

Gerd Fischer

Lineare Algebra

Gerd Fischer

Lineare Algebra

Eine Einführung für Studienanfänger

17., aktualisierte Auflage

STUDIUM



**VIEWEG+
TEUBNER**

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.

Prof. Dr. Gerd Fischer
Technische Universität München
Zentrum Mathematik
Boltzmannstr. 3
85748 Garching

gfisher@ma.tum.de

Der bisherige Titel der Reihe „Grundkurs Mathematik“ lautete „vieweg studium – Grundkurs Mathematik“.

- | | | |
|-----------------|------------------|------------------|
| 1. Auflage 1975 | 9. Auflage 1986 | 17. Auflage 2010 |
| 2. Auflage 1975 | 10. Auflage 1995 | |
| 3. Auflage 1976 | 11. Auflage 1997 | |
| 4. Auflage 1978 | 12. Auflage 2000 | |
| 5. Auflage 1979 | 13. Auflage 2002 | |
| 6. Auflage 1980 | 14. Auflage 2003 | |
| 7. Auflage 1981 | 15. Auflage 2005 | |
| 8. Auflage 1984 | 16. Auflage 2008 | |

Alle Rechte vorbehalten

© Vieweg+Teubner | GWV Fachverlage GmbH, Wiesbaden 2010

Lektorat: Ulrike Schmickler-Hirzebruch | Nastassja Vanselow

Vieweg+Teubner ist Teil der Fachverlagsgruppe Springer Science+Business Media.

www.viewegteubner.de



Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlags unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

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

Umschlaggestaltung: KünkelLopka Medienentwicklung, Heidelberg
Druck und buchbinderische Verarbeitung: Těšínská Tiskárna, a. s., Tschechien
Gedruckt auf säurefreiem und chlorfrei gebleichtem Papier.
Printed in Czech Republic

ISBN 978-3-8348-0996-4

*We must not accept the old blasphemous nonsense
that the ultimate justification of mathematical science
is the „glory of the human mind“.
Abstraction and generalization
are not more vital for mathematics
than individuality of phenomena
and, before all,
not more than inductive intuition.
Only the interplay of these forces and their synthesis
can keep mathematics alive
and prevent its drying out into a dead skeleton.*

RICHARD COURANT

Vorwort zur 10. Auflage

Die erste im Jahr 1975 veröffentlichte Auflage dieses Buches war entstanden aus meiner Vorlesung im Wintersemester 1972/73 an der Universität Regensburg und einer von Richard Schimpl angefertigten Ausarbeitung, die als Band 1 der Reihe „Der Regensburger Trichter“ erschienen war. Es freut mich, daß das Buch in den vergangenen 20 Jahren so viel Anklang gefunden hat.

Im Jahr 1994/95 hatte ich an der Universität Düsseldorf wieder einmal Gelegenheit, eine Anfängervorlesung über „Lineare Algebra“ zu halten. Dabei fand ich in dem alten Buch zahllose Dinge, die man wohl besser erklären kann. Dazu kam die Versuchung, die Möglichkeiten von $\text{\LaTeX}$ zu nutzen, was schließlich dazu geführt hat, daß ich fast das ganze Buch neu aufgeschrieben habe.

Geblichen ist die Überzeugung, daß am Anfang jeder Theorie Probleme stehen müssen, und daß die entwickelten Methoden danach zu bewerten sind, was sie zur Lösung der Probleme beigetragen haben. Dies deutlich zu machen, ist in der linearen Algebra eine vordringliche Aufgabe, weil hier die axiomatische Methode sehr ausgeprägt ist. Mit Hilfe eines wohlorganisierten Instrumentariums von Begriffen können Beweise kurz und klar durchgeführt werden, Rechnungen können weitgehend vermieden werden und erhalten – wo sie notwendig sind – eine Interpretation von einem abstrakteren Standpunkt aus.

Es hat lange gedauert, bis sich die lineare Algebra von einem Hilfsmittel der sogenannten „analytischen Geometrie“ (das ist die Lehre von den linearen und

quadratischen geometrischen Gebilden) zu einer selbständigen Disziplin entwickelt hat. Die größten Veränderungen gab es zu Anfang dieses Jahrhunderts, als die axiomatische Methode durch den Einfluß von D. HILBERT und speziell in der Algebra durch EMMY NOETHER ausgebaut wurde. Das zeigt ganz deutlich ein Blick in Lehrbücher aus dieser Zeit, etwa die „klassische“ Darstellung von KOWALEWSKI [Kow 2]* aus dem Jahr 1910 und die 1931 erschienene „moderne“ Version von SCHREIER-SPERNER [S-S]. Dieser Wandel ist vergleichbar mit dem Übergang vom Jugendstil zum Bauhaus. Inzwischen ist die lineare Algebra durchdrungen von einer Ökonomie der Gedanken sowie einer Ästhetik in der Darstellung, und sie ist unentbehrliches Hilfsmittel in vielen anderen Gebieten geworden, etwa der Analysis und der angewandten Mathematik.

Dieser eindrucksvolle Fortschritt ist nicht frei von Gefahren. Die Axiomatik beginnt mit den allgemeinsten Situationen und schreitet fort in Richtung zu spezielleren Sachverhalten. Dieser Weg wurde mit letzter Konsequenz in den Werken von N. BOURBAKI [Bo], [C] beschritten. Er läuft der historischen Entwicklung – die einem „natürlichen Wachstum“ der Mathematik entspricht – jedoch meist entgegen. So wurden etwa *Determinanten* schon von LEIBNIZ um 1690 benutzt, CAYLEY begann 1850 *Matrizen* als eigenständige Objekte anzusehen, der allgemeine Begriff des *Körpers* ist erstmals in dem 1895 bei Vieweg erschienenen „Lehrbuch der Algebra“ von H. WEBER [We] zu finden. Abstrakte Begriffe und ihre Axiome entstehen aus der Entdeckung von Gemeinsamkeiten, sie setzen lange Erfahrung im naiven Umgang und kreative Auseinandersetzung mit den Gegenständen der Mathematik voraus. Eine Darstellung, die mit den Axiomen beginnt, könnte den verhängnisvollen Eindruck erwecken, als seien die aufgestellten Regeln zufällig oder willkürlich. Einer solchen Gefahr entgegenzuwirken, ist das stete Bestreben dieses Buches. Die neue Auflage soll helfen, die abstrakten Begriffe noch mehr zu motivieren und die Beziehungen der linearen Algebra zu ihren Anwendungen deutlicher zu machen.

Viele theoretische Überlegungen der linearen Algebra dienen der Rechtfertigung oder der Entwicklung von Rechenverfahren, mit deren Hilfe man schließlich gegebene Probleme durch eine Iteration lösen kann. Dies wird hier in vielen Fällen bis zur Berechnung konkreter Beispiele vorgeführt. In der Praxis läßt man besser einen Computer rechnen, aber die Schwelle zur Beschreibung von Programmen dafür wurde in diesem Buch mit Vorsatz nicht überschritten. Für einen Anfänger erscheint es mir viel wichtiger, zunächst einmal ohne Ablenkung durch Probleme der Programmierung die Struktur des Lösungsweges zu verstehen und mit einfachsten, im Kopf berechenbaren Beispielen die unmittel-

*Eckige Klammern beziehen sich auf das Literaturverzeichnis

bare gute Erfahrung zu machen, daß ein Algorithmus funktioniert. Danach kann man getrost die Ausführung der Rechnungen einem fertigen Programmpaket wie Maple oder Mathematica überlassen. Etwa im Rahmen der numerischen Mathematik hat man Gelegenheit, die Rechenverfahren genauer zu studieren und dazu weitere Hilfsmittel der linearen Algebra kennen zu lernen (vgl. etwa [Str]).

Dieses Buch ist entstanden aus Vorlesungen für Studienanfänger in den Fächern Mathematik, Physik und Informatik; an Vorkenntnissen ist nur das sogenannte „Schulwissen“ (etwa im Umfang von [Sch]) nötig. Es enthält insgesamt genügend viel Material für zwei Semester, dabei gibt es zahlreiche Möglichkeiten für Auswahl und Reihenfolge. Der Text ist meist nach den Regeln der Logik angeordnet, in einer Vorlesung kann es gute Gründe geben, davon abzuweichen. Einige Abschnitte sind durch einen Stern * markiert, als Anregung, sie beim ersten Durchgang zu überspringen und später (etwa im zweiten Semester) darauf zurückzukommen. Die Anwendungen der linearen Algebra auf affine und projektive Geometrie sowie die lineare Optimierung sind in einem eigenen Band [Fi] enthalten, auch damit kann man den vorliegenden Text nach Belieben mischen.

Um Mathematik zu verstehen, genügt es nicht, ein Buch zu lesen oder eine Vorlesung zu hören, man muß selbst an Problemen arbeiten. Als Anregung dazu dienen die zahlreichen Aufgaben. Die dort eingestreuten Sterne sind nicht als Warnung, sondern als besonderer Ansporn zu verstehen.

Der durch diese Neuauflage abgelöste Text war durch zahllose Hinweise von Lesern fast restlos von Druckfehlern befreit worden. Nun gibt es sicher wieder reichlich Nachschub, ich möchte auch die neuen Leser ermuntern, mir „Ansichtskarten“ zu schreiben.

Mein Dank gilt all denen, die bei der Neubearbeitung beteiligt waren: In erster Linie Hannes Stoppel, durch dessen Begeisterung, Bücher zu $\text{\LaTeX}$ -en, ich in dieses Projekt geschlittert bin, Martin Gräf, der mit viel Sorgfalt die Übungsaufgaben zusammengestellt hat, Carsten Töller, dem einfallsreichen Meister der Bilder und dem Verlag für seine stetige Unterstützung.

Düsseldorf, im September 1995

Gerd Fischer

Vorwort zur 17. Auflage

Seit der 10. Auflage hat es nur wenige größere Änderungen gegeben. Sie betreffen Ergänzungen zu Quotientenräumen, Tensorprodukten und äußeren Produkten. Diese grundlegenden abstrakten Begriffe bereiten erfahrungsgemäß Studienanfängern einige Schwierigkeiten; man sollte ihnen jedoch nicht zu lange ausweichen, denn sie treten später als unentbehrliches Hilfsmittel an vielen Stellen wieder auf. Zudem sind sie im hier behandelten Fall von Vektorräumen mit Hilfe von Basen noch recht konkret zu beschreiben.

Birgit Griesse und Hannes Stoppel haben die Aufgaben teilweise überarbeitet und ein eigenes Buch mit Lösungen veröffentlicht. Das sollte die Leser nicht davon abhalten, zunächst selbst daran zu arbeiten.

Neu in dieser Auflage ist eine Einführung zum Thema *Warum Lineare Algebra?* Sie ist die überarbeitete Fassung einer einführenden Vorlesungs-Doppelstunde, die GÜNTER M. ZIEGLER im April 2006 an der TU Berlin gehalten hat. Zum genaueren Verständnis der darin ausgeführten Beispiele sind Kenntnisse über die einfachsten Begriffe und Techniken der Linearen Algebra nützlich. Wer damit noch gar nicht vertraut ist, kann Einzelheiten im Laufe seines Studiums der Linearen Algebra nachlesen.

Zu vielen Themen der Linearen Algebra findet man interaktive Visualisierungen unter

www.mathe-vital.de

Diese wurden von JÜRGEN RICHTER-GEBERT an der TU München im Rahmen des Projektes *mathe-vital* mit dem Programm *Cinderella* erstellt

Mein Dank gilt den zahlreichen Lesern, die mich auf Unklarheiten und Fehler hingewiesen haben, ganz besonders Kollegen und Studierenden aus Berlin, München und Regensburg. Ich habe versucht, die nötigen Verbesserungen einzuarbeiten. Es wird nun schwer sein, noch etwas zu finden; wer dennoch Erfolg gehabt hat, wird gebeten, mir das mitzuteilen.

München, im Juni 2009

Gerd Fischer

Warum Lineare Algebra?

(von Gerd Fischer und Günter M. Ziegler)

Was ist Lineare Algebra?

Die Lineare Algebra gehört neben der reellen Analysis zum Curriculum für Studierende der Mathematik und anderer Fächer, die mathematische Methoden intensiv benutzen. Das liegt daran, dass sie zu den Grundpfeilern der Mathematik zählt, auf denen alles andere aufbaut. Zu den darüberliegenden Gebäudeteilen der Mathematik gehören beispielsweise Algebra, Differentialgleichungen, Numerik, Differentialgeometrie, Funktionalanalysis usw. Die Beziehungen und Verwindungen zwischen all diesen Gebieten sind vielfältig und schwer schematisch zu skizzieren. Aber Konsens besteht, dass Lineare Algebra zur unverzichtbaren Basis gehört. Sie ist entstanden aus der Aufgabe, lineare Gleichungssysteme zu lösen und solche Aufgaben treten in allen Gebieten der Mathematik und ihren Anwendungen immer wieder auf. Wie schon in der Einleitung erwähnt wurde, hat es lange gedauert, bis die Lineare Algebra als eigenständiges Gebiet in die Lehrpläne aufgenommen wurde. Lange Zeit wurde sie vorwiegend als technisches Hilfsmittel der Geometrie angesehen, zur Beschreibung von linearen Gebilden wie Geraden und Ebenen, sowie Kegelschnitten. Eine der ersten zusammenfassenden aber wenig beachteten Darstellungen war die 1844 in Leipzig erschienene „*Ausdehnungslehre*“ von H. GRASSMANN. Erst in der Göttinger Schule wurden die abstrakten Hintergründe konsequent herausgearbeitet, und Vektorräume als die wesentlichen Objekte der Untersuchung eingeführt. Neben dem Buch von SCHREIER und SPERNER [S-S] ist hierzu auch die 1931 erstmals erschienene „*Moderne Algebra*“ von B. L. VAN DER WAERDEN hervorzuheben. Bis in die 50er Jahre des vorigen Jahrhunderts hatten die entsprechenden Anfängervorlesungen meist noch den Titel „*Analytische Geometrie*“, erst danach wurde die Geometrie als eine von mehreren möglichen Anwendungen in den Hintergrund gedrängt. Sorgfältige historische Hinweise zu dieser langen Entwicklung findet man bei BRIESKORN [B].

Seither wird in der Ausbildung in Linearer Algebra neben der Beschäftigung mit linearen Gleichungssystemen auch der Umgang mit abstrakten mathematischen Strukturen, wie Gruppen, Ringen, Körpern, Vektorräumen usw. geübt. Dabei muss man zunächst die mathematische Sprache lernen, d.h. präzise Formulierungen finden, mit denen Strukturen definiert sind, sowie korrekte Behauptungen darüber aufstellen, und stichhaltige Begründungen dafür erarbeiten. Die Anschauung kann dabei helfen, Beweise zu finden; aber dann beginnt die Knochenarbeit, sie präzise aufzuschreiben. Das ist erfahrungsgemäß die größte Hürde für Studienanfänger.

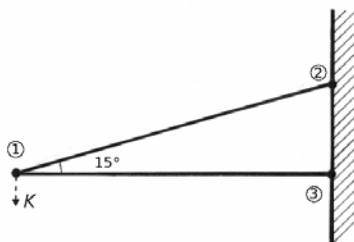
Anwendungen der Linearen Algebra

Man sollte sich nicht der Illusion hingeben, dass ein einzelnes mathematisches Teilgebiet, wie die Lineare Algebra, die Hilfsmittel zur Lösung großer praktischer Probleme liefern könnte. Wenn Mathematik in die Praxis geht, dann gehen da immer verschiedene mathematische Teilgebiete gemeinsam. Aber trotzdem: Es gibt sehr markante Beispiele für Anwendungen der Linearen Algebra - ein paar wollen wir im Folgenden beschreiben.

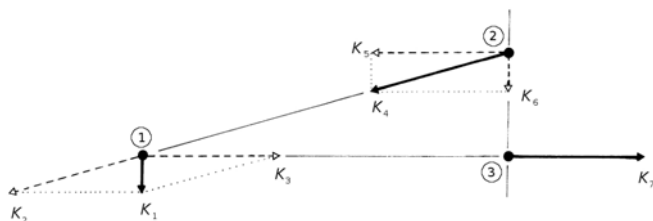
1 Statik von Gerüsten. Das Problem, Bauwerke und andere Konstruktionen auszuführen, die stabil sind, ist alt und bei weitem nicht trivial. Betrachten wir etwa ein **Gerüst**, d.h. ein Gebilde, das aus Streben und Knoten besteht (in der Baustatik spricht man von einem **Fachwerk**). Soll es stabil gebaut werden, so muss man wissen, welche Kräfte auf die Bauteile wirken. Grundlegende Untersuchungen dazu hat u.a. J. C. MAXWELL (1831 - 1879) geleistet, den man vor allem wegen seiner Beiträge zur Elektrodynamik kennt; dann aber auch C. CULMANN, von dem 1866 das Buch „*Die Graphische Statik*“ erschien. Die Methode benutzt Lineare Algebra, sein Schüler M. KOECHLIN hat als Ingenieur die Statik des Eiffel-Turms gerechnet, der seit der Weltausstellung 1889 noch heute steht.

Wir wollen die Methode in ihrer graphischen und ihrer rechnerischen Form an einem ganz einfachen, aber doch charakteristischen Beispiel illustrieren.

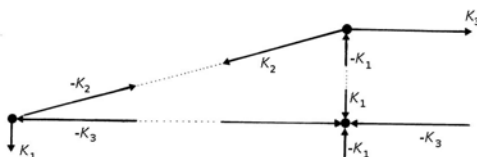
Zur Vereinfachung betrachten wir ein ebenes Problem, nämlich die Aufhängung eines Gewichtes (etwa einer Lampe) an einer Mauer mit Hilfe eines Gestänges in Form eines rechtwinkligen Dreiecks. Die Stangen und ihre Befestigungen müssen so ausgelegt sein, dass sie den entstehenden Zug- und Druckkräften standhalten. Zunächst wirkt im Punkt ① eine Kraft K senkrecht nach unten, sie soll groß sein im Vergleich zum Gewicht der Stangen.



Kräfte addieren sich wie Vektoren, im Punkt ① ist $K = K_1$ Summe von K_2 und K_3 ; K_2 verursacht einen Zug in Richtung ②, K_3 einen Druck in Richtung ③. Graphisch kann man K_2 und K_3 durch eine Parallelogrammkonstruktion ermitteln.



Im Punkt ② zerfällt die Zugkraft $K_4 = K_2$ in $K_5 = -K_3$ und $K_6 = K_1$, im Punkt ③ ist die Druckkraft $K_7 = K_3$. Der gesamte Fluss von Kräften und Gegenkräften sieht dann so aus:



Gleichgewicht bedeutet, dass in jedem Befestigungspunkt die Summe aller angreifenden Kräfte verschwindet; rechnerisch ergibt sich daraus ein System von linearen Gleichungen. Dazu beschreibt man jede der beteiligten Kräfte K_i als Vektor

$$K_i = (x_i, y_i, z_i)$$

mit reellen Komponenten x_i, y_i, z_i . Im obigen ebenen Beispiel sind die $z_i = 0$, das Gleichgewicht im Punkt ① ergibt folgende Bedingungen: K_1 ist vorgegeben, etwa $K_1 = (0, -1)$, das bedeutet

$$x_1 = 0 \quad \text{und} \quad y_1 = -1.$$

Aus der Geometrie des Dreiecks und $\tan(15^\circ) \approx 0.268$ folgt

$$0.268x_3 + y_2 = 0 \quad \text{und} \quad y_3 = 0$$

Schließlich ergibt die Gleichgewichtsbedingung $K_1 - K_2 - K_3 = 0$, dass

$$x_1 - x_2 - x_3 = 0 \quad \text{und} \quad y_1 - y_2 - y_3 = 0.$$

Das sind 6 lineare Gleichungen für die 6 Komponenten der drei Kräfte, die Lösung ist

$$K_1 = (0, -1), K_2 = (-3.732, -1), K_3 = (3.732, 0).$$

Daraus ergeben sich einfach die Kräfte in den Punkten ② und ③. Wie man sieht, ist die Zugkraft auf die Befestigung im Punkt ② fast viermal so groß wie das angehängte Gewicht.

In komplizierteren Fällen sind die Gleichungssysteme zu den einzelnen Punkten stärker gekoppelt, man kann sie dann nur gemeinsam lösen. Hat man n Punkte und in jedem Punkt drei räumliche Kräfte, so ergibt das insgesamt $9n$ zu bestimmende Koordinaten. Wie groß n sein kann, sieht man nicht nur am Eiffel-Turm, sondern schon an jedem Baukran.

Neben den statischen Kräften gibt es aber auch dynamische Effekte, da durch elastische Verformungen Schwingungen ausgelöst werden können. So wird berichtet, dass im Jahr 1850 eine Brücke über den Fluss Maine bei Angers einstürzte, während Soldaten im Gleichschritt darüber marschierten. Seither ist dem Militär verboten, Brücken auf diese Art zu überqueren.

Ein aktuelleres Beispiel ist die von SIR NORMAN FOSTER und Partnern entworfene *Millenium Bridge* über die Themse in London, eine Fußgängerbrücke, die St. Paul's Cathedral mit der Tate Modern Gallery verbindet. Sie ist konzipiert als „blade of light“, die Tragseile zwischen den 144 m entfernten Pylonen haben einen Durchhang von nur 2.3 m; das Ingenieurbüro ARUP berechnete die diffizile Statik. Nach der Einweihung durch Königin ELISABETH II wurde die Brücke am 10. Juni 2000 eröffnet - am 12. Juni 2000 musste sie wieder geschlossen werden, da sie bedrohlich zu wackeln anfang; seither heißt sie „the wobbly bridge“.



Nach vielen Experimenten und Rechnungen von ARUP wurde die Ursache gefunden: Die Statik war in Ordnung, aber es entstanden seitliche Schwingungen, die durch die Reaktionen der Fußgänger noch verstärkt wurden: Durch Resonanz schaukelten sich die Schwingungen gefährlich auf. Einzelheiten dazu findet man unter [ARUP]. Der Umbau und Einbau von Schwingungsdämpfern kostete 5 Millionen Pfund. Im Februar 2002 wurde das Ergebnis mit bis zu 20 000 Freiwilligen getestet und für gut befunden. Seither ist die Brücke wieder geöffnet.



Was hat das mit Linearer Algebra zu tun? Schwingungen und ihre Dämpfung hängen mit Matrizen und ihren Eigenwerten zusammen.

2 Linearisierung. In der Praxis gibt es kaum eine strikt lineare Beziehung zwischen zwei Größen; selbst für Preise wird bei Abnahme größerer Mengen ein Nachlass gewährt. Aber die Tangente an eine Funktion gibt in einem begrenzten Bereich wenigstens eine brauchbare Näherung. Dieses Prinzip der Analysis heißt **lineare Approximation**, hier helfen die Methoden der Linearen Algebra. Man kann es ausbauen und eine gegebene oder gesuchte differenzierbare Funktion in ihrem gesamten Definitionsbereich durch eine **stückweise lineare** Funktion approximieren, etwa bei der Lösung von Differentialgleichungen durch **Diskretisierung** und **stückweise Linearisierung**.

Als einfaches Beispiel sei die Berechnung einer Wärmeverteilung in der Ebene angegeben. Bezeichnet $f(x,y)$ die Temperatur im Punkt (x,y) , so erfüllt die Funktion f bei Temperaturgleichgewicht die partielle Differentialgleichung zweiten Grades

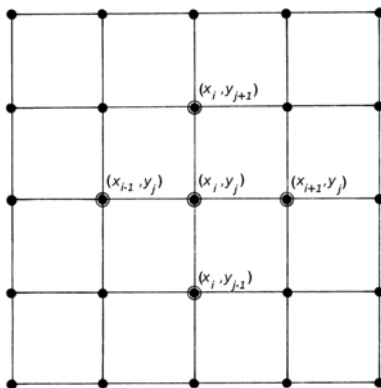
$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)f(x,y) = 0,$$

man nennt sie **Laplace-Gleichung**, die Lösungen heißen **harmonisch**. Wir nehmen an, dass die Temperaturverteilung am Rand eines Quadrats vorgegeben ist und unverändert bleibt. Eine übliche Grundlage für die approximative numerische Berechnung der Lösung f ist eine Diskretisierung: Man überzieht den quadratischen Bereich mit einem genügend feinen quadratischen Gitter von Messpunkten (x_i, y_j) . Ersetzt man in der Laplace-Gleichung Differentialquotienten durch Differenzenquotienten und schafft man die Nenner weg, so ergibt sich an der Stelle (x_i, y_j) die Bedingung

$$\begin{aligned} & (f(x_{i+1}, y_j) - f(x_i, y_j)) - (f(x_i, y_j) - f(x_{i-1}, y_j)) \\ & + (f(x_i, y_{j+1}) - f(x_i, y_j)) - (f(x_i, y_j) - f(x_i, y_{j-1})) = 0, \end{aligned}$$

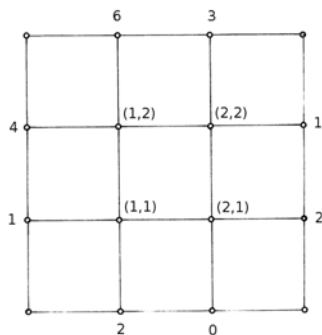
das ist gleichbedeutend mit

$$f(x_i, y_j) = \frac{1}{4}(f(x_{i+1}, y_j) + f(x_{i-1}, y_j) + f(x_i, y_{j+1}) + f(x_i, y_{j-1})). \quad (*)$$



Physikalisch bedeutet diese Bedingung, dass die Temperatur an jeder Stelle (x_i, y_j) im Inneren gleich dem Mittelwert der Temperaturen an den vier nächstgelegenen Gitterpunkten ist. Insgesamt erhält man mit Hilfe von $(*)$ für die Temperaturen $f(x_i, y_j)$ so viele lineare Gleichungen, wie man Gitterpunkte hat; dieses Gleichungssystem ist zu lösen.

Als Beispiel wählen wir ein relativ grobmaschiges Gitter, mit folgenden Werten von f auf den relevanten Gitterpunkten am Rand:



Für die vier gesuchten Werte

$$f_{ij} := f(x_i, y_j) \quad \text{mit} \quad 1 \leq i, j \leq 2$$

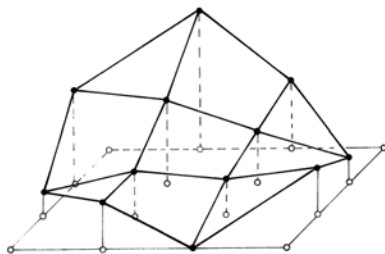
erhält man daraus die linearen Gleichungen

$$\begin{array}{rcccccl} 4f_{11} & - & f_{12} & - & f_{21} & = & 3 \\ -f_{11} & + & 4f_{12} & - & f_{22} & = & 10 \\ -f_{11} & & & + & 4f_{21} & - & f_{22} = 2 \\ & - & f_{12} & - & f_{21} & + & 4f_{22} = 4 \end{array}$$

mit den Lösungen

$$f_{11} = \frac{49}{24} \approx 2.042, \quad f_{12} = \frac{43}{12} \approx 3.583, \quad f_{21} = \frac{19}{12} \approx 1.583, \quad f_{22} = \frac{55}{24} \approx 2.292.$$

Das entstehende stückweise lineare „harmonische“ Funktionsgebirge sieht so aus:



3 Der Page Rank bei Google. Ein aktuelleres Beispiel für eine Anwendung der Linearen Algebra ist die Internet-Suchmaschine Google, die in ihrer ursprünglichen Form in den 90er Jahren von den beiden Studenten S. BRIN und L. PAGE entwickelt und 2001 patentiert wurde. Wichtiger Bestandteil ist eine Methode für die Anordnung der Suchergebnisse im Browser, Grundlage dafür ist der „PageRank“ $p(S)$ für jede Web-Site S . Er ist ein Maß dafür, wie stark die Seite mit anderen vernetzt ist, sagt allerdings nichts über die Qualität des Inhalts der Seite aus. Die Definition des PageRank kann wie folgt motiviert werden.

Man stellt sich einen Surfer vor, der einen Weg auf den vorhandenen Seiten $S_1, \dots, S_N$ des Internets zurücklegt. Er beginnt auf einer zufällig ausgewählten Seite und folgt in der Regel einem der angegebenen Links. Da er aber entmutigt werden kann (etwa weil die Links nicht mehr weiterhelfen), darf er gelegentlich auch einmal auf eine beliebige andere Seite springen. Um das zu präzisieren wird zunächst ein **Damping Faktor** d mit $0 \leq d \leq 1$ festgesetzt (meist wird $d = 0.85$ gewählt). Er hat folgenden Einfluss: Auf irgendeiner Seite angekommen folgt der Surfer mit der Wahrscheinlichkeit d einem zufällig ausgewählten Link, mit der Wahrscheinlichkeit $1 - d$ springt er vom Zufall gesteuert auf eine beliebige andere Seite des Netzes. Der **PageRank** $p(S)$ ist nun erklärt als die Wahrscheinlichkeit dafür, dass sich der Surfer auf einem sehr langen Weg zu einem zufällig gewählten Zeitpunkt auf der Seite S befindet. Da N sehr groß ist, wird $p(S)$ eine sehr kleine Zahl sein, auf jeden Fall gilt

$$0 \leq p(S) \leq 1.$$

Nach den elementaren Regeln für eine Wahrscheinlichkeitsverteilung ist

$$p(S_1) + \dots + p(S_N) = 1.$$

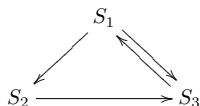
Die Wahrscheinlichkeitsrechnung ergibt nun eine Beziehung zwischen den verschiedenen PageRanks. Dazu betrachtet man für eine Seite S alle Seiten, die einen Link auf S enthalten, wir bezeichnen sie mit $T_1, \dots, T_n$, wobei $0 \leq n \leq N - 1$. Bezeichnet c_i die Anzahl der Links, die von T_i ausgehen, so gilt

$$p(S) = d \cdot \left(\frac{p(T_1)}{c_1} + \dots + \frac{p(T_n)}{c_n} \right) + \frac{1 - d}{N}.$$

Damit könnte man $p(S)$ nur ausrechnen, wenn die Werte $p(T_i)$ schon bekannt wären. Aber immerhin erhält man auf diese Weise ein System von N linearen Gleichungen für die N gesuchten Zahlen $p(S_1), \dots, p(S_N)$.

Nach der Theorie kann man ein solches Gleichungssystem lösen, aber in der Praxis benötigt man bei großem N sehr gute und schnelle numerische Verfahren. In der Gründerzeit des Internets rechnete man noch mit etwa 20 Millionen Seiten, inzwischen ist die Gesamtzahl N unüberschaubar geworden. Daher kann der PageRank nur noch für ausgewählte Seiten berechnet werden. Mehr dazu findet man bei [LM].

Um das Prinzip erläutern zu können, geben wir ein ganz einfaches Beispiel mit $N = 3$, das schematisch so aussieht:



Wie man daran erkennt, ist

$$c_1 = 2 \quad \text{und} \quad c_2 = c_3 = 1 .$$

Also lauten die drei Gleichungen für $p_i = p(S_i)$ mit $b := \frac{1}{3}(1-d)$:

$$\begin{aligned} p_1 & - dp_3 &= b \\ -\frac{d}{2}p_1 + p_2 & &= b \\ -\frac{d}{2}p_1 - dp_2 + p_3 & &= b \end{aligned}$$

Für $d = 0.85$ erhält man die Lösungen

$$p_1 \approx 0.388 , \quad p_2 \approx 0.215 , \quad p_3 \approx 0.397 .$$

Da S_2 weniger verlinkt ist, als S_1 und S_3 ist p_2 nur etwa halb so groß wie p_1 und p_3 . Bei kleinerem d haben die Links weniger Einfluss. Etwa für $d = 0.1$ wird

$$p_1 \approx 0.335 , p_2 \approx 0.317 , p_3 \approx 0.348 ,$$

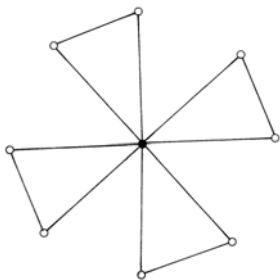
da sind die PageRanks schon beinahe gleichverteilt.

Eine Variante des Gleichungssystems erhält man mit Hilfe der **Linking Matrix** A . Bezeichnet c_i für $1 \leq i \leq N$ die Anzahl der Links, die von der Seite S_i ausgehen, so hat A für $1 \leq i, j \leq N$ die Einträge

$$a_{ij} := \begin{cases} \frac{d}{c_i} , & \text{falls } i \neq j \text{ und } S_i \text{ einen Link auf } S_j \text{ enthält;} \\ 0 & \text{sonst.} \end{cases}$$

Im Extremfall $d = 1$ ist dann $(p(S_1), \dots, p(S_N))$ ein Eigenvektor der Matrix A zum Eigenwert 1. Zur Lösung solcher Probleme gibt es auch schnelle numerische Verfahren.

4 Der Satz vom Politiker. Eine ganz andersartige Fragestellung betrifft **Graphen**, das sind Konfigurationen, die aus Punkten und Verbindungsgeraden (oder **Ecken** und **Kanten**) bestehen. Ein Beispiel ist der **Windmühlengraph**.



Er hat eine zentrale Ecke und eine gerade Zahl von Ecken am Rand, kann also für jede ungerade Zahl von Ecken konstruiert werden. Interpretiert man die Punkte als Personen und die Kanten als gegenseitige Freundschaften, so stehen am Rand befreundete Paare, und jeder ist mit der Person in der Mitte befreundet. Eine solche Person, die mit jedem befreundet ist, wird als **Politiker** bezeichnet, seine „Freundschaften“ sind berufsspezifisch. In dieser Interpretation hat der Windmühlengraph dann folgende Eigenschaft:

*Je zwei beliebige Personen haben immer
genau einen gemeinsamen Freund* (*)

Der **Satz vom Politiker** sagt nun aus, dass Bedingung (*) für n Personen nur dann erfüllt sein kann, wenn n ungerade ist und es einen Politiker gibt. Außerdem muss der entsprechende Graph ein Windmühlengraph sein.

Für diesen Satz gibt es elementare Beweise. Aber der klarste und überzeugendste wurde von P. ERDÖS, A. RENYI und V. SÓS mit Hilfe von Linearer Algebra, genauer Eigenwerten symmetrischer Matrizen gegeben; das findet man bei [A-Z, Kap. 34]. Der Schlüssel dazu ist die **Adjazenzmatrix** A des Graphen: Sind die Personen mit $1, \dots, n$ nummeriert, so sind die Einträge von A gegeben durch

$$a_{ij} := \begin{cases} 1, & \text{wenn } i \neq j \text{ und } i \text{ mit } j \text{ befreundet,} \\ 0 & \text{sonst.} \end{cases}$$

Fazit

Unsere kleine Liste von Beispielen für Fragen, hinter denen Lineare Algebra steckt, könnte man beliebig erweitern. Etwa in [A-B] kann man nachlesen, was in einer CD versteckt ist, in [A-Z, Kap. 15] findet man Ergebnisse zur berühmten **Borsuk-Vermutung** über die Zerlegung von Teilmengen des $\mathbb{R}^n$ mit beschränktem Durchmesser. Viele weitere Anwendungen findet man bei G. STRANG [St₁] und [St₂].

Um eine mathematische Theorie sachgemäß anwenden zu können, muss man sie zunächst sorgfältig studieren und genügend verstehen; das gilt auch für die Lineare Algebra. Daneben kann die Mathematik durch ihren klaren Aufbau und die Schönheit ihrer Strukturen begeistern; das zeigt sich zu Beginn des Studiums besonders in der Linearen Algebra, für die der Leser nun hoffentlich nachhaltig motiviert ist.

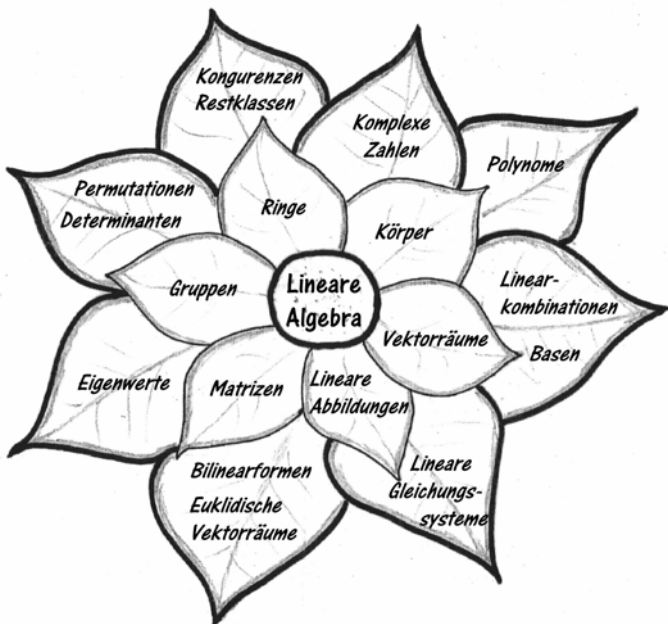
Literatur zur Einführung

- [A-B] Aigner, M. und E. Behrends (Hrsg.): *Alles Mathematik*. Vieweg 2008³
- [A-Z] Aigner, M. und G. M. Ziegler: *Das BUCH der Beweise*. Springer 2004²
- [ARUP] www.arup.com/MilleniumBridge/
- [B] Brieskorn, E.: *Lineare Algebra und Analytische Geometrie (mit historischen Anmerkungen von E. Scholz)*. Vieweg 1983
- [LM] Langville, A. N. and C. D. Meyer: *Google's PageRank and Beyond, The Science of Search Engine Rankings*. Princeton 2006
- [S-S] Schreier, O. und E. Sperner: *Einführung in die Analytische Geometrie und Algebra*. Teubner 1931
- [St₁] Strang, G.: *Linear Algebra and its Applications*. WB Saunders 2005⁴
- [St₂] Strang, G.: *Linear Algebra: A Happy Chance to Apply Mathematics*. Proc. Int. Congress on Math. Education (ICME - 10). Denmark 2004

Inhaltsverzeichnis

0	Lineare Gleichungssysteme	1
0.1	Der reelle n -dimensionale Raum	1
0.2	Geraden in der Ebene	4
0.3	Ebenen und Geraden im Standardraum $\mathbb{R}^3$	11
0.4	Das Eliminationsverfahren von GAUSS	20
1	Grundbegriffe	32
1.1	Mengen und Abbildungen	32
1.2	Gruppen	43
1.3	Ringe, Körper und Polynome	54
1.4	Vektorräume	75
1.5	Basis und Dimension	86
1.6	Summen von Vektorräumen*	100
2	Lineare Abbildungen	106
2.1	Beispiele und Definitionen	106
2.2	Bild, Fasern und Kern, Quotientenvektorräume*	114
2.3	Lineare Gleichungssysteme	129
2.4	Lineare Abbildungen und Matrizen	137
2.5	Multiplikation von Matrizen	143
2.6	Koordinatentransformationen	154
2.7	Elementarmatrizen und Matrizenumformungen	163
3	Determinanten	174
3.1	Beispiele und Definitionen	174
3.2	Existenz und Eindeutigkeit	186
3.3	Minoren*	201
3.4	Determinante eines Endomorphismus und Orientierung*	212
4	Eigenwerte	222
4.1	Beispiele und Definitionen	222
4.2	Das charakteristische Polynom	228
4.3	Diagonalisierung	234
4.4	Trigonalisierung*	242
4.5	Potenzen eines Endomorphismus*	250
4.6	Die Jordansche Normalform*	259

5	Euklidische und unitäre Vektorräume	274
5.1	Das kanonische Skalarprodukt im $\mathbb{R}^n$	274
5.2	Das Vektorprodukt im $\mathbb{R}^3$	282
5.3	Das kanonische Skalarprodukt im $\mathbb{C}^n$	286
5.4	Bilinearformen und Sesquilinearformen	288
5.5	Orthogonale und unitäre Endomorphismen	303
5.6	Selbstadjungierte Endomorphismen*	312
5.7	Hauptachsentransformation*	318
6	Dualität und Tensorprodukte*	331
6.1	Dualräume	331
6.2	Dualität und Skalarprodukte	340
6.3	Tensorprodukte	350
6.4	Multilineare Algebra	366
	Literaturverzeichnis	372
	Namensverzeichnis	374
	Sachwortverzeichnis	376
	Symbolverzeichnis	383



Diese *Blume der Linearen Algebra* wurde entworfen von Bettina Meserle, Claudia Jochum und Jonathan Zinsl.

Kapitel 0

Lineare Gleichungssysteme

Schon die Nummer dieses Kapitels deutet an, daß es zur Motivation und Vorbereitung der in Kapitel 1 beginnenden systematischen Darstellung dient. Wir haben dafür das wichtigste Problem der elementaren linearen Algebra gewählt, nämlich lineare Gleichungssysteme. Dabei kann man sehr schön die wesentlichen Aspekte vorführen: den geometrischen Hintergrund und die algorithmische Methode. Was auf die späteren Kapitel verschoben wird, sind die präzisen Beweise mit Hilfe der üblichen theoretischen Hilfsmittel.

Wer mit den verwendeten Notationen von Mengen und Abbildungen nicht vertraut ist, kann bei Bedarf in 1.1 nachsehen.

0.1 Der reelle n -dimensionale Raum

Ein großer Fortschritt in der Geometrie gelang durch Einführung von Koordinaten, die man zur Erinnerung an R. DESCARTES auch *kartesische* Koordinaten nennt. Dadurch kann man Punkte in den Räumen der Geometrie beschreiben durch Systeme von Zahlen, und mit den Zahlen kann man rechnen. Diese Methode zur Behandlung geometrischer Fragen nennt man *analytische Geometrie*. Die elementarsten Begriffe hierzu seien kurz erklärt.

0.1.1. Wir gehen aus von den *reellen Zahlen*, deren Gesamtheit wir mit $\mathbb{R}$ bezeichnen. Ihre Einführung ist Gegenstand der Analysis, in der Geometrie dienen sie als *Zahlengerade*, und diese Zahlen kann man nach den üblichen Regeln addieren und multiplizieren.

Punkte der Ebene sind festgelegt durch Paare, Punkte des gewöhnlichen Raumes durch Tripel von reellen Zahlen. Für die Theorie macht es keine Probleme, gleich n -Tupel zu betrachten, wobei n eine beliebige natürliche Zahl ist. Damit erhält man den *reellen Standardraum der Dimension n*

$$\mathbb{R}^n = \{x = (x_1, \dots, x_n) : x_1, \dots, x_n \in \mathbb{R}\},$$

d.h. die Menge der *geordneten n -Tupel* (oder *Vektoren*) von reellen Zahlen. Geordnet heißt, daß die Reihenfolge wichtig ist, d.h. zwei n -Tupel $(x_1, \dots, x_n)$ und $(y_1, \dots, y_n)$ sind genau dann gleich, wenn $x_1 = y_1, \dots, x_n = y_n$. Die Zahlen $x_1, \dots, x_n$ heißen *Komponenten* von x .

Der Fall $n = 0$ ist sinnlos, $\mathbb{R}^1$ ist die Zahlengerade, $\mathbb{R}^2$ entspricht der Ebene und $\mathbb{R}^3$ dem „Raum“. Für größere n hat man zunächst keine unmittelbare geometrische Vorstellung mehr, dafür aber eine ganz andere und sehr realistische Interpretation. Hat etwa eine Bank n Kunden, so kann man deren Kontostände zu einem bestimmten Zeitpunkt mit $x_1, \dots, x_n$ bezeichnen, alle zusammen (und

geordnet!) sind ein „Punkt“

$$x = (x_1, \dots, x_n) \in \mathbb{R}^n.$$

Die Entwicklung der Kontostände im Laufe der Zeit wird dann durch eine „Kurve“ im $\mathbb{R}^n$ beschrieben, ihre Beschreibung geht schon über den Rahmen der linearen Algebra hinaus. Eine lineare Operation ist etwa die Berechnung der augenblicklichen Bilanz. Haben die Einlagen neben dem Nennwert x_i einen Börsenkurs a_i , so ist das bewertete Kapital gegeben durch

$$a_1 x_1 + \dots + a_n x_n = b,$$

ein typischer Fall für eine *lineare Gleichung*.

In der Praxis hat man mehrere Bedingungen, und diese sind meist nicht durch Gleichungen, sondern durch Ungleichungen gegeben, von der Art, daß für die obige Summe Begrenzungen vorgegeben sind. Und das Problem besteht darin, einen „Gewinn“ zu optimieren. Zur Lösung solcher Aufgaben in der *linearen Optimierung* (vgl. etwa [Fi]) benötigt man genügende Kenntnisse über lineare Gleichungen, was vielleicht auch einen gewinnorientierten Leser eine Zeit lang bei der Stange halten kann.

0.1.2. In der linearen Algebra muß man mit n -Tupeln rechnen. Die grundlegenden Operationen sind eine *Addition*

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

und eine *Multiplikation mit einer Zahl* $\lambda \in \mathbb{R}$

$$\lambda \cdot (x_1, \dots, x_n) := (\lambda \cdot x_1, \dots, \lambda \cdot x_n).$$

Man kann diese Operationen geometrisch deuten, wenn man die n -Tupel als Vektoren ansieht, d.h. naiv als Pfeile vom *Ursprung* $0 = (0, \dots, 0)$ mit Spitze in $x = (x_1, \dots, x_n)$. Für $n = 2$ kann man das einfach zeichnen:

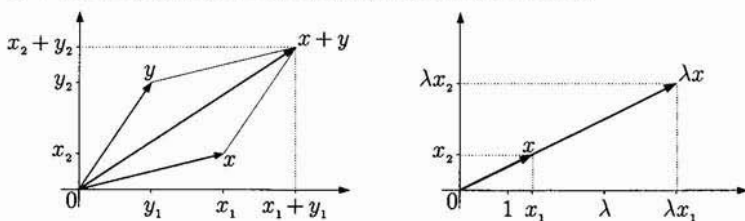


Bild 0.1

Der Ursprung selbst heißt auch *Nullvektor*, wenn man ihn addiert, hat das keine Wirkung. Multipliziert man mit $\lambda = 0$, so wird jedes x zum Nullvektor. Das Negative von x ist gegeben durch

$$-x := (-x_1, \dots, -x_n),$$

es gilt $x + (-x) = 0$. Statt $x + (-y)$ schreibt man kürzer $x - y$.

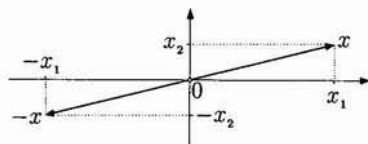


Bild 0.2

Nach diesen wenigen Formalitäten können wir nun die einfachsten Beispiele von linearen Gleichungen behandeln. Um die geometrische Anschauung dabei zu benutzen, betrachten wir zunächst ausführlich die Fälle $n = 2$ und $n = 3$.

0.2 Geraden in der Ebene

0.2.1. Durch zwei verschiedene Punkte geht genau eine *Gerade*, das gehört zu den wenigen Tatsachen der Geometrie, die auch Nicht-Mathematikern einleuchten. Mit Hilfe von Vektoren kann man das so beschreiben: Sind $v, v' \in \mathbb{R}^2$ die beiden Punkte, so ist $w := v' - v \neq 0$. Die Punkte auf der Geraden L durch v und v' kann man unter Benutzung eines reellen *Parameters* λ darstellen als

$$L = \{x \in \mathbb{R}^2 : \text{es gibt ein } \lambda \in \mathbb{R} \text{ mit } x = v + \lambda w\} =: v + \mathbb{R}w.$$

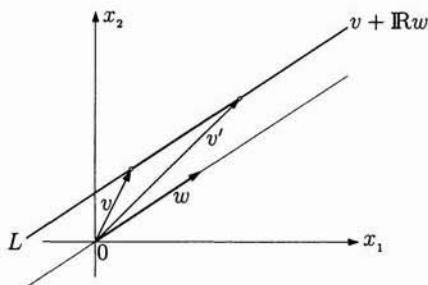


Bild 0.3

Man kann L auch ansehen als Bild der Zahlengerade $\mathbb{R}$ unter der Abbildung

$$\Phi: \mathbb{R} \rightarrow L \subset \mathbb{R}^2, \quad \lambda \mapsto v + \lambda w.$$

Das nennt man eine *Parametrisierung* der Geraden.

0.2.2. Die zweite Möglichkeit der Beschreibung benutzt eine *lineare Gleichung* der Form

$$a_1 x_1 + a_2 x_2 = b.$$

Dabei gelten x_1, x_2 als *Unbestimmte* und $a_1, a_2 \in \mathbb{R}$ als *Koeffizienten*. Die Unbestimmten sind variabel, die Koeffizienten fest. Man betrachtet die Menge der *Lösungen*

$$L := \{(x_1, x_2) \in \mathbb{R}^2 : a_1 x_1 + a_2 x_2 = b\}.$$

Ist $a_1 = a_2 = 0$, so ist $L = \emptyset$ für $b \neq 0$ und $L = \mathbb{R}^2$ für $b = 0$. Dieser Fall gilt als „entartet“. Andernfalls müßte man im Prinzip alle Paare (x_1, x_2) in die Gleichung einsetzen und feststellen, ob sie erfüllt ist. Wenn man das ohne System tut, wird man dabei sehr selten Glück haben.

Ein gutes System ist eine Parametrisierung, mit deren Hilfe sich alle Lösungen produzieren lassen. Das ist in diesem Fall leicht zu erhalten.

1) Ist $a_2 = 0$ und $a_1 \neq 0$, so wird die gegebene Gleichung zu

$$x_1 = \frac{b}{a_1},$$

das ist eine zur x_2 -Achse parallele Gerade, und eine Parametrisierung ist gegeben durch

$$\mathbb{R} \rightarrow L, \quad \lambda \mapsto \left(\frac{b}{a_1}, \lambda \right).$$

Hier ist also die erste Koordinate fest, die zweite frei wählbar.

Ist $a_1 = 0$, aber $a_2 \neq 0$, so hat man nur die Rollen von x_1 und x_2 zu vertauschen.

2) Ist $a_1 \neq 0$ und $a_2 \neq 0$, so kann man die Gerade leicht zeichnen, indem man die Schnittpunkte mit den Achsen $x_1 = 0$ und $x_2 = 0$ berechnet:

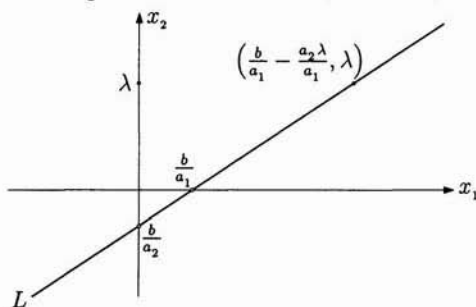


Bild 0.4

Wählt man wieder die x_2 -Koordinate eines Punktes der Geraden als Parameter λ , so kann man daraus

$$x_1 = \frac{b}{a_1} - \frac{a_2 \lambda}{a_1}$$

berechnen, und eine Parametrisierung der zunächst durch die Gleichung gegebenen Geraden ist gefunden:

$$\mathbb{R} \rightarrow L, \quad \lambda \mapsto \left(\frac{b}{a_1} - \frac{a_2 \lambda}{a_1}, \lambda \right).$$

0.2.3. Zwei Geraden in der Ebene schneiden sich in genau einem Punkt, es sei denn, sie sind gleich oder parallel. Sind sie durch Gleichungen gegeben, so stellt sich die Frage, wie man entscheiden kann, welcher Fall vorliegt, und wie man eventuell den eindeutigen Schnittpunkt findet. Dazu einige

Beispiele. a) Die Geraden seien gegeben durch

$$\begin{aligned}x_1 - x_2 &= 1, \\x_2 &= 2.\end{aligned}$$

Der Schnittpunkt p ist ganz einfach zu finden, indem man $x_2 = 2$ aus der zweiten Gleichung in die erste einsetzt: $x_1 = 1 + 2 = 3$, also $p = (3, 2)$.

Eine Variante sind die Gleichungen

$$\begin{aligned}x_1 - x_2 &= 1, \\x_1 + 3x_2 &= 9.\end{aligned}$$

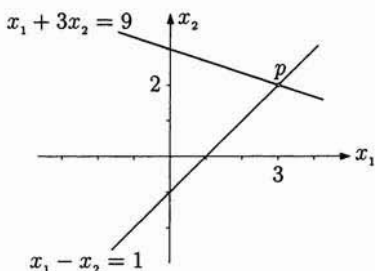
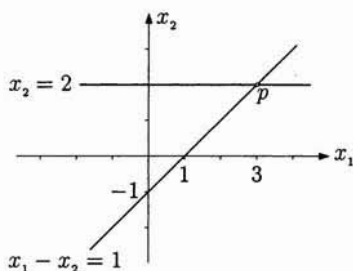


Bild 0.5

Zieht man die erste Gleichung von der zweiten ab und dividiert die Differenz durch vier, so erhält man wieder die obigen Gleichungen, und man sieht an den Zeichnungen, daß der Schnittpunkt der gleiche ist.

b) Die Geraden seien gegeben durch

$$\begin{aligned}x_1 - x_2 &= 1, \\2x_1 - 2x_2 &= b,\end{aligned}$$

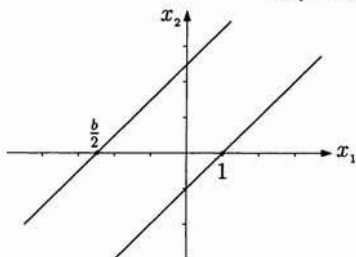


Bild 0.6

mit beliebigem b . Man sieht sofort, daß sie für $b = 2$ gleich und für $b \neq 2$ parallel, also ohne Schnittpunkt sind. Darauf kommt man auch durch formales

Rechnen, wenn man wieder die 2-fache erste Gleichung von der zweiten abzieht. Das ergibt

$$\begin{aligned}x_1 - x_2 &= 1, \\ 0x_1 - 0x_2 &= b - 2.\end{aligned}$$

Die zweite Gleichung lautet in jedem Fall $b = 2$. Ist b so gewählt, kann man sie weglassen und es bleibt die erste. Ist $b \neq 2$ gewählt, so ist die zweite Gleichung nie erfüllt und man kann die erste weglassen.

c) Nun nehmen wir zu den zwei Geraden aus Beispiel a) eine dritte hinzu:

$$\begin{aligned}x_1 - x_2 &= 1, & \text{I} \\ x_1 + 3x_2 &= 9, & \text{II} \\ x_1 + x_2 &= 2. & \text{III}\end{aligned}$$

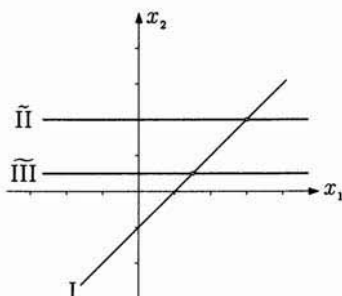
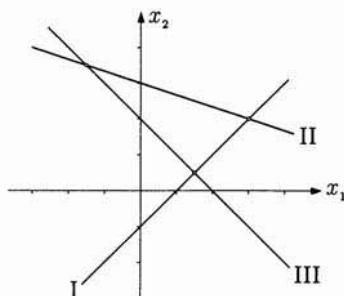


Bild 0.7

Daß sie keinen gemeinsamen Schnittpunkt haben, sieht man an Bild 0.7, wir wollen es auch durch formales Umformen zeigen. Wie umgeformt wurde, ist rechts vermerkt.

$$\begin{aligned}x_1 - x_2 &= 1, & \text{I} \\ 4x_2 &= 8, & \widetilde{\text{II}} = \text{II} - \text{I} \\ 2x_2 &= 1. & \widetilde{\text{III}} = \text{III} - \text{I}\end{aligned}$$

Die Gleichungen $\widetilde{\text{II}}$ und $\widetilde{\text{III}}$ verlangen $x_2 = 2$ und $x_2 = \frac{1}{2}$, das ist ein Widerspruch.

