitzt jede Untergruppe $H \subset G$ eine Ordnung, die ein Teiler von n ist.
2. Zu $n \in \mathbb{N}$ gibt es höchstens endlich viele Klassen isomorpher abelscher Gruppen der Ordnung n . Wie groß ist diese Anzahl für $n = 20$ bzw. $n = 30$?
3. Es seien $a_1, \dots, a_n \in \mathbb{N} - \{0\}$ paarweise teilerfremd und $r_1, \dots, r_n \in \mathbb{N}$ Zahlen mit $0 \leq r_i < a_i$ für $i = 1, \dots, n$. Man zeige: Es existiert eine Zahl $a \in \mathbb{N}$, so dass a bei ganzzahliger Division durch a_i jeweils den Rest r_i lässt.
4. Es sei M ein endlich erzeugter Modul über einem Hauptidealring R . Für den Fall, dass M kein Torsionsmodul ist, zeige man: Es existieren endlich viele freie Untermoduln $F_1, \dots, F_r \subset M$ mit $M = \sum_{i=1}^r F_i$.
5. Über einem Hauptidealring R betrachte man endliche freie Moduln F und F' mit Untermoduln $M \subset F$ und $M' \subset F'$, so dass die Restklassenmoduln F/M und F'/M' jeweils Torsionsmoduln sind. Man zeige: Es existiert genau dann ein Isomorphismus $F/M \xrightarrow{\sim} F'/M'$, wenn die Nichteinheiten unter den Elementarteilern von $M \subset F$ mit denen von $M' \subset F'$ übereinstimmen.
6. Es sei K ein Körper und $G \subset K^*$ eine endliche Untergruppe der multiplikativen Gruppe der Einheiten von K . Man zeige, dass G *zyklisch* ist, d. h. dass es ein $n \in \mathbb{Z}$ mit $G \simeq \mathbb{Z}/n\mathbb{Z}$ gibt. Hinweis: Man betrachte Nullstellen aus K von Polynomen des Typs $T^n - 1 \in K[T]$.

6.5 Allgemeine und Jordansche Normalform für Matrizen

Im Folgenden sei V stets ein Vektorraum über einem Körper K und $f: V \longrightarrow V$ ein Endomorphismus. Wie wir in Abschnitt 6.3 gesehen haben, lässt sich V als Modul über dem Polynomring $K[T]$ auffassen, indem man die Multiplikation der Variablen T mit Elementen von V durch Anwenden von f erklärt. Ist V von endlicher Dimension, so existiert ein eindeutig bestimmtes normiertes Polynom minimalen Grades $p_f \in K[T]$ mit $p_f \cdot v = 0$ für alle $v \in V$; man nennt p_f das *Minimalpolynom* zu f , vgl. 6.2/9. Es ist V dann insbesondere ein $K[T]$ -Torsionsmodul, sogar ein endlich erzeugter, da V als K -Vektorraum

endlich erzeugt ist. Wir können daher das Strukturresultat 6.4/3 auf V als $K[T]$ -Modul anwenden.

Definition 1. Es sei $f: V \longrightarrow V$ ein Endomorphismus eines K -Vektorraums V , sowie $U \subset V$ ein Untervektorraum.

- (i) U heißt f -invariant, wenn $f(U) \subset U$ gilt.
- (ii) U heißt f -zyklisch, wenn es einen Vektor $u \in U$ gibt, so dass die Folge $u, f(u), f^2(u), \dots$ ein Erzeugendensystem von U bildet; insbesondere ist U dann auch f -invariant.
- (iii) U heißt f -unzerlegbar, wenn U ein f -invarianter Untervektorraum $\neq 0$ ist, der sich nicht in eine direkte Summe zweier echter f -invarianter Untervektorräume zerlegen lässt.

Wir wollen die gerade eingeführten Eigenschaften von Unterräumen nun auch modultheoretisch interpretieren. Ein Modul M über einem Ring R wird als *monogen* bezeichnet, wenn er von einem einzigen Element erzeugt wird, d. h. wenn ein $a \in M$ existiert mit $M = Ra$.

Bemerkung 2. Es sei $f: V \longrightarrow V$ ein Endomorphismus eines K -Vektorraums V , wobei man V insbesondere auch als $K[T]$ -Modul bezüglich f betrachte. Für eine Teilmenge $U \subset V$ gilt dann:

- (i) U ist genau dann ein f -invarianter Untervektorraum von V , wenn U ein $K[T]$ -Untermodul von V ist.
- (ii) U ist genau dann ein f -zyklischer Untervektorraum von V , wenn U ein monogener $K[T]$ -Untermodul von V ist.
- (iii) U ist genau dann ein f -unzerlegbarer Untervektorraum von V , wenn U ein $K[T]$ -Untermodul $\neq 0$ von V ist, der sich nicht in eine direkte Summe zweier echter $K[T]$ -Untermoduln zerlegen lässt.

Beweis. Aussage (i) ist unmittelbar klar, denn für Untervektorräume $U \subset V$ bedeutet die Bedingung $f(u) \in U$ für $u \in U$ gerade $T \cdot u \in U$ und ist daher äquivalent zu $q \cdot u \in U$ für $q \in K[T]$, $u \in U$. In ähnlicher Weise erhält man Aussage (ii), denn $U = \sum_{i \in \mathbb{N}} K \cdot f^i(u)$ für ein $u \in U$ ist gleichbedeutend mit $U = \sum_{i \in \mathbb{N}} K \cdot T^i \cdot u$, also mit $U = K[T] \cdot u$. Aussage (iii) schließlich beruht auf der Tatsache, dass für f -invariante Untervektorräume bzw. $K[T]$ -Untermoduln $U', U'' \subset V$ die Summe $U = U' + U''$ genau dann direkt ist, wenn jedes $u \in U$ eine eindeutig bestimmte Darstellung $u = u' + u''$ mit $u' \in U'$, $u'' \in U''$ besitzt. Ob wir bei der skalaren Multiplikation Elemente aus K oder $K[T]$ zulassen, ist dabei ohne Belang. $\square$

Im Weiteren ist es auch wichtig zu wissen, wie man Homomorphismen von $K[T]$ -Moduln mittels Homomorphismen der unterliegenden K -Vektorräume interpretieren kann.

Bemerkung 3. Es seien V und V' zwei K -Vektorräume, welche man bezüglich gegebener Endomorphismen $f: V \rightarrow V$ und $f': V' \rightarrow V'$ als $K[T]$ -Moduln betrachte. Für Abbildungen $\varphi: V \rightarrow V'$ ist dann äquivalent:

- (i) φ ist ein Homomorphismus von $K[T]$ -Moduln.
- (ii) φ ist ein Homomorphismus von K -Vektorräumen, und das Diagramm

$$\begin{array}{ccc} V & \xrightarrow{f} & V, \\ \downarrow \varphi & & \downarrow \varphi \\ V' & \xrightarrow{f'} & V', \end{array} \quad \begin{array}{l} v \mapsto f(v) = T \cdot v, \\ v' \mapsto f'(v') = T \cdot v', \end{array}$$

ist kommutativ.

Eine entsprechende Äquivalenz gilt auch für Mono-, Epi- und Isomorphismen $\varphi: V \rightarrow V'$. Insbesondere stimmen in der Situation eines Isomorphismus φ in (i) bzw. (ii) die Minimalpolynome von f und f' überein. Gleiches gilt für die charakteristischen Polynome von f und f' .

Beweis. Jeder Homomorphismus von $K[T]$ -Moduln $\varphi: V \rightarrow V'$ ist $K[T]$ -linear und damit insbesondere K -linear, also ein Homomorphismus von K -Vektorräumen. Zudem bedeutet die Kommutativität des Diagramms in (ii) gerade die T -Linearität von φ , womit wir

$$\varphi(T \cdot v) = T \cdot \varphi(v) \quad \text{für alle } v \in V$$

meinen. Somit ergibt sich (ii) aus (i) und umgekehrt (i) auch aus (ii), wenn man benutzt, dass aus der T -Linearität die T^n -Linearität von φ für alle $n \in \mathbb{N}$ folgt, und aufgrund der Additivität sogar die $K[T]$ -Linearität von φ . Entsprechend schließt man im Falle von Mono-, Epi- oder Isomorphismen. $\square$

Satz 4. Es sei $f: V \rightarrow V$ ein Endomorphismus eines K -Vektorraums V der Dimension $r < \infty$. Man betrachte V als $K[T]$ -Modul bezüglich f . Dann ist äquivalent:

- (i) V ist f -zyklisch.
- (ii) V ist als $K[T]$ -Modul monogen.
- (iii) Es gibt ein normiertes Polynom $p \in K[T]$, so dass V als $K[T]$ -Modul isomorph zu $K[T]/(p)$ ist.

Sind diese Bedingungen erfüllt, so gilt weiter:

- (iv) Das Polynom p aus (iii) stimmt mit dem Minimalpolynom p_f sowie dem charakteristischen Polynom χ_f von f überein; insbesondere gilt $\text{grad } p = r$.

Beweis. Die Bedingungen (i) und (ii) sind aufgrund von Bemerkung 2 äquivalent. Wird weiter V gemäß Bedingung (ii) als $K[T]$ -Modul von dem Element $u \in V$ erzeugt, so betrachte man die $K[T]$ -lineare Abbildung

$$\varphi: K[T] \rightarrow V, \quad h \mapsto h \cdot u = h(f)(u);$$

φ ist surjektiv. Weiter ist $\ker \varphi$ ein $K[T]$ -Untermodul von $K[T]$, also ein Ideal, welches, da $K[T]$ ein Hauptidealring ist, von einem Polynom $p \in K[T]$ erzeugt wird. Der Homomorphiesatz für $K[T]$ -Moduln ergibt dann den in (iii) gewünschten Isomorphismus $K[T]/(p) \simeq V$. Wegen $\dim_K V < \infty$ und $\dim_K K[T] = \infty$ gilt $p \neq 0$, und man darf p als normiertes Polynom annehmen.

Ist andererseits Bedingung (iii) gegeben, so betrachte man die Restklassenabbildung $K[T] \longrightarrow K[T]/(p)$, $h \longmapsto \bar{h}$. Die Restklasse $\bar{1} \in K[T]/(p)$ erzeugt dann natürlich $K[T]/(p)$ als $K[T]$ -Modul, d. h. es folgt (ii).

Es bleibt nun noch Aussage (iv) zu begründen, wobei wir (i) bzw. (iii) als gegeben annehmen. Wir lesen zunächst aus (iii) ab, dass p das Minimalpolynom von

$$f': K[T]/(p) \longrightarrow K[T]/(p), \quad \bar{g} \longmapsto T \cdot \bar{g} = \overline{T} \cdot \bar{g},$$

ist; dabei bezeichne $\bar{g}$ für $g \in K[T]$ jeweils die Restklasse in $K[T]/(p)$, entsprechend für $\overline{T}$. Das Minimalpolynom von f' erzeugt nämlich gemäß 6.2/9 bzw. 6.2/10 das Ideal

$$\begin{aligned} & \{h \in K[T] ; h(f')(\bar{g}) = 0 \text{ für alle } \bar{g} \in K[T]/(p)\} \\ &= \{h \in K[T] ; h \cdot \bar{g} = 0 \text{ für alle } \bar{g} \in K[T]/(p)\} \\ &= \{h \in K[T] ; h \cdot \bar{1} = 0\} = \{h \in K[T] ; \bar{h} = 0\} \\ &= (p) \subset K[T]. \end{aligned}$$

Dann stimmt p aber gemäß Bemerkung 3 auch mit dem Minimalpolynom p_f des Endomorphismus $f: V \longrightarrow V$ überein.

Dass $p = p_f$ gilt, können wir in etwas ausführlicherer Argumentation auch auf dem Niveau des Endomorphismus f und des Vektorraums V einsehen: Es erzeugt p_f gemäß 6.2/9 bzw. 6.2/10 dasjenige Ideal in $K[T]$, welches aus allen Polynomen h mit $h(f) = 0$ besteht. Also gilt

$$(*) \quad (p_f) = \{h \in K[T] ; h(f)(v) = 0 \text{ für alle } v \in V\},$$

wobei es genügt, wenn man die Bedingung $h(f)(v) = 0$ für v aus einem K -Erzeugendensystem von V testet. Ist daher V als $K[T]$ -Modul monogen, etwa erzeugt von einem Element $u \in V$, so bilden die Elemente $u, f(u), f^2(u), \dots$ ein K -Erzeugendensystem von V , und die Beschreibung $(*)$ ist äquivalent zu

$$(**) \quad (p_f) = \{h \in K[T] ; h(f)(f^n(u)) = 0 \text{ für alle } n \in \mathbb{N}\}.$$

Nun zeigt aber die Rechnung

$$h(f)(f^n(u)) = h \cdot f^n(u) = h \cdot T^n \cdot u = T^n \cdot h \cdot u = T^n \cdot h(f)(u),$$

dass aus $h(f)(u) = 0$ bereits $h(f)(f^n(u)) = 0$ für alle $n \in \mathbb{N}$ folgt. Somit erhält man aus $(**)$

$$(p_f) = \{h \in K[T] ; h(f)(u) = 0\} = \{h \in K[T] ; h \cdot u = 0\} = (p)$$

und damit wegen der Normiertheit von p_f und p bereits $p_f = p$.

Insbesondere stimmt der Grad von p_f aufgrund des nachfolgenden Lemmas 5 mit der K -Vektorraumdimension von $K[T]/(p)$ überein, also mit der Dimension r von V . Da aber das charakteristische Polynom χ_f normiert vom Grad $\dim_K V$ ist und da aufgrund des Satzes von Cayley-Hamilton p_f ein Teiler von χ_f ist, vgl. 6.2/11, gilt bereits $\chi_f = p_f$. Die Gleichung $\chi_f = p$ lässt sich allerdings auch durch explizite Berechnung nachweisen, indem man die Multiplikation mit T auf $K[T]/(p)$ bezüglich der Basis $\bar{T}^0, \dots, \bar{T}^{r-1}$ durch eine Matrix beschreibt und deren charakteristisches Polynom ermittelt. $\square$

Schließlich bleibt noch die K -Vektorraumdimension von Quotienten des Typs $K[T]/(p)$ zu berechnen.

Lemma 5. *Es sei $p \in K[T]$ ein normiertes Polynom. Dann gilt*

$$\dim_K(K[T]/(p)) = \text{grad } p,$$

und die Restklassen $\bar{T}^0, \dots, \bar{T}^{s-1}$ mit $s = \text{grad } p$ bilden eine K -Basis von $K[T]/(p)$.

Beweis. Es genügt zu zeigen, dass für $s = \text{grad } p$ die Restklassen $\bar{T}^0, \dots, \bar{T}^{s-1}$ eine K -Basis von $K[T]/(p)$ bilden. Hierzu betrachte man ein beliebiges Element $\bar{h} \in K[T]/(p)$ und wähle ein Urbild $h \in K[T]$. Division mit Rest ergibt dann eine Zerlegung $h = qp + c$ mit Polynomen $q, c \in K[T]$, wobei c vom Grad $< s$ ist, etwa $c = \sum_{i=0}^{s-1} c_i T^i$ mit Koeffizienten $c_i \in K$. Es folgt

$$\bar{h} = \bar{c} = \sum_{i=0}^{s-1} c_i \bar{T}^i,$$

d. h. $K[T]/(p)$ wird von den Restklassen $\bar{T}^0, \dots, \bar{T}^{s-1}$ als K -Vektorraum erzeugt. Diese Elemente sind aber auch linear unabhängig über K . Hat man nämlich $\sum_{i=0}^{s-1} c_i \bar{T}^i = 0$ mit gewissen Koeffizienten $c_i \in K$, so bedeutet dies $\sum_{i=0}^{s-1} c_i T^i \in (p)$ und damit $c_i = 0$ für alle $i = 0, \dots, s-1$, da alle von Null verschiedenen Elemente in (p) einen Grad $\geq \text{grad } p = s$ haben. $\square$

Wir wollen nun den Struktursatz 6.4/3 umformulieren zu einem Struktursatz für $K[T]$ -Moduln oder, genauer, für die Situation eines endlich-dimensionalen K -Vektorraums V mit einem Endomorphismus $f: V \rightarrow V$.

Theorem 6. *Es sei V ein endlich-dimensionaler K -Vektorraum mit einem Endomorphismus $f: V \rightarrow V$. Dann existieren paarweise verschiedene normierte (und damit nicht-assozierte) Primpolynome $p_1, \dots, p_r \in K[T]$ sowie natürliche Zahlen $1 \leq n(i, 1) \leq \dots \leq n(i, s_i)$ und f -zyklische K -Untervektorräume $V_{ij} \subset V$, $i = 1, \dots, r$, $j = 1, \dots, s_i$, mit*

$$V = \bigoplus_{i=1}^r \bigoplus_{j=1}^{s_i} V_{ij}, \quad V_{ij} \simeq K[T]/(p_i^{n(i,j)}),$$

im Sinne von $K[T]$ -Moduln. Insbesondere ist $p_i^{n(i,j)}$ das Minimal- bzw. charakteristische Polynom von $f|_{V_{ij}}$; vgl. Satz 4.

In der vorstehenden Zerlegung sind die Primpolynome $p_1, \dots, p_r$, die Zahlen $n(i, j)$ und die Vektorräume $V_i := \bigoplus_{j=1}^{s_i} V_{ij}$ für $i = 1, \dots, r$ eindeutig bestimmt, nicht notwendig jedoch die Unterräume V_{ij} selbst; es gilt

$$V_i = \bigcup_{n \in \mathbb{N}} \ker p_i^n(f).$$

Weiter ist $p_f = p_1^{n(1,s_1)} \cdot \dots \cdot p_r^{n(r,s_r)}$ das Minimalpolynom von f , und man hat

$$\dim_K V_{ij} = n(i, j) \cdot \text{grad } p_i,$$

insbesondere also

$$\dim_K V = \sum_{i=1}^r \sum_{j=1}^{s_i} n(i, j) \cdot \text{grad } p_i.$$

Beweis. Die Aussagen zur Existenz und Eindeutigkeit einer Zerlegung von V in f -zyklische Unterräume V_{ij} der behaupteten Art ergeben sich unmittelbar aus 6.4/3 und Satz 4. Dabei ist $V_i = \bigoplus_{j=1}^{s_i} V_{ij}$ in der Sprache der $K[T]$ -Moduln der Untermodul der p_i -Torsion, und der freie Anteil aus 6.4/3 entfällt, da V ein $K[T]$ -Torsionsmodul ist.

Gemäß Satz 4 ist $p_i^{n(i,j)}$ das Minimalpolynom der Einschränkung von f auf V_{ij} . Dann erkennt man leicht, dass das kleinste gemeinsame Vielfache aller dieser Polynome, also

$$p_1^{n(1,s_1)} \cdot \dots \cdot p_r^{n(r,s_r)}$$

das Minimalpolynom von f auf V ist. Die Dimensionsformeln schließlich ergeben sich unter Benutzung von 1.6/6 aus Lemma 5. $\square$

Korollar 7. *Es sei $f: V \rightarrow V$ ein Endomorphismus eines endlich-dimensionalen K -Vektorraums V . Dann ist äquivalent:*

- (i) V ist f -unzerlegbar.
- (ii) V ist f -zyklisch, und $p_f = \chi_f$ ist Potenz eines Primpolynoms in $K[T]$.

Insbesondere sind die f -zyklischen K -Untervektorräume $V_{ij} \subset V$ aus Theorem 6 f -unzerlegbar.

Beweis. Man zerlege V gemäß Theorem 6 in eine direkte Summe f -zyklischer Unterräume, deren Minimalpolynom jeweils Potenz eines Primpolynoms aus $K[T]$ ist. Ist V dann f -unzerlegbar, so kann diese Zerlegung nur aus einem Summanden bestehen, und es folgt, dass V f -zyklisch ist.

Ist andererseits V f -zyklisch mit einem Minimalpolynom p_f , welches Potenz eines Primpolynoms in $K[T]$ ist, so gilt $V \simeq K[T]/(p_f)$ im Sinne von $K[T]$ -Moduln gemäß Satz 4, und die Eindeutigkeit der Zerlegung in Theorem 6 zeigt, dass V f -unzerlegbar ist. $\square$

Schließlich wollen wir das Zerlegungstheorem 6 noch umformulieren zu einer Aussage über Normalformen von quadratischen Matrizen. Wir beginnen mit einer trivialen Beobachtung, wobei wir folgende Schreibweise benutzen: Für Matrizen $A_i \in K^{n_i \times n_i}$, $i = 1, \dots, r$, bezeichne $\text{Diag}(A_1, \dots, A_r) \in K^{n \times n}$, $n = \sum_{i=1}^r n_i$, die "Diagonalmatrix" mit den Einträgen $A_1, \dots, A_r$, also

$$\text{Diag}(A_1, \dots, A_r) = \begin{pmatrix} A_1 & & & 0 \\ & A_2 & & \\ & & \ddots & \\ 0 & & & \ddots & \\ & & & & A_r \end{pmatrix}$$

Gilt $A_i = (\lambda_i) \in K^{1 \times 1}$ für $i = 1, \dots, r$, so ist $\text{Diag}(A_1, \dots, A_r)$ eine Diagonalmatrix im echten Sinne, und wir schreiben auch $\text{Diag}(\lambda_1, \dots, \lambda_r)$ anstelle von $\text{Diag}(A_1, \dots, A_r)$.

Lemma 8. *Es sei $f: V \longrightarrow V$ ein Endomorphismus eines endlich-dimensionalen K -Vektorraums V mit einer Zerlegung $V = \bigoplus_{i=1}^r V_i$ in f -invariante Unterräume. Es sei X_i jeweils eine Basis von V_i , sowie X die Basis von V , die sich aus den Basen $X_1, \dots, X_r$ zusammensetzt. Mit $f_i = f|_{V_i}$ gilt dann für die zu f gehörige Matrix sowie das charakteristische bzw. Minimalpolynom:*

- (i) $A_{f,X,X} = \text{Diag}(A_{f_1,X_1,X_1}, \dots, A_{f_r,X_r,X_r})$.
- (ii) $\chi_f = \chi_{f_1} \cdot \dots \cdot \chi_{f_r}$.
- (iii) $p_f = \text{kgV}(p_{f_1}, \dots, p_{f_r})$.

Lemma 9. *Es sei $f: V \longrightarrow V$ ein Endomorphismus eines K -Vektorraums V mit einer Basis $X = (x_1, \dots, x_n)$. Gilt dann $A_{f,X,X} = \text{Diag}(A_1, \dots, A_r)$ mit Matrizen $A_i \in K^{n_i \times n_i}$ (wobei $n = \sum_{i=1}^r n_i$), so sind für $m_i = \sum_{j=1}^i n_j$ die Unterräume*

$$V_i = \sum_{j=m_{i-1}+1}^{m_i} K \cdot x_j, \quad i = 1, \dots, r,$$

f -invariant, und es gilt $V = \bigoplus_{i=1}^r V_i$.

Beweis zu den Lemmata 8 und 9. Aussage (i) in Lemma 8 sowie Lemma 9 sind unmittelbar klar aufgrund der in 3.1/3 definierten Korrespondenz zwischen linearen Abbildungen und beschreibenden Matrizen. Weiter folgt Aussage (ii) in Lemma 8 aufgrund des Beispiels (2) am Ende von Abschnitt 4.3. Aussage (iii) schließlich ist gültig, da einerseits das kleinste gemeinsame Vielfache aller p_{f_i} den Endomorphismus f annulliert, sowie andererseits p_f ein Vielfaches eines jeden p_{f_i} sein muss. $\square$

Zu einem normierten Polynom $p = T^n + \sum_{i=0}^{n-1} c_i T^i \in K[T]$ vom Grad n nennt man die Matrix

$$A(p) = \begin{pmatrix} 0 & & & & -c_0 \\ 1 & 0 & & & -c_1 \\ & 1 & . & & -c_2 \\ & & . & . & \dots \\ & & & . & \dots \\ & & & . & 0 & -c_{n-2} \\ & & & & 1 & -c_{n-1} \end{pmatrix}$$

die Begleitmatrix zu p . Insbesondere gilt $A(p) = (-c_0)$ für $n = 1$. Auch den Sonderfall $n = 0$ wollen wir nicht ausschließen. Es ist $A(p)$ dann die leere Matrix.

Lemma 10. *Es sei $f: V \rightarrow V$ ein Endomorphismus eines K -Vektorraums der Dimension $n < \infty$ und $p \in K[T]$ ein normiertes Polynom vom Grad n . Dann ist äquivalent:*

- (i) V ist f -zyklisch mit Minimalpolynom $p_f = p$.
- (ii) Es existiert eine Basis X von V , so dass die Matrix $A_{f,X,X}$ die Begleitmatrix zu p ist.

Beweis. Für $V = 0$ ist die Aussage trivial, so dass wir $V \neq 0$ annehmen dürfen. Sei V zunächst f -zyklisch, gelte also $V \simeq K[T]/(p)$ im Sinne von $K[T]$ -Moduln, wobei $p = p_f$ das Minimalpolynom von f ist; vgl. Satz 4. Dann ergeben die Restklassen $\overline{T}^0, \dots, \overline{T}^{n-1}$ gemäß Lemma 5 eine K -Basis von $K[T]/(p)$, bzw. mit anderen Worten, es existiert ein Element $u \in V$ (nämlich dasjenige Element, welches unter dem Isomorphismus $V \simeq K[T]/(p)$ zu der Restklasse $\overline{1} \in K[T]/(p)$ korrespondiert), so dass $f^0(u), f^1(u), \dots, f^{n-1}(u)$ eine K -Basis X von V bilden. Für $p = T^n + c_{n-1}T^{n-1} + \dots + c_0$ folgt dann

$$f(f^{n-1}(u)) = f^n(u) = -\sum_{i=0}^{n-1} c_i f^i(u),$$

und man sieht, dass $A_{f,X,X}$ die Begleitmatrix zu p ist.

Sei nun umgekehrt $X = (x_1, \dots, x_n)$ eine Basis von V , so dass $A_{f,X,X}$ die Begleitmatrix zu einem normierten Polynom $p \in K[T]$ vom Grade n ist. Setzen wir dann $u = x_1$, so gilt $x_i = f^{i-1}(u)$ für $i = 1, \dots, n$, und es folgt, dass V f -zyklisch ist. Weiter liest man aus der Matrix $A_{f,X,X}$ die Beziehungen $p(f)(u) = 0$ bzw. $p \cdot u = 0$ im Sinne von V als $K[T]$ -Modul ab. Da man aber $V = K[T] \cdot u$ hat, ergibt sich auch $p \cdot V = 0$, so dass das Minimalpolynom p_f ein Teiler von p ist. Da p_f gemäß Satz 4 mit dem charakteristischen Polynom χ_f übereinstimmt und damit ebenso wie p den Grad n hat, erhält man damit bereits $\chi_f = p_f = p$. $\square$

Nun sind wir in der Lage, aus Theorem 6 die Existenz und Eindeutigkeit der so genannten *allgemeinen Normalform* für Matrizen herzuleiten.

Theorem 11. *Jede Matrix $A \in K^{n \times n}$ ist ähnlich zu einer Matrix der Form $\text{Diag}(A_1, \dots, A_r)$ mit quadratischen Matrizen A_i , wobei A_i jeweils die Begleitmatrix zu einer Potenz q_i eines normierten Primpolynoms aus $K[T]$ ist. Die*

Matrizen $A_1, \dots, A_r$ sind bis auf die Reihenfolge eindeutig durch A bestimmt. Es gilt $p_A = \text{kgV}(q_1, \dots, q_r)$ für das Minimalpolynom von A , sowie $\chi_A = q_1 \cdot \dots \cdot q_r$ für das charakteristische Polynom von A .

Beweis. Interpretieren wir A als Matrix eines Endomorphismus $f: V \longrightarrow V$ eines endlich-dimensionalen K -Vektorraums V , so zerfällt V nach Theorem 6 in eine direkte Summe f -zyklischer Unterräume $V_1, \dots, V_r$, etwa $V_i \simeq K[T]/(q_i)$, wobei q_i Potenz eines normierten Primpolynoms ist. Nach Lemma 8 zeigt dies, dass A zu einer Diagonalmatrix $\text{Diag}(A_1, \dots, A_r)$ ähnlich ist, wobei A_i jeweils den Endomorphismus $f_i = f|_{V_i}$ bezüglich einer geeigneten Basis von V_i beschreibt. Aus Satz 4 ergibt sich, dass q_i das Minimalpolynom bzw. das charakteristische Polynom von f_i ist, und aus Lemma 10, dass wir A_i als Begleitmatrix zu q_i annehmen dürfen.

Ist umgekehrt A zu einer Matrix der Form $\text{Diag}(A_1, \dots, A_r)$ ähnlich, wobei A_i jeweils die Begleitmatrix zu einer Potenz q_i eines normierten Primpolynoms ist, so korrespondiert hierzu eine Zerlegung $V \simeq \bigoplus_{i=1}^r K[T]/(q_i)$; vgl. Lemmata 9 und 10 in Verbindung mit Satz 4. Die Eindeutigkeitsaussage in Theorem 6 impliziert dann die Eindeutigkeit der Matrizen $A_1, \dots, A_r$, die ja durch ihre Minimalpolynome bzw. charakteristischen Polynome eindeutig festgelegt sind.

Die Aussagen über das Minimalpolynom p_A sowie das charakteristische Polynom χ_A wurden bereits in Lemma 8 hergeleitet. $\square$

Als Folgerung können wir nochmals den Satz von Cayley-Hamilton ablesen, sogar in einer etwas verbesserten Version:

Korollar 12 (Cayley-Hamilton). *Es sei $f: V \longrightarrow V$ ein Endomorphismus eines K -Vektorraums der Dimension $n < \infty$ mit Minimalpolynom p_f und charakteristischem Polynom χ_f . Dann ist p_f ein Teiler von χ_f , und jeder Primfaktor, der in der Primfaktorzerlegung von χ_f vorkommt, kommt auch in der Primfaktorzerlegung von p_f vor, im Allgemeinen allerdings mit geringerer Vielfachheit.*

Eine Diagonalmatrix $A = \text{Diag}(\lambda_1, \dots, \lambda_n) \in K^{n \times n}$ ist bereits von allgemeiner Normalform, da die Matrix $A_i = (\lambda_i) \in K^{1 \times 1}$ als Begleitmatrix des Polynoms $T - \lambda_i \in K[T]$ aufgefasst werden kann. Aus Theorem 11 ergibt sich daher folgendes Diagonalisierbarkeitskriterium:

Korollar 13. *Folgende Aussagen sind für einen Endomorphismus f eines endlich-dimensionalen K -Vektorraums (bzw. eine Matrix $A \in K^{n \times n}$) äquivalent:*

- (i) *f (bzw. A) ist diagonalisierbar.*
- (ii) *Das Minimalpolynom p_f (bzw. p_A) zerfällt in ein Produkt $\prod_{i=1}^r (T - \lambda_i)$ mit paarweise verschiedenen Nullstellen $\lambda_1, \dots, \lambda_r \in K$.*

Es soll als Nächstes die so genannte *Jordansche Normalform* für Matrizen hergeleitet werden. Diese Normalform kann nur in den Fällen konstruiert werden, in denen das charakteristische bzw. Minimalpolynom vollständig in lineare

Faktoren zerfällt, also beispielsweise dann, wenn der Körper K algebraisch abgeschlossen ist. Für $\lambda \in K$ bezeichnen wir die Matrix

$$J(\lambda, n) = \begin{pmatrix} \lambda & & & & 0 \\ 1 & \lambda & & & \\ & 1 & \lambda & & \\ & & \cdot & \cdot & \\ & & & \cdot & \cdot \\ 0 & & & & 1 & \lambda \end{pmatrix} \in K^{n \times n}$$

als Jordankästchen der Länge n zum Eigenwert λ .

Lemma 14. Für $A \in K^{n \times n}$ und $\lambda \in K$ ist äquivalent:

- (i) A ist ähnlich zu $J(\lambda, n)$.
- (ii) A ist ähnlich zur Begleitmatrix $A(q)$ des Polynoms $q = (T - \lambda)^n \in K[T]$.

Beweis. Man interpretiere A als beschreibende Matrix eines Endomorphismus $f: V \rightarrow V$, wobei V ein n -dimensionaler K -Vektorraum sei. Ist dann A ähnlich zu $J(\lambda, n)$, so existiert eine K -Basis $x_1, \dots, x_n$ von V mit

$$(*) \quad f(x_i) = \lambda x_i + x_{i+1}, \quad i = 1, \dots, n,$$

wenn wir $x_{n+1} = 0$ setzen. Mit Induktion ergibt sich hieraus

$$x_1, \dots, x_i \in \langle f^0(x_1), \dots, f^{i-1}(x_1) \rangle, \quad i = 1, \dots, n;$$

denn für $i = 1$ ist dies trivial, und unter Benutzung der Induktionsvoraussetzung zum Index i schließt man

$$\begin{aligned} x_{i+1} &= f(x_i) - \lambda x_i \\ &\in \langle f^1(x_1), \dots, f^i(x_1) \rangle + \langle f^0(x_1), \dots, f^{i-1}(x_1) \rangle \\ &= \langle f^0(x_1), \dots, f^i(x_1) \rangle. \end{aligned}$$

Insbesondere gilt daher

$$V = \langle f^0(x_1), \dots, f^{n-1}(x_1) \rangle,$$

d. h. V ist f -zyklisch und wird als $K[T]$ -Modul von x_1 erzeugt.

Man betrachte nun den surjektiven $K[T]$ -Modulhomomorphismus

$$\varphi: K[T] \rightarrow V, \quad h \mapsto h \cdot x_1 = h(f)(x_1).$$

Aus $(*)$ ergibt sich

$$(f - \lambda \text{id})(x_i) = x_{i+1}, \quad i = 1, \dots, n,$$

also

$$(f - \lambda \text{id})^i(x_1) = \begin{cases} x_{i+1} \neq 0 & \text{für } i = 0, \dots, n-1 \\ 0 & \text{für } i = n \end{cases}.$$

Der Kern von φ enthält daher das Polynom $(T - \lambda)^n$, nicht aber $(T - \lambda)^{n-1}$. Als Hauptideal wird $\ker \varphi$ von einem normierten Polynom $q \in K[T]$ erzeugt. Dieses teilt $(T - \lambda)^n$, nicht aber $(T - \lambda)^{n-1}$, und stimmt folglich mit $(T - \lambda)^n$ überein. Aufgrund des Isomorphiesatzes für Moduln induziert φ dann einen Isomorphismus $K[T]/((T - \lambda)^n) \xrightarrow{\sim} V$ im Sinne von $K[T]$ -Moduln, und man erkennt V gemäß Satz 4 als f -zyklisch mit Minimalpolynom $p_f = (T - \lambda)^n$. Schließlich folgt mit Lemma 10, dass A zur Begleitmatrix des Minimalpolynoms $(T - \lambda)^n$ ähnlich ist. Die Implikation von (i) nach (ii) ist daher bewiesen.

Zum Nachweis der umgekehrten Implikation sei A nun ähnlich zur Begleitmatrix des Polynoms $q = (T - \lambda)^n \in K[T]$. Dann ist V nach Lemma 10 f -zyklisch mit Minimalpolynom q zu f , und es gilt $V \simeq K[T]/(q)$ im Sinne von $K[T]$ -Moduln. Wir wollen zeigen, dass die Restklassen

$$(\overline{T} - \lambda)^0, (\overline{T} - \lambda)^1, \dots, (\overline{T} - \lambda)^{n-1}$$

eine K -Basis von $K[T]/(q)$ bilden. Gemäß Lemma 5 ist bereits bekannt, dass dies für die Potenzen $\overline{T}^0, \dots, \overline{T}^{n-1}$ gilt. Es genügt daher, wenn wir induktiv

$$\overline{T}^0, \dots, \overline{T}^{i-1} \in \langle (\overline{T} - \lambda)^0, \dots, (\overline{T} - \lambda)^{i-1} \rangle, \quad i = 1, \dots, n,$$

zeigen. Für $i = 1$ ist dies trivial, und unter Verwendung der Induktionsvoraussetzung zum Index i können wir wie folgt schließen:

$$\begin{aligned} \overline{T}^i &= (\overline{T} - \lambda) \cdot \overline{T}^{i-1} + \lambda \overline{T}^{i-1} \\ &\in \langle (\overline{T} - \lambda)^1, \dots, (\overline{T} - \lambda)^i \rangle + \langle (\overline{T} - \lambda)^0, \dots, (\overline{T} - \lambda)^{i-1} \rangle \\ &= \langle (\overline{T} - \lambda)^0, \dots, (\overline{T} - \lambda)^i \rangle \end{aligned}$$

Folglich bilden die Potenzen $(\overline{T} - \lambda)^{i-1}$ für $i = 1, \dots, n$ eine K -Basis von $K[T]/(q)$.

Wir betrachten nun die Gleichungen

$$\overline{T} \cdot (\overline{T} - \lambda)^{i-1} = \lambda \cdot (\overline{T} - \lambda)^{i-1} + (\overline{T} - \lambda)^i, \quad i = 1, \dots, n,$$

wobei $(\overline{T} - \lambda)^n = 0$ gilt. Sie besagen, dass der im Sinne von Bemerkung 3 zu f korrespondierende Endomorphismus

$$K[T]/(q) \longrightarrow K[T]/(q), \quad \overline{h} \longmapsto \overline{T} \cdot \overline{h},$$

bezüglich obiger Basis durch das Jordankästchen $J(\lambda, n)$ beschrieben wird. Dann gibt es aber auch in V eine Basis, bezüglich der f durch $J(\lambda, n)$ beschrieben wird, und die Implikation von (ii) nach (i) ist bewiesen. $\square$

Nun lassen sich Existenz und Eindeutigkeit der Jordanschen Normalform leicht aus Theorem 11 folgern.

Theorem 15. *Es sei $A \in K^{n \times n}$ eine Matrix, deren Minimal- bzw. charakteristisches Polynom vollständig in Linearfaktoren zerfällt. Dann ist A ähnlich zu einer so genannten Jordan-Matrix*

$$\text{Diag}(J(\lambda_1, n_1), \dots, J(\lambda_r, n_r)),$$

deren “Einträge” Jordankästchen sind. Die Elemente λ_i, n_i sind, abgesehen von der Reihenfolge, eindeutig durch A bestimmt, wobei die λ_i (unter eventueller Mehrfachaufzählung) gerade die Eigenwerte von A durchlaufen. Wählt man einen n -dimensionalen K -Vektorraum V und realisiert A als Matrix eines Endomorphismus $f: V \rightarrow V$, so ist

$$V \simeq \bigoplus_{i=1}^r K[T]/((T - \lambda_i)^{n_i})$$

gerade die Zerlegung aus 6.4/3 bzw. Theorem 6.

Beweis. Man benutze Lemma 14, um von der allgemeinen Normalform aus Theorem 11 von A zur Jordanschen Normalform zu gelangen bzw. umgekehrt von der Jordanschen Normalform zur allgemeinen Normalform. $\square$

Wir sagen, eine Matrix $A \in K^{n \times n}$ sei *trigonalisierbar*, wenn sie ähnlich zu einer Matrix der Form $B = (\beta_{ij}) \in K^{n \times n}$ ist mit $\beta_{ij} = 0$ für $i < j$; man nennt B eine (untere) *Dreiecksmatrix*. Aus der Existenz der Jordanschen Normalform kann man insbesondere ein Kriterium für Trigonalisierbarkeit ableiten.

Korollar 16. *Eine Matrix $A \in K^{n \times n}$ ist genau dann trigonalisierbar, wenn das Minimal- bzw. charakteristische Polynom von A vollständig in Linearfaktoren zerfällt.*

Beweis. Sei A trigonalisierbar, also ähnlich zu einer Matrix $B = (\beta_{ij}) \in K^{n \times n}$ mit $\beta_{ij} = 0$ für $i < j$. Dann gilt $\chi_A = \chi_B = \prod_{i=1}^n (T - \beta_{ii})$; insbesondere zerfällt χ_A und damit nach Korollar 12 auch das Minimalpolynom p_A vollständig in Linearfaktoren. Ist umgekehrt letztere Bedingung gegeben, so zeigt die Existenz der Jordanschen Normalform, dass A trigonalisierbar ist. $\square$

Wir wollen als Nächstes ein erstes (recht grobes) praktisches Verfahren zur expliziten Berechnung der Jordanschen Normalform einer Matrix angeben. Man betrachte also eine Matrix $A \in K^{n \times n}$, deren charakteristisches Polynom vollständig in Linearfaktoren zerfällt, etwa

$$\chi_A = (T - \lambda_1)^{n_1} \cdot \dots \cdot (T - \lambda_r)^{n_r};$$

im Unterschied zu Theorem 15 setzen wir hierbei voraus, dass die Eigenwerte $\lambda_1, \dots, \lambda_r$ paarweise verschieden sind. Die Jordansche Normalform J zu A ist dann eine Matrix, auf deren Diagonalen Jordankästchen des Typs $J(\lambda, s)$ stehen.

Für $i = 1, \dots, r$ und $j = 1, \dots, n_i$ bezeichne $k_{i,j}$ diejenige Anzahl, mit der das Jordankästchen $J(\lambda_i, j)$ in J vorkommt. Es gilt

$$k_{i,1} + 2k_{i,2} + \dots + n_i k_{i,n_i} = n_i,$$

da die Vielfachheit des Eigenwertes λ_i gerade n_i ist, das Element λ_i also genau n_i -mal auf der Diagonalen von J vorkommt.

Da die Matrizen A und J zueinander ähnlich sind, sind auch die Matrizen $A - \lambda E$ und $J - \lambda E$ sowie die Potenzen $(A - \lambda E)^\ell$ und $(J - \lambda E)^\ell$ mit $\ell \in \mathbb{N}$ zueinander ähnlich; dabei sei $\lambda \in K$ und $E \in K^{n \times n}$ die Einheitsmatrix. Insbesondere folgt dann mit 3.4/6

$$\operatorname{rg}(A - \lambda E)^\ell = \operatorname{rg}(J - \lambda E)^\ell \quad \text{für } \ell \in \mathbb{N},$$

und es ist $(J - \lambda E)^\ell$ wiederum eine "Diagonalmatrix", gebildet aus Kästchen, nämlich aus den ℓ -ten Potenzen der Jordankästchen von $J - \lambda E$. Nun berechnet sich für ein Jordankästchen des Typs $J(\lambda, m)$ der Rang einer ℓ -ten Potenz zu

$$\operatorname{rg} J(\lambda, m)^\ell = \begin{cases} \max\{0, m - \ell\} & \text{für } \lambda = 0 \\ m & \text{für } \lambda \neq 0 \end{cases}.$$

Daher bestehen folgende Gleichungen:

$$\begin{aligned} \operatorname{rg}(A - \lambda_i E)^{n_i} &= \operatorname{rg}(J - \lambda_i E)^{n_i} = n - n_i \\ \operatorname{rg}(A - \lambda_i E)^{n_i-1} &= \operatorname{rg}(J - \lambda_i E)^{n_i-1} = n - n_i + k_{i,n_i} \\ \operatorname{rg}(A - \lambda_i E)^{n_i-2} &= \operatorname{rg}(J - \lambda_i E)^{n_i-2} = n - n_i + 2k_{i,n_i} + k_{i,n_i-1} \\ &\dots \\ \operatorname{rg}(A - \lambda_i E)^1 &= \operatorname{rg}(J - \lambda_i E)^1 = n - n_i + (n_i - 1)k_{i,n_i} + \dots + k_{i,2} \end{aligned}$$

Ermittelt man also die Ränge der Potenzen von $A - \lambda_i E$, so lassen sich die Zahlen $k_{i,j}$, $j = n_i, n_i - 1, \dots, 2$, der Reihe nach berechnen. Weiter gilt aufgrund obiger Gleichung

$$k_{i,1} = n_i - 2k_{i,2} - \dots - n_i k_{i,n_i},$$

so dass man auch $k_{i,1}$ berechnen kann. Sind aber alle Zahlen $k_{i,j}$ bekannt, so hat man insgesamt die Jordansche Normalform J von A bestimmt.

Als Beispiel wollen wir die Jordansche Normalform der Matrix

$$A = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 2 & 0 & 0 \\ -1 & 1 & 2 & 1 \\ -1 & 1 & 0 & 3 \end{pmatrix} \in \mathbb{R}^{4 \times 4}$$

bestimmen. Es gilt $\chi_A = (T - 2)^4$, also $r = 1$ und $n_1 = 4$ in obiger Notation. Wir haben

$$\operatorname{rg}(A - 2E) = \operatorname{rg} \begin{pmatrix} -1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 1 \\ -1 & 1 & 0 & 1 \end{pmatrix} = 1,$$

sowie $\operatorname{rg}(A - 2E)^2 = \operatorname{rg}(0) = 0$ und damit auch $\operatorname{rg}(A - 2E)^s = 0$ für $s \geq 2$. Daher bestehen die Gleichungen

$$\begin{aligned} 0 &= \operatorname{rg}(A - 2E)^4 = n - n_1 = 0, \\ 0 &= \operatorname{rg}(A - 2E)^3 = k_{1,4}, \\ 0 &= \operatorname{rg}(A - 2E)^2 = 2k_{1,4} + k_{1,3}, \\ 1 &= \operatorname{rg}(A - 2E)^1 = 3k_{1,4} + 2k_{1,3} + k_{1,2}, \end{aligned}$$

und dies ergibt $k_{1,4} = k_{1,3} = 0$, $k_{1,2} = 1$, sowie

$$k_{1,1} = 4 - 2k_{1,2} = 2.$$

Folglich ist $J = \operatorname{Diag}(J(2, 1), J(2, 1), J(2, 2))$, also

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

die Jordansche Normalform zu A .

Wir behandeln abschließend noch ein weiteres viel effektiveres Verfahren, mit dessen Hilfe man neben der Jordanschen auch die allgemeine Normalform von Matrizen ermitteln kann. Das Vorgehen stellt sich wie folgt dar: Ausgehend von einer Matrix $A \in K^{n \times n}$ interpretieren wir diese als Endomorphismus f des n -dimensionalen K -Vektorraums $V = K^n$, so dass wir V unter f als $K[T]$ -Modul auffassen können. Sodann konstruieren wir zu A eine kanonische endliche Präsentation von V als $K[T]$ -Modul und bestimmen die zugehörigen Elementarteiler. Wir gelangen auf diese Weise zu der in 6.4/2 angegebenen Zerlegung von V in monogene $K[T]$ -Untermodule, aus der sich alle weiteren Zerlegungen und insbesondere auch die Normalformen von A ergeben.

Lemma 17. *Für eine Matrix $A \in K^{n \times n}$ betrachte man K^n als $K[T]$ -Modul unter dem durch $x \mapsto A \cdot x$ gegebenen Endomorphismus. Es sei*

$$\varphi: K[T]^n \longrightarrow K^n, \quad (p_1, \dots, p_n) \longmapsto \sum_{i=1}^n p_i(A) \cdot e_i,$$

diejenige $K[T]$ -lineare Abbildung, die die kanonische $K[T]$ -Basis von $K[T]^n$ auf die kanonische K -Basis $e_1, \dots, e_n$ von K^n abbildet; φ ist surjektiv. Weiter sei

$$\psi: K[T]^n \longrightarrow K[T]^n, \quad P \longmapsto (TE - A) \cdot P,$$

die durch die Matrix $TE - A \in K[T]^{n \times n}$ gegebene $K[T]$ -lineare Abbildung; $E \in K^{n \times n}$ sei die Einheitsmatrix. Dann ist die Sequenz

$$K[T]^n \xrightarrow{\psi} K[T]^n \xrightarrow{\varphi} K^n \longrightarrow 0$$

eine endliche Präsentation von K^n als $K[T]$ -Modul.

Beweis. Bezeichnet e_i in $K[T]^n$, wie auch in K^n , den i -ten Einheitsvektor, so gilt für $i = 1, \dots, n$

$$\psi(e_i) = (TE - A)e_i = Te_i - Ae_i, \quad \varphi(Te_i) = Ae_i, \quad \varphi(Ae_i) = Ae_i,$$

und damit

$$\varphi \circ \psi(e_i) = \varphi(Te_i - Ae_i) = Ae_i - Ae_i = 0.$$

Es folgt $\varphi \circ \psi = 0$, also im $\psi \subset \ker \varphi$. Aufgrund des Homomorphiesatzes zerlegt sich φ dann in eine Komposition

$$\varphi: K[T]^n \xrightarrow{\pi} K[T]^n / \operatorname{im} \psi \xrightarrow{\bar{\varphi}} K^n$$

surjektiver Abbildungen, und wir behaupten, dass $\bar{\varphi}$ sogar bijektiv ist. Da $K[T]^n$ als K -Vektorraum von den Elementen $T^\nu e_i$ mit $i = 1, \dots, n$ und $\nu \in \mathbb{N}$ erzeugt wird, erzeugen entsprechend die Bilder $\pi(T^\nu e_i)$ den Quotienten $K[T]^n / \operatorname{im} \psi$ als K -Vektorraum. Nun gilt aber für alle $\nu \in \mathbb{N}$

$$T^{\nu+1}e_j - T^\nu Ae_j = \psi(T^\nu e_j) \in \operatorname{im} \psi, \quad j = 1, \dots, n,$$

und deshalb

$$T^{\nu+1}e_j \in T^\nu \sum_{i=1}^n Ke_i + \operatorname{im} \psi, \quad j = 1, \dots, n.$$

Per Induktion folgt hieraus für $\nu \in \mathbb{N}$

$$T^\nu e_j \in \sum_{i=1}^n Ke_i + \operatorname{im} \psi, \quad j = 1, \dots, n,$$

und damit

$$K[T]^n = \sum_{i=1}^n Ke_i + \operatorname{im} \psi.$$

Dies bedeutet, dass $K[T]^n / \operatorname{im} \psi$ als K -Vektorraum von n Elementen erzeugt wird und damit über K eine Dimension $\leq n$ besitzt. Aufgrund der Dimensionsformel 2.1/10 ist $\bar{\varphi}$ dann notwendigerweise injektiv und damit bijektiv, wie behauptet. Dann gilt aber $\operatorname{im} \psi = \ker \pi = \ker \varphi$, und es folgt, dass die angegebene Sequenz eine endliche Präsentation von K^n als $K[T]$ -Modul darstellt. $\square$

Theorem 18. *Es sei $f: V \longrightarrow V$ ein Endomorphismus eines endlich-dimensionalen K -Vektorraums V . Sei $A = A_{f,X,X} \in K^{n \times n}$ die Matrix, welche f bezüglich einer gegebenen Basis X von V beschreibt, und seien $\alpha_1, \dots, \alpha_s \in K[T]$ mit $\alpha_j | \alpha_{j+1}$ diejenigen Elementarteiler der Matrix $TE - A \in K[T]^{n \times n}$, die nicht invertierbar sind; wir nehmen die α_j als normierte Polynome an. Weiter betrachten wir die Primfaktorzerlegungen $\alpha_j = p_1^{n(1,j)} \dots p_r^{n(r,j)}$, $j = 1, \dots, s$, der α_j mit paarweise verschiedenen normierten Primpolynomen $p_1, \dots, p_r \in K[T]$ und Exponenten $n(i, j) \geq 0$. Dann gilt:*

(i) Fasst man V als $K[T]$ -Modul unter f auf, so folgt

$$V \simeq \bigoplus_{j=1}^s K[T]/\alpha_j K[T] \simeq \bigoplus_{i=1}^r \bigoplus_{j=1}^s K[T]/p_i^{n(i,j)} K[T],$$

und die letztere Zerlegung gibt Anlass zu einer Zerlegung von V in f -zyklische Untervektorräume gemäß Theorem 6.

(ii) Die Kästchen der allgemeinen Normalform von A sind die Begleitmatrizen zu den Primpotenzen $p_i^{n(i,j)}$ mit $n(i,j) > 0$, $i = 1, \dots, r$, $j = 1, \dots, s$. Falls alle p_i linear sind, erhält man hieraus die Jordansche Normalform mittels Lemma 14.

(iii) Charakteristisches Polynom χ_f und Minimalpolynom p_f von f berechnen sich zu

$$\chi_f = \alpha_1 \dots \alpha_s, \quad p_f = \alpha_s.$$

Beweis. Wir können $V = K^n$ und f als den durch $x \mapsto Ax$ gegebenen Endomorphismus von V annehmen. Unter Benutzung der in Lemma 17 bereitgestellten endlichen Präsentation von V erhalten wir dann mit 6.4/2 die erste Zerlegung aus (i) und mittels des Chinesischen Restsatzes 6.4/1 auch die zweite, wobei der freie Anteil in 6.4/2 entfällt, da V ein $K[T]$ -Torsionsmodul ist. Die restlichen Aussagen ergeben sich mittels Satz 4 sowie mit den Lemmata 8 und 10, wie im Beweis zu Theorem 11. $\square$

Als Beispiel betrachten wir nochmals die Matrix

$$A = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 2 & 0 & 0 \\ -1 & 1 & 2 & 1 \\ -1 & 1 & 0 & 3 \end{pmatrix} \in \mathbb{R}^{4 \times 4}$$

und bestimmen zunächst die Elementarteiler der Matrix $TE - A \in \mathbb{R}^{4 \times 4}$, indem wir das Verfahren aus dem Beweis zu 6.3/5 anwenden:

$$\begin{aligned} & \begin{pmatrix} T-1 & -1 & 0 & -1 \\ 0 & T-2 & 0 & 0 \\ 1 & -1 & T-2 & -1 \\ 1 & -1 & 0 & T-3 \end{pmatrix} \mapsto \begin{pmatrix} 1 & -1 & 0 & T-3 \\ 0 & T-2 & 0 & 0 \\ 1 & -1 & T-2 & -1 \\ T-1 & -1 & 0 & -1 \end{pmatrix} \\ & \mapsto \begin{pmatrix} 1 & -1 & 0 & T-3 \\ 0 & T-2 & 0 & 0 \\ 0 & 0 & T-2 & -(T-2) \\ 0 & T-2 & 0 & -(T-2)^2 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & T-2 & 0 & 0 \\ 0 & 0 & T-2 & 0 \\ 0 & 0 & 0 & (T-2)^2 \end{pmatrix} \end{aligned}$$

Damit erhalten wir $T-2, T-2, (T-2)^2$ als die nicht-invertierbaren Elementarteiler von $TE - A$ und gemäß Theorem 18 dann $\chi_A = (T-2)^4$ als charakteristisches Polynom, sowie $p_A = (T-2)^2$ als Minimalpolynom zu A . Da die

Elementarteiler bereits Primpotenzen sind, entfällt die Anwendung des Chinesischen Restsatzes und allgemeine bzw. Jordansche Normalform von A ergeben sich zu

$$\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & -4 \\ 0 & 0 & 1 & 4 \end{pmatrix}, \quad \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 2 \end{pmatrix}.$$

Aufgaben

Im Folgenden sei K ein Körper.

1. Es sei $A \in \mathbb{R}^{7 \times 7}$ eine Matrix mit charakteristischem Polynom

$$\chi_A = (T^2 + 1)^2(T - 2)(T^2 - 1) \in \mathbb{R}[T].$$

Man untersuche, welche Gestalt die allgemeine Normalform von A haben kann.

2. Man zeige, dass zwei Matrizen $A, B \in \mathbb{R}^{3 \times 3}$ genau dann ähnlich sind, wenn ihre Minimalpolynome sowie ihre charakteristischen Polynome übereinstimmen.
3. Für die Matrizen

$$\begin{pmatrix} 2 & 2 & 0 & -3 \\ 1 & 1 & 0 & -1 \\ 1 & 2 & -1 & -1 \\ 1 & 2 & 0 & -2 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 1 & -2 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & -2 \\ -1 & 1 & 1 & -1 \end{pmatrix}, \quad \begin{pmatrix} 3 & 0 & 1 & -2 \\ 4 & 1 & 1 & -3 \\ 0 & 0 & -1 & 0 \\ 4 & 0 & 2 & -3 \end{pmatrix} \in \mathbb{R}^{4 \times 4}$$

berechne man jeweils charakteristisches und Minimalpolynom sowie, falls existent, die Jordansche Normalform.

4. Es seien $A, B \in K^{n \times n}$ Matrizen mit den Minimalpolynomen $p_A, p_B \in K[T]$ und mit Primfaktorzerlegung $p_A = p_1^{n_1} \dots p_r^{n_r}$, wobei $p_1, \dots, p_r \in K[T]$ paarweise verschiedene normierte Primpolynome seien. Man zeige, dass A und B genau dann ähnlich sind, wenn $p_A = p_B$ sowie $\operatorname{rg} p_i^{k_i}(A) = \operatorname{rg} p_i^{k_i}(B)$ für $i = 1, \dots, r$ und $1 \leq k_i \leq n_i$ gilt.
5. Man betrachte normierte Primpolynome $p_1, \dots, p_n \in K[T]$, die paarweise verschieden seien, sowie natürliche Zahlen $1 \leq r_i \leq s_i$, $i = 1, \dots, n$. Gibt es einen Endomorphismus $f: V \rightarrow V$ eines endlich-dimensionalen K -Vektorraums V mit Minimalpolynom $p_f = p_1^{r_1} \dots p_n^{r_n}$ und mit charakteristischem Polynom $\chi_f = p_1^{s_1} \dots p_n^{s_n}$?
6. Es sei V ein endlich-dimensionaler K -Vektorraum mit einem Endomorphismus $f: V \rightarrow V$ und $U \subset V$ ein f -invarianter Unterraum. Man zeige:
 - (i) f induziert einen Endomorphismus $\bar{f}: V/U \rightarrow V/U$.
 - (ii) Es gilt $p_{\bar{f}} | p_f$ für die Minimalpolynome von $\bar{f}$ und f .
 - (iii) Es gilt $\chi_f = \chi_{f|U} \cdot \chi_{\bar{f}}$ für die charakteristischen Polynome von f , von $f|_U$ (der Einschränkung von f auf U) und von $\bar{f}$.
7. Es sei V ein endlich-dimensionaler K -Vektorraum mit einem Endomorphismus $f: V \rightarrow V$. Das Minimalpolynom p_f sei Potenz eines Primpolynoms $p \in K[T]$, etwa $p_f = p^r$ mit $r > 0$. Man zeige:
 - (i) Es existiert ein Vektor $u \in V$ mit $p^{r-1}(f)(u) \neq 0$.

- (ii) Ist $u \in V$ wie in (i), und ist $U \subset V$ der von u erzeugte f -zyklische Untervektorraum, so existiert ein f -invarianter Untervektorraum $U' \subset V$ mit $V = U \oplus U'$.
8. Es sei V ein endlich-dimensionaler f -zyklischer K -Vektorraum unter einem Endomorphismus $f: V \rightarrow V$. Man zeige, dass jeder f -invariante Unterraum $U \subset V$ wiederum f -zyklisch ist.
9. Es sei V ein endlich-dimensionaler f -zyklischer K -Vektorraum unter einem Endomorphismus $f: V \rightarrow V$, und es sei $u \in V$ ein erzeugendes Element. Man zeige:
- (i) Es gibt auf V eine eindeutig bestimmte Ringstruktur, deren Addition mit der Vektorraumaddition auf V übereinstimmt und deren Multiplikation $(af^i(u)) \cdot (bf^j(u)) = abf^{i+j}(u)$ für $a, b \in K$ und $i, j \in \mathbb{N}$ erfüllt.
 - (ii) Unter dieser Ringstruktur ist V genau dann ein Körper, wenn das Minimalpolynom zu f prim ist.