Wie man sieht, sind derartige Umformungen von Gleichungssystemen sehr wirksam, wir werden in 0.4.6 eine Rechtfertigung dafür geben. In obigem Beispiel sollte man bemerken, daß die gemeinsamen Schnittpunkte von I mit II und I mit III erhalten bleiben.

0.2.4. Den Begriff Gerade hatten wir bisher noch nicht präzise erklärt, das soll schleunigst nachgeholt werden:

Definition. Eine Teilmenge $L \subset \mathbb{R}^2$ heißt *Gerade*, wenn es $a_1, a_2, b \in \mathbb{R}$ mit $(a_1, a_2) \neq (0, 0)$ gibt, so daß

$$L = \{(x_1, x_2) \in \mathbb{R}^2 : a_1 x_1 + a_2 x_2 = b\}.$$

Daß man eine Gerade genauso gut mit Hilfe einer Parametrisierung beschreiben kann, ist die Aussage von folgendem

Satz. Eine Teilmenge $L \subset \mathbb{R}^2$ ist genau dann eine Gerade, wenn es $v, w \in \mathbb{R}^2$ mit $w \neq 0$ gibt, so daß

$$L = v + \mathbb{R}w.$$

Dieser Satz folgt später sehr leicht aus der allgemeinen Theorie. Ohne weitere Hilfsmittel ist der Beweis etwas mühsam, wir wollen es dennoch vorführen:

1) Ist L eine Gerade im Sinne der obigen Definition, so gibt es v und w mit den angegebenen Eigenschaften.

Sei also L gegeben durch a_1, a_2, b mit $(a_1, a_2) \neq (0, 0)$. Wir führen den Fall $a_1 \neq 0$ aus, der Fall $a_2 \neq 0$ geht analog. Indem man im Ergebnis von 0.2.2 $\lambda = 0$ und $\lambda = a_1$ setzt (siehe Bild 0.8), kommt man zu der Definition

$$v := \left(\frac{b}{a_1}, 0 \right), \quad w := (-a_2, a_1), \quad L' := v + \mathbb{R}w,$$

und es ist zu zeigen, daß $L = L'$. Dazu ist $L \subset L'$ und $L' \subset L$ zu beweisen.

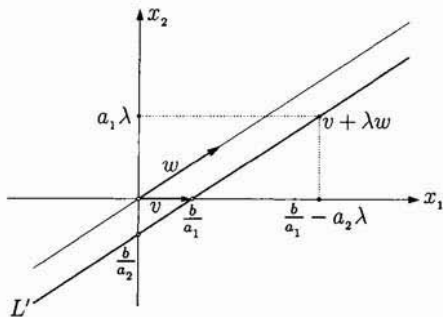


Bild 0.8

a) $L \subset L'$: Ist $(x_1, x_2) \in L$, so ist $a_1x_1 + a_2x_2 = b$. Also gilt für $\lambda := x_2/a_1$

$$(x_1, x_2) = \left(\frac{b}{a_1} - a_2\lambda, a_1\lambda \right) = v + \lambda w.$$

Somit ist $(x_1, x_2) \in L'$.

b) $L' \subset L$: Ist $x \in L'$, so gibt es ein λ mit

$$x = v + \lambda w = \left(\frac{b}{a_1} - a_2\lambda, a_1\lambda \right) = (x_1, x_2).$$

Setzt man $x_1 = \frac{b}{a_1} - a_2\lambda$ und $x_2 = a_1\lambda$ in die Gleichung von L ein, so erhält man

$$a_1x_1 + a_2x_2 = b - a_1a_2\lambda + a_1a_2\lambda = b.$$

Also ist $x \in L$.

2) Ist $L = v + \mathbb{R}w$, so ist eine Gleichung zu finden. Ist $v = (v_1, v_2)$ und $w = (w_1, w_2)$ mit $w_1 \neq 0$, so zeigt eine einfache Überlegung (siehe Bild 0.9), daß

$$\frac{x_2 - v_2}{x_1 - v_1} = \frac{w_2}{w_1}, \quad \text{also} \quad w_2x_1 - w_1x_2 = w_2v_1 - w_1v_2$$

sein muß. Wir definieren daher $a_1 := w_2$, $a_2 := -w_1$, $b := w_2v_1 - w_1v_2$ und

$$L' := \{(x_1, x_2) \in \mathbb{R}^2 : a_1x_1 + a_2x_2 = b\}.$$

a) $L \subset L'$: Dazu muß man $x = v + \lambda w = (v_1 + \lambda w_1, v_2 + \lambda w_2) \in L$ in die Gleichung einsetzen. Das ergibt

$$w_2(v_1 + \lambda w_1) - w_1(v_2 + \lambda w_2) = w_2v_1 - w_1v_2.$$

Also ist $x \in L'$.

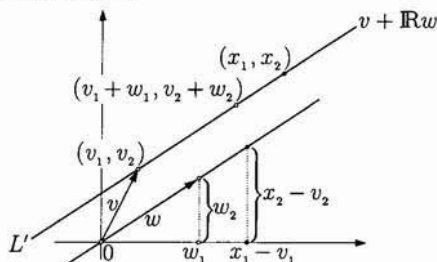


Bild 0.9

b) $L' \subset L$: Sei $x = (x_1, x_2) \in L'$. Dazu muß man ein $\lambda \in \mathbb{R}$ finden, so daß

$$x = (x_1, x_2) = v + \lambda w = (v_1 + \lambda w_1, v_2 + \lambda w_2).$$

Wegen $w_1 \neq 0$ kann man aus der ersten Komponente ausrechnen, daß

$$\lambda = \frac{x_1 - v_1}{w_1}$$

sein muß. Definiert man λ durch diese Gleichung, so folgt aus $x \in L'$, daß auch

$$x_2 = \frac{b - a_1 x_1}{a_2} = \frac{w_2 v_1 - w_1 v_2 - w_2 x_1}{-w_1} = v_2 + \lambda w_2.$$

Also ist $x \in L$.

□

0.3 Ebenen und Geraden im Standardraum $\mathbb{R}^3$

0.3.1. Als „Raum“ betrachten wir den dreidimensionalen reellen Raum. Wie in der Ebene geht durch zwei verschiedene Punkte $v, v' \in \mathbb{R}^3$ genau eine Gerade L . Ist $w := v' - v$, so ist wie in 0.2.1

$$L = \{v + \lambda w : \lambda \in \mathbb{R}\} = v + \mathbb{R}w$$

eine *Parameterdarstellung* von L . Es gibt jedoch zwei wesentliche Unterschiede zur Ebene:

- 1) Zwei Geraden sind im allgemeinen *windschief*, d.h. ohne Schnittpunkt und nicht parallel.
- 2) Eine lineare Gleichung im $\mathbb{R}^3$ beschreibt eine Ebene, zur Beschreibung einer Geraden benötigt man zwei lineare Gleichungen (d.h. man stellt sie als Schnitt von zwei Ebenen dar).

Das wird im folgenden genauer ausgeführt.

0.3.2. Wir betrachten eine lineare Gleichung der Form

$$a_1x_1 + a_2x_2 + a_3x_3 = b$$

und die Menge der *Lösungen*

$$E = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : a_1x_1 + a_2x_2 + a_3x_3 = b\}.$$

Ist $a_1 = a_2 = a_3 = 0$, so ist $E = \emptyset$ für $b \neq 0$ und $E = \mathbb{R}^3$ für $b = 0$. Für $a_1 = a_2 = 0$, aber $a_3 \neq 0$ lautet die Gleichung

$$x_3 = \frac{b}{a_3},$$

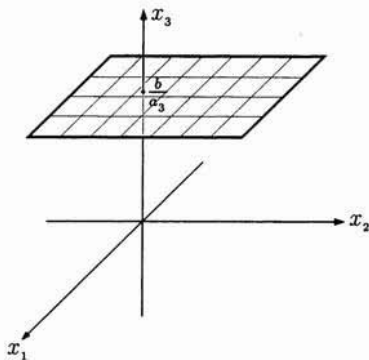


Bild 0.10

das ist eine Ebene parallel zur Ebene $x_3 = 0$.

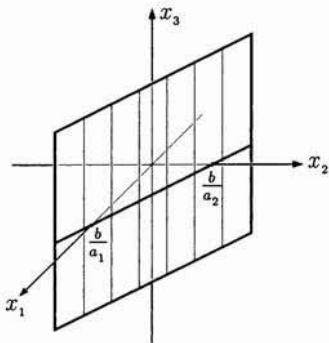


Bild 0.11

Ist $a_3 = 0$ und $(a_1, a_2) \neq (0, 0)$, so haben wir eine Gerade

$$L = \{(x_1, x_2) \in \mathbb{R}^2 : a_1 x_1 + a_2 x_2 = b\},$$

und E entsteht daraus, indem man alle im Raum darüber- und darunterliegenden Punkte dazunimmt, d.h. (siehe Bild 0.11)

$$E = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : (x_1, x_2) \in L\}.$$

Sind alle drei Koeffizienten $a_i \neq 0$, so schneidet E die Achsen an den Stellen b/a_i .

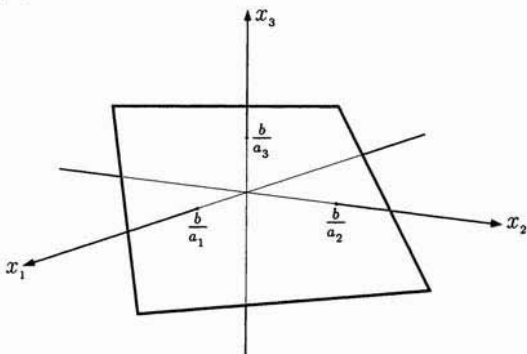


Bild 0.12

Nach diesen Beispielen eine präzise

Definition. Eine Teilmenge $E \subset \mathbb{R}^3$ heißt *Ebene*, wenn es $a_1, a_2, a_3, b \in \mathbb{R}$ mit $(a_1, a_2, a_3) \neq (0, 0, 0)$ gibt, so daß

$$E = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : a_1 x_1 + a_2 x_2 + a_3 x_3 = b\}.$$

0.3.3. Zur Parametrisierung einer Geraden reicht ein Parameter, bei einer Ebene geht es mit zweien, und zwar wie folgt. In der Ebenengleichung nehmen wir $a_3 \neq 0$ an (andernfalls vertausche man die Indizes). Wir betrachten die *Standardebene* $\mathbb{R}^2$ und bezeichnen die Punkte mit (λ_1, λ_2) . Setzt man $x_1 = \lambda_1, x_2 = \lambda_2$ in die Ebenengleichung ein, so erhält man

$$x_3 = \frac{1}{a_3} (b - a_1 \lambda_1 - a_2 \lambda_2).$$

Abstrakter ausgedrückt erhält man eine Abbildung

$$\Phi: \mathbb{R}^2 \rightarrow \mathbb{R}^3, \quad (\lambda_1, \lambda_2) \mapsto \left(\lambda_1, \lambda_2, \frac{b}{a_3} - \frac{a_1 \lambda_1}{a_3} - \frac{a_2 \lambda_2}{a_3} \right).$$

Eine triviale Rechnung ergibt

$$a_1 \lambda_1 + a_2 \lambda_2 + a_3 \left(\frac{b}{a_3} - \frac{a_1 \lambda_1}{a_3} - \frac{a_2 \lambda_2}{a_3} \right) = b,$$

also ist $\Phi(\mathbb{R}^2) \subset E$, und mit etwas mehr Rechnung als in 0.2.4 kann man zeigen, daß

$$\Phi: \mathbb{R}^2 \rightarrow E$$

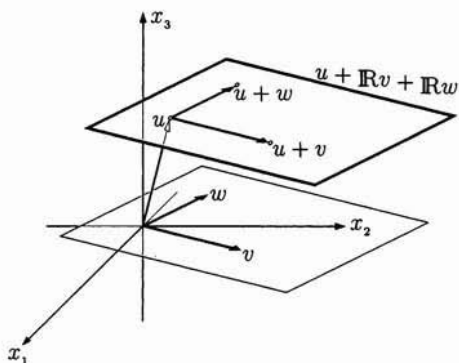
sogar bijektiv ist. Wir verzichten darauf, das hier vorzurechnen, weil es in 2.3 aus der allgemeinen Theorie folgt. Die so erhaltene Abbildung Φ von der Standardebene $\mathbb{R}^2$ auf die in $\mathbb{R}^3$ liegende Ebene E heißt *Parametrisierung*. In der oben angegebenen Form hat sie eine einfache geometrische Interpretation: Φ ist die Umkehrung der Projektion

$$\pi: E \rightarrow \mathbb{R}^2, \quad (x_1, x_2, x_3) \mapsto (x_1, x_2),$$

auf die Koordinatenebene. Natürlich gibt es viele andere Parametrisierungen von E .

Die oben mit Hilfe der Abbildung Φ angegebene Parametrisierung der Ebene E kann man noch etwas anders schreiben. Sind $u, v, w \in \mathbb{R}^3$ gegeben, so sei

$$u + \mathbb{R}v + \mathbb{R}w := \{x \in \mathbb{R}^3 : \text{es gibt } \lambda_1, \lambda_2 \in \mathbb{R} \text{ mit } x = u + \lambda_1 v + \lambda_2 w\}.$$

**Bild 0.13**

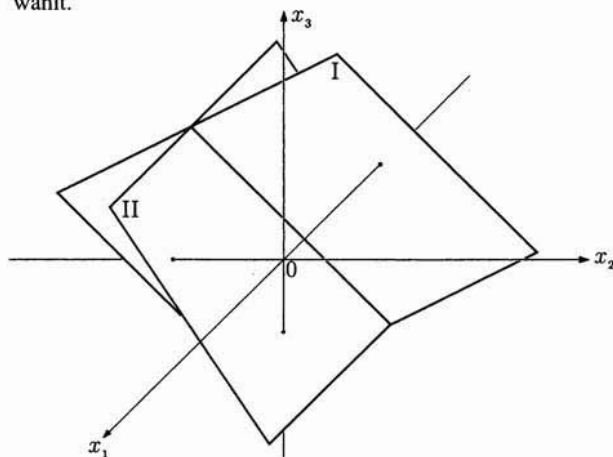
Das ist das Bild von Φ , wenn man

$$u = \Phi(0, 0) = \left(0, 0, \frac{b}{a_3}\right),$$

$$v = \Phi(1, 0) - u = \left(1, 0, -\frac{a_1}{a_3}\right),$$

$$w = \Phi(0, 1) - u = \left(0, 1, -\frac{a_2}{a_3}\right)$$

wählt.

**Bild 0.14**

0.3.4. Nun wollen wir zwei Ebenen schneiden. Dazu zunächst ein

Beispiel. Wir betrachten die Gleichungen

$$x_1 + x_2 + x_3 = -6, \quad \text{I}$$

$$x_1 + 2x_2 + 3x_3 = -10, \quad \text{II}$$

und formen sie um zu

$$x_1 + x_2 + x_3 = -6, \quad \text{I}$$

$$x_2 + 2x_3 = -4. \quad \tilde{\text{II}} = \text{II} - \text{I}$$

Der Schnitt der beiden Ebenen ist eine Gerade L , ein Punkt darauf ist festgelegt durch seine x_3 -Koordinate. Übersetzt in eine Rechnung bedeutet das: man wählt einen reellen Parameter λ , setzt $x_3 = \lambda$ und berechnet erst mit $\tilde{\text{II}}$ und dann mit I

$$x_2 = -2\lambda - 4,$$

$$x_1 = \lambda - 2.$$

Das ergibt eine Parametrisierung von L

$$\Phi: \mathbb{R} \rightarrow L \subset \mathbb{R}^3, \quad \lambda \mapsto (\lambda - 2, -2\lambda - 4, \lambda).$$

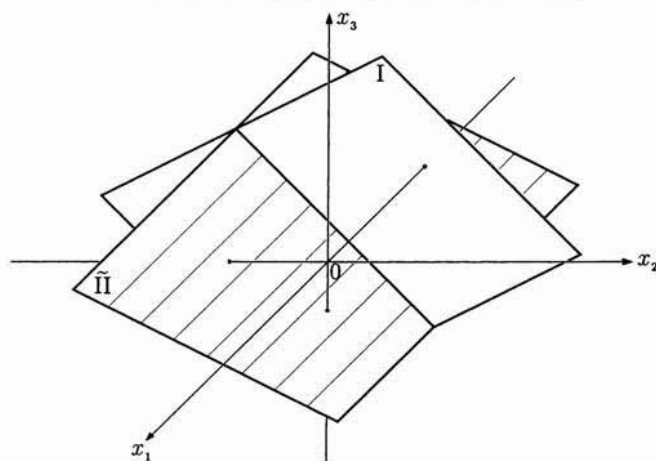


Bild 0.15

Hat man allgemein zwei Ebenengleichungen

$$a_1x_1 + a_2x_2 + a_3x_3 = b, \quad \text{I}$$

$$a'_1x_1 + a'_2x_2 + a'_3x_3 = b', \quad \text{II}$$

und ist $a_1 \neq 0$, so führt eine Umformung wie oben zu einem Unglück, wenn es ein $\varrho \in \mathbb{R}$ gibt mit

$$(a'_1, a'_2, a'_3) = (\varrho a_1, \varrho a_2, \varrho a_3). \quad (*)$$

Dann wird nämlich die linke Seite von $\tilde{\text{II}} := \text{II} - \varrho \cdot \text{I}$ gleich Null. Das hat einen geometrischen Hintergrund, denn $(*)$ bedeutet, daß die durch I und II beschriebenen Ebenen parallel oder gleich sind. Ein präziser Beweis davon kann später leicht nachgeholt werden (vgl. Aufgabe 8 zu 2.3).

0.3.5. Nun schneiden wir schließlich drei Ebenen im $\mathbb{R}^3$.

Beispiel. Wir nehmen zu dem Beispiel aus 0.3.4 noch eine Gleichung dazu:

$$\begin{array}{rcl} x_1 + x_2 + x_3 & = & -6, \quad \text{I} \\ x_1 + 2x_2 + 3x_3 & = & -10, \quad \text{II} \\ 2x_1 + 3x_2 + 6x_3 & = & -18. \quad \text{III} \end{array}$$

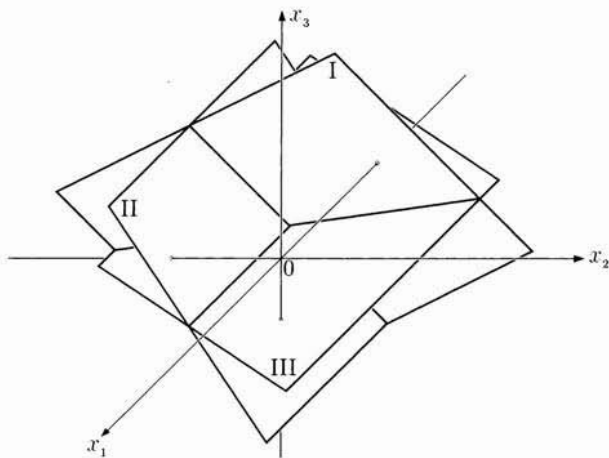


Bild 0.16

Die drei Ebenen schneiden sich in einem Punkt. Um ihn zu berechnen, formen wir wieder um. Die erste Runde ergibt

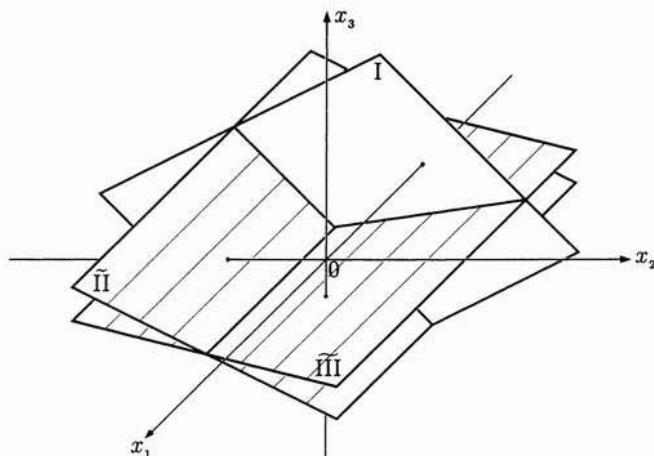


Bild 0.17

$$\begin{array}{rcl}
 x_1 + x_2 + x_3 & = & -6, \quad \text{I} \\
 x_2 + 2x_3 & = & -4, \quad \tilde{\text{II}} = \text{II} - \text{I} \\
 x_2 + 4x_3 & = & -6, \quad \tilde{\text{III}} = \text{III} - 2\text{I}
 \end{array}$$

Um den Schnittpunkt ganz schematisch ausrechnen zu können, formen wir $\tilde{\text{III}}$ noch einmal um zu

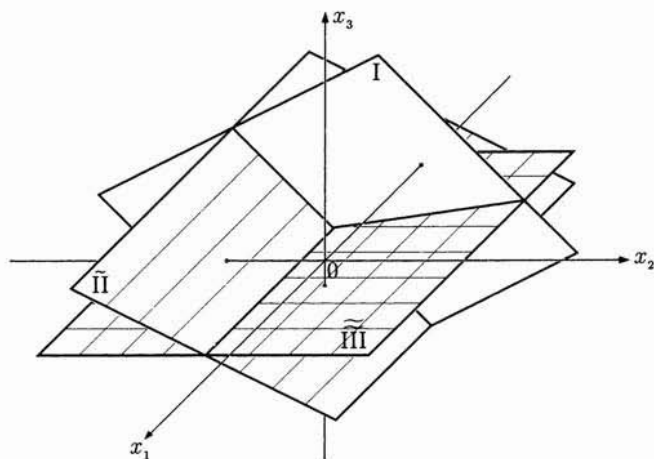
$$2x_3 = -2. \quad \tilde{\tilde{\text{III}}} = \tilde{\text{III}} - \tilde{\text{II}}$$

Damit ergibt sich durch Einsetzen von unten nach oben

$$\begin{array}{rcl}
 x_3 & = & -1 \quad \text{nach } \tilde{\tilde{\text{III}}}, \\
 x_2 & = & -2 \quad \text{nach } \tilde{\text{II}}, \\
 x_1 & = & -3 \quad \text{nach I.}
 \end{array}$$

Der einzige Schnittpunkt der drei Ebenen ist also $(-3, -2, -1)$.

Ob dieses Verfahren im allgemeinen einen einzigen Schnittpunkt liefert, hängt ganz von den Koeffizienten der drei Gleichungen ab (vgl. 2.3.6).

**Bild 0.18**

Aufgaben zu 0.3

1. Zeigen Sie, dass für zwei Punkte $v, w \in \mathbb{R}^n$ die folgenden Bedingungen äquivalent sind:

- i) $v \neq 0$, und es gibt kein $\varrho \in \mathbb{R}$ mit $w = \varrho \cdot v$.
- ii) $w \neq 0$, und es gibt kein $\varrho \in \mathbb{R}$ mit $v = \varrho \cdot w$.
- iii) Sind $\lambda, \mu \in \mathbb{R}$ mit $\lambda v + \mu w = 0$, so folgt notwendigerweise $\lambda = \mu = 0$.

Man nennt v und w *linear unabhängig*, falls eine der obigen Bedingungen erfüllt ist. v und w heißen *linear abhängig*, falls sie nicht linear unabhängig sind. Im untenstehenden Bild sind v und w linear unabhängig, v und w' linear abhängig.

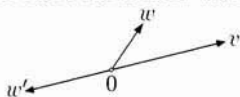


Bild 0.19

2. a) Beweisen Sie, dass eine Teilmenge E des $\mathbb{R}^3$ genau dann eine Ebene ist, wenn es Vektoren $u, v, w \in \mathbb{R}^3$ gibt, so dass v und w linear unabhängig sind und

$$E = u + \mathbb{R}v + \mathbb{R}w.$$

b) Finden Sie für die Ebene $E = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : 3x_1 - 2x_2 + x_3 = -1\}$ eine Parametrisierung.

c) Geben Sie für die in Parameterdarstellung gegebene Ebene

$$E = (1, 2, 3) + \mathbb{R} \cdot (4, 5, 6) + \mathbb{R} \cdot (7, 8, 9)$$

eine beschreibende lineare Gleichung an.

3. Zeige Sie: Sind $x, y, z \in \mathbb{R}^3$ drei Punkte, die nicht auf einer Geraden liegen, so gibt es genau eine Ebene $E \subset \mathbb{R}^3$, die x, y und z enthält, nämlich

$$E = x + \mathbb{R} \cdot (x - y) + \mathbb{R} \cdot (x - z).$$

0.4 Das Eliminationsverfahren von GAUSS

0.4.1. Nach den vielen Spezialfällen und Beispielen ist der Leser hoffentlich gerüstet für den allgemeinen Fall. Die Zahl $n \in \mathbb{N}$ ist die Zahl der Unbestimmten, davon unabhängig kann man eine Zahl $m \in \mathbb{N}$ für die Anzahl der linearen Bedingungen wählen. Da m die Anzahl der verfügbaren Buchstaben überschreiten kann, verwenden wir für die Koeffizienten a_{ij} doppelte Indizes. Das Gleichungssystem lautet also:

$$\begin{array}{ccccccc} a_{11}x_1 + & \dots & + a_{1n}x_n & = & b_1 \\ \vdots & & \vdots & & \vdots \\ a_{m1}x_1 + & \dots & + a_{mn}x_n & = & b_m, \end{array} \quad (*)$$

und gesucht ist die Menge der $(x_1, \dots, x_n) \in \mathbb{R}^n$, die alle Gleichungen erfüllen.

Das System $(*)$ ist mühsam aufzuschreiben. Ein Meister in übersichtlichen Rechenverfahren war A. CAYLEY, der auch erstmals systematisch *Matrizen* verwendete. Das hilft hier sofort. Die Koeffizienten a_{ij} schreibt man, wie sie in $(*)$ vorkommen, als rechteckiges Schema (*Matrix* genannt)

$$A := \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}.$$

Nun ist der Kniff, nicht die *liegenden* Vektoren (oder *Zeilen*) $(x_1, \dots, x_n)$ und $(b_1, \dots, b_m)$ zu betrachten, sondern entsprechend der Anordnung der b_i in $(*)$ die *stehenden* Vektoren (oder *Spalten*)

$$x := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{und} \quad b := \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Zwischen der Matrix A und der Spalte x der Höhe n erklärt man ein Produkt, das eine Spalte der Höhe m ergibt:

$$A \cdot x := \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{pmatrix}.$$

Dabei ist entscheidend, daß x so viele Zeilen wie A Spalten hat. Das *lineare Gleichungssystem* $(*)$ kann man dann in der Form

$$A \cdot x = b \quad (*')$$

schreiben, wobei das eine Gleichheit von zwei Spalten mit jeweils m Zeilen bedeutet. Diese geschickte Schreibweise ist gewöhnungsbedürftig und auch etwas gefährlich, weil man leicht vergessen kann, was sie explizit bedeutet.

Man nennt A die *Koeffizientenmatrix* des linearen Gleichungssystems. Hängt man die Spalte b noch an, so erhält man die Matrix

$$(A, b) := \begin{pmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{pmatrix},$$

sie heißt *erweiterte Koeffizientenmatrix*. Darin ist alle Information über das Gleichungssystem enthalten.

Hat eine Matrix A insgesamt m Zeilen und n Spalten, so spricht man zur Abkürzung von einer $(m \times n)$ -Matrix. Man schreibt dafür $A = (a_{ij})$, die reellen Zahlen a_{ij} heißen *Einträge* von A .

Eine andere Methode, das Gleichungssystem $(*)$ kürzer aufzuschreiben, benutzt das *Summenzeichen* $\sum$. Allgemein ist

$$\sum_{j=1}^n c_j := c_1 + \dots + c_n.$$

Dabei heißt j der *Summationsindex*, man kann ihn durch jeden anderen Buchstaben ersetzen. In dieser Schreibweise lautet die i -te Gleichung

$$\sum_{j=1}^n a_{ij} x_j = b_i,$$

das ganze System also

$$\sum_{j=1}^n a_{ij} x_j = b_i \quad \text{für } i = 1, \dots, m. \quad (*'')$$

Welche Schreibweise man bevorzugt, ist nebensächlich. Wir werden vorwiegend $A \cdot x = b$ benutzen, weil dabei die geringste Anzahl von Buchstaben erforderlich ist.

0.4.2. Nachdem eine energiesparende Schreibweise für Gleichungssysteme vereinbart ist, können wir das Problem der Lösung in Angriff nehmen. Die *Lösungsmenge* ist nach Definition gleich

$$\text{Lös}(A, b) := \{x \in \mathbb{R}^n : A \cdot x = b\},$$

wobei x als Spaltenvektor geschrieben ist. (Diese suggestive Bezeichnungsweise habe ich bei [J] entdeckt; ich konnte nicht der Versuchung widerstehen, sie zu

übernehmen.) Das System zu *lösen* heißt, eine effiziente Methode zur Beschreibung der Menge $\text{Lös}(A, b)$ anzugeben. Was wir schließlich erhalten werden, ist eine Zahl $k \in \mathbb{N}$ und eine explizit angebbare bijektive Abbildung

$$\Phi: \mathbb{R}^k \rightarrow \text{Lös}(A, b) \subset \mathbb{R}^n,$$

sie heißt *Parametrisierung*. Die Berechnung von Φ mit Hilfe des nach C.F. GAUSS benannten *Eliminationsverfahrens* ist recht einfach, das ist Ziel dieses Kapitels. Der Nachweis der guten Eigenschaften von Φ erfordert etwas Theorie und wird in Kapitel 2 nachgeholt. Der abstrakte Hintergrund von linearen Gleichungssystemen wird schließlich in Kapitel 6 erläutert.

0.4.3. In den Beispielen aus 0.2 und 0.3 hatten wir Gleichungssysteme so lange umgeformt, bis eine Parametrisierung schrittweise „von unten nach oben“ berechnet werden konnte. Beispiele für so umgeformte Koeffizientenmatrizen A waren

$$\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{pmatrix}.$$

Die Nullen zu Beginn der Zeilen haben dabei eine typische Staffelung, die Trennlinie von den anderen Einträgen hat Stufenform.

Definition. Eine $m \times n$ -Matrix $A = (a_{ij})$ heißt in *Zeilenstufenform*, wenn sie von der folgenden Form ist:

$$A = \left(\begin{array}{ccccccc} & & \circledast & & & & \\ & & & \circledast & & & \\ & & & & \circledast & & \\ & & & & & \ddots & \\ & & & & & & \circledast \\ & 0 & & & & & \\ & & & & & & \end{array} \right) \left. \vphantom{\begin{pmatrix} \\ \\ \\ \\ \\ \end{pmatrix}} \right\} r.$$

Dabei müssen die Einträge an den mit $\circledast$ markierten Stellen ungleich Null sein, und unterhalb der eingezeichneten „Stufenlinie“ dürfen nur Nullen stehen.

Damit auch die Grenzfälle klar geregelt sind, kann man diese Definition noch präziser aufschreiben. A ist in Zeilenstufenform, wenn folgendes gilt:

1. Es gibt eine Zahl r mit $0 \leq r \leq m$, so daß in den Zeilen mit Index 1 bis r jeweils nicht nur Nullen stehen und in den Zeilen mit Index $r + 1$ bis m nur Nullen stehen.

2. Für jedes i mit $1 \leq i \leq r$ betrachten wir den niedrigsten Index j_i der Spalte, in der ein Eintrag ungleich Null steht, in Zeichen

$$j_i := \min\{j: a_{ij} \neq 0\}.$$

Offensichtlich ist $1 \leq j_i \leq n$, und die zusätzliche *Stufenbedingung* lautet

$$j_1 < j_2 < \dots < j_r.$$

Man beachte, daß der Fall $r = 0$ zugelassen ist; dann sind alle Einträge von A gleich Null. Die besonders ausgezeichneten und oben durch $\circledast$ gekennzeichneten Einträge

$$a_{1j_1}, \dots, a_{rj_r}$$

heißen *Pivots* (auf deutsch *Angelpunkte*) von A . Sie sind nach Definition von Null verschieden.

Beispiel. Für

$$A = \begin{pmatrix} 0 & 2 & 0 & 4 & 6 & 0 & 5 \\ 0 & 0 & 1 & 3 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

ist $m = 4$, $n = 7$, $r = 3$, $j_1 = 2$, $j_2 = 3$, $j_3 = 6$.

Besonders einfach aufzuschreiben ist der Spezialfall, in dem

$$j_1 = 1, \quad j_2 = 2, \quad \dots, \quad j_r = r.$$

Dann hat A die Form

$$A = \left(\begin{array}{ccccccc} \circledast & & & & & & \\ & \circledast & & & & & \\ & & \ddots & & & & \\ & & & 0 & & \circledast & \\ & & & & & & \end{array} \right) \Bigg\}^r.$$

Durch eine Umordnung der Spalten von A , d.h. eine andere Numerierung der Unbekannten des entsprechenden Gleichungssystems, kann man das stets erreichen. Für die Praxis ist das nebensächlich, aber für die Theorie kann man sich dadurch die lästigen Doppelindizes j_i ersparen.

0.4.4. Nun geben wir ein Lösungsverfahren für ein lineares Gleichungssystem an, bei dem die Koeffizientenmatrix A in Zeilenstufenform ist. Zur Vereinfachung nehmen wir an, daß die Pivots in den ersten r Spalten sitzen. Dann hat die

erweiterte Koeffizientenmatrix die Gestalt

$$(A, b) = \left(\begin{array}{cccc|c} a_{11} & & & & b_1 \\ & a_{22} & & & \vdots \\ & & \ddots & & \vdots \\ & & & a_{rr} & b_r \\ & & & & b_{r+1} \\ & 0 & & & \vdots \\ & & & & b_m \end{array} \right)$$

mit $a_{11} \neq 0, \dots, a_{rr} \neq 0$. Die Einträge $b_{r+1}, \dots, b_m$ sind entscheidend für die Frage, ob es überhaupt eine Lösung gibt.

Bemerkung. Gibt es ein $b_i \neq 0$ mit $r+1 \leq i \leq m$, so ist $\text{Lös}(A, b)$ leer.

Beweis. Die i -te Gleichung lautet

$$0 \cdot x_1 + \dots + 0 \cdot x_n = b_i \neq 0.$$

Diese Bedingung kann kein x erfüllen. □

Im gegenteiligen Fall

$$b_{r+1} = \dots = b_m = 0$$

geben wir nun eine Methode an, Lösungen zu konstruieren. Dazu unterscheiden wir zwischen zwei Arten von Variablen:

$x_{r+1}, \dots, x_n$ sind *freie Variablen*, sie können alle beliebigen Werte annehmen.

$x_1, \dots, x_r$ sind *gebundene Variablen*, sie sind eindeutig dadurch festgelegt, für welche Werte sich die freien Variablen entschieden haben.

Das kann man so beschreiben. Man setzt $k := n - r$, das ist die Zahl der freien Variablen, wählt $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ als *Parameter*, und setzt

$$x_{r+1} = \lambda_1, \quad x_{r+2} = \lambda_2, \dots, x_n = \lambda_k.$$

Zur Berechnung der $x_1, \dots, x_r$ daraus beginnt man mit der r -ten Gleichung

$$a_{rr}x_r + a_{r,r+1}\lambda_1 + \dots + a_{rn}\lambda_k = b_r.$$

Daraus erhält man

$$x_r = \frac{1}{a_{rr}}(b_r - a_{r,r+1}\lambda_1 - \dots - a_{rn}\lambda_k).$$

Setzt man das in die $(r-1)$ -te Gleichung ein, erhält man analog

$$x_{r-1} = d_{r-1,r-1}b_{r-1} + d_{r-1,r}b_r + c_{r-1,1}\lambda_1 + \dots + c_{r-1,k}\lambda_k,$$

wobei die auftretenden Zahlen c und d von den Einträgen der Zeilen $r-1$ und r aus der Matrix A abhängen. Führt man so fort, erhält man schließlich

$$x_1 = d_{11}b_1 + \dots + d_{1r}b_r + c_{11}\lambda_1 + \dots + c_{1k}\lambda_k.$$

Insgesamt ergibt sich eine Abbildung

$$\begin{aligned}\Phi: \mathbb{R}^k &\rightarrow \text{Lös}(A, b) \subset \mathbb{R}^n, \\ (\lambda_1, \dots, \lambda_k) &\mapsto (x_1, \dots, x_r, \lambda_1, \dots, \lambda_k),\end{aligned}$$

wobei für $x_1, \dots, x_r$ die oben berechneten von $\lambda_1, \dots, \lambda_k$, den Einträgen von A und $b_1, \dots, b_r$ abhängigen Ausdrücke einzusetzen sind. Da hierfür nach den obigen Rechnungen alle r Gleichungen erfüllt sind, liegen die Werte von Φ in der Lösungsmenge $\text{Lös}(A, b)$. Man hat also für beliebig gewählte Parameter $\lambda_1, \dots, \lambda_k$ eine Lösung des gegebenen Gleichungssystems erhalten. $\square$

Beispiel. Ist A die (4×7) -Matrix aus 0.4.3, und wählt man

$$b_1 = 3, \quad b_2 = 1, \quad b_3 = 2, \quad b_4 = 0,$$

so sind die Lösungen von $A \cdot x = b$ also in Spalten geschrieben gegeben durch

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ \frac{3}{2} - 2\lambda_2 - 3\lambda_3 - \frac{5}{2}\lambda_4 \\ \frac{1}{3} - 3\lambda_2 - 2\lambda_3 + \frac{1}{3}\lambda_4 \\ \lambda_2 \\ \lambda_3 \\ \frac{2}{3} - \frac{1}{3}\lambda_4 \\ \lambda_4 \end{pmatrix} \\ = \begin{pmatrix} 0 \\ \frac{3}{2} \\ \frac{1}{3} \\ 0 \\ 0 \\ \frac{2}{3} \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ -2 \\ -3 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0 \\ -3 \\ -2 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_4 \begin{pmatrix} 0 \\ -\frac{5}{2} \\ \frac{1}{3} \\ 0 \\ 0 \\ -\frac{1}{3} \\ 1 \end{pmatrix}.$$

0.4.5. Einen wichtigen Spezialfall wollen wir noch erwähnen: Ist die Matrix A quadratisch, so hat man ebensoviele Gleichungen wie Unbekannte. Ist speziell

A auf Zeilenstufenform mit $r = n$, so ist

$$A = \begin{pmatrix} \boxed{\circledast} & & & \\ & \boxed{\circledast} & & \\ & & \ddots & \\ 0 & & & \boxed{\circledast} \end{pmatrix},$$

und es gibt wegen $k = n - r = 0$ keinen freien Parameter, also eine einzige Lösung

$$x = (x_1, \dots, x_n),$$

die man wieder von unten nach oben berechnet. Ist überdies

$$b_1 = \dots = b_n = 0, \quad \text{so ist} \quad x_n = \dots = x_1 = 0,$$

man erhält also nur die *triviale Lösung*. Beispiele für eindeutig lösbare Gleichungssysteme findet man in 0.2.3 a), 0.3.5 und Aufgabe 2.

0.4.6. Nachdem wir gesehen haben, wie sich ein Gleichungssystem in Zeilenstufenform lösen läßt, versuchen wir nun, ein beliebiges System auf diese Form zu bringen. Dazu benutzen wir zwei Arten von *elementaren Zeilenumformungen* der erweiterten Koeffizientenmatrix:

- 1) Vertauschung von zwei Zeilen.
- 2) Addition der λ -fachen i -ten Zeile zur k -ten Zeile, wobei $0 \neq \lambda \in \mathbb{R}$ und $i \neq k$ ist.

Diese Umformungen sind gerechtfertigt durch den

Satz. Sei (A, b) die erweiterte Koeffizientenmatrix eines linearen Gleichungssystems, und $(\tilde{A}, \tilde{b})$ aus (A, b) durch endlich viele elementare Zeilenumformungen entstanden. Dann haben die Systeme $A \cdot x = b$ und $\tilde{A} \cdot x = \tilde{b}$ gleiche Lösungsräume, in Zeichen $\text{Lös}(A, b) = \text{Lös}(\tilde{A}, \tilde{b})$.

Vorsicht! Man beachte, daß Spaltenumformungen eine völlig andere Wirkung haben, weil dadurch die Unbekannten „gemischt“ werden. Das ist unerwünscht. Nur Vertauschungen in den ersten n Spalten sind ungefährlich, sie bewirken lediglich eine Umnummerierung der Unbekannten.

Beweis. Es genügt zu beweisen, daß der Lösungsraum bei einer einzigen elementaren Zeilenumformung unverändert bleibt, denn dann ändert auch Wiederholung nichts.

Typ 1) ist völlig unproblematisch, weil alle Gleichungen simultan erfüllt sein müssen, die Reihenfolge ist gleichgültig.

Bei Typ 2) muß man etwas rechnen. Da nur die Zeilen i und k betroffen sind, genügt es zu zeigen, daß die beiden aus jeweils zwei Gleichungen bestehenden Systeme

$$\begin{aligned} a_{i1}x_1 + \dots + a_{in}x_n &= b_i \\ a_{k1}x_1 + \dots + a_{kn}x_n &= b_k \end{aligned} \quad (*)$$

und

$$\begin{aligned} a_{i1}x_1 + \dots + a_{in}x_n &= b_i \\ (a_{k1} + \lambda a_{i1})x_1 + \dots + (a_{kn} + \lambda a_{in})x_n &= b_k + \lambda b_i \end{aligned} \quad (\tilde{*})$$

gleiche Lösungsräume haben. Erfüllt $x = (x_1, \dots, x_n)$ die Gleichungen $(*)$, so auch die erste von $(\tilde{*})$, und durch Addition der λ -fachen ersten Gleichung von $(*)$ zur zweiten die zweite Gleichung von $(\tilde{*})$. Umgekehrt folgt durch Subtraktion der λ -fachen ersten Gleichung aus $(\tilde{*})$ von der zweiten auch die zweite Gleichung aus $(*)$. $\square$

Was bei Umformungen vom Typ 2) geometrisch vorgeht, sieht man am einfachsten in der Ebene. Zwei Gleichungen beschreiben zwei Geraden, die Lösungsmenge besteht aus den Schnittpunkten (keiner, einer, oder eine ganze Gerade, vgl. 0.2). Was verschiedene Faktoren λ bewirken, wollen wir am besten an einem Beispiel zeigen: Gegeben seien Geraden

L_i durch $x_1 = 1$ und L_k durch $x_1 - x_2 = 2$.

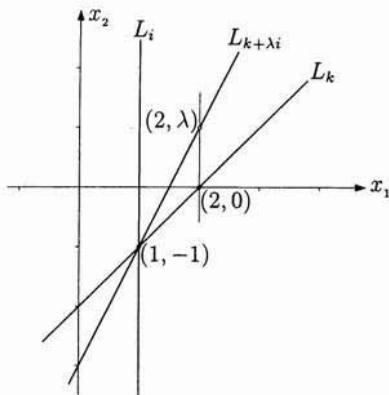


Bild 0.20

Dann ist

$$L_{k+\lambda i} \text{ gegeben durch } (1 + \lambda)x_1 - x_2 = 2 + \lambda.$$

Diese Schar von Geraden mit Parameter λ geht durch $(1, -1)$, sie enthält alle Geraden durch $(1, -1)$ mit Ausnahme von L_i , und die Zahl λ ist am Schnittpunkt mit der Geraden $x_1 = 2$ zu sehen.

0.4.7. Der letzte und am schwierigsten in allgemeiner Form aufzuschreibende Schritt ist enthalten in dem

Satz. Jede Matrix A kann man durch elementare Zeilenumformungen in eine Matrix $\tilde{A}$ in Zeilenstufenform überführen.

Beweis. Wir geben ein konkretes Verfahren an, das schrittweise durchgeführt wird und so aufgebaut ist, daß daraus ohne große Schwierigkeiten ein Computerprogramm gemacht werden kann. Wer durch die vielen Indizes verwirrt ist, möge zunächst das unten angegebene Beispiel studieren, bei dem drei Runden nötig sind.

Sei A eine $m \times n$ -Matrix. Ist $A = 0$, so hat A nach Definition schon Zeilenstufenform mit $r = 0$.

Ist $A \neq 0$, so gibt es mindestens einen Eintrag $\neq 0$. Also gibt es mindestens eine von Null verschiedene Spalte, wir wählen die mit dem kleinsten Index j_1 , in Zeichen

$$j_1 = \min\{j : \text{es gibt ein } i \text{ mit } a_{ij} \neq 0\}.$$

Ist $a_{1j_1} \neq 0$, so können wir es als Pivot wählen. Andernfalls suchen wir uns ein $a_{i_1j_1} \neq 0$ und vertauschen die Zeile 1 mit der Zeile i_1 . Das ist schon die erste Zeile von $\tilde{A}$, also gilt für den ersten Pivot

$$\tilde{a}_{1j_1} = a_{i_1j_1}.$$

Durch Umformungen vom Typ 2) kann man alle unterhalb von $\tilde{a}_{1j_1}$ stehenden Einträge zu Null machen. Ist a einer davon, so soll

$$a + \lambda \tilde{a}_{1j_1} = 0$$

werden, also hat man

$$\lambda = -\frac{a}{\tilde{a}_{1j_1}} \quad (*)$$

zu wählen. Das Ergebnis dieser Umformungen ist von der Gestalt

$$\tilde{A}_1 = \left(\begin{array}{cccc|cccc} 0 & \cdots & 0 & \tilde{a}_{1j_1} & * & \cdots & * & * \\ \vdots & & \vdots & 0 & \hline \vdots & & \vdots & \vdots & & & & \\ 0 & \cdots & 0 & 0 & & & & A_2 \end{array} \right),$$

wobei an den mit * markierten Stellen irgendwelche Einträge stehen. Die Matrix A_2 hat $m - 1$ Zeilen und $n - j_1$ Spalten.

Im zweiten Schritt macht man mit A_2 das Gleiche wie oben im ersten Schritt mit $A = A_1$: Ist $A_2 = 0$, so hat $\tilde{A}_1$ schon Zeilenstufenform; andernfalls suche man $j_2 > j_1$ und den Pivot $\tilde{a}_{2j_2}$. Die dabei nötigen Zeilenumformungen von A_2 kann man auf die Zeilen 2 bis m von $\tilde{A}_1$ ausdehnen, ohne daß sich in den Spalten 1 bis j_1 etwas ändert, denn dort stehen nur Nullen.

Ist A_2 umgeformt, so erhält man A_3 , u.s.w. Das Verfahren muß abbrechen, weil die Zeilen- und Spaltenzahlen der Matrizen A_k abnehmen, oder weil im Lauf des Verfahrens eine Matrix $A_k = 0$ entsteht. Das Endergebnis ist

$$\tilde{A} = \begin{pmatrix} & \tilde{a}_{1j_1} & * & \cdots & \cdots & \cdots & * \\ & & \tilde{a}_{2j_2} & & & & \vdots \\ & & & \ddots & & & \vdots \\ & 0 & & & \tilde{a}_{rj_r} & \cdots & * \end{pmatrix}.$$

□

Beispiel. Damit der Gang der Rechnung mit dem bloßen Auge zu erkennen ist, sind die Einträge so gewählt, daß sie ganzzahlig bleiben.

$$A = \begin{array}{ccccc|ccccc} 0 & 0 & 1 & 2 & 9 & 0 & 3 & 4 & 5 & 9 & \overline{0} & \underline{3} & \underline{4} & \underline{5} & \underline{9} \\ 0 & 3 & 4 & 5 & 9 & \rightsquigarrow & 0 & 0 & 1 & 2 & 9 & 0 & 0 & 1 & 2 & 9 \\ 0 & 6 & 7 & 8 & 9 & & 0 & 6 & 7 & 8 & 9 & 0 & 0 & -1 & -2 & -9 \\ 0 & 9 & 9 & 9 & 9 & & 0 & 9 & 9 & 9 & 9 & 0 & 0 & -3 & -6 & -18 \end{array}$$

$$\rightsquigarrow \begin{array}{ccccc|ccccc} \overline{0} & \underline{3} & \underline{4} & \underline{5} & \underline{9} & \overline{0} & \underline{3} & \underline{4} & \underline{5} & \underline{9} \\ 0 & 0 & \underline{1} & \underline{2} & \underline{9} & 0 & 0 & \underline{1} & \underline{2} & \underline{9} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 9 & 0 & 0 & 0 & 0 & 0 \end{array} = \tilde{A}$$

Bei dem oben allgemein beschriebenen Verfahren wird aus r verschiedenen Spalten jeweils ein Eintrag als Pivot ausgewählt, Kandidaten sind alle von Null verschiedenen Einträge. Für die Theorie wird sich später zeigen, daß das Ergebnis nicht von der Wahl abhängt. Für die Praxis ist es vorteilhaft, den vom Betrag her größten Eintrag zu wählen, weil entsprechend (*) durch den Pivot dividiert wird, und kleine Nenner zu großen Schwankungen führen können (vgl. Aufgabe 4).

0.4.8. Nun ist das *Eliminationsverfahren* von GAUSS für ein System von m linearen Gleichungen und n Unbestimmten mit reellen Koeffizienten komplett, wir fassen die einzelnen Schritte noch einmal zusammen:

- 1) Man schreibe die erweiterte Koeffizientenmatrix (A, b) auf.
- 2) Man bringe A auf Zeilenstufenform und forme dabei die Spalte b mit um. Ergebnis ist $(\tilde{A}, \tilde{b})$, insbesondere die Zahl r . Beachte, daß in der b -Spalte kein Pivot gesucht wird!
- 3) Man lese an $\tilde{b}$ ab, ob es Lösungen gibt und wenn ja, berechne man die Parametrisierung

$$\Phi: \mathbb{R}^{n-r} \rightarrow \text{Lös}(\tilde{A}, \tilde{b}) = \text{Lös}(A, b) \subset \mathbb{R}^n$$

der Lösungsmenge.

Nun zu der entscheidenden Frage nach den Eigenschaften der so erhaltenen Abbildung Φ , der *Parametrisierung* der Lösungen. Zu verschiedenen k -Tupeln von Parametern gehören auch verschiedene Lösungen, da die Parameter direkt in die letzten k Komponenten der Lösung eingetragen werden (vgl. 0.4.4). Also ist Φ injektiv. Bleibt die Frage, ob Φ surjektiv ist, d.h. ob alle Lösungen von der Parametrisierung erreicht werden.

Zur Beantwortung der letzten Frage muß man etwas arbeiten. Insbesondere ist zu zeigen, daß die mit Hilfe der Umformungen erhaltene Zahl r nur von der Matrix A und nicht vom Verfahren (etwa der Auswahl der Pivots) abhängt. Man kann das alles relativ direkt angehen, aber viel schöner geht es mit Hilfe von etwas Theorie, die ja ohnehin schmackhaft gemacht werden soll. Die Zahl r wird sich als *Rang* der Matrix A erweisen, und die Lösungsmenge $\text{Lös}(A, b)$ hat die Struktur eines *affinen Raumes* der *Dimension* $k = n - r$. In dieser Situation folgt die Surjektivität von Φ aus der Injektivität, die Begründungen werden in 2.3 nachgeholt. Das soll aber nicht daran hindern, schon vorher Gleichungssysteme zu lösen, wenn sie auftreten.

Aufgaben zu 0.4

1. Lösen Sie folgende lineare Gleichungssysteme:

a)

$$\begin{array}{rrrrrcl} & x_2 & +2x_3 & +3x_4 & = & 0 \\ x_1 & +2x_2 & +3x_3 & +4x_4 & = & 0 \\ 2x_1 & +3x_2 & +4x_3 & +5x_4 & = & 0 \\ 3x_1 & +4x_2 & +5x_3 & +6x_4 & = & 0 \end{array}$$

b)

$$\begin{array}{rrrrr} -6x_1 & +6x_2 & +2x_3 & -2x_4 & = & 2 \\ -9x_1 & +8x_2 & +3x_3 & -2x_4 & = & 3 \\ -3x_1 & +2x_2 & +x_3 & & = & 1 \\ -15x_1 & +14x_2 & +5x_3 & -4x_4 & = & 5 \end{array}$$

2. Geben Sie die Lösung des linearen Gleichungssystems an, das durch die folgende erweiterte Koeffizientenmatrix gegeben ist:

$$\left(\begin{array}{cccc|c} 1 & -1 & 2 & -3 & 7 \\ 4 & 0 & 3 & 1 & 9 \\ 2 & -5 & 1 & 0 & -2 \\ 3 & -1 & -1 & 2 & -2 \end{array} \right).$$

3. Bestimmen Sie, für welche $t \in \mathbb{R}$ das folgende lineare Gleichungssystem in Matrixdarstellung lösbar ist und geben Sie gegebenenfalls die Lösung an.

$$\left(\begin{array}{ccc|c} 2 & 4 & 2 & 12t \\ 2 & 12 & 7 & 12t + 7 \\ 1 & 10 & 6 & 7t + 8 \end{array} \right)$$

4. Lösen Sie das folgende lineare Gleichungssystem auf einem Taschenrechner mit einer Rechengenauigkeit von n Stellen hinter dem Komma (Abschneiden weiterer Stellen ohne Rundung!) für $\varepsilon = 10^{-k}$ für größer werdendes $k \leq n$, und zwar einmal mit dem Pivot ε und einmal mit dem „maximalen Zeilenpivot“ 1 der ersten Spalte.

$$\begin{array}{rcl} x + y & = & 2, \\ \varepsilon x + y & = & 1. \end{array}$$

Beschreiben Sie den geometrischen Hintergrund dieser Umformungen.

Kapitel 1

Grundbegriffe

Zu Beginn dieses Kapitels erklären wir die grundlegenden Begriffe der Algebra, die an den verschiedensten Stellen der Mathematik auftauchen. Wie intensiv man dies bei der ersten Lektüre studieren soll, ist eine Frage des Geschmacks. Wer gar keinen Appetit darauf verspürt, kann sofort bis nach 1.4 weiterblättern, wo die wirkliche lineare Algebra, nämlich die Theorie der Vektorräume, beginnt, und das, was er von den Grundbegriffen später benötigt, bei Bedarf nachschlagen.

1.1 Mengen und Abbildungen

Wir wollen hier nicht auf die schwierige Frage eingehen, wie der Begriff „Menge“ erklärt werden kann; das ist Gegenstand der mathematischen Grundlagenforschung. Für die Zwecke der linearen Algebra genügt es, einige elementare Regeln und Bezeichnungen zu erläutern. Wer an einer fundierten Darstellung interessiert ist, möge zum Beispiel [F-P] konsultieren.

1.1.1. Die *endlichen Mengen* kann man im Prinzip durch eine vollständige Liste ihrer Elemente angeben. Man schreibt dafür

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

wobei der Doppelpunkt neben dem Gleichheitszeichen bedeutet, daß das links stehende Symbol X durch den rechts stehenden Ausdruck *definiert* wird. Die x_i heißen *Elemente* von X , in Zeichen $x_i \in X$. Man beachte, daß die Elemente x_i nicht notwendig verschieden sein müssen, und daß die Reihenfolge gleichgültig ist. Man nennt die Elemente $x_1, \dots, x_n \in X$ *paarweise verschieden*, wenn

$x_i \neq x_j$ für $i \neq j$. In diesem Fall ist n die *Anzahl der Elemente* von X .

Die *leere Menge* $\emptyset$ ist dadurch gekennzeichnet, daß sie kein Element enthält.

Eine Menge X' heißt *Teilmenge* von X , in Zeichen $X' \subset X$, wenn aus $x \in X'$ immer $x \in X$ folgt. Es gilt

$$X = Y \Leftrightarrow X \subset Y \text{ und } Y \subset X.$$

Die einfachste *unendliche Menge* ist die Menge

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

der *natürlichen Zahlen*. Man kann sie charakterisieren durch die PEANO-Axiome (vgl. [Z]). Diese enthalten das *Prinzip der vollständigen Induktion*:

Sei $M \subset \mathbb{N}$ eine Teilmenge mit folgenden Eigenschaften:

a) $0 \in M$,

b) $n \in M \Rightarrow n + 1 \in M$.

Dann ist $M = \mathbb{N}$.

Mancher Leser wird sich an dieser Stelle zum ersten Mal – aber sicher insgesamt nicht zum letzten Mal – darüber wundern, daß die Bezeichnungen in der Mathematik nicht einheitlich sind. So wird die Null manchmal nicht als natürliche Zahl angesehen. Es gibt Versuche, hier durch DIN-Normen Ordnung zu schaffen (vgl. [DIN]), aber viele Mathematiker lieben mehr ihre Freiheit, als Normblätter.

Durch *Erweiterungen von Zahlbereichen* erhält man ausgehend von $\mathbb{N}$ die *ganzen Zahlen*

$$\mathbb{Z} = \{0, +1, -1, +2, -2, \dots\},$$

die *rationalen Zahlen*

$$\mathbb{Q} = \left\{ \frac{p}{q} : p, q \in \mathbb{Z}, q \neq 0 \right\},$$

und etwa als Dezimalbrüche oder Cauchy-Folgen rationaler Zahlen die *reellen Zahlen* $\mathbb{R}$. Es ist

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C},$$

wobei die in gewisser Weise abschließende Erweiterung zu den komplexen Zahlen $\mathbb{C}$ in 1.3.3 und 1.3.9 behandelt wird. Einem Leser, der mehr über den Aufbau der „Zahlen“ wissen möchte, sei [Z] empfohlen.

1.1.2. In der linearen Algebra werden wir mit solchen Mengen auskommen, die sich aus den in 1.1.1 angegebenen Mengen mit Hilfe einiger elementarer Operationen ableiten lassen.

Aus einer gegebenen Menge X kann man *Teilmengen* auswählen, die durch gewisse Eigenschaften der Elemente charakterisiert sind, in Zeichen

$$X' := \{x \in X : x \text{ hat die Eigenschaft } E\},$$

zum Beispiel $X' := \{n \in \mathbb{N} : n \text{ ist gerade}\}$.

Sind $X_1, \dots, X_n$ Mengen, so hat man eine *Vereinigung*

$$X_1 \cup \dots \cup X_n := \{x : \text{es gibt ein } i \in \{1, \dots, n\} \text{ mit } x \in X_i\}$$

und den *Durchschnitt*

$$X_1 \cap \dots \cap X_n := \{x : x \in X_i \text{ für alle } i \in \{1, \dots, n\}\}.$$

An einigen Stellen wird es nötig sein, Vereinigungen und Durchschnitte von mehr als endlich vielen Mengen zu betrachten. Dazu verwendet man eine Menge I , die *Indexmenge* heißt, so daß für jedes $i \in I$ eine Menge X_i gegeben ist. Dann sind *Vereinigung* und *Durchschnitt* erklärt durch

$$\bigcup_{i \in I} X_i := \{x: \text{ es gibt ein } i \in I \text{ mit } x \in X_i\},$$

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

Ist etwa $I = \mathbb{N}$ und $X_i := [-i, i] \subset \mathbb{R}$ für jedes $i \in \mathbb{N}$ ein Intervall, so ist

$$\bigcup_{i \in \mathbb{N}} X_i = \mathbb{R} \quad \text{und} \quad \bigcap_{i \in \mathbb{N}} X_i = \{0\}.$$

Ist $X' \subset X$ eine Teilmenge, so nennt man

$$X \setminus X' := \{x \in X: x \notin X'\}$$

die *Differenzmenge* (oder das *Komplement*).

1.1.3. Zur Beschreibung von Beziehungen zwischen verschiedenen Mengen verwendet man „Abbildungen“. Sind X und Y Mengen, so versteht man unter einer *Abbildung* von X nach Y eine Vorschrift f , die jedem $x \in X$ eindeutig ein $f(x) \in Y$ zuordnet. Man schreibt dafür

$$f: X \rightarrow Y, \quad x \mapsto f(x).$$

Man beachte den Unterschied zwischen den beiden Pfeilen: „ $\rightarrow$ “ steht zwischen den Mengen und „ $\mapsto$ “ zwischen den Elementen.

Zwei Abbildungen $f: X \rightarrow Y$ und $g: X \rightarrow Y$ heißen *gleich*, in Zeichen $f = g$, wenn $f(x) = g(x)$ für alle $x \in X$. Mit

$$\text{Abb}(X, Y)$$

bezeichnen wir die Menge aller Abbildungen von X nach Y .

Ein Problem bei dieser Definition einer Abbildung ist, daß nicht präzisiert ist, in welcher Form die Abbildungsvorschrift gegeben sein soll (genauso wenig war festgelegt worden, in welcher Form die charakterisierende Eigenschaft einer Teilmenge vorliegen soll). Besonders einfach zu beschreiben sind Abbildungen, bei denen $f(x)$ durch eine explizite Formel angegeben werden kann, etwa im Fall $X = Y = \mathbb{R}$ durch

$$f(x) = ax, \quad f(x) = x^2, \quad f(x) = \sqrt{x}.$$

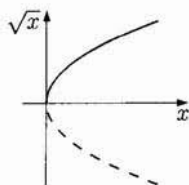
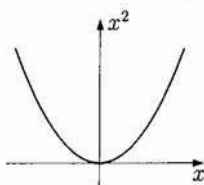
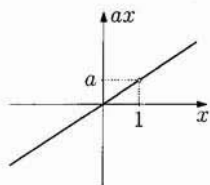


Bild 1.1

Die letzte Vorschrift ist schon problematisch, weil es für positive reelle Zahlen zwei und für negative keine Quadratwurzel gibt. Eine Abbildung im Sinn der Definition liegt nur vor, wenn man negative x ausschließt und für positive x eine Wurzel (etwa die positive) auswählt. Dann hat man eine Abbildung

$$f: \mathbb{R}_+ \rightarrow \mathbb{R}_+, \quad x \mapsto +\sqrt{x},$$

wobei $\mathbb{R}_+ := \{x \in \mathbb{R}: x \geq 0\}$.

Mit einer Abbildung kann man nicht nur Elemente, sondern auch Teilmengen bewegen. Sei also $f: X \rightarrow Y$ und $M \subset X$, $N \subset Y$. Dann heißt

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

das *Bild* von M (in Y), insbesondere $f(X)$ das Bild von X in Y , und

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

das *Urbild* von N in X . Man beachte, daß für eine einelementige Menge $N = \{y\}$ das Urbild

$$f^{-1}(y) := f^{-1}(\{y\}) \subset X$$

eine Teilmenge ist, die aus mehreren Elementen bestehen, aber auch leer sein kann (etwa bei $f(x) = x^2$). Daher ist f^{-1} im allgemeinen keine Abbildung von Y nach X im Sinn der Definition.

Noch eine Bezeichnungsweise: Ist $f: X \rightarrow Y$ eine Abbildung und $M \subset X$ eine Teilmenge, so nennt man

$$f|_M: M \rightarrow Y, \quad x \mapsto f(x),$$

die *Beschränkung* von f auf M . Sie unterscheidet sich von f nur durch den eingeschränkten Definitionsbereich. Ist $Y \subset Y'$ eine Teilmenge, so ist es üblich, mit $f: X \rightarrow Y'$ auch die Abbildung mit dem ausgedehnten Bildbereich zu bezeichnen.

1.1.4. Besonders wichtige Eigenschaften von Abbildungen haben eigene Namen. Eine Abbildung $f: X \rightarrow Y$ heißt

injektiv, falls aus $x, x' \in X$ und $f(x) = f(x')$ stets $x = x'$ folgt,

surjektiv, falls $f(X) = Y$,

d.h. falls es zu jedem $y \in Y$ ein $x \in X$ gibt mit $y = f(x)$,

bijektiv, falls f injektiv und surjektiv ist.

Ist f bijektiv, so gibt es zu jedem $y \in Y$ genau ein $x \in X$ mit $f(x) = y$. In diesem Fall kann man also eine *Umkehrabbildung*

$$f^{-1}: Y \rightarrow X, \quad y \mapsto x = f^{-1}(y) \quad \text{mit} \quad y = f(x)$$

erklären. Man beachte, daß das Symbol f^{-1} in verschiedenen Bedeutungen vorkommt:

- bei einer beliebigen Abbildung ist für jede Teilmenge $N \subset Y$ das Urbild $f^{-1}(N) \subset X$ eine Teilmenge,
- ist f bijektiv, so besteht für die einelementige Teilmenge $\{y\} \subset Y$ das Urbild $f^{-1}(\{y\})$ aus einem Element x , in Zeichen

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

dafür schreibt man dann $f^{-1}(y) = x$.

Durch systematisches Zählen beweist man den folgenden einfachen

Satz. Ist X eine endliche Menge, so sind für eine Abbildung $f: X \rightarrow X$ folgende Bedingungen äquivalent:

- i) f ist injektiv,
- ii) f ist surjektiv,
- iii) f ist bijektiv.

Beweis. X bestehe aus n Elementen, also $X = \{x_1, \dots, x_n\}$, wobei $n \in \mathbb{N}$ ist und die x_i paarweise verschieden sind.

i) $\Rightarrow$ ii): Wir zeigen „nicht surjektiv“ $\Rightarrow$ „nicht injektiv“. Ist $f(X) \neq X$, so besteht $f(X)$ aus $m < n$ Elementen. Nun besagt das sogenannte *Schubladenprinzip von DIRICHLET*: Verteilt man n Objekte irgendwie in m Schubladen, wobei $m < n$, so gibt es mindestens eine Schublade, in der mehr als ein Objekt liegt. Also kann f nicht injektiv sein.

ii) $\Rightarrow$ i): Ist f nicht injektiv, so gibt es x_i, x_j mit $x_i \neq x_j$, aber $f(x_i) = f(x_j)$. Dann kann $f(X)$ höchstens $n - 1$ Elemente enthalten, also ist f nicht surjektiv.

Wegen ii) $\Rightarrow$ i) folgt auch ii) $\Rightarrow$ iii), iii) $\Rightarrow$ i) ist klar. Damit sind alle möglichen Implikationen bewiesen. $\square$

1.1.5. Sind X, Y, Z Mengen und $f: X \rightarrow Y$ sowie $g: Y \rightarrow Z$ Abbildungen, so heißt die Abbildung

$$g \circ f: X \rightarrow Z, \quad x \mapsto g(f(x)) =: (g \circ f)(x)$$

die *Komposition* (oder *Hintereinanderschaltung*) von f und g (man sagt g *kompontiert mit f* für $g \circ f$). Man beachte dabei, daß die zuerst angewandte Abbildung rechts steht, im Gegensatz zum Diagramm

$$X \xrightarrow{f} Y \xrightarrow{g} Z.$$

Daher könnte man das Diagramm besser umgekehrt schreiben:

$$Z \xleftarrow{g} Y \xleftarrow{f} X.$$

Bemerkung. Die Komposition von Abbildungen ist assoziativ, d.h. für Abbildungen $f: X \rightarrow Y$, $g: Y \rightarrow Z$ und $h: Z \rightarrow W$ ist

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

Beweis. Ist $x \in X$, so ist

$$\begin{aligned} ((h \circ g) \circ f)(x) &= (h \circ g)(f(x)) = h(g(f(x))) \\ &= h((g \circ f)(x)) = (h \circ (g \circ f))(x). \end{aligned} \quad \square$$

Vorsicht! Die Komposition von Abbildungen ist i.a. *nicht* kommutativ. Ist

$$\begin{aligned} f: \mathbb{R} &\rightarrow \mathbb{R}, & x &\mapsto x + 1, \\ g: \mathbb{R} &\rightarrow \mathbb{R}, & x &\mapsto x^2, \end{aligned}$$

so ist $(f \circ g)(x) = x^2 + 1$ und $(g \circ f)(x) = (x + 1)^2$, also $f \circ g \neq g \circ f$.

Um eine etwas andersartige Charakterisierung von Injektivität und Surjektivität zu erhalten, erklären wir für jede Menge X die identische Abbildung

$$\text{id}_X: X \rightarrow X, \quad x \mapsto x.$$

Lemma. Sei $f: X \rightarrow Y$ eine Abbildung zwischen den nicht leeren Mengen X und Y . Dann gilt:

- 1) f ist genau dann injektiv, wenn es eine Abbildung $g: Y \rightarrow X$ gibt, so daß $g \circ f = \text{id}_X$.
- 2) f ist genau dann surjektiv, wenn es eine Abbildung $g: Y \rightarrow X$ gibt, so daß $f \circ g = \text{id}_Y$.
- 3) f ist genau dann bijektiv, wenn es eine Abbildung $g: Y \rightarrow X$ gibt, so daß $g \circ f = \text{id}_X$ und $f \circ g = \text{id}_Y$. In diesem Fall ist $g = f^{-1}$.

Beweis. 1) Sei f injektiv. Dann gibt es zu jedem $y \in f(X)$ genau ein $x \in X$ mit $f(x) = y$, und wir definieren $g(y) := x$. Ist $x_0 \in X$ beliebig, so definieren wir weiter $g(y) = x_0$ für alle $y \in Y \setminus f(X)$. Das ergibt eine Abbildung $g: Y \rightarrow X$ mit $g \circ f = \text{id}_X$.

Ist umgekehrt $g: Y \rightarrow X$ mit $g \circ f = \text{id}_X$ gegeben, und ist $f(x) = f(x')$ für $x, x' \in X$, so ist $x = \text{id}_X(x) = g(f(x)) = g(f(x')) = \text{id}_X(x') = x'$. Also ist f injektiv.

2) Sei f surjektiv. Zu jedem $y \in Y$ wählen wir ein festes $x \in X$ mit $f(x) = y$ und setzen $g(y) := x$. Die so erklärte Abbildung $g: Y \rightarrow X$ hat die Eigenschaft $f \circ g = \text{id}_Y$.

Ist umgekehrt $g: Y \rightarrow X$ mit $f \circ g = \text{id}_Y$ gegeben, und ist $y \in Y$, so ist $y = f(g(y))$, also Bild von $g(y)$, und f ist surjektiv.

3) Ist f bijektiv, so erfüllt $g := f^{-1}$ die beiden Gleichungen. Ist umgekehrt $g: Y \rightarrow X$ mit $g \circ f = \text{id}_X$ und $f \circ g = \text{id}_Y$ gegeben, so ist f nach 1) und 2) bijektiv, und es gilt $g = f^{-1}$. $\square$

1.1.6. Schon bei der Einführung des Raumes $\mathbb{R}^n$ hatten wir ein „direktes Produkt“ betrachtet. Sind allgemeiner $X_1, \dots, X_n$ Mengen, so betrachten wir die *geordneten n -Tupel*

$$x = (x_1, \dots, x_n) \quad \text{mit} \quad x_i \in X_1, \dots, x_n \in X_n.$$

Genauer kann man x als Abbildung

$$x: \{1, \dots, n\} \rightarrow X_1 \cup \dots \cup X_n \quad \text{mit} \quad x(i) \in X_i$$

ansetzen (man nennt x auch *Auswahlfunktion*), und zur Vereinfachung der Schreibweise $x_i := x(i)$ und $x := (x_1, \dots, x_n)$ setzen. Im Sinne der Gleichheit von Abbildungen gilt dann

$$(x_1, \dots, x_n) = (x'_1, \dots, x'_n) \Leftrightarrow x_1 = x'_1, \dots, x_n = x'_n.$$

Nun erklären wir das *direkte Produkt*

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

als Menge der geordneten n -Tupel. Offensichtlich ist $X_1 \times \dots \times X_n \neq \emptyset$, wenn $X_i \neq \emptyset$ für alle $i \in \{1, \dots, n\}$.

Für jedes i hat man eine *Projektion* auf die i -te *Komponente*

$$\pi_i: X_1 \times \dots \times X_n \rightarrow X_i, \quad (x_1, \dots, x_i, \dots, x_n) \mapsto x_i.$$

Ist speziell $X_1 = \dots = X_n = X$, so schreibt man

$$X^n = X \times \dots \times X.$$

Ein Element von X^n ist also eine Abbildung $\{1, \dots, n\} \rightarrow X$. Ist allgemeiner I irgendeine Menge (man nennt sie *Indexmenge*), so nennt man eine Abbildung

$$\varphi: I \rightarrow X, \quad i \mapsto x_i = \varphi(i),$$

eine *Familie* von Elementen in X . Man beachte den Unterschied zu einer Teilmenge $X' \subset X$, die man mit $\varphi(I)$ vergleichen kann: die Elemente $i \in I$ kann man als Etiketten (oder noch näher am Familienleben als Pflichten) ansehen, die unter den Elementen von X verteilt werden. Jedes Etikett i hat einen eindeutigen Empfänger x_i , die Elemente von $X' = \varphi(I)$ erhalten mindestens ein Etikett,

und es ist möglich, daß manche $x \in X'$ mehrere Etiketten erhalten (was zur oft gehörten Klage „immer ich“ führt).

Zur Abkürzung bezeichnet man eine Familie $I \rightarrow X$ oft mit $(x_i)_{i \in I}$.

Für die Indexmenge $I = \mathbb{N}$ der natürlichen Zahlen nennt man $(x_i)_{i \in \mathbb{N}}$ eine *Folge*, das ist ein grundlegender Begriff der Analysis.

Vorsicht! Die Frage der Existenz der oben betrachteten Auswahlfunktionen ist nicht unproblematisch. Das führt zum *Auswahlaxiom*, vgl. [F-P], das auch im Beweis von Lemma 1.1.5 schon stillschweigend verwendet wurde.

1.1.7. Für eine Abbildung $f: X \rightarrow Y$ nennt man die Menge

$$\Gamma_f := \{(x, f(x)) \in X \times Y\}$$

den *Graphen* von f . Damit kann man oft eine Abbildung durch eine Skizze veranschaulichen.

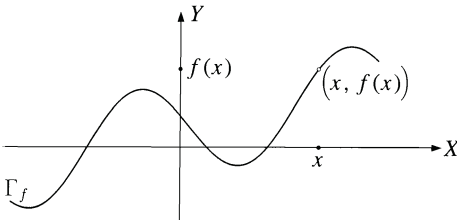


Bild 1.2

Nach der Definition einer Abbildung ist die Einschränkung der Projektion auf X

$$\pi: \Gamma_f \rightarrow X$$

bijektiv. Daraus folgt, daß das „Funktionengebirge“ keine „Überhänge“ hat, wie im folgenden Bild:



Bild 1.3 Gebirge mit Überhängen: In der Fränkischen Schweiz

1.1.8. Das direkte Produkt ist auch nützlich, um Beziehungen (oder Relationen) zwischen je zwei Elementen $x, y \in X$ zu studieren. Man schreibt dafür allgemein $x \sim y$.

Beispiele. a) $X =$ Menge der Menschen, $x \sim y :\Leftrightarrow x$ kennt y .

b) $X = \mathbb{R}$, $x \sim y :\Leftrightarrow x \leq y$.

c) $X = \mathbb{R}^n$, $(x_1, \dots, x_n) \sim (y_1, \dots, y_n) :\Leftrightarrow x_1^2 + \dots + x_n^2 = y_1^2 + \dots + y_n^2$.

d) $X = \mathbb{Z}$, $0 \neq m \in \mathbb{N}$, $x \sim y :\Leftrightarrow y - x$ durch m teilbar.

Eine Relation ist beschrieben durch ihren *Graphen* $R \subset X \times X$, wobei

$$(x, y) \in R \Leftrightarrow x \sim y. \quad (*)$$

Man kann also eine *Relation* auf X definieren als Teilmenge $R \subset X \times X$, und das Zeichen $\sim$ durch $(*)$ erklären.

Definition. Eine Relation $\sim$ auf X heißt *Äquivalenzrelation*, wenn für beliebige $x, y, z \in X$ gilt:

A1 $x \sim x$, (reflexiv)

A2 $x \sim y \Rightarrow y \sim x$, (symmetrisch)

A3 $x \sim y$ und $y \sim z \Rightarrow x \sim z$. (transitiv)

In diesem Fall sagt man x ist *äquivalent* zu y für $x \sim y$.

Der Leser möge zur Übung die obigen Beispiele auf diese Eigenschaften überprüfen. Vor allem die Reflexivität in Beispiel a) ist etwas Nachdenken wert.

Hat man auf einer Menge eine Äquivalenzrelation eingeführt, so kann man – wie wir sehen werden – zu einer neuen Menge übergehen, in der äquivalente Elemente der ursprünglichen Menge zu „Repräsentanten“ desselben neuen Elementes werden. Dabei wird – mit den Worten von HERMANN WEYL – *alles im Sinne des eingenommenen Standpunktes Unwesentliche an den untersuchten Objekten abgestreift*. Übersetzt ins Mengen-Latein, liest sich das wie folgt:

Ist eine Äquivalenzrelation auf einer Menge X gegeben, so heißt eine Teilmenge $A \subset X$ *Äquivalenzklasse* (bezüglich R), falls gilt:

1. $A \neq \emptyset$.

2. $x, y \in A \Rightarrow x \sim y$.

3. $x \in A, y \in X, x \sim y \Rightarrow y \in A$.

Man überlege sich, wie die Äquivalenzklassen in obigen Beispielen c) und d) aussehen.

Bemerkung. Ist R eine Äquivalenzrelation auf einer Menge X , so gehört jedes Element $a \in X$ zu genau einer Äquivalenzklasse. Insbesondere gilt für zwei beliebige Äquivalenzklassen A, A' entweder $A = A'$ oder $A \cap A' = \emptyset$.

Beweis. Für ein fest gegebenes $a \in X$ definieren wir

$$A := \{x \in X : x \sim a\}.$$

Wir zeigen, daß A eine Äquivalenzklasse ist, die a enthält. Wegen $a \sim a$ ist $a \in A$, und es folgt $A \neq \emptyset$. Sind $x, y \in A$, so ist $x \sim a$ und $y \sim a$, also $x \sim y$ nach A2 und A3.

Ist $x \in A$, $y \in X$ und $x \sim y$, so ist $x \sim a$, also nach A2 und A3 auch $y \sim a$ und daher $y \in A$. Damit ist gezeigt, daß a in mindestens einer Äquivalenzklasse enthalten ist.

Es bleibt zu zeigen, daß zwei Äquivalenzklassen A und A' entweder gleich oder disjunkt sind. Angenommen, es ist $A \cap A' \neq \emptyset$ und $a \in A \cap A'$. Ist $x \in A$, so ist $x \sim a$, und wegen $a \in A'$ folgt auch $x \in A'$. Also ist $A \subset A'$. Ebenso beweist man $A' \subset A$, woraus $A = A'$ folgt. $\square$

Jede Äquivalenzrelation R auf einer Menge X liefert also eine *Zerlegung* von X in disjunkte Äquivalenzklassen. Diese Äquivalenzklassen betrachtet man als Elemente einer neuen Menge, die man mit X/R bezeichnet. Man nennt sie die *Quotientenmenge* von X nach der Äquivalenzrelation R . Die *Elemente* von X/R sind also spezielle *Teilmengen* von X . Indem man jedem Element $a \in X$ die Äquivalenzklasse A_a zuordnet, in der es enthalten ist, erhält man eine *kanonische* (d.h. in der gegebenen Situation eindeutig festgelegte) Abbildung

$$X \rightarrow X/R, \quad a \mapsto A_a.$$

Das Urbild eines Elementes $A \in X/R$ ist dabei wieder A , aber aufgefaßt als *Teilmenge* von X .

Jedes $a \in A$ heißt ein *Repräsentant* der Äquivalenzklasse A . Im allgemeinen gibt es keine Möglichkeit, spezielle Repräsentanten besonders auszuzeichnen. Das wäre auch gar nicht im Sinne der durchgeführten Konstruktion.

Als Beispiel aus dem realen Leben kann eine Schule dienen: die Menge der Schüler wird in Klassen eingeteilt, und für manche Fragen, etwa die Gestaltung des Stundenplans, ist nur noch die Menge der Klassen von Bedeutung.

Aufgaben zu 1.1

1. Beweisen Sie die folgenden Rechenregeln für die Operationen mit Mengen:

- $X \cap Y = Y \cap X$, $X \cup Y = Y \cup X$,
- $X \cap (Y \cap Z) = (X \cap Y) \cap Z$, $X \cup (Y \cup Z) = (X \cup Y) \cup Z$,

- c) $X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z)$, $X \cup (Y \cap Z) = (X \cup Y) \cap (X \cup Z)$,
 d) $X \setminus (M_1 \cap M_2) = (X \setminus M_1) \cup (X \setminus M_2)$, $X \setminus (M_1 \cup M_2) = (X \setminus M_1) \cap (X \setminus M_2)$.

2. Sei $f: X \rightarrow Y$ eine Abbildung. Zeigen Sie:

- a) Ist $M_1 \subset M_2 \subset X$, so folgt $f(M_1) \subset f(M_2)$.
 Ist $N_1 \subset N_2 \subset Y$, so folgt $f^{-1}(N_1) \subset f^{-1}(N_2)$.
 b) $M \subset f^{-1}(f(M))$ für $M \subset X$, $f(f^{-1}(N)) \subset N$ für $N \subset Y$.
 c) $f^{-1}(Y \setminus N) = X \setminus f^{-1}(N)$ für $N \subset Y$.
 d) Für $M_1, M_2 \subset X$ und $N_1, N_2 \subset Y$ gilt:

$$f^{-1}(N_1 \cap N_2) = f^{-1}(N_1) \cap f^{-1}(N_2), \quad f^{-1}(N_1 \cup N_2) = f^{-1}(N_1) \cup f^{-1}(N_2),$$

$$f(M_1 \cup M_2) = f(M_1) \cup f(M_2), \quad f(M_1 \cap M_2) \subset f(M_1) \cap f(M_2).$$

Finden Sie ein Beispiel, in dem $f(M_1 \cap M_2) \neq f(M_1) \cap f(M_2)$ gilt!

3. Seien $f: X \rightarrow Y$, $g: Y \rightarrow Z$ Abbildungen und $g \circ f: X \rightarrow Z$ die Komposition von f und g . Dann gilt:

- a) Sind f und g injektiv (surjektiv), so ist auch $g \circ f$ injektiv (surjektiv).
 b) Ist $g \circ f$ injektiv (surjektiv), so ist auch f injektiv (g surjektiv).

4. Untersuchen Sie die folgenden Abbildungen auf Injektivität und Surjektivität:

- a) $f_1: \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x, y) \mapsto x + y$, b) $f_2: \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x, y) \mapsto x^2 + y^2 - 1$,
 c) $f_3: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $(x, y) \mapsto (x + 2y, 2x - y)$,

5. Zwei Mengen X und Y heißen gleichmächtig genau dann, wenn es eine bijektive Abbildung $f: X \rightarrow Y$ gibt. Eine Menge X heißt abzählbar unendlich, falls X und $\mathbb{N}$ gleichmächtig sind.

- a) Zeigen Sie, dass $\mathbb{Z}$ und $\mathbb{Q}$ abzählbar unendlich sind.
 b) Zeigen Sie, dass $\mathbb{R}$ nicht abzählbar unendlich ist.
 c) Für eine nichtleere Menge M sei $\text{Abb}(M, \{0, 1\})$ die Menge aller Abbildungen von M nach $\{0, 1\}$. Zeigen Sie, dass M und $\text{Abb}(M, \{0, 1\})$ nicht gleichmächtig sind.

6. Ein Konferenzhotel für Mathematiker hat genau $\mathbb{N}$ Betten. Das Hotel ist bereits voll belegt, aber die Mathematiker lassen sich nach Belieben innerhalb des Hotels umquartieren. Das Hotel soll aus wirtschaftlichen Gründen stets voll belegt sein, und wenn möglich, sollen alle neu ankommenden Gäste untergebracht werden. Was macht man in folgenden Fällen?

- a) Ein weiterer Mathematiker trifft ein.
 b) Die Insassen eines Kleinbusses mit n Plätzen suchen Unterkunft.
 c) Ein Großraumbus mit $\mathbb{N}$ Personen kommt an.
 d) n Großraumbusse treffen ein.
 e) $\mathbb{N}$ Großraumbusse fahren vor.

1.2 Gruppen

1.2.1. Unter einer *Verknüpfung* (oder *Komposition*) auf einer Menge G versteht man eine Vorschrift $*$, die zwei gegebenen Elementen $a, b \in G$ ein neues Element $*(a, b) \in G$ zuordnet, d.h. eine Abbildung

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

Wir geben einige Beispiele für solche Vorschriften $*$ und schreiben dabei zur Abkürzung $a * b$ statt $*(a, b)$:

a) Ist X eine Menge und $G = \text{Abb}(X, X)$ die Menge aller Abbildungen

$$f: X \rightarrow X,$$

so ist für $f, g \in G$ nach 1.1.5 auch $f \circ g \in G$, also kann man

$$f * g := f \circ g$$

setzen.

b) In $G = \mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ oder $\mathbb{R}_+^*$ hat man Addition und Multiplikation, also kann man für $a, b \in G$

$$a * b := a + b \quad \text{oder} \quad a * b := a \cdot b$$

setzen. Im Gegensatz zu Beispiel a) ist die Reihenfolge von a und b hier egal.

c) In $G = \mathbb{Q}$ oder $\mathbb{R}$ kann man

$$a * b := \frac{1}{2}(a + b)$$

als das *arithmetische Mittel* von a und b erklären.

1.2.2. Verknüpfungen mit besonders guten Eigenschaften haben eigene Namen.

Definition. Eine Menge G zusammen mit einer Verknüpfung $*$ heißt *Gruppe*, wenn folgende Axiome erfüllt sind:

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

G2 Es gibt ein $e \in G$ (*neutrales Element* genannt) mit den folgenden Eigenschaften:

a) $e * a = a$ für alle $a \in G$.

b) Zu jedem $a \in G$ gibt es ein $a' \in G$ (*inverses Element von a* genannt) mit $a' * a = e$.

Die Gruppe heißt *abelsch* (oder *kommutativ*), falls außerdem $a * b = b * a$ für alle $a, b \in G$.

Falls das keine Verwirrung stiftet, schreibt man die Verknüpfung zur Vereinfachung meist als Multiplikation, also $a \cdot b$ oder nur ab statt $a * b$.

Ist die Verknüpfung als Addition geschrieben, so setzt man stillschweigend voraus, daß sie kommutativ ist. Das neutrale Element 0 heißt dann *Nullelement*, das Inverse von a wird mit $-a$ bezeichnet und heißt *Negatives*.

Wenn das Assoziativgesetz erfüllt ist, kann man bei mehrfachen Produkten die Klammern weglassen, also schreiben:

$$abc = (ab)c = a(bc).$$

Zunächst wollen wir die Gruppenaxiome bei den Beispielen aus 1.2.1 nachprüfen.

a) In $G = \text{Abb}(X, X)$ ist die Komposition nach 1.1.5 assoziativ, die identische Abbildung id_X erfüllt G2a, aber aus der Existenz eines g mit $g \circ f = \text{id}_X$ folgt, daß f injektiv sein muß. Also ist G2b im allgemeinen nicht erfüllt.

Das kann man aber reparieren. Die Hintereinanderschaltung ist auch eine Verknüpfung in der Teilmenge

$$S(X) := \{f \in \text{Abb}(X, X) : f \text{ bijektiv}\},$$

und $S(X)$ wird damit zu einer Gruppe. Sie heißt die *symmetrische Gruppe* der Menge X . Ist $X = \{1, \dots, n\}$, so schreibt man S_n statt $S(X)$. Jedes $\sigma \in S_n$ heißt *Permutation* der Zahlen $1, \dots, n$, und S_n nennt man auch *Permutationsgruppe*. In der linearen Algebra ist sie wichtig bei der Berechnung von Determinanten, daher verschieben wir alles Weitere hierüber auf Kapitel 3. Insbesondere werden wir dort sehen, daß S_n für $n \geq 3$ nicht abelsch ist.

b) Hier sind nur $\mathbb{Z}$, $\mathbb{Q}$ und $\mathbb{R}$ mit der Addition und $\mathbb{R}_+^*$ mit der Multiplikation Gruppen. Der Leser möge zur Übung nachprüfen, welche Axiome in den anderen Fällen verletzt sind.

c) Das arithmetische Mittel ist nur kommutativ, aber nicht assoziativ, und es gibt kein neutrales Element.

1.2.3. Bei der Aufstellung von Axiomen versucht man, so wenig wie möglich zu fordern und die weiteren Eigenschaften daraus abzuleiten. Insbesondere haben wir weder bei e noch bei a' die Eindeutigkeit postuliert. Derartigen Kleinkram kann man nachträglich beweisen.

Bemerkung. Ist G eine Gruppe, so gilt:

a) Das neutrale Element $e \in G$ ist eindeutig bestimmt und hat auch die Eigenschaft $a \cdot e = a$ für alle $a \in G$.

b) Das inverse Element a' ist für jedes $a \in G$ eindeutig bestimmt und hat auch die Eigenschaft $a \cdot a' = e$ für alle $a \in G$.

c) Da das Inverse zu a nach b) eindeutig bestimmt ist, kann man es mit a^{-1} bezeichnen. Es gilt für $a, b \in G$:

$$a^{-1} \cdot a = a \cdot a^{-1} = e, \quad (a^{-1})^{-1} = a, \quad (ab)^{-1} = b^{-1}a^{-1}.$$

d) Es gelten die folgenden Kürzungsregeln:

$$a \cdot \tilde{x} = a \cdot x \Rightarrow x = \tilde{x} \quad \text{und} \quad y \cdot a = \tilde{y} \cdot a \Rightarrow y = \tilde{y}.$$

Beweis. Wir betrachten ein neutrales e und ein $a \in G$. Zu a' gibt es ein a'' mit $a''a' = e$. Daraus folgt

$$\begin{aligned} aa' &= e(aa') = (a''a')(aa') = a''(a'(aa')) \\ &= a''((a'a)a') = a''(ea') = a''a' = e, \end{aligned}$$

und somit $ae = a(a'a) = (aa')a = ea = a$.

Ist $\tilde{e}$ ein eventuelles anderes neutrales Element, so ist

$$e\tilde{e} = e \quad \text{und} \quad e\tilde{e} = \tilde{e}, \quad \text{also} \quad e = \tilde{e}.$$

Damit ist a) und die erste Gleichung von c) bewiesen.

Ist $\tilde{a}'$ ein eventuelles anderes Inverses, so folgt

$$\tilde{a}' = \tilde{a}'e = \tilde{a}'(aa') = (\tilde{a}'a)a' = ea' = a'$$

unter Verwendung der bereits vorher bewiesenen Eigenschaften. Damit ist auch b) bewiesen.

Aus $aa^{-1} = e$ folgt, daß a inverses Element zu a^{-1} ist, d.h. $(a^{-1})^{-1} = a$. Weiter gilt

$$(b^{-1}a^{-1})(ab) = b^{-1}(a^{-1}(ab)) = b^{-1}((a^{-1}a)b) = b^{-1}(eb) = b^{-1}b = e.$$

Schließlich folgen die Kürzungsregeln durch Multiplikation der jeweils ersten Gleichung von links bzw. rechts mit a^{-1} . $\square$

1.2.4. Auf der Suche nach Beispielen für Gruppen kann es helfen, das Axiom G2 etwas anders zu formulieren. Dazu betrachten wir für ein festes $a \in G$ die Abbildungen

$$\begin{aligned} \tau_a: G &\rightarrow G, \quad x \mapsto x \cdot a, \quad (\text{Rechtstranslation}), \text{ und} \\ {}_a\tau: G &\rightarrow G, \quad x \mapsto a \cdot x, \quad (\text{Linkstranslation}). \end{aligned}$$

Lemma. Ist G eine Gruppe, so sind für jedes $a \in G$ die Abbildungen τ_a und ${}_a\tau$ bijektiv.

Ist umgekehrt G eine Menge mit einer assoziativen Verknüpfung, so folgt G2 aus der Surjektivität der Abbildungen τ_a und ${}_a\tau$ für alle $a \in G$.

Beweis. Die Bijektivität von τ_a und ${}_a\tau$ bedeutet, daß es zu $b \in G$ genau ein $x \in G$ und genau ein $y \in G$ gibt mit

$$x \cdot a = b \quad \text{und} \quad a \cdot y = b,$$

d.h. daß diese Gleichungen mit x und y als Unbestimmten eindeutig lösbar sind. Die Existenz einer Lösung ist klar, denn es genügt,

$$x = b \cdot a^{-1} \quad \text{und} \quad y = a^{-1} \cdot b$$

zu setzen. Sind $\tilde{x}$ und $\tilde{y}$ weitere Lösungen, so gilt

$$x \cdot a = \tilde{x} \cdot a \quad \text{und} \quad a \cdot y = a \cdot \tilde{y},$$

also $x = \tilde{x}$ und $y = \tilde{y}$ nach der Kürzungsregel in 1.2.3.

Seien umgekehrt die Gleichungen

$$x \cdot a = b \quad \text{und} \quad a \cdot y = b$$

für beliebige $a, b \in G$ lösbar. Dann gibt es zu a ein e mit $e \cdot a = a$. Ist $b \in G$ beliebig, so ist

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

also existiert ein neutrales Element. Durch Lösen der Gleichung $x \cdot a = e$ erhält man das inverse Element von a . $\square$

1.2.5. Eine Verknüpfung auf einer endlichen Menge $G = \{a_1, \dots, a_n\}$ kann man im Prinzip dadurch angeben, daß man die Werte aller Produkte $a_i \cdot a_j$ in einem quadratischen Schema ähnlich einer Matrix aufschreibt. Dabei steht $a_i \cdot a_j$ in der i -ten Zeile und der j -ten Spalte der *Verknüpfungstafel* (oder im Fall einer Gruppe der *Gruppentafel*):

*	...	a_j	...
$\vdots$			
a_i		$a_i * a_j$	
$\vdots$			

Ob das Gruppenaxiom G2 erfüllt ist, kann man dann nach obigem Lemma daran erkennen, ob jede Zeile und jede Spalte der Tafel eine Permutation von $a_1, \dots, a_n$ ist.

Daraus folgt sofort, daß es nur eine Möglichkeit gibt, die 2-elementige Menge $G_2 = \{a_1, a_2\}$ zu einer Gruppe zu machen: Ein Element, etwa $a_1 = e$, muß

neutral sein, das andere ist $a_2 = a$, und die Gruppentafel ist

$\cdot$	e	a
e	e	a
a	a	e

Die Kommutativität erkennt man an der Symmetrie der Tafel. Ob das Assoziativgesetz erfüllt ist, kann man der Tafel nicht direkt ansehen, das muß man (leider) für alle möglichen n^3 Tripel nachprüfen.

Für $n = 3$ und $G = \{e, a, b\}$ erhält man ebenfalls nur eine mögliche Gruppentafel, nämlich

$\cdot$	e	a	b
e	e	a	b
a	a	b	e
b	b	e	a

und man stellt fest, daß diese Multiplikation assoziativ ist. Für $n = 4$ und $G = \{e, a, b, c\}$ findet man leicht zwei wesentlich verschiedene Möglichkeiten, nämlich

$\cdot$	e	a	b	c
e	e	a	b	c
a	a	b	c	e
b	b	c	e	a
c	c	e	a	b

und

$\odot$	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

Wieder ist die Kommutativität offensichtlich, die Assoziativität etwas mühsam zu überprüfen. Um diese beiden verschiedenen Multiplikationen unterscheiden zu können, schreiben wir $G_4^{\odot}$ für G_4 mit der Multiplikation $\odot$.

Es ist klar, daß dieses nur auf Ausprobieren beruhende Verfahren für größere n zu kompliziert und unbefriedigend ist.

1.2.6. Um eine bessere Methode zum Studium von Gruppen zu erhalten, benötigt man geeignete Begriffe, die Beziehungen der Gruppen untereinander regeln.

Definition. Sei G eine Gruppe mit Verknüpfung $\cdot$ und $G' \subset G$ eine nichtleere Teilmenge. G' heißt *Untergruppe*, wenn für $a, b \in G'$ auch $a \cdot b \in G'$ und $a^{-1} \in G'$.

Sind G und H Gruppen mit Verknüpfungen $\cdot$ und $\odot$, so heißt eine Abbildung $\varphi: G \rightarrow H$ *Homomorphismus (von Gruppen)*, wenn

$$\varphi(a \cdot b) = \varphi(a) \odot \varphi(b) \quad \text{für alle } a, b \in G.$$

Ein Homomorphismus heißt *Isomorphismus*, wenn er bijektiv ist.

Zunächst einige unmittelbare Folgerungen aus den Definitionen.

Bemerkung 1. Ist G eine Gruppe und $G' \subset G$ Untergruppe, so ist G' mit der Verknüpfung aus G wieder eine Gruppe.

Man nennt diese Verknüpfung in G' die von G induzierte Verknüpfung.

Beweis. Die induzierte Verknüpfung ist assoziativ, weil sie aus G kommt. Da es ein $a \in G'$ gibt, ist $a^{-1} \in G'$ und $a \cdot a^{-1} = e \in G'$. $\square$

Bemerkung 2. Sei $\varphi: G \rightarrow H$ ein Homomorphismus von Gruppen. Dann gilt:

- a) $\varphi(e) = \mathbf{e}$, wenn $e \in G$ und $\mathbf{e} \in H$ die neutralen Elemente bezeichnen.
- b) $\varphi(a^{-1}) = \varphi(a)^{-1}$ für alle $a \in G$.
- c) Ist φ Isomorphismus, so ist auch die Umkehrabbildung $\varphi^{-1}: H \rightarrow G$ ein Homomorphismus.

Beweis. a) folgt aus der Kürzungsregel in H , da

$$\mathbf{e} \odot \varphi(e) = \varphi(e) = \varphi(e \cdot e) = \varphi(e) \odot \varphi(e),$$

und daraus ergibt sich b), weil

$$\mathbf{e} = \varphi(e) = \varphi(a^{-1} \cdot a) = \varphi(a^{-1}) \odot \varphi(a).$$

Zum Beweis von c) nehmen wir $c, d \in H$. Ist $c = \varphi(a)$ und $d = \varphi(b)$, so ist

$$\varphi(a \cdot b) = \varphi(a) \odot \varphi(b) = c \odot d, \quad \text{also} \quad \varphi^{-1}(c \odot d) = a \cdot b = \varphi^{-1}(c) \cdot \varphi^{-1}(d).$$

$\square$

Im folgenden werden wir die Verknüpfung und die neutralen Elemente in G und H nur noch dann verschieden bezeichnen, wenn das erforderlich ist (etwa in dem Beispiel b) weiter unten).

Beispiele. a) Zunächst betrachten wir die in 1.2.5 konstruierten Beispiele. Für die Mengen gilt

$$G_2 \subset G_3, \quad G_3 \subset G_4 \quad \text{und} \quad G_2 \subset G_4,$$

aber bei keiner der Teilmengen handelt es sich um eine Untergruppe. Dagegen ist

$$G_2 \subset G_4^\circ$$

eine Untergruppe.

Die identische Abbildung $G_4 \rightarrow G_4^\circ$ ist kein Homomorphismus. Beispiele von Homomorphismen $\varphi: G_4 \rightarrow G_4^\circ$ sind gegeben durch

$$\varphi(e) = \varphi(a) = \varphi(b) = \varphi(c) = e,$$

$$\varphi(e) = e, \quad \varphi(a) = a, \quad \varphi(b) = e, \quad \varphi(c) = a,$$

$$\varphi(e) = e, \quad \varphi(a) = b, \quad \varphi(b) = e, \quad \varphi(c) = b,$$

$$\varphi(e) = e, \quad \varphi(a) = c, \quad \varphi(b) = e, \quad \varphi(c) = c.$$

Die einfachen Begründungen seien dem Leser zur Übung empfohlen.

b) Ist $G = \mathbb{R}$ mit der Addition und $H = \mathbb{R}_+^*$ mit der Multiplikation als Verknüpfung, so ist die Exponentialfunktion

$$\exp: \mathbb{R} \rightarrow \mathbb{R}_+^*, \quad x \mapsto e^x,$$

ein Isomorphismus, da $e^{x+y} = e^x \cdot e^y$.

c) Wir betrachten die abelsche Gruppe $\mathbb{Z}$ mit der Addition als Verknüpfung. Für jedes $m \in \mathbb{Z}$ ist die Abbildung

$$\varphi_m: \mathbb{Z} \rightarrow \mathbb{Z}, \quad a \mapsto m \cdot a,$$

ein Homomorphismus, denn $m(a+b) = ma + mb$. Sein Bild

$$m\mathbb{Z} := \{m \cdot a : a \in \mathbb{Z}\} \subset \mathbb{Z}$$

ist eine Untergruppe, weil $ma + mb = m(a+b)$ und $-mb = m(-b)$.

1.2.7. Im Anschluß an das letzte Beispiel kann man nun eine sogenannte *zyklische Gruppe* mit m Elementen konstruieren. Wir nehmen $m > 0$ an und teilen die ganzen Zahlen in m Klassen (d.h. disjunkte Teilmengen) folgendermaßen ein: Zu jedem

$$r \in \{0, 1, \dots, m-1\}$$

betrachten wir die Menge

$$r + m\mathbb{Z} := \{r + m \cdot a : a \in \mathbb{Z}\},$$

die man sich als die um r verschobene Untergruppe $m\mathbb{Z}$ vorstellen kann (vgl. 1.1.8). Für $m = 2$ ist

$0 + 2\mathbb{Z}$ die Menge der geraden und

$1 + 2\mathbb{Z}$ die Menge der ungeraden Zahlen.

Im allgemeinen Fall ist

$$\mathbb{Z} = (0 + m\mathbb{Z}) \cup (1 + m\mathbb{Z}) \cup \dots \cup (m-1 + m\mathbb{Z}),$$

und diese Vereinigung ist disjunkt. Zu welcher Klasse eine Zahl $a \in \mathbb{Z}$ gehört, kann man durch Division mit Rest entscheiden. Ist

$$\frac{a}{m} = k + \frac{r}{m} \quad \text{mit } k \in \mathbb{Z} \text{ und } r \in \{0, \dots, m-1\},$$

so ist $a \in r + m\mathbb{Z}$, denn $a = r + mk$. Daher nennt man die Klassen $r + m\mathbb{Z}$ auch *Restklassen modulo m* : sie bestehen aus all den Zahlen, die bei Division durch m den gleichen Rest hinterlassen. Zwei Zahlen a, a' liegen genau dann in derselben Restklasse, wenn $a - a'$ durch m teilbar ist. Man schreibt dafür auch

$$a \equiv a' \pmod{m}$$

und sagt „ a ist kongruent a' modulo m “. Für $m = 0$ ist die Kongruenz die Gleichheit.

Zu jedem $a \in \mathbb{Z}$ bezeichnen wir mit $\bar{a} = a + m\mathbb{Z}$ seine Restklasse. Die Zahl a heißt *Repräsentant* von $\bar{a}$. In der Menge der Restklassen kann man nun eine Addition definieren durch

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

Diese Definition ist nur sinnvoll, wenn sie nicht von der Auswahl der Repräsentanten abhängt. Man sagt dazu, die Addition ist *wohldefiniert*. Das ist aber einfach zu sehen:

Ist $\bar{a} = \bar{a}'$ und $\bar{b} = \bar{b}'$, so ist $a - a' = mk$ und $b - b' = ml$ mit $k, l \in \mathbb{Z}$, also

$$a + b = a' + b' + m(k + l), \quad \text{d.h.} \quad \overline{a + b} = \overline{a' + b'}.$$

Nun können wir das Ergebnis formulieren:

Satz. Ist $m \in \mathbb{Z}$ und $m > 0$, so ist die Menge

$$\mathbb{Z}/m\mathbb{Z} := \{\bar{0}, \bar{1}, \dots, \overline{m-1}\}$$

der Restklassen modulo m mit der oben erklärten Addition eine abelsche Gruppe, und die Abbildung

$$\mathbb{Z} \rightarrow \mathbb{Z}/m\mathbb{Z}, \quad a \mapsto a + m\mathbb{Z},$$

ist ein surjektiver Homomorphismus.

Beweis. Die Assoziativität vererbt sich von $\mathbb{Z}$ nach $\mathbb{Z}/m\mathbb{Z}$:

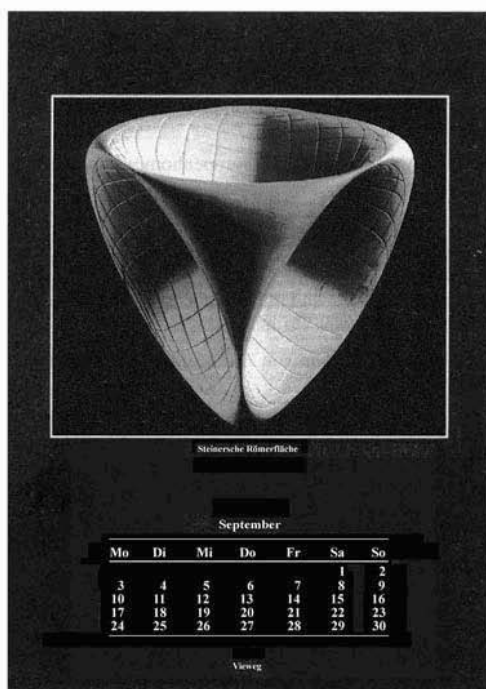
$$\begin{aligned} (\bar{a} + \bar{b}) + \bar{c} &= \overline{a + b} + \bar{c} = \overline{(a + b) + c} = \overline{a + (b + c)} \\ &= \overline{a + b + c} = \bar{a} + (\bar{b} + \bar{c}). \end{aligned}$$

Neutrales Element ist $\bar{0}$, denn $\bar{0} + \bar{a} = \overline{0 + a} = \bar{a}$, und Inverses von $\bar{a}$ ist $\overline{-a}$. Auch die Kommutativität wird von $\mathbb{Z}$ vererbt. $\square$

Man nennt $\mathbb{Z}/m\mathbb{Z}$ für $m > 0$ die *zyklische Gruppe der Ordnung m* , für $m = 0$ ist $\mathbb{Z}/0\mathbb{Z} = \mathbb{Z}$, diese Gruppe heißt *unendlich zyklisch*.

Das etwas ungewohnt erscheinende Rechnen mit Restklassen ist im täglichen Leben höchst vertraut: Wer um 10 Uhr weggeht und 3 Stunden unterwegs ist, kommt um 1 Uhr zurück. Denn der Stundenzeiger der Uhr rechnet modulo 12. Die Zeitrechnung insgesamt mit ihren durch Sekunden, Minuten, Stunden, Tage, Monate und Jahre verursachten Kongruenzen ist weit komplizierter als dieser letzte Abschnitt über Gruppen.

Teile der Restklassen modulo 7 findet man auf jedem Blatt eines Monatskalenders. Betrachtet man die Wochentage als Restklassen, so ergibt deren Addition in diesem Monat z.B. Mittwoch + Samstag = Donnerstag. Die ebenfalls abgebildete *Römerfläche* wird in der projektiven Geometrie wieder auftauchen [Fi].



Ein Leser, der mit Kongruenzen immer noch Schwierigkeiten hat, ist in guter Gesellschaft mit GOETHE'S Faust, der zu Mephisto sagt: „Mich dünkt, die Alte spricht im Fieber“, als die Hexe aus ihrem Rechenbuche vorliest:

*Du must verstehn!
Aus Eins mach Zehn,
Und Zwei laß gehn,
Und Drei mach gleich,
So bist du reich.
Verlier die Vier!*

*Aus Fünf und Sechs,
So sagt die Hex',
Mach Sieben und Acht,
So ist's vollbracht:
Und Neun ist Eins,
Und Zehn ist keins.
Das ist das Hexen-Einmaleins.*

Alles klar: die Hexe rechnet schließlich modulo 2. Nur am Anfang holperst noch etwas, da scheint der Reim vor der Rechnung zu stehen.

Aufgaben zu 1.2

1. Sei G eine Gruppe mit $aa = e$ für alle $a \in G$, wobei e das neutrale Element von G bezeichnet. Zeigen Sie, dass G abelsch ist.

2. Bestimmen Sie (bis auf Isomorphie) alle Gruppen mit höchstens vier Elementen. Welche davon sind abelsch?

3. Welche der folgenden Abbildungen sind Gruppenhomomorphismen?

a) $f_1: \mathbb{Z} \rightarrow \mathbb{Z}, z \mapsto 2z$, b) $f_2: \mathbb{Z} \rightarrow \mathbb{Z}, z \mapsto z + 1$,

c) $f_3: \mathbb{Z} \rightarrow \mathbb{Q}^*, z \mapsto z^2 + 1$, d) $f_4: \mathbb{C}^* \rightarrow \mathbb{R}^*, z \mapsto |z|$,

e) $f_5: \mathbb{C} \rightarrow \mathbb{R}^*, z \mapsto |z|$, f) $f_6: \mathbb{Z}/p\mathbb{Z} \rightarrow \mathbb{Z}/p\mathbb{Z}, z \mapsto z^p$.

Dabei ist die Verknüpfung in $\mathbb{Z}$, $\mathbb{C}$ und $\mathbb{Z}/p\mathbb{Z}$ jeweils die Addition, in $\mathbb{Q}^*$, $\mathbb{R}^*$ und $\mathbb{C}^*$ jeweils die Multiplikation und p eine Primzahl.

4. Sei G eine Gruppe und $A \subset G$. Die von A erzeugte Untergruppe $\text{erz}(A)$ ist definiert durch

$$\text{erz}(A) = \{a_1 \cdot \dots \cdot a_n : n \in \mathbb{N}, a_i \in A \text{ oder } a_i^{-1} \in A\}.$$

$\text{erz}(A)$ ist somit die Menge aller endlichen Produkte von Elementen aus A bzw. deren Inversen. Zeigen Sie, dass $\text{erz}(A)$ die „kleinste“ Untergruppe von G ist, die A enthält, d.h.

i) $\text{erz}(A) \subset G$ ist eine Untergruppe.

ii) Ist $U \subset G$ eine Untergruppe mit $A \subset U$, so folgt $\text{erz}(A) \subset U$.

Wie sieht $\text{erz}(A)$ aus für den Fall, dass A einelementig ist?

5. Für eine natürliche Zahl $n \geq 3$ sei $d \in S(\mathbb{R}^2)$ die Drehung um den Winkel $2\pi/n$ und $s \in S(\mathbb{R}^2)$ die Spiegelung an der x -Achse. Die *Diedergruppe* D_n ist definiert durch

$$D_n := \text{erz}(\{s, d\}).$$

a) Wie viele Elemente hat D_n ?

b) Geben Sie eine Gruppentafel von D_3 an.

6. Eine Gruppe G heißt *zyklisch*, falls es ein $g \in G$ gibt mit $G = \text{erz}(\{g\})$.

a) Wie sieht die Gruppentafel einer endlichen zyklischen Gruppe aus?

b)* Zeigen Sie, dass jede zyklische Gruppe entweder isomorph zu $\mathbb{Z}$ oder $\mathbb{Z}/n\mathbb{Z}$ ($n \in \mathbb{N}$ geeignet) ist.

7. Zeigen Sie: Ist G eine abelsche Gruppe und $H \subset G$ eine Untergruppe, so ist durch

$$x \sim y \Leftrightarrow xy^{-1} \in H$$

eine Äquivalenzrelation auf G erklärt. Sei $G/H := G/\sim$ die Menge der Äquivalenzklassen, und die zu $x \in G$ gehörige Äquivalenzklasse sei mit $\bar{x}$ bezeichnet. Sind $x, x', y, y' \in G$ mit $x \sim x'$ und $y \sim y'$, so ist $xy \sim x'y'$. Somit kann man auf G/H durch

$$\bar{x} \cdot \bar{y} := \overline{xy}$$

eine Verknüpfung erklären.

Zeigen Sie, dass G/H auf diese Weise zu einer abelschen Gruppe wird und für $G = \mathbb{Z}$, $H = n\mathbb{Z}$ genau die in 1.2.7 definierten zyklischen Gruppen $\mathbb{Z}/n\mathbb{Z}$ entstehen.

8. Man gebe ein Beispiel einer nicht assoziativen Verknüpfung auf der Menge $G = \{1, 2, 3\}$, so dass für alle $a \in G$ die Translationen τ_a und ${}_a\tau$ aus 1.2.4 surjektiv sind.

1.3 Ringe, Körper und Polynome

1.3.1. Bei Gruppen hat man eine einzige Verknüpfung; ob man sie als Addition oder als Multiplikation bezeichnet, ist nebensächlich. In der linearen Algebra braucht man aber Addition und Multiplikation, also zwei Arten der Verknüpfung, die miteinander in geregelter Beziehung stehen. Damit beschäftigt sich dieser Abschnitt.