7. Euklidische und unitäre Vektorräume

Vorbemerkungen

In den Vorbemerkungen zu Kapitel 1 hatten wir überlegt, auf welche Weise man den $\mathbb{R}$ -Vektorraum $\mathbb{R}^2$ als Modell einer anschaulichen Ebene interpretieren kann. Will man in diesem Modell auch Abstände und damit letztendlich Winkel korrekt reflektieren, so muss man das Modell mit einer Abstandsfunktion ausstatten. Beispielsweise ist im $\mathbb{R}^2$ der gewöhnliche euklidische Abstand für zwei Punkte $P_1 = (x_1, y_1)$, $P_2 = (x_2, y_2)$ durch

$$d(P_1, P_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

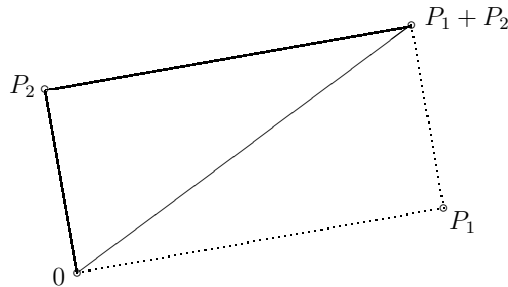
gegeben, bzw. die Länge oder der *Betrag* des Vektors $\overrightarrow{OP_1}$ durch

$$|\overrightarrow{OP_1}| := |P_1| := \sqrt{x_1^2 + y_1^2}.$$

Hieraus gewinnt man mit

$$\langle P_1, P_2 \rangle := \frac{1}{2}(|P_1 + P_2|^2 - |P_1|^2 - |P_2|^2) = x_1 x_2 + y_1 y_2$$

eine Funktion in zwei Variablen, eine so genannte *symmetrische Bilinearform*, in unserem Falle sogar ein *Skalarprodukt*, welches dann $|P_1| = \sqrt{\langle P_1, P_1 \rangle}$ erfüllt. Als Indiz, dass man mit einem solchen Skalarprodukt auch Winkel charakterisieren kann, mag folgende Beobachtung dienen: Das Skalarprodukt $\langle P_1, P_2 \rangle$ verschwindet genau dann, wenn $|P_1 + P_2|^2 = |P_1|^2 + |P_2|^2$ gilt. Für nicht-triviale Punkte P_1, P_2 ist dies aufgrund der Umkehrung des Satzes von Pythagoras äquivalent dazu, dass die Vektoren $\overrightarrow{OP_1}$ und $\overrightarrow{OP_2}$ aufeinander senkrecht stehen:



Thema des vorliegenden Kapitels ist das Studium von endlich-dimensionalen $\mathbb{R}$ - und $\mathbb{C}$ -Vektorräumen V mit einer Abstandsfunktion, die durch ein Skalarprodukt $\langle \cdot, \cdot \rangle: V \times V \longrightarrow \mathbb{K}$ gegeben ist; für $\mathbb{K}$ ist dabei der jeweils betrachtete Grundkörper $\mathbb{R}$ bzw. $\mathbb{C}$ einzusetzen. Wir sprechen dann von *euklidischen* bzw. *unitären Vektorräumen*. Wichtigstes Beispiel ist der Vektorraum $\mathbb{K}^n$ mit dem Skalarprodukt $\langle x, y \rangle = \sum_{i=1}^n x_i \overline{y_i}$, wobei $\overline{y_i}$ jeweils die zu y_i konjugierte komplexe Zahl bezeichne. Jedes Skalarprodukt erfüllt die *Schwarzsche Ungleichung* $|\langle x, y \rangle| \leq |x| |y|$ für $x, y \in V$, aus der sich insbesondere die Dreiecksungleichung $|x + y| \leq |x| + |y|$ ergibt. Im Übrigen bietet die Schwarzsche Ungleichung im Falle $\mathbb{K} = \mathbb{R}$ die Möglichkeit, für $x, y \in V - \{0\}$ den Winkel ω zwischen x und y mittels

$$\cos \omega = \frac{\langle x, y \rangle}{|x| \cdot |y|}$$

zu definieren. Als fundamentales Hilfsmittel werden wir das nach E. Schmidt benannte Orthonormalisierungsverfahren behandeln. Es erlaubt, aus einer beliebigen Basis $x_1, \dots, x_n$ von V durch sukzessives Abändern der x_i eine so genannte *Orthonormalbasis* $e_1, \dots, e_n$ zu konstruieren, die $|e_i| = 1$ für alle i und $\langle e_i, e_j \rangle = 0$ für alle $i \neq j$ erfüllt. Eine solche Orthonormalbasis ist damit als ein rechtwinkliges Koordinatensystem in V anzusehen. Im Übrigen lässt sich das Skalarprodukt von V aufgrund seiner Bilinearität leicht mittels einer Orthonormalbasis von V rekonstruieren, so dass ein Vektorraum V mit einer Abstandsfunktion in Form eines Skalarprodukts bereits dann festgelegt ist, wenn man in V eine Orthonormalbasis kennt.

In Kapitel 6 haben wir für endlich-dimensionale Vektorräume V über einem beliebigen Körper K Normalformen von Endomorphismen $f: V \longrightarrow V$ studiert. Ist f beispielsweise durch eine Matrix $A_{f,X,X}$ bezüglich einer gewissen Basis X von V gegeben, so ging es darum, eine Basiswechselmatrix $A_{\text{id},X',X}$ zu finden, so dass die Matrix

$$A_{f,X',X'} = (A_{\text{id},X',X})^{-1} \cdot A_{f,X,X} \cdot A_{\text{id},X',X}$$

von möglichst "einfacher" Gestalt war. Mit anderen Worten, wir hatten mit der Normalformtheorie die Ähnlichkeitsklassen von Matrizen in $K^{n \times n}$ bestimmt, wobei zwei Matrizen $A, B \in K^{n \times n}$ ähnlich heißen, wenn es eine invertierbare Matrix $S \in \text{Gl}(n, K)$ mit $B = S^{-1} \cdot A \cdot S$ gibt.

Wir wollen ein analoges Problem nun auch im Rahmen euklidischer und unitärer Vektorräume V behandeln. Allerdings müssen wir hier die Abstandsfunktion auf V mit einbeziehen. Wir werden deshalb nur solche Basiswechselmatrizen $A_{\text{id},X',X}$ zulassen, die längenerhaltend sind und damit einen Wechsel zwischen Orthonormalbasen X und X' definieren. Letzteres ist genau dann der Fall, wenn

$$(A_{\text{id},X',X})^{-1} = (A_{\text{id},X',X})^* := \overline{(A_{\text{id},X',X})}^t$$

gilt, wobei für eine Matrix $A = (\alpha_{ij}) \in \mathbb{K}^{n \times n}$ die zugehörige *komplex konjugierte* Matrix $\overline{A}$ durch $\overline{A} = (\overline{\alpha_{ij}})$ gegeben sei; eine Matrix $S = A_{\text{id},X',X}$ mit der vorstehenden Eigenschaft wird als *orthogonal* (für $\mathbb{K} = \mathbb{R}$) bzw. *unitär* (für $\mathbb{K} = \mathbb{C}$)

bezeichnet. Die mittels solcher Matrizen definierte Relation der Äquivalenz von Matrizen in $\mathbb{K}^{n \times n}$ ist somit viel enger gefasst als die in Kapitel 6 betrachtete allgemeine Äquivalenz von Matrizen. Es ist deshalb sinnvoll, sich für das Normalformenproblem auf gewisse Teilklassen von Matrizen in $\mathbb{K}^{n \times n}$ zu beschränken. Wir werden hier im Wesentlichen nur *symmetrische* bzw. *hermitesche* Matrizen betrachten, d. h. Matrizen $A \in \mathbb{K}^{n \times n}$ mit $A = A^*$; bezüglich Orthonormalbasen stellen diese gerade die so genannten *selbstadjungierten* Endomorphismen $f: V \rightarrow V$ dar, die der Relation $\langle f(x), y \rangle = \langle x, f(y) \rangle$ für $x, y \in V$ genügen. Dann können wir allerdings die erstaunliche Tatsache zeigen, dass eine solche Matrix unter der strengeren Äquivalenz mittels orthogonaler bzw. unitärer Matrizen stets äquivalent zu einer *reellen* Diagonalmatrix ist.

Die bei dem vorstehend beschriebenen Klassifikationsproblem gewonnenen Erkenntnisse lassen sich mit Gewinn auch auf die Klassifikation von *symmetrischen Bilinearformen* (für $\mathbb{K} = \mathbb{R}$) bzw. *hermiteschen Formen* (für $\mathbb{K} = \mathbb{C}$) auf endlich-dimensionalen $\mathbb{K}$ -Vektorräumen V anwenden, also auf entsprechenden Abbildungen $\Phi: V \times V \rightarrow \mathbb{K}$. Denn auch diese werden bezüglich einer Basis X von V durch eine Matrix $A_{\Phi, X}$ beschrieben, welche die Relation $A_{\Phi, X} = A_{\Phi, X}^*$ erfüllt. Allerdings ist das Transformationsverhalten bei Basiswechsel ein anderes als bei Endomorphismen, nämlich

$$A_{\Phi, X'} = A_{\text{id}, X', X}^t \cdot A_{\Phi, X} \cdot \overline{A_{\text{id}, X', X}}.$$

Wenn wir uns aber im Rahmen euklidischer bzw. unitärer Vektorräume bewegen und lediglich orthogonale bzw. unitäre Basiswechselmatrizen zulassen, so gilt $A_{\text{id}, X', X}^t = \overline{A_{\text{id}, X', X}}^{-1}$, und damit ein analoges Transformationsverhalten wie bei Endomorphismen von V . Wir können damit die Klassifikation selbstadjungierter Endomorphismen verwenden und erhalten für euklidische bzw. unitäre Vektorräume, dass die Klassen symmetrischer Bilinearformen bzw. hermitescher Formen durch *reelle* Diagonalmatrizen repräsentiert werden (Satz über die *Hauptachsentransformation*.) Es ist dann relativ leicht einzusehen, dass die entsprechenden Klassen bezüglich allgemeiner Äquivalenz (mittels invertierbarer Matrizen in $\text{Gl}(n, \mathbb{K})$) durch Diagonalmatrizen repräsentiert werden, deren Diagonaleinträge die Werte $1, -1, 0$ annehmen können (*Sylvesterscher Trägheitssatz*).

Die Bezeichnung Hauptachsentransformation weist auf einen konkreten geometrischen Sachverhalt hin. Für $\mathbb{K} = \mathbb{R}$ betrachte man beispielsweise eine quadratische Form auf $\mathbb{R}^n$, gegeben durch

$$q(x) = \sum_{i \leq j; i, j=1, \dots, n} \alpha_{ij} x_i x_j$$

mit gewissen Konstanten $\alpha_{ij} \in \mathbb{R}$, sowie das durch $q(x) = c$ für ein $c \in \mathbb{R}$ gegebene geometrische Gebilde. Dann kann man q die symmetrische Bilinearform

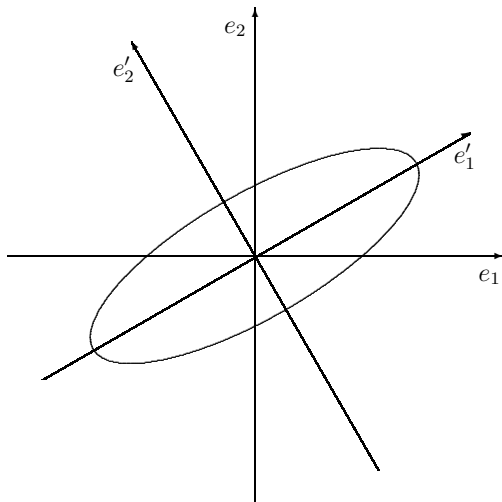
$$\langle \cdot, \cdot \rangle: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}, \quad \langle x, y \rangle = \frac{1}{2}(q(x+y) - q(x) - q(y)),$$

zuordnen; für diese gilt $q(x) = \langle x, x \rangle$. In dieser Situation besagt der Satz über die Hauptachsentransformation, dass es ein neues rechtwinkliges Koordinatensystem $e'_1, \dots, e'_n$ in $\mathbb{R}^n$ gibt, so dass sich q bezüglich der neuen Koordinaten in

der Form

$$q'(x') = \sum_{i=1}^n \lambda_i x_i'^2$$

mit gewissen $\lambda_i \in \mathbb{R}$ beschreibt. Nehmen wir etwa $n = 2$ und $\lambda_1, \lambda_2, c > 0$ an, so wird durch $q(x) = c$ eine Ellipse beschrieben, deren Achsen zunächst noch nicht mit den Achsen des gegebenen Koordinatensystems e_1, e_2 übereinstimmen. Die Hauptachsentransformation besagt gerade, dass man durch orthogonalen Basiswechsel erreichen kann, dass die Achsen des neuen Koordinatensystems e'_1, e'_2 mit den Achsen der Ellipse übereinstimmen:



7.1 Sesquilinearformen

Im Folgenden werden wir ausschließlich Vektorräume über den Körpern $\mathbb{R}$ oder $\mathbb{C}$ betrachten und meist $\mathbb{K}$ anstelle von $\mathbb{R}$ oder $\mathbb{C}$ schreiben. Speziell für $\mathbb{C}$ wird die komplexe *Konjugationsabbildung*

$$\mathbb{C} \longrightarrow \mathbb{C}, \quad a \longmapsto \bar{a},$$

von Bedeutung sein, wobei für $a = \alpha + i\beta$ mit $\alpha, \beta \in \mathbb{R}$ die zugehörige *komplex konjugierte Zahl* durch $\bar{a} = \alpha - i\beta$ gegeben ist. Die Konjugationsabbildung $a \longmapsto \bar{a}$ stellt eine so genannte *Involution* auf $\mathbb{C}$ dar, d. h. einen Automorphismus $\sigma: \mathbb{C} \longrightarrow \mathbb{C}$ mit $\sigma^2 = \text{id}$, und es gilt zusätzlich $\sigma|_{\mathbb{R}} = \text{id}$. Für $a = \alpha + i\beta \in \mathbb{C}$ mit $\alpha, \beta \in \mathbb{R}$ bezeichnet man

$$\alpha = \frac{1}{2}(a + \bar{a}) = \text{Re}(a)$$

als den *Realteil* von a ,

$$\beta = \frac{1}{2i}(a - \bar{a}) = \text{Im}(a)$$

als den *Imaginärteil* von a , sowie

$$|a| = \sqrt{\alpha^2 + \beta^2} = \sqrt{a \cdot \bar{a}}$$

als den *Absolutbetrag* von a . Dabei bestehen die Äquivalenzen

$$a \in \mathbb{R} \iff \operatorname{Re}(a) = a \iff \operatorname{Im}(a) = 0 \iff a = \bar{a}.$$

Für spezielle Elemente $a \in \mathbb{K}$ werden wir im Folgenden auch Bedingungen des Typs $a \geq 0$ oder $a > 0$ betrachten. Hiermit ist gemeint, dass a *reell* ist, also $a \in \mathbb{R} \subset \mathbb{K}$ erfüllt, und zudem der Bedingung $a \geq 0$ bzw. $a > 0$ genügt.

Definition 1. Es sei V ein $\mathbb{K}$ -Vektorraum. Eine Abbildung $\Phi: V \times V \longrightarrow \mathbb{K}$ heißt *Sesquilinearform* (*sesqui* = *eineinhalb*), wenn für $x, x_1, x_2, y, y_1, y_2 \in V$ sowie $\alpha \in \mathbb{K}$ gilt:

- (i) $\Phi(x_1 + x_2, y) = \Phi(x_1, y) + \Phi(x_2, y)$,
- (ii) $\Phi(\alpha x, y) = \alpha \Phi(x, y)$,
- (iii) $\Phi(x, y_1 + y_2) = \Phi(x, y_1) + \Phi(x, y_2)$,
- (iv) $\Phi(x, \alpha y) = \bar{\alpha} \Phi(x, y)$.

Im Falle $\mathbb{K} = \mathbb{R}$ gilt stets $\alpha = \bar{\alpha}$, und man spricht dann auch von einer *Bilinearform* Φ . Man bezeichnet Φ als *nicht- α -geartet*, wenn gilt:

$$\begin{aligned} \Phi(x, y) &= 0 \text{ für alle } y \in V \text{ impliziert } x = 0, \\ \Phi(x, y) &= 0 \text{ für alle } x \in V \text{ impliziert } y = 0. \end{aligned}$$

Um ein einfaches Beispiel zu erhalten, setze man $V = \mathbb{K}$. Dann wird für festes $a \in \mathbb{K}$ durch

$$\Phi(x, y) = a \cdot x \cdot \bar{y}$$

eine Sesquilinearform auf V definiert, und diese ist genau dann nicht α -geartet, wenn $a \neq 0$ gilt.

Definition 2. Eine Sesquilinearform $\Phi: V \times V \longrightarrow \mathbb{K}$, welche

$$\Phi(x, y) = \overline{\Phi(y, x)} \quad \text{für alle } x, y \in V$$

erfüllt, wird im Falle $\mathbb{K} = \mathbb{R}$ als *symmetrische Bilinearform* (oder kurz *sBF*) und im Falle $\mathbb{K} = \mathbb{C}$ als *Hermiteische Form* (oder kurz *HF*) bezeichnet. Für solche Formen verwendet man anstelle von $\Phi(x, y)$ häufig auch die Notation $\langle x, y \rangle$. Insbesondere gilt dann $\Phi(x, x) = \langle x, x \rangle \in \mathbb{R}$ für alle $x \in V$.

Beispielsweise definiert

$$\Phi: \mathbb{K}^n \times \mathbb{K}^n \longrightarrow \mathbb{K}, \quad (x, y) \longmapsto x^t \cdot \bar{y},$$

eine sBF bzw. HF, wobei diese Abbildung in ausführlicher Schreibweise wie folgt gegeben ist:

$$((x_1, \dots, x_n)^t, (y_1, \dots, y_n)^t) \mapsto \sum_{i=1}^n x_i \cdot \overline{y_i}$$

Man erkennt leicht, dass die Form Φ nicht ausgeartet ist, da etwa $\Phi(x, x) > 0$ für $x \neq 0$ gilt. Andererseits wird durch die Vorschrift

$$((x_1, \dots, x_n)^t, (y_1, \dots, y_n)^t) \mapsto \sum_{i=1}^r x_i \cdot \overline{y_i},$$

wobei man nur bis zu einer Zahl $r < n$ summiere, eine ausgeartete Form auf $\mathbb{K}^n$ erklärt.

Definition 3. Eine sBF bzw. HF $\Phi: V \times V \longrightarrow \mathbb{K}$ heißt positiv semi-definit, falls $\Phi(x, x) \geq 0$ für alle $x \in V$ gilt. Gilt sogar $\Phi(x, x) > 0$ für alle $x \in V - \{0\}$, so bezeichnet man Φ als positiv definit. Man spricht dann auch von einem Skalarprodukt auf V und nennt das Paar (V, Φ) im Falle $\mathbb{K} = \mathbb{R}$ einen euklidischen Vektorraum, sowie im Falle $\mathbb{K} = \mathbb{C}$ einen unitären Vektorraum.

Die oben betrachtete Form $\mathbb{K}^n \times \mathbb{K}^n \longrightarrow \mathbb{K}$, $(x, y) \mapsto x^t \cdot \overline{y}$ definiert beispielsweise ein Skalarprodukt auf $\mathbb{K}^n$, und zwar das so genannte *kanonische* Skalarprodukt.

Satz 4 (Schwarzsche Ungleichung). Sei $\Phi: V \times V \longrightarrow \mathbb{K}$, $(x, y) \mapsto \langle x, y \rangle$, eine positiv semi-definite sBF bzw. HF auf einem $\mathbb{K}$ -Vektorraum V . Dann besteht für $x, y \in V$ folgende Ungleichung:

$$|\langle x, y \rangle|^2 \leq \langle x, x \rangle \cdot \langle y, y \rangle$$

Ist Φ sogar positiv definit, so gilt in dieser Formel genau dann das Gleichheitszeichen, wenn x und y linear abhängig sind.

Beweis. Man wähle $\alpha, \beta \in \mathbb{K}$ beliebig. Dann gilt

$$\begin{aligned} 0 &\leq \langle \alpha x + \beta y, \alpha x + \beta y \rangle \\ &= \alpha \overline{\alpha} \langle x, x \rangle + \alpha \overline{\beta} \langle x, y \rangle + \beta \overline{\alpha} \langle y, x \rangle + \beta \overline{\beta} \langle y, y \rangle \\ &= \alpha \overline{\alpha} \langle x, x \rangle + 2\operatorname{Re}(\alpha \overline{\beta} \langle x, y \rangle) + \beta \overline{\beta} \langle y, y \rangle. \end{aligned}$$

Hat man nun $\langle x, x \rangle = \langle y, y \rangle = 0$, so setze man speziell $\alpha = -1$ und $\beta = \langle x, y \rangle$. Es folgt

$$0 \leq 2\operatorname{Re}(\alpha \overline{\beta} \langle x, y \rangle) = -2|\langle x, y \rangle|^2 \leq 0$$

und damit $\langle x, y \rangle = 0$, also insbesondere

$$|\langle x, y \rangle|^2 \leq \langle x, x \rangle \cdot \langle y, y \rangle.$$

Sei nun $\langle x, x \rangle > 0$. In diesem Falle setze man

$$\alpha = -\overline{\langle x, y \rangle}, \quad \beta = \langle x, x \rangle.$$

Dann folgt

$$\begin{aligned} 0 &\leq |\langle x, y \rangle|^2 \langle x, x \rangle - 2|\langle x, y \rangle|^2 \langle x, x \rangle + \langle x, x \rangle^2 \langle y, y \rangle \\ &= -|\langle x, y \rangle|^2 \langle x, x \rangle + \langle x, x \rangle^2 \langle y, y \rangle, \end{aligned}$$

also wie gewünscht

$$|\langle x, y \rangle|^2 \leq \langle x, x \rangle \cdot \langle y, y \rangle.$$

Der Fall $\langle y, y \rangle > 0$ lässt sich entsprechend behandeln.

Sind schließlich $x, y \in V$ linear abhängig, so ist einer dieser Vektoren ein Vielfaches des anderen, und es ergibt sich ohne Schwierigkeiten

$$|\langle x, y \rangle|^2 = \langle x, x \rangle \cdot \langle y, y \rangle.$$

Ist umgekehrt diese Gleichung gegeben, so erhält man mit

$$\alpha = -\overline{\langle x, y \rangle}, \quad \beta = \langle x, x \rangle$$

wie oben

$$\begin{aligned} &\langle \alpha x + \beta y, \alpha x + \beta y \rangle \\ &= |\langle x, y \rangle|^2 \langle x, x \rangle - 2|\langle x, y \rangle|^2 \langle x, x \rangle + \langle x, x \rangle^2 \langle y, y \rangle = 0. \end{aligned}$$

Ist nun Φ positiv definit, so ergibt sich $\alpha x + \beta y = 0$, und man sieht, dass x, y linear abhängig sind. Für $x = 0$ ist dies nämlich trivialerweise klar und für $x \neq 0$ ebenfalls, da dann der Koeffizient $\beta = \langle x, x \rangle$ aufgrund der positiven Definitheit von Φ nicht verschwindet. $\square$

Korollar 5. *Sei $\Phi: V \times V \longrightarrow \mathbb{K}$ positiv semi-definite sBF bzw. HF. Es ist Φ genau dann positiv definit, wenn Φ nicht ausgeartet ist.*

Ist $\Phi: V \times V \longrightarrow \mathbb{K}$ positiv semi-definite sBF bzw. HF, so gilt $\langle x, x \rangle \geq 0$ für alle $x \in V$, und man bezeichnet mit $|x| = \sqrt{\langle x, x \rangle}$ die *Länge* oder den *Betrag* eines Vektors $x \in V$ (bezüglich Φ). Der so definierte Betrag von Vektoren erfüllt die gewöhnliche Dreiecksungleichung:

Korollar 6. *Sei $\Phi: V \times V \longrightarrow \mathbb{K}$ positiv semi-definite sBF bzw. HF. Dann gilt für $x, y \in V$*

$$|x + y| \leq |x| + |y|.$$

Beweis. Es gilt

$$\begin{aligned} |x + y|^2 &= \langle x + y, x + y \rangle \\ &= \langle x, x \rangle + 2\operatorname{Re}(\langle x, y \rangle) + \langle y, y \rangle \\ &\leq \langle x, x \rangle + 2|\langle x, y \rangle| + \langle y, y \rangle \\ &\leq |x|^2 + 2|x||y| + |y|^2 \\ &= (|x| + |y|)^2 \end{aligned}$$

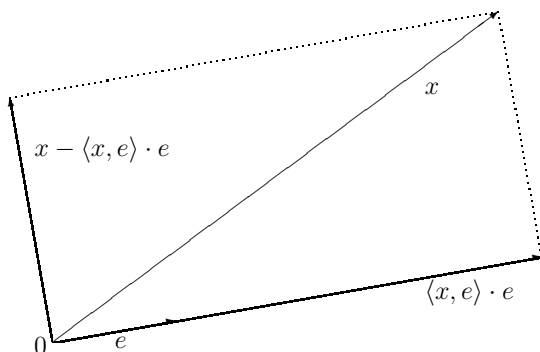
und damit $|x + y| \leq |x| + |y|$. $\square$

An weiteren Eigenschaften des Betrages von Vektoren können wir anführen: Für $\alpha \in \mathbb{K}$, $x \in V$ gilt $|\alpha x| = |\alpha||x|$. Ist Φ sogar positiv definit, so ist $|x| = 0$ äquivalent zu $x = 0$. Weiter bezeichnet man einen Vektor $x \in V$ als *normiert*, falls $\langle x, x \rangle = 1$ und damit $|x| = 1$ gilt. Für $x \in V$ mit $|x| \neq 0$ ist beispielsweise $\frac{x}{|x|}$ normiert.

Im Falle des kanonischen Skalarprodukts auf dem $\mathbb{R}^n$ stimmt der Betrag eines Vektors $x \in \mathbb{R}^n$ gerade mit dem üblichen euklidischen Abstand des Punktes x vom Nullpunkt überein. Ist $e \in \mathbb{R}^n$ ein weiterer Vektor mit $|e| = 1$, so gilt

$$|x|^2 = \langle x, e \rangle^2 + |x - \langle x, e \rangle e|^2,$$

wie man leicht nachrechnet. Geometrisch bedeutet dies aufgrund der Umkehrung des Satzes von Pythagoras, dass das Dreieck mit den Seitenlängen $|x|$, $|\langle x, e \rangle|$ und $|x - \langle x, e \rangle e|$ rechtwinklig ist:



Somit entsteht der Vektor $\langle x, e \rangle \cdot e$, indem man x senkrecht auf den durch e gegebenen Unterraum $\mathbb{R} \cdot e \subset \mathbb{R}^n$ projiziert. Mit anderen Worten, das Skalarprodukt zweier Vektoren $x, y \in \mathbb{R}^n - \{0\}$ ist, abgesehen vom Vorzeichen, gleich dem Produkt von $|y|$ mit dem Betrag der senkrechten Projektion von x auf den Unterraum $\mathbb{R} \cdot y \subset \mathbb{R}^n$.

Die Möglichkeit, mit $x - \langle x, e \rangle \cdot e$ aus x einen Vektor zu konstruieren, der senkrecht auf e steht, ist im Übrigen die zentrale Idee des Orthonormalisierungsverfahrens von E. Schmidt, das wir im nachfolgenden Abschnitt 7.2 behandeln werden.

Aufgaben

Falls nicht anders bestimmt, sei V ein endlich-dimensionaler $\mathbb{R}$ -Vektorraum und $\Phi: V \times V \rightarrow \mathbb{R}$ eine symmetrische Bilinearform auf V .

1. Man definiere den *Kern* von Φ durch

$$\ker \Phi = \{x \in V ; \Phi(x, y) = 0 \text{ für alle } y \in V\}$$

und zeige, dass Φ eine nicht-ausgeartete symmetrische Bilinearform auf $V/\ker \Phi$ induziert.

2. Man definiere die zu Φ gehörige *quadratische Form* $q: V \rightarrow \mathbb{R}$, indem man $q(x) = \Phi(x, x)$ für $x \in V$ setze, und zeige, dass Φ durch q eindeutig bestimmt ist.
3. Es sei Φ positiv definit. Für $x, y \in V$, $y \neq 0$, betrachte man die polynomiale Funktion $p(t) = |x + ty|^2$ in $t \in \mathbb{R}$. Man bestimme sämtliche Nullstellen von $p(t)$ und folgere die Schwarzsche Ungleichung in der Version von Satz 4 für den Fall $\mathbb{K} = \mathbb{R}$.
4. Es sei Φ positiv definit. Für $x, y \in V - \{0\}$ zeige man:
 - (i) Es gilt $-1 \leq \frac{\Phi(x, y)}{|x| \cdot |y|} \leq 1$. Aus der Infinitesimalrechnung ergibt sich damit die Existenz eines Winkels $0 \leq \omega \leq \pi$ mit $\frac{\Phi(x, y)}{|x| \cdot |y|} = \cos \omega$.
 - (ii) Es gilt der Cosinus-Satz:

$$|x - y|^2 = |x|^2 + |y|^2 - 2|x||y|\cos \omega$$

5. Es sei V ein Vektorraum über einem Körper K und V^* sein Dualraum, sowie W der K -Vektorraum aller K -bilinearen Abbildungen $V \times V \rightarrow K$. Man zeige, dass die Abbildung $\text{Hom}_K(V, V^*) \rightarrow W$, welche einem Homomorphismus $\varphi: V \rightarrow V^*$ die Abbildung

$$V \times V \rightarrow K, \quad (x, y) \mapsto \varphi(x)(y),$$

zuordnet, ein Isomorphismus von K -Vektorräumen ist.

7.2 Orthogonalität

Definition 1. Sei $\Phi: V \times V \rightarrow \mathbb{K}$, $(x, y) \mapsto \langle x, y \rangle$, eine sBF bzw. HF. Zwei Vektoren $x, y \in V$ heißen *orthogonal* bzw. *senkrecht zueinander*, wenn $\langle x, y \rangle = 0$ gilt. Ein System M von Vektoren aus V mit $0 \notin M$ heißt *Orthogonalsystem*, wenn je zwei Vektoren aus M zueinander orthogonal sind. Gilt zusätzlich $\langle x, x \rangle = 1$ für alle $x \in M$, so spricht man auch von einem *Orthonormalsystem*.

Bemerkung 2. Sei $\Phi: V \times V \rightarrow \mathbb{K}$ ein Skalarprodukt. Ist dann M ein Orthogonalsystem von Vektoren aus V , so ist M linear unabhängig.

Beweis. Zu einem endlichen Teilsystem $x_1, \dots, x_n$ von M betrachte man Koeffizienten $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ mit $\sum_{i=1}^n \alpha_i x_i = 0$. Dann ergibt sich für $j = 1, \dots, n$

$$0 = \left\langle \sum_{i=1}^n \alpha_i x_i, x_j \right\rangle = \sum_{i=1}^n \alpha_i \langle x_i, x_j \rangle = \alpha_j \langle x_j, x_j \rangle$$

und damit $\alpha_j = 0$ wegen $\langle x_j, x_j \rangle > 0$. □

Definition 3. Sei $\Phi: V \times V \rightarrow \mathbb{K}$ ein Skalarprodukt. Ein System M von Vektoren $e_1, \dots, e_n \in V$ wird als *Orthogonalbasis* (bzw. *Orthonormalbasis*) von V bezeichnet, wenn M folgende Bedingungen erfüllt:

- (i) M ist eine Basis von V .
- (ii) M ist ein Orthogonalsystem (bzw. Orthonormalsystem).