Definition. Eine Menge R zusammen mit zwei Verknüpfungen

$$\begin{aligned} +: R \times R &\rightarrow R, & (a, b) &\mapsto a + b, & \text{ und} \\ \cdot: R \times R &\rightarrow R, & (a, b) &\mapsto a \cdot b, \end{aligned}$$

heißt *Ring*, wenn folgendes gilt:

R1 R zusammen mit der *Addition* $+$ ist eine abelsche Gruppe.

R2 Die *Multiplikation* $\cdot$ ist assoziativ.

R3 Es gelten die *Distributivgesetze*, d.h. für alle $a, b, c \in R$ gilt

$$a \cdot (b + c) = a \cdot b + a \cdot c \quad \text{und} \quad (a + b) \cdot c = a \cdot c + b \cdot c.$$

Ein Ring heißt *kommutativ*, wenn $a \cdot b = b \cdot a$ für alle $a, b \in R$. Ein Element $1 \in R$ heißt *Einselement*, wenn $1 \cdot a = a \cdot 1 = a$ für alle $a \in R$.

Wie üblich soll dabei zur Einsparung von Klammern die Multiplikation stärker binden als die Addition.

Bemerkung. Ist R ein Ring und $0 \in R$ das Nullelement, so gilt für alle $a \in R$

$$0 \cdot a = a \cdot 0 = 0.$$

Beweis. $0 \cdot a = (0 + 0) \cdot a = 0 \cdot a + 0 \cdot a.$ □

Beispiele. a) Die Mengen $\mathbb{Z}$ der ganzen Zahlen, $\mathbb{Q}$ der rationalen Zahlen und $\mathbb{R}$ der reellen Zahlen sind zusammen mit der üblichen Addition und Multiplikation kommutative Ringe.

b) Ist $I \subset \mathbb{R}$ ein Intervall und

$$R := \{f: I \rightarrow \mathbb{R}\}$$

die Menge der reellwertigen Funktionen, so sind durch

$$(f + g)(x) := f(x) + g(x) \quad \text{und} \quad (f \cdot g)(x) := f(x) \cdot g(x)$$

Verknüpfungen erklärt, und R wird damit zu einem kommutativen Ring. Das folgt ganz leicht aus den Ringeigenschaften von $\mathbb{R}$.

c) In der Gruppe $\mathbb{Z}/m\mathbb{Z}$ der Restklassen modulo m aus 1.2.7 kann man durch

$$\overline{a} \cdot \overline{b} := \overline{a \cdot b}$$

auch eine Multiplikation erklären. Denn ist wieder $a - a' = mk$ und $b - b' = ml$, so folgt

$$a \cdot b = (a' + mk) \cdot (b' + ml) = a' \cdot b' + m(b'k + a'l + mkl).$$

Also ist die Definition der Multiplikation unabhängig von der Auswahl der Repräsentanten.

Die Regeln R2 und R3 und die Kommutativität der Multiplikation folgen ganz einfach aus den entsprechenden Regeln in $\mathbb{Z}$.

Die Multiplikationstafeln wollen wir für $m = 2, 3, 4$ explizit aufschreiben. Dabei lassen wir zur Vereinfachung bei den Restklassen die Querstriche weg.

·	0	1
0	0	0
1	0	1

·	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

·	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

Die Multiplikation mit 0 ist uninteressant, wir betrachten also in diesen drei Fällen die Menge $R \setminus \{0\}$. Für $m = 2$ und $m = 3$ ist sie zusammen mit der Multiplikation wieder eine Gruppe, für $m = 4$ nicht, denn hier ist

$$1 \cdot 2 = 3 \cdot 2 \quad \text{und} \quad 2 \cdot 2 = 0.$$

Also ist die Kürzungsregel verletzt, und das Produkt $2 \cdot 2$ liegt nicht in $R \setminus \{0\}$.

1.3.2. Das vorangehende Beispiel motiviert die

Definition. Ein Ring R heißt *nullteilerfrei*, wenn für alle $a, b \in R$ aus $a \cdot b = 0$ stets $a = 0$ oder $b = 0$ folgt.

Bemerkung. Der Restklassenring $\mathbb{Z}/m\mathbb{Z}$ ist genau dann nullteilerfrei, wenn m eine Primzahl ist.

Beweis. Ist m keine Primzahl, also $m = k \cdot l$ mit $1 < k, l < m$, so ist

$$\bar{k}, \bar{l} \neq \bar{0}, \quad \text{aber} \quad \bar{0} = \bar{m} = \bar{k} \cdot \bar{l}.$$

Ist umgekehrt m Primzahl und $\bar{k} \cdot \bar{l} = \bar{0}$, so ist

$$k \cdot l = r \cdot m$$

für ein $r \in \mathbb{Z}$. Also hat entweder k oder l einen Primfaktor m , d.h. $\bar{k} = \bar{0}$ oder $\bar{l} = \bar{0}$. $\square$

Als Vorsorge für später noch eine

Definition. Ist R ein Ring und $R' \subset R$ eine Teilmenge, so heißt R' *Unterring*, wenn R' bezüglich der Addition Untergruppe ist (also entsprechend 1.2.6 für $a, b \in R'$ auch $a + b \in R'$ und $-a \in R'$), und bezüglich der Multiplikation für $a, b \in R'$ auch $a \cdot b \in R'$.

Sind R und S Ringe mit Verknüpfungen $+$, $\cdot$ und $\oplus$, $\odot$, so heißt eine Abbildung $\varphi: R \rightarrow S$ ein *Homomorphismus (von Ringen)*, wenn für alle $a, b \in R$ gilt:

$$\varphi(a + b) = \varphi(a) \oplus \varphi(b) \quad \text{und} \quad \varphi(a \cdot b) = \varphi(a) \odot \varphi(b).$$

Zum Beispiel ist $m\mathbb{Z} \subset \mathbb{Z}$ ein Unterring und $\mathbb{Z} \rightarrow \mathbb{Z}/m\mathbb{Z}$, $a \mapsto a + m\mathbb{Z}$, ein Homomorphismus.

1.3.3. In einem nullteilerfreien Ring R ist für $a, b \in R \setminus \{0\}$ auch das Produkt $a \cdot b \in R \setminus \{0\}$, also induziert die Multiplikation von R eine assoziative Multiplikation in $R \setminus \{0\}$. Das Gruppenaxiom G2 braucht jedoch keineswegs erfüllt zu sein: Im Ring $\mathbb{Z}$ gibt es außer für 1 und -1 kein multiplikatives Inverses, im Ring $2\mathbb{Z}$ der geraden Zahlen nicht einmal ein Einselement. Ist $R \setminus \{0\}$ mit der Multiplikation eine Gruppe und darüber hinaus der Ring kommutativ, so nennt man ihn *Körper*. Das wollen wir noch einmal direkter aufschreiben:

Definition. Eine Menge K zusammen mit zwei Verknüpfungen

$$\begin{aligned} +: K \times K &\rightarrow K, & (a, b) &\mapsto a + b, & \text{und} \\ \cdot: K \times K &\rightarrow K, & (a, b) &\mapsto a \cdot b, \end{aligned}$$

heißt *Körper*, wenn folgendes gilt:

K1 K zusammen mit der Addition $+$ ist eine abelsche Gruppe.

(Ihr neutrales Element wird mit 0, das zu $a \in K$ inverse Element mit $-a$ bezeichnet.)

K2 Bezeichnet $K^* := K \setminus \{0\}$, so gilt für $a, b \in K^*$ auch $a \cdot b \in K^*$, und K^* zusammen mit der so erhaltenen Multiplikation ist eine abelsche Gruppe.

(Ihr neutrales Element wird mit 1, das zu $a \in K^*$ inverse Element mit a^{-1} oder $1/a$ bezeichnet. Man schreibt $b/a = a^{-1}b = ba^{-1}$.)

K3 Es gelten die *Distributivgesetze*, d.h. für $a, b, c \in K$ ist

$$a \cdot (b + c) = a \cdot b + a \cdot c \quad \text{und} \quad (a + b) \cdot c = a \cdot c + b \cdot c.$$

Bemerkung. In einem Körper K gelten die folgenden weiteren Rechenregeln (dabei sind $a, b, x, \tilde{x} \in K$ beliebig):

a) $1 \neq 0$ (also hat ein Körper mindestens zwei Elemente).

$$b) 0 \cdot a = a \cdot 0 = 0.$$

$$c) a \cdot b = 0 \Rightarrow a = 0 \text{ oder } b = 0.$$

$$d) a(-b) = -(ab) \text{ und } (-a)(-b) = ab.$$

$$e) x \cdot a = \tilde{x} \cdot a \text{ und } a \neq 0 \Rightarrow x = \tilde{x}.$$

Beweis. a) ist ganz klar, denn $1 \in K^*$, aber $0 \notin K^*$. b) sieht man wie in 1.3.1. Die Nullteilerfreiheit c) ist in K2 enthalten. d) folgt aus

$$ab + (-a)b = (a + (-a))b = 0 \cdot b = 0 \quad \text{und}$$

$$(-a)(-b) = -((-a)b) = -(-ab) = ab \quad \text{nach 1.2.3, Bem. c).}$$

Die Kürzungsregel e) gilt in K^* , also im Fall $x, \tilde{x} \in K^*$. Ist $x = 0$, so muß auch $\tilde{x} = 0$, also $x = \tilde{x}$ sein. $\square$

1.3.4. Beispiele für Körper. a) Die rationalen Zahlen $\mathbb{Q}$ und die reellen Zahlen $\mathbb{R}$ sind Körper. Das lernt man in der Analysis (vgl. etwa [Fo1], §2).

b) Zur Konstruktion der komplexen Zahlen $\mathbb{C}$ führt man in der reellen Ebene $\mathbb{R} \times \mathbb{R}$ eine Addition und Multiplikation ein. Die naheliegende Frage, warum das im $\mathbb{R}^n$ mit $n > 2$ nicht mehr so geht, wird in [Z] behandelt.

Durch einfaches Nachprüfen der Körperaxiome beweist man:

$$\mathbb{R} \times \mathbb{R} = \{(a, b) : a, b \in \mathbb{R}\}$$

zusammen mit der durch

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

definierten Addition und der durch

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

definierten Multiplikation ist ein Körper mit $(0,0)$ als neutralem Element der Addition, $(-a, -b)$ als Negativem von (a, b) , $(1,0)$ als neutralem Element der Multiplikation und

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

als multiplikativem Inversen.

Wir nennen $\mathbb{R} \times \mathbb{R}$ mit diesen Verknüpfungen den Körper der komplexen Zahlen und bezeichnen ihn mit $\mathbb{C}$.

Die Abbildung

$$\mathbb{R} \rightarrow \mathbb{R} \times \mathbb{R} = \mathbb{C}, \quad a \mapsto (a, 0),$$

ist injektiv. Da

$$\begin{aligned}(a, 0) + (a', 0) &= (a + a', 0) \quad \text{und} \\ (a, 0) \cdot (a', 0) &= (a \cdot a', 0)\end{aligned}$$

gilt, braucht man zwischen $\mathbb{R}$ und

$$\mathbb{R} \times \{0\} = \{(a, b) \in \mathbb{C} : b = 0\}$$

auch hinsichtlich Addition und Multiplikation nicht zu unterscheiden. Man kann also $\mathbb{R}$ mit $\mathbb{R} \times \{0\}$, d.h. a mit $(a, 0)$ „identifizieren“ und $\mathbb{R}$ als Teilmenge von $\mathbb{C}$ betrachten.

Dies wird noch einleuchtender durch folgende übliche Konventionen. Man definiert $i := (0, 1)$ als *imaginäre Einheit*. Dann ist $i^2 = -1$, und für jedes $(a, b) \in \mathbb{C}$ gilt

$$(a, b) = (a, 0) + (0, b) = a + bi.$$

Für $\lambda = (a, b) = a + bi \in \mathbb{C}$ nennt man $\operatorname{re} \lambda := a \in \mathbb{R}$ den *Realteil* und $\operatorname{im} \lambda := b \in \mathbb{R}$ den *Imaginärteil*,

$$\bar{\lambda} := a - bi \in \mathbb{C}$$

heißt die zu λ *konjugiert komplexe Zahl*.

Für die komplexe Konjugation gelten folgende, ganz einfach nachzuweisende Regeln: Für alle $\lambda, \mu \in \mathbb{C}$ ist

$$\begin{aligned}\overline{\lambda + \mu} &= \bar{\lambda} + \bar{\mu}, \\ \overline{\lambda \cdot \mu} &= \bar{\lambda} \cdot \bar{\mu}, \\ \lambda \in \mathbb{R} &\Leftrightarrow \lambda = \bar{\lambda}.\end{aligned}$$

Da für $\lambda = a + bi \in \mathbb{C}$

$$\lambda \cdot \bar{\lambda} = (a + bi) \cdot (a - bi) = a^2 + b^2 \in \mathbb{R}_+,$$

kann man den *Absolutbetrag*

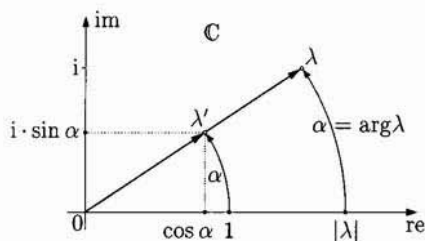
$$|\lambda| := \sqrt{\lambda \cdot \bar{\lambda}} = \sqrt{a^2 + b^2}$$

definieren. Wie man leicht nachrechnet, ist für alle $\lambda, \mu \in \mathbb{C}$

$$|\lambda + \mu| \leq |\lambda| + |\mu| \quad \text{Dreiecksungleichung} \quad \text{und} \quad |\lambda \cdot \mu| = |\lambda| \cdot |\mu|.$$

Vorsicht! Die in $\mathbb{R}$ vorhandene $\leq$ -Relation läßt sich nicht in sinnvoller Weise auf $\mathbb{C}$ fortsetzen. Für komplexe Zahlen kann man daher i.a. nur die Absolutbeträge vergleichen, d.h. für $\lambda, \mu \in \mathbb{C}$ ist

$$|\lambda| \leq |\mu| \quad \text{oder} \quad |\mu| \leq |\lambda|.$$

**Bild 1.4**

Wir wollen noch eine geometrische Beschreibung von Addition und Multiplikation komplexer Zahlen geben. Die Addition entspricht der Addition von Vektoren im $\mathbb{R}^2$ (Bild 1.5, vgl. auch Bild 0.1). Ist λ eine von Null verschiedene komplexe Zahl und $\lambda' = \frac{1}{|\lambda|} \cdot \lambda$, so ist $|\lambda'| = 1$. Es gibt also ein eindeutig bestimmtes $\alpha \in [0, 2\pi[$, so daß

$$\lambda' = \cos \alpha + i \cdot \sin \alpha = e^{i\alpha},$$

wie man in der Analysis lernt.

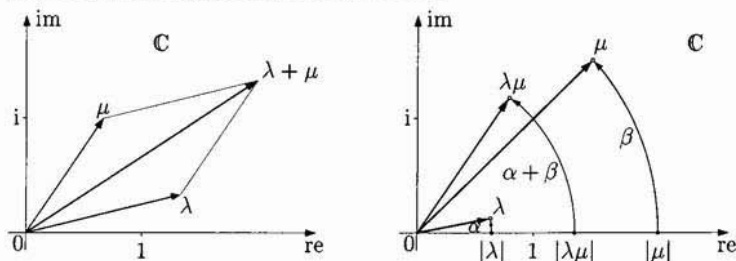
Man nennt $\arg \lambda := \alpha$ das *Argument* von λ , und es ist

$$\lambda = |\lambda| \cdot e^{i \arg \lambda}.$$

Ist $\mu = |\mu| \cdot e^{i \arg \mu}$ eine weitere von Null verschiedene komplexe Zahl, so ist

$$\lambda \cdot \mu = |\lambda| \cdot |\mu| \cdot e^{i \arg \lambda} \cdot e^{i \arg \mu} = |\lambda| \cdot |\mu| \cdot e^{i(\arg \lambda + \arg \mu)}.$$

Bei der Multiplikation komplexer Zahlen werden also die Absolutbeträge multipliziert und die Argumente addiert (Bild 1.5).

**Bild 1.5**

c) Wie wir gesehen haben, gibt es in jedem Körper zwei verschiedene Elemente 0 und 1. Daraus kann man schon einen Körper machen, indem man in $K = \{0, 1\}$ Addition und Multiplikation einführt durch die Tafeln

+	0	1
0	0	1
1	1	0

·	0	1
0	0	0
1	0	1

Offensichtlich ist das auch die einzige Möglichkeit, als Ring war dieser Körper schon in 1.3.1 in der Form $\mathbb{Z}/2\mathbb{Z}$ aufgetreten. Diese Verknüpfungen kann man elektronisch leicht realisieren, daher ist dieser Körper der elementare Baustein aller Computer.

d) In 1.3.2 hatten wir für jedes $m \in \mathbb{N} \setminus \{0\}$ den Restklassenring $\mathbb{Z}/m\mathbb{Z}$ eingeführt und bewiesen, daß er genau dann nullteilerfrei ist, wenn m eine Primzahl ist. Dies ist also eine notwendige Bedingung dafür, ein Körper zu sein. Daß es auch hinreichend ist, folgt aus der etwas allgemeineren

Bemerkung. Ein nullteilerfreier, kommutativer Ring K mit endlich vielen Elementen und Eins ist ein Körper.

Beweis. Nach 1.2.4 genügt es, für jedes $a \in K^*$ zu zeigen, daß die Multiplikation

$$K^* \rightarrow K^*, \quad x \mapsto a \cdot x,$$

eine surjektive Abbildung ist. Wenn K und damit auch K^* endlich ist, genügt dafür die Injektivität (vgl. 1.1.4). Die ist aber klar, denn für $x \neq \tilde{x}$ und $a \cdot x = a \cdot \tilde{x}$ würde

$$a(x - \tilde{x}) = 0 \quad \text{und} \quad a \neq 0, \quad x - \tilde{x} \neq 0$$

gelten. □

Im Ring $\mathbb{Z}$ gilt für jedes $n \in \mathbb{N}$ mit $n \geq 1$

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

In $\mathbb{Z}/m\mathbb{Z}$ mit $m \geq 1$ ist die Restklasse $\bar{1}$ das Einselement, es gilt

$$m \cdot \bar{1} := \underbrace{\bar{1} + \dots + \bar{1}}_{m\text{-mal}} = \bar{1} + \dots + \bar{1} = \bar{m} = \bar{0}.$$

Das zeigt einen grundlegenden Unterschied zwischen den beiden Ringen $\mathbb{Z}$ und $\mathbb{Z}/m\mathbb{Z}$ und motiviert die

Definition. Ist R ein Ring mit Einselement 1, so ist seine *Charakteristik* erklärt durch

$$\text{char}(R) := \begin{cases} 0, & \text{falls } n \cdot 1 \neq 0 \text{ für alle } n \geq 1, \\ \min \{n \in \mathbb{N} \setminus \{0\} : n \cdot 1 = 0\} & \text{sonst.} \end{cases}$$

Lemma. Ist K ein Körper, so ist $\text{char}(K)$ entweder Null oder eine Primzahl.

Beweis. Angenommen, $\text{char}(K) = m = k \cdot \ell \neq 0$ mit $1 < k, \ell < m$. Aus

$$0 = m \cdot 1 = (k \cdot \ell) \cdot 1 = (k \cdot 1)(\ell \cdot 1)$$

folgt wegen der Nullteilerfreiheit $k \cdot 1 = 0$ oder $\ell \cdot 1 = 0$ im Widerspruch zur Minimalität von m . $\square$

Ist p eine Primzahl, so nennt man $\mathbb{Z}/p\mathbb{Z}$ den *Primkörper der Charakteristik p* . In der Algebra lernt man, daß es zu jedem endlichen Körper K eine Primzahl p gibt derart, daß $\mathbb{Z}/p\mathbb{Z}$ Unterkörper von K ist, und die Anzahl der Elemente von K eine Potenz von p ist.

1.3.5. Spätestens bei der Suche nach Eigenwerten in Kapitel 4 wird es sich nicht mehr vermeiden lassen, Polynome beliebigen Grades zu Hilfe zu nehmen. Weil es von der Systematik paßt, wollen wir die dann benötigten Tatsachen schon hier zusammenstellen.

Wir nehmen einen Körper K und eine *Unbestimmte* t . Eine Unbestimmte soll dabei einfach ein Buchstabe sein, für den man alles einsetzen darf, was sinnvoll ist (das kann man präziser formulieren, aber damit wollen wir uns hier nicht aufhalten, vgl. [F-S]). Ein *Polynom* mit *Koeffizienten* in K (oder *Polynom über K*) ist dann ein formaler Ausdruck der Gestalt

$$f(t) = a_0 + a_1 t + \dots + a_n t^n,$$

wobei $a_0, \dots, a_n \in K$. Meist schreibt man statt $f(t)$ nur f . Mit $K[t]$ bezeichnen wir die Menge all solcher Polynome. Sind alle Koeffizienten $a_v = 0$, so spricht man vom *Nullpolynom* und schreibt $f = 0$. Der *Grad* von f ist erklärt als

$$\deg f := \begin{cases} -\infty, & \text{falls } f = 0, \\ \max\{v \in \mathbb{N} : a_v \neq 0\}, & \text{sonst.} \end{cases}$$

Schließlich heißt f *normiert*, wenn $a_n = 1$.

Das nächstliegende, was man für die Unbestimmte t einsetzen kann, sind Elemente aus K . Ist $\lambda \in K$, so ist auch

$$f(\lambda) := a_0 + a_1 \lambda + \dots + a_n \lambda^n \in K,$$

aus dem Polynom f erhält man also eine Abbildung

$$\tilde{f}: K \rightarrow K, \quad \lambda \mapsto f(\lambda),$$

insgesamt also (mit der Notation aus 1.1.3) eine Abbildung

$$\tilde{} : K[t] \rightarrow \text{Abb}(K, K), \quad f \mapsto \tilde{f}.$$

Die etwas pedantische Unterscheidung von f und $\tilde{f}$ ist leider nötig, wenn man sich einmal mit endlichen Körpern eingelassen hat.

Beispiel. Ist $K = \{0, 1\}$ der Körper mit zwei Elementen aus 1.3.4 und

$$f = t^2 + t, \quad \text{so ist} \quad f(0) = 0 + 0 = 0 \quad \text{und} \quad f(1) = 1 + 1 = 0,$$

also $\tilde{f}$ die Nullabbildung, obwohl $f \neq 0$, weil $a_1 = a_2 = 1$. Die obige Abbildung $\tilde{}$ ist also nicht injektiv.

1.3.6. Die Menge $K[t]$ hat viel Struktur, insbesondere eine natürliche Addition und Multiplikation. Dazu nehmen wir $f, g \in K[t]$. Ist

$$f = a_0 + a_1 t + \dots + a_n t^n, \quad g = b_0 + b_1 t + \dots + b_m t^m,$$

so können wir zur Definition der Addition $m = n$ annehmen (ist etwa $m < n$, so setze man $b_{m+1} = \dots = b_n = 0$). Dann ist

$$f + g := (a_0 + b_0) + (a_1 + b_1)t + \dots + (a_n + b_n)t^n.$$

Die Multiplikation ist dadurch erklärt, daß man formal ausmultipliziert, also

$$f \cdot g := c_0 + c_1 t + \dots + c_{n+m} t^{n+m} \quad \text{mit} \quad c_k = \sum_{i+j=k} a_i b_j.$$

Insbesondere ist

$$\begin{aligned} c_0 &= a_0 b_0, \\ c_1 &= a_0 b_1 + a_1 b_0, \\ c_2 &= a_0 b_2 + a_1 b_1 + a_2 b_0, \\ &\vdots \\ c_{n+m} &= a_n b_m. \end{aligned}$$

Ist $f \cdot g = h$, so nennt man f und g *Teiler* von h .

Bemerkung. Ist K ein Körper, so ist die Menge $K[t]$ der Polynome über K zusammen mit den oben definierten Verknüpfungen ein kommutativer Ring. Weiter gilt

$$\deg(f \cdot g) = \deg f + \deg g$$

für $f, g \in K[t]$. Dabei soll formal $n - \infty = -\infty + m = -\infty + (-\infty) = -\infty$ sein.

Man nennt $K[t]$ den *Polynomring* über K .

Beweis. Der Nachweis der Axiome erfordert nur geduldiges Rechnen. Die Aussage über den Grad folgt aus $a_n b_m \neq 0$, falls $a_n, b_m \neq 0$. $\square$

Es sei angemerkt, daß man analog für einen kommutativen Ring R einen kommutativen Polynomring $R[t]$ erhält. Die Aussage über den Grad des Produktpolynoms gilt nur über einem nullteilerfreien Ring.

1.3.7. Der Mangel eines Ringes gegenüber einem Körper ist der, daß man im allgemeinen nicht dividieren kann. Bei ganzen Zahlen hat man als Ersatz eine *Division mit Rest* (vgl. 1.2.7), die ganz analog auch für Polynome erklärt werden kann.

Satz. Sind $f, g \in K[t]$, und ist $g \neq 0$, so gibt es dazu eindeutig bestimmte Polynome $q, r \in K[t]$ derart, daß

$$f = q \cdot g + r \quad \text{und} \quad \deg r < \deg g.$$

Man kann die Beziehung auch in der nicht ohne weiteres erlaubten, aber sehr suggestiven Form

$$\frac{f}{g} = q + \frac{r}{g}$$

schreiben. Der Buchstabe q steht für „Quotient“, r für „Rest“.

Beweis. Wir zeigen zunächst die Eindeutigkeit. Seien $q, r, q', r' \in K[t]$ mit

$$f = q \cdot g + r = q' \cdot g + r', \quad \text{wobei} \quad \deg r, \deg r' < \deg g.$$

Durch Subtraktion folgt

$$0 = (q - q') \cdot g + (r - r'), \quad \text{also} \quad (q - q') \cdot g = r' - r. \quad (*)$$

Da zwei gleiche Polynome gleichen Grad haben, folgt aus $q \neq q'$ wegen $g \neq 0$

$$\deg(r' - r) = \deg(q - q') + \deg g \geq \deg g,$$

was nicht sein kann. Also ist $q = q'$, und aus $(*)$ folgt auch $r = r'$.

Zum Beweis der Existenz von q und r geben wir ein Verfahren an, mit dem man diese Polynome schrittweise berechnen kann. Dazu schreibt man die Polynome am besten nach fallenden Potenzen, also

$$f = a_n t^n + \dots + a_1 t + a_0, \quad g = b_m t^m + \dots + b_1 t + b_0,$$

wobei $a_n, b_m \neq 0$. Ist $n < m$, so kann man $q = 0$ und $r = f$ wählen, denn

$$f = 0 \cdot g + f \quad \text{und} \quad \deg f < \deg g.$$

Im Fall $n \geq m$ teilt man zunächst die höchsten Terme von f und g , das ergibt

$$q_1 := \frac{a_n}{b_m} t^{n-m}$$

als höchsten Term von q . Im nächsten Schritt betrachtet man

$$f_1 := f - q_1 \cdot g.$$

Nach Definition ist $\deg f_1 < \deg f$. Ist $\deg f_1 < m$, so kann man $q = q_1$ und $r = f_1$ setzen. Andernfalls wiederholt man den ersten Schritt mit f_1 statt f , d.h. man erhält den nächsten Term q_2 von q und

$$f_2 := f_1 - q_2 \cdot g \quad \text{mit} \quad \deg f_2 < \deg f_1.$$

Da die Grade der f_i bei jedem Schritt um mindestens eins abnehmen, erhält man schließlich ein $k \leq n - m + 1$, so daß für

$$f_k := f_{k-1} - q_k \cdot g \quad \text{erstmal} \quad \deg f_k < \deg g,$$

und damit bricht das Verfahren ab: Setzt man die Gleichungen ineinander ein, so ergibt sich

$$f = q_1 g + f_1 = (q_1 + q_2)g + f_2 = \dots = (q_1 + \dots + q_k)g + f_k,$$

also ist eine Lösung unseres Problems gegeben durch

$$q := q_1 + \dots + q_k \quad \text{und} \quad r := f_k. \quad \square$$

Da bei der Konstruktion von q immer nur durch b_m dividiert wird, kann man im Fall $b_m = 1$ den Körper K durch einen Ring ersetzen.

Beispiel. Sei $K = \mathbb{R}$, $f = 3t^3 + 2t + 1$, $g = t^2 - 4t$.

Die Rechnung verläuft nach folgendem Schema:

$$\begin{array}{r}
 (3t^3 \quad + 2t + 1) : (t^2 - 4t) = 3t + 12 + (50t + 1) : (t^2 - 4t) \\
 \underline{-3t^3 \quad + 12t^2} \\
 12t^2 \quad + 2t + 1 \\
 \underline{-12t^2 \quad + 48t} \\
 50t + 1
 \end{array}$$

Es ist also $q = 3t + 12$ und $r = 50t + 1$.

1.3.8. Nach diesen formalen Vorbereitungen kommen wir nun zum eigentlichen Thema, nämlich der Frage nach der Existenz von *Nullstellen* (d.h. $\lambda \in K$ mit $f(\lambda) = 0$) bei Polynomen. Darauf kann man nämlich viele andere Fragen zurückführen, etwa in Kapitel 4 die Frage nach der Existenz von Eigenwerten. Neben der reinen Existenzfrage ist es für die Praxis wichtig, Verfahren für die näherungsweise Berechnung von Nullstellen zu haben. Das lernt man in der numerischen Mathematik (vgl. etwa [O]).

Beispiele. a) Ist $K = \mathbb{R}$ und $f = t^2 + 1$, so ist $f(\lambda) \geq 1$ für alle $\lambda \in \mathbb{R}$. Also gibt es keine Nullstelle.

b) Ist $K = \{a_0, a_1, \dots, a_n\}$ ein endlicher Körper und

$$f = (t - a_0) \cdot \dots \cdot (t - a_n) + 1,$$

so ist $f(\lambda) = 1$ für alle $\lambda \in K$, also hat f keine Nullstelle.

Zum Glück sind diese beiden Beispiele von ausgewählter Bosheit; im allgemeinen hat man doch etwas mehr Chancen, eine Nullstelle zu finden. Hat man eine gefunden, so genügt es zur Suche nach weiteren, ein Polynom von kleinerem Grad zu betrachten:

Lemma. Ist $\lambda \in K$ eine Nullstelle von $f \in K[t]$, so gibt es ein eindeutig bestimmtes $g \in K[t]$ mit folgenden Eigenschaften:

1) $f = (t - \lambda) \cdot g.$

2) $\deg g = (\deg f) - 1.$

Beweis. Wir dividieren f durch $(t - \lambda)$ mit Rest; es gibt also eindeutig bestimmte $g, r \in K[t]$ mit

$$f = (t - \lambda)g + r \quad \text{und} \quad \deg r < \deg(t - \lambda) = 1.$$

Also ist $r = a_0$ mit $a_0 \in K$. Aus $f(\lambda) = 0$ folgt durch Einsetzen von λ

$$0 = (\lambda - \lambda) \cdot g(\lambda) + r = 0 + a_0,$$

also ist $a_0 = r = 0$, und 1) ist bewiesen. Wegen

$$\deg f = \deg(t - \lambda) + \deg g = 1 + \deg g$$

folgt 2). □

Korollar 1. Sei K ein beliebiger Körper, $f \in K[t]$ ein Polynom und k die Anzahl der Nullstellen von f . Ist f vom Nullpolynom verschieden, so gilt

$$k \leq \deg f.$$

Beweis. Wir führen Induktion über den Grad von f . Für $\deg f = 0$ ist $f = a_0 \neq 0$ ein konstantes Polynom. Dieses hat gar keine Nullstelle, also ist unsere Behauptung richtig.

Sei $\deg f = n \geq 1$, und sei die Aussage schon für alle Polynome $g \in K[t]$ mit $\deg g \leq n - 1$ bewiesen. Wenn f keine Nullstelle hat, ist die Behauptung richtig. Ist $\lambda \in K$ eine Nullstelle, so gibt es nach dem Lemma ein $g \in K[t]$ mit

$$f = (t - \lambda) \cdot g \quad \text{und} \quad \deg g = n - 1.$$

Alle von λ verschiedenen Nullstellen von f müssen auch Nullstellen von g sein. Ist l die Anzahl der Nullstellen von g , so ist nach Induktionsannahme

$$l \leq n - 1, \quad \text{also} \quad k \leq l + 1 \leq n.$$

□

Korollar 2. Ist K unendlich, so ist die Abbildung

$$\sim : K[t] \rightarrow \text{Abb}(K, K), \quad f \mapsto \tilde{f},$$

injektiv.

Beweis. Seien $f_1, f_2 \in K[t]$ und $g := f_2 - f_1$. Ist $\tilde{f}_1 = \tilde{f}_2$, so folgt $\tilde{g} = 0$, d.h. $g(\lambda) = 0$ für alle $\lambda \in K$. Also hat g unendlich viele Nullstellen, und aus Korollar 1 folgt $g = 0$, somit ist $f_1 = f_2$. $\square$

Ist λ Nullstelle von f , also $f = (t - \lambda) \cdot g$, so kann λ auch Nullstelle von g sein. Man spricht dann von einer *mehrfachen* Nullstelle.

Definition. Ist $f \in K[t]$ vom Nullpolynom verschieden und $\lambda \in K$, so heißt

$$\mu(f; \lambda) := \max\{r \in \mathbb{N} : f = (t - \lambda)^r \cdot g \text{ mit } g \in K[t]\}$$

die *Vielfachheit der Nullstelle* λ von f .

Nach dem Lemma gilt $\mu(f; \lambda) = 0 \Leftrightarrow f(\lambda) \neq 0$. Ist

$$f = (t - \lambda)^r \cdot g \quad \text{mit} \quad r = \mu(f; \lambda),$$

so folgt $g(\lambda) \neq 0$. Die Vielfachheit der Nullstelle λ gibt also an, wie oft der Linearfaktor $(t - \lambda)$ in f enthalten ist.

Ist $K = \mathbb{R}$ oder $\mathbb{C}$, so kann man die Vielfachheit der Nullstelle mit den r -ten Ableitungen $f^{(r)}$ von f in Beziehung bringen. Es gilt

$$\mu(f; \lambda) = \max\{r \in \mathbb{N} : f(\lambda) = f'(\lambda) = \dots = f^{(r-1)}(\lambda) = 0\},$$

wie man leicht nachrechnet.

1.3.9. Sind $\lambda_1, \dots, \lambda_k \in K$ die verschiedenen Nullstellen eines Polynoms $f \in K[t]$, und ist $r_i = \mu(f; \lambda_i)$, so ist

$$f = (t - \lambda_1)^{r_1} \cdot \dots \cdot (t - \lambda_k)^{r_k} \cdot g,$$

wobei g ein Polynom vom Grad $n - (r_1 + \dots + r_k)$ ohne Nullstellen ist. Der schlimmste Fall ist $g = f$, der beste $\deg g = 0$, d.h. f zerfällt in *Linearfaktoren*. Die wichtigste Existenzaussage für Nullstellen von Polynomen macht der sogenannte

Fundamentalsatz der Algebra. Jedes Polynom $f \in \mathbb{C}[t]$ mit $\deg f > 0$ hat mindestens eine Nullstelle.

Dieser Satz wurde von C.F. GAUSS erstmals 1799 bewiesen. Es gibt dafür sehr viele Beweise, die aber alle Hilfsmittel aus der Analysis benutzen, denn $\mathbb{C}$ entsteht aus $\mathbb{R}$, und die reellen Zahlen sind ein Produkt der Analysis. Der wohl kürzeste Beweis verwendet Hilfsmittel aus der Theorie der holomorphen Funktionen (vgl. etwa [F-L]).

Hat $f \in \mathbb{C}[t]$ eine Nullstelle λ , so kann man sie herausdividieren, also

$$f = (t - \lambda) \cdot g$$

schreiben. Ist $\deg g > 0$, so hat auch g eine komplexe Nullstelle, und indem man das Verfahren so lange wiederholt, bis das verbleibende Polynom den Grad Null hat, ergibt sich das

Korollar. Jedes Polynom $f \in \mathbb{C}[t]$ zerfällt in Linearfaktoren, d.h. es gibt a und $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ mit $n = \deg f$, so daß

$$f = a(t - \lambda_1) \cdot \dots \cdot (t - \lambda_n).$$

1.3.10. Nun wollen wir aus dem Fundamentalsatz der Algebra Aussagen über die Nullstellen reeller Polynome folgern. Dazu betrachten wir $\mathbb{R}$ als Teilmenge von $\mathbb{C}$ (vgl. 1.3.4). Dann ist auch $\mathbb{R}[t]$ Teilmenge von $\mathbb{C}[t]$.

Lemma. Ist $f \in \mathbb{R}[t]$ und $\lambda \in \mathbb{C}$ eine Nullstelle von f , so ist auch die konjugiert komplexe Zahl $\bar{\lambda} \in \mathbb{C}$ eine Nullstelle von f . Es gilt sogar

$$\mu(f; \lambda) = \mu(f; \bar{\lambda}).$$

Die komplexen Nullstellen eines Polynoms mit reellen Koeffizienten liegen also symmetrisch zur reellen Achse.

Beweis. Ist

$$f = a_0 + a_1 t + \dots + a_n t^n,$$

so ist wegen $a_0 = \bar{a}_0, \dots, a_n = \bar{a}_n$ nach den Rechenregeln für die komplexe Konjugation

$$f(\bar{\lambda}) = a_0 + a_1 \bar{\lambda} + \dots + a_n (\bar{\lambda})^n = \bar{a}_0 + \bar{a}_1 \bar{\lambda} + \dots + \bar{a}_n \bar{\lambda}^n = \overline{f(\lambda)} = \overline{0} = 0.$$

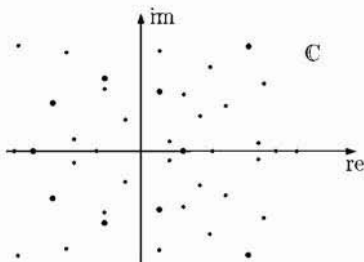


Bild 1.6

Also ist auch $\bar{\lambda}$ Nullstelle von f . Um die Vielfachheiten von λ und $\bar{\lambda}$ zu vergleichen, genügt es für jedes $k \in \mathbb{N}$

$$\mu(f; \lambda) \geq k \Rightarrow \mu(f; \bar{\lambda}) \geq k$$

zu beweisen. Für $\lambda = \bar{\lambda}$ ist die Aussage trivial. Für $\lambda \neq \bar{\lambda}$ verwenden wir den folgenden

Hilfssatz. Sei $f \in \mathbb{R}[t]$ und $\lambda \in \mathbb{C}$ eine nicht reelle Nullstelle von f , sowie $g := (t - \lambda)(t - \bar{\lambda}) \in \mathbb{C}[t]$. Dann gilt:

1) $g \in \mathbb{R}[t]$.

2) Es gibt ein $q \in \mathbb{R}[t]$ mit $f = g \cdot q$.

Nun genügt es, durch Induktion über k aus $\mu(f; \lambda) \geq k$ die Existenz eines Polynoms $f_k \in \mathbb{R}[t]$ mit

$$f = g^k \cdot f_k$$

zu folgern. Dabei ist g wie im Hilfssatz erklärt.

Für $k = 0$ ist nichts zu beweisen. Sei also die Aussage für $k \geq 0$ bewiesen und $\mu(f; \lambda) \geq k + 1$. Dann ist $f = g^k \cdot f_k$, und es muß $f_k(\lambda) = 0$ sein; aus dem Hilfssatz folgt die Existenz eines $f_{k+1} \in \mathbb{R}[t]$ mit

$$f_k = g \cdot f_{k+1}, \quad \text{also ist} \quad f = g^{k+1} \cdot f_{k+1}.$$

Es bleibt der Hilfssatz zu beweisen. Dazu setzen wir

$$\lambda = \alpha + i\beta \quad \text{mit} \quad \alpha, \beta \in \mathbb{R}.$$

Dann ist

$$g = (t - \lambda)(t - \bar{\lambda}) = (t - \alpha - i\beta)(t - \alpha + i\beta) = t^2 - 2\alpha t + \alpha^2 + \beta^2 \in \mathbb{R}[t].$$

Durch Division mit Rest in $\mathbb{R}[t]$ erhält man $q, r \in \mathbb{R}[t]$ mit

$$f = g \cdot q + r \quad \text{und} \quad \deg r \leq (\deg g) - 1 = 1.$$

Diese Gleichung gilt selbstverständlich auch in $\mathbb{C}[t]$. Durch Einsetzen von λ und $\bar{\lambda}$ folgt

$$r(\lambda) = r(\bar{\lambda}) = 0.$$

Nach Korollar 1 aus 1.3.8 folgt $r = 0$ wegen $\lambda \neq \bar{\lambda}$. Damit ist der Hilfssatz und auch das Lemma bewiesen. $\square$

Nun wenden wir auf ein Polynom $f \in \mathbb{R}[t]$ den Fundamentalsatz der Algebra an. Danach gibt es $a, \lambda_1, \dots, \lambda_n \in \mathbb{C}$, so daß

$$f = a(t - \lambda_1) \cdots (t - \lambda_n).$$

Da a der höchste Koeffizient ist, gilt $a \in \mathbb{R}$.

Seien $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ und $\lambda_{k+1}, \dots, \lambda_n \in \mathbb{C}$. Nach dem Lemma ist $n - k$, d.h. die mit Vielfachheit gezählte Anzahl der nicht reellen Nullstellen, gerade, und durch Umnummerierung kann man erreichen, daß

$$\lambda_{k+1} = \bar{\lambda}_{k+2}, \dots, \lambda_{n-1} = \bar{\lambda}_n$$

gilt. Jedes Paar $\lambda, \bar{\lambda}$ konjugiert komplexer Nullstellen mit $\lambda = \alpha + i\beta$ kann man also (wie wir im Hilfssatz gesehen haben) zu einem normierten quadratischen Faktor

$$g = (t - \lambda)(t - \bar{\lambda}) = t^2 - 2\alpha t + (\alpha^2 + \beta^2) \in \mathbb{R}[t]$$

zusammenfassen. g hat keine reelle Nullstelle, was man auch an der *Diskriminante* ablesen kann:

$$4\alpha^2 - 4(\alpha^2 + \beta^2) = -4\beta^2 < 0.$$

Damit ist folgendes Ergebnis bewiesen:

Theorem. Jedes Polynom $f \in \mathbb{R}[t]$ mit $\deg f = n \geq 1$ gestattet eine Zerlegung

$$f = a(t - \lambda_1) \cdot \dots \cdot (t - \lambda_r) \cdot g_1 \cdot \dots \cdot g_m,$$

wobei $a, \lambda_1, \dots, \lambda_r$ reell sind, mit $a \neq 0$, und $g_1, \dots, g_m \in \mathbb{R}[t]$ normierte Polynome vom Grad 2 ohne reelle Nullstellen sind. Insbesondere ist $n = r + 2m$.

Korollar. Jedes Polynom $f \in \mathbb{R}[t]$ von ungeradem Grad hat mindestens eine reelle Nullstelle.

Dies folgt sofort aus der Gleichung $n = r + 2m$. Natürlich kann man die Aussage des Korollars mit Hilfe des Zwischenwertsatzes viel einfacher direkt beweisen.

Es sei erwähnt, daß man den Fundamentalsatz der Algebra ohne weitere Hilfsmittel der Analysis aus diesem Korollar ableiten kann (vgl. etwa [F-S]).

Der Fundamentalsatz der Algebra ist eine reine Existenzaussage, d.h. man erhält daraus kein Verfahren zur praktischen Bestimmung der Nullstellen. Für ein Polynom

$$at^2 + bt + c \in \mathbb{C}[t]$$

von Grad 2 kann man die Nullstelle nach der Formel

$$\frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

berechnen. Etwas kompliziertere Formeln dieser Art gibt es auch für die Nullstellen von Polynomen vom Grad 3 und 4.

Wie erstmals N.H. ABEL im Jahre 1826 zeigte, kann es solche allgemeine Formeln für Polynome vom Grad größer als 4 aus algebraischen Gründen nicht geben. Man ist daher weitgehend auf Näherungsverfahren zur Approximation der Nullstellen angewiesen.

1.3.11. Wie gerade erläutert, gibt es keine allgemein gültige Formel zur Berechnung der Nullstellen eines Polynoms aus den Koeffizienten. Die umgekehrte Aufgabe ist jedoch ganz einfach: Ist $f \in K[t]$ und

$$f(t) = t^n + \alpha_{n-1}t^{n-1} + \dots + \alpha_1t + \alpha_0 = (t - \lambda_1) \cdot \dots \cdot (t - \lambda_n),$$

d.h. zerfällt f in Linearfaktoren, so folgt durch ausmultiplizieren der rechten Seite

$$\begin{aligned}\alpha_0 &= (-1)^n \lambda_1 \cdot \dots \cdot \lambda_n, \\ \alpha_1 &= (-1)^{n-1} (\lambda_2 \cdot \dots \cdot \lambda_n + \lambda_1 \lambda_3 \cdot \dots \cdot \lambda_n + \dots + \lambda_1 \cdot \dots \cdot \lambda_{n-1}), \\ &\vdots \\ \alpha_{n-2} &= \lambda_1 \lambda_2 + \dots + \lambda_1 \lambda_n + \lambda_2 \lambda_3 + \dots + \lambda_2 \lambda_n + \dots + \lambda_{n-1} \lambda_n, \\ \alpha_{n-1} &= -(\lambda_1 + \dots + \lambda_n).\end{aligned}$$

Um das formal besser aufschreiben zu können, definieren wir für $k = 1, \dots, n$ die *elementarsymmetrischen Funktionen*

$$s_k(\lambda_1, \dots, \lambda_n) := \sum_{1 \leq i_1 < \dots < i_k \leq n} \lambda_{i_1} \cdot \dots \cdot \lambda_{i_k}.$$

Die Summe hat so viele Summanden, wie es Möglichkeiten gibt, k verschiedene Indizes $i_1, \dots, i_k \in \{1, \dots, n\}$ auszuwählen, also $\binom{n}{k}$ Summanden. Für die Koeffizienten von f gilt dann

$$\alpha_k = (-1)^{n-k} s_{n-k}(\lambda_1, \dots, \lambda_n).$$

Diese Aussage nennt man den *Wurzelsatz von VIETA*. Er gibt an, wie sich die Koeffizienten aus den Nullstellen berechnen.

Wie schon gesagt, ist umgekehrt die Bestimmung der Nullstellen aus den Koeffizienten im Allgemeinen ein sehr schwieriges Problem. Im Spezialfall eines reellen Polynoms gibt es immerhin einen Zusammenhang zwischen den Vorzeichen der Koeffizienten und den Nullstellen. Die Tendenz ist klar: Die Vorzeichen der Koeffizienten beeinflussen den Verlauf der Werte und damit die Lage der reellen Nullstellen eines Polynoms.

Vorweg behandeln wir den elementarsten Fall, der später (in 5.7.3) beim Test der Definitheit einer Matrix nützlich sein wird. Wir machen dabei die sehr einschränkende Voraussetzung, daß das Polynom $f \in \mathbb{R}[t]$ (mit Vielfachheit gezählt) so viele reelle Nullstellen hat, wie sein Grad angibt.

Vorzeichenregel. *Angenommen, das reelle Polynom*

$$f(t) = t^n + \alpha_{n-1}t^{n-1} + \dots + \alpha_1t + \alpha_0$$

hat reelle Nullstellen $\lambda_1, \dots, \lambda_n$. Dann gilt:

- a) *Genau dann sind alle Nullstellen λ_i negativ, wenn alle Koeffizienten α_j positiv sind.*

b) Genau dann sind alle Nullstellen λ_i positiv, wenn die Vorzeichen der Koeffizienten α_j alternierend sind, d.h. es ist

$$(-1)^{n-j} \alpha_j > 0 \text{ für } j = 0, \dots, n-1.$$

Beweis. Es genügt, den Fall a) zu beweisen; Fall b) folgt daraus sofort, indem man das Polynom f_- mit $f_-(t) := f(-t)$ betrachtet.

Sind alle $\alpha_j > 0$ und ist $\lambda \geq 0$, so ist

$$f(\lambda) = \lambda^n + \alpha_{n-1} \lambda^{n-1} + \dots + \alpha_0 \geq \alpha_0 > 0.$$

Also sind alle Nullstellen negativ.

Sind alle $\lambda_i < 0$, so ist s_k positiv für gerades k und negativ für ungerades k . Also ist

$$\alpha_j = (-1)^{n-j} \cdot s_{n-j} > 0$$

für alle j . □

Um eine allgemeinere Aussage zu erhalten, definiert man zunächst für ein beliebiges Polynom

$$g(t) := t^m + \beta_{m-1} t^{m-1} + \dots + \beta_0 \in \mathbb{R}[t]$$

die Zahl $Z(g)$ der *Zeichenwechsel* von g wie folgt: Man schreibt die Koeffizienten

$$1, \beta_{m-1}, \beta_{m-2}, \dots, \beta_1, \beta_0$$

auf und zählt von links beginnend ab, wie oft die Vorzeichen der β_i wechseln. Koeffizienten Null werden dabei einfach übergangen. Für den gleich folgenden Beweis schreiben wir das noch etwas präziser auf. Es gibt eindeutig bestimmte Indizes k_p mit

$$m > k_1 > k_2 > \dots > k_r \geq 0$$

derart, daß k_1 (von oben beginnend) der erste Index mit $\beta_{k_1} < 0$ ist, k_2 der erste darauf folgende mit $\beta_{k_2} > 0$ usw. Schließlich findet bei β_{k_r} der letzte Vorzeichenwechsel statt. Dann ist offensichtlich

$$\beta_{k_p} \begin{cases} > 0 \text{ für } p \text{ gerade,} \\ < 0 \text{ für } p \text{ ungerade,} \end{cases} \quad \text{und } r = Z(g) \leq m = \deg g.$$

Ohne Voraussetzung über die Gesamtzahl der reellen Nullstellen eines Polynoms $f \in \mathbb{R}[t]$ kann man nun mit Hilfe der Vorzeichen der Koeffizienten eine Abschätzung für die Zahlen

$N_+(f) := \text{Anzahl der positiven Nullstellen von } f$ und

$N_-(f) := \text{Anzahl der negativen Nullstellen von } f$

geben. Dabei sei wie oben $f_- \in \mathbb{R}[t]$ definiert durch $f_-(t) := f(-t)$.

Vorzeichenregel von DESCARTES. Für ein Polynom

$$f(t) = t^n + \alpha_{n-1} t^{n-1} + \dots + \alpha_0 \in \mathbb{R}[t]$$

mit $\alpha_0 \neq 0$ gilt

$$N_+(f) \leq Z(f) \text{ und } N_-(f) \leq Z(f_-).$$

Mit etwas zusätzlichem Aufwand kann man zeigen, daß die Differenz

$$Z(f) - N_+(f) \text{ immer gerade ist.}$$

Beweis. Da beim Übergang von f zu f_- die Vorzeichen der Nullstellen umgekehrt werden, genügt es, die erste Ungleichung zu zeigen. Dazu benutzen wir das

Lemma. Ist $g(t) = t^m + \beta_{m-1}t^{m-1} + \dots + \beta_0 \in \mathbb{R}[t]$ mit $\beta_0 \neq 0$ und $0 < \lambda \in \mathbb{R}$, so folgt

$$Z((t - \lambda) \cdot g) \geq Z(g) + 1.$$

Bei Multiplikation mit einem Linearfaktor mit positiver Nullstelle nehmen also die Zeichenwechsel zu.

Sind $\lambda_1, \dots, \lambda_k$ mit $k = N_+(f)$ die positiven reellen Nullstellen von f , so gibt es eine Zerlegung

$$f(t) = (t - \lambda_1) \cdot \dots \cdot (t - \lambda_k) \cdot g(t)$$

mit $g(t) \in \mathbb{R}[t]$. Wegen $Z(g) \geq 0$ folgt durch k -malige Anwendung des Lemmas $Z(f) \geq Z(g) + k \geq k = N_+(f)$. □

Es bleibt der etwas knifflige *Beweis des Lemmas*. Für die Koeffizienten des Produktpolynoms

$$(t - \lambda)(t^m + \beta_{m-1}t^{m-1} + \dots + \beta_0) = t^{m+1} + \gamma_m t^m + \dots + \gamma_0$$

gilt:

$$\gamma_m = \beta_{m-1} - \lambda,$$

$$\gamma_j = \beta_{j-1} - \lambda \beta_j, \quad \text{für } j = 1, \dots, m-1,$$

$$\gamma_0 = -\lambda \beta_0.$$

Wir betrachten die oben angegebenen kritischen Indizes $k_1, \dots, k_r$. Ist ρ ungerade, so folgt

$$\beta_{k_\rho} < 0, \quad \beta_{k_\rho+1} \geq 0, \quad \text{also } \gamma_{k_\rho+1} = \beta_{k_\rho} - \lambda \beta_{k_\rho+1} < 0.$$

Für gerade ρ ist

$$\beta_{k_\rho} > 0, \quad \beta_{k_\rho+1} \leq 0, \quad \text{also } \gamma_{k_\rho+1} = \beta_{k_\rho} - \lambda \beta_{k_\rho+1} > 0.$$

Das bedeutet, daß β_{k_ρ} und $\gamma_{k_\rho+1}$ für $\rho = 1, \dots, r$ dasselbe Vorzeichen haben.

Wegen $\gamma_0 = -\lambda\beta_0$ haben γ_0 und β_0 entgegengesetzte Vorzeichen. Daher hat $(t - \lambda) \cdot g$ mindestens einen Vorzeichenwechsel mehr als g . $\square$

Noch allgemeiner kann man mit Hilfe so genannter STURMScher Ketten die Anzahl der Nullstellen abschätzen, die in einem vorgegebenen Intervall liegen (vgl. dazu etwa [D], [St], [Wi]).

Aufgaben zu 1.3

1. Bestimmen Sie (bis auf Isomorphie) alle Körper mit 3 bzw. 4 Elementen.

2. K und K' seien zwei Körper und $\varphi: K \rightarrow K'$ ein Ringhomomorphismus. Zeigen Sie, dass φ entweder injektiv oder der Nullhomomorphismus ist.

3. Ist R ein Ring, M eine beliebige nichtleere Menge und $S = \text{Abb}(M; R)$ die Menge aller Abbildungen von M nach R , so ist auf S durch

$$(f + g)(m) := f(m) + g(m), (f \cdot g)(m) := f(m) \cdot g(m),$$

eine Addition und eine Multiplikation erklärt.

a) Zeigen Sie, dass S auf diese Weise zu einem Ring wird.

b) Ist S ein Körper, falls R ein Körper ist?

4.* Sei $p \in \mathbb{N}$ eine Primzahl und $n \in \mathbb{N} \setminus \{0\}$. Zeigen Sie, dass es einen Körper mit p^n Elementen gibt.

5. Sei K' ein Körper und K ein Unterkörper von K' .

Zeigen Sie: Sind $f, g \in K[t]$, $q \in K'[t]$ mit $f = qg$, so folgt bereits $q \in K[t]$.

6. Sei K ein Körper und $x_0, \dots, x_n, y_0, \dots, y_n \in K$ mit $x_i \neq x_j$ für alle $i \neq j$. Zeigen Sie, dass es genau ein Polynom $f \in K[t]$ vom Grad $\leq n$ gibt, so dass $f(x_i) = y_i$ für $i = 0, \dots, n$. Hinweis: Konstruieren Sie zuerst Polynome $g_k \in K[t]$ vom Grad $\leq n$ mit

$$g_k(x_i) = \begin{cases} 1 & \text{für } i = k, \\ 0 & \text{für } i \neq k. \end{cases}$$

7. Seien $f, g \in \mathbb{C}[t]$ Polynome mit $\mu(f, \lambda) \leq \mu(g, \lambda)$ für alle $\lambda \in \mathbb{C}$. Zeigen Sie, dass dann f ein Teiler von g ist. Gilt diese Aussage auch in $\mathbb{R}[t]$?

8. Sei K ein Körper und $\sim: K[t] \rightarrow \text{Abb}(K, K)$, $f \mapsto \tilde{f}$, die Abbildung aus 1.3.5, die jedem Polynom f die zugehörige Abbildung $\tilde{f}$ zuordnet. Zeigen Sie, dass $\sim$ surjektiv, aber nicht injektiv ist, falls der Körper K endlich ist.

9. Analog zu 1.3.5 definiert man ein Polynom mit Koeffizienten über einem Körper K in n Unbestimmten $t_1, \dots, t_n$ als einen formalen Ausdruck der Gestalt

$$f(t_1, \dots, t_n) = \sum_{0 \leq i_1, \dots, i_n \leq k} a_{i_1 \dots i_n} \cdot t_1^{i_1} \cdot \dots \cdot t_n^{i_n},$$

wobei $k \in \mathbb{N}$ und $a_{i_1 \dots i_n} \in K$. $K[t_1, \dots, t_n]$ bezeichne die Menge all solcher Polynome. Wie für Polynome in einer Unbestimmten kann auch in $K[t_1, \dots, t_n]$ eine Addition und eine Multiplikation erklärt werden. Sind $f, g \in K[t_1, \dots, t_n]$, so erfolgt die Addition von f und g koeffizientenweise und die Multiplikation wieder durch formales Ausmultiplizieren.

- a) Finden Sie Formeln für die Addition und Multiplikation von Polynomen in $K[t_1, \dots, t_n]$, und zeigen Sie, dass $K[t_1, \dots, t_n]$ auf diese Weise zu einem nullteilerfreien, kommutativen Ring wird.

Ein Polynom $h \in K[t_1, \dots, t_n] \setminus \{0\}$ heißt *homogen* (vom Grad d), falls

$$h = \sum_{i_1 + \dots + i_n = d} a_{i_1 \dots i_n} \cdot t_1^{i_1} \cdot \dots \cdot t_n^{i_n}.$$

- b) Für ein homogenes Polynom $h \in K[t_1, \dots, t_n]$ vom Grad d gilt:

$$h(\lambda t_1, \dots, \lambda t_n) = \lambda^d \cdot h(t_1, \dots, t_n) \quad \text{für alle } \lambda \in K.$$

- c) Ist K unendlich und $f \in K(t_1, \dots, t_n) \setminus \{0\}$, so folgt aus

$$f(\lambda t_1, \dots, \lambda t_n) = \lambda^d \cdot f(t_1, \dots, t_n) \quad \text{für alle } \lambda \in K,$$

dass f homogen vom Grad d ist.

- d) Ist h_1 homogen vom Grad d_1 und h_2 homogen vom Grad d_2 , so ist $h_1 \cdot h_2$ homogen vom Grad $d_1 + d_2$.

10. Sei K ein Körper und $K[t]$ der Polynomring in einer Unbestimmten.

- a) Zeigen Sie, dass in der Menge $K[t] \times (K[t] \setminus \{0\})$ durch

$$(g, h) \sim (g', h') \Leftrightarrow gh' = g'h$$

eine Äquivalenzrelation gegeben ist.

$K(t)$ sei die Menge der Äquivalenzklassen. Die zu (g, h) gehörige Äquivalenzklasse sei mit $\frac{g}{h}$ bezeichnet. Somit ist $\frac{g}{h} = \frac{g'}{h'} \Leftrightarrow gh' = g'h$.

- b) Zeigen Sie, dass in $K(t)$ die Verknüpfungen

$$\frac{g}{h} + \frac{g'}{h'} := \frac{gh' + hg'}{hh'}, \quad \frac{g}{h} \cdot \frac{g'}{h'} := \frac{gg'}{hg'},$$

wohl definiert sind (vgl. 1.2.7).

- c) Zeigen Sie schließlich, dass $K(t)$ mit diesen Verknüpfungen zu einem Körper wird. Man nennt $K(t)$ den Körper der rationalen Funktionen.

11. Was folgt aus der Vorzeichenregel von DESCARTES für das Polynom $t^n + 1$?

12. Man folgere die spezielle Vorzeichenregel aus der Vorzeichenregel von DESCARTES.

1.4 Vektorräume

In diesem ganzen Abschnitt bezeichnet K einen Körper. Wer die vorhergehenden Definitionen übersprungen hat, kann sich zunächst mit dem äußerst wichtigen Spezialfall $K = \mathbb{R}$ begnügen.

1.4.1. Bevor wir den allgemeinen Begriff des Vektorraums einführen, einige

Beispiele. a) Das Standardbeispiel ist der *Standardraum*

$$K^n = \{x = (x_1, \dots, x_n) : x_i \in K\}.$$

Mit Hilfe der Addition und Multiplikation in K erhält man zwei neue Verknüpfungen

$$\begin{aligned} \dot{+} : K^n \times K^n &\rightarrow K^n, & (x, y) &\mapsto x \dot{+} y, & \text{ und} \\ \cdot : K \times K^n &\rightarrow K^n, & (\lambda, x) &\mapsto \lambda \cdot x, \end{aligned}$$

durch

$$\begin{aligned} (x_1, \dots, x_n) \dot{+} (y_1, \dots, y_n) &:= (x_1 + y_1, \dots, x_n + y_n) \quad \text{und} \\ \lambda \cdot (x_1, \dots, x_n) &:= (\lambda x_1, \dots, \lambda x_n). \end{aligned}$$

Zur vorübergehenden Unterscheidung sind die neuen Verknüpfungen mit $\dot{+}$ und $\cdot$ bezeichnet. In K wird die Addition durch $+$ und die Multiplikation ohne Symbol ausgedrückt. Man beachte, daß nur $\dot{+}$ eine Verknüpfung im Sinn von 1.2.1 ist (solche Verknüpfungen nennt man manchmal auch *innere* im Gegensatz zur *äußeren* $\cdot$).

b) In der Menge $M(m \times n; K)$ der Matrizen mit m Zeilen, n Spalten und Einträgen aus K kann man addieren und mit Skalaren multiplizieren:

Ist $A = (a_{ij})$, $B = (b_{ij}) \in M(m \times n; K)$ und $\lambda \in K$, so sind

$$A \dot{+} B := (a_{ij} + b_{ij}) \quad \text{und} \quad \lambda \cdot A := (\lambda a_{ij}) \in M(m \times n; K).$$

Bei der Addition werden also die Einträge an den entsprechenden Stellen addiert, bei Multiplikation mit λ alle Einträge gleich multipliziert. Bis auf die andere Art, die Einträge aufzuschreiben, ist dieses Beispiel gleich $K^{m \cdot n}$ aus a).

c) Im Körper $\mathbb{C}$ der komplexen Zahlen kann man mit reellen Zahlen multiplizieren, das ergibt eine Abbildung

$$\mathbb{R} \times \mathbb{C} \rightarrow \mathbb{C}, \quad (\lambda, a + ib) \mapsto \lambda a + i\lambda b.$$

d) Im Polynomring $K[t]$ kann man neben Addition und Multiplikation von Polynomen eine weitere Multiplikation

$$\cdot : K \times K[t] \rightarrow K[t], \quad (\lambda, f) \mapsto \lambda \cdot f,$$

mit Elementen aus K erklären durch

$$\lambda \cdot (a_0 + a_1 t + \dots + a_n t^n) := \lambda a_0 + (\lambda a_1) t + \dots + (\lambda a_n) t^n.$$

e) Ist M eine beliebige Menge und

$$V = \{f: M \rightarrow K\} = \text{Abb}(M, K)$$

die Menge aller Abbildungen, so sind für $f, g \in V$ und $\lambda \in K$

$$f \dot{+} g \in V \quad \text{und} \quad \lambda \cdot f \in V$$

erklärt durch

$$(f \dot{+} g)(x) := f(x) + g(x) \quad \text{und} \quad (\lambda \cdot f)(x) := \lambda f(x).$$

Offensichtlich erhält man im Spezialfall $M = \{1, \dots, n\}$ wieder das Standardbeispiel a).

Als Extrakt aus diesen Beispielen führen wir nun den wichtigsten Begriff der linearen Algebra ein (zur Entstehung vgl. [Koe], Kap. 1, §2):

Definition. Sei K ein Körper. Eine Menge V zusammen mit einer inneren Verknüpfung

$$\dot{+}: V \times V \rightarrow V, \quad (v, w) \mapsto v \dot{+} w,$$

(Addition genannt) und einer äußeren Verknüpfung

$$\cdot: K \times V \rightarrow V, \quad (\lambda, v) \mapsto \lambda \cdot v,$$

(Multiplikation mit Skalaren oder skalare Multiplikation genannt) heißt K -Vektorraum (oder Vektorraum über K), wenn folgendes gilt:

V1 V zusammen mit der Addition ist eine abelsche Gruppe (das neutrale Element heißt Nullvektor, es wird mit 0 , und das Negative wird mit $-v$ bezeichnet).

V2 Die Multiplikation mit Skalaren muß in folgender Weise mit den anderen Verknüpfungen verträglich sein:

$$(\lambda + \mu) \cdot v = \lambda \cdot v \dot{+} \mu \cdot v, \quad \lambda \cdot (v \dot{+} w) = \lambda \cdot v \dot{+} \lambda \cdot w,$$

$$\lambda \cdot (\mu \cdot v) = (\lambda\mu) \cdot v, \quad 1 \cdot v = v,$$

für alle $\lambda, \mu \in K$ und $v, w \in V$.

Man beachte, daß, wie in Beispiel a) erläutert, die Verknüpfungen in K und V vorübergehend verschieden bezeichnet werden.

Durch Einsetzen der Definitionen und elementarste Rechnungen sieht man, daß in den obigen Beispielen a) bis e) die Vektorraumaxiome erfüllt sind. Dabei ist in Beispiel a) der Nullvektor gegeben durch $0 = (0, \dots, 0)$ und das Negative durch

$$-(x_1, \dots, x_n) = (-x_1, \dots, -x_n).$$

In Beispiel c) wird $\mathbb{C}$ zu einem $\mathbb{R}$ -Vektorraum.

Bemerkung. In einem K -Vektorraum V hat man die folgenden weiteren Rechenregeln:

a) $0 \cdot v = 0$.

b) $\lambda \cdot 0 = 0$.

c) $\lambda \cdot v = 0 \Rightarrow \lambda = 0 \text{ oder } v = 0$.

d) $(-1) \cdot v = -v$.

Beweis. a) $0 \cdot v = (0 + 0) \cdot v = 0 \cdot v + 0 \cdot v$.

b) $\lambda \cdot 0 = \lambda \cdot (0 + 0) = \lambda \cdot 0 + \lambda \cdot 0$.

c) Ist $\lambda \cdot v = 0$, aber $\lambda \neq 0$, so folgt

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

d) $v + (-1) \cdot v = 1 \cdot v + (-1) \cdot v = (1 - 1) \cdot v = 0 \cdot v = 0$. □

Die Axiome und die daraus abgeleiteten Regeln zeigen insbesondere, daß es völlig ungefährlich ist, wenn man die Verknüpfungen in K und V gleich bezeichnet und auch den Nullvektor 0 abmagert zu 0 . Das wollen wir ab sofort tun. Was gemeint ist, wird jeweils aus dem Zusammenhang klar werden.

1.4.2. In Kapitel 0 hatten wir homogene lineare Gleichungssysteme der Form $A \cdot x = 0$ mit reellen Koeffizienten betrachtet. Die Lösungen sind Teilmengen $W \subset \mathbb{R}^n$, ihre präzise Beschreibung ist unser Ziel. Schlüssel dafür ist die Beobachtung, daß W von $\mathbb{R}^n$ eine Vektorraumstruktur erbt. Allgemeiner ist die Frage, wann das für eine Teilmenge $W \subset V$ eines Vektorraumes der Fall ist.

Definition. Sei V ein K -Vektorraum und $W \subset V$ eine Teilmenge. W heißt *Untervektorraum von V* , falls folgendes gilt:

UV1 $W \neq \emptyset$.

UV2 $v, w \in W \Rightarrow v + w \in W$
(d.h. W ist abgeschlossen gegenüber der Addition).

UV3 $v \in W, \lambda \in K \Rightarrow \lambda v \in W$
(d.h. W ist abgeschlossen gegenüber der Multiplikation mit Skalaren).

Beispiele. a) In $V = \mathbb{R}^2$ betrachten wir die Teilmengen

$$\begin{aligned} W_1 &= \{0\}, \\ W_2 &= \{(x_1, x_2) \in \mathbb{R}^2 : a_1 x_1 + a_2 x_2 = b\}, \\ W_3 &= \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1\}, \\ W_4 &= \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \geq 0, x_2 \geq 0\}, \\ W_5 &= \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \cdot x_2 \geq 0\}. \end{aligned}$$

Der Leser prüfe die Bedingungen UV2 und UV3 nach. Nur für W_1 und für W_2 mit $b = 0$ sind beide erfüllt.

b) Ist A eine reelle $m \times n$ -Matrix, so ist die Lösungsmenge

$$W := \{x \in \mathbb{R}^n : Ax = 0\}$$

des zugehörigen homogenen linearen Gleichungssystems (vgl. 0.4.2) ein Untervektorraum des $\mathbb{R}^n$. Das rechnet man ganz leicht nach.

c) Im Vektorraum $V = \text{Abb}(\mathbb{R}, \mathbb{R})$ (vgl. 1.4.1, Beispiel e)) hat man die Untervektorräume

$$\mathbb{R}[t]_d \subset \mathbb{R}[t] \subset \mathcal{D}(\mathbb{R}, \mathbb{R}) \subset \mathcal{C}(\mathbb{R}, \mathbb{R}) \subset \text{Abb}(\mathbb{R}, \mathbb{R})$$

der Polynome vom Grad $\leq d$, aller Polynome, der differenzierbaren Funktionen und der stetigen Funktionen.

1.4.3. Aus den Eigenschaften UV2 und UV3 folgt, daß Addition und Multiplikation mit Skalaren von V auf W induziert werden.

Satz. Ein Untervektorraum $W \subset V$ ist zusammen mit der induzierten Addition und Multiplikation mit Skalaren wieder ein Vektorraum.

Mit Hilfe dieses Satzes kann man sich in vielen Fällen (etwa Beispiel b) und c) aus 1.4.2) den langweiligen Nachweis aller Vektorraumaxiome für W sparen, wenn man das ein für alle mal schon in einem größeren V getan hatte.

Beweis. Die Eigenschaften V2 sowie Kommutativ- und Assoziativgesetz der Addition gelten in W , da sie in V gelten. Der Nullvektor 0 liegt in W , da wegen UV1 ein $v \in W$ existiert, woraus $0 = 0 \cdot v \in W$ mit UV3 folgt. Zu jedem $v \in W$ ist wegen UV3 auch $-v = (-1) \cdot v \in W$. $\square$

Noch eine kleine Pflichtübung zur Erzeugung neuer Untervektorräume aus alten:

Lemma. Sei V ein Vektorraum, I eine beliebige Indexmenge, und für jedes $i \in I$ sei ein Untervektorraum W_i gegeben. Dann ist der Durchschnitt

$$W := \bigcap_{i \in I} W_i \subset V$$

wieder ein Untervektorraum.

Beweis. Da 0 in allen W_i enthalten ist, ist auch $0 \in W$, also $W \neq \emptyset$. Sind $v, w \in W$, so sind v, w in allen W_i enthalten. Da dann auch $v + w$ in allen W_i enthalten ist, ist $v + w$ in W enthalten. Ganz analog beweist man, daß mit $\lambda \in K$ und $v \in W$ auch $\lambda v \in W$ gilt. $\square$

Beispiel. Ist $V = \mathbb{R}[t]$ der Vektorraum aller reellen Polynome, $I = \mathbb{N}$ und $W_d := \mathbb{R}[t]_d$ der Untervektorraum der Polynome vom Grad $\leq d$, so ist

$$\bigcap_{d \in \mathbb{N}} W_d = W_0 = \mathbb{R}.$$

Vorsicht! Die Vereinigung von Untervektorräumen ist im allgemeinen kein Untervektorraum. Man mache sich das an Beispielen klar (etwa zwei Geraden im $\mathbb{R}^2$). Es gilt sogar die

Bemerkung. Sind $W, W' \subset V$ Untervektorräume derart, daß $W \cup W'$ Untervektorraum ist, so ist $W \subset W'$ oder $W' \subset W$.

Beweis. Angenommen, es ist $W \not\subset W'$. Dann ist $W' \subset W$ zu zeigen. Ist $w' \in W'$ und $w \in W \setminus W'$, so sind $w, w' \in W \cup W'$, also auch $w + w' \in W \cup W'$. $w + w'$ kann nicht Element von W' sein, denn sonst wäre $w = w + w' - w' \in W'$. Also ist $w + w' \in W$ und auch $w' = w + w' - w \in W$. $\square$

1.4.4. Eine Teilmenge eines Vektorraumes, die kein Untervektorraum ist, kann man zu einem solchen *abschließen*. Wie das geht, wird nun beschrieben.

Wir betrachten eine noch etwas allgemeinere Situation, nämlich einen K -Vektorraum V und eine Familie $(v_i)_{i \in I}$ von Vektoren $v_i \in V$ (vgl. 1.1.6). Ist $I = \{1, \dots, r\}$, so hat man Vektoren $v_1, \dots, v_r$. Ein $v \in V$ heißt *Linearkombination* von $v_1, \dots, v_r$, wenn es $\lambda_1, \dots, \lambda_r \in K$ gibt, so daß

$$v = \lambda_1 v_1 + \dots + \lambda_r v_r.$$

Für allgemeines I definiert man

$$\text{span}_K(v_i)_{i \in I}$$

als die Menge aller $v \in V$, die sich aus einer (von v abhängigen) endlichen Teilfamilie von $(v_i)_{i \in I}$ linear kombinieren lassen. Um das präzise aufzuschreiben

zu können, benötigt man Doppelindizes: Zu $v \in V$ muß es eine Zahl $r \in \mathbb{N}$ sowie Indizes $i_1, \dots, i_r \in I$ und Skalare $\lambda_1, \dots, \lambda_r$ geben, so daß

$$v = \lambda_1 v_{i_1} + \dots + \lambda_r v_{i_r}.$$

Man nennt $\text{span}_K(v_i)_{i \in I}$ den von der Familie *aufgespannten* (oder *erzeugten*) Raum. Ist $I = \emptyset$, so setzt man

$$\text{span}_K(v_i)_{i \in \emptyset} := \{0\}.$$

Für eine endliche Familie $(v_1, \dots, v_r)$ verwendet man oft die suggestivere Notation

$$\begin{aligned} K v_1 + \dots + K v_r &:= \text{span}_K(v_1, \dots, v_r) \\ &= \{v \in V : \text{es gibt } \lambda_1, \dots, \lambda_r \in K \text{ mit } v = \lambda_1 v_1 + \dots + \lambda_r v_r\}. \end{aligned}$$

Falls klar ist, welcher Körper gemeint ist, schreibt man nur span statt span_K .

Bemerkung. Sei V ein K -Vektorraum und $(v_i)_{i \in I}$ eine Familie von Elementen aus V . Dann gilt:

- a) $\text{span}(v_i) \subset V$ ist Untervektorraum.
- b) Ist $W \subset V$ Untervektorraum, und gilt $v_i \in W$ für alle $i \in I$, so ist $\text{span}(v_i) \subset W$.

Kurz ausgedrückt: $\text{span}(v_i)$ ist der *kleinste* Untervektorraum von V , der alle v_i enthält.

Beweis. a) ist ganz klar. Sind alle v_i in W enthalten, so sind auch alle endlichen Linearkombinationen aus den v_i in W enthalten, denn W ist Untervektorraum. Daraus folgt b). $\square$

Ist $M \subset V$ eine Teilmenge, so ist entsprechend $\text{span}(M)$ erklärt als die Menge aller endlichen Linearkombinationen von Vektoren aus M , und das ist der kleinste Untervektorraum mit

$$M \subset \text{span}(M) \subset V.$$

Es mag auf den ersten Blick nicht recht einleuchten, warum man bei der Erzeugung eines Untervektorraumes allgemeiner von einer Familie von Vektoren ausgeht. Das hat den Vorteil, daß es bei einer Familie (im Gegensatz zu einer Menge) sinnvoll ist, wenn man sagt „ein Vektor kommt mehrfach vor“. Ist $I = \{1, \dots, n\}$, so haben die Vektoren der Familie außerdem eine natürliche Reihenfolge.

Beispiele. a) Sei $V = \mathbb{R}^3$. Sind $v_1, v_2 \in \mathbb{R}^3$, so ist $\text{span}(v_1)$ die Gerade durch 0 und v_1 , wenn $v_1 \neq 0$. $\text{span}(v_1, v_2)$ ist die Ebene durch 0, v_1 und v_2 , falls $v_2 \notin \text{span}(v_1)$.

b) Im K^n mit $I = \{1, \dots, n\}$ erklären wir für $i \in I$

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

wobei die 1 an der i -ten Stelle steht. Dann ist $\text{span}(e_i)_{i \in I} = K^n$.

c) Ist $V = K[t]$ der Polynomring, $I = \mathbb{N}$ und $v_n = t^n$, so ist

$$\text{span}(v_n)_{n \in \mathbb{N}} = K[t].$$

d) Betrachten wir in Beispiel a) aus 1.4.2 die W_j mit $j = 3, 4, 5$, so ist $\text{span}(W_j) = \mathbb{R}^2$.

1.4.5. Ein Untervektorraum kann von sehr vielen verschiedenen Familien erzeugt werden, und das kann mit unterschiedlicher Effizienz geschehen. Dazu betrachten wir die Beispiele aus 1.4.4. Bei b) ist die Situation optimal, denn es folgt für $x = (x_1, \dots, x_n) \in K^n$ und

$$x = \lambda_1 e_1 + \dots + \lambda_n e_n, \quad \text{daß} \quad \lambda_1 = x_1, \dots, \lambda_n = x_n.$$

Die Linearkombination ist also für jedes x eindeutig bestimmt, entsprechend in c). In den Beispielen aus d) hat man für jedes $x \in \mathbb{R}^2$ jeweils unendlich viele Möglichkeiten, es linear aus Elementen von W_j zu kombinieren.

Die Eindeutigkeit ist besonders einfach beim Nullvektor zu überprüfen. Bei beliebigen $v_1, \dots, v_r$ hat man die *triviale Linearkombination*

$$0 = 0v_1 + \dots + 0v_r.$$

Gibt es eine andere Linearkombination des Nullvektors, so ist die Eindeutigkeit der Darstellung verletzt. Das motiviert die

Definition. Sei V ein K -Vektorraum. Eine endliche Familie $(v_1, \dots, v_r)$ von Vektoren aus V heißt *linear unabhängig*, falls gilt: Sind $\lambda_1, \dots, \lambda_r \in K$ und ist

$$\lambda_1 v_1 + \dots + \lambda_r v_r = 0,$$

so folgt

$$\lambda_1 = \dots = \lambda_r = 0.$$

Anders ausgedrückt bedeutet das, daß sich der Nullvektor nur trivial aus den $v_1, \dots, v_r$ linear kombinieren läßt.

Eine beliebige Familie $(v_i)_{i \in I}$ von Vektoren aus V heißt *linear unabhängig*, falls jede endliche Teilfamilie linear unabhängig ist.

Die Familie $(v_i)_{i \in I}$ heißt *linear abhängig*, falls sie nicht linear unabhängig ist, d.h. falls es eine endliche Teilfamilie $(v_{i_1}, \dots, v_{i_r})$ und $\lambda_1, \dots, \lambda_r \in K$ gibt, die nicht alle gleich Null sind, so daß

$$\lambda_1 v_{i_1} + \dots + \lambda_r v_{i_r} = 0.$$

Zur Bequemlichkeit sagt man meist anstatt „die Familie $(v_1, \dots, v_r)$ von Vektoren aus V ist linear (un-)abhängig“ einfacher „die Vektoren $v_1, \dots, v_r \in V$ sind linear (un-)abhängig“.

Es ist vorteilhaft, auch die leere Familie, die den Nullvektorraum aufspannt, linear unabhängig zu nennen.

Die Definition der linearen Unabhängigkeit ist grundlegend für die ganze lineare Algebra, aber man muß sich etwas daran gewöhnen. Was sie geometrisch bedeutet, sieht man sehr gut an der Bedingung aus Aufgabe 1 zu 0.3 für zwei Vektoren im $\mathbb{R}^n$.

In der Definition der linearen Unabhängigkeit spielt der Nullvektor scheinbar eine besondere Rolle. Daß dem nicht so ist, zeigt das

Lemma. Für eine Familie $(v_i)_{i \in I}$ von Vektoren eines K -Vektorraumes sind folgende Bedingungen äquivalent:

- i) (v_i) ist linear unabhängig.
- ii) Jeder Vektor $v \in \text{span}(v_i)$ läßt sich in eindeutiger Weise aus Vektoren der Familie (v_i) linear kombinieren.

Beweis. ii) $\Rightarrow$ i) ist klar, denn bei einer linear abhängigen Familie hat der Nullvektor verschiedene Darstellungen.

i) $\Rightarrow$ ii): Sei ein $v \in \text{span}(v_i)$ auf zwei Arten linear kombiniert, also

$$v = \sum_{i \in I} \lambda_i v_i = \sum_{i \in I} \mu_i v_i, \quad (*)$$

wobei in beiden Summen jeweils nur endlich viele der Skalare λ_i und μ_i von Null verschieden sind. Es gibt also eine endliche Teilmenge $J \subset I$, so daß für jedes $\lambda_i \neq 0$ oder $\mu_i \neq 0$ der Index in J enthalten ist. Aus (*) folgt dann

$$\sum_{i \in I} (\lambda_i - \mu_i) v_i = 0,$$

und wegen der vorausgesetzten linearen Unabhängigkeit folgt $\lambda_i = \mu_i$ für alle $i \in J$ und somit auch für alle $i \in I$, da ja die restlichen λ_i und μ_i ohnehin Null waren. Damit ist die Eindeutigkeit der Linearkombination bewiesen. $\square$

Beispiele. a) Im K^n sind die Vektoren $e_1, \dots, e_n$ linear unabhängig.

b) Ist $A = (a_{ij}) \in M(m \times n; K)$ eine Matrix in Zeilenstufenform (vgl. 0.4.3), so sind die ersten r Zeilen $v_1, \dots, v_r$ von A linear unabhängig. Ist nämlich

$$A = \begin{pmatrix} & a_{1j_1} & & & \\ & & a_{2j_2} & & \\ & & & \ddots & \\ & 0 & & & a_{rj_r} \end{pmatrix},$$

so folgt aus $\lambda_1 v_1 + \dots + \lambda_r v_r = 0$ zunächst $\lambda_1 a_{1j_1} = 0$, also $\lambda_1 = 0$ wegen $a_{1j_1} \neq 0$. Im zweiten Schritt folgt daraus analog $\lambda_2 = 0$ und weiter so $\lambda_3 = \dots = \lambda_r = 0$. Analog zeigt man, daß die Spalten von A mit den Indizes $j_1, j_2, \dots, j_r$ linear unabhängig sind.

c) Im Polynomring $K[t]$ ist die Familie $(t^n)_{n \in \mathbb{N}}$ linear unabhängig.

Weitere Beispiele finden sich in den Aufgaben 8 und 9. Noch ein paar weitere Kleinigkeiten:

Bemerkung. In jedem K -Vektorraum V gilt:

- a) Ein einziger Vektor $v \in V$ ist genau dann linear unabhängig, wenn $v \neq 0$.
- b) Gehört der Nullvektor zu einer Familie, so ist sie linear abhängig.
- c) Kommt der gleiche Vektor in einer Familie mehrmals vor, so ist sie linear abhängig.
- d) Ist $r \geq 2$, so sind die Vektoren $v_1, \dots, v_r$ genau dann linear abhängig, wenn einer davon Linearkombination der anderen ist.

Diese letzte Charakterisierung ist plausibler als die Definition, aber formal nicht so bequem zu handhaben.

Beweis. a) Ist v linear abhängig, so gibt es ein $\lambda \in K^*$ mit $\lambda v = 0$, also ist $v = 0$ nach Bemerkung c) in 1.4.1. Umgekehrt ist 0 linear abhängig, da $1 \cdot 0 = 0$.

b) $1 \cdot 0 = 0$.

c) Gibt es $i_1, i_2 \in I$ mit $i_1 \neq i_2$ aber $v_{i_1} = v_{i_2}$, so ist $1 \cdot v_{i_1} + (-1)v_{i_2} = 0$.

d) Sind die Vektoren $v_1, \dots, v_r$ linear abhängig, so gibt es $\lambda_1, \dots, \lambda_r \in K$ mit $\lambda_1 v_1 + \dots + \lambda_r v_r = 0$ und ein $k \in \{1, \dots, r\}$ mit $\lambda_k \neq 0$. Dann ist

$$v_k = -\frac{\lambda_1}{\lambda_k} v_1 - \dots - \frac{\lambda_{k-1}}{\lambda_k} v_{k-1} - \frac{\lambda_{k+1}}{\lambda_k} v_{k+1} - \dots - \frac{\lambda_r}{\lambda_k} v_r.$$

Ist umgekehrt $v_k = \mu_1 v_1 + \dots + \mu_{k-1} v_{k-1} + \mu_{k+1} v_{k+1} + \dots + \mu_r v_r$, so ist

$$\mu_1 v_1 + \dots + \mu_{k-1} v_{k-1} + (-1) v_k + \mu_{k+1} v_{k+1} + \dots + \mu_r v_r = 0. \quad \square$$

Aufgaben zu 1.4

1. Welche der folgenden Mengen sind Untervektorräume der angegebenen Vektorräume?

- $\{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 = x_2 = 2x_3\} \subset \mathbb{R}^3$.
- $\{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 = 0\} \subset \mathbb{R}^2$.
- $\{(\mu + \lambda, \lambda^2) \in \mathbb{R}^2 : \mu, \lambda \in \mathbb{R}\} \subset \mathbb{R}^2$.
- $\{f \in \text{Abb}(\mathbb{R}, \mathbb{R}) : f(x) = f(-x) \text{ für alle } x \in \mathbb{R}\} \subset \text{Abb}(\mathbb{R}, \mathbb{R})$.
- $\{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 \geq x_2\} \subset \mathbb{R}^3$.
- $\{A \in M(m \times n; \mathbb{R}) : A \text{ ist in Zeilenstufenform}\} \subset M(m \times n; \mathbb{R})$.

2. Seien V und W zwei K -Vektorräume. Zeigen Sie, dass das direkte Produkt $V \times W$ durch die Verknüpfungen

$$(v, w) + (v', w') := (v + v', w + w'), \quad \lambda \cdot (v, w) := (\lambda v, \lambda w),$$

ebenfalls zu einem K -Vektorraum wird.

3. Ist X eine nichtleere Menge, V ein K -Vektorraum und $\text{Abb}(X, V)$ die Menge aller Abbildungen von X nach V , so ist auf $\text{Abb}(X, V)$ durch

$$(f + g)(x) := f(x) + g(x), \quad (\lambda \cdot f)(x) := \lambda f(x),$$

eine Addition und eine skalare Multiplikation erklärt.

Zeigen Sie, dass $\text{Abb}(X, V)$ mit diesen Verknüpfungen zu einem K -Vektorraum wird.

4. Eine Abbildung $f: \mathbb{R} \rightarrow \mathbb{R}$ heißt 2π -periodisch, falls $f(x) = f(x + 2\pi)$ für alle $x \in \mathbb{R}$.

- Zeigen Sie, dass $V = \{f \in \text{Abb}(\mathbb{R}, \mathbb{R}) : f \text{ ist } 2\pi\text{-periodisch}\} \subset \text{Abb}(\mathbb{R}, \mathbb{R})$ ein Untervektorraum ist.
- Zeigen Sie, dass $W = \text{span}(\cos nx, \sin mx)_{n,m \in \mathbb{N}}$ ein Untervektorraum von V ist. (Man nennt W den Vektorraum der *trigonometrischen Polynome*.)

5. Seien

$$\begin{aligned}\ell^1 &:= \left\{ (x_i)_{i \in \mathbb{N}} : \sum_{i=0}^{\infty} |x_i| < \infty \right\} \subset \text{Abb}(\mathbb{N}, \mathbb{R}), \\ \ell^2 &:= \left\{ (x_i)_{i \in \mathbb{N}} : \sum_{i=0}^{\infty} |x_i|^2 < \infty \right\} \subset \text{Abb}(\mathbb{N}, \mathbb{R}), \\ \ell &:= \{ (x_i)_{i \in \mathbb{N}} : (x_i)_{i \in \mathbb{N}} \text{ konvergiert} \} \subset \text{Abb}(\mathbb{N}, \mathbb{R}), \\ \ell_{\infty} &:= \{ (x_i)_{i \in \mathbb{N}} : (x_i)_{i \in \mathbb{N}} \text{ beschränkt} \} \subset \text{Abb}(\mathbb{N}, \mathbb{R}).\end{aligned}$$

Zeigen Sie, dass $\ell^1 \subset \ell^2 \subset \ell \subset \ell_{\infty} \subset \text{Abb}(\mathbb{N}, \mathbb{R})$ eine aufsteigende Kette von Untervektorräumen ist.

6. Kann eine abzählbar unendliche Menge M eine $\mathbb{R}$ -Vektorraumstruktur besitzen?

7. Gibt es eine $\mathbb{C}$ -Vektorraumstruktur auf $\mathbb{R}$, so dass die skalare Multiplikation $\mathbb{C} \times \mathbb{R} \rightarrow \mathbb{R}$ eingeschränkt auf $\mathbb{R} \times \mathbb{R}$ die übliche Multiplikation reeller Zahlen ist?

8. Sind die folgenden Vektoren linear unabhängig?

- a) $1, \sqrt{2}, \sqrt{3}$ im $\mathbb{Q}$ -Vektorraum $\mathbb{R}$.
- b) $(1, 2, 3), (4, 5, 6), (7, 8, 9)$ im $\mathbb{R}^3$.
- c) $\left(\frac{1}{n+x}\right)_{n \in \mathbb{N}}$ in $\text{Abb}(\mathbb{R}_+^*, \mathbb{R})$.
- d) $(\cos nx, \sin mx)_{n, m \in \mathbb{N} \setminus \{0\}}$ in $\text{Abb}(\mathbb{R}, \mathbb{R})$.

9. Für welche $t \in \mathbb{R}$ sind die folgenden Vektoren aus $\mathbb{R}^3$ linear abhängig?

$$(1, 3, 4), \quad (3, t, 11), \quad (-1, -4, 0).$$

10. Stellen Sie den Vektor w jeweils als Linearkombination der Vektoren u_1, v_2, v_3 dar:

- a) $w = (6, 2, 1), v_1 = (1, 0, 1), v_2 = (7, 3, 1), v_3 = (2, 5, 8)$.
- b) $w = (2, 1, 1), v_1 = (1, 5, 1), v_2 = (0, 9, 1), v_3 = (3, -3, 1)$.

1.5 Basis und Dimension

Nun muß die erste ernsthafte technische Schwierigkeit überwunden werden, um einem Vektorraum eindeutig eine Zahl (genannt Dimension) als Maß für seine Größe zuzuordnen zu können.

1.5.1. Zunächst die wichtigste

Definition. Eine Familie $\mathcal{B} = (v_i)_{i \in I}$ in einem Vektorraum V heißt *Erzeugendensystem* von V , wenn

$$V = \text{span}(v_i)_{i \in I},$$

d.h. wenn jedes $v \in V$ Linearkombination von endlich vielen v_i ist.

Eine Familie $\mathcal{B} = (v_i)_{i \in I}$ in V heißt *Basis* von V , wenn sie ein linear unabhängiges Erzeugendensystem ist.

V heißt *endlich erzeugt*, falls es ein endliches Erzeugendensystem (d.h. eine endliche Familie $\mathcal{B} = (v_1, \dots, v_n)$ mit $V = \text{span}(v_i)$ gibt. Ist $\mathcal{B}$ eine endliche Basis, so nennt man die Zahl n die *Länge der Basis*.

Beispiele. a) Die Begriffe sind so erklärt, daß die leere Familie eine Basis des Nullvektorraumes ist. Diese kleine Freude kann man ihr gönnen.

b) $\mathcal{K} := (e_1, \dots, e_n)$ ist eine Basis des K^n , sie heißt die *kanonische Basis* oder *Standardbasis* (vgl. Beispiel b) in 1.4.4).

c) Im Vektorraum $M(m \times n; K)$ hat man die Matrizen

$$E_i^j = \begin{pmatrix} & & & 0 & & & \\ & & & \vdots & & & \\ & & & 0 & & & \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ & & & 0 & & & \\ & & & \vdots & & & \\ & & & 0 & & & \end{pmatrix}$$

mit einer Eins in der i -ten Zeile und j -ten Spalte und sonst Nullen. Diese $m \times n$ -Matrizen bilden eine Basis von $M(m \times n; K)$. Das ist wieder nur eine Variante von Beispiel b).

d) $(1, i)$ ist eine Basis des $\mathbb{R}$ -Vektorraumes $\mathbb{C}$.

e) $(1, t, t^2, \dots)$ ist eine Basis unendlicher Länge des Polynomrings $K[t]$.

1.5.2. Das sieht alles recht einfach aus, bis auf eine zunächst spitzfindig erscheinende Frage: wenn man im K^n neben der Standardbasis irgendeine andere Basis findet, ist es gar nicht klar, daß sie die gleiche Länge hat. Im K^n wäre das noch zu verschmerzen, aber schon bei Untervektorräumen $W \subset K^n$ gibt es keine Standardbasis mehr. Es ist nicht einmal ohne weiteres klar, daß jedes solche W endlich erzeugt ist (Korollar 3 in 1.5.5). Daher kann man es nicht umgehen, die Längen verschiedener Basen zu vergleichen. Bevor wir das in Angriff nehmen, noch einige oft benutzte Varianten der Definition einer Basis. Zur Vereinfachung der Bezeichnungen betrachten wir dabei nur endliche Familien.

Satz. Für eine Familie $\mathcal{B} = (v_1, \dots, v_n)$ von Vektoren eines K -Vektorraumes $V \neq \{0\}$ sind folgende Bedingungen gleichwertig:

- i) $\mathcal{B}$ ist eine Basis, d.h. ein linear unabhängiges Erzeugendensystem.
- ii) $\mathcal{B}$ ist ein „unverkürzbares“ Erzeugendensystem, d.h.

$$(v_1, \dots, v_{r-1}, v_{r+1}, \dots, v_n)$$

ist für jedes $r \in \{1, \dots, n\}$ kein Erzeugendensystem mehr.

- iii) Zu jedem $v \in V$ gibt es eindeutig bestimmte $\lambda_1, \dots, \lambda_n \in K$ mit

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n,$$

d.h. $\mathcal{B}$ ist ein Erzeugendensystem mit der zusätzlichen Eindeigkeitseigenschaft.

- iv) $\mathcal{B}$ ist „unverlängerbar“ linear unabhängig, d.h. $\mathcal{B}$ ist linear unabhängig, und für jedes $v \in V$ wird die Familie $(v_1, \dots, v_n, v)$ linear abhängig.

Beweis. i) $\Rightarrow$ ii). Gegeben sei ein Erzeugendensystem $\mathcal{B}$. Ist $\mathcal{B}$ verkürzbar, also zur Vereinfachung der Notation mit $r = 1$

$$v_1 = \lambda_2 v_2 + \dots + \lambda_n v_n, \quad \text{so folgt} \quad (-1)v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n = 0.$$

Also ist $\mathcal{B}$ linear abhängig.

ii) $\Rightarrow$ iii). Sei wieder $\mathcal{B}$ ein Erzeugendensystem. Ist die Eindeigkeitseigenschaft verletzt, so gibt es ein $v \in V$ mit

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n = \mu_1 v_1 + \dots + \mu_n v_n,$$

und o.B.d.A. $\lambda_1 \neq \mu_1$. Subtraktion der Linearkombinationen und Division durch $\lambda_1 - \mu_1$ ergibt

$$v_1 = \frac{\mu_2 - \lambda_2}{\lambda_1 - \mu_1} v_2 + \dots + \frac{\mu_n - \lambda_n}{\lambda_1 - \mu_1} v_n,$$

also ist $\mathcal{B}$ verkürzbar.

iii) $\Rightarrow$ iv). Aus iii) folgt, daß $\mathcal{B}$ linear unabhängig ist. (Lemma in 1.4.5). Ist $v \in V$, so ist

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n, \quad \text{also} \quad \lambda_1 v_1 + \dots + \lambda_n v_n + (-1)v = 0,$$

d.h. $(v_1, \dots, v_n, v)$ ist linear abhängig.

iv) $\Rightarrow$ i). Sei $\mathcal{B}$ unverlängerbar linear unabhängig. Für jedes $v \in V$ gibt es $\lambda_1, \dots, \lambda_n, \lambda \in K$ mit

$$\lambda_1 v_1 + \dots + \lambda_n v_n + \lambda v = 0.$$

Da $\mathcal{B}$ linear unabhängig ist, muß $\lambda \neq 0$ sein, also ist

$$v = -\frac{\lambda_1}{\lambda} v_1 - \dots - \frac{\lambda_n}{\lambda} v_n,$$

und es ist bewiesen, daß $\mathcal{B}$ ein Erzeugendensystem ist. $\square$

Der Beweis von iv) $\Rightarrow$ i) ergibt den

Zusatz. Ist V nicht endlich erzeugt, so gibt es eine unendliche linear unabhängige Familie.

Beweis. Es genügt zu zeigen, daß es für beliebiges n zu linear unabhängigen Vektoren $v_1, \dots, v_n$ einen weiteren Vektor v gibt, so daß auch $(v_1, \dots, v_n, v)$ linear unabhängig ist. Wäre $(v_1, \dots, v_n, v)$ für jedes $v \in V$ linear abhängig, so wäre nach obigem Argument $(v_1, \dots, v_n)$ ein Erzeugendensystem, was der Voraussetzung widerspricht. $\square$

1.5.3. Die Bedeutung des obigen Satzes erkennt man schon an seinem gar nicht selbstverständlichen Korollar, dem

Basisauswahlsatz. Aus jedem endlichen Erzeugendensystem eines Vektorraumes kann man eine Basis auswählen. Insbesondere hat jeder endlich erzeugte Vektorraum eine endliche Basis.

Beweis. Von dem gegebenen Erzeugendensystem nehme man so lange einzelne Vektoren weg, bis es unverkürzbar geworden ist. Da am Anfang nur endlich viele da waren, führt das Verfahren zum Ziel. $\square$

Allgemeiner gilt das

Theorem. Jeder Vektorraum besitzt eine Basis.

Der Beweis ist wesentlich schwieriger, wenn es kein endliches Erzeugendensystem gibt, weil man möglicherweise unendlich viele Vektoren weglassen muß, bis die Unverkürzbarkeit erreicht ist. Ein Beweis des Theorems erfordert Hilfsmittel aus der Mengenlehre, etwa das ZORNsche Lemma. Darauf gehen wir hier nicht ein (vgl. etwa [B1], p.261).

Im Falle nicht endlich erzeugter Vektorräume sind Basen im hier definierten Sinn von geringer Bedeutung. Hier ist es für die Anwendungen in der Analysis wichtiger, konvergente unendliche Linearkombinationen zu untersuchen. Damit beschäftigt sich die Funktionalanalysis (vgl. etwa [M-V]).

1.5.4. Um die Längen verschiedener Basen zu vergleichen, muß man systematisch Vektoren austauschen. Dieses Verfahren entwickelte E. STEINITZ im Jahre 1910 bei einem ähnlichen Problem in der Theorie der Körpererweiterungen. Ein einzelner Schritt des Verfahrens wird geregelt durch das

Austauschlemma. *Gegeben sei ein K -Vektorraum V mit der Basis*

$$\mathcal{B} = (v_1, \dots, v_r) \quad \text{und} \quad w = \lambda_1 v_1 + \dots + \lambda_r v_r \in V.$$

Ist $k \in \{1, \dots, r\}$ mit $\lambda_k \neq 0$, so ist

$$\mathcal{B}' := (v_1, \dots, v_{k-1}, w, v_{k+1}, \dots, v_r)$$

wieder eine Basis von V . Man kann also v_k gegen w austauschen.

Beweis. Zur Vereinfachung der Schreibweise können wir annehmen, daß $k = 1$ ist (durch Umnummerierung kann man das erreichen). Es ist also zu zeigen, daß $\mathcal{B}' = (w, v_2, \dots, v_r)$ eine Basis von V ist. Ist $v \in V$, so ist

$$v = \mu_1 v_1 + \dots + \mu_r v_r$$

mit $\mu_1, \dots, \mu_r \in K$. Wegen $\lambda_1 \neq 0$ ist

$$v_1 = \frac{1}{\lambda_1} w - \frac{\lambda_2}{\lambda_1} v_2 - \dots - \frac{\lambda_r}{\lambda_1} v_r \quad \text{also}$$

$$v = \frac{\mu_1}{\lambda_1} w + \left(\mu_2 - \frac{\mu_1 \lambda_2}{\lambda_1} \right) v_2 + \dots + \left(\mu_r - \frac{\mu_1 \lambda_r}{\lambda_1} \right) v_r,$$

womit gezeigt ist, daß $\mathcal{B}'$ ein Erzeugendensystem ist.

Zum Nachweis der linearen Unabhängigkeit von $\mathcal{B}'$ sei

$$\mu w + \mu_2 v_2 + \dots + \mu_r v_r = 0,$$

wobei $\mu, \mu_2, \dots, \mu_r \in K$. Setzt man $w = \lambda_1 v_1 + \dots + \lambda_r v_r$ ein, so ergibt sich

$$\mu \lambda_1 v_1 + (\mu \lambda_2 + \mu_2) v_2 + \dots + (\mu \lambda_r + \mu_r) v_r = 0,$$

also $\mu \lambda_1 = \mu \lambda_2 + \mu_2 = \dots = \mu \lambda_r + \mu_r = 0$, da $\mathcal{B}$ linear unabhängig war. Wegen $\lambda_1 \neq 0$ folgt $\mu = 0$ und damit $\mu_2 = \dots = \mu_r = 0$. $\square$

Durch Iteration erhält man den

Austauschsatz. In einem K -Vektorraum V seien eine Basis

$$\mathcal{B} = (v_1, \dots, v_r)$$

und eine linear unabhängige Familie $(w_1, \dots, w_n)$ gegeben. Dann ist $n \leq r$, und es gibt Indizes $i_1, \dots, i_n \in \{1, \dots, r\}$ derart, daß man nach Austausch von v_{i_1} gegen $w_1, \dots, v_{i_n}$ gegen w_n wieder eine Basis von V erhält. Numeriert man so um, daß $i_1 = 1, \dots, i_n = n$ ist, so bedeutet das, daß

$$\mathcal{B}^* = (w_1, \dots, w_n, v_{n+1}, \dots, v_r)$$

eine Basis von V ist.

Vorsicht! Die Ungleichung $n \leq r$ wird nicht vorausgesetzt, sondern gefolgert.

Beweis durch Induktion nach n . Für $n = 0$ ist nichts zu beweisen. Sei also $n \geq 1$, und sei der Satz schon für $n - 1$ bewiesen (Induktionsannahme).

Da auch $(w_1, \dots, w_{n-1})$ linear unabhängig ist, ergibt die Induktionsannahme, daß (bei geeigneter Numerierung) $(w_1, \dots, w_{n-1}, v_n, \dots, v_r)$ eine Basis von V ist. Da nach Induktionsannahme $n - 1 \leq r$ gilt, muß zum Nachweis von $n \leq r$ nur noch der Fall $n - 1 = r$ ausgeschlossen werden. Dann wäre aber $(w_1, \dots, w_{n-1})$ schon eine Basis von V , was Aussage iv) aus Satz 1.5.2 widerspricht. Wir schreiben

$$w_n = \lambda_1 w_1 + \dots + \lambda_{n-1} w_{n-1} + \lambda_n v_n + \dots + \lambda_r v_r$$

mit $\lambda_1, \dots, \lambda_r \in K$. Wäre $\lambda_n = \dots = \lambda_r = 0$, so hätte man einen Widerspruch zur linearen Unabhängigkeit von $w_1, \dots, w_n$. Bei erneuter geeigneter Numerierung können wir also $\lambda_n \neq 0$ annehmen, und wie wir im Austauschlemma gesehen haben, läßt sich daher v_n gegen w_n austauschen. Also ist $\mathcal{B}^*$ eine Basis von V . $\square$

1.5.5. Nach Überwindung dieser kleinen technischen Schwierigkeiten läuft die Theorie wieder wie von selbst. Wir notieren die wichtigsten Folgerungen.

Korollar 1. Hat ein K -Vektorraum V eine endliche Basis, so ist jede Basis von V endlich.

Beweis. Sei $(v_1, \dots, v_r)$ eine endliche Basis und $(w_i)_{i \in I}$ eine beliebige Basis von V . Wäre I unendlich, so gäbe es $i_1, \dots, i_{r+1} \in I$ derart, daß $w_{i_1}, \dots, w_{i_{r+1}}$ linear unabhängig wären. Das widerspricht aber dem Austauschsatz. $\square$

Korollar 2. Je zwei endliche Basen eines K -Vektorraumes haben gleiche Länge.

Beweis. Sind $(v_1, \dots, v_r)$ und $(w_1, \dots, w_k)$ zwei Basen, so kann man den Austauschsatz zweimal anwenden, was $k \leq r$ und $r \leq k$, also $r = k$ ergibt. $\square$

Mit Hilfe dieser Ergebnisse können wir nun in sinnvoller Weise die Dimension eines Vektorraumes erklären.

Definition. Ist V ein K -Vektorraum, so definieren wir

$$\dim_K V := \begin{cases} \infty, & \text{falls } V \text{ keine endliche Basis besitzt,} \\ r, & \text{falls } V \text{ eine Basis der Länge } r \text{ besitzt.} \end{cases}$$

$\dim_K V$ heißt die *Dimension* von V über K . Falls klar ist, welcher Körper gemeint ist, schreibt man auch $\dim V$.

Korollar 3. Ist $W \subset V$ Untervektorraum eines endlich erzeugten Vektorraumes V , so ist auch W endlich erzeugt, und es gilt $\dim W \leq \dim V$.

Aus $\dim W = \dim V$ folgt $W = V$.

Beweis. Wäre W nicht endlich erzeugt, so gäbe es nach dem Zusatz aus 1.5.2 eine unendliche linear unabhängige Familie, was dem Austauschsatz widerspricht. Also hat W eine endliche Basis, und wieder nach dem Austauschsatz ist ihre Länge höchstens gleich $\dim V$.

Sei $n = \dim W = \dim V$ und $w_1, \dots, w_n$ Basis von W . Ist $W \neq V$, so gibt es ein $v \in V \setminus W$ und $w_1, \dots, w_n, v$ sind linear unabhängig im Widerspruch zum Austauschsatz. $\square$

In 1.5.3 hatten wir gesehen, daß man aus einem endlichen Erzeugendensystem eine Basis auswählen kann. Manchmal ist die Konstruktion „aus der anderen Richtung“ wichtig:

Basisergänzungssatz. In einem endlich erzeugten Vektorraum V seien linear unabhängige Vektoren $w_1, \dots, w_n$ gegeben. Dann kann man $w_{n+1}, \dots, w_r$ finden, so daß

$$\mathcal{B} = (w_1, \dots, w_n, w_{n+1}, \dots, w_r)$$

eine Basis von V ist.

Beweis. Sei $(v_1, \dots, v_m)$ ein Erzeugendensystem. Nach 1.5.3 kann man daraus eine Basis auswählen, etwa $(v_1, \dots, v_r)$ mit $r \leq m$. Nun wendet man den Austauschsatz an und sieht, daß bei geeigneter Numerierung durch

$$w_{n+1} := v_{n+1}, \dots, w_r := v_r$$

die gesuchte Ergänzung gefunden ist. $\square$

Beispiele. a) $\dim K^n = n$, denn K^n hat die kanonische Basis $(e_1, \dots, e_n)$. Nach Korollar 2 hat auch jede andere Basis von K^n die Länge n , was gar nicht selbstverständlich ist.

b) Geraden (bzw. Ebenen) durch den Nullpunkt des $\mathbb{R}^n$ sind Untervektorräume der Dimension 1 (bzw. 2).

- c) Für den Polynomring gilt $\dim_K K[t] = \infty$.
 d) $\dim_{\mathbb{R}} \mathbb{C} = 2$ denn 1 und i bilden eine Basis. Dagegen ist $\dim_{\mathbb{C}} \mathbb{C} = 1$.
 e) $\dim_{\mathbb{Q}} \mathbb{R} = \infty$ (vgl. Aufgabe 4).

1.5.6. Bei der Definition eines Vektorraumes in 1.4.1 hatten wir einen Körper K zugrundegelegt. Zur Formulierung der Axiome genügt ein kommutativer Ring R mit Einselement, man spricht dann von einem *Modul über R* . Die Begriffe wie Linearkombination, Erzeugendensystem und lineare Unabhängigkeit kann man in dieser allgemeineren Situation analog erklären.

In den vorangegangenen Beweisen wird immer wieder durch Skalare dividiert, was einen Skalarenkörper voraussetzt. Über einem Ring ist von den erhaltenen Aussagen über Basis und Dimension wenig zu retten. Daher beschränken wir uns auf zwei Aufgaben (8 und 9), die als Warnung vor diesen Gefahren dienen sollen.

1.5.7. Im Basisauswahlsatz 1.5.3 hatten wir bewiesen, daß man aus jedem endlichen Erzeugendensystem eine Basis auswählen kann. Für die Praxis ist das Verfahren des Weglassens (und die Kontrolle, ob ein Erzeugendensystem übrig bleibt) nicht gut geeignet. Weitaus einfacher ist es, aus einem Erzeugendensystem eine Basis linear zu kombinieren. Wir behandeln hier den Spezialfall eines Untervektorraumes $W \subset K^n$; in 2.4.2 werden wir sehen, daß sich der allgemeine Fall darauf zurückführen läßt.

Seien also $a_1, \dots, a_m \in K^n$ gegeben, und sei $W = \text{span}(a_1, \dots, a_m)$. Sind die Vektoren a_i Zeilen, so ergeben sie untereinandergeschrieben eine Matrix

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \in M(m \times n; K),$$

d.h. es ist $a_i = (a_{i1}, \dots, a_{in})$.

Beispiel. Aus der kanonischen Basis $(e_1, \dots, e_n)$ von K^n erhält man die Matrix

$$E_n := \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \in M(n \times n; K),$$

man nennt sie die n -reihige *Einheitsmatrix*. Dieser Name ist durch ihre Wirkung bei der Matrizenmultiplikation erklärt (vgl. 2.5.4). Die Einträge von

$E_n = (\delta_{ij})$ sind die sogenannten **KRONECKER-Symbole**

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

Nun kommen wir zurück auf die schon in Kapitel 0 benutzten Zeilenumformungen. Anstatt der reellen Zahlen stehen Einträge aus einem beliebigen Körper K , und wir betrachten vier verschiedene Arten von *elementaren Zeilenumformungen*:

I Multiplikation der i -ten Zeile mit $\lambda \in K^*$:

$$A = \begin{pmatrix} \vdots \\ a_i \\ \vdots \end{pmatrix} \mapsto \begin{pmatrix} \vdots \\ \lambda a_i \\ \vdots \end{pmatrix} =: A_I.$$

II Addition der j -ten Zeile zur i -ten Zeile:

$$A = \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} \mapsto \begin{pmatrix} \vdots \\ a_i + a_j \\ \vdots \\ a_j \\ \vdots \end{pmatrix} =: A_{II}.$$

III Addition der λ -fachen j -ten Zeile zur i -ten Zeile ($\lambda \in K^*$):

$$A = \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} \mapsto \begin{pmatrix} \vdots \\ a_i + \lambda a_j \\ \vdots \\ a_j \\ \vdots \end{pmatrix} =: A_{III}.$$

IV Vertauschen der i -ten Zeile mit der j -ten Zeile:

$$A = \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} \mapsto \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \end{pmatrix} =: A_{IV}.$$

Dabei bezeichnen jeweils $a_1, \dots, a_m$ die Zeilen von A , es ist stets $i \neq j$ vorausgesetzt, und an den mit Punkten markierten Zeilen ändert sich nichts.