Das Skalarprodukt von Vektoren, die als Linearkombinationen von Elementen einer Orthonormalbasis $M = (e_1, \dots, e_n)$ dargestellt sind, lässt sich in einfacher Weise berechnen:

$$\begin{aligned} \left\langle \sum_{i=1}^n \alpha_i e_i, \sum_{i=1}^n \beta_i e_i \right\rangle &= \sum_{i,j=1}^n \alpha_i \bar{\beta}_j \langle e_i, e_j \rangle \\ &= \sum_{i,j=1}^n \alpha_i \bar{\beta}_j \delta_{ij} = \sum_{i=1}^n \alpha_i \bar{\beta}_i \end{aligned}$$

Betrachtet man beispielsweise das kanonische Skalarprodukt auf $\mathbb{K}^n$, so bildet die kanonische Basis von $\mathbb{K}^n$ eine Orthonormalbasis.

Als Nächstes wollen wir das so genannte *Orthonormalisierungsverfahren* von E. Schmidt besprechen, mit dessen Hilfe man Basen in euklidischen bzw. unitären Vektorräumen zu Orthonormalbasen abändern kann. Zunächst behandeln wir den Kernschritt dieses Verfahrens, einen Schritt, den wir für das kanonische Skalarprodukt auf dem $\mathbb{R}^n$ bereits zum Ende von Abschnitt 7.1 geometrisch motiviert hatten.

Lemma 4. Sei $\Phi: V \times V \longrightarrow \mathbb{K}$ ein Skalarprodukt, und sei $e_1, \dots, e_k$ eine Orthonormalbasis eines Untervektorraums $U \subset V$. Dann ist

$$p_U: V \longrightarrow U, \quad x \longmapsto \sum_{j=1}^k \langle x, e_j \rangle e_j,$$

eine surjektive $\mathbb{K}$ -lineare Abbildung, die so genannte orthogonale Projektion auf U . Diese beschränkt sich auf U zur identischen Abbildung und erfüllt im Übrigen die Gleichung $\langle x - p_U(x), y \rangle = 0$ für $x \in V$ und $y \in U$. Durch diese Gleichung ist p_U als $\mathbb{K}$ -lineare Abbildung $V \longrightarrow U$ eindeutig bestimmt.

Beweis. Die Abbildung p_U ist $\mathbb{K}$ -linear, da die Form $\langle x, y \rangle$ $\mathbb{K}$ -linear in x ist. Gilt weiter $x = \sum_{i=1}^k \alpha_i e_i \in U$, so hat man

$$\langle x, e_j \rangle = \left\langle \sum_{i=1}^k \alpha_i e_i, e_j \right\rangle = \sum_{i=1}^k \alpha_i \langle e_i, e_j \rangle = \alpha_j,$$

und es folgt insbesondere $p_U|_U = \text{id}_U$. Weiter ergibt sich

$$\begin{aligned} \langle x - p_U(x), e_j \rangle &= \langle x, e_j \rangle - \left\langle \sum_{i=1}^k \langle x, e_i \rangle e_i, e_j \right\rangle \\ &= \langle x, e_j \rangle - \langle x, e_j \rangle \langle e_j, e_j \rangle = 0, \end{aligned}$$

d. h. $x - p_U(x)$ ist orthogonal zu $e_1, \dots, e_k$ und damit zu U .

Ist $q_U: V \longrightarrow U$ irgendeine lineare Abbildung mit $\langle x - q_U(x), y \rangle = 0$ für $x \in V$ und $y \in U$, so hat man insbesondere $\langle p_U(x) - q_U(x), y \rangle = 0$ für $x \in V$ und $y \in U$. Da sich das Skalarprodukt von V zu einem Skalarprodukt auf U beschränkt, gilt notwendig $p_U(x) - q_U(x) = 0$ für $x \in V$ und damit $p_U = q_U$. $\square$

Ist nun $x_1, \dots, x_n$ eine Basis eines euklidischen bzw. unitären $\mathbb{K}$ -Vektorraums, so kann man zunächst x_1 normieren, also den Vektor $e_1 = |x_1|^{-1} \cdot x_1$ betrachten. Sodann kann man gemäß Lemma 4 die Projektion p_1 von V auf den Untervektorraum $U_1 = \mathbb{K}e_1$ bilden. Der Vektor $e'_2 = x_2 - p_1(x_2)$ ist dann orthogonal zu e_1 , und e_1 bildet zusammen mit $|e'_2|^{-1} \cdot e'_2$ ein Orthonormalsystem e_1, e_2 . Führt man in dieser Weise fort, so kann man die Basis $x_1, \dots, x_n$ orthonormalisieren, d. h. insgesamt in eine Orthonormalbasis überführen. Wir wollen dieses Resultat hier noch genauer formulieren und beweisen.

Satz 5. *Sei $\Phi: V \times V \longrightarrow \mathbb{K}$ ein Skalarprodukt, und sei $x_1, \dots, x_n$ eine Basis von V . Dann existieren eindeutig Vektoren $e_1, \dots, e_n \in V$, so dass gilt:*

- (i) $e_1, \dots, e_n$ ist eine Orthonormalbasis von V .
- (ii) $\bigoplus_{i=1}^k \mathbb{K}e_i = \bigoplus_{i=1}^k \mathbb{K}x_i$ für $k = 1, \dots, n$.
- (iii) Für die Basiswechselmatrix $A_k = A_{\text{id}, E_k, X_k}$ mit $E_k = (e_1, \dots, e_k)$ und $X_k = (x_1, \dots, x_k)$ gilt $\det A_k > 0$, $k = 1, \dots, n$.¹

Genauer lassen sich die Vektoren e_k für $k = 1, \dots, n$ in induktiver Weise wie folgt konstruieren:

$$e_k = |x_k - \sum_{j=1}^{k-1} \langle x_k, e_j \rangle e_j|^{-1} \cdot (x_k - \sum_{j=1}^{k-1} \langle x_k, e_j \rangle e_j).$$

Beweis. Wir zeigen zunächst die Existenzaussage und verwenden dabei Induktion nach n . Der Induktionsanfang $n = 0$ ist trivial. Sei also $n > 0$, und sei $e_1, \dots, e_{n-1}$ eine Orthonormalbasis von $U_{n-1} = \bigoplus_{i=1}^{n-1} \mathbb{K}x_i$, welche die gewünschten Eigenschaften besitzt. Man hat dann $x_n \notin U_{n-1} = \bigoplus_{i=1}^{n-1} \mathbb{K}e_i$, so dass $x_n - p_{n-1}(x_n)$ von Null verschieden ist; $p_{n-1}: V \longrightarrow U_{n-1}$ sei die Projektion gemäß Lemma 4. Dann ist

$$e_n = |x_n - p_{n-1}(x_n)|^{-1} \cdot (x_n - p_{n-1}(x_n))$$

wohldefiniert, und es folgt mit Bemerkung 2 sowie Lemma 4, dass $e_1, \dots, e_n$ eine Orthonormalbasis von V bilden. Weiter besteht gemäß Definition von e_n eine Gleichung des Typs

$$e_n = \alpha \cdot x_n + y$$

mit einer Konstanten $\alpha > 0$ und einem Vektor $y \in U_{n-1}$. Hieraus ergibt sich

$$A_n = \begin{pmatrix} A_{n-1} & * \\ 0 & \alpha \end{pmatrix},$$

und es folgt $\det A_n = \alpha \cdot \det A_{n-1} > 0$ wegen $\alpha, \det A_{n-1} > 0$.

¹ Gemäß unserer Konvention schließt $\det A_k > 0$ die Bedingung $\det A_k \in \mathbb{R}$ mit ein.

Zum Beweis der Eindeutigkeitsaussage sei $f_1, \dots, f_n$ eine Orthonormalbasis von V , die den Eigenschaften (ii) und (iii) genügt. Wiederum verwenden wir Induktion nach n , um $f_i = e_i$ für $i = 1, \dots, n$ zu zeigen, wobei der Induktionsanfang $n = 0$ trivial ist. Sei also $n > 0$. Nach Induktionsvoraussetzung dürfen wir $f_i = e_i$ für $i = 1, \dots, n-1$ annehmen. Dann existiert eine Gleichung des Typs

$$f_n = \alpha \cdot x_n + y$$

mit $\alpha \in \mathbb{K}$, $y \in U_{n-1}$. Folglich hat die Basiswechselmatrix $A_{\text{id}, E', X}$, wobei wir die Basen $E' = (f_1, \dots, f_n) = (e_1, \dots, e_{n-1}, f_n)$ und $X = (x_1, \dots, x_n)$ verwenden, die Gestalt

$$A = \begin{pmatrix} A_{n-1} & * \\ 0 & \alpha \end{pmatrix},$$

und es folgt $\alpha > 0$ wegen $\det A, \det A_{n-1} > 0$. Nun gilt aber $U_{n-1} = \bigoplus_{i=1}^{n-1} \mathbb{K}e_i$, und wir können y daher als Linearkombination der $e_1, \dots, e_{n-1}$ schreiben. Folglich gibt es Konstanten $\beta_1, \dots, \beta_{n-1} \in \mathbb{K}$ mit

$$f_n = \alpha(x_n - \sum_{i=1}^{n-1} \beta_i e_i).$$

Für $i = 1, \dots, n-1$ gilt dann

$$0 = \langle f_n, e_i \rangle = \alpha(\langle x_n, e_i \rangle - \beta_i),$$

also $\beta_i = \langle x_n, e_i \rangle$, also

$$f_n = \alpha(x_n - \sum_{i=1}^{n-1} \langle x_n, e_i \rangle e_i).$$

Aus der Gleichung

$$1 = |f_n| = \alpha |x_n - \sum_{i=1}^{n-1} \langle x_n, e_i \rangle e_i|$$

ergibt sich dann

$$\alpha = |x_n - \sum_{i=1}^{n-1} \langle x_n, e_i \rangle e_i|^{-1}$$

und somit $f_n = e_n$, was zu zeigen war. $\square$

Korollar 6. *Jeder endlich-dimensionale euklidische bzw. unitäre Vektorraum V besitzt eine Orthonormalbasis. Jede Orthonormalbasis eines Untervektorraums $U \subset V$ lässt sich zu einer Orthonormalbasis von V ergänzen.*

Korollar 7. *Sei V ein endlich-dimensionaler euklidischer bzw. unitärer Vektorraum, und sei $V_1 \subset V_2 \subset \dots \subset V_r = V$ eine Kette von $\mathbb{K}$ -Untervektorräumen. Sei $\dim_K V_i = n_i$. Dann existiert eine Orthonormalbasis $e_1, \dots, e_n$ von V , so dass $e_1, \dots, e_{n_i}$ jeweils eine Orthonormalbasis von V_i ist, $i = 1, \dots, r$.*

Ist V ein $\mathbb{K}$ -Vektorraum mit einer sBF oder HF, so heißen zwei Teilmengen $M, N \subset V$ orthogonal, in Zeichen $M \perp N$, wenn stets $\langle x, y \rangle = 0$ für $x \in M$, $y \in N$ gilt. Man schreibt dabei auch $x \perp y$ anstelle von $\langle x, y \rangle = 0$, wobei $x \perp y$ äquivalent zu $y \perp x$ ist. Außerdem kann man zu einer Teilmenge $M \subset V$ den $\mathbb{K}$ -Untervektorraum

$$M^\perp = \{x \in V; x \perp y \text{ für alle } y \in M\}$$

betrachten. Für einen Untervektorraum $W \subset V$ bezeichnet man $W^\perp$ als das *orthogonale Komplement* von W in V .

Korollar 8. *Sei V ein endlich-dimensionaler euklidischer bzw. unitärer Vektorraum, und sei $W \subset V$ ein Untervektorraum. Dann gilt:*

- (i) $V = W \oplus W^\perp$, insbesondere $\dim_K W^\perp = \dim_K V - \dim_K W$.
- (ii) $(W^\perp)^\perp = W$.

Beweis. Man wähle eine Orthonormalbasis $e_1, \dots, e_r$ von W , und ergänze diese durch Elemente $e_{r+1}, \dots, e_n$ zu einer Orthonormalbasis von V ; vgl. Korollar 6. Für $W' = \bigoplus_{i=r+1}^n \mathbb{K}e_i$ gilt dann $W \perp W'$ und deshalb $W' \subset W^\perp$. Sei nun $x = \sum_{i=1}^n \alpha_i e_i \in W^\perp$. Die Gleichungen $\langle x, e_i \rangle = 0$ für $i = 1, \dots, r$ zeigen dann $\alpha_i = 0$ für $i = 1, \dots, r$ und somit $x \in W'$. Es gilt also $W^\perp = W'$ und damit $V = W \oplus W^\perp$. Die gleiche Argumentation, angewandt auf $W^\perp$ anstelle von W , ergibt $(W^\perp)^\perp = W$. $\square$

Als Anwendung wollen wir noch auf das Volumen eines Parallelotops im $\mathbb{R}^n$ eingehen. Für ein linear unabhängiges System von Vektoren $x_1, \dots, x_r \in \mathbb{R}^n$ bezeichne

$$P(x_1, \dots, x_r) = \{x \in \mathbb{R}^n; x = \sum_{i=1}^r \alpha_i x_i \text{ mit } 0 \leq \alpha_i \leq 1\}$$

das von diesen Vektoren aufgespannte r -dimensionale *Parallelotop*. Handelt es sich für $r = n$ bei $x_1, \dots, x_n$ beispielsweise um die Einheitsvektoren im $\mathbb{R}^n$, so ist $P(x_1, \dots, x_n)$ gerade der n -dimensionale Einheitswürfel.

Wir fassen, wie üblich, $\mathbb{R}^n$ als euklidischen Vektorraum mit dem kanonischen Skalarprodukt auf. Zu einem r -dimensionalen Parallelotop $P(x_1, \dots, x_r)$ betrachte man den von $x_1, \dots, x_r$ erzeugten r -dimensionalen Untervektorraum $U \subset \mathbb{R}^n$ und wähle eine Orthonormalbasis $M = (e_1, \dots, e_r)$ in U . Sodann sei $\det_M$ diejenige Determinantenfunktion auf U , die auf der Basis M den Wert 1 annimmt; vgl. 4.2/8. In dieser Situation wird das *Volumen* des Parallelotops definiert durch

$$\text{Vol}(P(x_1, \dots, x_r)) = |\det_M(x_1, \dots, x_r)|.$$

Natürlich ist zu zeigen, dass diese Definition unabhängig von der Wahl der Orthonormalbasis M von U ist. Auch wollen wir plausibel machen, dass das Volumen mit der anschaulichen Vorstellung des Volumens eines Körpers im $\mathbb{R}^n$ übereinstimmt.

Satz 9. Für Vektoren $x_1, \dots, x_r \in \mathbb{K}^n$ sei die zugehörige Gramsche Determinante definiert durch

$$G(x_1, \dots, x_r) = \det \begin{pmatrix} \langle x_1, x_1 \rangle & \dots & \langle x_1, x_r \rangle \\ \vdots & \ddots & \vdots \\ \langle x_r, x_1 \rangle & \dots & \langle x_r, x_r \rangle \end{pmatrix}$$

mit $\langle \cdot, \cdot \rangle$ als kanonischem Skalarprodukt auf $\mathbb{K}^n$. Es gilt $G(x_1, \dots, x_r) \geq 0$, wobei $G(x_1, \dots, x_r)$ genau dann verschwindet, wenn $x_1, \dots, x_r$ linear abhängig sind.

Sind $x_1, \dots, x_r$ linear unabhängig und ist $M = (e_1, \dots, e_r)$ eine Orthonormalbasis des von $x_1, \dots, x_r$ erzeugten linearen Unterraums im $\mathbb{K}^n$, so besteht die Beziehung

$$G(x_1, \dots, x_r) = |\det_M(x_1, \dots, x_r)|^2.$$

Man beachte, dass die Aussage für $r = 2$ gerade die Schwarzsche Ungleichung 7.1/4 ergibt, und zwar unabhängig von dem Beweis, der in 7.1/4 gegeben wurde. Weiter können wir feststellen:

Korollar 10. Das Volumen eines Parallelotops $P(x_1, \dots, x_r) \subset \mathbb{R}^n$ ist wohldefiniert, es gilt

$$\text{Vol}(P(x_1, \dots, x_r)) = G(x_1, \dots, x_r)^{\frac{1}{2}}.$$

Beweis zu Satz 9. Sind $x_1, \dots, x_r$ linear abhängig, so sieht man unmittelbar, dass die Spalten bzw. Zeilen in der Matrix der Gramschen Determinante linear abhängig sind, also $G(x_1, \dots, x_r) = 0$ gilt. Seien daher $x_1, \dots, x_r$ linear unabhängig. Dann erzeugen diese Vektoren einen r -dimensionalen Untervektorraum $U \subset \mathbb{K}^n$, der wiederum mit einem Skalarprodukt versehen ist, und wir können gemäß Satz 5 eine Orthonormalbasis $M = (e_1, \dots, e_r)$ in U wählen. Bezeichnen wir dann mit $x_{1,M}, \dots, x_{r,M}$ die Koordinatenspaltenvektoren von $x_1, \dots, x_r$ bezüglich der Basis M , so gilt

$$(x_{1,M}, \dots, x_{r,M})^t \cdot (\bar{x}_{1,M}, \dots, \bar{x}_{r,M}) = (\langle x_i, x_j \rangle)_{i,j=1,\dots,r}.$$

Insbesondere folgt

$$|\det_M(x_1, \dots, x_r)|^2 = |\det(x_{1,M}, \dots, x_{r,M})|^2 = G(x_1, \dots, x_r)$$

und damit $G(x_1, \dots, x_r) > 0$, da $x_1, \dots, x_r$ linear unabhängig sind. $\square$

Wir wollen nun noch plausibel machen, dass das definierte Volumen eines Parallelotops

$$\text{Vol}(P(x_1, \dots, x_r)) = |\det_M(x_1, \dots, x_r)| = G(x_1, \dots, x_r)^{\frac{1}{2}}$$

(mit einer Orthonormalbasis $M = (e_1, \dots, e_r)$ des von $x_1, \dots, x_r$ erzeugten Untervektorraums von $\mathbb{R}^n$) auch in anschaulicher Weise dem r -dimensionalen Volumen von $P(x_1, \dots, x_r)$ entspricht. Für $r = 1$ ist dies unmittelbar klar. Für

$r > 1$ betrachte man die orthogonale Projektion $p_U: V \longrightarrow U := \sum_{i=1}^{r-1} \mathbb{R}x_i$. Es ist dann $x'_r = x_r - p_U(x_r)$ orthogonal zu U , also zu $x_1, \dots, x_{r-1}$. Folglich gilt

$$\begin{aligned}\text{Vol}(P(x_1, \dots, x_r)) &= |\det_M(x_1, \dots, x_r)| \\ &= |\det_M(x_1, \dots, x_{r-1}, x'_r)| = \text{Vol}(P(x_1, \dots, x_{r-1}, x'_r))\end{aligned}$$

und weiter

$$\begin{aligned}\text{Vol}(P(x_1, \dots, x_r))^2 &= \text{Vol}(P(x_1, \dots, x_{r-1}, x'_r))^2 \\ &= \det \begin{pmatrix} \langle x_1, x_1 \rangle & \dots & \langle x_1, x_{r-1} \rangle & 0 \\ \vdots & \ddots & \vdots & \vdots \\ \langle x_{r-1}, x_1 \rangle & \dots & \langle x_{r-1}, x_{r-1} \rangle & 0 \\ 0 & \dots & 0 & \langle x'_r, x'_r \rangle \end{pmatrix} \\ &= \det \begin{pmatrix} \langle x_1, x_1 \rangle & \dots & \langle x_1, x_{r-1} \rangle \\ \vdots & \ddots & \vdots \\ \langle x_{r-1}, x_1 \rangle & \dots & \langle x_{r-1}, x_{r-1} \rangle \end{pmatrix} \cdot |x'_r|^2 \\ &= \text{Vol}(P(x_1, \dots, x_{r-1}))^2 \cdot |x'_r|^2.\end{aligned}$$

Nun ist $|x'_r|$ als senkrechter Abstand von x_r zu U zu interpretieren; vgl. hierzu auch Aufgabe 4. Somit ergibt sich das Volumen des r -dimensionalen Parallelotops $P(x_1, \dots, x_r)$ als Produkt aus dem Volumen der $(r-1)$ -dimensionalen “Grundfläche” $P(x_1, \dots, x_{r-1})$ mit der “Höhe” von x_r über dieser Grundfläche. In induktiver Weise folgt daher, dass das definierte Volumen eines Parallelotops mit dem eines Quaders übereinstimmt, der die gleichen Höhenverhältnisse hat, was mit der anschaulichen Vorstellung übereinstimmt.

Aufgaben

1. Man betrachte $\mathbb{R}^3$ als euklidischen Vektorraum mit dem kanonischen Skalarprodukt und wende das Orthonormalisierungsverfahren von E. Schmidt auf die Basis $(1, 1, 0), (1, 0, 1), (0, 1, 1) \in \mathbb{R}^3$ an.
2. Für $n \in \mathbb{N}$ sei $\mathbb{R}[T]_n \subset \mathbb{R}[T]$ der $\mathbb{R}$ -Untervektorraum aller Polynome vom Grad $\leq n$. Man zeige, dass durch

$$\langle f, g \rangle = \int_0^1 f(t)g(t)dt, \quad f, g \in \mathbb{R}[T]_n,$$

ein positiv definites Skalarprodukt auf $\mathbb{R}[T]_n$ definiert wird. Für $n = 2$ wende man das E. Schmidtsche Orthonormalisierungsverfahren auf die Basis $1, T, T^2$ von $\mathbb{R}[T]_2$ an.

3. Es sei V ein euklidischer bzw. unitärer $\mathbb{K}$ -Vektorraum mit Basis $x_1, \dots, x_n$, aus der man durch Anwenden des E. Schmidtschen Orthonormalisierungsverfahrens die Orthonormalbasis $e_1, \dots, e_n$ erhalte. Man zeige für Konstanten $\varepsilon_1, \dots, \varepsilon_n \in \mathbb{K}$ mit $|\varepsilon_i| = 1$, dass das Orthonormalisierungsverfahren die Basis $\varepsilon_1 x_1, \dots, \varepsilon_n x_n$ in die Orthonormalbasis $\varepsilon_1 e_1, \dots, \varepsilon_n e_n$ überführt.

4. Es sei V ein euklidischer bzw. unitärer $\mathbb{K}$ -Vektorraum, $U \subset V$ ein Untervektorraum und $u \in V - U$. Man zeige:
- (i) Es existiert genau ein $u_0 \in U$ mit $v - u_0 \in U^\perp$.
 - (ii) Für alle $u \in U$ mit $u \neq u_0$ gilt $|v - u| > |v - u_0|$.
5. Es sei $V = \mathbb{R}^{n \times n}$ der $\mathbb{R}$ -Vektorraum aller reellen $(n \times n)$ -Matrizen. Man zeige:
- (i) Durch $\Phi(A, B) = \text{Spur}(A \cdot B)$ wird auf V eine nicht-ausgeartete symmetrische Bilinearform erklärt.
 - (ii) Sei $U_+ = \{U \in V; U^t = U\}$ der Untervektorraum aller symmetrischen und $U_- = \{U \in V; U^t = -U\}$ der Untervektorraum aller schiefsymmetrischen Matrizen. Es gilt

$$V = U_+ \oplus U_-, \quad U_+^\perp = U_-, \quad U_-^\perp = U_+.$$

- (iii) Es ist Φ positiv definit auf U_+ und negativ definit auf U_- , d. h. es gilt $\Phi(A, A) > 0$ für alle $A \in V_+ - \{0\}$ und $\Phi(A, A) < 0$ für alle $A \in V_- - \{0\}$.

7.3 Sesquilinearformen und Matrizen

Für eine Matrix $A = (\alpha_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} \in \mathbb{K}^{m \times n}$ bezeichnet man mit

$$\overline{A} = (\overline{\alpha_{ij}})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} \in \mathbb{K}^{m \times n}$$

die *konjugierte* Matrix, mit

$$A^t = (\alpha_{ij})_{\substack{j=1,\dots,n \\ i=1,\dots,m}} \in \mathbb{K}^{n \times m}$$

die *transponierte* Matrix, sowie mit

$$A^* = \overline{A}^t = (\overline{\alpha_{ij}})_{\substack{j=1,\dots,n \\ i=1,\dots,m}} \in \mathbb{K}^{n \times m}$$

die *adjungierte* Matrix zu A . Dabei ist zu beachten, dass hier die Bezeichnung “adjungiert” in einem anderen Sinne als in Abschnitt 4.4 gemeint ist. Für das Rechnen mit konjugierten Matrizen gelten folgende Regeln; A, B seien Matrizen, $c \in \mathbb{K}$ sei eine Konstante:

$$\begin{aligned} \overline{A + B} &= \overline{A} + \overline{B} \\ \overline{c \cdot A} &= \overline{c} \cdot \overline{A} \\ \overline{A \cdot B} &= \overline{A} \cdot \overline{B} \\ \overline{A^{-1}} &= (\overline{A})^{-1} \\ \det \overline{A} &= \overline{\det A} \end{aligned}$$

Für das Transponieren von Matrizen hatten wir bereits die folgenden Regeln kennen gelernt:

$$\begin{aligned}
(A+B)^t &= A^t + B^t \\
(c \cdot A)^t &= c \cdot A^t \\
(A \cdot B)^t &= B^t \cdot A^t \\
(A^{-1})^t &= (A^t)^{-1} \\
\det A^t &= \det A
\end{aligned}$$

Und zwar ergeben sich die ersten beiden Gleichungen aus 3.2/6, die dritte, sowie als leichte Folgerung auch die vierte aus 3.2/8, und schließlich die letzte aus 4.3/4. Somit ergeben sich folgende Regeln für das Rechnen mit adjungierten Matrizen:

$$\begin{aligned}
(A+B)^* &= A^* + B^* \\
(c \cdot A)^* &= \bar{c} \cdot A^* \\
(A \cdot B)^* &= B^* \cdot A^* \\
(A^{-1})^* &= (A^*)^{-1} \\
\det A^* &= \overline{\det A}
\end{aligned}$$

Wir wollen im Folgenden Sesquilinearformen mit Hilfe von Matrizen beschreiben.

Definition 1. Sei $\Phi: V \times V \longrightarrow \mathbb{K}$ eine Sesquilinearform auf einem $\mathbb{K}$ -Vektorraum V mit Basis $X = (x_1, \dots, x_n)$. Dann heißt

$$A_{\Phi, X} = (\Phi(x_i, x_j))_{i,j=1, \dots, n} \in \mathbb{K}^{n \times n}$$

die zu Φ gehörige Matrix bezüglich der Basis X .

Satz 2. Sei $\Phi: V \times V \longrightarrow \mathbb{K}$ eine Sesquilinearform auf einem $\mathbb{K}$ -Vektorraum V mit Basis $X = (x_1, \dots, x_n)$. Bezeichnet dann wie üblich a_X den Koordinatenspaltenvektor zu einem Vektor $a \in V$, so gilt für $a, b \in V$

$$\Phi(a, b) = a_X^t \cdot A_{\Phi, X} \cdot \bar{b}_X,$$

und die Matrix $A_{\Phi, X} \in \mathbb{K}^{n \times n}$ ist durch diese Beziehung eindeutig charakterisiert. Weiter ist Φ genau dann nicht ausgeartet, wenn $\det A_{\Phi, X} \neq 0$ gilt.

Beweis. Sei $a = \sum_{i=1}^n \alpha_i x_i$, $b = \sum_{j=1}^n \beta_j x_j$. Dann folgt

$$\Phi(a, b) = \sum_{i,j=1}^n \alpha_i \cdot \Phi(x_i, x_j) \cdot \bar{\beta}_j = a_X^t \cdot A_{\Phi, X} \cdot \bar{b}_X,$$

wie behauptet. Hat man andererseits eine Matrix $A = (\alpha_{ij})_{i,j} \in \mathbb{K}^{n \times n}$ mit

$$\Phi(a, b) = a_X^t \cdot A \cdot \bar{b}_X,$$

für $a, b \in V$, so ergibt sich, indem man $x_1, \dots, x_n$ für a bzw. b einsetzt,

$$\Phi(x_i, x_j) = \alpha_{ij}$$

und damit $A = A_{\Phi, X}$, d. h. die Matrix $A = A_{\Phi, X}$ ist durch obige Beziehung eindeutig bestimmt.

Sei nun Φ ausgeartet, etwa ausgeartet im ersten Argument. Sei also $a \in V$ von Null verschieden mit $\Phi(a, b) = 0$ für alle $b \in V$. Dann gilt

$$a_X^t \cdot A_{\Phi, X} \cdot \bar{b}_X = 0 \quad \text{für alle } b \in V.$$

Indem man dies für $b = x_1, \dots, x_n$ anwendet, erhält man $a_X^t \cdot A_{\Phi, X} = 0$. Die $\mathbb{K}$ -lineare Abbildung

$$\mathbb{K}^n \longrightarrow \mathbb{K}^n, \quad x \longmapsto A_{\Phi, X}^t \cdot x,$$

hat daher einen nicht-trivialen Kern, und es ergibt sich

$$\det A_{\Phi, X} = \det A_{\Phi, X}^t = 0$$

mit 4.3/4. Umgekehrt folgt mittels 2.1/11 und 4.3/4 aus einer solchen Gleichung, dass Φ im ersten Argument ausgeartet ist. Der Fall, dass Φ im zweiten Argument ausgeartet ist, lässt sich entsprechend behandeln. $\square$

Der vorstehende Beweis zeigt genauer:

Korollar 3. *Sei V ein endlich-dimensionaler $\mathbb{K}$ -Vektorraum mit Basis X . Dann definiert die Zuordnung $\Phi \mapsto A_{\Phi, X}$ eine bijektive Abbildung zwischen der Menge aller Sesquilinearformen $\Phi: V \times V \longrightarrow \mathbb{K}$ und der Menge der $(n \times n)$ -Matrizen $\mathbb{K}^{n \times n}$. Weiter ist für eine solche Sesquilinearform Φ äquivalent:*

- (i) $\Phi(a, b) = 0$ für alle $b \in V$ impliziert $a = 0$.
- (ii) $\Phi(a, b) = 0$ für alle $a \in V$ impliziert $b = 0$.
- (iii) Φ ist nicht-ausgeartet.
- (iv) $\det A_{\Phi, X} \neq 0$.

Korollar 4. *Sei $\Phi: V \times V \longrightarrow \mathbb{K}$ eine Sesquilinearform auf einem endlich-dimensionalen $\mathbb{K}$ -Vektorraum V mit Basis X . Dann ist äquivalent:*

- (i) Φ ist eine sBF bzw. HF, d. h. für alle $a, b \in V$ gilt $\Phi(a, b) = \overline{\Phi(b, a)}$.
- (ii) $A_{\Phi, X} = (A_{\Phi, X})^*$.

Beweis. Für $a, b \in V$ gilt

$$\begin{aligned} \overline{\Phi(b, a)} &= \overline{(b_X^t \cdot A_{\Phi, X} \cdot \bar{a}_X)} = \bar{b}_X^t \cdot \overline{A_{\Phi, X}} \cdot a_X \\ &= (\bar{b}_X^t \cdot \overline{A_{\Phi, X}} \cdot a_X)^t = a_X^t \cdot (A_{\Phi, X})^* \cdot \bar{b}_X. \end{aligned}$$

Nach Satz 2 ist die Gleichung $\Phi(a, b) = \overline{\Phi(b, a)}$ für $a, b \in V$ daher äquivalent zu $A_{\Phi, X} = (A_{\Phi, X})^*$. $\square$

Als Nächstes wollen wir untersuchen, wie sich ein Basiswechsel in V auf die beschreibende Matrix $A_{\Phi,X}$ einer Sesquilinearform $\Phi: V \times V \longrightarrow \mathbb{K}$ auswirkt. In Abschnitt 3.4 hatten wir Basiswechselmatrizen der Form $A_{\text{id},Y,X}$ zu gegebenen Basen Y und X von V betrachtet. Wir werden im Folgenden anstelle von $A_{\text{id},Y,X}$ abkürzend $A_{Y,X}$ schreiben.

Satz 5. *Sei $\Phi: V \times V \longrightarrow \mathbb{K}$ eine Sesquilinearform auf einem endlich-dimensionalen $\mathbb{K}$ -Vektorraum V mit Basen X und Y . Dann gilt*

$$A_{\Phi,Y} = A_{Y,X}^t \cdot A_{\Phi,X} \cdot \overline{A_{Y,X}}.$$

Beweis. Man hat $a_X = A_{Y,X} \cdot a_Y$ für $a \in V$, folglich für $a, b \in V$

$$\begin{aligned} \Phi(a, b) &= a_X^t \cdot A_{\Phi,X} \cdot \overline{b_X} \\ &= (A_{Y,X} \cdot a_Y)^t \cdot A_{\Phi,X} \cdot (\overline{A_{Y,X} \cdot b_Y}) \\ &= a_Y^t \cdot A_{Y,X}^t \cdot A_{\Phi,X} \cdot \overline{A_{Y,X}} \cdot \overline{b_Y} \end{aligned}$$

und deshalb $A_{\Phi,Y} = A_{Y,X}^t \cdot A_{\Phi,X} \cdot \overline{A_{Y,X}}$ gemäß Satz 2. $\square$

Korollar 6. *Sei $\Phi: V \times V \longrightarrow \mathbb{K}$ eine Sesquilinearform auf einem $\mathbb{K}$ -Vektorraum V mit Basis $X = (x_1, \dots, x_n)$. Dann ist äquivalent:*

- (i) Φ ist ein Skalarprodukt, also eine positiv definite sBF bzw. HF.
- (ii) Es existiert eine Matrix $S \in \text{Gl}(n, \mathbb{K})$, so dass gilt:

$$S^t \cdot A_{\Phi,X} \cdot \overline{S} = E$$

Dabei ist $E \in \mathbb{K}^{n \times n}$ die Einheitsmatrix.

Beweis. Sei zunächst Φ ein Skalarprodukt. Dann besitzt V nach 7.2/5 eine Orthonormalbasis Y . Es gilt also $A_{\Phi,Y} = E$, und man erhält $S^t \cdot A_{\Phi,X} \cdot \overline{S} = E$ mit $S = A_{Y,X}$ aus Satz 5.

Gilt umgekehrt $S^t \cdot A_{\Phi,X} \cdot \overline{S} = E$ für ein $S \in \text{Gl}(n, \mathbb{K})$, so kann man S als Basiswechselmatrix des Typs $A_{Y,X}$ auffassen, so dass also $A_{\Phi,Y} = E$ gilt. Φ ist dann ein Skalarprodukt. $\square$

Korollar 7. *Ist $\Phi: V \times V \longrightarrow \mathbb{K}$ ein Skalarprodukt auf einem endlich-dimensionalen $\mathbb{K}$ -Vektorraum V , so gilt $\det A_{\Phi,X} > 0$ für alle Basen X von V .*

Beweis. Sei X eine Orthonormalbasis von V ; vgl. 7.2/5, es gilt $A_{\Phi,X} = E$. Für eine weitere Basis Y von V erhält man unter Verwendung der Gleichung aus Satz 5

$$\begin{aligned} \det A_{\Phi,Y} &= \det A_{Y,X}^t \cdot \det A_{\Phi,X} \cdot \det \overline{A_{Y,X}} \\ &= \det A_{Y,X} \cdot \overline{\det A_{Y,X}} = |\det A_{Y,X}|^2 > 0, \end{aligned}$$

da $\det A_{Y,X} \neq 0$. $\square$

Abschließend wollen wir noch ein Determinantenkriterium für die positive Definitheit einer sBF bzw. HF geben.

Satz 8. *Sei $\Phi: V \times V \longrightarrow \mathbb{K}$ eine sBF bzw. HF auf einem $\mathbb{K}$ -Vektorraum V mit Basis $X = (x_1, \dots, x_n)$. Man betrachte die Matrizen*

$$A_r = (\Phi(x_i, x_j))_{i,j=1,\dots,r} \in \mathbb{K}^{r \times r}, \quad r = 1, \dots, n.$$

Dann ist äquivalent:

- (i) Φ ist positiv definit und damit ein Skalarprodukt.
- (ii) $\det A_r > 0$ für $r = 1, \dots, n$.

Beweis. Die Implikation (i) $\implies$ (ii) ist leicht einzusehen. Man schränke Φ auf die Untervektorräume $V_r = \sum_{i=1}^r \mathbb{K}x_i$ ein, $r = 1, \dots, n$. Korollar 7 zeigt dann $\det A_r > 0$.

Zum Nachweis der Umkehrung nehmen wir $\det A_r > 0$ an für $r = 1, \dots, n$ und zeigen mit Induktion nach n , dass Φ ein Skalarprodukt ist. Der Fall $n = 1$ ist trivial, da dann x_1 eine Basis von V ist und $\Phi(x_1, x_1) = \det A_1 > 0$ gilt. Sei also $n > 1$. Dann ist $\Phi|_{V_{n-1}}$ positiv definit nach Induktionsvoraussetzung, und es besitzt V_{n-1} nach 7.2/5 eine Orthonormalbasis $e_1, \dots, e_{n-1}$. Weiter ist mit 7.2/4 leicht zu sehen, dass $e_1, \dots, e_{n-1}$ zusammen mit

$$x'_n = x_n - \sum_{i=1}^{n-1} \Phi(x_n, e_i) e_i$$

eine Orthogonalbasis Y von V bilden, wobei

$$A_{\Phi,Y} = \begin{pmatrix} 1 & & & 0 \\ & \ddots & & \\ & & \ddots & \\ 0 & & & 1 & 0 \\ & & & 0 & \Phi(x'_n, x'_n) \end{pmatrix}$$

gilt. Mit $S = A_{Y,X}$ folgt dann $A_{\Phi,Y} = S^t A_n \overline{S}$ aus Satz 5 und damit

$$\Phi(x'_n, x'_n) = \det A_{\Phi,Y} = |\det S|^2 \cdot \det A_n > 0.$$

Setzen wir daher

$$e_n = \frac{1}{\sqrt{\Phi(x'_n, x'_n)}} \cdot x'_n,$$

so gilt $\Phi(e_n, e_n) = 1$, und es bilden $e_1, \dots, e_n$ eine Orthonormalbasis von V . Insbesondere ist Φ positiv definit. $\square$

In der Situation von Satz 8 bezeichnet man die Determinanten $\det A_r$, $r = 1, \dots, n$, auch als die *Hauptunterdeterminanten* der Matrix $A_{\Phi,X}$. Durch Kombination von Korollar 4 mit Satz 8 ergibt sich dann:

Korollar 9. Für eine Matrix $A \in \mathbb{K}^{n \times n}$ betrachte man die Sesquilinearform

$$\Phi: \mathbb{K}^n \times \mathbb{K}^n \longrightarrow \mathbb{K}, \quad (a, b) \longmapsto a^t \cdot A \cdot \bar{b}.$$

Dann ist äquivalent:

- (i) Φ ist ein Skalarprodukt.
- (ii) $A = A^*$ und alle Hauptunterdeterminanten von A sind positiv.

Aufgaben

V sei stets ein $\mathbb{K}$ -Vektorraum der Dimension $n < \infty$.

- Man zeige, dass die Menge aller Sesquilinearformen $V \times V \longrightarrow \mathbb{K}$ unter der Addition und skalaren Multiplikation von $\mathbb{K}$ -wertigen Funktionen auf $V \times V$ einen $\mathbb{K}$ -Vektorraum und insbesondere auch einen $\mathbb{R}$ -Vektorraum bildet. Man berechne jeweils die Dimension. Welche der folgenden Teilmengen bilden lineare Unterräume über $\mathbb{R}$ bzw. $\mathbb{C}$? Man berechne gegebenenfalls die zugehörige Dimension.
 - (i) Symmetrische Bilinearformen im Falle $\mathbb{K} = \mathbb{R}$
 - (ii) Hermiteische Formen im Falle $\mathbb{K} = \mathbb{C}$
 - (iii) Skalarprodukte
- Im Falle $\dim_{\mathbb{K}} V \geq 2$ konstruiere man eine nicht-ausgeartete sBF bzw. HF Φ auf V sowie einen nicht-trivialen linearen Unterraum $U \subset V$, so dass $\Phi|_{U \times U}$ ausgeartet ist.
- Es sei Φ eine positiv semi-definite sBF bzw. HF auf V . Man zeige, dass es eine Basis X von V gibt, so dass die zugehörige Matrix $A_{\Phi, X}$ eine Diagonalmatrix mit Diagonaleinträgen 1 oder 0 ist, also mit $A_{\Phi, X} = \text{Diag}(1, \dots, 1, 0, \dots, 0)$.
- Es sei Φ eine sBF bzw. HF auf V und X eine Basis von V . Man zeige, dass alle Hauptunterdeterminanten von $A_{\Phi, X}$ reelle Zahlen ≥ 0 sind, falls Φ positiv semi-definit ist. Gilt auch die Umkehrung?
- Für eine Matrix $A \in \mathbb{K}^{n \times n}$ gilt genau dann $A^* = A^{-1}$, wenn die Spalten (bzw. Zeilen) von A eine Orthonormalbasis in $\mathbb{K}^n$ bilden. Eine solche Matrix wird als *orthogonal* (für $\mathbb{K} = \mathbb{R}$) bzw. *unitär* (für $\mathbb{K} = \mathbb{C}$) bezeichnet.
- Es sei Φ ein Skalarprodukt auf V . Man zeige für jedes weitere Skalarprodukt Ψ auf V : Es existiert ein Endomorphismus $f: V \longrightarrow V$ mit $\Psi(x, y) = \Phi(f(x), f(y))$ für alle $x, y \in V$.
- Für eine Matrix $A \in \mathbb{K}^{n \times n}$ zeige man: Durch $\langle a, b \rangle = a^t \cdot A \cdot \bar{b}$ wird genau dann ein Skalarprodukt auf $\mathbb{K}^n$ definiert, wenn es eine Matrix $S \in \text{Gl}(n, \mathbb{K})$ mit $A = S^t \cdot \bar{S}$ gibt.

7.4 Die adjungierte Abbildung

Als Nächstes soll zu einem Endomorphismus $\varphi: V \longrightarrow V$ eines endlich-dimensionalen euklidischen bzw. unitären Vektorraums V die so genannte *adjungierte* Abbildung $\varphi^*: V \longrightarrow V$ definiert werden. Diese Abbildung lässt sich

in einem gewissen Sinne als duale Abbildung zu φ interpretieren, weshalb sie meist mit φ^* bezeichnet wird. Allerdings werden wir zusätzlich aus Konstruktionsgründen auch die “echte” duale Abbildung zu φ (im Sinne von 2.3/2) benötigen, für die wir im Folgenden anstelle von $\varphi^*: V^* \longrightarrow V^*$ die Notation $\varphi': V' \longrightarrow V'$ verwenden werden. Wir beginnen mit einer technischen Vorbetrachtung.

Zu einem $\mathbb{K}$ -Vektorraum V kann man wie folgt einen $\mathbb{K}$ -Vektorraum $\overline{V}$ bilden. Man setze $\overline{V} = V$ als additive abelsche Gruppe, definiere aber die skalare Multiplikation von $\overline{V}$ durch

$$\mathbb{K} \times \overline{V} \longrightarrow \overline{V}, \quad (\alpha, v) \longmapsto \alpha \bullet v := \overline{\alpha} \cdot v,$$

wobei zur Bildung des Produktes $\overline{\alpha} \cdot v$ die skalare Multiplikation von V verwendet werden soll. Man prüft leicht nach, dass $\overline{V}$ auf diese Weise ein $\mathbb{K}$ -Vektorraum ist und dass die $\mathbb{K}$ -Endomorphismen von V mit den $\mathbb{K}$ -Endomorphismen von $\overline{V}$ übereinstimmen. Weiter gilt $\overline{(\overline{V})} = V$, $\dim_{\mathbb{K}}(\overline{V}) = \dim_{\mathbb{K}}(V)$ und natürlich $V = \overline{V}$ für $\mathbb{K} = \mathbb{R}$.

Lemma 1. *Sei $\Phi: V \times V \longrightarrow \mathbb{K}$ eine nicht-ausgeartete Sesquilinearform auf einem endlich-dimensionalen $\mathbb{K}$ -Vektorraum V . Ist dann $V' = \text{Hom}_{\mathbb{K}}(V, \mathbb{K})$ der Dualraum von V , so wird durch*

$$\tau: \overline{V} \longrightarrow V', \quad x \longmapsto \Phi(\cdot, x),$$

ein Isomorphismus von $\mathbb{K}$ -Vektorräumen erklärt.

Beweis. Für $x, y \in \overline{V}$ gilt

$$\tau(x + y) = \Phi(\cdot, x + y) = \Phi(\cdot, x) + \Phi(\cdot, y) = \tau(x) + \tau(y),$$

sowie für $\alpha \in \mathbb{K}$, $x \in \overline{V}$

$$\tau(\alpha \bullet x) = \tau(\overline{\alpha} \cdot x) = \Phi(\cdot, \overline{\alpha} \cdot x) = \alpha \cdot \Phi(\cdot, x) = \alpha \cdot \tau(x),$$

d. h. φ ist $\mathbb{K}$ -linear. Weiter ist τ injektiv, da Φ nicht ausgeartet ist, und es gilt $\dim_{\mathbb{K}}(\overline{V}) = \dim_{\mathbb{K}}(V) = \dim_{\mathbb{K}}(V')$ gemäß 2.3/6. Dann ist φ aber aufgrund von 2.1/11 ein Isomorphismus. $\square$

Satz 2. *Sei $\Phi: V \times V \longrightarrow \mathbb{K}$ eine nicht-ausgeartete Sesquilinearform auf einem endlich-dimensionalen $\mathbb{K}$ -Vektorraum V , und sei $\varphi \in \text{End}_{\mathbb{K}}(V)$. Dann existiert eine eindeutig bestimmte Abbildung $\varphi^* \in \text{End}_{\mathbb{K}}(V)$ mit*

$$\Phi(\varphi(x), y) = \Phi(x, \varphi^*(y))$$

für alle $x, y \in V$. Man nennt φ^ den zu φ adjungierten Endomorphismus.*

Beweis. Wir betrachten den Dualraum V' zu V sowie die von φ induzierte duale Abbildung

$$\varphi': V' \longrightarrow V', \quad f \longmapsto f \circ \varphi.$$

Indem wir den Isomorphismus $\tau: \overline{V} \xrightarrow{\sim} V'$ aus Lemma 1 benutzen, können wir durch $\varphi^* := \tau^{-1} \circ \varphi' \circ \tau$ einen $\mathbb{K}$ -Endomorphismus φ^* von $\overline{V}$ bzw. V definieren, so dass folglich das Diagramm

$$\begin{array}{ccc} \overline{V} & \xrightarrow{\tau} & V' \\ \varphi^* \downarrow & & \downarrow \varphi' \\ \overline{V} & \xrightarrow{\tau} & V' \end{array}$$

kommutiert. Somit gilt für $y \in V$

$$\Phi(\varphi(\cdot), y) = \tau(y) \circ \varphi = \varphi'(\tau(y)) = \tau(\varphi^*(y)) = \Phi(\cdot, \varphi^*(y)),$$

also wie gewünscht

$$\Phi(\varphi(x), y) = \Phi(x, \varphi^*(y))$$

für alle $x, y \in V$. Dass $\varphi^* \in \text{End}_{\mathbb{K}}(V)$ durch diese Beziehung eindeutig bestimmt ist, folgert man leicht aus der Tatsache, dass Φ nicht ausgeartet ist. $\square$

Korollar 3. Sei $\Phi: V \times V \longrightarrow \mathbb{K}$ eine nicht-ausgeartete sBF bzw. HF eines endlich-dimensionalen $\mathbb{K}$ -Vektorraums V .

(i) Die Abbildung

$$*: \text{End}_{\mathbb{K}}(V) \longrightarrow \text{End}_{\mathbb{K}}(V), \quad \varphi \longmapsto \varphi^*,$$

ist semi-linear, d. h. man hat

$$(\alpha\varphi_1 + \beta\varphi_2)^* = \overline{\alpha}\varphi_1^* + \overline{\beta}\varphi_2^*$$

für $\alpha, \beta \in \mathbb{K}, \varphi_1, \varphi_2 \in \text{End}_{\mathbb{K}}(V)$. Weiter gilt $\text{id}^* = \text{id}$, sowie $\varphi^{**} = \varphi$ für alle $\varphi \in \text{End}_{\mathbb{K}}(V)$.

(ii) Es gilt

$$\ker \varphi^* = (\text{im } \varphi)^\perp, \quad \ker \varphi = (\text{im } \varphi^*)^\perp$$

für $\varphi \in \text{End}_{\mathbb{K}}(V)$.

(iii) Es gilt $\text{rg } \varphi = \text{rg } \varphi^*$ für $\varphi \in \text{End}_{\mathbb{K}}(V)$. Insbesondere ist φ genau dann bijektiv, wenn φ^* bijektiv ist.

Beweis. (i) Für $x, y \in V$ gilt

$$\begin{aligned} \langle (\alpha\varphi_1 + \beta\varphi_2)(x), y \rangle &= \alpha\langle \varphi_1(x), y \rangle + \beta\langle \varphi_2(x), y \rangle \\ &= \alpha\langle x, \varphi_1^*(y) \rangle + \beta\langle x, \varphi_2^*(y) \rangle \\ &= \langle x, (\overline{\alpha}\varphi_1 + \overline{\beta}\varphi_2)(y) \rangle \end{aligned}$$

und damit $(\alpha\varphi_1 + \beta\varphi_2)^* = \overline{\alpha}\varphi_1^* + \overline{\beta}\varphi_2^*$ gemäß Satz 2. In gleicher Weise zeigt

$$\langle x, y \rangle = \langle \text{id}(x), y \rangle = \langle x, \text{id}^*(y) \rangle$$

die Gleichung $\text{id}^* = \text{id}$, sowie

$$\langle \varphi^*(x), y \rangle = \overline{\langle y, \varphi^*(x) \rangle} = \overline{\langle \varphi(y), x \rangle} = \langle x, \varphi(y) \rangle$$

die Gleichung $\varphi^{**} = \varphi$.

(ii) Sei $x \in V$. Da Φ nicht ausgeartet ist, ist die Bedingung $x \in \ker \varphi^*$ äquivalent zu $\langle v, \varphi^*(x) \rangle = 0$ für alle $v \in V$, wegen $\langle v, \varphi^*(x) \rangle = \langle \varphi(v), x \rangle$ aber auch zu $\langle \varphi(v), x \rangle = 0$ für alle $v \in V$ und damit zu $x \in (\text{im } \varphi)^\perp$. Weiter hat man dann aufgrund von (i)

$$\ker \varphi = \ker \varphi^{**} = (\text{im } \varphi^*)^\perp.$$

(iii) Berücksichtigen wir die Konstruktion von φ^* im Beweis zu Satz 2, so stimmt der Rang von φ^* mit dem Rang der zu φ dualen Abbildung überein. Dieser ist jedoch nach 2.3/7 identisch mit dem Rang von φ . Alternativ können wir aber unter Benutzung von (ii) auch wie folgt rechnen:

$$\begin{aligned} \text{rg } \varphi &= \dim V - \dim(\ker \varphi) = \dim V - \dim(\text{im } \varphi^*)^\perp \\ &= \dim V - (\dim V - \text{rg } \varphi^*) = \text{rg } \varphi^* \end{aligned}$$

Dabei wurde allerdings für den Unterraum $U = \text{im } \varphi^* \subset V$ die Formel

$$\dim U + \dim U^\perp = \dim V$$

benutzt, welche wir in 7.2/8 nur für euklidische bzw. unitäre Vektorräume bewiesen hatten. $\square$

Wir wollen nun adjungierte Abbildungen auch mittels Matrizen beschreiben. Hierzu fixieren wir für den Rest dieses Abschnitts einen endlich-dimensionalen $\mathbb{K}$ -Vektorraum V mit einem *Skalarprodukt* $\Phi: V \times V \longrightarrow \mathbb{K}$, also einen *euklidischen* bzw. *unitären* Vektorraum V .

Bemerkung 4. Sei X eine Orthonormalbasis von V . Für $\varphi \in \text{End}_{\mathbb{K}}(V)$ und die zugehörige adjungierte Abbildung φ^* gilt dann

$$A_{\varphi^*, X, X} = (A_{\varphi, X, X})^*.$$

Beweis. Die Matrix $A_{\Phi, X}$, welche Φ bezüglich der Orthonormalbasis X beschreibt, ist die Einheitsmatrix. Folglich gilt für $a, b \in V$ gemäß 3.1/7 und 7.3/2

$$\langle \varphi(a), b \rangle = (A_{\varphi, X, X} \cdot a_X)^t \cdot A_{\Phi, X} \cdot \bar{b}_X = a_X^t \cdot A_{\varphi, X, X}^t \cdot \bar{b}_X,$$

sowie

$$\langle a, \varphi^*(b) \rangle = a_X^t \cdot A_{\Phi, X} \cdot (\bar{A}_{\varphi^*, X, X} \cdot \bar{b}_X) = a_X^t \cdot \bar{A}_{\varphi^*, X, X} \cdot \bar{b}_X.$$

Betrachten wir nun $\langle \varphi(a), b \rangle = \langle a, \varphi^*(b) \rangle$ als Sesquilinearform in $a, b \in V$, so ergibt sich $A_{\varphi, X, X}^t = \overline{A_{\varphi^*, X, X}}$ bzw. $A_{\varphi^*, X, X} = A_{\varphi, X, X}^*$ mittels 7.3/2. $\square$

Definition 5. Eine Abbildung $\varphi \in \text{End}_{\mathbb{K}}(V)$ heißt normal, wenn φ mit der zugehörigen adjungierten Abbildung kommutiert, d. h. wenn $\varphi \circ \varphi^* = \varphi^* \circ \varphi$ gilt.

Satz 6. $\varphi \in \text{End}_{\mathbb{K}}(V)$ ist genau dann normal, wenn

$$\langle \varphi(x), \varphi(y) \rangle = \langle \varphi^*(x), \varphi^*(y) \rangle$$

für alle $x, y \in V$ gilt.

Beweis. Für $x, y \in V$ hat man

$$\begin{aligned} \langle \varphi(x), \varphi(y) \rangle &= \langle x, \varphi^* \circ \varphi(y) \rangle, \\ \langle \varphi^*(x), \varphi^*(y) \rangle &= \langle x, \varphi \circ \varphi^*(y) \rangle, \end{aligned}$$

wobei wir $\varphi^{**} = \varphi$ ausgenutzt haben. Die Gleichung

$$\langle \varphi(x), \varphi(y) \rangle = \langle \varphi^*(x), \varphi^*(y) \rangle, \quad x, y \in V,$$

ist daher äquivalent zu

$$\langle x, \varphi^* \circ \varphi(y) \rangle = \langle x, \varphi \circ \varphi^*(y) \rangle, \quad x, y \in V,$$

sowie aufgrund der Tatsache, dass die Form $\langle x, y \rangle$ auf V nicht ausgeartet ist, äquivalent zu

$$\varphi^* \circ \varphi(y) = \varphi \circ \varphi^*(y), \quad y \in V,$$

also zu $\varphi^* \circ \varphi = \varphi \circ \varphi^*$. $\square$

Korollar 7. Sei $\varphi \in \text{End}_{\mathbb{K}}(V)$ normal.

(i) Es gilt $\ker \varphi = \ker \varphi^*$.

(ii) Ein Vektor $x \in V$ ist genau dann Eigenvektor von φ zum Eigenwert λ , wenn x Eigenvektor von φ^* zum Eigenwert $\overline{\lambda}$ ist.

Beweis. Aussage (i) ergibt sich mittels Satz 6 aus der Gleichung

$$|\varphi(x)|^2 = \langle \varphi(x), \varphi(x) \rangle = \langle \varphi^*(x), \varphi^*(x) \rangle = |\varphi^*(x)|^2, \quad x \in V.$$

Weiter gilt für $\lambda \in \mathbb{K}$ gemäß Korollar 3

$$(\lambda \text{id} - \varphi)^* = \overline{\lambda} \text{id} - \varphi^*,$$

und man erkennt aufgrund der Normalität von φ , dass auch $\lambda \text{id} - \varphi$ wieder normal ist. Folglich gilt nach (i)

$$\ker(\lambda \operatorname{id} - \varphi) = \ker((\lambda \operatorname{id} - \varphi)^*) = \ker(\bar{\lambda} \operatorname{id} - \varphi^*),$$

also Aussage (ii). $\square$

Wir wollen diese Information verwenden, um den so genannten *Spektralsatz für normale Abbildungen* zu beweisen; Spektralsätze geben Auskunft über Eigenwerte und Eigenvektoren von Homomorphismen von Vektorräumen.

Satz 8. *Es sei $\varphi \in \operatorname{End}_{\mathbb{K}}(V)$ ein Endomorphismus, dessen charakteristisches Polynom $\chi_{\varphi} \in \mathbb{K}[T]$ vollständig in Linearfaktoren zerfällt. Dann ist äquivalent:*

- (i) φ ist normal.
- (ii) *Es existiert eine Orthonormalbasis von V , bestehend aus Eigenvektoren bezüglich φ .*

Beweis. Sei zunächst Bedingung (i) gegeben, also φ normal. Um (ii) zu zeigen, verwenden wir Induktion nach $n = \dim_{\mathbb{K}} V$, wobei der Fall $n = 0$ trivial ist. Sei also $n > 0$. Da χ_{φ} vollständig in Linearfaktoren zerfällt, besitzt φ mindestens einen Eigenwert λ und damit auch einen zugehörigen Eigenvektor e_1 . Indem wir e_1 normieren, dürfen wir $|e_1| = 1$ annehmen. Man betrachte nun die Zerlegung

$$V = \mathbb{K}e_1 \oplus (\mathbb{K}e_1)^{\perp};$$

vgl. 7.2/8. Dabei ist $(\mathbb{K}e_1)^{\perp}$ ein φ -invarianter Unterraum von V , denn für $x \in (\mathbb{K}e_1)^{\perp}$ gilt unter Benutzung von Korollar 7 (ii)

$$\langle \varphi(x), e_1 \rangle = \langle x, \varphi^*(e_1) \rangle = \langle x, \bar{\lambda}e_1 \rangle = \lambda \langle x, e_1 \rangle = 0.$$

Die vorstehende Zerlegung ist also eine Zerlegung in φ -invariante Unterräume von V , und wir können auf $\tilde{V} = (\mathbb{K}e_1)^{\perp}$ die Induktionsvoraussetzung anwenden. Es existiert daher eine Orthonormalbasis $e_2, \dots, e_n$ von $\tilde{V}$, die aus Eigenvektoren bezüglich φ besteht. Insgesamt ist dann $e_1, \dots, e_n$ eine Orthonormalbasis von V , die aus Eigenvektoren bezüglich φ besteht, d. h. Bedingung (ii) ist erfüllt.

Sei nun Bedingung (ii) gegeben, sei also X eine Orthonormalbasis von V , die aus Eigenvektoren zu φ besteht. Dann ist die Matrix $A_{\varphi, X, X}$ eine Diagonalmatrix und folglich auch die Matrix $A_{\varphi^*, X, X} = (A_{\varphi, X, X})^*$; vgl. Bemerkung 4. Da Diagonalmatrizen miteinander kommutieren, gilt dasselbe für φ und φ^* , und wir sehen, dass (i) gilt, φ also normal ist. $\square$

Aufgaben

V sei stets ein euklidischer bzw. unitärer $\mathbb{K}$ -Vektorraum endlicher Dimension, φ ein Endomorphismus von V und φ^* die zugehörige adjungierte Abbildung.

1. Man zeige $\operatorname{Spur}(\varphi \circ \varphi^*) \geq 0$, wobei $\operatorname{Spur}(\varphi \circ \varphi^*)$ genau für $\varphi = 0$ verschwindet.
2. Man zeige:
 - (i) Ist φ normal, so gilt $\operatorname{im} \varphi^* = \operatorname{im} \varphi$.
 - (ii) Ist ψ ein weiterer Endomorphismus von V und sind φ, ψ normal, so ist $\varphi \circ \psi = 0$ äquivalent zu $\psi \circ \varphi = 0$.