Die Typen III und IV entsprechen 1) und 2) aus 0.4.6. Die Typen I und II sind noch elementarer, denn man kann III und IV daraus durch Kombination erhalten,

und zwar nach folgendem Rezept:

$$\begin{pmatrix} a_i \\ a_j \end{pmatrix} \xrightarrow{\text{I}} \begin{pmatrix} a_i \\ \lambda a_j \end{pmatrix} \xrightarrow{\text{II}} \begin{pmatrix} a_i + \lambda a_j \\ \lambda a_j \end{pmatrix} \xrightarrow{\text{I}} \begin{pmatrix} a_i + \lambda a_j \\ a_j \end{pmatrix} \quad \text{bzw.}$$

$$\begin{pmatrix} a_i \\ a_j \end{pmatrix} \xrightarrow{\text{I}} \begin{pmatrix} a_i \\ -a_j \end{pmatrix} \xrightarrow{\text{II}} \begin{pmatrix} a_i \\ a_i - a_j \end{pmatrix} \xrightarrow{\text{III}}$$

$$\begin{pmatrix} a_i - (a_i - a_j) \\ a_i - a_j \end{pmatrix} = \begin{pmatrix} a_j \\ a_i - a_j \end{pmatrix} \xrightarrow{\text{II}} \begin{pmatrix} a_j \\ a_i \end{pmatrix}.$$

Zum Verständnis der Wirkung von Zeilenumformungen hilft ein weiterer Begriff:

Definition. Ist $A \in M(m \times n; K)$ mit Zeilen $a_1, \dots, a_m$, so heißt

$$\text{ZR}(A) := \text{span}(a_1, \dots, a_m) \subset K^n$$

der Zeilenraum von A .

Lemma. Ist B aus A durch elementare Zeilenumformungen entstanden, so ist $\text{ZR}(B) = \text{ZR}(A)$.

Beweis. Nach der obigen Bemerkung genügt es, die Typen I und II zu betrachten. Ist $B = A_{\text{I}}$ und $v \in \text{ZR}(A)$, so ist

$$v = \dots + \mu_i a_i + \dots = \dots + \frac{\mu_i}{\lambda} (\lambda a_i) + \dots,$$

also auch $v \in \text{ZR}(B)$. Analog folgt $v \in \text{ZR}(A)$ aus $v \in \text{ZR}(B)$.

Ist $B = A_{\text{II}}$ und $v \in \text{ZR}(A)$, so ist

$$v = \dots + \mu_i a_i + \dots + \mu_j a_j + \dots = \dots + \mu_i (a_i + a_j) + \dots + (\mu_j - \mu_i) a_j + \dots,$$

also $v \in \text{ZR}(B)$ und analog umgekehrt. $\square$

Wie in 0.4.7 beweist man den

Satz. Jede Matrix $A \in M(m \times n; K)$ kann man durch elementare Zeilenumformungen auf Zeilenstufenform bringen. $\square$

Damit ist das zu Beginn dieses Abschnitts formulierte Problem gelöst: Hat man aus den gegebenen Vektoren $a_1, \dots, a_m$ die Matrix A aufgestellt und diese zu B in Zeilenstufenform umgeformt, so sind die von Null verschiedenen Zeilen $b_1, \dots, b_r$ von B eine Basis von $W = \text{ZR}(A) = \text{ZR}(B)$, denn $b_1, \dots, b_r$ sind nach Beispiel b) in 1.4.5 linear unabhängig.

Beispiel. Im $\mathbb{R}^5$ seien die Vektoren

$$\begin{aligned}a_1 &= (0, 0, 0, 2, -1), \\a_2 &= (0, 1, -2, 1, 0), \\a_3 &= (0, -1, 2, 1, -1), \\a_4 &= (0, 0, 0, 1, 2)\end{aligned}$$

gegeben. Dann verläuft die Rechnung wie folgt:

$$\begin{aligned}A &= \begin{array}{ccccc|ccccc} 0 & 0 & 0 & 2 & -1 & 0 & 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 & 0 & 0 & 2 & -1 \\ 0 & -1 & 2 & 1 & -1 & 0 & -1 & 2 & 1 & -1 \\ 0 & 0 & 0 & 1 & 2 & 0 & 0 & 0 & 1 & 2 \end{array} \rightsquigarrow \begin{array}{ccccc|ccccc} 0 & 1 & -2 & 1 & 0 & 0 & 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 1 & 2 & 0 & -1 & 2 & 1 & -1 \\ 0 & 0 & 0 & 1 & 2 & 0 & 0 & 0 & 1 & 2 \end{array} \\ &\rightsquigarrow \begin{array}{ccccc|ccccc} 0 & 1 & -2 & 1 & 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 2 & -1 & 0 & 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 2 & -1 & 0 & 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 1 & 2 & 0 & 0 & 0 & 2 & -1 \end{array} \\ &\rightsquigarrow \begin{array}{ccccc|ccccc} 0 & 1 & -2 & 1 & 0 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 & 2 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & -5 & 0 & 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 0 & -5 & 0 & 0 & 0 & 0 & 0 \end{array} = B.\end{aligned}$$

Also ist eine Basis von $W = \text{span}(a_1, a_2, a_3, a_4)$ gegeben durch

$$\begin{aligned}b_1 &= (0, 1, -2, 1, 0), \\b_2 &= (0, 0, 0, 1, 2), \\b_3 &= (0, 0, 0, 0, -5).\end{aligned}$$

1.5.8. Ob man Vektoren als Zeilen oder Spalten schreibt ist willkürlich, aber der Übergang von der einen zur anderen Konvention wirft Fragen auf. Macht man in einer Matrix Zeilen zu Spalten, so werden Spalten zu Zeilen; man nennt das *Transposition*: Zu

$$A = (a_{ij}) \in M(m \times n; K) \quad \text{ist} \quad {}^t A = (a'_{ij}) \in M(n \times m; K) \quad \text{mit} \quad a'_{ij} := a_{ji}.$$

tA heißt die *Transponierte* von A . Zum Beispiel ist für $m = 2$ und $n = 3$

$$\begin{pmatrix} 1 & 0 & 3 \\ 0 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \\ 3 & 1 \end{pmatrix}.$$

Eine quadratische Matrix wird dabei an der Diagonale gespiegelt. Abstrakt gesehen ist die Transposition eine bijektive Abbildung

$$M(m \times n; K) \rightarrow M(n \times m; K), \quad A \mapsto {}^tA,$$

und es gelten die folgenden Rechenregeln:

- 1) ${}^t(A + B) = {}^tA + {}^tB$,
- 2) ${}^t(\lambda \cdot A) = \lambda \cdot {}^tA$,
- 3) ${}^t({}^tA) = A$.

Nun kann man ganz analog zu 1.5.7 für eine Matrix A elementare Spaltenumformungen, Spaltenraum $\text{SR}(A)$, Spaltenstufenform etc. erklären, was durch Transposition auf die entsprechenden Begriffe für Zeilen zurückgeführt werden kann. Das sind einfache Spielereien, aber ein ernsthaftes Problem wird deutlich durch die folgende

Definition. Für eine Matrix $A \in M(m \times n; K)$ sei

$$\begin{aligned} \text{Zeilenrang } A &:= \dim \text{ZR}(A) \quad \text{und} \\ \text{Spaltenrang } A &:= \dim \text{SR}(A). \end{aligned}$$

Man beachte dabei, daß Zeilen- und Spaltenraum in verschiedenen Vektorräumen liegen:

$$\text{ZR}(A) \subset K^n \quad \text{und} \quad \text{SR}(A) \subset K^m.$$

Zum Beispiel für $A = E_n$ ist $\text{ZR}(E_n) = \text{SR}(E_n) = K^n$, also

$$\text{Zeilenrang } E_n = n = \text{Spaltenrang } E_n.$$

Bei der Behandlung linearer Gleichungssysteme benötigt man die etwas überraschende Tatsache, daß diese beiden Zahlen für jede Matrix gleich sind. In 2.6.6 steht genügend viel Theorie für einen indexfreien Beweis zur Verfügung, in Kapitel 6 wird der abstrakte Hintergrund beleuchtet. Der Schlüssel für einen direkten Beweis mit den Hilfsmitteln dieses Kapitels ist das

Lemma. In der Matrix $A \in M(m \times n; K)$ sei die letzte Zeile Linearkombination der vorhergehenden, $\bar{A} \in M(m-1 \times n; K)$ entstehe aus A durch weglassen der letzten Zeile. Dann ist

$$\text{Spaltenrang } \bar{A} = \text{Spaltenrang } A.$$

Beweis. Wir bezeichnen die Zeilen von A mit $a_1, \dots, a_m$ und die Spalten mit $a^1, \dots, a^n$ (man beachte, daß der obere Index keine Potenz bedeuten soll). Nach Voraussetzung gibt es $\mu_1, \dots, \mu_{m-1} \in K$ mit

$$a_m = \mu_1 a_1 + \dots + \mu_{m-1} a_{m-1}.$$

Das bedeutet, daß alle Spalten von A enthalten sind im Untervektorraum

$$W = \{(x_1, \dots, x_m) \in K^m : x_m = \mu_1 x_1 + \dots + \mu_{m-1} x_{m-1}\},$$

also ist $\text{SR}(A) \subset W$. Nun betrachten wir die „Projektion“

$$\pi : W \rightarrow K^{m-1}, \quad x = (x_1, \dots, x_m) \mapsto \bar{x} = (x_1, \dots, x_{m-1}).$$

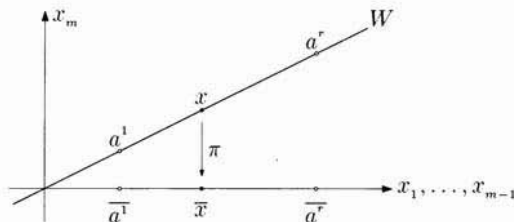


Bild 1.7

Nach Definition von $\bar{A}$ gilt $\text{SR}(\bar{A}) = \pi(\text{SR}(A))$, zu zeigen ist $\dim \text{SR}(\bar{A}) = \dim \text{SR}(A)$.

Nach 1.5.3 können wir annehmen, daß $(a^1, \dots, a^r)$ für ein r mit $0 \leq r \leq n$ eine Basis von $\text{SR}(A)$ ist. Offensichtlich ist $\bar{a}^1, \dots, \bar{a}^r$ ein Erzeugendensystem von $\text{SR}(\bar{A})$. Ist

$$\lambda_1 \bar{a}^1 + \dots + \lambda_r \bar{a}^r = 0, \quad (*)$$

so folgt daraus für die m -ten Komponenten von $a^1, \dots, a^r$

$$\begin{aligned} \lambda_1 a_m^1 + \dots + \lambda_r a_m^r &= \lambda_1 \sum_{i=1}^{m-1} \mu_i a_i^1 + \dots + \lambda_r \sum_{i=1}^{m-1} \mu_i a_i^r \\ &= \mu_1 \sum_{j=1}^r \lambda_j a_j^1 + \dots + \mu_{m-1} \sum_{j=1}^r \lambda_j a_j^{m-1} \\ &= \mu_1 \cdot 0 + \dots + \mu_{m-1} \cdot 0 = 0, \end{aligned}$$

also folgt aus $(*)$, daß $\lambda_1 a^1 + \dots + \lambda_r a^r = 0$ und somit $\lambda_1 = \dots = \lambda_r = 0$. Daher ist $\bar{a}^1, \dots, \bar{a}^r$ eine Basis von $\text{SR}(\bar{A})$.

Selbstverständlich bleibt der Zeilenraum ganz unverändert. Das folgt etwa aus dem Lemma in 1.5.7. $\square$

Der Rest des Beweises für die Gleichheit von Zeilenrang und Spaltenrang ist Aufgabe 13.

Aufgaben zu 1.5

1. Gegeben seien im $\mathbb{R}^5$ die Vektoren $v_1 = (4, 1, 1, 0, -2)$, $v_2 = (0, 1, 4, -1, 2)$, $v_3 = (4, 3, 9, -2, 2)$, $v_4 = (1, 1, 1, 1, 1)$, $v_5 = (0, -2, -8, 2, -4)$.

- Bestimmen Sie eine Basis von $V = \text{span}(v_1, \dots, v_5)$.
- Wählen Sie alle möglichen Basen von V aus den Vektoren $v_1, \dots, v_5$ aus, und kombinieren Sie jeweils $v_1, \dots, v_5$ daraus linear.

2. Geben Sie für folgende Vektorräume jeweils eine Basis an:

- $\{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 = x_3\}$,
- $\{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 : x_1 + 3x_2 + 2x_4 = 0, 2x_1 + x_2 + x_3 = 0\}$,
- $\text{span}(t^2, t^2 + t, t^2 + 1, t^2 + t + 1, t^7 + t^5) \subset \mathbb{R}[t]$,
- $\{f \in \text{Abb}(\mathbb{R}, \mathbb{R}) : f(x) = 0 \text{ bis auf endlich viele } x \in \mathbb{R}\}$.

3. Für $d \in \mathbb{N}$ sei

$K[t_1, \dots, t_n]_{(d)} := \{F \in K[t_1, \dots, t_n] : F \text{ ist homogen vom Grad } d \text{ oder } F = 0\}$
(vgl. Aufgabe 9 zu 1.3). Beweisen Sie, dass $K[t_1, \dots, t_n]_{(d)} \subset K[t_1, \dots, t_n]$ ein Untervektorraum ist und bestimmen Sie $\dim K[t_1, \dots, t_n]_{(d)}$.

4. Zeigen Sie, dass $\mathbb{C}$ endlich erzeugt über $\mathbb{R}$ ist, aber $\mathbb{R}$ nicht endlich erzeugt über $\mathbb{Q}$.

5. Ist $(v_i)_{i \in I}$ eine Basis des Vektorraumes V und $(w_j)_{j \in J}$ eine Basis des Vektorraumes W , so ist $((v_i, 0))_{i \in I} \cup ((0, w_j))_{j \in J}$ eine Basis von $V \times W$ (vgl. Aufgabe 2 zu 1.4). Insbesondere gilt

$$\dim V \times W = \dim V + \dim W,$$

falls $\dim V, \dim W < \infty$.

6. Sei V ein reeller Vektorraum und $a, b, c, d, e \in V$. Zeigen Sie, dass die folgenden Vektoren linear abhängig sind:

$$\begin{aligned} v_1 &= a + b + c, & v_2 &= 2a + 2b + 2c - d, & v_3 &= a - b - e, \\ v_4 &= 5a + 6b - c + d + e, & v_5 &= a - c + 3e, & v_6 &= a + b + d + e. \end{aligned}$$

7. Für einen endlichdimensionalen Vektorraum V definieren wir

$$h(V) := \sup \{n \in \mathbb{N} : \text{es gibt eine Kette } V_0 \subset V_1 \subset \dots \subset V_{n-1} \subset V_n \text{ von Untervektorräumen } V_i \subset V\}.$$

Zeigen Sie $h(V) = \dim V$.

8. Sei $R = \mathcal{C}(\mathbb{R}, \mathbb{R})$ der Ring der stetigen Funktionen und

$$W := \{f \in R : \text{es gibt ein } \varrho \in \mathbb{R} \text{ mit } f(x) = 0 \text{ für } x \geq \varrho\} \subset R.$$

Für $k \in \mathbb{N}$ definieren wir die Funktion

$$f_k(x) := \begin{cases} 0 & \text{für alle } x \geq k, \\ k - x & \text{für } x \leq k. \end{cases}$$

a) $W = \text{span}_R(f_k)_{k \in \mathbb{N}}$.

b) W ist über R nicht endlich erzeugt (aber R ist über R endlich erzeugt).

c) Ist die Familie $(f_k)_{k \in \mathbb{N}}$ linear abhängig über R ?

9. Zeigen Sie $\mathbb{Z} = 2\mathbb{Z} + 3\mathbb{Z}$ und folgern Sie daraus, dass es in $\mathbb{Z}$ unverkürzbare Erzeugendensysteme verschiedener Längen gibt.

10. Wie viele Elemente hat ein endlichdimensionaler Vektorraum über einem endlichen Körper?

11.* a) Ist K ein Körper mit $\text{char } K = p > 0$, so enthält K einen zu $\mathbb{Z}/p\mathbb{Z}$ isomorphen Körper und kann somit als $\mathbb{Z}/p\mathbb{Z}$ -Vektorraum aufgefasst werden.

b) Zeigen Sie: Ist K ein endlicher Körper mit $\text{char } K = p$, so hat K genau p^n Elemente, wobei $n = \dim_{\mathbb{Z}/p\mathbb{Z}} K$.

12. Zeigen Sie: Zeilenrang = Spaltenrang für Matrizen mit sehr kleiner Zeilenzahl (etwa $m = 1, 2$) und beliebig großer Spaltenzahl n .

13. Folgern Sie aus Lemma 1.5.8, dass für eine Matrix $A \in M(m \times n; K)$

a) Zeilenrang $A \leq$ Spaltenrang A ,

b) Zeilenrang $A \geq$ Spaltenrang A ,

und somit insgesamt Zeilenrang $A =$ Spaltenrang A gilt.

1.6 Summen von Vektorräumen*

Für die Menge der Linearkombinationen aus Vektoren $v_1, \dots, v_r$ hatten wir in 1.4.4 auch die Schreibweise

$$Kv_1 + \dots + Kv_r$$

angegeben. Das soll andeuten, daß dies die Menge der Summen von Vektoren aus den für $v_j \neq 0$ eindimensionalen Räumen Kv_j ist. Wir betrachten nun den Fall, daß die Summanden aus beliebigen Untervektorräumen stammen, das wird in Kapitel 4 nützlich sein. Einem eiligen Leser raten wir, dies zunächst zu überblättern.

1.6.1. Ausgangspunkt ist die folgende

Definition. Gegeben sei ein K -Vektorraum V mit Untervektorräumen $W_1, \dots, W_r \subset V$. Dann heißt

$$W_1 + \dots + W_r := \{v \in V : \text{es gibt } v_j \in W_j \text{ mit } v = v_1 + \dots + v_r\}$$

die *Summe* von $W_1, \dots, W_r$.

Ohne Schwierigkeit beweist man die

Bemerkung. Für die oben definierte Summe gilt:

a) $W_1 + \dots + W_r \subset V$ ist ein Untervektorraum.

b) $W_1 + \dots + W_r = \text{span}(W_1 \cup \dots \cup W_r)$.

c) $\dim(W_1 + \dots + W_r) \leq \dim W_1 + \dots + \dim W_r$. □

Nun ist die Frage naheliegend, wie unscharf die letzte Ungleichung ist. Im Fall $r = 2$ ist das einfach zu beantworten:

Dimensionsformel für Summen. Für endlichdimensionale Untervektorräume $W_1, W_2 \subset V$ gilt

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2).$$

Beweis. Wir beginnen mit einer Basis $(v_1, \dots, v_m)$ von $W_1 \cap W_2$ und ergänzen sie entsprechend 1.5.5 zu Basen

$$(v_1, \dots, v_m, w_1, \dots, w_k) \text{ von } W_1 \text{ und } (v_1, \dots, v_m, w'_1, \dots, w'_l) \text{ von } W_2.$$

Die Behauptung ist bewiesen, wenn wir gezeigt haben, daß

$$\mathcal{B} := (v_1, \dots, v_m, w_1, \dots, w_k, w'_1, \dots, w'_l)$$

eine Basis von $W_1 + W_2$ ist. Daß $W_1 + W_2$ von $\mathcal{B}$ erzeugt wird ist klar. Zum Beweis der linearen Unabhängigkeit sei

$$\lambda_1 v_1 + \dots + \lambda_m v_m + \mu_1 w_1 + \dots + \mu_k w_k + \mu'_1 w'_1 + \dots + \mu'_l w'_l = 0. \quad (*)$$

Setzen wir

$$v := \lambda_1 v_1 + \dots + \lambda_m v_m + \mu_1 w_1 + \dots + \mu_k w_k,$$

so ist $v \in W_1$ und $-v = \mu'_1 w'_1 + \dots + \mu'_l w'_l \in W_2$, also $v \in W_1 \cap W_2$. Also ist

$$v = \lambda'_1 v_1 + \dots + \lambda'_m v_m$$

mit $\lambda'_1, \dots, \lambda'_m \in K$, und wegen der Eindeutigkeit der Linearkombinationen folgt insbesondere $\mu_1 = \dots = \mu_k = 0$. Setzt man das in (*) ein, so folgt auch

$$\lambda_1 = \dots = \lambda_m = \mu'_1 = \dots = \mu'_l = 0. \quad \square$$

1.6.2. Der Korrekturterm in der Dimensionsformel wird durch die Dimension des Durchschnitts verursacht. Diesen Mangel einer Summendarstellung kann man auch anders charakterisieren.

Lemma. Ist $V = W_1 + W_2$, so sind folgende Bedingungen äquivalent:

- i) $W_1 \cap W_2 = \{0\}$.
- ii) Jedes $v \in V$ ist eindeutig darstellbar als $v = w_1 + w_2$ mit $w_1 \in W_1$, $w_2 \in W_2$.
- iii) Zwei von Null verschiedene Vektoren $w_1 \in W_1$ und $w_2 \in W_2$ sind linear unabhängig.

Beweis. i) $\Rightarrow$ ii): Ist $v = w_1 + w_2 = \tilde{w}_1 + \tilde{w}_2$, so folgt

$$w_1 - \tilde{w}_1 = \tilde{w}_2 - w_2 \in W_1 \cap W_2.$$

ii) $\Rightarrow$ iii): Sind w_1, w_2 linear abhängig, so erhält man verschiedene Darstellungen des Nullvektors.

iii) $\Rightarrow$ i): Ist $0 \neq v \in W_1 \cap W_2$, so erhält man einen Widerspruch zu iii) durch

$$1v + (-1)v = 0. \quad \square$$

Da Bedingung i) am kürzesten aufzuschreiben ist, die folgende

Definition. Ein Vektorraum V heißt *direkte Summe* von zwei Untervektorräumen W_1 und W_2 , in Zeichen

$$V = W_1 \oplus W_2, \quad \text{wenn} \quad V = W_1 + W_2 \quad \text{und} \quad W_1 \cap W_2 = \{0\}.$$

1.6.3. Im endlichdimensionalen Fall hat man einfacher nachzuprüfende Bedingungen für die Direktheit einer Summe.

Satz. Ist V endlichdimensional mit Untervektorräumen W_1 und W_2 , so sind folgende Bedingungen gleichwertig:

- i) $V = W_1 \oplus W_2$.
- ii) Es gibt Basen $(w_1, \dots, w_k)$ von W_1 und $(w'_1, \dots, w'_l)$ von W_2 , so daß $(w_1, \dots, w_k, w'_1, \dots, w'_l)$ eine Basis von V ist.
- iii) $V = W_1 + W_2$ und $\dim V = \dim W_1 + \dim W_2$.

Beweis. i) $\Rightarrow$ ii) $\Rightarrow$ iii) folgt aus der Dimensionsformel (einschließlich Beweis) in 1.6.1 im Spezialfall $W_1 \cap W_2 = \{0\}$.

iii) $\Rightarrow$ i): Nach der Dimensionsformel ist $\dim(W_1 \cap W_2) = 0$, also

$$W_1 \cap W_2 = \{0\}. \quad \square$$

Daraus folgt sofort die Existenz von „direkten Summanden“:

Korollar. Ist V endlichdimensional und $W \subset V$ Untervektorraum, so gibt es dazu einen (im allgemeinen nicht eindeutig bestimmten) Untervektorraum $W' \subset V$, so daß

$$V = W \oplus W'.$$

W' heißt *direkter Summand* von V zu W .

Beweis. Man nehme eine Basis $(v_1, \dots, v_m)$ von W , ergänze sie nach 1.5.4 zu einer Basis $(v_1, \dots, v_m, v_{m+1}, \dots, v_n)$ von V und definiere

$$W' := \text{span}(v_{m+1}, \dots, v_n). \quad \square$$

1.6.4. In Kapitel 4 werden wir direkte Summen von mehreren Unterräumen antreffen. Zur Vorsorge dafür die

Definition. Ein Vektorraum V heißt *direkte Summe* von Untervektorräumen $W_1, \dots, W_k$, in Zeichen

$$V = W_1 \oplus \dots \oplus W_k,$$

wenn folgende Bedingungen erfüllt sind:

DS1 $V = W_1 + \dots + W_k$.

DS2 Sind $w_1 \in W_1, \dots, w_k \in W_k$ gegeben mit $w_1 + \dots + w_k = 0$, so folgt $w_1 = \dots = w_k = 0$.

Vorsicht! Bedingung DS2 darf man für $k > 2$ nicht ersetzen durch

$$W_1 \cap \dots \cap W_k = \{0\} \quad \text{oder} \quad W_i \cap W_j = \{0\} \quad \text{für alle } i \neq j$$

(vgl. Aufgabe 1).

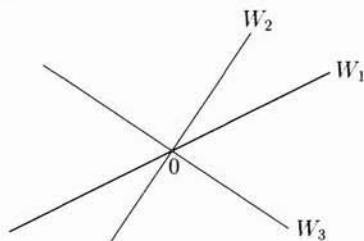


Bild 1.8

Beispiel. Ist $(v_1, \dots, v_n)$ eine Basis von V , so ist $V = K v_1 \oplus \dots \oplus K v_n$.

Satz. Für Untervektorräume $W_1, \dots, W_k$ eines endlichdimensionalen Vektorraumes V sind folgende Bedingungen äquivalent:

i) $V = W_1 \oplus \dots \oplus W_k$.

ii) Ist für jedes $i \in \{1, \dots, k\}$ eine Basis $(v_1^{(i)}, \dots, v_{r_i}^{(i)})$ von W_i gegeben, so ist

$$\mathcal{B} := (v_1^{(1)}, \dots, v_{r_1}^{(1)}, \dots, v_1^{(k)}, \dots, v_{r_k}^{(k)})$$

eine Basis von V .

iii) $V = W_1 + \dots + W_k$ und $\dim V = \dim W_1 + \dots + \dim W_k$.

Man beachte die Klammern bei den oberen Indizes. Sie dienen zur Unterscheidung von Exponenten, deren Stammpatz an dieser Stelle ist.

Beweis. i) $\Rightarrow$ ii). Offensichtlich ist $\mathcal{B}$ ein Erzeugendensystem. Zum Beweis der linearen Unabhängigkeit sei

$$\mu_1^{(1)} v_1^{(1)} + \dots + \mu_{r_1}^{(1)} v_{r_1}^{(1)} + \dots + \mu_1^{(k)} v_1^{(k)} + \dots + \mu_{r_k}^{(k)} v_{r_k}^{(k)} = 0.$$

Setzen wir $w_i := \mu_1^{(i)} v_1^{(i)} + \dots + \mu_{r_i}^{(i)} v_{r_i}^{(i)}$, so bedeutet das

$$w_1 + \dots + w_k = 0,$$

und wegen DS2 muß $w_1 = \dots = w_k = 0$ sein. Also ist

$$\mu_1^{(i)} v_1^{(i)} + \dots + \mu_{r_i}^{(i)} v_{r_i}^{(i)} = 0 \quad \text{für } i = 1, \dots, k,$$

und daraus folgt $\mu_1^{(i)} = \dots = \mu_{r_i}^{(i)} = 0$.

ii) $\Leftrightarrow$ iii) ist klar.

ii) $\Rightarrow$ i). Zum Nachweis von DS2 stellen wir jedes $w_i \in W_i$ durch die gegebene Basis von W_i dar:

$$w_i = \mu_1^{(i)} v_1^{(i)} + \dots + \mu_{r_i}^{(i)} v_{r_i}^{(i)}.$$

Aus der Annahme $w_1 + \dots + w_k = 0$ folgt

$$\sum_{i=1}^k \sum_{\varrho=1}^{r_i} \mu_{\varrho}^{(i)} v_{\varrho}^{(i)} = 0.$$

Da $\mathcal{B}$ eine Basis ist, folgt $\mu_{\varrho}^{(i)} = 0$ für alle ϱ und i , also ist auch

$$w_1 = \dots = w_k = 0.$$

□

Aufgaben zu 1.6

1. Beweisen Sie, dass für einen Vektorraum V folgende Bedingungen äquivalent sind:

i) $V = W_1 \oplus \dots \oplus W_k$.

ii) Jedes $v \in V$ ist eindeutig darstellbar als $v = w_1 + \dots + w_k$ mit $w_i \in W_i$.

iii) $V = W_1 + \dots + W_k$, $W_i \neq \{0\}$ für alle i und von Null verschiedene Vektoren $w_1 \in W_1, \dots, w_k \in W_k$ sind linear unabhängig.

Vorsicht! Die Voraussetzung $W_i \neq \{0\}$ ist wesentlich. Ist etwa $W_1 = \{0\}$, so ist die angegebene zweite Bedingung stets erfüllt!

iv) $V = W_1 + \dots + W_k$ und $W_i \cap \sum_{\substack{j=1 \\ j \neq i}}^k W_j = \{0\}$ für alle $i \in \{1, \dots, k\}$.

v) $V = W_1 + \dots + W_k$ und $W_i \cap (W_{i+1} + \dots + W_k) = \{0\}$ für alle $i \in \{1, \dots, k-1\}$.

Zeigen Sie anhand von Gegenbeispielen, dass die obigen Bedingungen für $k > 2$ im Allgemeinen nicht äquivalent sind zu $W_1 \cap \dots \cap W_k = \{0\}$ bzw. $W_i \cap W_j = \{0\}$ für alle $i \neq j$.

2. Sind V und W Vektorräume, so gilt

$$V \times W = (V \times \{0\}) \oplus (\{0\} \times W).$$

3. Eine Matrix $A \in M(n \times n; K)$ heißt *symmetrisch*, falls $A = {}^t A$.

a) Zeigen Sie, dass die symmetrischen Matrizen einen Untervektorraum $\text{Sym}(n; K)$ von $M(n \times n; K)$ bilden. Geben Sie die Dimension und eine Basis von $\text{Sym}(n; K)$ an.

Ist $\text{char } K \neq 2$, so heißt $A \in M(n \times n; K)$ *schief-symmetrisch* (oder *alternierend*), falls ${}^tA = -A$. Im Folgenden sei stets $\text{char } K \neq 2$.

- b) Zeigen Sie, dass die alternierenden Matrizen einen Untervektorraum $\text{Alt}(n; K)$ von $M(n \times n; K)$ bilden. Bestimmen Sie auch für $\text{Alt}(n; K)$ die Dimension und eine Basis.
- c) Für $A \in M(n \times n; K)$ sei $A_s := \frac{1}{2}(A + {}^tA)$ und $A_a := \frac{1}{2}(A - {}^tA)$. Zeigen Sie: A_s ist symmetrisch, A_a ist alternierend, und es gilt $A = A_s + A_a$.
- d) Es gilt: $M(n \times n; K) = \text{Sym}(n; K) \oplus \text{Alt}(n; K)$.

Kapitel 2

Lineare Abbildungen

2.1 Beispiele und Definitionen

2.1.1. a) Besonders wichtige Funktionen $f: \mathbb{R} \rightarrow \mathbb{R}$ sind die Polynome mit

$$f(x) = a_0 + a_1x + \dots + a_nx^n,$$

denn die ganze Abbildungsvorschrift ist durch Angabe der Koeffizienten $a_0, \dots, a_n$, d.h. von $n+1$ Zahlen, festgelegt. Ist $a_n \neq 0$, so nennt man n den Grad von f . Für Grad 0 hat man die konstanten und für Grad 1 die linearen Funktionen. In der Analysis untersucht man, ob sich eine beliebige Funktion bei kleinen Veränderungen des Arguments x nahezu wie eine lineare Funktion verhält; das führt zum Begriff der Differenzierbarkeit. Daher sind die linearen Funktionen unentbehrliches Hilfsmittel der Analysis. Wir betrachten also ein

$$f: \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto ax + b = f(x),$$

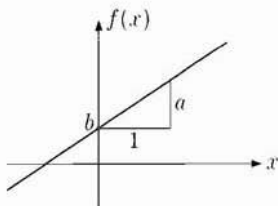


Bild 2.1

mit $a, b \in \mathbb{R}$. Ihr Graph ist eine Gerade mit Steigung a , die durch $(0, b)$ geht. Setzt man den konstanten Anteil $b = 0$, so bleibt eine Funktion $f(x) = ax$, die durch eine einzige reelle Zahl a festgelegt ist, und offensichtlich die Eigenschaften

$$f(x + x') = f(x) + f(x') \quad \text{und} \quad f(\lambda x) = \lambda f(x) \quad (*)$$

für beliebige $x, x', \lambda \in \mathbb{R}$ hat.

b) Viel interessanter und anschaulich einfach zu verstehen ist die Situation in der Ebene. Zunächst betrachten wir eine *Drehung* um den Nullpunkt mit dem Winkel ϑ . Das ist eine Abbildung

$$F: \mathbb{R}^2 \rightarrow \mathbb{R}^2,$$

die sich wie folgt beschreiben läßt. Es ist

$$F(0) = 0, \quad F(e_1) = (\cos \vartheta, \sin \vartheta), \quad F(e_2) = (-\sin \vartheta, \cos \vartheta),$$

und für ein beliebiges $x = (x_1, x_2) = x_1 e_1 + x_2 e_2$ ist

$$F(x) = x_1 F(e_1) + x_2 F(e_2) = (x_1 \cos \vartheta - x_2 \sin \vartheta, x_1 \sin \vartheta + x_2 \cos \vartheta).$$

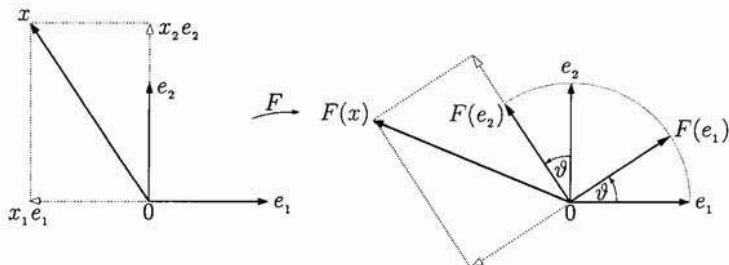


Bild 2.2

Benutzen wir die Matrix

$$A = \begin{pmatrix} \cos \vartheta & -\sin \vartheta \\ \sin \vartheta & \cos \vartheta \end{pmatrix},$$

und schreiben wir x und $F(x)$ als Spaltenvektoren, so ist mit der in 0.4.1 erklärten Multiplikation $F(x) = A \cdot x$.

An dieser Überlegung ist zu sehen, daß man für $F(e_1)$ und $F(e_2)$ beliebige Vektoren $a = (a_1, a_2)$, $b = (b_1, b_2)$ vorschreiben kann. Dann läßt sich F mit der Konstruktion aus Bild 2.3 zu einer Abbildung von $\mathbb{R}^2$ auf sich ausdehnen.

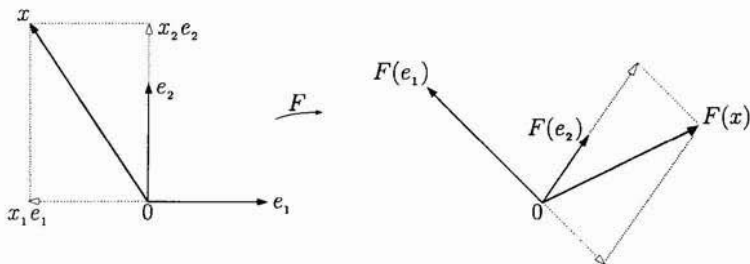


Bild 2.3

Mit Hilfe der Matrix

$$A = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$$

ist F wieder beschrieben durch $F(x) = A \cdot x$.

c) Betrachtet man nun anstelle von $\mathbb{R}$ einen beliebigen Körper K , und nimmt man m und n anstelle von 2, so ist man sofort in einer recht allgemeinen Situation. Ist nämlich $A = (a_{ij})$ eine $m \times n$ -Matrix, so erhält man daraus eine Abbildung

$$F: K^n \rightarrow K^m, \quad \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}.$$

Wie man leicht nachrechnet, hat sie wieder die Eigenschaft

$$F(x + x') = F(x) + F(x'), \quad F(\lambda x) = \lambda F(x) \quad (*)$$

für $x, x' \in K^n$ und $\lambda \in K$.

Mit Hilfe der in 0.4.1 erklärten Multiplikation einer Matrix A mit einer Spalte kann man diese Abbildung einfacher beschreiben als

$$F: K^n \rightarrow K^m, \quad x \mapsto A \cdot x,$$

wenn man die Vektoren von K^n und K^m als Spalten entsprechender Höhe schreibt. Setzt man für x die Basisvektoren $e_1, \dots, e_n$ als Spalten ein und rechnet man die Produkte aus, so stellt man fest, daß

$$A \cdot e_1, \dots, A \cdot e_n$$

die Spalten von A in dieser Reihenfolge sind. Diese Beobachtung:

Die Spaltenvektoren der Matrix sind die Bilder der Basisvektoren

sollte man sich einprägen.

d) Die Transposition von Matrizen als Abbildung

$$F: M(m \times n; K) \rightarrow M(n \times m; K)$$

hat nach den Rechenregeln in 1.5.8 die zu (*) analogen Eigenschaften.

e) Der Vektorraum $V = \mathcal{C}(I; \mathbb{R})$ der auf einem Intervall $I = [a, b]$ stetigen Funktionen ist unendlichdimensional, und die Abbildung

$$S: \mathcal{C}(I; \mathbb{R}) \rightarrow \mathbb{R}, \quad f \mapsto \int_a^b f(x) dx,$$

hat nach den Rechenregeln für Integrale die zu (*) analogen Eigenschaften

$$S(f + g) = S(f) + S(g), \quad S(\lambda f) = \lambda S(f).$$

f) Ist $V = \mathcal{D}(I; \mathbb{R})$ der Vektorraum der beliebig oft differenzierbaren Funktionen, so hat die durch Differentiation erklärte Abbildung

$$D: V \rightarrow V, \quad f \mapsto f',$$

die zu (*) analoge Eigenschaft.

2.1.2. Die obigen Beispiele motivieren die folgende

Definition. Eine Abbildung $F: V \rightarrow W$ zwischen K -Vektorräumen V und W heißt *linear* (genauer *K -linear* oder *Homomorphismus von K -Vektorräumen*), wenn

$$\text{L1} \quad F(v + w) = F(v) + F(w),$$

$$\text{L2} \quad F(\lambda v) = \lambda F(v)$$

für alle $v, w \in V$ und alle $\lambda \in K$. Diese beiden Bedingungen kann man zusammenfassen zu einer:

$$\text{L} \quad F(\lambda v + \mu w) = \lambda F(v) + \mu F(w)$$

für alle $v, w \in V$ und $\lambda, \mu \in K$. Man überlegt sich ganz leicht, daß L1 und L2 zusammen mit L gleichwertig sind.

Es ist üblich, den Begriff Homomorphismus zu verschärfen. Man nennt eine lineare Abbildung $F: V \rightarrow W$ einen

Isomorphismus, wenn F bijektiv ist,

Endomorphismus, wenn $V = W$,

Automorphismus, wenn $V = W$ und F bijektiv ist.

Wir notieren einige einfache Folgerungen aus den Axiomen:

Bemerkung. Ist $F: V \rightarrow W$ linear, so gilt:

$$a) \quad F(0) = 0 \text{ und } F(v - w) = F(v) - F(w).$$

$$b) \quad F(\lambda_1 v_1 + \dots + \lambda_n v_n) = \lambda_1 F(v_1) + \dots + \lambda_n F(v_n).$$

c) Ist die Familie $(v_i)_{i \in I}$ in V linear abhängig, so ist $(F(v_i))_{i \in I}$ in W linear abhängig.

d) Sind $V' \subset V$ und $W' \subset W$ Untervektorräume, so sind auch $F(V') \subset W$ und $F^{-1}(W') \subset V$ Untervektorräume.

$$e) \quad \dim F(V) \leq \dim V.$$

f) Ist F ein Isomorphismus, so ist auch $F^{-1}: W \rightarrow V$ linear.

Beweis. a) $F(0) = F(0 \cdot 0) = 0 \cdot F(0) = 0$ und

$$F(v - w) = F(v + (-1)w) = F(v) + (-1)F(w) = F(v) - F(w).$$

b) folgt durch wiederholte Anwendung der Regel L.

c) Ist für $i_1, \dots, i_n \in I$ und $\lambda_1, \dots, \lambda_n \in K$

$$\lambda_1 v_{i_1} + \dots + \lambda_n v_{i_n} = 0,$$

so folgt nach b)

$$\lambda_1 F(v_{i_1}) + \dots + \lambda_n F(v_{i_n}) = 0.$$

d) Wegen $0 \in V'$ ist $0 = F(0) \in F(V')$. Sind $w, w' \in F(V')$, so gibt es $v, v' \in V'$ mit $F(v) = w$ und $F(v') = w'$. Also ist

$$w + w' = F(v) + F(v') = F(v + v') \in F(V'),$$

denn $v + v' \in V'$. Für $\lambda \in K$ gilt wegen $\lambda v \in V'$

$$\lambda w = \lambda F(v) = F(\lambda v) \in F(V').$$

Sind $v, v' \in F^{-1}(W')$, so bedeutet das $F(v), F(v') \in W'$. Also ist

$$F(v + v') = F(v) + F(v') \in W',$$

also $v + v' \in F^{-1}(W')$ und analog sieht man $\lambda v \in F^{-1}(W')$.

e) Hat man $w_1 = F(v_1), \dots, w_n = F(v_n) \in F(V)$ linear unabhängig, so sind nach c) auch $v_1, \dots, v_n$ in V linear unabhängig.

f) Seien $w, w' \in W$ und $\lambda, \mu \in K$. Ist $w = F(v)$ und $w' = F(v')$, so ist $v = F^{-1}(w)$, $v' = F^{-1}(w')$, und es folgt aus

$$F(\lambda v + \mu v') = \lambda w + \mu w'$$

durch Anwendung von F^{-1} auf beiden Seiten

$$\lambda F^{-1}(w) + \mu F^{-1}(w') = F^{-1}(\lambda w + \mu w').$$

□

2.1.3. Lineare Abbildungen kann man in verschiedener Weise miteinander verknüpfen. Zunächst betrachten wir die Hintereinanderschaltung.

Bemerkung 1. Sind U, V, W Vektorräume und $G: U \rightarrow V$, $F: V \rightarrow W$ lineare Abbildungen, so ist auch

$$F \circ G: U \rightarrow W \quad \text{linear.}$$

Man beachte dabei wie immer die Reihenfolge der Abbildungen, was man sich in einem Diagramm aufzeichnen kann:

$$\begin{array}{ccccc} & & G & \rightarrow & V & \xrightarrow{F} & & \\ & & & & & & & \\ U & \xrightarrow{F \circ G} & & & W & & & \end{array}$$

Beweis. Für $u, u' \in U$ ist

$$\begin{aligned} (F \circ G)(u + u') &= F(G(u + u')) = F(G(u) + G(u')) \\ &= F(G(u)) + F(G(u')) \\ &= (F \circ G)(u) + (F \circ G)(u'). \end{aligned}$$

Ganz analog zeigt man $(F \circ G)(\lambda u) = \lambda(F \circ G)(u)$. □

In Aufgabe 3 zu 1.4 hatten wir gesehen, wie man für eine Menge X und einen Vektorraum W die Menge $\text{Abb}(X, W)$ zu einem Vektorraum machen kann. Ist auch $X = V$ ein K -Vektorraum, so definiert man

$$\text{Hom}_K(V, W) := \{F: V \rightarrow W: F \text{ ist } K\text{-linear}\}.$$

Falls klar ist, welcher Körper K gemeint ist, schreibt man einfacher $\text{Hom}(V, W)$.

Bemerkung 2. Für Vektorräume V und W über demselben Körper K ist

$$\text{Hom}_K(V, W) \subset \text{Abb}(V, W)$$

ein Untervektorraum.

Beweis. Für $F, G \in \text{Hom}_K(V, W)$ und $\lambda \in K$ ist zu zeigen, daß $F + G$ und λF wieder K -linear sind. Das ist aber klar, denn für alle $\sigma, \tau \in K$ und $v, w \in V$ ist

$$\begin{aligned} (F + G)(\sigma v + \tau w) &= F(\sigma v + \tau w) + G(\sigma v + \tau w) \\ &= \sigma F(v) + \tau F(w) + \sigma G(v) + \tau G(w) \\ &= \sigma(F(v) + G(v)) + \tau(F(w) + G(w)) \\ &= \sigma(F + G)(v) + \tau(F + G)(w) \end{aligned}$$

und

$$\begin{aligned} (\lambda \cdot F)(\sigma v + \tau w) &= \lambda F(\sigma v + \tau w) = \lambda(\sigma F(v) + \tau F(w)) \\ &= \sigma \lambda F(v) + \tau \lambda F(w) = \sigma(\lambda \cdot F)(v) + \tau(\lambda \cdot F)(w). \end{aligned}$$

Wir vermerken noch, daß der Nullvektor in $\text{Hom}_K(V, W)$ die Nullabbildung

$$0: V \rightarrow W \quad \text{mit} \quad 0(v) := 0 \quad \text{für alle } v \in V$$

ist, und die zu $F: V \rightarrow W$ negative Abbildung gegeben ist durch

$$-F: V \rightarrow W \quad \text{mit} \quad (-F)(v) := -F(v) \quad \text{für alle } v \in V. \quad \square$$

Die Dimension von $\text{Hom}(V, W)$ werden wir in 2.4.2 berechnen (siehe auch Aufgabe 6 zu 2.4).

2.1.4. Im Spezialfall $V = W$ setzt man

$$\text{End}(V) := \text{Hom}(V, V),$$

das sind die *Endomorphismen* von V . Diese Menge wird, wie wir in 2.1.3 gesehen haben, zu einem Vektorraum mit Addition und Multiplikation mit Skalaren. Man kann überdies eine *Multiplikation* durch die Hintereinanderschaltung erklären, in Zeichen

$$F \cdot G := F \circ G \quad \text{für} \quad F, G \in \text{End}(V).$$

Satz. Ist V ein K -Vektorraum, so ist $\text{End}(V)$ zusammen mit der oben erklärten Addition und Multiplikation ein Ring.

Der einfache *Beweis* sei dem Leser überlassen. In 2.6.4 werden wir sehen, daß dieser *Endomorphismenring* für endlichdimensionales V zu einem Matrizenring isomorph ist. Damit erhält man eine Methode, viele der interessanten Unterringe von $\text{End}(V)$ durch die Gestalt der entsprechenden Matrizen zu beschreiben.

Aufgaben zu 2.1

1. Sei X eine Menge und V der $\mathbb{R}$ -Vektorraum aller Funktionen $f: X \rightarrow \mathbb{R}$. Beweisen Sie: Ist $\varphi: X \rightarrow X$ eine beliebige Abbildung, so ist die Abbildung

$$F_\varphi: V \rightarrow V, \quad f \mapsto f \circ \varphi$$

$\mathbb{R}$ -linear.

2. Untersuchen Sie die folgenden Abbildungen auf Linearität:

- a) $\mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad (x, y) \mapsto (3x + 2y, x), \quad$ b) $\mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto ax + b,$
 c) $\mathbb{Q}^2 \rightarrow \mathbb{R}, \quad (x, y) \mapsto x + \sqrt{2}y$ (über $\mathbb{Q}$), d) $\mathbb{C} \rightarrow \mathbb{C}, \quad z \mapsto \bar{z},$
 e) $\text{Abb}(\mathbb{R}, \mathbb{R}) \rightarrow \mathbb{R}, \quad f \mapsto f(1), \quad$ f) $\mathbb{C} \rightarrow \mathbb{C}, \quad z \mapsto \bar{z}$ (über $\mathbb{R}$).

3. Für einen Endomorphismus $F: V \rightarrow V$ ist die Menge der *Fixpunkte* von F definiert durch $\text{Fix } F := \{v \in V: F(v) = v\}.$

- a) Zeigen Sie, dass $\text{Fix } F \subset V$ ein Untervektorraum ist.
 b) Sei der Endomorphismus F gegeben durch

$$\text{i) } F: \mathbb{R}^3 \rightarrow \mathbb{R}^3, x \mapsto \begin{pmatrix} 1 & 2 & 2 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix} \cdot x,$$

$$\text{ii) } F: \mathbb{R}[t] \rightarrow \mathbb{R}[t], P \mapsto P',$$

$$\text{iii) } F: \mathcal{D}(\mathbb{R}, \mathbb{R}) \rightarrow \mathcal{D}(\mathbb{R}, \mathbb{R}), f \mapsto f'.$$

Bestimmen Sie jeweils eine Basis von $\text{Fix } F$.

4. Zeigen Sie, dass die Menge $\text{Aut}(V)$ der Automorphismen eines Vektorraums V mit der Komposition von Abbildungen als Verknüpfung eine Gruppe ist.

5. Sei $F: V \rightarrow V$ ein Endomorphismus des Vektorraums V und $v \in V$, so dass für eine natürliche Zahl n gilt:

$$F^n(v) \neq 0 \quad \text{und} \quad F^{n+1}(v) = 0.$$

Beweisen Sie, dass dann $v, F(v), \dots, F^n(v)$ linear unabhängig sind.

6. Ist $F: V \rightarrow W$ ein Isomorphismus und $V = U_1 \oplus U_2$, so ist $W = F(U_1) \oplus F(U_2)$.

2.2 Bild, Fasern und Kern, Quotientenvektorräume*

Nachdem wir eine ganze Reihe formaler Eigenschaften von linearen Abbildungen behandelt haben, soll nun versucht werden, die „Geometrie“ solcher Abbildungen etwas besser zu verstehen.

2.2.1. Ist $F: V \rightarrow W$ eine lineare Abbildung, so nennen wir

$$\begin{aligned} \operatorname{Im} F &:= F(V) && \text{das Bild von } F \text{ (vgl. 1.1.3),} \\ F^{-1}(w) &:= \{v \in V : F(v) = w\} && \text{die Faser über } w \in W \text{ und} \\ \operatorname{Ker} F &= F^{-1}(0) && \text{den Kern von } F. \end{aligned}$$

Bemerkung. Ist $F: V \rightarrow W$ linear, so gilt:

- a) $\operatorname{Im} F \subset W$ und $\operatorname{Ker} F \subset V$ sind Untervektorräume.
- b) F surjektiv $\Leftrightarrow \operatorname{Im} F = W$.
- c) F injektiv $\Leftrightarrow \operatorname{Ker} F = \{0\}$.
- d) Ist F injektiv, und sind $v_1, \dots, v_n \in V$ linear unabhängig, so sind auch die Bilder $F(v_1), \dots, F(v_n)$ linear unabhängig.

Beweis. a) und b) und die „Hinrichtung“ von c) sind ganz klar. Gibt es umgekehrt zwei verschiedene $v, v' \in V$ mit $F(v) = F(v')$, so folgt $F(v - v') = 0$, also $0 \neq v - v' \in \operatorname{Ker} F$.

Zu d) nehmen wir an, daß

$$\mu_1 F(v_1) + \dots + \mu_n F(v_n) = 0.$$

Die linke Seite hat wegen der Linearität das Urbild $\mu_1 v_1 + \dots + \mu_n v_n$, wegen der Injektivität folgt

$$\mu_1 v_1 + \dots + \mu_n v_n = 0,$$

also $\mu_1 = \dots = \mu_n = 0$. □

Eine besonders wichtige Zahl für eine lineare Abbildung ist die Dimension ihres Bildes. Man nennt sie den *Rang*, in Zeichen

$$\operatorname{rang} F := \dim \operatorname{Im} F.$$

Insbesondere beschreibt eine Matrix $A \in M(m \times n; K)$ eine lineare Abbildung

$$A: K^n \rightarrow K^m, \quad x \mapsto y = Ax,$$

wobei die Vektoren $x \in K^n$ und $y \in K^m$ als Spalten geschrieben sind, und als *Rang von A*, in Zeichen

$$\operatorname{rang} A,$$

bezeichnet man den Rang dieser linearen Abbildung. Ist $(e_1, \dots, e_n)$ die kanonische Basis des K^n , so sind

$$Ae_1, \dots, Ae_n$$

die Spalten von A , also ist

$$\operatorname{Im} A = A(K^n) = \operatorname{span}(Ae_1, \dots, Ae_n),$$

das ist der Spaltenraum von A . Also ist der gerade erklärte Rang von A gleich dem in 1.5.8 eingeführten Spaltenrang. Will man ihn berechnen, so genügt es, A auf Spaltenstufenform, d.h. A auf Zeilenstufenform zu bringen (vgl. 1.5.7).

2.2.2. Die Begriffe *Bild* und *Faser* hat man analog für eine beliebige Abbildung $F: X \rightarrow Y$ zwischen Mengen, und X wird durch die Fasern in disjunkte Teilmengen zerlegt:

$$X = \bigcup_{y \in \operatorname{Im} F} F^{-1}(y).$$

Wir wollen untersuchen, wie diese *Faserung* im Fall einer linearen Abbildung aussieht. Dazu zunächst ein einfaches, aber typisches

Beispiel. Wir betrachten die Abbildung

$$F: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} -2x_1 + 2x_2 \\ -x_1 + x_2 \end{pmatrix}.$$

Es ist $\operatorname{Im} F = \mathbb{R} \cdot (2, 1)$, $\operatorname{Ker} F = \mathbb{R} \cdot (1, 1)$, und für $(2b, b) \in \operatorname{Im} F$ ist die Faser die Gerade mit der Gleichung $x_2 = x_1 + b$, also

$$\begin{aligned} F^{-1}(2b, b) &= (0, b) + \mathbb{R} \cdot (1, 1) \\ &= \{(\lambda, b + \lambda) : \lambda \in \mathbb{R}\}. \end{aligned}$$

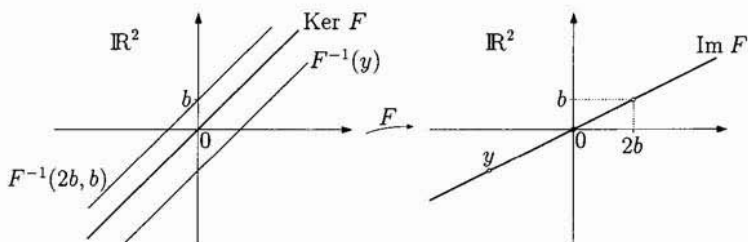


Bild 2.4

Die Fasern sind also parallele Geraden, der Kern ist die einzige Faser durch den Nullpunkt. Allgemein gilt die

Bemerkung. Ist $F: V \rightarrow W$ linear, $w \in \text{Im } F$ und $u \in F^{-1}(w)$ beliebig, so ist

$$F^{-1}(w) = u + \text{Ker } F = \{u + v : v \in \text{Ker } F\}.$$

Beweis. Ist $v' \in F^{-1}(w)$, so folgt

$$\begin{aligned} F(v') &= F(u) \Rightarrow F(v' - u) = 0 \Rightarrow v := v' - u \in \text{Ker } F \\ &\Rightarrow v' = u + v \in u + \text{Ker } F \end{aligned}$$

Ist umgekehrt $v' = u + v \in u + \text{Ker } F$, so ist $F(v') = F(u) = w$, also $v' \in F^{-1}(w)$. $\square$

2.2.3. Teilmengen, die durch „Parallelverschiebung“ eines Untervektorraumes entstehen, erhalten einen eigenen Namen.

Definition. Eine Teilmenge X eines K -Vektorraumes V heißt ein *affiner Unterraum*, falls es ein $v \in V$ und einen Untervektorraum $W \subset V$ gibt, so daß

$$X = v + W := \{u \in V : \text{es gibt ein } w \in W \text{ mit } u = v + w\}$$

(Bild 2.5). Es ist vorteilhaft, auch die leere Menge einen affinen Unterraum zu nennen.

Beispiele für affine Unterräume des $\mathbb{R}^n$ sind Punkte, Geraden und Ebenen (vgl. Kap. 0).