3. Für $\varphi = \varphi^2$ zeige man: Es gilt genau dann $\varphi = \varphi^*$, wenn $\ker \varphi$ und $\operatorname{im} \varphi$ orthogonal zueinander sind.
4. Für $\mathbb{K} = \mathbb{C}$ zeige man: φ ist genau dann normal, wenn es ein Polynom $p \in \mathbb{C}[T]$ mit $\varphi^* = p(\varphi)$ gibt.
5. Für $\mathbb{K} = \mathbb{C}$ und φ normal zeige man: Sind $x, y \in V$ zwei Eigenvektoren zu verschiedenen Eigenwerten, so gilt $x \perp y$.

7.5 Isometrien, orthogonale und unitäre Matrizen

Generell sei V in diesem Abschnitt ein endlich-dimensionaler $\mathbb{K}$ -Vektorraum mit einem Skalarprodukt $\Phi: V \times V \longrightarrow \mathbb{K}$, also ein *euklidischer* bzw. *unitärer* Vektorraum. Wir wollen im Folgenden Endomorphismen $V \longrightarrow V$ studieren, die als solche nicht nur mit der Vektorraumstruktur von V verträglich sind, sondern zusätzlich auch das Skalarprodukt respektieren und damit längenerhaltend bzw., soweit definiert, auch winkelerhaltend sind. Es handelt sich um die so genannten *Isometrien*.

Satz 1. Für $\varphi \in \operatorname{End}_{\mathbb{K}}(V)$ ist äquivalent:

- (i) Es gilt $\langle \varphi(x), \varphi(y) \rangle = \langle x, y \rangle$ für alle $x, y \in V$.
- (ii) φ ist ein Isomorphismus mit $\varphi^* = \varphi^{-1}$.
- (iii) Es gilt $|\varphi(x)| = |x|$ für alle $x \in V$.
- (iv) Ist X eine Orthonormalbasis von V , so ist deren Bild $\varphi(X)$ ebenfalls eine Orthonormalbasis von V .
- (v) Es existiert eine Orthonormalbasis X von V , so dass $\varphi(X)$ eine Orthonormalbasis von V ist.

Sind diese Bedingungen erfüllt, so bezeichnet man φ als eine *Isometrie*. Im Falle $\mathbb{K} = \mathbb{R}$ nennt man φ auch eine *orthogonale* und im Falle $\mathbb{K} = \mathbb{C}$ eine *unitäre* Abbildung.

Aus den Eigenschaften (ii) bzw. (iv) und (v) liest man sofort ab:

Korollar 2. Ist $\varphi \in \operatorname{End}_{\mathbb{K}}(V)$ eine Isometrie, so ist φ ein Isomorphismus, und φ^{-1} ist ebenfalls eine Isometrie. Die Komposition zweier Isometrien ergibt wiederum eine Isometrie. Insbesondere bilden die Isometrien von V eine Untergruppe der Automorphismengruppe $\operatorname{Aut}_{\mathbb{K}}(V)$.

Beweis zu Satz 1. Wir beginnen mit der Implikation (i) $\implies$ (ii). Sei also (i) gegeben. Dann hat man für $x, y \in V$

$$\langle x, y \rangle = \langle \varphi(x), \varphi(y) \rangle = \langle x, \varphi^* \varphi(y) \rangle.$$

Da die Form $\langle \cdot, \cdot \rangle$ nicht ausgeartet ist, gilt $\varphi^* \circ \varphi(y) = y$ für alle $y \in V$ und damit $\varphi^* \circ \varphi = \operatorname{id}_V$. Insbesondere ist φ injektiv und damit nach 2.1/11 ein Isomorphismus, so dass $\varphi^* = \varphi^{-1}$ und damit (ii) folgt.

Ist andererseits (ii) gegeben, also φ ein Isomorphismus mit $\varphi^* = \varphi^{-1}$, so gilt

$$\langle \varphi(x), \varphi(x) \rangle = \langle x, \varphi^* \varphi(x) \rangle = \langle x, \varphi^{-1} \varphi(x) \rangle = \langle x, x \rangle$$

für $x \in V$, also (iii).

Um die Implikation (iii) $\implies$ (iv) nachzuweisen, nehmen wir (iii) als gegeben an und betrachten eine Orthonormalbasis $X = (e_1, \dots, e_n)$ von V . Für $\mu \neq \nu$ und $\alpha \in \mathbb{K}$, $|\alpha| = 1$, gilt dann

$$\begin{aligned} 2 &= |e_\mu|^2 + |\alpha|^2 |e_\nu|^2 = |e_\mu + \alpha e_\nu|^2 \\ &= |\varphi(e_\mu) + \alpha \varphi(e_\nu)|^2 \\ &= |\varphi(e_\mu)|^2 + 2\operatorname{Re}(\overline{\alpha} \langle \varphi(e_\mu), \varphi(e_\nu) \rangle) + |\varphi(e_\nu)|^2 \\ &= |e_\mu|^2 + 2\operatorname{Re}(\overline{\alpha} \langle \varphi(e_\mu), \varphi(e_\nu) \rangle) + |e_\nu|^2 \\ &= 2 + 2\operatorname{Re}(\overline{\alpha} \langle \varphi(e_\mu), \varphi(e_\nu) \rangle), \end{aligned}$$

also

$$\operatorname{Re}(\overline{\alpha} \langle \varphi(e_\mu), \varphi(e_\nu) \rangle) = 0.$$

Setzen wir $\alpha = 1$, so folgt $\operatorname{Re}(\langle \varphi(e_\mu), \varphi(e_\nu) \rangle) = 0$ und damit $\langle \varphi(e_\mu), \varphi(e_\nu) \rangle = 0$ im Falle $\mathbb{K} = \mathbb{R}$. Für $\mathbb{K} = \mathbb{C}$ kann man aber auch $\alpha = i$ setzen und erhält dann

$$\operatorname{Im}(\langle \varphi(e_\mu), \varphi(e_\nu) \rangle) = \operatorname{Re}(-i \langle \varphi(e_\mu), \varphi(e_\nu) \rangle) = 0,$$

also insgesamt ebenfalls $\langle \varphi(e_\mu), \varphi(e_\nu) \rangle = 0$.

Da weiter $|\varphi(e_\mu)| = |e_\mu|$ für $\mu = 1, \dots, n$ gilt, ist $(\varphi(e_1), \dots, \varphi(e_n))$ ein Orthonormalsystem in V . Da dieses nach 7.2/2 linear unabhängig ist, handelt es sich um eine Orthonormalbasis. Das Bild einer Orthonormalbasis von V unter φ ist folglich wieder eine Orthonormalbasis.

Da es in V stets eine Orthonormalbasis gibt, vgl. 7.2/5, ist die Implikation (iv) $\implies$ (v) trivial. Es bleibt daher lediglich noch die Implikation (v) $\implies$ (i) zu zeigen. Sei also $X = (e_1, \dots, e_n)$ eine Orthonormalbasis von V , so dass $\varphi(X)$ ebenfalls eine Orthonormalbasis von V ist. Sind dann

$$x = \sum_{\mu=1}^n \alpha_\mu e_\mu, \quad y = \sum_{\nu=1}^n \beta_\nu e_\nu$$

zwei Vektoren in V , so gilt

$$\begin{aligned} \langle \varphi(x), \varphi(y) \rangle &= \sum_{\mu, \nu=1}^n \alpha_\mu \overline{\beta}_\nu \langle \varphi(e_\mu), \varphi(e_\nu) \rangle \\ &= \sum_{\mu, \nu=1}^n \alpha_\mu \overline{\beta}_\nu \delta_{\mu\nu} = \sum_{\mu, \nu=1}^n \alpha_\mu \overline{\beta}_\nu \langle e_\mu, e_\nu \rangle = \langle x, y \rangle, \end{aligned}$$

d. h. Bedingung (i) ist erfüllt. $\square$

Um Beispiele von Isometrien zu geben, fassen wir $\mathbb{R}^2$ als euklidischen Vektorraum mit dem kanonischen Skalarprodukt auf. Die kanonische Basis e_1, e_2 ist dann eine Orthonormalbasis. Sei nun ϑ , $0 \leq \vartheta < 2\pi$, ein Winkel und

$$\begin{aligned}\varphi: \mathbb{R}^2 &\longrightarrow \mathbb{R}^2, & e_1 &\longmapsto \cos \vartheta \cdot e_1 + \sin \vartheta \cdot e_2, \\ & & e_2 &\longmapsto -\sin \vartheta \cdot e_1 + \cos \vartheta \cdot e_2,\end{aligned}$$

die Drehung mit ϑ um den Nullpunkt in $\mathbb{R}^2$, also die durch die Matrix

$$R(\vartheta) := \begin{pmatrix} \cos \vartheta & -\sin \vartheta \\ \sin \vartheta & \cos \vartheta \end{pmatrix} \in \mathbb{R}^{2 \times 2}$$

gegebene $\mathbb{R}$ -lineare Abbildung. Für $x = \alpha e_1 + \beta e_2 \in \mathbb{R}^2$ gilt dann

$$\begin{aligned}|\varphi(x)|^2 &= |(\alpha \cdot \cos \vartheta - \beta \cdot \sin \vartheta) \cdot e_1 + (\alpha \cdot \sin \vartheta + \beta \cdot \cos \vartheta) \cdot e_2|^2 \\ &= (\alpha \cdot \cos \vartheta - \beta \cdot \sin \vartheta)^2 + (\alpha \cdot \sin \vartheta + \beta \cdot \cos \vartheta)^2 \\ &= \alpha^2 + \beta^2 = |x|^2,\end{aligned}$$

d. h. φ ist eine Isometrie. Aber auch Spiegelungen, wie etwa für eine beliebige Orthonormalbasis u_1, u_2 von $\mathbb{R}^2$ die durch

$$u_1 \longmapsto u_1, \quad u_2 \longmapsto -u_2,$$

gegebene $\mathbb{R}$ -lineare Abbildung $\mathbb{R}^2 \longrightarrow \mathbb{R}^2$, sind Isometrien.

Korollar 3. Für $\varphi \in \text{End}_{\mathbb{K}}(V)$ ist äquivalent:

- (i) φ ist eine Isometrie.
- (ii) Ist X eine Orthonormalbasis von V , so ist $A_{\varphi, X, X}$ invertierbar, und es gilt

$$(A_{\varphi, X, X})^* = (A_{\varphi, X, X})^{-1}.$$

- (iii) Es existiert eine Orthonormalbasis X von V , so dass $A_{\varphi, X, X}$ invertierbar ist und

$$(A_{\varphi, X, X})^* = (A_{\varphi, X, X})^{-1}$$

gilt.

Beweis. Man benutze 7.4/4 und Satz 1 (ii). □

Schließlich kann man leicht erkennen, dass die Eigenschaft $A^* = A^{-1}$ aus Korollar 3 (ii) bzw. (iii) auch zur Charakterisierung von Basiswechselmatrizen verwendet werden kann, die Orthonormalbasen in Orthonormalbasen von V überführen.

Satz 4. Es sei X eine Orthonormalbasis, sowie Y eine weitere Basis von V . Dann ist äquivalent:

- (i) Y ist eine Orthonormalbasis von V .
- (ii) Für die Basiswechselmatrix $A = A_{Y, X}$ gilt $A^* = A^{-1}$.

Beweis. Sei $\varphi: V \longrightarrow V$ die $\mathbb{K}$ -lineare Abbildung, die die Basis X auf die Basis Y abbildet; es gilt

$$A_{\varphi, X, X} = A = A_{Y, X}.$$

Nach Satz 1 (iv) bzw. (v) ist Y genau dann eine Orthonormalbasis, wenn φ eine Isometrie ist, und nach Korollar 3 ist dies äquivalent zu $(A_{\varphi, X, X})^* = (A_{\varphi, X, X})^{-1}$, also zu $A^* = A^{-1}$. $\square$

Eine Matrix $A \in \mathbb{R}^{n \times n}$ heißt *orthogonal*, wenn die $\mathbb{R}$ -lineare Abbildung

$$\mathbb{R}^n \longrightarrow \mathbb{R}^n, \quad x \longmapsto Ax,$$

orthogonal, also eine Isometrie ist. Dabei betrachte man $\mathbb{R}^n$ mittels des kanonischen Skalarprodukts als euklidischen Vektorraum. Nach Korollar 3 ist die Orthogonalität einer Matrix $A \in \mathbb{R}^{n \times n}$ äquivalent zu der Bedingung $A^t = A^{-1}$ bzw. $A^t A = E$, wobei $E \in \mathbb{R}^{n \times n}$ die Einheitsmatrix sei. Sind $a_1, \dots, a_n \in \mathbb{R}^n$ die Spalten von A , so bedeutet diese Bedingung gerade

$$a_i^t \cdot a_j = \delta_{ij}, \quad i, j = 1, \dots, n.$$

Eine entsprechende Überlegung lässt sich auch für die Zeilen von A anstellen. Eine Matrix $A \in \mathbb{R}^{n \times n}$ ist daher genau dann orthogonal, wenn ihre Spalten (bzw. Zeilen) ein Orthonormalsystem und damit eine Orthonormalbasis von $\mathbb{R}^n$ bilden.

Über dem Körper $\mathbb{C}$ verfährt man in ähnlicher Weise. Eine Matrix $A \in \mathbb{C}^{n \times n}$ heißt *unitär*, wenn die $\mathbb{C}$ -lineare Abbildung

$$\mathbb{C}^n \longrightarrow \mathbb{C}^n, \quad x \longmapsto Ax,$$

unitär, also eine Isometrie ist. Dabei betrachte man $\mathbb{C}^n$ mittels des kanonischen Skalarprodukts als unitären Vektorraum. Gemäß Korollar 3 ist $A \in \mathbb{C}^{n \times n}$ genau dann unitär, wenn $A^* = A^{-1}$ bzw. $A^* A = E$ gilt, und diese Bedingung ist wiederum dazu äquivalent, dass die Spalten (bzw. Zeilen) von A eine Orthonormalbasis von $\mathbb{C}^n$ bilden.

Aus den Korollaren 2 und 3 ergibt sich unmittelbar, dass die orthogonalen bzw. unitären Matrizen $A \in \mathbb{K}^{n \times n}$ eine Untergruppe der Gruppe $\text{Gl}(n, \mathbb{K})$ aller invertierbaren Matrizen in $\mathbb{K}^{n \times n}$ bilden. Dabei nennt man

$$\text{O}(n) = \{A \in \text{Gl}(n, \mathbb{R}) ; A^t = A^{-1}\}$$

die *orthogonale Gruppe* zum Index n , sowie

$$\text{U}(n) = \{A \in \text{Gl}(n, \mathbb{C}) ; A^* = A^{-1}\}$$

die *unitäre Gruppe* zum Index n , wobei man $\text{O}(n)$ als Untergruppe von $\text{U}(n)$ auffassen kann. Für $A \in \text{U}(n)$ ist $A^* \cdot A = A^{-1} \cdot A$ die Einheitsmatrix, so dass man für jede orthogonale bzw. unitäre Matrix $|\det(A)| = 1$ erhält. Daher lässt sich insbesondere

$$\text{SO}(n) = \{A \in \text{O}(n) ; \det(A) = 1\}$$

als Untergruppe der orthogonalen Gruppe betrachten; man nennt diese Gruppe die *spezielle orthogonale Gruppe*.

Wir wollen im Falle $n = 2$ die orthogonale Gruppe $O(2)$ genauer beschreiben. Bereits oben wurde gezeigt, dass die Drehmatrizen

$$R(\vartheta) = \begin{pmatrix} \cos \vartheta & -\sin \vartheta \\ \sin \vartheta & \cos \vartheta \end{pmatrix} \in \mathbb{R}^{2 \times 2}, \quad 0 \leq \vartheta < 2\pi,$$

zu Isometrien von $\mathbb{R}^2$ Anlass geben und folglich zu $O(2)$ gehören, ja sogar zu der Untergruppe $SO(2)$, da jeweils $\det(R(\vartheta)) = 1$ gilt. Wir wollen zunächst zeigen, dass $SO(2)$ keine weiteren Matrizen enthält. Hierzu betrachten wir eine Matrix

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} \in O(2).$$

Dann ergibt die Relation

$$A^t \cdot A = \begin{pmatrix} \alpha_{11}^2 + \alpha_{21}^2 & \alpha_{11}\alpha_{12} + \alpha_{21}\alpha_{22} \\ \alpha_{12}\alpha_{11} + \alpha_{22}\alpha_{21} & \alpha_{12}^2 + \alpha_{22}^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

das Gleichungssystem

$$\begin{aligned} \alpha_{11}^2 + \alpha_{21}^2 &= 1, \\ \alpha_{12}^2 + \alpha_{22}^2 &= 1, \\ \alpha_{11}\alpha_{12} + \alpha_{21}\alpha_{22} &= 0. \end{aligned}$$

Die komplexen Zahlen

$$\alpha_{11} + i\alpha_{21}, \quad \alpha_{22} + i\alpha_{12}$$

sind daher vom Betrag 1. Gleiches gilt für ihr Produkt, und es folgt

$$\begin{aligned} (\alpha_{11} + i\alpha_{21}) \cdot (\alpha_{22} + i\alpha_{12}) &= (\alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12}) + i(\alpha_{11}\alpha_{12} + \alpha_{21}\alpha_{22}) \\ &= \alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12} \\ &= \det(A) = \pm 1 \in \mathbb{R}. \end{aligned}$$

Benutzen wir nun die Identität $z\bar{z} = 1$ für komplexe Zahlen z vom Betrag 1, so ergibt sich für $\det(A) = 1$, also im Falle $A \in SO(2)$

$$\alpha_{22} + i\alpha_{12} = \overline{\alpha_{11} + i\alpha_{21}} = \alpha_{11} - i\alpha_{21},$$

und damit

$$\alpha_{22} = \alpha_{11}, \quad \alpha_{12} = -\alpha_{21}.$$

Wir benutzen nun aus der Analysis, dass es wegen $\alpha_{11}^2 + \alpha_{21}^2 = 1$ genau einen Winkel ϑ , $0 \leq \vartheta < 2\pi$, mit

$$\alpha_{11} + i\alpha_{21} = \cos \vartheta + i \sin \vartheta = e^{i\vartheta},$$

also mit

$$A = \begin{pmatrix} \alpha_{11} & -\alpha_{21} \\ \alpha_{21} & \alpha_{11} \end{pmatrix} = \begin{pmatrix} \cos \vartheta & -\sin \vartheta \\ \sin \vartheta & \cos \vartheta \end{pmatrix} = R(\vartheta)$$

gibt. Dies bedeutet, dass A eine Drehung um den Nullpunkt mit dem Winkel ϑ beschreibt.

Das charakteristische Polynom einer Drehung $R(\vartheta)$ berechnet sich zu

$$\chi_{R(\vartheta)} = T^2 - 2(\cos \vartheta)T + 1 = (T - e^{i\vartheta})(T - e^{-i\vartheta}),$$

hat also nur für $\sin \vartheta = 0$, d. h. nur für $\vartheta \in \{0, \pi\}$ reelle Nullstellen. Somit erhält man:

Satz 5. *Es besteht $\text{SO}(2)$ aus allen reellen (2×2) -Matrizen, die zu Drehungen um den Nullpunkt im $\mathbb{R}^2$ Anlass geben, also*

$$\text{SO}(2) = \{R(\vartheta) ; 0 \leq \vartheta < 2\pi\}.$$

Abgesehen von den trivialen Fällen $R(0) = \text{id}$ und $R(\pi) = -\text{id}$ besitzen die Matrizen $R(\vartheta)$ keine reellen Eigenwerte und sind folglich auch nicht diagonalisierbar.

Für Matrizen $A = (\alpha_{ij})_{i,j=1,2} \in \text{O}(2)$ mit $\det(A) = -1$ ergibt die obige Rechnung in entsprechender Weise die Identitäten

$$\alpha_{22} + i\alpha_{12} = -\overline{(\alpha_{11} + i\alpha_{21})} = -\alpha_{11} + i\alpha_{21},$$

also

$$\alpha_{22} = -\alpha_{11}, \quad \alpha_{12} = \alpha_{21},$$

so dass es auch in diesem Falle genau einen Winkel ϑ , $0 \leq \vartheta < 2\pi$, mit

$$A = \begin{pmatrix} \cos \vartheta & \sin \vartheta \\ \sin \vartheta & -\cos \vartheta \end{pmatrix}$$

gibt. Natürlich gehören alle Matrizen A dieses Typs zu $\text{O}(2)$. Das charakteristische Polynom einer solchen Matrix ist von der Form $\chi_A = T^2 - 1$, so dass A die reellen Eigenwerte 1 und -1 besitzt und folglich diagonalisierbar ist. Die Gleichung $A^t = A^{-1}$ zeigt, dass A einen normalen Endomorphismus von $\mathbb{R}^2$ beschreibt. Gemäß 7.4/8 gibt es dann in $\mathbb{R}^2$ eine Orthonormalbasis bestehend aus Eigenvektoren von A , und man erkennt A als Spiegelung. Somit folgt:

Satz 6. *Es besteht $\text{O}(2)$ aus allen reellen (2×2) -Matrizen, die zu Drehungen bzw. Spiegelungen im $\mathbb{R}^2$ Anlass geben. Genauer gilt*

$$\text{O}(2) = \left\{ \begin{pmatrix} \cos \vartheta & -\sin \vartheta \\ \sin \vartheta & \cos \vartheta \end{pmatrix}, \begin{pmatrix} \cos \vartheta & \sin \vartheta \\ \sin \vartheta & -\cos \vartheta \end{pmatrix} \in \mathbb{R}^{2 \times 2} ; 0 \leq \vartheta < 2\pi \right\}.$$

Die Struktur der Gruppen $\text{O}(2)$ und $\text{SO}(2)$ ist damit vollständig geklärt. Man sollte noch vermerken (und dies auch mittels elementarer Rechnung überprüfen), dass $\text{SO}(2)$ kommutativ ist, nicht aber $\text{O}(2)$.

Wir wollen nun Normalformen für Isometrien allgemeinen Typs in euklidischen und unitären Vektorräumen herleiten. Mit Satz 1 (ii) folgt insbesondere, dass jede Isometrie normal ist. Dementsprechend ist zumindest über dem Körper $\mathbb{C}$ der komplexen Zahlen der Spektralsatz 7.4/8 anwendbar.

Satz 7. *Es sei $\varphi \in \text{End}_{\mathbb{C}}(V)$ ein Endomorphismus eines unitären $\mathbb{C}$ -Vektorraums V endlicher Dimension. Dann ist äquivalent:*

- (i) φ ist eine Isometrie.
- (ii) Für alle Eigenwerte λ von φ gilt $|\lambda| = 1$, und es besitzt V eine Orthonormalbasis, bestehend aus Eigenvektoren von φ .

Beweis. Sei zunächst (i) gegeben, φ also eine Isometrie. Dann kann φ als längenerhaltende Abbildung nur Eigenwerte vom Betrag 1 haben. Da $\mathbb{C}$ im Übrigen algebraisch abgeschlossen ist, zerfällt das charakteristische Polynom χ_{φ} vollständig in Linearfaktoren. Weiter ist φ normal. Also folgt mit 7.4/8 die Existenz einer Orthonormalbasis X von V , die aus lauter Eigenvektoren von φ besteht, d. h. es gilt (ii).

Ist umgekehrt X eine Orthonormalbasis von V , die aus Eigenvektoren von φ besteht, und sind alle Eigenwerte vom Betrag 1, so wird X durch φ offenbar wiederum in eine Orthonormalbasis überführt. φ ist damit eine Isometrie gemäß Satz 1. $\square$

Insbesondere ist jede Isometrie eines unitären Vektorraums endlicher Dimension diagonalisierbar. Auf ein entsprechendes Resultat für euklidische Vektorräume kann man allerdings nicht hoffen, wie das Beispiel der Drehungen um den Nullpunkt im $\mathbb{R}^2$ zeigt. Die Situation ist hier etwas komplizierter.

Satz 8. *Es sei $\varphi \in \text{End}_{\mathbb{R}}(V)$ ein Endomorphismus eines euklidischen $\mathbb{R}$ -Vektorraums V endlicher Dimension. Dann ist äquivalent:*

- (i) φ ist eine Isometrie.
- (ii) Es existiert eine Orthonormalbasis X von V , so dass φ bezüglich X durch eine Matrix des Typs

$$A = A_{\varphi, X, X} = \text{Diag}(E_k, -E_{\ell}, R(\vartheta_1), \dots, R(\vartheta_m))$$

$$= \begin{pmatrix} E_k & & & & \\ & -E_{\ell} & & & \\ & & R(\vartheta_1) & & \\ & & & \ddots & \\ & & & & R(\vartheta_m) \end{pmatrix}$$

beschrieben wird. Dabei ist $E_k \in \mathbb{R}^{k \times k}$ (bzw. $E_{\ell} \in \mathbb{R}^{\ell \times \ell}$) die Einheitsmatrix mit einer gewissen Zeilen- und Spaltenzahl k (bzw. ℓ), und $R(\vartheta_1), \dots, R(\vartheta_m) \in \mathbb{R}^{2 \times 2}$ sind Drehmatrizen zu gewissen Winkeln $0 < \vartheta_1 \leq \vartheta_2 \leq \dots \leq \vartheta_m < \pi$.

Ist φ eine Isometrie, so ist die Normalform in (ii) mit den Zahlen k, ℓ, m und den Winkeln $\vartheta_1, \dots, \vartheta_m$ eindeutig durch φ bestimmt, und zwar gilt

$$\chi_\varphi = (T - 1)^k (T + 1)^\ell \prod_{j=1}^m (T - e^{i\vartheta_j})(T - e^{-i\vartheta_j}).$$

Beweis. Sei zunächst φ eine Isometrie. Um die in (ii) behauptete Normalform herzuleiten, verwenden wir Induktion nach der Dimension n von V . Im Falle $n = 0$ ist nichts zu zeigen, sei also $n > 0$. Wir behaupten, dass dann V einen φ -unzerlegbaren linearen Unterraum U der Dimension 1 oder 2 enthält. Mit 6.5/6 folgt nämlich, dass V jedenfalls einen φ -invarianten Unterraum U' besitzt, der, betrachtet als $K[T]$ -Modul unter φ , isomorph zu einem Quotienten $\mathbb{R}[T]/(p^c)$ ist; dabei ist $p \in \mathbb{R}[T]$ ein Primpolynom und $c \geq 1$ ein gewisser Exponent. Die Multiplikation mit p^{c-1} definiert nun einen $\mathbb{R}[T]$ -Modulisomorphismus $\mathbb{R}[T] \xrightarrow{\sim} p^{c-1}\mathbb{R}[T]$ und durch Quotientenbildung einen $\mathbb{R}[T]$ -Modulisomorphismus $\mathbb{R}[T]/(p) \xrightarrow{\sim} p^{c-1}\mathbb{R}[T]/(p^c)$, wobei wir $p^{c-1}\mathbb{R}[T]/(p^c)$ als $\mathbb{R}[T]$ -Untermodule von $\mathbb{R}[T]/(p^c)$ auffassen können. Folglich enthält U' und damit V einen φ -invarianten linearen Unterraum U , der als $\mathbb{R}[T]$ -Modul isomorph zu $\mathbb{R}[T]/(p)$ ist, also nach 6.5/4 bzw. 6.5/7 φ -unzerlegbar ist und nach 6.5/5 die Dimension $\deg p$ besitzt. Nun haben aber Primpolynome in $\mathbb{R}[T]$ den Grad 1 oder 2, vgl. Aufgabe 3 aus Abschnitt 5.3, so dass wir $U \subset V$ als φ -invarianten linearen Unterraum der Dimension 1 oder 2 erkennen.

Wir betrachten nun die Zerlegung $V = U \oplus U^\perp$, vgl. 7.2/8, und zeigen, dass mit U auch das orthogonale Komplement $U^\perp$ ein φ -invarianter Unterraum von V ist. Hierzu bemerken wir zunächst, dass φ als Isomorphismus eine lineare Abbildung $\varphi|_U: U \rightarrow U$ induziert, die zumindest injektiv, dann aber aufgrund von 2.1/11 sogar bijektiv ist und damit $\varphi^{-1}(U) = U$ erfüllt. Sei nun $y \in U^\perp$. Für alle $x \in U$ gilt dann

$$\langle x, \varphi(y) \rangle = \langle \varphi^*(x), y \rangle = \langle \varphi^{-1}(x), y \rangle = 0,$$

da mit x , wie wir gesehen haben, auch $\varphi^{-1}(x)$ zu U gehört. Dies bedeutet aber $\varphi(y) \in U^\perp$, und wir erkennen auf diese Weise, dass in der Tat $U^\perp$ ein φ -invarianter Unterraum von V ist.

Man wähle nun in U eine Orthonormalbasis X' . Im Falle $\dim_{\mathbb{R}}(U) = 1$ besteht X' lediglich aus einem einzigen Vektor, der dann notwendig ein Eigenvektor von φ ist. Der zugehörige Eigenwert ist gemäß Satz 1 (iii) vom Betrag 1, also gleich 1 oder -1 . Für $\dim_{\mathbb{R}}(U) = 2$ besteht X' aus 2 Vektoren, und es folgt mit Korollar 3, dass die Matrix $A_{\varphi|_U, X', X'}$ zu $O(2)$ gehört, allerdings nicht diagonalisierbar sein kann, da wir U als φ -unzerlegbar angenommen hatten. Dann gilt aufgrund der obigen Beschreibung der Matrizen in $O(2)$ notwendig $A_{\varphi|_U, X', X'} = R(\vartheta)$ mit einem Winkel $\vartheta \neq \pi$, $0 < \vartheta < 2\pi$. Indem wir notfalls die Reihenfolge der Vektoren in X' ändern, was einem Ersetzen des Winkels ϑ durch $2\pi - \vartheta$ entspricht, können wir sogar $0 < \vartheta < \pi$ annehmen.

Nach Induktionsvoraussetzung gibt es nun in $U^\perp$ eine Orthonormalbasis X'' , so dass $A_{\varphi|_{U^\perp}, X'', X''}$ von der in (ii) beschriebenen Gestalt ist. Man kann dann X' und X'' zu einer Orthonormalbasis X von V zusammensetzen, und es folgt nach geeigneter Umnummerierung der Basisvektoren von X , dass die Matrix $A_{\varphi, X, X}$ die gewünschte Gestalt hat. Die Implikation (i) $\implies$ (ii) ist somit bewiesen. Umgekehrt ist unmittelbar klar, dass jede Matrix des Typs (ii) orthogonal ist, φ in diesem Falle daher eine Isometrie ist.

Zum Nachweis der Eindeutigkeitsaussage hat man lediglich zu beachten, dass sich für eine Isometrie φ das charakteristische Polynom einer beschreibenden Matrix A wie in (ii) zu

$$\chi_\varphi = \chi_A = (T - 1)^k (T + 1)^\ell \prod_{j=1}^m (T - e^{i\vartheta_j})(T - e^{-i\vartheta_j})$$

berechnet. Somit bestimmt χ_φ zunächst die Zahlen k, ℓ und m in A . Da bei den Paaren komplex konjugierter Nullstellen $e^{i\vartheta_j}, e^{-i\vartheta_j}$ der Realteil $\cos \vartheta_j$ jeweils gleich ist, und da dieser eindeutig den Winkel im Bereich $0 < \vartheta_j < \pi$ bestimmt, ist insgesamt die Gestalt der Matrix A durch das Polynom χ_φ festgelegt. $\square$

Aufgaben

Falls nicht anderweitig bestimmt, sei V stets ein euklidischer bzw. unitärer $\mathbb{K}$ -Vektorraum endlicher Dimension.

1. Man zeige: Zu $x, y \in V$ gibt es genau dann eine Isometrie $\varphi \in \text{End}_{\mathbb{K}}(V)$ mit $\varphi(x) = y$, wenn $|x| = |y|$ gilt.
2. Man zeige für Drehmatrizen $R(\vartheta_1), R(\vartheta_2) \in \mathbb{R}^{2 \times 2}$ mit Winkeln $0 \leq \vartheta_1 < \vartheta_2 < 2\pi$, dass diese genau dann ähnlich sind, wenn $\vartheta_1 + \vartheta_2 = 2\pi$ gilt.
3. Es seien $e_1, \dots, e_n \in \mathbb{R}^n$ die kanonischen Einheitsvektoren, aufgefasst als Spaltenvektoren. Eine Matrix $A \in \mathbb{R}^{n \times n}$ heißt *Permutationsmatrix*, wenn es eine Permutation $\pi \in \mathfrak{S}_n$ mit $A = E_\pi := (e_{\pi(1)}, \dots, e_{\pi(n)})$ gibt. Man zeige:
 - (i) Jede Permutationsmatrix ist invertierbar, und die kanonische Abbildung $\sigma: \mathfrak{S}_n \longrightarrow \text{Gl}(n, \mathbb{R}), \pi \longmapsto E_\pi$, definiert einen injektiven Gruppenhomomorphismus.
 - (ii) Es gilt $E_\pi \in O(n)$ für alle $\pi \in \mathfrak{S}_n$.
 - (iii) Es sei $A = (\alpha_{ij}) \in O(n)$ eine orthogonale Matrix mit Koeffizienten $\alpha_{ij} \geq 0$. Dann ist A bereits eine Permutationsmatrix.
4. Es sei $A = (\alpha_{ij}) \in O(n)$ eine orthogonale Matrix in unterer Dreiecksgestalt, d. h. es gelte $\alpha_{ij} = 0$ für $i < j$. Man zeige, dass A sogar eine Diagonalmatrix ist. Gilt eine entsprechende Aussage auch für Matrizen $A \in U(n)$?
5. Es seien $A, B \in \mathbb{C}^{n \times n}$ hermitesche Matrizen. Man zeige, dass A und B genau dann ähnlich sind, wenn sie hermitesch ähnlich sind, d. h. wenn es eine hermitesche Matrix $S \in \mathbb{C}^{n \times n}$ mit $B = S^{-1}AS$ gibt. Gilt die entsprechende Aussage auch über dem Körper $\mathbb{K} = \mathbb{R}$ mit "symmetrisch" anstelle von "hermitesch"?

7.6 Selbstadjungierte Abbildungen

Es sei V weiterhin ein endlich-dimensionaler $\mathbb{K}$ -Vektorraum mit einem Skalarprodukt $\Phi: V \times V \longrightarrow \mathbb{K}$, also ein *euklidischer* bzw. *unitärer* Vektorraum.

Definition 1. Ein Endomorphismus $\varphi \in \text{End}_{\mathbb{K}}(V)$ heißt selbstadjungiert, wenn $\varphi^* = \varphi$ gilt.

Eine Matrix $A \in \mathbb{K}^{n \times n}$ heißt symmetrisch (im Falle $\mathbb{K} = \mathbb{R}$) bzw. hermitesch (im Falle $\mathbb{K} = \mathbb{C}$), wenn $A^* = A$ gilt.

Für einen beliebigen Endomorphismus $\varphi: V \longrightarrow V$ ist die Komposition $\varphi \circ \varphi^*$ selbstadjungiert, wie die Gleichung

$$\langle \varphi \circ \varphi^*(x), y \rangle = \langle \varphi^*(x), \varphi^*(y) \rangle = \langle x, \varphi \circ \varphi^*(y) \rangle, \quad x, y \in V,$$

zeigt. Entsprechend ist für jede Matrix $A \in \mathbb{K}^{n \times n}$ das Produkt AA^* symmetrisch bzw. hermitesch. Weiter ist jeder selbstadjungierte Endomorphismus von V insbesondere normal. Mittels 7.4/4 ergibt sich folgende Charakterisierung selbstadjungierter Endomorphismen:

Bemerkung 2. Sei X eine Orthonormalbasis von V . Für $\varphi \in \text{End}_{\mathbb{K}}(V)$ ist dann äquivalent:

- (i) φ ist selbstadjungiert.
- (ii) $(A_{\varphi, X, X})^* = A_{\varphi, X, X}$, d. h. $A_{\varphi, X, X}$ ist symmetrisch bzw. hermitesch.

Aufgrund unserer Untersuchungen normaler Endomorphismen können wir nun folgende wichtige Beobachtung machen:

Satz 3. Sei $\varphi \in \text{End}_{\mathbb{K}}(V)$ selbstadjungiert. Dann besitzt das charakteristische Polynom χ_{φ} reelle Koeffizienten und zerfällt in $\mathbb{R}[T]$ vollständig in Linearfaktoren. Insbesondere besitzt φ nur reelle Eigenwerte.

Beweis. Sei zunächst $\mathbb{K} = \mathbb{C}$. Mittels 7.4/7 ergibt sich $\lambda = \overline{\lambda}$ und damit $\lambda \in \mathbb{R}$ für alle Eigenwerte λ von φ bzw. alle Nullstellen von χ_{φ} . Benutzen wir, dass $\mathbb{C}$ algebraisch abgeschlossen ist, so sehen wir, dass χ_{φ} vollständig in Linearfaktoren zerfällt und dass diese alle reell sind. Insbesondere ist χ_{φ} ein reelles Polynom.

Im Falle $\mathbb{K} = \mathbb{R}$ wähle man eine Orthonormalbasis X von V und betrachte die Matrix $A = A_{\varphi, X, X}$, wobei dann $A^* = A^t = A$ gilt; vgl. 7.4/4. Betrachtet man nun $\mathbb{C}^n$ mit $n = \dim_{\mathbb{R}} V$ als unitären Vektorraum unter dem kanonischen Skalarprodukt, so definiert A als symmetrische reelle Matrix insbesondere auch einen selbstadjungierten Endomorphismus

$$\mathbb{C}^n \longrightarrow \mathbb{C}^n, \quad x \longmapsto Ax,$$

und wir können wie oben schließen, dass das zugehörige charakteristische Polynom, nämlich χ_A , vollständig in reelle Linearfaktoren zerfällt. Da aber χ_A mit χ_{φ} übereinstimmt, sind wir fertig. $\square$

Für selbstadjungierte Endomorphismen können wir daher den Spektralsatz für normale Endomorphismen 7.4/8 wie folgt verschärfen:

Korollar 4. *Für einen Endomorphismus $\varphi \in \text{End}_{\mathbb{K}}(V)$ ist äquivalent:*

- (i) *φ ist selbstadjungiert.*
- (ii) *Alle Eigenwerte von φ sind reell, und V besitzt eine Orthonormalbasis, die aus lauter Eigenvektoren von φ besteht.*
- (iii) *Es gibt in V eine Orthonormalbasis X , so dass $A_{\varphi, X, X}$ eine reelle Diagonalmatrix ist.*

Beweis. Sei zunächst Bedingung (i) gegeben, also φ selbstadjungiert. Dann besitzt das charakteristische Polynom χ_{φ} nach Satz 3 reelle Koeffizienten und zerfällt über $\mathbb{R}$ vollständig in Linearfaktoren; alle Eigenwerte sind daher reell. Da φ insbesondere normal ist, existiert nach 7.4/8 eine Orthonormalbasis X von V , bestehend aus Eigenvektoren zu φ , d. h. Bedingung (ii) ist erfüllt. Weiter folgt aus (ii), dass die Matrix $A_{\varphi, X, X}$ Diagonalgestalt hat, wobei auf der Diagonalen die Eigenwerte von φ stehen. Somit ist $A_{\varphi, X, X}$ eine reelle Diagonalmatrix, wie in (iii) behauptet.

Gibt es schließlich eine Orthonormalbasis X von V , so dass $A_{\varphi, X, X}$ eine reelle Diagonalmatrix ist, so gilt

$$(A_{\varphi, X, X})^* = A_{\varphi, X, X},$$

und es folgt $A_{\varphi^*, X, X} = A_{\varphi, X, X}$ bzw. $\varphi^* = \varphi$ mit 7.4/4. Bedingung (iii) impliziert daher (i). $\square$

Da man eine Matrix $A \in \mathbb{K}^{n \times n}$ mit $A = A^*$ stets als selbstadjungierten Endomorphismus von $\mathbb{K}^n$, versehen mit dem kanonischen Skalarprodukt, interpretieren kann, erhält man als unmittelbare Folgerung:

Korollar 5. *Es sei $A \in \mathbb{K}^{n \times n}$ eine Matrix, welche der Bedingung $A^* = A$ genügt, also eine symmetrische bzw. hermitesche Matrix. Dann ist A ähnlich zu einer reellen Diagonalmatrix. Insbesondere gilt $\chi_A \in \mathbb{R}[T]$, und dieses Polynom zerfällt in $\mathbb{R}[T]$ vollständig in Linearfaktoren.*

Im Spezialfall $\mathbb{K} = \mathbb{R}$ ist also eine symmetrische Matrix $A \in \mathbb{R}^{n \times n}$ stets diagonalisierbar und besitzt daher (für $n \geq 1$) mindestens einen (reellen) Eigenwert.

Wir haben in Korollar 5 noch nicht ausgenutzt, dass die Basiswechselmatrix, welche die Ähnlichkeit zwischen einer symmetrischen bzw. hermiteschen Matrix A und der zugehörigen reellen Diagonalmatrix vermittelt, nach Konstruktion in Korollar 4 einen Basiswechsel zwischen Orthonormalbasen beschreibt und folglich aufgrund von 7.5/4 orthogonal bzw. unitär ist. Wir wollen dies nun berücksichtigen und gelangen auf diese Weise zum so genannten Satz über die *Hauptachsentransformation*, den wir in zwei verschiedenen Versionen herleiten werden.

Theorem 6. *Es sei $A \in \mathbb{K}^{n \times n}$ eine Matrix mit $A^* = A$. Dann existiert eine orthogonale bzw. unitäre Matrix $S \in \text{Gl}(n, \mathbb{K})$ mit der Eigenschaft, dass*

$$D = S^{-1}AS = S^*AS = (\overline{S})^t A \overline{S}$$

eine reelle Diagonalmatrix ist. Dabei ist mit S auch $\overline{S} \in \text{Gl}(n, \mathbb{K})$ orthogonal bzw. unitär, und die Diagonaleinträge von D sind gerade die Eigenwerte von A .

Die Gleichung $D = S^{-1}AS$ besagt, dass A als (selbstadjungierte) lineare Abbildung $\mathbb{K}^n \rightarrow \mathbb{K}^n$ mittels des durch S gegebenen Basiswechsels auf Diagonalgestalt transformiert werden kann. Entsprechend bedeutet $D = (\overline{S})^t A \overline{S}$, dass die durch A gegebene sBF bzw. HF mittels des durch $\overline{S}$ gegebenen Basiswechsels auf Diagonalgestalt transformiert werden kann.

Theorem 7. *Sei V ein euklidischer bzw. unitärer Vektorraum mit Skalarprodukt Φ , und sei $\Psi: V \times V \rightarrow \mathbb{K}$ eine beliebige sBF bzw. HF auf V . Dann existiert eine Orthonormalbasis X von V , so dass die Matrix $A_{\Psi, X}$ eine reelle Diagonalmatrix ist.*

Beweis zu Theorem 6. Wir fassen $\mathbb{K}^n$ zusammen mit dem kanonischen Skalarprodukt als euklidischen bzw. unitären Vektorraum auf. Dann ist die lineare Abbildung

$$\varphi: \mathbb{K}^n \rightarrow \mathbb{K}^n, \quad x \mapsto Ax,$$

selbstadjungiert, denn es gilt $A^* = A$; vgl. Bemerkung 2. Nach Korollar 4 existiert daher eine Orthonormalbasis X von $\mathbb{K}^n$, so dass die beschreibende Matrix $D = A_{\varphi, X, X}$ eine reelle Diagonalmatrix ist. Ist nun $S = A_{X, E} \in \text{Gl}(n, \mathbb{K})$ die Matrix des Basiswechsels zwischen X und der kanonischen Basis E von $\mathbb{K}^n$, so erhalten wir gemäß 3.4/5

$$D = S^{-1} \cdot A \cdot S.$$

Da S einen Basiswechsel zwischen zwei Orthonormalbasen beschreibt, ist S nach 7.5/4 orthogonal bzw. unitär, und es gilt $S^* = S^{-1}$. Durch Konjugieren erhält man $\overline{S}^* = (\overline{S})^{-1}$, d. h. $\overline{S}$ ist ebenfalls orthogonal bzw. unitär. $\square$

Beweis zu Theorem 7. Wir gehen von einer beliebigen Orthonormalbasis X von V aus; vgl. 7.2/6. Für die zu Ψ gehörige Matrix gilt dann $A_{\Psi, X} = A_{\Psi, X}^*$ nach 7.3/4, und es existiert, wie wir in Theorem 6 gesehen haben, eine orthogonale bzw. unitäre Matrix $S \in \text{Gl}(n, \mathbb{K})$, so dass $D = S^t \cdot A_{\Psi, X} \cdot \overline{S}$ eine reelle Diagonalmatrix ist. Wir können nun S als Basiswechselmatrix der Form $A_{Y, X}$ mit einer neuen Basis Y von V auffassen. Nach 7.5/4 ist Y wiederum eine Orthonormalbasis, und es gilt

$$A_{\Psi, Y} = S^t \cdot A_{\Psi, X} \cdot \overline{S} = D,$$

wobei man 7.3/5 benutze. Y anstelle von X erfüllt also die Behauptung. $\square$

Um die geometrische Bedeutung der Hauptachsentransformation zu erläutern, betrachte man beispielsweise die Kurve C , die in $\mathbb{R}^2$ durch die Gleichung

$$41x_1^2 - 24x_1x_2 + 34x_2^2 = 25$$

gegeben ist. Mit $x = (x_1, x_2)^t$, sowie

$$A = \begin{pmatrix} 41 & -12 \\ -12 & 34 \end{pmatrix}$$

lässt sich C dann auch durch die Gleichung

$$x^t \cdot A \cdot x = 25$$

beschreiben. Wenden wir nun Theorem 6 an, so existiert eine orthogonale Matrix $S \in O(2)$ mit der Eigenschaft, dass

$$D = S^t A S = S^{-1} A S$$

eine reelle Diagonalmatrix ist, wobei auf der Diagonalen von D gerade die Eigenwerte von A stehen. Da das charakteristische Polynom von A die Gestalt

$$\chi_A = (T - 25)(T - 50)$$

hat, können wir etwa

$$S^t A S = S^{-1} A S = \begin{pmatrix} 25 & 0 \\ 0 & 50 \end{pmatrix}$$

annehmen. Ist nun $X = E$ die kanonische Basis von $\mathbb{R}^2$, so lässt sich S als Basiswechselmatrix der Form $A_{Y,X}$ mit einer neuen Basis Y auffassen, wobei Y gemäß 7.5/4 wiederum eine Orthonormalbasis ist. Es besteht dann nach 3.4/3 die Beziehung

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = S \cdot \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

zwischen Koordinaten x_1, x_2 bezüglich X und y_1, y_2 bezüglich Y , so dass sich die Kurve C in den neuen Koordinaten durch die Gleichung

$$25y_1^2 + 50y_2^2 = 25$$

bzw. durch

$$y_1^2 + 2y_2^2 = 1$$

beschreibt. Insbesondere sehen wir, dass es sich um eine Ellipse handelt. In der Gleichung kommen nunmehr keine gemischten Terme mehr vor, was bedeutet, dass man C bezüglich seiner ‘Hauptachsen’ beschrieben hat.

Um C noch genauer zu charakterisieren, sollte man natürlich die Hauptachsen von C explizit bestimmen, also die Orthonormalbasis Y und die zugehörige

Transformationsmatrix $S = A_{Y,X}$.² Dies kann wie folgt geschehen. Man betrachte

$$\mathbb{R}^2 \longrightarrow \mathbb{R}^2, \quad x \longmapsto Ax,$$

als selbstadjungierten Endomorphismus von $\mathbb{R}^2$, versehen mit dem kanonischen Skalarprodukt, und bestimme eine Orthonormalbasis Y von $\mathbb{R}^2$, die aus Eigenvektoren zu A besteht. Hierzu ermittelt man zunächst Basen der Eigenräume zu A (durch Lösen der entsprechenden linearen Gleichungssysteme) und orthonormalisiert die Basen anschließend nach E. Schmidt. In unserem Fall ist dies sehr simpel. Es sind

$$(3, 4)^t, \quad (-4, 3)^t$$

Eigenvektoren zu den Eigenwerten 25 und 50. Nach Normierung ergibt sich

$$\frac{1}{5}(3, 4)^t, \quad \frac{1}{5}(-4, 3)^t$$

als Orthonormalbasis Y von $\mathbb{R}^2$ und

$$S = A_{Y,X} = \frac{1}{5} \begin{pmatrix} 3 & -4 \\ 4 & 3 \end{pmatrix}$$

als zugehörige Transformationsmatrix. Dabei ist die Isometrie

$$\mathbb{R}^2 \longrightarrow \mathbb{R}^2, \quad x \longmapsto Sx,$$

als Drehung $R(\vartheta)$ mit einem Winkel ϑ von ungefähr 53° zu interpretieren.

Abschließend wollen wir noch einige Folgerungen aus dem Theorem über die Hauptachsentransformation ziehen.

Korollar 8. *Für eine sBF bzw. HF Ψ auf einem endlich-dimensionalen $\mathbb{K}$ -Vektorraum V ist äquivalent:*

- (i) Ψ ist positiv definit.
- (ii) *Es existiert eine Basis X von V , so dass die Matrix $A_{\Psi,X}$ nur positive Eigenwerte hat.*

Beweis. Ist Ψ positiv definit, so gibt es bezüglich Ψ eine Orthonormalbasis X von V . Die zugehörige Matrix $A_{\Psi,X}$ ist dann die Einheitsmatrix.

Sei nun umgekehrt X eine Basis von V , so dass die Matrix $A_{\Psi,X}$ nur positive Eigenwerte hat. Es gilt $A_{\Psi,X}^* = A_{\Psi,X}$ nach 7.3/4. Aufgrund von Theorem 6 existiert dann eine Matrix $S \in \text{Gl}(n, \mathbb{K})$ mit $S^* = S^{-1}$, so dass

$$(\overline{S})^t \cdot A_{\Psi,X} \cdot \overline{\overline{S}} = S^{-1} A_{\Psi,X} \cdot S = D$$

eine reelle Diagonalmatrix ist. Dabei sind die Diagonalelemente von D gerade die Eigenwerte von $A_{\Psi,X}$ und folglich größer als 0. Interpretiert man nun $\overline{S}$

² Man beachte jedoch, dass die Basis Y nicht im strengen Sinne als eindeutig bezeichnet werden kann, denn deren Elemente lassen sich noch bezüglich Reihenfolge und Vorzeichen abändern. Für eine Kreislänge C besitzt sogar jede Orthonormalbasis von $\mathbb{R}^2$ die Eigenschaft von Hauptachsen.

als Basiswechselmatrix des Typs $A_{Y,X}$ mit einer neuen Basis Y von V , so gilt $A_{\Psi,Y} = D$, und man erkennt, dass Ψ positiv definit ist. $\square$

Korollar 9 (Sylvesterscher Trägheitssatz). *Sei Ψ eine sBF bzw. HF auf einem endlich-dimensionalen $\mathbb{K}$ -Vektorraum V . Für eine Basis X von V sei weiter k die Anzahl der positiven, ℓ die Anzahl der negativen, sowie m die Anzahl der Eigenwerte von $A_{\Psi,X}$, die verschwinden, jeweils gezählt mit Vielfachheiten. Dann gilt $k + \ell + m = \dim_{\mathbb{K}} V$, und die Zahlen k, ℓ, m sind eindeutig durch Ψ bestimmt, insbesondere unabhängig von der Wahl von X .*

Es existiert eine Basis X von V , so dass die Matrix $A_{\Psi,X}$ die folgende Diagonalgestalt

$$\begin{pmatrix} 1 & & & & & & & & & 0 \\ & \ddots & & & & & & & & \\ & & \ddots & & & & & & & \\ & & & 1 & & & & & & \\ & & & & -1 & & & & & \\ & & & & & \ddots & & & & \\ & & & & & & \ddots & & & \\ & & & & & & & -1 & & \\ & & & & & & & & 0 & \\ & & & & & & & & & \ddots \\ & & & & & & & & & & \ddots \\ 0 & & & & & & & & & & 0 \end{pmatrix}$$

besitzt, wobei auf der Diagonalen k -mal 1, ℓ -mal -1 und m -mal 0 steht.

Beweis. Wir gehen aus von einer Basis $X = (x_1, \dots, x_n)$ von V und betrachten die Matrix $A_{\Psi,X}$, sowie die zugehörigen Zahlen k, ℓ, m . Dann existiert nach Theorem 6 eine orthogonale bzw. unitäre Matrix $S \in \text{Gl}(n, \mathbb{K})$, so dass $D = S^t A \bar{S}$ eine reelle Diagonalmatrix ist; die Diagonalelemente von D sind gerade die Eigenwerte von A . Indem wir S als Basiswechselmatrix auffassen und X durch die entsprechende neue Basis ersetzen, können wir $\Psi(x_i, x_j) = \delta_{ij} \alpha_i$ mit $\alpha_i \in \mathbb{R}$ annehmen, $i, j = 1, \dots, n$. Weiter kann man für $\Psi(x_i, x_i) \neq 0$ den Basisvektor x_i mit $|\Psi(x_i, x_i)|^{-1/2}$ multiplizieren und somit $\alpha_i \in \{1, -1, 0\}$ annehmen. Nummerieren wir die x_i dann noch in der Weise um, dass

$$\alpha_i = \begin{cases} 1 & \text{für } i = 1, \dots, k \\ -1 & \text{für } i = k+1, \dots, k+\ell \\ 0 & \text{für } i = k+\ell+1, \dots, k+\ell+m \end{cases}$$

gilt, so hat die Matrix $A_{\Psi,X}$ die behauptete Gestalt.

Es bleibt noch zu zeigen, dass die Zahlen k, ℓ, m unabhängig von der Wahl der Basis X sind. Seien also X_i , $i = 1, 2$, zwei Basen von V , und seien k_i, ℓ_i, m_i die zugehörigen Zahlen, wobei wir annehmen dürfen, wie wir soeben gesehen haben, dass die Matrizen A_{Ψ,X_i} reelle Diagonalmatrizen sind. Sei nun $V_i^+ \subset V$

der von allen $x \in X_i$ mit $\Psi(x, x) > 0$ erzeugte Unterraum, $V_i^- \subset V$ der von allen $x \in X_i$ mit $\Psi(x, x) < 0$ erzeugte Unterraum, sowie $V_i^0 \subset V$ der von allen $x \in X_i$ mit $\Psi(x, x) = 0$ erzeugte Unterraum. Dann gilt für $i = 1, 2$

$$\begin{aligned} V &= V_i^+ \oplus V_i^- \oplus V_i^0, \\ \dim V_i^+ &= k_i, \quad \dim V_i^- = \ell_i, \quad \dim V_i^0 = m_i, \\ k_i + \ell_i + m_i &= \dim V. \end{aligned}$$

Dabei erkennt man

$$V_1^0 = V_2^0 = \{x \in V ; \Psi(x, y) = 0 \text{ für alle } y \in V\}$$

als “Entartungsraum” von Ψ , es gilt folglich $m_1 = m_2$. Weiter hat man offenbar

$$V_1^+ \cap (V_2^- \oplus V_2^0) = 0,$$

da Ψ auf V_1^+ positiv definit und auf $V_2^- \oplus V_2^0$ negativ semi-definit ist. Dies bedeutet

$$k_1 + \ell_2 + m_2 \leq \dim V,$$

und dies ergibt wegen $k_2 + \ell_2 + m_2 = \dim V$ dann $k_1 \leq k_2$, sowie aus Symmetriegründen $k_1 = k_2$. Dann ergibt sich aber auch

$$\ell_1 = \dim V - k_i - m_i = \ell_2, \quad i = 1, 2,$$

und damit die Behauptung. $\square$

Aufgaben

Falls nicht anderweitig bestimmt, sei V stets ein euklidischer bzw. unitärer $\mathbb{K}$ -Vektorraum endlicher Dimension.

1. Es sei $\varphi \in \text{End}_{\mathbb{K}}(V)$ selbstadjungiert. Man zeige: φ besitzt genau dann lauter positive reelle Eigenwerte, wenn $\langle \varphi(x), x \rangle > 0$ für alle $x \in V - \{0\}$ gilt.
2. Für einen normalen Endomorphismus $\varphi \in \text{End}_{\mathbb{K}}(V)$ zeige man: $\varphi \circ \varphi^*$ besitzt lauter reelle Eigenwerte ≥ 0 .
3. Für $A \in \mathbb{R}^{n \times n}$ zeige man: A ist genau dann symmetrisch, wenn es eine Matrix $S \in \mathbb{C}^{n \times n}$ mit $A = S^t S$ gibt.
4. Es sei $A \in \mathbb{K}^{n \times n}$ symmetrisch bzw. hermitesch. Man zeige, dass A genau dann eine positiv semi-definite Sesquilinearform auf $\mathbb{K}^n$ definiert, wenn A von der Form $S^* S$ mit einer Matrix $S \in \mathbb{K}^{n \times n}$ ist.
5. Man zeige: Zwei symmetrische bzw. hermitesche Matrizen $A, B \in \mathbb{K}^{n \times n}$ sind genau dann ähnlich, wenn es eine orthogonale bzw. unitäre Matrix $S \in \text{Gl}(n, \mathbb{K})$ mit $B = S^t A \overline{S}$ gibt.