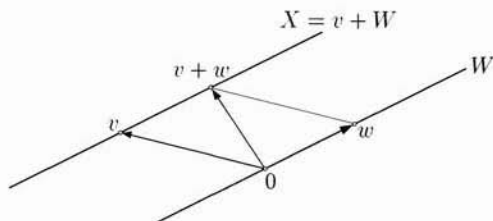


Bild 2.5

Bemerkung. Sei $X = v + W \subset V$ ein affiner Unterraum. Dann gilt:

- Für ein beliebiges $v' \in X$ ist $X = v' + W$.
- Ist $v' \in V$ und $W' \subset V$ ein Untervektorraum mit $v + W = v' + W'$, so folgt $W = W'$ und $v' - v \in W$.

Kurz ausgedrückt: Zu einem affinen Unterraum $v + W$ ist der Untervektorraum W eindeutig bestimmt, und der Aufhängepunkt v kann beliebig in X gewählt werden.

Beweis. a) Wir schreiben $v' = v + w'$.

$$\begin{aligned} X \subset v' + W: \quad u \in X &\Rightarrow u = v + w \quad \text{mit } w \in W \\ &\Rightarrow u = v' + (w - w') \\ &\Rightarrow u \in v' + W \end{aligned}$$

$$v' + W \subset X: \quad u = v' + w \in v' + W \Rightarrow u = v + (w + w') \in v + W.$$

b) Definiert man

$$X - X := \{u - u' : u, u' \in X\}$$

als die Menge der Differenzen (man beachte den Unterschied zu der in 1.1.2 definierten Differenzmenge $X \setminus X = \emptyset$), so sieht man ganz leicht, daß

$$X - X = W \quad \text{und} \quad X - X = W'$$

sein muß. Also ist $W = W'$.

Wegen $v + W = v' + W$ gibt es ein $w \in W$ mit $v' = v + w$. Also ist $v' - v = w \in W$. $\square$

Da für einen affinen Unterraum $X = v + W \subset V$ der Untervektorraum W eindeutig bestimmt ist, können wir durch

$$\dim X := \dim W$$

die *Dimension* von X erklären.

2.2.4. Wir zeigen nun, wie man Basen wählen kann, die maßgeschneidert sind für eine lineare Abbildung (vgl. 2.4.3).

Satz. Sei $F: V \rightarrow W$ linear und $\dim V < \infty$. Sind Basen

$$(v_1, \dots, v_k) \quad \text{von } \operatorname{Ker} F, \quad (w_1, \dots, w_r) \quad \text{von } \operatorname{Im} F,$$

sowie beliebige Vektoren $u_1 \in F^{-1}(w_1), \dots, u_r \in F^{-1}(w_r)$ gegeben, so ist

$$\mathcal{A} := (u_1, \dots, u_r, v_1, \dots, v_k)$$

eine Basis von V . Insbesondere gilt die Dimensionsformel

$$\dim V = \dim \operatorname{Im} F + \dim \operatorname{Ker} F.$$

Beweis. Für $v \in V$ sei

$$F(v) = \mu_1 w_1 + \dots + \mu_r w_r \quad \text{und} \quad v' := \mu_1 u_1 + \dots + \mu_r u_r.$$

Wegen $F(v) = F(v')$ folgt $v - v' \in \operatorname{Ker} F$, also

$$v - v' = \lambda_1 v_1 + \dots + \lambda_k v_k$$

und

$$v = \mu_1 u_1 + \dots + \mu_r u_r + \lambda_1 v_1 + \dots + \lambda_k v_k.$$

Also wird V durch $\mathcal{A}$ erzeugt. Ist

$$\mu_1 u_1 + \dots + \mu_r u_r + \lambda_1 v_1 + \dots + \lambda_k v_k = 0, \quad (*)$$

so folgt durch Anwendung von F

$$\mu_1 w_1 + \dots + \mu_r w_r = 0, \quad \text{also} \quad \mu_1 = \dots = \mu_r = 0,$$

da $w_1, \dots, w_r$ linear unabhängig sind. In $(*)$ eingesetzt ergibt sich

$$\lambda_1 v_1 + \dots + \lambda_k v_k = 0, \quad \text{also} \quad \lambda_1, \dots, \lambda_k = 0$$

da $v_1, \dots, v_k$ linear unabhängig sind. $\square$

Als unmittelbare Folgerung aus der Dimensionsformel notieren wir:

Korollar 1. Ist V endlichdimensional und $F: V \rightarrow W$ linear, so gilt für alle nichtleeren Fasern

$$\dim F^{-1}(w) = \dim V - \dim \operatorname{Im} F. \quad \square$$

Korollar 2. Zwischen zwei endlichdimensionalen Vektorräumen V und W gibt es genau dann einen Isomorphismus, wenn $\dim V = \dim W$. $\square$

Korollar 3. Sei $\dim V = \dim W < \infty$ und $F: V \rightarrow W$ linear. Dann sind folgende Bedingungen gleichwertig:

- i) F injektiv
- ii) F surjektiv
- iii) F bijektiv $\square$

2.2.5. Durch weiteres Spielen mit den Basen aus Satz 2.2.4 erhält man folgenden

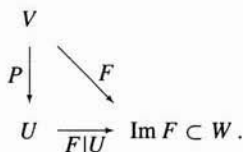
Faktorisierungssatz. Sei $F: V \rightarrow W$ linear und

$$\mathcal{A} = (u_1, \dots, u_r, v_1, \dots, v_k) \quad \text{eine Basis von } V$$

mit $\operatorname{Ker} F = \operatorname{span}(v_1, \dots, v_k)$. Definieren wir $U = \operatorname{span}(u_1, \dots, u_r)$, so gilt

- 1) $V = U \oplus \operatorname{Ker} F$.
- 2) Die Einschränkung $F|_U: U \rightarrow \operatorname{Im} F$ ist ein Isomorphismus.
- 3) Bezeichnet $P: V = U \oplus \operatorname{Ker} F \rightarrow U$, $v = u + v' \mapsto u$, die Projektion auf den ersten Summanden, so ist $F = (F|_U) \circ P$.

In Form eines Diagrammes hat man



Insbesondere hat jede nichtleere Faser $F^{-1}(w)$ mit U genau einen Schnittpunkt, und es ist

$$P(v) = F^{-1}(F(v)) \cap U.$$

Man kann also $F: V \rightarrow W$ zerlegen (oder *faktorisieren*) in drei Anteile: eine *Parallelprojektion*, einen *Isomorphismus* und die *Inklusion* des Bildes. Der zur Konstruktion erforderliche direkte Summand U ist allerdings nicht eindeutig bestimmt, er hängt ab von der Wahl der Basisvektoren $u_1, \dots, u_r$. Wenn in V eine Winkelmessung möglich ist (vgl. Kapitel 5), kann man U eindeutig machen durch die Vorschrift, auf $\text{Ker } F$ senkrecht zu stehen. Die Umkehrung $\text{Im } F \rightarrow U$ von $F|U$ nennt man einen *Schnitt*, da sie aus jeder Faser genau einen Punkt ausschneidet. Als gute Illustration kann Beispiel 2.2.2 mit $k = r = 1$ dienen.

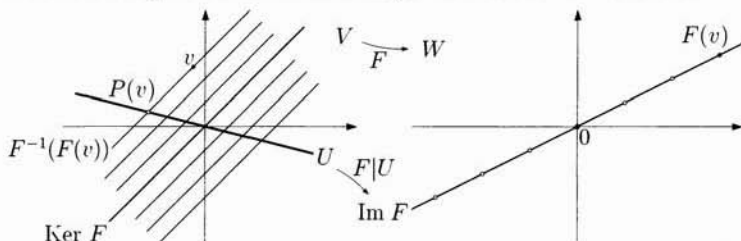


Bild 2.6

Beweis. 1) folgt aus der Charakterisierung direkter Summen in 1.6.3.

Wegen $\text{Ker } F|U = (\text{Ker } F) \cap U = \{0\}$ ist $F|U$ auch injektiv, also Isomorphismus mit Bild $\text{Im } F$. 3) folgt aus der Konstruktion von P . Ist schließlich

$$v \in V \text{ und } v = u + v' \text{ mit } u \in U \text{ und } v' \in \text{Ker } F,$$

so ist $u = P(v)$, also $F(v) = F(u) = F(P(v)) =: w$. Ist überdies $w_1, \dots, w_r$ eine Basis von $\text{Im } F$ mit $F(u_i) = w_i$,

$$w = \mu_1 w_1 + \dots + \mu_r w_r \text{ und } v \in F^{-1}(w) \cap U,$$

so folgt $v = \mu_1 u_1 + \dots + \mu_r u_r$. □

Zur Vorbereitung auf den gleich folgenden Abschnitt über lineare Gleichungssysteme ist es nützlich, die gerade beschriebene allgemeine Situation für eine

durch eine Matrix $A \in M(m \times n; K)$ in Zeilenstufenform gegebene Abbildung

$$A: K^n \rightarrow K^m$$

zu betrachten. Sind (in der Notation von 0.4.3) $j_1, \dots, j_r$ die Indizes der Pivotspalten, und sind $e_{j_1}, \dots, e_{j_r}$ die zu diesen Indizes gehörigen Basisvektoren des K^n , so sind die Bilder

$$A(e_{j_1}), \dots, A(e_{j_r}) \in K^m$$

(das sind gerade die Pivotspalten) eine Basis von

$$\text{Im}(A) = \text{span}(e'_1, \dots, e'_r).$$

Dabei ist mit $(e'_1, \dots, e'_r)$ die kanonische Basis des K^r bezeichnet. Also ist

$$U := \text{span}(e_{j_1}, \dots, e_{j_r})$$

in diesem Fall ein direkter Summand zum Kern von A im Sinn von 1.6.3. Der Leser möge das zur Übung präzise begründen.

Für die erste Lektüre wird empfohlen, den Rest dieses Abschnittes zu überblättern und bei 2.3 wieder einzusteigen.

2.2.6. Ist $F: V \rightarrow W$ eine lineare Abbildung, so sind die Fasern von F nach 2.2.2 die zum Untervektorraum $\text{Ker } F \subset V$ parallelen affinen Räume. Wir wollen nun umgekehrt zu jedem vorgegebenen Untervektorraum $U \subset V$ eine lineare Abbildung mit Kern U konstruieren. Dazu benötigt man man einen Vektorraum W als Bild; wir zeigen, daß es dafür einen kanonischen Kandidaten gibt, den „Quotientenvektorraum“ $W = V/U$. Da die Konstruktion ziemlich abstrakt ist, wollen wir zunächst etwas inhaltlichen Hintergrund bereitstellen.

Beispiel 1. Sei $V = \mathbb{R}^2$ und $U \subset V$ eine Gerade durch den Ursprung. Man nennt zwei Punkte $v, v' \in \mathbb{R}^2$ *äquivalent*, wenn die Differenz in U liegt, in Zeichen

$$v \sim v' \Leftrightarrow v' - v \in U.$$

Geometrisch bedeutet das, daß v und v' *gleich weit entfernt* von U sind, wobei die Entfernung von Punkten links von U negativ und rechts von U positiv gerechnet sein soll.

Es ist ganz einfach zu sehen, daß dadurch in V eine Äquivalenzrelation im Sinn von 1.1.8 erklärt wird. Die Äquivalenzklassen sind die zu U parallelen Geraden, das sind die affinen Räume aus 2.2.3.

Der Leser mache sich auch die Analogie zu den Restklassen modulo m aus 1.2.7 klar: Dort wurde die *Gleichheit* abgeschwächt zur *Kongruenz*, hier wird *gleich* ersetzt durch *gleich weit entfernt*.

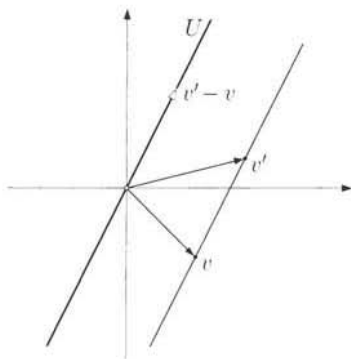


Bild 2.7

Beispiel 2. a) Wir betrachten den unendlich-dimensionalen Vektorraum

$$\mathcal{C}(\mathbb{R}) = \{f: \mathbb{R} \rightarrow \mathbb{R}: f \text{ stetig}\}.$$

Eine beliebige Teilmenge $X \subset \mathbb{R}$ sei vorgegeben, ihr Komplement $A := \mathbb{R} \setminus X$ soll die Rolle einer *Ausnahmemenge* spielen, d.h. die Werte von f auf A werden als unwesentlich angesehen. Damit können wir den Untervektorraum

$$\mathcal{I}(X) := \{f \in \mathcal{C}(\mathbb{R}): f(x) = 0 \text{ für alle } x \in X\} \subset \mathcal{C}(\mathbb{R})$$

der „unwesentlichen“ Funktionen betrachten und für $f, g \in \mathcal{C}(\mathbb{R})$

$$f \sim g : \Leftrightarrow g - f \in \mathcal{I}(X)$$

erklären. In Worten bedeutet das, f und g sind im wesentlichen (d.h. außerhalb A) gleich. Auch diese Äquivalenz ist eine kontrollierte (von A abhängige) Abschwächung der Gleichheit.

b) Eine Variante davon ist die folgende: Man benutzt auf $\mathbb{R}$ ein Integral (etwa das Riemann- oder besser das Lebesgue-Integral), d.h. ein Integral, mit dem möglichst viele Funktionen integrierbar sind. Sei

$$\mathcal{L}(\mathbb{R}) := \{f: \mathbb{R} \rightarrow \mathbb{R}: f \text{ integrierbar}\}$$

und

$$\mathcal{N} := \{f \in \mathcal{L}(\mathbb{R}): \int_{\mathbb{R}} |f(t)| dt = 0\} \subset \mathcal{L}(\mathbb{R}).$$

Nach den Rechenregeln für ein Integral folgt, daß $\mathcal{N} \subset \mathcal{L}(\mathbb{R})$ ein Untervektorraum ist. Man beachte, daß $\mathcal{N}$ unendliche Dimension hat, denn etwa die Funktionen f_i mit $f_i(t) = 0$ für $t \neq i$ und $f_i(i) = 1$ sind für $i \in \mathbb{N}$ in $\mathcal{N}$ linear unabhängig.

Für $f, g \in \mathcal{L}(\mathbb{R})$ bedeutet $f \sim_{\mathcal{N}} g$ dann

$$\int_{\mathbb{R}} |f(t) - g(t)| dt = 0.$$

Dafür sagt man auch, f und g sind „fast überall“ gleich, denn die Menge

$$\{t \in \mathbb{R} : f(t) \neq g(t)\}$$

muß sehr klein sein.

2.2.7. Sei nun ganz allgemein V ein K -Vektorraum und $U \subset V$ ein Untervektorraum. Für $v, v' \in V$ erklären wir die Äquivalenz modulo U

$$v \sim_U v' :\Leftrightarrow v' - v \in U.$$

Aus den Eigenschaften eines Untervektorraumes folgt ganz einfach, daß die Bedingungen für eine Äquivalenzrelation aus 1.1.8 erfüllt sind.

Die Äquivalenzklasse eines $v \in V$ ist gleich dem affinen Unterraum, also

$$\{v' \in V : v' \sim_U v\} = v + U,$$

denn

$$v' \sim_U v \Leftrightarrow v' - v \in U \Leftrightarrow \text{es gibt ein } u \in U \text{ mit } v' = v + u.$$

Die Menge der Äquivalenzklassen wird mit V/U bezeichnet, die kanonische Abbildung sei

$$\varrho: V \rightarrow V/U = \{v + U : v \in V\}, \quad v \mapsto \varrho(v) = v + U.$$

Dabei wird jedem Punkt der ihn enthaltende affine Raum zugeordnet, oder anders ausgedrückt wird jeder Vektor ersetzt durch die Menge all der zu ihm gleichwertigen Vektoren. Im Extremfall $U = 0$ ist die Äquivalenz die Gleichheit und ϱ wird bijektiv. Für $U = V$ ist alles äquivalent, und V/U besteht nur aus einem Element.

Nun kommt der entscheidende Schritt, nämlich die Beobachtung, daß man mit den affinen Räumen rechnen kann wie mit Vektoren.

Satz. Sei V ein K -Vektorraum und $U \subset V$ ein Untervektorraum. Dann kann man die Menge V/U auf genau eine Weise so zu einem K -Vektorraum machen, daß die kanonische Abbildung

$$\varrho: V \rightarrow V/U, \quad v \mapsto v + U,$$

linear wird. Weiter gilt:

1) ϱ ist surjektiv.

2) $\text{Ker } \varrho = U$.

3) $\dim V/U = \dim V - \dim U$, falls $\dim V < \infty$.

4) Der Quotientenvektorraum V/U hat die folgende universelle Eigenschaft:
Ist $F: V \rightarrow W$ eine lineare Abbildung mit $U \subset \text{Ker } F$, so gibt es genau eine lineare Abbildung $\bar{F}: V/U \rightarrow W$ mit $F = \bar{F} \circ \varrho$. Das kann man in Form eines kommutativen Diagramms schreiben:

$$\begin{array}{ccc} V & \xrightarrow{F} & W \\ \varrho \downarrow & \nearrow \bar{F} & \\ V/U & & \end{array}$$

Weiter ist $\text{Ker } \bar{F} = (\text{Ker } F)/U$.

Man nennt V/U den Quotientenvektorraum von V nach U . Diese Bezeichnung entspricht der Vorstellung, daß man U aus V „herausdividiert“, weil U in V/U zur Null wird.

Beweis. Zur vorübergehenden Unterscheidung werden die neu zu definierenden Verknüpfungen in V/U mit $\dot{+}$ und $\cdot$, die alten in V mit $+$ und ohne Symbol bezeichnet. Soll ϱ linear werden, so muß

$$(v + U) \dot{+} (w + U) = \varrho(v) \dot{+} \varrho(w) = \varrho(v + w) = (v + w) + U,$$

$$\lambda \cdot (v + U) = \lambda \cdot \varrho(v) = \varrho(\lambda v) = \lambda v + U$$

gelten. Also gibt es nur eine Möglichkeit, die gesuchten Verknüpfungen in V/U zu erklären:

$$(v + U) \dot{+} (w + U) := (v + w) + U, \quad \lambda \cdot (v + U) := \lambda v + U.$$

Es ist jedoch keineswegs klar, daß diese „Definition“ sinnvoll ist. Man muß noch zeigen, daß sie von der Wahl der Repräsentanten v und w unabhängig ist; dann sagt man, durch $\dot{+}$ und $\cdot$ seien Verknüpfungen in V/U wohldefiniert.

Seien also weitere Repräsentanten v', w' gegeben, d.h.

$$v + U = v' + U \quad \text{und} \quad w + U = w' + U,$$

Dann ist $v' - v \in U$ und $w' - w \in U$, also $(v' + w') - (v + w) \in U$, und somit (siehe Bild 2.8)

$$(v + w) + U = (v' + w') + U.$$

Analog zeigt man, daß

$$\lambda v + U = \lambda v' + U,$$

d.h. auch die Multiplikation mit Skalaren ist in V/U wohldefiniert.

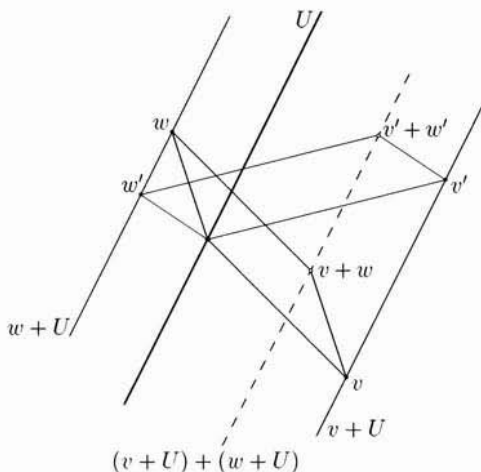


Bild 2.8

Der Nachweis der Vektorraumaxiome in V/U mit Hilfe der entsprechenden Rechenregeln in V bereitet keinerlei Probleme, das sei dem Leser zur Übung empfohlen. Nullvektor in V/U ist U , denn

$$(v + U) \dot{+} U = (v + U) \dot{+} (0 + U) = (v + 0) + U = v + U,$$

und der zu $v + U$ negative Vektor ist $-v + U$. Diese Rechnungen zeigen, daß die Unterscheidung von $+$ und $\dot{+}$ überflüssig ist.

Die zusätzlichen Aussagen sind ganz einfach. 1) folgt aus der Definition von ϱ . Ist $v + U = U$, so ist $v \in U$, also folgt 2). 3) folgt aus der Dimensionsformel in 2.2.4.

Zu 4) bemerkt man zunächst, daß wegen der Forderung $F = \overline{F} \circ \varrho$ für alle $v \in V$

$$F(v) = \overline{F}(\varrho(v)) = \overline{F}(v + U)$$

sein muß. Dadurch ist $\overline{F}$ auch wohldefiniert: denn ist $v + U = v' + U$, so folgt

$$v' - v \in U \subset \text{Ker } F, \quad \text{also} \quad F(v) = F(v').$$

Die Linearität von $\overline{F}$ ist klar. Die Gleichung $\text{Ker } \overline{F} = \text{Ker } F/U$ folgt aus

$$v + U \in \text{Ker } \overline{F} \Leftrightarrow v \in \text{Ker } F \Leftrightarrow v + U \in \text{Ker } F/U,$$

wobei zu bedenken ist, daß $\text{Ker } F/U \subset V/U$ ein Untervektorraum ist. $\square$

Eine Schwierigkeit beim Verständnis der Quotientenstruktur besteht wohl darin, daß Mengen von Vektoren (in diesem Fall affine Räume) zu neuen Vektoren

werden. Aber Vektor zu sein hat keine individuelle Bedeutung; ein Vektor muß sich nur innerhalb einer Gesamtheit von Vektoren (d.h. in einem Vektorraum) nach den dort geltenden Spielregeln (den Axiomen) verhalten. In diesem Sinne ist z.B. auch eine Funktion ein Vektor, d.h. ein Element oder „Punkt“ eines Vektorraumes (vgl. 1.4.1, Beispiel e).

2.2.8. Manchmal mag es beruhigend sein, wenn man einen abstrakten Quotientenvektorraum durch etwas Konkretes ersetzen kann. Dazu betrachten wir noch einmal die Beispiele aus 2.2.6.

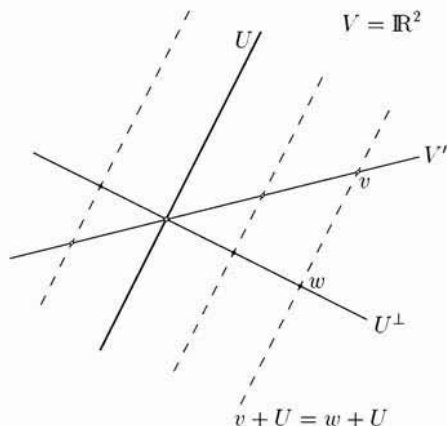


Bild 2.9

Beispiel 1. Für eine Gerade $U \subset V = \mathbb{R}^2$ ist der Quotient V/U eindimensional. Jeder affine Raum $v+U \in V/U$ kann durch einen Repräsentanten $v \in V$ gegeben werden, und man kann die Repräsentanten alle auf einen Streich in folgender Weise erhalten: Ist $V' \subset V$ eine von U verschiedene Gerade durch 0, so schneidet V' jeden affinen Raum $v+U$ in genau einem Punkt (Bild 2.9). Bezeichnet man mit

$$q': V' \rightarrow V/U, \quad v \mapsto v+U,$$

die Beschränkung der kanonischen Abbildung q , so wird q' zu einem Isomorphismus. Man kann also in gewisser Weise den abstrakten Quotientenvektorraum V/U durch einen konkreten Untervektorraum V' ersetzen. V' ist direkter Summand im Sinne von 1.6.3, d.h. es ist

$$V = U \oplus V',$$

und die Umkehrung von q' ist ein Schnitt im Sinn von 2.2.5. Aber V' hat den Nachteil, nicht eindeutig zu sein. Ein besonders ausgezeichneter direkter Summand ist die zu U senkrechte Gerade $U^\perp$ (vgl. dazu 5.4.8).

Daß die elementargeometrische Vorstellung hier nicht immer hilfreich ist, sieht man an

Beispiel 2. a) Die Elemente aus $\mathcal{C}(\mathbb{R})/\mathcal{I}(X)$ sind Klassen auf $\mathbb{R}$ stetiger Funktionen, die auf X gleich sind. Eine solche Klasse kann man stetige Funktion auf X nennen, damit hat man Stetigkeit auch auf nicht-offenen Teilmengen $X \subset \mathbb{R}$ erklärt.

Das geht zum Glück auch etwas weniger abstrakt. Sei

$$\mathcal{F}(X) = \{\varphi: X \rightarrow \mathbb{R}\}$$

der Vektorraum aller auf X definierten Funktionen und

$$\sigma: \mathcal{C}(\mathbb{R}) \rightarrow \mathcal{F}(X), \quad f \mapsto f|X,$$

der Einschränkungshomomorphismus. Wir definieren

$$\mathcal{C}(X) := \text{Im } \sigma = \{\varphi \in \mathcal{F}(X): \text{es gibt ein } f \in \mathcal{C}(\mathbb{R}) \text{ mit } \varphi = f|X\} \subset \mathcal{F}(X)$$

als den Vektorraum der auf X stetigen, d.h. auf $\mathbb{R}$ stetig fortsetzbaren Funktionen. Offenbar ist $\text{Ker } \sigma = \mathcal{I}(X)$, also hat man nach der universellen Eigenschaft des Quotientenvektorraumes ein Diagramm

$$\begin{array}{ccc} \mathcal{C}(\mathbb{R}) & \xrightarrow{\sigma} & \mathcal{F}(X) \\ q \downarrow & \nearrow \bar{\sigma} & \\ \mathcal{C}(\mathbb{R})/\mathcal{I}(X) & & \end{array}$$

wobei $\bar{\sigma}$ wegen $\text{Ker } \bar{\sigma} = \text{Ker } \sigma / \mathcal{I}(X) = 0$ injektiv ist. Der abstrakte Quotientenvektorraum $\mathcal{C}(\mathbb{R})/\mathcal{I}(X)$ kann also als Untervektorraum des konkreteren Vektorraumes $\mathcal{F}(X)$ aufgefaßt werden.

b) Der Quotientenvektorraum

$$L(\mathbb{R}) := \mathcal{C}(\mathbb{R})/\mathcal{N}$$

besteht aus den Klassen fast überall gleicher Funktionen. Im Gegensatz zu a) kann man ihn nicht als Untervektorraum von $\mathcal{F}(\mathbb{R})$ realisieren (warum?). In Aufgabe 6 zu 5.1 wird er mit einer Norm versehen. Das ergibt einen brauchbaren Begriff der Konvergenz; der Preis dafür ist, daß man Funktionen durch Äquivalenzklassen ersetzen muß.

2.2.9. Nach diesen Beispielen wieder zurück zur allgemeinen Theorie. Wir zeigen, daß man den Quotientenvektorraum weitgehend durch einen direkten Sum-

manden ersetzen kann. Dessen Existenz war im endlich-dimensionalen Fall in 1.6.3 gezeigt worden. Im allgemeinen Fall ist das zwar auch noch richtig, aber für die Praxis nutzlos.

Satz. Sei $V = V_1 \oplus V_2$ und $\varrho: V \rightarrow V/V_2$ die kanonische Abbildung. Dann ist

$$\varrho' := \varrho|_{V_1}: V_1 \rightarrow V/V_2$$

ein Isomorphismus.

Beweis. Jedes $v \in V$ hat eine eindeutige Darstellung $v = v_1 + v_2$ mit $v_1 \in V_1$ und $v_2 \in V_2$. Weiter ist

$$\varrho(v) = \varrho(v_1 + v_2) = v_1 + v_2 + V_2 = v_1 + V_2 = \varrho'(v_1).$$

Daraus folgt sofort, daß ϱ' bijektiv ist. □

Aufgaben zu 2.2

1. Sei $F: \mathbb{R}^n \rightarrow \mathbb{R}^m$ gegeben durch die folgenden Matrizen:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \end{pmatrix}.$$

Bestimmen Sie jeweils Basen von $\text{Ker } F$ und $\text{Im } F$.

2. Sei $I \subset \mathbb{R}$ ein Intervall und

$$d: \mathcal{D}(I; \mathbb{R}) \rightarrow \mathcal{D}(I; \mathbb{R}), \quad f \mapsto f'.$$

Zeigen Sie, dass d eine $\mathbb{R}$ -lineare Abbildung ist, und geben Sie eine Basis von $\text{Ker } d$ an. Wie sieht $\text{Ker } d$ aus im Fall, dass I disjunkte Vereinigung von Intervallen ist?

3. Sei V ein endlichdimensionaler Vektorraum und $F: V \rightarrow V$ ein Endomorphismus. Es sei definiert: $W_0 := V$ und $W_{i+1} := F(W_i)$ für $i \in \mathbb{N}$. Dann gilt: Es gibt ein $m \in \mathbb{N}$ mit $W_{m+i} = W_m$ für alle $i \in \mathbb{N}$.

4. Sei $F: V \rightarrow V$ linear mit $F^2 = F$. Zeigen Sie, dass es Untervektorräume U, W von V gibt mit $V = U \oplus W$ und $F(W) = 0, F(u) = u$ für alle $u \in U$.

5. Sei $F: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ gegeben durch die Matrix

$$\begin{pmatrix} 2 & 1 & 3 \\ -4 & -2 & -6 \end{pmatrix}.$$

a) Bestimmen Sie Basen $\mathcal{A} = (u, v_1, v_2)$ des $\mathbb{R}^3$ und $\mathcal{B} = (w, w')$ des $\mathbb{R}^2$, so dass

$$\text{Ker } F = \text{span}(v_1, v_2), \text{Im } F = \text{span}(w) \text{ und } F(u) = w.$$

- b) Geben Sie für $x \in \operatorname{Im} F$ eine Parametrisierung der Faser $F^{-1}(x)$ an und zeigen Sie, dass jede nichtleere Faser $F^{-1}(x)$ genau einen Schnittpunkt mit $U = \operatorname{span}(u)$ hat (vgl. 2.2.5).
6. Beweisen Sie das Lemma aus 1.5.8 noch einmal, aber benutzen Sie nun, dass die Projektion $\pi: W \rightarrow K^{m-1}$ linear und injektiv ist.
7. Sei $F: V \rightarrow W$ linear und $U \subset W$ ein Untervektorraum. Zeigen Sie, dass dann
- $$\dim F^{-1}(U) = \dim(U \cap \operatorname{Im} F) + \dim \operatorname{Ker} F.$$
8. Geben Sie einen neuen Beweis von Teil a) der Bemerkung aus 2.2.3 unter Benutzung der Äquivalenzrelation $\sim_W$ in V .
9. Zeigen Sie mit Hilfe der universellen Eigenschaft des Quotientenvektorraumes, dass für Vektorräume V, W sowie einen Untervektorraum $U \subset V$ die lineare Abbildung
- $$\{F \in \operatorname{Hom}(V, W): F|_U = 0\} \rightarrow \operatorname{Hom}(V/U, W) \quad \text{mit} \quad F \mapsto \bar{F}$$
- (vgl. Satz 2.2.7) ein Isomorphismus von Vektorräumen ist.

2.3 Lineare Gleichungssysteme

Mit Hilfe der bisher entwickelten Techniken von Vektorräumen und linearen Abbildungen können wir nun lineare Gleichungssysteme noch einmal von einem etwas abstrakteren Standpunkt aus behandeln und dabei auch die in Kapitel 0 versprochenen Begründungen nachliefern.

2.3.1. In Kapitel 0 hatten wir nur den reellen Vektorraum $\mathbb{R}^n$ betrachtet. Man kann ihn für einen beliebigen Körper K durch K^n ersetzen, da bei der Behandlung von linearen Gleichungssystemen nur die Körpereigenschaften der reellen Zahlen verwendet wurden.

Wir betrachten jetzt allgemein eine Matrix $A = (a_{ij}) \in M(m \times n; K)$ und eine Spalte $b = {}^t(b_1, \dots, b_m) \in M(m \times 1; K)$. Daraus ergibt sich das Gleichungssystem

$$A \cdot x = b, \quad \text{d.h.} \quad \sum_{j=1}^n a_{ij} x_j = b_i \quad \text{für } i = 1, \dots, m. \quad (**)$$

Man nennt

$$A \cdot x = 0, \quad \text{d.h.} \quad \sum_{j=1}^n a_{ij} x_j = 0 \quad \text{für } i = 1, \dots, m. \quad (*)$$

das zu (**) gehörige *homogene System*; ist $b \neq 0$, so nennt man das System (**) *inhomogen*. Die Mengen

$$\text{Lös}(A, b) := \{x \in K^n : A \cdot x = b\} \subset K^n$$

nennt man *Lösungsräume*. Entscheidend für die Theorie ist die Beziehung zu der durch A beschriebenen linearen Abbildung

$$F: K^n \rightarrow K^m, \quad x \mapsto A \cdot x,$$

denn es ist $\text{Lös}(A, b) = F^{-1}(b)$, also insbesondere $\text{Lös}(A, 0) = \text{Ker } F$.

Nach 2.2 ist die „Größe“ der Lösungsräume festgelegt durch die Zahl

$$r := \text{rang } F = \text{rang } A = \text{Spaltenrang } A.$$

Genauer folgt aus 2.2.3 und 2.2.4 das

Korollar. Gegeben sei das lineare Gleichungssystem $A \cdot x = b$ mit m Gleichungen und n Unbekannten, es sei $r = \text{rang } A$. Dann gilt für die Lösungsräume:

- 1) $\text{Lös}(A, 0) \subset K^n$ ist ein Untervektorraum der Dimension $n - r$.
- 2) $\text{Lös}(A, b) \subset K^n$ ist entweder leer oder ein affiner Raum der Dimension $n - r$.
Ist $v \in \text{Lös}(A, b)$ beliebig, so ist

$$\text{Lös}(A, b) = v + \text{Lös}(A, 0).$$

□

Anders ausgedrückt sagt man dafür: Die *allgemeine Lösung* eines inhomogenen linearen Gleichungssystems erhält man durch die Addition einer *speziellen Lösung* des inhomogenen Gleichungssystems und der *allgemeinen Lösung* des zugehörigen homogenen Gleichungssystems.

2.3.2. Der Lösungsraum eines homogenen linearen Gleichungssystems enthält als Untervektorraum stets die *triviale* Lösung 0, für ein inhomogenes System gibt es keine triviale Lösung. In 0.4.4 hatten wir durch Umformung der erweiterten Koeffizientenmatrix am Ergebnis ablesen können, ob es Lösungen gibt. Auch das kann man etwas abstrakter ansehen.

Wir vergleichen dazu die Matrizen A vom Rang r und (A, b) . Die erweiterte Matrix hat eine Spalte mehr, also ist

$$r \leq \text{rang}(A, b) \leq r + 1.$$

Satz. Der Lösungsraum des linearen Gleichungssystems $A \cdot x = b$ ist genau dann nicht leer, wenn

$$\text{rang } A = \text{rang}(A, b).$$

Diese Bedingung wurde in den Jahren 1875/76 von G. FONTENÉ, E. ROUCHÉ und F.G. FROBENIUS gefunden (vgl. [Fr]).

Beweis. A beschreibt eine lineare Abbildung

$$A: K^n \rightarrow K^m, \quad x \mapsto A \cdot x,$$

und (A, b) beschreibt eine lineare Abbildung

$$A': K^{n+1} \rightarrow K^m, \quad x' \mapsto A' \cdot x'.$$

Sind $(e_1, \dots, e_n)$ und $(e'_1, \dots, e'_{n+1})$ die kanonischen Basen, so gilt

$$A(e_1) = A'(e'_1), \dots, A(e_n) = A'(e'_n) \text{ und } A'(e'_{n+1}) = b.$$

Bei der Abbildung A' kommt also b nach Konstruktion im Bild vor, während das bei A gerade zu entscheiden ist. Wegen $\text{Im } A \subset \text{Im } A'$ ist stets

$$\text{rang } A \leq \text{rang } A'.$$

Also ist $\text{rang } A = \text{rang } A'$ gleichbedeutend mit $\text{Im } A = \text{Im } A'$, d.h. nach Definition von A' mit $b \in \text{Im } A$. $\square$

Hat man (A, b) auf Zeilenstufenform gebracht, so ist der Zeilenrang gleich der Anzahl der von Null verschiedenen Zeilen. Verwenden wir die Gleichheit von Zeilenrang und Spaltenrang (vgl. 1.5.8 und 2.6.6), so folgt die

Bemerkung. Sei

$$(A, b) = \left(\begin{array}{ccc|c} \circledast & & & \\ & \circledast & & \\ & & \ddots & \\ & & & \circledast & \\ 0 & & & & b_{r+1} \\ & & & & \vdots \\ & & & & b_m \end{array} \right)$$

Dann ist $\text{rang } A = \text{rang } (A, b)$ genau dann, wenn

$$b_{r+1} = \dots = b_m = 0.$$

□

2.3.3. Wie im reellen Fall zeigt man, daß sich jede Matrix durch elementare Zeilenumformungen auf Zeilenstufenform $\tilde{A}$ bringen läßt (vgl. 0.4.7), und daß sich der Lösungsraum nicht ändert, wenn man dabei die Spalte b zu $\tilde{b}$ mit umformt, d.h.

$$\text{Lös}(\tilde{A}, \tilde{b}) = \text{Lös}(A, b)$$

(vgl. 0.4.6).

2.3.4. Schließlich betrachten wir noch einmal den Fall, daß die erweiterte Koeffizientenmatrix in Zeilenstufenform und der Lösungsraum $\text{Lös}(A, b)$ nicht leer ist. Nach eventueller Umordnung der Spalten von A können wir annehmen, daß

$$(A, b) = \left(\begin{array}{ccc|c} a_{11} & & & b_1 \\ & a_{22} & & \vdots \\ & & \ddots & \vdots \\ & & & a_{rr} & b_r \\ 0 & & & & b_{r+1} \\ & & & & \vdots \\ & & & & b_m \end{array} \right)$$

mit $a_{ii} \neq 0$ für $i = 1, \dots, r$ und $b_{r+1} = \dots = b_m = 0$.

Die in 0.4.4 konstruierte Parametrisierung müssen wir nun genauer ansehen.

Die dort berechneten Koeffizienten tragen wir in Matrizen ein:

$$D' := (d_{ij}) \in M(r \times r; K), \quad C' := (c_{ij}) \in M(r \times k; K).$$

Diese werden vergrößert zu

$$C := \begin{pmatrix} C' \\ E_k \end{pmatrix} \in M(n \times k; K) \quad \text{und} \quad D := \begin{pmatrix} D' \\ 0 \end{pmatrix} \in M(n \times r; K),$$

das ergibt lineare Abbildungen

$$\varphi: K^r \rightarrow K^n, \quad b \mapsto D \cdot b, \quad \text{und} \quad \Phi_0: K^k \rightarrow K^n, \quad \lambda \mapsto C \cdot \lambda,$$

wobei $b = (b_1, \dots, b_r)$, und nach der in 0.4.4 ausgeführten Rechnung ist

$$\Phi_b: K^k \rightarrow K^n \quad \text{gegeben durch} \quad \Phi_b(\lambda) = \varphi(b) + \Phi_0(\lambda).$$

Nach Konstruktion von Φ_b gilt für alle $b \in K^r$

$$\Phi_b(K^k) \subset \text{Lös}(A, b) = \varphi(b) + \text{Lös}(A, 0).$$

Für $b = 0$ sieht man an der Matrix C , daß Φ_0 injektiv ist. Weiter ist nach 2.3.1 $\dim \text{Lös}(A, 0) = k$, also folgt

$$\Phi_0(K^k) = \text{Lös}(A, 0),$$

und durch Verschiebung um $\varphi(b)$ ergibt sich daraus

$$\Phi_b(K^k) = \text{Lös}(A, b).$$

Man beachte, daß hierbei wieder die Gleichheit von Zeilenrang und Spaltenrang benutzt wird: Die Dimension von $\text{Lös}(A, b)$ ist nach der Dimensionsformel aus 2.2.4 durch den Spaltenrang bestimmt, an der Zeilenstufenform liest man den Zeilenrang ab. Insgesamt ist folgendes bewiesen:

Satz. Sei (A, b) in Zeilenstufenform mit $r = \text{Zeilenrang } A$ und $b \in K^r$. Dann hat die in 0.4.4 konstruierte von b abhängige Parametrisierung

$$\Phi_b: K^{n-r} \rightarrow \text{Lös}(A, b) \subset K^n$$

folgende Eigenschaften:

- 1) $\Phi_0: K^{n-r} \rightarrow \text{Lös}(A, 0) \subset K^n$ ist ein Vektorraumisomorphismus.
- 2) $\Phi_b: K^{n-r} \rightarrow \text{Lös}(A, b) \subset K^n$ ist für jedes $b \in K^r$ bijektiv.
- 3) Es gibt einen Homomorphismus $\varphi: K^r \rightarrow K^n$, so daß für alle $b \in K^r$

$$\Phi_b = \varphi(b) + \Phi_0 \quad \text{und} \quad \text{Lös}(A, b) = \varphi(b) + \text{Lös}(A, 0). \quad \square$$

Damit ist die Frage nach Lösbarkeit von linearen Gleichungssystemen beantwortet und auch die Abhängigkeit der Lösungsmenge von der „rechten Seite“ b explizit beschrieben. Die oben angegebene Abbildung $\varphi: K^r \rightarrow K^n$ ist ein Schnitt im Sinne von 2.2.5.

Ist das System nicht in Zeilenstufenform, so kann man zeigen, daß der Übergang vom ursprünglichen b zum umgeformten $\tilde{b}$ durch einen Isomorphismus

$$S: K^m \rightarrow K^m, \quad b \mapsto \tilde{b},$$

beschrieben wird. In 2.7.7 zeigen wir, wie man S berechnen kann.

2.3.5. Zur Beschreibung der Lösungen eines linearen Gleichungssystems benutzt man oft eine weitere Sprechweise. Eine Basis $(w_1, \dots, w_k)$ von $\text{Lös}(A, 0)$ heißt *Fundamentalsystem von Lösungen des homogenen Systems*. Ein beliebiges $v \in \text{Lös}(A, b)$ heißt *spezielle Lösung des inhomogenen Systems*. Dann hat man eine Darstellung

$$\text{Lös}(A, b) = v + K w_1 + \dots + K w_k,$$

und die Linearkombinationen der Lösungen sind eindeutig. Man erhält die Vektoren $w_1, \dots, w_k$ als Spalten der in 2.3.4 konstruierten Matrix C . Dabei ist jedoch Vorsicht geboten, weil die Zeilen von C anders angeordnet sind, wenn die Pivotspalten nicht die Voraussetzung $j_i = i$ erfüllen.

Es lohnt sich nicht, den allgemeinen Fall aufzuschreiben, wir geben lieber ein typisches

Beispiel. Wir betrachten das Gleichungssystem in Zeilenstufenform und mit beliebiger rechter Seite:

$$\left(\begin{array}{cccccc|c} 0 & 1 & 0 & 2 & -1 & -4 & 0 & b_1 \\ 0 & 0 & 1 & -1 & -1 & 2 & 1 & b_2 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & b_3 \\ 0 & 0 & 0 & 0 & 0 & 1 & 2 & b_4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

mit $n = 7$, $m = 5$, $r = 4$, $j_1 = 2$, $j_2 = 3$, $j_3 = 5$, $j_4 = 6$ und $k = 3$. Dann ist $x_1 = \lambda_1$, $x_4 = \lambda_2$, $x_7 = \lambda_3$, und daraus ergibt sich

$$\begin{aligned} x_6 &= b_4 - 2x_7 = b_4 - 2\lambda_3, \\ x_5 &= b_3 - x_6 = b_3 - b_4 + 2\lambda_3, \\ x_3 &= b_2 + x_4 + x_5 - 2x_6 - x_7 = b_2 + b_3 - 3b_4 + \lambda_2 + 5\lambda_3, \\ x_2 &= b_1 - 2x_4 + x_5 + 4x_6 = b_1 + b_3 + 3b_4 - 2\lambda_2 - 6\lambda_3. \end{aligned}$$

Trägt man die erhaltenen Koeffizienten in die entsprechenden Matrizen ein, so

wird

$$D = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 3 \\ 0 & 1 & 1 & -3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & -6 \\ 0 & 1 & 5 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & -2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Die Spalten von C sind das Fundamentalsystem, $D \cdot b$ ist für jedes $b \in K^4$ eine spezielle Lösung.

2.3.6. Zwei wichtige Spezialfälle linearer Gleichungssysteme haben eigene Namen. Besteht der Lösungsraum aus genau einem Element, so nennt man das System *eindeutig lösbar*. Aus dem bisher Bewiesenen folgt sofort die

Bemerkung. Für $A \in M(m \times n; K)$ und $b \in K^m$ sind folgende Bedingungen gleichwertig:

- i) Das lineare Gleichungssystem $A \cdot x = b$ ist *eindeutig lösbar*.
- ii) $\text{rang } A = \text{rang } (A, b) = n$. □

In diesem Fall besitzt das zugehörige homogene Gleichungssystem $A \cdot x = 0$ nur die *triviale Lösung* 0.

Ist $m = n$, so kann man ii) ersetzen durch $\text{rang } A = n$. Dies bedeutet, daß die lineare Abbildung $A: K^n \rightarrow K^n$ surjektiv, also nach Korollar 3 aus 2.2.4 sogar bijektiv ist. Bezeichnet A^{-1} die inverse Abbildung, so ist die eindeutige Lösung x gegeben durch

$$x = A^{-1}(b).$$

Nach 2.5.5 ist A^{-1} beschrieben durch die inverse Matrix.

Ist die Matrix $A \in M(m \times n; K)$ vom Rang m , so ist die lineare Abbildung $A: K^n \rightarrow K^m$ surjektiv, also ist der Lösungsraum von $A \cdot x = b$ für jedes $b \in K^m$ nicht leer. Ein solches Gleichungssystem nennt man *universell lösbar*.

Ist der Rang von A kleiner als m , so ist das System nur für spezielle b lösbar. Ein Rechenverfahren, dies bei festem A für ein gegebenes b zu entscheiden, leiten wir in 2.7.7 ab.

Aufgaben zu 2.3

1. Ein Nahrungsmittel enthält Schadstoffe $S_1, \dots, S_5$, die bei der Produktion und Lagerung als Bestandteile von Pflanzenschutzmitteln auftreten. Auf den einzelnen Stationen werden die folgenden Pflanzenschutzmittel benutzt:

Station	Mittel
1. Landwirt	A
2. Rohproduktlagerung	B
3. Veredelungsbetrieb	C
4. Grossist und Transport	D
5. Einzelhändler	E

Die folgende Tabelle gibt die prozentuale Zusammensetzung der Mittel A, ..., E wieder:

	S_1	S_2	S_3	S_4	S_5
A	0.2	0.5	0	0.3	0
B	0.1	0.6	0.3	0	0
C	0.1	0.2	0.2	0.3	0.2
D	0	0	0.1	0.4	0.5
E	0	0.1	0.3	0.3	0.3

Für das fertige Produkt ergibt die Nahrungsmittelanalyse die folgenden Werte (in Gewichtseinheiten):

S_1	S_2	S_3	S_4	S_5
0.75	2.25	0.65	1.60	0.75

Ermitteln Sie, wieviel (in Gewichtseinheiten) die einzelnen Stationen zur Schadstoffbelastung beitragen.

2. Es seien Metall-Legierungen M_1, M_2 und M_3 gegeben, die alle Kupfer, Silber und Gold enthalten, und zwar in folgenden Prozentsätzen:

	Kupfer	Silber	Gold
M_1	20	60	20
M_2	70	10	20
M_3	50	50	0

Kann man diese Legierungen so mischen, dass eine Legierung entsteht, die 40% Kupfer, 50% Silber und 10% Gold enthält?

3. Zeigen Sie: Ist die Matrix $A \in M(m \times n; K)$ in Zeilenstufenform und r der Rang von A , so ist $(e_1, \dots, e_r)$ eine Basis von $\text{Im } A \subset K^m$.

4. Bestimmen Sie für das folgende Gleichungssystem in Zeilenstufenform mit beliebiger rechter Seite Matrizen C und D wie in 2.3.4, so dass die Spalten von C ein Fundamentalsystem bilden und $D \cdot b$ für jedes $b \in \mathbb{R}^5$ eine spezielle Lösung ist.

$$\left(\begin{array}{cccccc|c} 0 & 1 & -1 & 2 & 0 & 3 & b_1 \\ 0 & 0 & 2 & -1 & 2 & 0 & b_2 \\ 0 & 0 & 0 & -1 & 4 & 0 & b_3 \\ 0 & 0 & 0 & 0 & 0 & -7 & b_4 \\ 0 & 0 & 0 & 0 & 0 & 0 & b_5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

5. Gegeben seien die Matrizen

$$A = \begin{pmatrix} 3 & 5 & 7 \\ 4 & 6 & 8 \\ 1 & 3 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 3 & 2 & 6 & 3 \\ 2 & 1 & 3 & 2 \\ 2 & 3 & 1 & 4 \end{pmatrix}.$$

a) Untersuchen Sie die folgenden Gleichungssysteme darauf, ob sie eindeutig lösbar sind:

$$Ax = \begin{pmatrix} 2 \\ 4 \\ 9 \end{pmatrix}, \quad Bx = \begin{pmatrix} 4 \\ 1 \\ 7 \end{pmatrix}.$$

b) Untersuchen Sie die Gleichungssysteme $Ax = b$ und $Bx = b$ für beliebige $b \in \mathbb{R}^3$ darauf, ob sie universell lösbar sind.

6. Sei der Untervektorraum $W \subset \mathbb{R}^n$ gegeben durch m lineare Gleichungen $\varphi_1, \dots, \varphi_m$, d. h.

$$W = \{x \in \mathbb{R}^n : \varphi_1(x) = \dots = \varphi_m(x) = 0\}.$$

Zeigen Sie, dass dann W bereits durch eine einzige (nicht notwendig lineare) Gleichung beschrieben werden kann. Genauer gilt: Es existiert ein Polynom $f \in \mathbb{R}[t_1, \dots, t_n]$ mit

$$W = \{(x_1, \dots, x_n) \in \mathbb{R}^n : f(x_1, \dots, x_n) = 0\}.$$

Zeigen Sie, dass diese Aussage auch gilt, falls man $\mathbb{R}$ durch einen endlichen Körper K ersetzt.

7. Finden Sie neue (kürzere) Beweise für Satz 0.2.4 und Aufgabe 2a) zu 0.3.

8. Zeigen Sie, dass eine Teilmenge L des $\mathbb{R}^3$ eine Gerade ist (d. h. es existieren $v, w \in \mathbb{R}^3$, $w \neq 0$, mit $L = v + \mathbb{R}w$) genau dann, wenn es eine Matrix $A \in M(2 \times 3; \mathbb{R})$ mit $\text{rang } A = 2$ und ein $b \in \mathbb{R}^2$ gibt, so dass $L = \{x \in \mathbb{R}^3 : Ax = b\}$. Was bedeutet das geometrisch?

2.4 Lineare Abbildungen und Matrizen

Eine Abbildung $F: X \rightarrow Y$ von Mengen ist nach Definition eine Vorschrift, die jedem Argument $x \in X$ einen Wert $F(x) \in Y$ zuordnet. Dabei ist im allgemeinen nichts darüber ausgesagt, wie die Vorschrift auszusehen hat, also kann man die Bilder verschiedener Argumente völlig unabhängig voneinander wählen.

Ganz anders ist die Situation für eine lineare Abbildung $F: V \rightarrow W$ zwischen Vektorräumen. Kennt man einen Wert $F(v)$, so ist F auf der ganzen Geraden Kv festgelegt. Will man für einen weiteren Vektor $v' \in V$ den Wert beliebig vorschreiben, so darf also v' nicht auf der Geraden Kv liegen. Durch $F(v)$ und $F(v')$ ist dann F auf der ganzen Ebene $Kv + Kv'$ festgelegt, usw.

2.4.1. Die Frage, durch wieviele Vorgaben eine lineare Abbildung festgelegt ist, hat eine einfache Antwort:

Satz. Gegeben seien endlichdimensionale Vektorräume V und W , sowie Vektoren $v_1, \dots, v_r \in V$ und $w_1, \dots, w_r \in W$. Dann gilt

- 1) Sind $v_1, \dots, v_r$ linear unabhängig, so gibt es mindestens eine lineare Abbildung

$$F: V \rightarrow W \quad \text{mit} \quad F(v_i) = w_i \quad \text{für} \quad i = 1, \dots, r.$$

- 2) Ist $(v_1, \dots, v_r)$ eine Basis, so gibt es genau eine lineare Abbildung

$$F: V \rightarrow W \quad \text{mit} \quad F(v_i) = w_i \quad \text{für} \quad i = 1, \dots, r.$$

Dieses F hat folgende Eigenschaften:

- a) $\text{Im } F = \text{span}(w_1, \dots, w_r)$.
 b) F injektiv $\Leftrightarrow w_1, \dots, w_r$ linear unabhängig.

Beweis. Wir beginnen mit Teil 2). Jedes $v \in V$ hat eine eindeutige Darstellung

$$v = \lambda_1 v_1 + \dots + \lambda_r v_r,$$

wegen $F(v_i) = w_i$ und der Linearität von F muß also

$$F(v) = \lambda_1 w_1 + \dots + \lambda_r w_r \tag{*}$$

sein. Also gibt es höchstens ein solches F , nämlich das durch (*) erklärte. Man darf nun allerdings nicht versäumen zu zeigen, daß die durch (*) erklärte Abbildung wirklich linear ist. Das folgt aus den Rechnungen

$$\begin{aligned} F(v + v') &= F(\lambda_1 v_1 + \dots + \lambda_r v_r + \lambda'_1 v_1 + \dots + \lambda'_r v_r) \\ &= F((\lambda_1 + \lambda'_1)v_1 + \dots + (\lambda_r + \lambda'_r)v_r) \\ &= (\lambda_1 + \lambda'_1)w_1 + \dots + (\lambda_r + \lambda'_r)w_r \\ &= \lambda_1 w_1 + \dots + \lambda_r w_r + \lambda'_1 w_1 + \dots + \lambda'_r w_r \\ &= F(v) + F(v') \end{aligned}$$

und

$$F(\lambda v) = F(\lambda \lambda_1 v_1 + \dots + \lambda \lambda_r v_r) = \lambda \lambda_1 w_1 + \dots + \lambda \lambda_r w_r = \lambda F(v).$$

Die Inklusion $\text{Im } F \subset \text{span}(w_1, \dots, w_r)$ ist klar. Ist umgekehrt

$$w = \mu_1 w_1 + \dots + \mu_r w_r, \quad \text{so folgt} \quad w = F(\mu_1 v_1 + \dots + \mu_r v_r).$$

Zu b) nehmen wir an, $w_1, \dots, w_r$ sei linear abhängig. Dann gibt es $(\mu_1, \dots, \mu_r) \neq (0, \dots, 0)$ mit

$$\mu_1 w_1 + \dots + \mu_r w_r = 0,$$

und es folgt $F(\mu_1 v_1 + \dots + \mu_r v_r) = 0$; also ist F nicht injektiv. Umgekehrt sei $F(v) = 0$. Wir schreiben

$$v = \lambda_1 v_1 + \dots + \lambda_r v_r, \quad \text{dann ist} \quad \lambda_1 w_1 + \dots + \lambda_r w_r = 0.$$

Wegen der linearen Unabhängigkeit von $w_1, \dots, w_r$ folgt $\lambda_1 = \dots = \lambda_r = 0$, also $v = 0$. Damit ist 2) bewiesen.