Symbolverzeichnis

$\overrightarrow{0P}$	Vektor von 0 nach P 3
$x \in X$	Element einer Menge 9
$\emptyset$	leere Menge 9
$\mathbb{N}$	natürliche Zahlen 9
$\mathbb{Z}$	ganze Zahlen 9
$\mathbb{Q}$	rationale Zahlen 9
$\mathbb{R}$	reelle Zahlen 9
$\{x_1, \dots, x_n\}$	Menge 9
$Y \subset X$	Teilmenge einer Menge 10
$\mathbb{R}_{>0}$	positive reelle Zahlen 10
$\mathfrak{P}(X)$	Potenzmenge einer Menge 10
$\bigcup_{i \in I} X_i$	Vereinigung von Mengen 10
$\bigcap_{i \in I} X_i$	Durchschnitt von Mengen 10
$X_1 \cup \dots \cup X_n$	endliche Vereinigung von Mengen 10
$X_1 \cap \dots \cap X_n$	endlicher Durchschnitt von Mengen 10
$\prod_{i \in I} X_i$	disjunkte Vereinigung von Mengen 10
$X_1 - X_2$	Differenz von Mengen 10
$\prod_{i=1}^n X_i$	endliches kartesisches Produkt von Mengen 11
$X_1 \times \dots \times X_n$	endliches kartesisches Produkt von Mengen 11
X^n	n -faches kartesisches Produkt einer Menge 11
$(x_1, \dots, x_n)$	n -Tupel von Elementen 11
$\prod_{i \in I} X_i$	kartesisches Produkt von Mengen 11
$(x_i)_{i \in I}$	Familie von Elementen 11
X^I	kartesisches Produkt einer Menge 11
$f: X \longrightarrow Y$	Abbildung zwischen Mengen 11
id_X	identische Abbildung 11
$g \circ f$	Komposition von Abbildungen 11
$f(M)$	Bild einer Menge 11
$f^{-1}(N)$	Urbild einer Menge 11
$f^{-1}(y)$	Urbild eines Elementes 11
$f^{-1}: Y \longrightarrow X$	Umkehrabbildung zu einer Abbildung 12
$a \cdot b$	Produkt von Elementen einer Gruppe 13
$a + b$	Summe von Elementen einer kommutativen Gruppe 13
G	Gruppe 13
$e \in G$	Einselement einer Gruppe 13
$\mathbb{Q}^*$	von 0 verschiedene rationale Zahlen 13
$\mathbb{R}^*$	von 0 verschiedene reelle Zahlen 13

$\text{Bij}(X, X)$	Gruppe bijektiver Selbstabbildungen 13
$1 \in G$	Einselement einer Gruppe 14
a^{-1}	inverses Element 14
$\prod_{i=1}^n a_i$	Produkt von Elementen 14
$\prod_{i=1}^0 a_i$	leeres Produkt 15
$0 \in G$	Nullelement einer kommutativen Gruppe 15
$-a$	inverses Element in einer kommutativen Gruppe 15
$\sum_{i=1}^n a_i$	Summe von Elementen in einer kommutativen Gruppe 15
$\sum_{i=1}^0 a_i$	leere Summe 15
$A \implies B$	Implikation von Aussagen 15
$A \iff B$	Äquivalenz von Aussagen 15
K	Körper 17
$0 \in K$	Nullelement eines Körpers 18
$1 \in K$	Einselement eines Körpers 18
$n \cdot a$	n -fache Summe eines Elements 18
K^*	multiplikative Gruppe eines Körpers 18
$\mathbb{F}_2$	Körper mit 2 Elementen 19
$\mathbb{Q}(\sqrt{2})$	von $\mathbb{Q}$ und $\sqrt{2}$ erzeugter Körper 19
$\mathbb{C}$	Körper der komplexen Zahlen 21
$i \in \mathbb{C}$	komplexe Zahl mit $i^2 = -1$ 22
a^n	n -fache Potenz 22
$\binom{n}{i}$	Binomialkoeffizient 23
V	Vektorraum 26
$0 \in V$	Nullvektor eines Vektorraums 26
$V = 0$	Nullvektorraum 26
$U \subset V$	Untervektorraum 27
$K \cdot a \subset V$	von a erzeugter Untervektorraum 27
K^n	n -faches kartesisches Produkt als K -Vektorraum 28
V^n, W^n	n -faches kartesisches Produkt von Vektorräumen 29
$\prod_{i \in I} V_i$	kartesisches Produkt von Vektorräumen 29
$\text{Abb}(X, K)$	Vektorraum von K -wertigen Funktionen auf einer Menge 29
$\bigcap_{i \in I} U_i$	Durchschnitt von Untervektorräumen 29
$\langle A \rangle \subset V$	von einer Menge von Vektoren erzeugter Untervektorraum 30
$\langle a_1, \dots, a_n \rangle \subset V$	von Vektoren erzeugter Untervektorraum 30
$\langle \mathfrak{A} \rangle \subset V$	von einer Familie von Vektoren erzeugter Untervektorraum 30
$e_i \in K^n$	i -ter Einheitsvektor 31
δ_{ij}	Kronecker-Symbol 31
$\sum_{i=1}^n \alpha_i a_i$	Linearkombination von Vektoren 32
$a_1, \dots, a_n$	Basis eines Vektorraums 34
$\dim_K V$	Dimension eines Vektorraums 40
$\sum_{i=1}^r U_i$	Summe von Untervektorräumen 44
$\bigoplus_{i=1}^r U_i$	direkte Summe von Untervektorräumen 45
$U_1 + \dots + U_r$	Summe von Untervektorräumen 45
$U_1 \oplus \dots \oplus U_r$	direkte Summe von Untervektorräumen 45
$f: V \longrightarrow V'$	lineare Abbildung zwischen Vektorräumen 57
$\ker f$	Kern einer linearen Abbildung 59
$\text{im } f$	Bild einer linearen Abbildung 59

$\text{Hom}_K(V, V')$	Vektorraum linearer Abbildungen 61
$f + g$	Summe linearer Abbildungen 61
$\alpha \cdot f$	skalares Produkt mit einer linearen Abbildung 61
$\text{rg } f$	Rang einer linearen Abbildung 62
$A = a + U$	affiner Unterraum 65
$f^{-1}(a')$	Faser einer linearen Abbildung 65
$R \subset M \times M$	Relation auf einer Menge 66
$a \sim b$	Relation, Äquivalenzrelation 66
$[a]$	Äquivalenzklasse eines Elementes 66
$M/\sim$	Menge von Äquivalenzklassen 67
$M \longrightarrow M/\sim$	kanonische Abbildung zu Menge von Äquivalenzklassen 67
$[a] = a + U$	Nebenklasse modulo eines linearen Unterraums 68
V/U	Quotientenvektorraum 68
$\pi: V \longrightarrow V/U$	kanonischer Epimorphismus zu Quotientenvektorraum 69
V^*	Dualraum eines Vektorraums 75
$f^*: V^* \longrightarrow U^*$	duale Abbildung 76
$\varphi_1, \dots, \varphi_n \in V^*$	duale Basis 79
V^{**}	doppelt dualer Vektorraum 80
$\text{coker } f$	Kokern einer linearen Abbildung 82
$A = (\alpha_{ij})_{ij}$	Matrix 91
$A + B$	Summe von Matrizen 91
λA	skalares Produkt mit einer Matrix 91
$0 \in K^{m \times n}$	Nullmatrix 91
$-A = (-\alpha_{ij})_{i,j}$	negatives Element zu einer Matrix 91
$E_{ij} = (\delta_{i\mu} \delta_{j\nu})_{\mu\nu}$	kanonische Basis von Matrizen 91
$a_Y \in K^m$	Koordinatenspaltenvektor 92
$X = (x_1, \dots, x_n)$	Basis eines Vektorraums 92
$A_{f,X,Y}$	beschreibende Matrix einer linearen Abbildung 92
$E = E_m = (\delta_{ij})_{ij}$	Einheitsmatrix 93
$A \cdot B$	Produkt von Matrizen 94
$\text{rg}_s A$	Spaltenrang einer Matrix 100
$\text{rg}_z A$	Zeilenrang einer Matrix 100
$\text{rg } A$	Rang einer Matrix 100
$A^t = (\alpha_{ij})$	transponierte Matrix 105
X^*, Y^*	duale Basis eines Dualraums 105
R	Ring 109
$0 \in R$	Nullelement eines Rings 109
$1 \in R$	Einselement eines Rings 109
$R = 0$	Nullring 109
$\text{End}_K(V)$	Endomorphismenring eines Vektorraums 110
$K^{n \times n}$	Matrizenring quadratischer Matrizen 110
A^{-1}	inverse Matrix 111
R^*	Einheitengruppe eines Rings 111
$\text{Aut}_K(V)$	Automorphismengruppe eines Vektorraums 111
$\text{Gl}(n, K)$	allgemeine lineare Gruppe 111
$A_{\text{id},Y,X}$	Matrix eines Basiswechsels 116
$A \cdot x = b$	lineares Gleichungssystem 119
$M_{A,b}$	Lösungsraum eines linearen Gleichungssystems 120

$\mathfrak{S}_n$	Permutationsgruppe 134
$\begin{pmatrix} 1 & \dots & n \\ a_1 & \dots & a_n \end{pmatrix}$	Permutation 135
$n!$	Fakultät einer natürlichen Zahl 135
$\operatorname{sgn} \pi$	Signum einer Permutation 136
$\mathfrak{A}_n$	alternierende Gruppe 138
$\Delta: V^n \longrightarrow K$	Determinantenfunktion 139
$\det(A)$	Determinante einer Matrix 141
$\det f$	Determinante eines Endomorphismus 144
A^{ad}	adjungierte Matrix 152
$\bigwedge^r V$	r -fache äußere Potenz eines Vektorraums 156
$V^r \longrightarrow \bigwedge^r V$	kanonische Abbildung in äußere Potenz 156
$a_1 \wedge \dots \wedge a_r$	äußeres Produkt von Elementen eines Vektorraums 156
$\det_{X,H}(a_1, \dots, a_r)$	Determinante einer Untermatrix 159
$\bigwedge^r f$	äußeres Produkt einer linearen Abbildung 160
$\bigwedge^r V = \bigoplus_{r \in \mathbb{N}} \bigwedge^r V$	äußere Algebra eines Vektorraums 162
R	Ring 166
$R[T]$	Polynomring über einem kommutativen Ring 167
$\operatorname{grad} f$	Grad eines Polynoms 168
$\varphi: R \longrightarrow R'$	Ringhomomorphismus 169
$R \longrightarrow A$	R -Algebra 169
$\Phi: A \longrightarrow B$	Homomorphismus von R -Algebren 170
$S \subset R$	Unterring 173
$\mathfrak{a} \subset R$	Ideal eines Rings 173
$\ker \varphi$	Kern eines Ringhomomorphismus 173
$(a) = Ra$	von einem Element erzeugtes Hauptideal 174
$(1) = R$	Einheitsideal 174
$0 \subset R$	Nullideal 174
$a + \mathfrak{a}$	Nebenklasse modulo eines Ideals 174
$R/\mathfrak{a}$	Restklassenring modulo eines Ideals 174
$R \longrightarrow R/\mathfrak{a}$	kanonische Projektion zu Restklassenring 175
$\delta: R - \{0\} \longrightarrow \mathbb{N}$	Gradfunktion auf einem euklidischen Ring 178
$a b$	a teilt b 179
$a \nmid b$	a teilt nicht b 179
$a = \varepsilon \prod_{p \in P} p^{\mu_p(a)}$	Primfaktorzerlegung 182
$\operatorname{ggT}(a, b)$	größter gemeinsamer Teiler 182
$\operatorname{kgV}(a, b)$	kleinstes gemeinsames Vielfaches 183
$\lambda \in K$	Eigenwert 194
V_λ	Eigenraum zu einem Eigenwert 196
χ_A	charakteristisches Polynom einer Matrix 200
$\operatorname{Spur} A$	Spur einer Matrix 200
$K(T)$	rationaler Funktionenkörper 200
χ_f	charakteristisches Polynom einer linearen Abbildung 201
$\operatorname{Spur} f$	Spur einer linearen Abbildung 201
p_f	Minimalpolynom einer linearen Abbildung 203
p_A	Minimalpolynom einer Matrix 203
M	Modul 206
$N \subset M$	Untermodul 207

$\varphi: M \longrightarrow N$	Homomorphismus von Moduln 207
$\sum_{i \in I} Ra_i \subset M$	von Elementen erzeugter Untermodul 207
$\sum_{i=1}^n M_i \subset M$	Summe von Untermoduln 208
$\bigoplus_{i=1}^n M_i$	direkte Summe von Moduln 208
$N_1 \times \dots \times N_n$	kartesisches Produkt von Moduln 208
M/N	Restklassenmodul 208
$[a] = a + N$	Nebenklasse modulo eines Untermoduls 208
$\ell_R(M)$	Länge eines Moduls 209
$E^{ij}(\sigma, \tau, \sigma', \tau')$	verallgemeinerte Elementarmatrix 213
$T \subset M$	Torsionsuntermodul eines Moduls 220
$R^n \xrightarrow{\psi} R^m \xrightarrow{\varphi} M$	endliche Präsentation eines Moduls 220
$M_p \subset M$	Untermodul der p -Torsion 222
$M = Ra$	monogener Modul 225
$\text{Diag}(A_1, \dots, A_r)$	“Diagonalmatrix” mit Matrizen als Diagonaleinträgen 230
$A(p)$	Begleitmatrix zu einem Polynom 231
$ 0\vec{P}_1 , P_1 $	Betrag eines Vektors bzw. Punktes 243
$\mathbb{K}$	Körper der reellen oder komplexen Zahlen 246
$\bar{a} = \alpha - i\beta$	konjugiert komplexe Zahl 246
$\text{Re}(a)$	Realteil einer komplexen Zahl 246
$\text{Im}(a)$	Imaginärteil einer komplexen Zahl 247
$ a = \sqrt{\alpha^2 + \beta^2}$	Absolutbetrag einer komplexen Zahl 247
$\Phi: V \times V \longrightarrow \mathbb{K}$	Sesquilinearform 247
$\langle x, y \rangle$	symmetrische Bilinearform oder hermitesche Form 247
$\langle x, y \rangle = \overline{x^t \cdot \bar{y}}$	kanonisches Skalarprodukt auf $\mathbb{K}^n$ 248
$ x = \sqrt{\langle x, x \rangle}$	Betrag bzw. Länge eines Vektors 249
$e_1, \dots, e_n \in V$	Orthonormalbasis 251
$M \perp N$	orthogonale Teilmengen 255
$M^\perp$	orthogonales Komplement 255
$P(x_1, \dots, x_r)$	von Vektoren aufgespanntes Parallelotop 255
$\text{Vol}(P(x_1, \dots, x_r))$	Volumen eines Parallelotops 255
$G(x_1, \dots, x_r)$	Gramsche Determinante 256
$\bar{A} = (\bar{\alpha}_{ij})_{ij}$	konjugierte Matrix 258
$A^t = (\alpha_{ij})_{ji}$	transponierte Matrix 258
$A^* = \bar{A}^t = (\bar{\alpha}_{ij})_{ji}$	adjungierte Matrix 258
$A_{\Phi, X}$	beschreibende Matrix zu einer Sesquilinearform 259
$A_{Y, X}$	Matrix eines Basiswechsels 261
$\bar{V}$	komplex konjugierter Vektorraum 264
φ^*	adjungierter Endomorphismus 264
$R(\vartheta)$	Drehmatrix 271
$O(n)$	orthogonale Gruppe 272
$U(n)$	unitäre Gruppe 272
$SO(n)$	spezielle orthogonale Gruppe 272

Namen- und Sachverzeichnis

- Abbildung, 11–12
 - adjungierte,
 - siehe* Endomorphismus, adjungierter
 - affine, 54
 - alternierende, 139, 155
 - bijektive, 12
 - Bild, 11
 - Bildbereich, 11
 - Definitionsbereich, 11
 - identische, 11
 - injektive, 11
 - Komposition, 11
 - lineare,
 - siehe* lineare Abbildung
 - multilineare, 139, 155
 - normale,
 - siehe* Endomorphismus, normaler
 - selbstadjungierte,
 - siehe* Endomorphismus, selbstadjungierter
 - skalares Produkt, 75
 - surjektive, 11
 - Urbild, 11
 - Wertebereich, 11
- Absolutbetrag, 243, 247, 249
- ähnlich, 193–195
- Äquivalenz, 66
 - von Aussagen, 15
- Äquivalenzklasse, 66–69
 - Menge von, 67
- Äquivalenzrelation, 66–69, 174, 208
- äußere Algebra, 162
- R -Algebra, 169–173
- allgemeine lineare Gruppe, 111, 272
- alternierende Gruppe, 138
- analytische Geometrie, 1
- Assoziativgesetz, 94, 109
- Automorphismengruppe, 111, 269
- Automorphismus
 - eines Rings, 169
 - eines Vektorraums, 59
- Basis, 34–35, 37–40, 42, 91, 208
 - äquivalente Charakterisierung, 37, 116
 - duale, 79, 105, 160
 - eines äußeren Produkts, 159
- Basisergänzungssatz, 38, 43
- Basiswechsel, 90, 96, 116–118, 261, 280, 282, 283
- Begleitmatrix, 191, 231
- Betrag,
 - siehe* Absolutbetrag
- Bild, 59, 265
- Bilinearform, 247
 - nicht ausgeartete, 249
 - symmetrische,
 - siehe* symmetrische Bilinearform
- Binomialkoeffizient, 23
- binomische Formel, 22–25
- Cauchysche Determinante, 151
- Cayley-Hamilton, Satz von, 191, 204, 232
- Charakteristik, 25, 184
- Chinesischer Restsatz, 219–220
- Cramersche Regel, 132–134, 152–154
- Dachprodukt, 156
- Descartes, R., 1
- Determinante, 143–150, 262
 - Beispiele, 147–150
 - Eigenschaften, 144
 - einer Matrix, 131–134, 141, 152–154
 - einer Untermatrix, 146
 - eines Endomorphismus, 144, 160

- Entwicklungssatz, 153, 162
- Rechenregeln, 146
- Regel von Sarrus, 150
- Determinantenfunktion, 139–144
- diagonalisierbar, 189, 193, 194, 196–197, 199, 201, 203, 232, 279, 280
- Diagonalmatrix, 193, 197, 230, 280
- Dimension, 39–42, 55, 94, 120
- Dimensionsformel
 - für lineare Abbildungen, 62, 69
 - für Untervektorräume, 46, 47
- disjunkt, 10
- Distributivgesetze, 109
- Division mit Rest, 176–178
- Drehung, 54, 271
- Dreiecksmatrix, 202, 235
- Dreiecksungleichung, 244, 249
- Dualraum, 57, 75–81, 264
- Durchschnitt, 10

- Ebene, 7, 71
- Eigenraum, 189, 196, 202
- Eigenvektor, 189, 194
 - einer normalen Abbildung, 267, 268
- Eigenwert, 189, 194–197, 199, 278, 282
 - einer normalen Abbildung, 267
- Einheit, 111, 166
- Einheitengruppe, 111, 166
- Einheitsideal, 174
- Einheitsmatrix, 90, 93, 110
- Einheitsvektor, 31, 33, 35
- Einselement, 14, 18
- Einsetzungshomomorphismus, 171
- Element
 - assoziiertes, 179
 - inverses, 13
 - invertierbares, 111
 - irreduzibles, 179
 - maximales, 42
 - neutrales, 13
 - primes, 179–183
 - reduzibles, 179
- Elementarmatrix, 101–102, 104, 112–114, 212
- Elementarteiler
 - einer endlichen Präsentation, 221
 - einer Matrix, 211–214, 217
 - eines Untermoduls, 210–217
- Elementarteilersatz, 210
- Elementarteilertheorie, 191
- Endomorphismenring, 110, 165, 166, 172
- Endomorphismus
 - adjungierter, 264–268
 - diagonalisierbarer, 189, 194, 196–197, 201, 232, 279
 - eines Rings, 169
 - eines Vektorraums, 59
 - invertierbarer, 111–112
 - normaler, 267–268
 - selbstadjungierter, 245, 278–280
 - Struktur, 189
- Epimorphismus
 - von Moduln, 207
 - von Ringen, 169
 - von Vektorräumen, 58, 63, 77
- Erzeugendensystem, 31, 207
 - freies, 208
- euklidischer Ring, 178

- Fakultät, 135
- Familie, 11
- Faser, 55, 65
- Fibonacci-Zahlen, 206
- Form
 - hermitesche, *siehe* hermitesche Form
 - quadratische, 251
- Funktionenkörper, 200

- Gaußsches Eliminationsverfahren, 85, 90, 101–104, 107
 - zur Invertierung von Matrizen, 113–114
 - zur Lösung linearer Gleichungssysteme, 120–128
 - zur Rangbestimmung, 102
- Gerade, 5–6, 27, 71, 85, 87
- Gleichung, algebraische, 165
- Gleichungssystem, lineares, *siehe* lineares Gleichungssystem
- Gradfunktion, 178
- Gramsche Determinante, 256
- Gruppe, 8, 12–16
 - abelsche, 13, 26, 109, 207
 - bijektiver Selbstabbildungen, 13
 - Einselement, 14

- inverses Element, 13, 14
- kommutative, 13
- neutrales Element, 13
- Nullelement, 15
- Ordnung, 224
- orthogonale, 272
- spezielle orthogonale, 272
- unitäre, 272
- zyklische, 224

Hauptachsentransformation, 245–246,
279–283

Hauptideal, 174

Hauptidealring, 178

Hauptsatz

- für endlich erzeugte abelsche Gruppen,
223

- für endlich erzeugte Moduln über
Hauptidealringen, 221–224

Hauptunterdeterminanten, 262

hermitesche Form, 245, 247–250, 260

- positiv definite, 248, 249, 282
- positiv semi-definite, 248

HF,

siehe hermitesche Form

Homomorphiesatz

- für Moduln, 209
- für Ringe, 175
- für Vektorräume, 56, 69

Homomorphismus

- von R -Algebren, 170
- von Gruppen, 137
- von Moduln, 207
- von Ringen, 169
- von Vektorräumen,

siehe lineare Abbildung

Hülle, lineare, 30

Hyperebene, 71

Ideal, 173

Imaginärteil, 22, 247

Implikation von Aussagen, 15

Induktion, vollständige, 23

Integritätsring, 167, 169

Involution, 246

Isometrie, 269–272, 275–277

- Beispiele, 270
- Eigenwerte, 275

Isomorphiesatz

- für Moduln, 219

- für Ringe, 176

- für Vektorräume, 70, 74

Isomorphismus

- und Dimension von Vektorräumen, 62
- von Einheitengruppen, 111
- von Moduln, 207
- von Ringen, 110, 169
- von Vektorräumen, 51, 58, 62, 63, 92,
93, 105

kanonisch, 28

Kern, 56, 59, 173, 265

Koeffizient, 90

Koeffizientenmatrix, 88

Körper, 8, 17–25, 166, 183

- Addition, 17
- algebraisch abgeschlossener, 186
- Assoziativgesetz, 17
- Charakteristik, 25
- der komplexen Zahlen, 187
- Distributivgesetze, 17
- Einselement, 18, 26
- inverses Element, 17
- Kommutativgesetz, 17
- komplexe Zahlen, 21–22
- mit 2 Elementen, 19
- Multiplikation, 17
- multiplikative Gruppe, 18
- neutrales Element, 17
- Nullelement, 18

Kokern, 82

Komplement, 46, 66

- orthogonales, 255, 265

Komplementärmatrix, 152

Komplex, 78

komplexe Zahlen, 21–22

Kongruenz, 67

Konjugationsabbildung, 246

Koordinaten, 1

Koordinatenspaltenvektor, 92, 95, 142,
144

Kronecker-Symbol, 31

Kürzungsregeln, 15

Länge

- eines Moduls, 209–210
- eines Vektors, 249

Laplacescher Entwicklungssatz, 153

- allgemeiner, 162
- linear abhängig, 7, 8, 32, 33, 142
 - äquivalente Charakterisierung, 35
- lineare Abbildung, 51, 57–63
 - Beispiele, 54–55
 - beschreibende Matrix, 62, 89, 92, 95, 116, 208, 266
 - Bild, 59, 265
 - Dimensionsformel, 55
 - duale, 76–81, 105
 - Faser, 55
 - injektive, 59
 - Kern, 56, 59, 265
 - orthogonale, 269
 - Rang, 55, 62, 80, 99–100, 106, 192, 265
 - skalares Produkt, 61
 - Summe, 61, 75
 - und Basen, 60
 - und Dimension, 60
 - und Erzeugendensysteme, 59
 - und lineare Abhängigkeit, 60
 - und lineare Unabhängigkeit, 60
 - unitäre, 269
 - von Moduln, 207
- lineares Gleichungssystem, 35, 85–89, 119–128
 - Cramersche Regel, 132–134
 - homogenes, 119
 - inhomogenes, 119
 - Koeffizientenmatrix, 88
 - Lösbarkeit, 124, 125
 - Lösungsraum, 120
 - partikuläre Lösung, 125
- Linearform, 75
- Linearkombination, 32, 33
- linear unabhängig, 7, 8, 32–34, 131, 195

- Matrix, 55, 58, 62, 85, 90ff., 208
 - adjungierte(1), 151, 152
 - adjungierte(2), 258
 - als lineare Abbildung, 95
 - Cofaktor, 152
 - diagonalisierbare, 193, 199, 203, 232, 279, 280
 - eines Basiswechsels, 90, 116
 - hermitesche, 245, 278
 - inverse, 111
 - invertierbare, 90, 111–114, 131
 - konjugierte, 258
 - leere, 91, 143, 231
 - orthogonale, 244, 263, 272–280
 - Produkt, 94
 - Rang, 100–104, 107, 117–118
 - Rechenregeln, 97
 - skalares Produkt, 91
 - Spalte, 91
 - Spaltenindex, 91
 - Spaltenrang, 100, 107, 118
 - Summe, 91
 - symmetrische, 245, 278
 - transponierte, 105–106, 258
 - trigonalisierbare, 235
 - unitäre, 244, 263, 272–280
 - Zeile, 91
 - Zeilenindex, 91
 - Zeilenrang, 100, 107, 118
- Matrizenring, 110, 166, 172
- Menge, 9–12
 - Differenz, 10
 - disjunkte Vereinigung, 10
 - Durchschnitt, 10
 - gleichmächtige, 43
 - kartesisches Produkt, 11
 - obere Schranke, 42
 - streng geordnete, 42
 - teilweise geordnete, 42
 - Vereinigung, 10
- Minimalpolynom, 172, 191, 203–205, 224, 232
- Modell, 2
- Modul, 191, 206ff.
 - Addition, 206
 - Assoziativgesetz, 207
 - Beispiele, 207
 - direkte Summe, 208
 - Distributivgesetze, 207
 - endlicher, 207
 - endlicher freier, 208
 - endlich erzeugter, 207
 - freier, 208
 - Länge, 209–210
 - monogener, 225
 - Rang, 217, 221
 - skalare Multiplikation, 206
- Monomorphismus
 - von Moduln, 207

- von Ringen, 169
- von Vektorräumen, 58, 63, 77
- Nebenklasse, 68, 174, 208
- Nilradikal, 176
- Normalform, 230
- allgemeine, 192, 231
- Berechnung, 235–240
- für orthogonale Abbildungen, 275
- für unitäre Abbildungen, 275
- Jordansche, 192, 232–237
- Normalteiler, 138
- Nullelement, 15, 18
- Nullideal, 174
- Nullmatrix, 91
- Nullraum, 26
- Nullring, 109, 166
- Nullstelle, 185–187
- Nullteiler, 109, 115, 167
- Nullvektor, 26
- Ordnung, 224
- orthogonal, 251, 255
- Orthogonalbasis, 251
- Orthogonalsystem, 251
- Orthonormalbasis, 244, 251, 266, 269, 271
- Orthonormalsystem, 251
- Paradoxon von Russel, 9
- Parallelotop, 255–257
- Permutation, 134–138, 140, 141, 155
- gerade, 136
- Signum, 136
- ungerade, 136
- Permutationsgruppe, 134
- Anzahl der Elemente, 135
- Permutationsmatrix, 277
- Polynom, 165, 167–169, 171–173, 203
- charakteristisches, 189, 200–205
- Grad, 168–169
- höchster Koeffizient, 168
- Koeffizient, 168
- normiertes, 182
- Nullstelle, 185–187
- Polynomring, 166–169, 171–173
- Potenz, 22
- äußere, 156–163
- symmetrische, 163

- Potenzmenge, 10
- Potenzreihen
- formale, 175
- Präsentation, endliche
- eines Moduls, 220, 237
- Elementarteiler, 221
- Primelement, 179–183
- Primfaktorzerlegung, 176, 181–183
- Produkt
- äußeres, 156–163
- kartesisches, 11, 29, 166, 208
- leeres, 15
- symmetrisches, 163
- Projektion, 54–55, 66
- orthogonale, 252
- senkrechte, 250
- Pythagoras, Satz von, 243, 250
- Quotientenring, 174
- Quotientenvektorraum, 56, 66, 68–71
- Rang
- einer linearen Abbildung, 55, 62, 80, 99–100, 106, 192, 265
- einer Matrix, 100–104, 107, 117–118
- eines Moduls, 217, 221
- Realteil, 22, 246
- Reflexivität, 66
- Relation, 66
- Repräsentant, 66
- Restklasse, 68
- Restklassenmodul, 208
- Restklassenring, 174
- Restklassenvektorraum, 56, 68–71
- Ring, 109, 166, 206
- Einheit, 111, 166
- Einheitsgruppe, 111, 166
- Einselement, 109
- faktorieller, 181
- invertierbares Element, 111
- kartesisches Produkt, 166
- kommutativer, 109
- Nullelement, 109
- sBF,
- siehe symmetrische Bilinearform
- Schmidtsches Orthonormalisierungsverfahren, 244, 250, 252–255
- Schwarzsche Ungleichung, 244, 248

- senkrecht, 243, 251
- Sequenz, 77
 - duale, 78
 - exakte, 78
 - kurze exakte, 78
- Sesquilinearform, 247–250
 - beschreibende Matrix, 259–263
 - nicht ausgeartete, 247, 259
- Skalarprodukt, 243, 248, 261–263
 - kanonisches, 248
- Spalten-Vektor, 88, 92
- Spaltenstufenform, 104
- Spaltenumformungen, elementare, 104, 211
- Spektralsatz
 - für normale Abbildungen, 268
 - für selbstadjungierte Abbildungen, 279
- Spiegelung, 54, 271
- Spur
 - einer linearen Abbildung, 201
 - einer Matrix, 200, 201
- Summe
 - direkte, 45–48, 208
 - konstruierte direkte, 46, 208
 - leere, 15
 - von Untermoduln, 208
 - von Untervektorräumen, 44–48
- Sylvesterscher Trägheitssatz, 245, 283
- Symmetrie, 66
- symmetrische Bilinearform, 243, 245, 247–250, 260
 - Kern, 250
 - positiv definite, 248, 249, 282
 - positiv semi-definite, 248
- symmetrische Gruppe, 134
- System, 32
 - freies, 208
 - Länge, 39
 - leeres, 33
 - linear abhängiges, 32
 - linear unabhängiges, 32, 34, 208
 - maximales linear unabhängiges, 37, 42
 - minimales Erzeugenden-, 37
- Teiler, 179
 - größter gemeinsamer, 182, 183
- Teilkörper, 19–21
- Teilmenge, 10
- Torsion, 191
- Torsionsmodul, 191, 220, 224
- Torsionsuntermodul, 220
- Transitivität, 66
- Transposition, 135
- n -Tupel, 11
- Umkehrabbildung, 12, 58
- universelle Eigenschaft, 56, 155
- Untergruppe, 15
- Untermodul, 207
 - der p -Torsion, 222
 - direkte Summe, 208
 - Summe, 208
- Unterraum
 - affiner, 56, 65–66, 71–73, 87, 120
 - kleinster affiner, 72
 - linearer,
 - siehe* Untervektorraum
- Unterring, 173
- Untervektorraum, 27–29, 65, 71, 120
 - aufgespannter, 7, 30
 - direkte Summe, 45–48
 - erzeugter, 7, 30
 - invarianter, 190, 225ff.
 - Komplement, 46
 - Summe, 44–48
 - unzerlegbarer, 191, 225ff.
 - zyklischer, 190, 225ff.
- Vandermondesche Determinante, 149
- Variable, 165, 173
- Vektor, 3, 5
 - Addition, 3–4
 - normierter, 250
 - skalare Multiplikation, 3
 - Skalarprodukt, 248
 - Subtraktion, 5
- Vektorraum, 3, 26–31
 - Addition, 26
 - Assoziativgesetz, 26
 - Automorphismengruppe, 111
 - Basis, 34–35, 37–40, 42, 116
 - Beispiele, 28, 29
 - Dimension, 39–42, 55, 94, 120
 - direkte Summe, 45–48
 - Distributivgesetze, 26
 - doppelt dualer, 80

- endlich erzeugter, 31
- Endomorphismenring, 110
- Erzeugendensystem, 31
- euklidischer, 244, 248
- kartesisches Produkt, 29
- konstruierte direkte Summe, 46
- neutrales Element, 26
- nicht endlich erzeugter, 42–43
- Nullvektor, 26
- Rechenregeln, 27
- skalare Multiplikation, 26
- unitärer, 244, 248
- von linearen Abbildungen, 61, 93
- K -wertiger Funktionen, 29, 41
- Vereinigung, 10
- Verknüpfung
 - äußere, 26, 206
 - assoziative, 13

- innere, 12, 17, 26, 109, 206
- inverses Element, 13
- kommutative, 13
- neutrales Element, 13

Vielfaches

- kleinstes gemeinsames, 183

vollständige Induktion, 23

Volumen, 255–257

Winkel, 244

Zahl, konjugiert komplexe, 246

Zeilen-Vektor, 88, 92

Zeilenstufenform, 102, 107, 121

Zeilenumformungen, elementare,
101–104, 107, 112–114, 211

Zornsches Lemma, 42

Zyklus, 138