Ist $v_1, \dots, v_r$ nun linear unabhängig, so können wir diese Familie zu einer Basis

$$(v_1, \dots, v_r, v_{r+1}, \dots, v_n)$$

ergänzen und durch Vorgabe beliebiger weiterer Werte $w_{r+1}, \dots, w_n$ entsprechend 2) ein F mit

$$F(v_i) = w_i \quad \text{für } i = 1, \dots, n$$

finden. An dieser Konstruktion kann man erkennen, wie weit F von der Eindeutigkeit entfernt ist: ein Maß dafür ist die Zahl $n - r$. $\square$

2.4.2. Der Satz aus 2.4.1 über die *Erzeugung von linearen Abbildungen* hat zahlreiche Folgerungen.

Korollar 1. Ist V ein Vektorraum mit einer Basis $\mathcal{B} = (v_1, \dots, v_n)$, so gibt es dazu genau einen Isomorphismus

$$\Phi_{\mathcal{B}}: K^n \rightarrow V \quad \text{mit} \quad \Phi_{\mathcal{B}}(e_j) = v_j \quad \text{für } j = 1, \dots, n,$$

wobei $(e_1, \dots, e_n)$ die kanonische Basis von K^n bezeichnet. $\square$

$\Phi_{\mathcal{B}}$ heißt *Koordinatensystem*, damit werden wir uns in 2.6.1 weiter beschäftigen.

Korollar 2. Zu jeder linearen Abbildung $F: K^n \rightarrow K^m$ gibt es genau eine Matrix $A \in M(m \times n; K)$, so daß

$$F(x) = A \cdot x$$

für alle Spaltenvektoren $x \in K^n$.

Man braucht also in diesem Fall zwischen linearen Abbildungen und Matrizen nicht mehr zu unterscheiden.

Beweis. Man schreibe $F(e_1), \dots, F(e_n)$ als Spaltenvektoren nebeneinander, das ergibt A . $\square$

Einen solchen Zusammenhang zwischen linearen Abbildungen und Matrizen gibt es nicht nur in den Standardräumen:

Satz. Gegeben seien K -Vektorräume

V mit Basis $\mathcal{A} = (v_1, \dots, v_n)$ und W mit Basis $\mathcal{B} = (w_1, \dots, w_m)$.

Dann gibt es zu jeder linearen Abbildung $F: V \rightarrow W$ genau eine Matrix $A = (a_{ij}) \in M(m \times n; K)$, so daß

$$F(v_j) = \sum_{i=1}^m a_{ij} w_i \quad \text{für } j = 1, \dots, n, \quad (*)$$

und die so erhaltene Abbildung

$$M_{\mathcal{B}}^{\mathcal{A}}: \text{Hom}(V, W) \rightarrow M(m \times n; K), \quad F \mapsto A = M_{\mathcal{B}}^{\mathcal{A}}(F),$$

ist ein Isomorphismus von K -Vektorräumen. Insbesondere gilt

$$M_{\mathcal{B}}^{\mathcal{A}}(F+G) = M_{\mathcal{B}}^{\mathcal{A}}(F) + M_{\mathcal{B}}^{\mathcal{A}}(G) \quad \text{und} \quad M_{\mathcal{B}}^{\mathcal{A}}(\lambda F) = \lambda M_{\mathcal{B}}^{\mathcal{A}}(F).$$

Kurz gesagt: Nach Wahl fester Basen kann man lineare Abbildungen durch Matrizen (d.h. relativ abstrakte durch konkrete Objekte) ersetzen. Man bezeichnet $M_{\mathcal{B}}^{\mathcal{A}}(F)$ als die Matrix, die F bezüglich der Basen $\mathcal{A}$ und $\mathcal{B}$ darstellt.

Beweis. Da $\mathcal{B}$ Basis ist, sind die Linearkombinationen aus (*) und somit auch die Spalten der Matrix A eindeutig bestimmt. Gehört zur Abbildung G die Matrix $B = (b_{ij})$, so ist

$$(F+G)(v_j) = F(v_j) + G(v_j) = \sum_{i=1}^m a_{ij} w_i + \sum_{i=1}^m b_{ij} w_i = \sum_{i=1}^m (a_{ij} + b_{ij}) w_i,$$

und für $\lambda \in K$ ist

$$(\lambda F)(v_j) = \lambda \cdot F(v_j) = \lambda \sum_{i=1}^m a_{ij} w_i = \sum_{i=1}^m (\lambda a_{ij}) w_i.$$

Daher ist die Abbildung $M_{\mathcal{B}}^{\mathcal{A}}$ linear. Da $\mathcal{A}$ eine Basis ist, gibt es nach 2.4.1 genau ein F , das darauf die durch Bedingung (*) festgelegten Werte annimmt. Also ist $M_{\mathcal{B}}^{\mathcal{A}}$ bijektiv. $\square$

Zusatz. Im Spezialfall $V = K^n$ und $W = K^m$ mit den kanonischen Basen $\mathcal{K}$ und $\mathcal{K}'$ ist der kanonische Isomorphismus

$$M_{\mathcal{K}'}^{\mathcal{K}}: \text{Hom}(K^n, K^m) \rightarrow M(m \times n; K), \quad F \mapsto A,$$

die in Korollar 2 beschriebene Beziehung. □

Mit Hilfe einer Basis kann man Vektoren eindeutig als Linearkombinationen darstellen. Der obige Satz zeigt, wie man Abbildungen als Vektoren betrachten und mit Hilfe zweier Basen analog verfahren kann. Dazu sei

$$F_i^j: V \rightarrow W \quad \text{erklärt durch} \quad F_i^j(v_k) := \begin{cases} w_i & \text{für } k = j, \\ 0 & \text{sonst.} \end{cases}$$

Dann ist $M_{\mathcal{B}}^{\mathcal{A}}(F_i^j) = E_j^i$ (mit der Bezeichnung aus 1.5.1), und die $m \cdot n$ Abbildungen F_i^j bilden eine Basis von $\text{Hom}(V, W)$. Die zur Linearkombination eines beliebigen F nötigen Skalare stehen an passender Stelle in $M_{\mathcal{B}}^{\mathcal{A}}(F)$.

Die naheliegende Frage, wie sich die Matrix A ändert, wenn man in V und W neue Basen einführt, wird in 2.6.5 beantwortet.

2.4.3. Als Folgerung aus 2.2.4 erhält man, daß bei Benutzung der dort konstruierten Basen auch die darstellende Matrix besonders einfach wird.

Korollar. Sei $F: V \rightarrow W$ linear, $n = \dim V$, $m = \dim W$ und $r = \dim \text{Im } F$. Dann gibt es Basen $\mathcal{A}$ von V und $\mathcal{B}$ von W , so daß

$$M_{\mathcal{B}}^{\mathcal{A}}(F) = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Dabei bezeichnet E_r die in 1.5.7 eingeführte r -reihige Einheitsmatrix.

Beweis. Es genügt, die in 2.2.4 gewählte Basis von $\text{Im } F$ zu einer Basis

$$\mathcal{B} = (w_1, \dots, w_r, w_{r+1}, \dots, w_m)$$

von W zu ergänzen. □

2.4.4. Ist der Bildraum W gleich dem Urbildraum V (d.h. hat man einen Endomorphismus), so setzt man am besten auch $\mathcal{A} = \mathcal{B}$ und zur Vereinfachung der Notation $M_{\mathcal{B}} := M_{\mathcal{B}}^{\mathcal{B}}$, sowie $\text{End}(V) = \text{Hom}(V, V)$. Der Vektorraumisomorphismus

$$M_{\mathcal{B}}: \text{End}(V) \rightarrow M(n \times n; K)$$

ist dann charakterisiert durch die Gleichungen

$$F(v_j) = \sum_{i=1}^n a_{ij} v_i \quad \text{für } j = 1, \dots, n,$$

wenn $\mathcal{B} = (v_1, \dots, v_n)$ und $A = (a_{ij}) = M_{\mathcal{B}}(F)$.

Die n -reihige Einheitsmatrix $E_n = (\delta_{ij})$ beschreibt dabei die identische Abbildung, in Zeichen

$$M_{\mathcal{B}}(\text{id}_V) = E_n.$$

Die zu 2.4.3 analoge Frage, für einen Endomorphismus *eine* besonders einfache Basis zu finden, ist weit schwieriger zu beantworten. Sie ist Gegenstand von Kapitel 4.

Aufgaben zu 2.4

1. Gibt es eine lineare Abbildung $F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ mit

$$F(2, 0) = (0, 1), \quad F(1, 1) = (5, 2), \quad F(1, 2) = (2, 3) ?$$

2. Sei $\mathcal{B} = (\sin, \cos, \sin \cdot \cos, \sin^2, \cos^2)$ und $V = \text{span } \mathcal{B} \subset \text{Abb}(\mathbb{R}, \mathbb{R})$. Betrachten Sie den Endomorphismus $F: V \rightarrow V$, $f \mapsto f'$, wobei f' die erste Ableitung von f bezeichnet.

a) Zeigen Sie, dass $\mathcal{B}$ eine Basis von V ist.

b) Bestimmen Sie die Matrix $M_{\mathcal{B}}(F)$.

c) Bestimmen Sie Basen von $\text{Ker } F$ und $\text{Im } F$.

3. Für $n \in \mathbb{N}$ sei $V_n = \text{span}(1, \dots, t^n) \subset \mathbb{R}[t]$ mit der Basis $\mathcal{B}_n = (1, \dots, t^n)$ und

$$\mathcal{D}_n: V_n \rightarrow V_{n-1}, \quad f \mapsto f'$$

der Ableitungshomomorphismus.

a) Bestimmen Sie die Matrix $M_{\mathcal{B}_{n-1}}^{\mathcal{B}_n}(\mathcal{D}_n)$.

b) Zeigen Sie, dass es eine lineare Abbildung $\mathcal{I}_n: V_{n-1} \rightarrow V_n$ gibt mit $\mathcal{D}_n \circ \mathcal{I}_n = \text{id}$, und bestimmen Sie $M_{\mathcal{B}_n}^{\mathcal{B}_{n-1}}(\mathcal{I}_n)$.

4. Sei $V = \{f \in \mathbb{R}[t]: \deg f \leq 3\}$ mit der Basis $\mathcal{B} = (1, t, t^2, t^3)$. Wir betrachten die linearen Abbildungen

$$F: V \rightarrow \mathbb{R}, \quad f \mapsto \int_{-1}^1 f(t) dt \quad \text{und} \quad G: V \rightarrow \mathbb{R}^3, \quad f \mapsto (f(-1), f(0), f(1)).$$

a) Es seien $\mathcal{K}$ und $\mathcal{K}'$ die kanonischen Basen von $\mathbb{R}$ und $\mathbb{R}^3$. Bestimmen Sie die Matrizen

$$M_{\mathcal{K}}^{\mathcal{B}}(F) \text{ und } M_{\mathcal{K}'}^{\mathcal{B}}(G).$$

b) Zeigen Sie: $\text{Ker } G \subset \text{Ker } F$.

c) Es gibt eine lineare Abbildung $H: \mathbb{R}^3 \rightarrow \mathbb{R}$ mit $H \circ G = F$.

5. Seien V und W endlichdimensionale Vektorräume mit $V = V_1 \oplus V_2$, $W = W_1 \oplus W_2$ sowie $F: V \rightarrow W$ linear mit $F(V_i) \subset W_i$ für $i = 1, 2$. Zeigen Sie, dass es Basen $\mathcal{A}$ von V und $\mathcal{B}$ von W gibt mit

$$M_{\mathcal{B}}^{\mathcal{A}}(F) = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix},$$

wobei $A \in M(\dim W_1 \times \dim V_1; K)$, $B \in M(\dim W_2 \times \dim V_2; K)$.

6. Zeigen Sie ohne Verwendung von Matrizen, dass die in 2.4.2 definierten Abbildungen $F_i^j: V \rightarrow W$ eine Basis von $\text{Hom}(V, W)$ bilden.

7. Sei

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

und $F: \mathbb{R}^4 \rightarrow \mathbb{R}^3$ die durch $F(x) = Ax$ definierte lineare Abbildung. Bestimmen Sie Basen $\mathcal{A}$ von $\mathbb{R}^4$ und $\mathcal{B}$ von $\mathbb{R}^3$ mit

$$M_{\mathcal{B}}^{\mathcal{A}}(F) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

8. Sei V ein endlichdimensionaler Vektorraum und $F: V \rightarrow V$ linear mit $F^2 = F$. Zeigen Sie, dass es eine Basis $\mathcal{B}$ von V gibt mit

$$M_{\mathcal{B}}(F) = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Hinweis: Aufgabe 5 und Aufgabe 4 zu 2.2.

9. Zeigen Sie: Ist $F: V \rightarrow V$ ein Endomorphismus des endlichdimensionalen Vektorraums V mit $\dim \text{Fix } F = r$ (vgl. Aufgabe 3 zu 2.1), so existiert eine Basis $\mathcal{B}$ von V mit

$$M_{\mathcal{B}}(F) = \begin{pmatrix} E_r & * \\ 0 & * \end{pmatrix}.$$

2.5 Multiplikation von Matrizen

2.5.1. In 2.1.3 hatten wir gesehen, daß für lineare Abbildungen $G: U \rightarrow V$ und $F: V \rightarrow W$ die Komposition $F \circ G: U \rightarrow W$ wieder linear ist. Im Spezialfall $U = K^r$, $V = K^n$, $W = K^m$ sind F und G entsprechend 2.4.2 durch Matrizen A und B gegeben, und wir wollen ausrechnen, welche Matrix C die lineare Abbildung $F \circ G$ beschreibt. Das folgende Diagramm soll die Übersicht erleichtern:

$$\begin{array}{ccccc}
 K^r & \xrightarrow{B} & K^n & \xrightarrow{A} & K^m, \\
 x = \begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix} & \mapsto & y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} & \mapsto & z = \begin{pmatrix} z_1 \\ \vdots \\ z_m \end{pmatrix}.
 \end{array}$$

Ist $B = (b_{jk}) \in M(n \times r; K)$ und $y = B(x)$, so folgt

$$y_j = b_{j1}x_1 + \dots + b_{jr}x_r \quad \text{für } j = 1, \dots, n, \quad (\text{b})$$

$A = (a_{ij}) \in M(m \times n; K)$ und $z = A(y)$ bedeutet

$$z_i = a_{i1}y_1 + \dots + a_{in}y_n \quad \text{für } i = 1, \dots, m, \quad (\text{a})$$

und schließlich $C = (c_{ik}) \in M(m \times r; K)$, $z = A(B(x))$ bedeutet

$$z_i = c_{i1}x_1 + \dots + c_{ir}x_r \quad \text{für } i = 1, \dots, m. \quad (\text{c})$$

Setzt man (b) in (a) ein, so erhält man

$$\begin{aligned}
 z_i &= a_{i1}(b_{11}x_1 + \dots + b_{1r}x_r) + \dots + a_{in}(b_{n1}x_1 + \dots + b_{nr}x_r) \\
 &= (a_{i1}b_{11} + \dots + a_{in}b_{n1})x_1 + \dots + (a_{i1}b_{1r} + \dots + a_{in}b_{nr})x_r.
 \end{aligned} \quad (\text{c}')$$

Vergleich von (c) und (c') ergibt

$$c_{i1} = a_{i1}b_{11} + \dots + a_{in}b_{n1}, \dots, c_{ir} = a_{i1}b_{1r} + \dots + a_{in}b_{nr}.$$

Unter Verwendung von Summenzeichen schreibt sich das so:

$$\begin{aligned}
 z_i &= \sum_{j=1}^n a_{ij}y_j = \sum_{j=1}^n a_{ij} \left(\sum_{k=1}^r b_{jk}x_k \right) = \sum_{j=1}^n \left(\sum_{k=1}^r a_{ij}b_{jk}x_k \right) \\
 &= \sum_{k=1}^r \left(\sum_{j=1}^n a_{ij}b_{jk} \right) x_k = \sum_{k=1}^r c_{ik}x_k,
 \end{aligned}$$

also ist

$$c_{ik} = \sum_{j=1}^n a_{ij}b_{jk} \quad \text{für } i = 1, \dots, m \text{ und } k = 1, \dots, r.$$

Diese kleine Rechnung, die nur auf der Umordnung einer Summe beruht, hat wichtige Konsequenzen. Man kann damit eine Multiplikation von Matrizen passender Größe erklären.

2.5.2. Zur Definition der *Multiplikation von Matrizen* nehmen wir

$$A = (a_{ij}) \in M(m \times n; K) \quad \text{und} \quad B = (b_{jk}) \in M(n \times r; K),$$

d.h. die Spaltenzahl von A muß mit der Zeilenzahl von B übereinstimmen. Dann ist das Produkt

$$A \cdot B = (c_{ik}) \in M(m \times r; K) \quad \text{erklärt durch} \quad c_{ik} := \sum_{j=1}^n a_{ij} b_{jk}.$$

Die Matrix $A \cdot B$ hat also so viele Zeilen wie A und so viele Spalten wie B , die gemeinsame Zahl n verschwindet bei der Multiplikation. Diese Größenverhältnisse kann man durch folgendes Schema zum Ausdruck bringen:

$$\underbrace{\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ \boxed{a_{i1}} & \cdots & \boxed{a_{in}} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}}_n \cdot \underbrace{\begin{pmatrix} b_{11} & \cdots & \boxed{b_{1k}} & \cdots & b_{1r} \\ \vdots & & \vdots & & \vdots \\ b_{n1} & \cdots & \boxed{b_{nk}} & \cdots & b_{nr} \end{pmatrix}}_r \left. \begin{matrix} \left. \begin{pmatrix} c_{11} & \cdots & c_{1r} \\ \vdots & \boxed{c_{ik}} & \vdots \\ c_{m1} & \cdots & c_{mr} \end{pmatrix} \right\} m \\ \left. \begin{pmatrix} b_{11} & \cdots & \boxed{b_{1k}} & \cdots & b_{1r} \\ \vdots & & \vdots & & \vdots \\ b_{n1} & \cdots & \boxed{b_{nk}} & \cdots & b_{nr} \end{pmatrix} \right\} n \end{matrix} \right\}$$

Hieran sieht man auch gut, wie

$$c_{ik} = a_{i1}b_{1k} + \cdots + a_{in}b_{nk}$$

aus der i -ten Zeile von A und der k -ten Spalte von B entsteht. Ansonsten ist es aber recht unpraktisch, Matrizen bei der Multiplikation so anzuschreiben. Ein Beispiel der üblichen Schreibweise ist

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

Ist speziell $m = r = 1$ und n beliebig, so ist

$$A \cdot B = (a_1 \dots a_n) \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = a_1 b_1 + \dots + a_n b_n \in K = M(1 \times 1; K).$$

Man beachte, daß dagegen

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \cdot (b_1 \dots b_n) = \begin{pmatrix} a_1 b_1 & \dots & a_1 b_n \\ \vdots & & \vdots \\ a_n b_1 & \dots & a_n b_n \end{pmatrix} \in M(n \times n; K)$$

gilt.

Ist speziell $m = n = r$, so kann man für $A, B \in M(m \times m; K)$ sowohl $A \cdot B$ als auch $B \cdot A$ bilden, im allgemeinen ist aber $A \cdot B \neq B \cdot A$. Zum Beispiel ist

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

aber

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

Daran sieht man außerdem, daß das Produkt von zwei Matrizen die Nullmatrix ergeben kann, obwohl beide Matrizen von der Nullmatrix verschieden waren.

Beispiel. Ist $f = (f_1, \dots, f_m): \mathbb{R}^n \rightarrow \mathbb{R}^m$ eine differenzierbare Abbildung (das heißt $f_1, \dots, f_m: \mathbb{R}^n \rightarrow \mathbb{R}$ sind differenzierbare Funktionen) mit $f(0) = 0$ (diese Annahme dient nur zur Vereinfachung der Bezeichnungen), und sind $y_1, \dots, y_n$ die Koordinaten im $\mathbb{R}^n$, so sei

$$A := \begin{pmatrix} \frac{\partial f_1}{\partial y_1}(0) & \dots & \frac{\partial f_1}{\partial y_n}(0) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial y_1}(0) & \dots & \frac{\partial f_m}{\partial y_n}(0) \end{pmatrix}$$

die sogenannte **JACOBI-Matrix** von f im Punkte Null.

Ist $g = (g_1, \dots, g_r): \mathbb{R}^r \rightarrow \mathbb{R}^n$ eine weitere differenzierbare Abbildung mit $g(0) = 0$, und sind $x_1, \dots, x_r$ Koordinaten im $\mathbb{R}^r$, so bezeichnen wir mit

$$B := \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(0) & \dots & \frac{\partial g_1}{\partial x_r}(0) \\ \vdots & & \vdots \\ \frac{\partial g_n}{\partial x_1}(0) & \dots & \frac{\partial g_n}{\partial x_r}(0) \end{pmatrix}$$

die Jacobimatrix von g im Punkte Null. Ist $h := f \circ g: \mathbb{R}^r \rightarrow \mathbb{R}^m$ und $h = (h_1, \dots, h_m)$, so gilt für die Jacobimatrix von h im Punkt Null

$$\begin{pmatrix} \frac{\partial h_1}{\partial x_1}(0) & \dots & \frac{\partial h_1}{\partial x_r}(0) \\ \vdots & & \vdots \\ \frac{\partial h_m}{\partial x_1}(0) & \dots & \frac{\partial h_m}{\partial x_r}(0) \end{pmatrix} = A \cdot B.$$

Das folgt sofort aus den Rechenregeln für die partiellen Ableitungen. Diese Verknüpfungseigenschaft der Systeme partieller Ableitungen war im 19. Jahrhundert einer der Ausgangspunkte für die Entwicklung des Matrizenkalküls gewesen.

2.5.3. Zwei Spezialfälle der Matrizenmultiplikation sollen noch besonders erwähnt werden.

a) Drehungen des $\mathbb{R}^2$ um die Winkel α und β werden nach 2.1.1 beschrieben durch

$$A = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} \cos \beta & -\sin \beta \\ \sin \beta & \cos \beta \end{pmatrix}.$$

Die Hintereinanderausführung ist eine Drehung um den Winkel $\alpha + \beta$, und sie wird beschrieben durch $B \cdot A$, das bedeutet

$$\begin{pmatrix} \cos(\alpha + \beta) & -\sin(\alpha + \beta) \\ \sin(\alpha + \beta) & \cos(\alpha + \beta) \end{pmatrix} = \begin{pmatrix} \cos \alpha \cos \beta - \sin \alpha \sin \beta & -(\sin \alpha \cos \beta + \cos \alpha \sin \beta) \\ \cos \alpha \sin \beta + \sin \alpha \cos \beta & \cos \alpha \cos \beta - \sin \alpha \sin \beta \end{pmatrix}.$$

Daß die Einträge auf beiden Seiten gleich sind, ist die Aussage der sogenannten *Additionstheoreme* für Sinus und Cosinus. Dem Leser sei empfohlen, in diesem Fall die Gleichheit $A \cdot B = B \cdot A$ nachzuweisen, denn das ist bei Matrizen ein bemerkenswertes Ereignis.

b) Die Multiplikation einer $m \times n$ -Matrix A mit einer $n \times 1$ -Matrix, d.h. einer Spalte x haben wir schon lange benutzt zur Beschreibung der linearen Abbildung

$$A: K^n \rightarrow K^m, \quad x \mapsto A \cdot x.$$

Mit Hilfe der Matrizenmultiplikation geschrieben bedeutet die in 2.5.1 durchgeführte Rechnung dann

$$A \cdot (B \cdot x) = (A \cdot B) \cdot x$$

für jede $(r \times 1)$ -Matrix x . Das ist ein Spezialfall des Assoziativgesetzes für die Matrizenmultiplikation, das in der folgenden Sammlung von Regeln enthalten ist.

2.5.4. Rechenregeln für Matrizen. Sind Matrizen $A, A' \in M(m \times n; K)$ und $B, B' \in M(n \times r; K)$, $C \in M(r \times s; K)$ und $\lambda \in K$ gegeben, so gilt:

$$1) A \cdot (B + B') = A \cdot B + A \cdot B' \text{ und } (A + A') \cdot B = A \cdot B + A' \cdot B. \\ \text{(Distributivgesetze)}$$

$$2) A \cdot (\lambda B) = (\lambda A) \cdot B = \lambda(A \cdot B).$$

$$3) (A \cdot B) \cdot C = A \cdot (B \cdot C). \quad \text{(Assoziativgesetz)}$$

$$4) {}^t(A \cdot B) = {}^t B \cdot {}^t A.$$

$$5) E_m \cdot A = A \cdot E_n = A. \quad \text{(Neutralität der Einheitsmatrix)}$$

Beweis. 1), 2) und 5) sind ganz einfach und erfordern höchstens etwas Schreibarbeit. Für den Beweis von 4) muß man sorgfältig mit den Buchstaben umgehen: Ist $A = (a_{ij})$ und $B = (b_{jk})$, so ist

$$A \cdot B = (c_{ik}) \quad \text{mit } c_{ik} = \sum_{j=1}^n a_{ij} b_{jk}, \quad \text{also } {}^t(A \cdot B) = (c'_{ki}) \quad \text{mit } c'_{ki} = c_{ik}.$$

Weiter ist

$${}^t B = (b'_{kj}) \quad \text{mit } b'_{kj} = b_{jk} \quad \text{und} \quad {}^t A = (a'_{ji}) \quad \text{mit } a'_{ji} = a_{ij},$$

also

$${}^t B \cdot {}^t A = (d_{ki}) \quad \text{mit } d_{ki} = \sum_{j=1}^n b'_{kj} a'_{ji} = \sum_{j=1}^n b_{jk} a_{ij} = \sum_{j=1}^n a_{ij} b_{jk}.$$

Also ist $c'_{ki} = d_{ki}$.

Bleibt der Beweis des Assoziativgesetzes 3). Dazu betrachten wir die beteiligten Matrizen als lineare Abbildungen:

$$K^s \xrightarrow{C} K^r \xrightarrow{B} K^n \xrightarrow{A} K^m.$$

Nach 1.1.5 gilt das Assoziativgesetz für die Hintereinanderschaltung von Abbildungen, also ist

$$(A \circ B) \circ C = A \circ (B \circ C).$$

In 2.5.1 haben wir gezeigt, daß die Hintereinanderschaltung der Abbildungen durch das Produkt der Matrizen ausgedrückt wird. Das liefert die Behauptung.

Wer bei diesem Kniff Unwohlsein empfindet, möge zur Linderung noch einmal Summen umordnen: Sei $A = (a_{ij})$, $B = (b_{jk})$ und $C = (c_{kl})$. Dann ist

$$A \cdot B = (\alpha_{ik}) \quad \text{mit } \alpha_{ik} = \sum_{j=1}^n a_{ij} b_{jk},$$

also

$$(A \cdot B) \cdot C = (d_{il}) \quad \text{mit } d_{il} = \sum_{k=1}^r \left(\sum_{j=1}^n a_{ij} b_{jk} \right) c_{kl}.$$

Weiter ist

$$B \cdot C = (\beta_{jl}) \quad \text{mit } \beta_{jl} = \sum_{k=1}^r b_{jk} c_{kl},$$

also

$$A \cdot (B \cdot C) = (d'_{il}) \quad \text{mit } d'_{il} = \sum_{j=1}^n a_{ij} \cdot \left(\sum_{k=1}^r b_{jk} c_{kl} \right).$$

Die beiden Summen für d_{il} und d'_{il} enthalten genau die gleichen Summanden, also ist $d_{il} = d'_{il}$. $\square$

Anstelle von $A \cdot B$ schreibt man für das Produkt von Matrizen meist nur AB . Nach den Regeln 2) und 3) kann man auch Klammern weglassen und einfach

$$\lambda AB \quad \text{bzw.} \quad ABC$$

schreiben.

Vorsicht! In der Rechenregel 4) steht auf der rechten Seite der Gleichung *nicht* ${}^t A \cdot {}^t B$. In dieser Reihenfolge könnte man die Matrizen im allgemeinen nicht einmal miteinander multiplizieren. Aber selbst wenn $A, B \in M(n \times n; K)$ gilt, ist im allgemeinen

$${}^t(A \cdot B) \neq {}^t A \cdot {}^t B.$$

Man kontrolliere das an Beispielen nach. Auch von der Richtigkeit des Assoziativgesetzes sollte man sich anhand von einigen Beispielen überzeugen, denn es ist gar nicht selbstverständlich (und eine gute Kontrollmöglichkeit für die Rechnung).

Im Spezialfall quadratischer Matrizen folgt aus diesen Regeln das

Korollar. Die Menge $M(n \times n; K)$ mit der Addition aus 1.4.1 und der Multiplikation aus 2.5.2 ist ein Ring. $\square$

2.5.5. Es ist eine naheliegende Frage, wie der Rang der Produktmatrix von den Rängen der Faktoren abhängt. Man hat folgende Abschätzungen:

Lemma. Ist $A \in M(m \times n; K)$ und $B \in M(n \times r; K)$, so gilt

$$\text{rang } A + \text{rang } B - n \leq \text{rang } (A \cdot B) \leq \min \{ \text{rang } A, \text{rang } B \} .$$

Beweis. Wir betrachten die Matrizen als lineare Abbildungen, das ergibt ein Diagramm

$$\begin{array}{ccc} K^r & \xrightarrow{A \cdot B} & K^m \\ & \searrow B & \nearrow A \\ & K^n & \end{array}$$

Weiter definieren wir $F' := A|_{\text{Im } B}$. Dann ist

$$\text{Im } F' = \text{Im } (A \cdot B) \quad \text{und} \quad \text{Ker } F' = \text{Ker } A \cap \text{Im } B .$$

Die Dimensionsformel aus 2.2.4 angewandt auf F' ergibt

$$\text{rang } (A \cdot B) = \text{rang } F' = \dim \text{Im } B - \dim \text{Ker } F' = \text{rang } B - \dim \text{Ker } F' . \quad (*)$$

Daraus folgt $\text{rang } (A \cdot B) \leq \text{rang } B$. Da $\text{Im } (A \cdot B) \subset \text{Im } A$, folgt die Abschätzung nach oben. Wegen $\text{Ker } F' \subset \text{Ker } A$ folgt aus (*) weiter

$$\text{rang } (A \cdot B) \geq \text{rang } B - \dim \text{Ker } A = \text{rang } B + \text{rang } A - n ,$$

wobei die letzte Gleichung nach der Dimensionsformel für die Abbildung A gilt. Das ergibt die Abschätzung nach unten.

Sie ist scharf, wenn $\text{Ker } F' = \text{Ker } A$, also wenn $\text{Ker } A \subset \text{Im } B$. $\square$

2.5.6. Wir betrachten den kanonischen Isomorphismus aus 2.4.2 im Fall $m = n$, also

$$\text{Hom}(K^n, K^n) \rightarrow M(n \times n; K)$$

und fragen, welche quadratischen Matrizen die Isomorphismen

$$F: K^n \rightarrow K^n$$

beschreiben. Ein Isomorphismus hat eine Umkehrung F^{-1} mit

$$F \circ F^{-1} = F^{-1} \circ F = \text{id}_{K^n} .$$

Übersetzt in die entsprechenden Matrizen ergibt das die

Definition. Eine Matrix $A \in M(n \times n; K)$ heißt *invertierbar*, wenn es ein $A' \in M(n \times n; K)$ gibt mit

$$A \cdot A' = A' \cdot A = E_n .$$

Bemerkung 1. Die Menge

$$\mathrm{GL}(n; K) = \{A \in \mathrm{M}(n \times n; K) : A \text{ invertierbar}\}$$

mit der Multiplikation von Matrizen als Verknüpfung ist eine Gruppe mit neutralem Element E_n .

Sie heißt die *allgemeine lineare Gruppe* (general linear).

Beweis. Zunächst ist zu zeigen, daß die Matrizenmultiplikation tatsächlich eine Multiplikation in $\mathrm{GL}(n; K)$ induziert, d.h. daß für $A, B \in \mathrm{GL}(n; K)$ auch $A \cdot B \in \mathrm{GL}(n; K)$ gilt. Seien A' und B' so gewählt, daß

$$AA' = A'A = E_n = BB' = B'B$$

gilt. Dann ist

$$(B'A')(AB) = E_n = (AB)(B'A')$$

nach dem Assoziativgesetz für die Matrizenmultiplikation, also ist AB invertierbar. Es bleiben die Gruppenaxiome G1 und G2 nachzuweisen (vgl. 1.2.2).

Das Assoziativgesetz gilt in $\mathrm{GL}(n; K)$, denn die Multiplikation ist sogar in $\mathrm{M}(n \times n; K)$ assoziativ.

Die n -reihige Einheitsmatrix hat die Eigenschaft eines neutralen Elementes, und zu $A \in \mathrm{GL}(n; K)$ gibt es nach Definition ein Inverses A' . $\square$

Wie wir in Abschnitt 1.2.3 gesehen haben, ist das Inverse A' eindeutig bestimmt, und wie üblich schreibt man dafür A^{-1} . Es gilt dann

$$(A^{-1})^{-1} = A \quad \text{und} \quad (AB)^{-1} = B^{-1}A^{-1}.$$

Bemerkung 2. Für eine Matrix $A \in \mathrm{M}(n \times n; K)$ sind folgende Bedingungen gleichwertig:

- i) A ist invertierbar.
- ii) $'A$ ist invertierbar.
- iii) Spaltenrang $A = n$.
- iv) Zeilenrang $A = n$.

Außerdem ist $('A)^{-1} = '(A^{-1})$.

Beweis. i) $\Rightarrow$ ii) folgt aus $'(A^{-1})'A = '(AA^{-1}) = 'E_n = E_n$, und ii) $\Rightarrow$ i) ergibt sich daraus durch Transposition. i) $\Leftrightarrow$ iii) ist eine Folgerung aus 2.2.4, und ii) $\Leftrightarrow$ iv) ergibt sich wieder durch Transposition. $\square$

Aufgaben zu 2.5

1. Gegeben seien die Matrizen

$$A := \begin{pmatrix} 1 & -1 & 2 \\ 0 & 3 & 5 \\ 1 & 8 & -7 \end{pmatrix}, \quad B := \begin{pmatrix} -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ 1 & 0 & -1 & 0 \end{pmatrix}, \quad C := \begin{pmatrix} 1 \\ 0 \\ 8 \\ -7 \end{pmatrix},$$

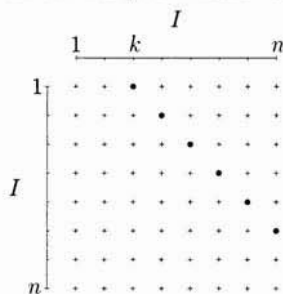
$$D := \begin{pmatrix} -1 & 2 & 0 & 8 \end{pmatrix}, \quad E := \begin{pmatrix} 1 & 4 \\ 0 & 5 \\ 6 & 8 \end{pmatrix}.$$

Berechnen Sie alle möglichen Produkte.

2. In dieser Aufgabe betrachten wir Eigenschaften „dünn besetzter“ Matrizen, in denen viele Einträge null sind.

a) Sei $n \in \mathbb{N} \setminus \{0\}$ und $I = \{1, \dots, n\}$. Wir betrachten die Menge $I \times I \subset \mathbb{N} \times \mathbb{N}$.

Finden Sie für $k \in \mathbb{N}$ Gleichungen für die „Gerade“ L in $I \times I$ durch $(1, k)$ und $(2, k+1)$ sowie für die Gerade L' durch $(k, 1)$ und $(k+1, 2)$. Finden Sie weiter Ungleichungen für den Halbraum H in $I \times I$, der oberhalb von L liegt und den Halbraum H' , der unterhalb von L' liegt.



b) Formulieren und beweisen Sie folgende Aussagen:

$$\begin{pmatrix} \diagdown & 0 \\ & * \\ 0 & \diagup \end{pmatrix} \cdot \begin{pmatrix} \diagdown & 0 \\ & * \\ 0 & \diagup \end{pmatrix} = \begin{pmatrix} \diagdown & 0 \\ & * \\ 0 & \diagup \end{pmatrix}$$

$$\begin{pmatrix} & * \\ 0 & \end{pmatrix} \cdot \begin{pmatrix} & * \\ 0 & \end{pmatrix} = \begin{pmatrix} & * \\ 0 & \end{pmatrix}$$

- c) Eine Matrix $A = (a_{ij}) \in M(n \times n; K)$ heißt echte obere Dreiecksmatrix, falls $a_{ij} = 0$ für $i \geq j$. Zeigen Sie, dass eine echte obere Dreiecksmatrix A nilpotent ist, d.h. es existiert ein $m \in \mathbb{N}$ mit $A^m = 0$.

3. Sind die folgenden Teilmengen Unterringe?

- a) $\{(a_{ij}) \in M(n \times n; K) : a_{ij} = 0 \text{ für } i \geq j\} \subset M(n \times n; K)$
 b) $\{(a_{ij}) \in M(n \times n; K) : a_{ij} = 0 \text{ für } i \geq j+k \text{ oder } j \geq i+k\} \subset M(n \times n; K)$, wobei $k \in \mathbb{N}$
 c) $\left\{ \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \in M(2 \times 2; \mathbb{R}) : a \in \mathbb{Q}, b, c \in \mathbb{R} \right\} \subset M(2 \times 2; \mathbb{R})$
 d) $\left\{ \begin{pmatrix} 0 & a \\ 0 & b \end{pmatrix} \in M(2 \times 2; K) : a, b \in K \right\} \subset M(2 \times 2; K)$
 e) $\{(a_{ij}) \in M(n \times n; K) : a_{ij} = 0 \text{ für } i \neq j \text{ oder } i \geq k\} \subset M(n \times n; K)$, wobei $k \in \mathbb{N}$.

4. Sei K ein Körper und $n \in \mathbb{N} \setminus \{0\}$.

- a) Für $\lambda \in K$ gilt: $(\lambda E_n)B = B(\lambda E_n)$ für alle $B \in M(n \times n; K)$.
 b) Zeigen Sie: Ist $A \in M(n \times n; K)$ mit $AB = BA$ für alle $B \in M(n \times n; K)$, so existiert ein $\lambda \in K$ mit $A = \lambda E_n$.

5. Sei $C = \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} : a, b \in \mathbb{R} \right\} \subset M(2 \times 2; \mathbb{R})$.

- a) Zeigen Sie, dass C ein Körper ist.
 b) In C ist die Gleichung $X^2 + 1 = 0$ lösbar.
 c) C ist als Körper isomorph zu $\mathbb{C}$.

6. Zeigen Sie, dass für eine Matrix $B \in M(n \times k; \mathbb{R})$ die Abbildung

$$\Phi: M(m \times n; \mathbb{R}) \rightarrow M(m \times k; \mathbb{R}), \quad A \mapsto A \cdot B,$$

stetig ist.

7. Zeigen Sie, dass die Abschätzung

$$\text{rang} A + \text{rang} B - n \leq \text{rang}(AB) \leq \min\{\text{rang} A, \text{rang} B\}$$

aus 2.5.5 für den Rang der Produktmatrix in beide Richtungen scharf ist, d. h. finden Sie Beispiele für

$$\text{rang} A + \text{rang} B - n = \text{rang}(AB) \quad \text{und} \quad \text{rang}(AB) = \min\{\text{rang} A, \text{rang} B\}.$$

8. Wir wollen eine Methode angeben, um die Inverse einer Matrix auszurechnen:

Sei dazu $A \in M(n \times n; K)$ invertierbar, d. h. $\text{rang } A = n$. Zeigen Sie: Ist

$$x^i = \begin{pmatrix} x_{1i} \\ \vdots \\ x_{ni} \end{pmatrix}$$

die Lösung des Gleichungssystems $Ax = e_i$, so ist

$$A^{-1} = \begin{pmatrix} x_{11} & \dots & x_{1n} \\ \vdots & & \vdots \\ x_{n1} & \dots & x_{nn} \end{pmatrix}.$$

Berechnen Sie auf diese Weise die inverse Matrix von

$$A = \begin{pmatrix} 1 & 1 & 2 & 4 \\ 1 & 3 & 4 & -2 \\ 0 & 1 & 3 & 6 \\ 1 & 3 & 5 & 3 \end{pmatrix}.$$

9. Für eine differenzierbare Abbildung

$$f: \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad x \mapsto (f_1(x), \dots, f_m(x)),$$

ist die Jacobi-Matrix von f im Punkt x definiert durch

$$\text{Jac}_x f := \left(\frac{\partial f_i}{\partial x_j}(x) \right).$$

Ist $m = 1$ und f zweimal stetig partiell differenzierbar, so versteht man unter der Hesse-Matrix von f im Punkt x die Matrix

$$\text{Hess}_x f := \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right).$$

a) Berechnen Sie die Jacobi-Matrix einer linearen Abbildung $F: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $x \mapsto Ax$, wobei $A \in M(m \times n; \mathbb{R})$.

b) Sei

$$P: \mathbb{R}^n \rightarrow \mathbb{R}, \quad (x_1, \dots, x_n) \mapsto \sum_{i \leq j} a_{ij} x_i x_j + \sum_{i=1}^n b_i x_i,$$

wobei $a_{ij}, b_i \in \mathbb{R}$. Berechnen Sie die Jacobi-Matrix und die Hesse-Matrix von P .

2.6 Koordinatentransformationen

Eine immer wiederkehrende Methode in der linearen Algebra ist es, durch Anwendung passender Begriffe langweilige Rechnungen mit Schlachten gegen Indizes zu vermeiden. Das ist etwas gefährlich, weil dadurch ein Trainingsrückstand im Rechnen entstehen kann. Daher vertrauen wir in diesem Abschnitt besonders darauf, daß der Leser nebenbei zur Übung genügend viele Beispiele rechnet.

2.6.1. Sei wieder V ein Vektorraum mit einer Basis $\mathcal{B} = (v_1, \dots, v_n)$. Entsprechend 2.4.2 gehört dazu genau ein Isomorphismus

$$\Phi_{\mathcal{B}}: K^n \rightarrow V \quad \text{mit} \quad \Phi_{\mathcal{B}}(e_j) = v_j \quad \text{für } j = 1, \dots, n,$$

wobei $(e_1, \dots, e_n)$ wie immer die kanonische Basis des K^n bezeichnet. Nach Definition ist

$$\Phi_{\mathcal{B}}(x_1, \dots, x_n) = x_1 v_1 + \dots + x_n v_n.$$

Man nennt $\Phi_{\mathcal{B}}$ das durch $\mathcal{B}$ bestimmte *Koordinatensystem* in V und

$$x = (x_1, \dots, x_n) = \Phi_{\mathcal{B}}^{-1}(v) \in K^n$$

die *Koordinaten* von $v = x_1 v_1 + \dots + x_n v_n$.

2.6.2. Für Anwendungen ist es wichtig, verschiedene Koordinaten ineinander umzurechnen. Dazu nimmt man an, es seien in V zwei Basen $\mathcal{A} = (v_1, \dots, v_n)$ und $\mathcal{B} = (w_1, \dots, w_n)$ gegeben. Dann hat man ein Diagramm von Isomorphismen

$$\begin{array}{ccc} K^n & & \\ & \searrow \Phi_{\mathcal{A}} & \\ T_{\mathcal{B}}^{\mathcal{A}} := \Phi_{\mathcal{B}}^{-1} \circ \Phi_{\mathcal{A}} & \downarrow & V \\ & \nearrow \Phi_{\mathcal{B}} & \\ K^n & & \end{array}$$

Man nennt die als lineare Abbildung angesehene Matrix $T_{\mathcal{B}}^{\mathcal{A}} \in \text{GL}(n; K)$ die *Transformationsmatrix* des Basiswechsels. Sie hat nach Definition die folgende Eigenschaft: ist

$$v = x_1 v_1 + \dots + x_n v_n = y_1 w_1 + \dots + y_n w_n \in V, \quad \text{so ist}$$

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = T_{\mathcal{B}}^{\mathcal{A}} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Kennt man die Matrix $T_{\mathcal{B}}^{\mathcal{A}}$, so kann man also die „neuen“ Koordinaten y aus den „alten“ x berechnen.

Das wichtigste Beispiel ist $V = K^n$. Sind A und B die Matrizen mit den Vektoren aus $\mathcal{A}$ und $\mathcal{B}$ als Spalten, so wird obiges Diagramm zu

$$\begin{array}{ccc} K^n & \xrightarrow{A} & K^n \\ T \downarrow & & \nearrow B \\ K^n & & \end{array} \quad \text{also} \quad T = B^{-1} \cdot A.$$

Ist insbesondere $\mathcal{A}$ die kanonische Basis, so folgt $T = B^{-1}$.

Ganz analog behandelt man den Fall, daß für allgemeines V die Vektoren aus $\mathcal{B}$ durch Linearkombinationen aus $\mathcal{A}$ gegeben sind. Sei

$$w_j = s_{1j}v_1 + \dots + s_{nj}v_n$$

und $S = (s_{ij})$ die Matrix mit diesen Koeffizienten als Spalten. Es gilt

$$\Phi_B = \Phi_A \circ S,$$

denn die Werte der beiden Abbildungen auf der kanonischen Basis des K^n sind gleich:

$$(\Phi_A \circ S)(e_j) = \Phi_A(s_{1j}, \dots, s_{nj}) = s_{1j}v_1 + \dots + s_{nj}v_n = w_j = \Phi_B(e_j).$$

Als Ergebnis erhalten wir

$$S = \Phi_A^{-1} \circ \Phi_B = (T_B^A)^{-1}, \quad \text{also} \quad T_B^A = S^{-1}.$$

Das kann man in Form eines Diagramms schreiben:

$$\begin{array}{ccc} K^n & \xrightarrow{\Phi_A} & V \\ S \uparrow & & \nearrow \Phi_B \\ K^n & & \end{array}.$$

Damit ist das Problem wieder auf die Bestimmung einer inversen Matrix zurückgeführt. Ein allgemeines Verfahren dafür wird in 2.7.4 angegeben. Wenigstens ein ganz einfaches Beispiel rechnen wir direkt aus.

Beispiel. Im $\mathbb{R}^2$ betrachten wir neben der kanonischen Basis $\mathcal{K} = (e_1, e_2)$ die Basis

$$\mathcal{B} = (w_1, w_2) \quad \text{mit} \quad w_1 = {}^t(2, 1), \quad w_2 = {}^t(1, 3).$$

Die Einträge von B^{-1} betrachten wir als Unbestimmte, das ergibt die Bedingung

$$\begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} x_1 & x_3 \\ x_2 & x_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Sie ist gleichwertig mit den linearen Gleichungssystemen

$$\begin{array}{rcl} 2x_1 + x_2 & = & 1 \\ x_1 + 3x_2 & = & 0 \end{array} \quad \text{und} \quad \begin{array}{rcl} 2x_3 + x_4 & = & 0 \\ x_3 + 3x_4 & = & 1. \end{array}$$

Die eindeutigen Lösungen sind

$$\begin{array}{ll} x_1 = \frac{3}{5}, & x_3 = -\frac{1}{5}, \\ x_2 = -\frac{1}{5}, & x_4 = \frac{2}{5}, \end{array} \quad \text{also ist} \quad B^{-1} = \frac{1}{5} \begin{pmatrix} 3 & -1 \\ -1 & 2 \end{pmatrix}.$$

Für $v = -e_1 - e_2$ ist $x = {}^t(-1, -1)$, also $y = B^{-1} \cdot x = {}^t(-\frac{2}{5}, -\frac{1}{5})$, d.h. $v = -\frac{2}{5}w_1 - \frac{1}{5}w_2$.

2.6.3. Der Zusammenhang zwischen Koordinatensystemen und darstellenden Matrizen ergibt sich aus der

Bemerkung. Sei $F: V \rightarrow W$ eine lineare Abbildung, und seien $\mathcal{A}$ und $\mathcal{B}$ Basen von V und W . Dann hat man ein Diagramm linearer Abbildungen

$$\begin{array}{ccc} K^n & \xrightarrow{\Phi_{\mathcal{A}}} & V \\ \downarrow M_{\mathcal{B}}^{\mathcal{A}}(F) & & \downarrow F \\ K^m & \xrightarrow{\Phi_{\mathcal{B}}} & W, \end{array}$$

und es gilt

$$\Phi_{\mathcal{B}} \circ M_{\mathcal{B}}^{\mathcal{A}}(F) = F \circ \Phi_{\mathcal{A}}, \quad \text{also} \quad M_{\mathcal{B}}^{\mathcal{A}}(F) = \Phi_{\mathcal{B}}^{-1} \circ F \circ \Phi_{\mathcal{A}}.$$

Die darstellenden Matrizen sind offenbar eine Verallgemeinerung der Transformationsmatrizen, denn für $V = W$ und $F = \text{id}_V$ gilt

$$M_{\mathcal{B}}^{\mathcal{A}}(\text{id}_V) = T_{\mathcal{B}}^{\mathcal{A}}.$$

Beweis. Es genügt zu zeigen, daß die beiden Abbildungen auf der kanonischen Basis $(e_1, \dots, e_n)$ übereinstimmen. Ist $M_{\mathcal{B}}^{\mathcal{A}}(F) = A = (a_{ij})$, so ist

$$\Phi_{\mathcal{B}}(M_{\mathcal{B}}^{\mathcal{A}}(F)(e_j)) = \Phi_{\mathcal{B}}(a_{1j}, \dots, a_{mj}) = \sum_{i=1}^m a_{ij} w_i,$$

$$F(\Phi_{\mathcal{A}}(e_j)) = F(v_j) = \sum_{i=1}^m a_{ij} w_i.$$

Die zweite Gleichung folgt aus der ersten durch Multiplikation von links mit $\Phi_{\mathcal{B}}^{-1}$. $\square$

2.6.4. Im Diagramm aus 2.6.3 hat man zwei verschiedene Wege, mit Hilfe der Abbildungen in Pfeilrichtung von K^n nach W zu gelangen, und die Aussage ist, daß auf den verschiedenen Wegen die gleiche Abbildung herauskommt. Ein Diagramm mit dieser Eigenschaft heißt *kommutativ*. Wie nützlich dieser Begriff ist, werden wir gleich sehen:

Satz. Gegeben seien Vektorräume U, V und W mit Basen $\mathcal{A}, \mathcal{B}$ und $\mathcal{C}$, sowie lineare Abbildungen $G: U \rightarrow V$ und $F: V \rightarrow W$. Dann gilt:

$$M_{\mathcal{C}}^{\mathcal{A}}(F \circ G) = M_{\mathcal{C}}^{\mathcal{B}}(F) \cdot M_{\mathcal{B}}^{\mathcal{A}}(G).$$

Kurz ausgedrückt: Der Hintereinanderschaltung von linearen Abbildungen entspricht das Produkt der darstellenden Matrizen.

Beweis. Für die Standardräume mit den kanonischen Basen wurde das schon in 2.5.1 ausgerechnet. Der allgemeine Fall folgt daraus durch Betrachtung des Diagramms

$$\begin{array}{ccccc}
 K^n & & \xrightarrow{\Phi_{\mathcal{A}}} & & U \\
 & \searrow B & & \nearrow G & \\
 A \cdot B & K^m & \xrightarrow{\Phi_{\mathcal{B}}} & V & \\
 & \nearrow A & & \searrow F & \\
 K^n & & \xrightarrow{\Phi_{\mathcal{C}}} & & W,
 \end{array}$$

$F \circ G$ (rechts neben dem Pfeil von V nach W)

wobei $A = M_{\mathcal{C}}^{\mathcal{B}}(F)$ und $B = M_{\mathcal{B}}^{\mathcal{A}}(G)$. Alle Teildiadramme sind kommutativ (man mache sich klar warum), daher ist das ganze Diagramm kommutativ, und insbesondere folgt die Behauptung.

Wer lieber etwas rechnet, betrachte einen Vektor $u \in U$ und seine Koordinaten $x = \Phi_{\mathcal{A}}^{-1}(u)$. Wegen 2.5.4 ist

$$\Phi_{\mathcal{B}}^{-1}(G(u)) = B \cdot x, \quad \Phi_{\mathcal{C}}^{-1}(F(G(u))) = A \cdot (B \cdot x) = (A \cdot B) \cdot \Phi_{\mathcal{A}}^{-1}(u),$$

also $\Phi_{\mathcal{C}}^{-1} \circ (F \circ G) \circ \Phi_{\mathcal{A}} = A \cdot B$. Auch daraus folgt die Behauptung. $\square$

Für den Spezialfall von Endomorphismen (vgl. 2.4.4) ergibt sich das

Korollar. In V seien eine Basis B sowie Endomorphismen F, G gegeben. Dann ist

$$M_B(F \circ G) = M_B(F) \cdot M_B(G).$$

□

Insbesondere folgt daraus, daß

$$M_B: \text{End}(V) \rightarrow M(n \times n; K)$$

ein Ringisomorphismus ist (vgl. 1.3.2, 2.1.4, 2.4.4 und 2.5.4).

2.6.5. Nun kommen wir zum wichtigsten Ergebnis dieses Abschnittes, nämlich der Antwort auf die Frage, wie sich die darstellende Matrix bei Einführung neuer Basen ändert.

Transformationsformel. Ist $F: V \rightarrow W$ eine lineare Abbildung, sind $\mathcal{A}, \mathcal{A}'$ Basen von V und $\mathcal{B}, \mathcal{B}'$ Basen von W , so ist das Diagramm

$$\begin{array}{ccc}
 K^n & \xrightarrow{M_{\mathcal{B}}^{\mathcal{A}}(F)} & K^m \\
 \downarrow T_{\mathcal{A}'}^{\mathcal{A}} & \searrow \Phi_{\mathcal{A}} & \swarrow \Phi_{\mathcal{B}} \\
 & V \xrightarrow{F} W & \\
 \uparrow \Phi_{\mathcal{A}'} & & \searrow \Phi_{\mathcal{B}'} \\
 K^n & \xrightarrow{M_{\mathcal{B}'}^{\mathcal{A}'}(F)} & K^m \\
 & \downarrow T_{\mathcal{B}'}^{\mathcal{B}} &
 \end{array}$$

kommutativ. Insbesondere gilt für die beteiligten Matrizen

$$M_{\mathcal{B}'}^{\mathcal{A}'}(F) = T_{\mathcal{B}'}^{\mathcal{B}} \cdot M_{\mathcal{B}}^{\mathcal{A}}(F) \cdot (T_{\mathcal{A}'}^{\mathcal{A}})^{-1}.$$

Anders ausgedrückt: Sind

$$A = M_{\mathcal{B}}^{\mathcal{A}}(F) \quad \text{und} \quad B = M_{\mathcal{B}'}^{\mathcal{A}'}(F)$$

die beiden Matrizen, die F bezüglich verschiedener Paare von Basen darstellen, und sind

$$T = T_{\mathcal{A}'}^{\mathcal{A}}, \quad S = T_{\mathcal{B}'}^{\mathcal{B}}$$

die Transformationsmatrizen zwischen den verschiedenen Basen, so gilt

$$B = S \cdot A \cdot T^{-1}.$$

Zum *Beweis* genügt es zu bemerken, daß nach 2.6.2 und 2.6.3 die dreieckigen und viereckigen Teile des Diagramms kommutativ sind. Also ist das Gesamtdiagramm kommutativ. $\square$

Wer diesen Beweis als Hokusfokus ansieht, möge die Formel $B = SAT^{-1}$ direkt durch Multiplikation der drei Matrizen nachrechnen (Viel Spaß mit den Indizes!). Dabei wird sich zeigen, daß nur Rechnungen wiederholt und ineinander eingesetzt werden, die vorher schon einmal ausgeführt worden waren. Der Trick besteht also darin, sich dieses zu ersparen.

Für den Spezialfall eines Endomorphismus ergibt sich mit der Notation aus 2.4.4 das

Korollar. Sind in V zwei Basen $\mathcal{A}$ und $\mathcal{B}$ sowie ein Endomorphismus F gegeben, so ist

$$M_{\mathcal{B}}(F) = T_{\mathcal{B}}^{\mathcal{A}} \cdot M_{\mathcal{A}}(F) \cdot T_{\mathcal{A}}^{\mathcal{B}},$$

oder anders ausgedrückt

$$B = SAS^{-1},$$

wenn $A = M_{\mathcal{A}}(F)$, $B = M_{\mathcal{B}}(F)$ und $S = T_{\mathcal{B}}^{\mathcal{A}}$. $\square$

2.6.6. Nun endlich können wir noch einmal (vgl. auch 1.5.8) die Gleichheit von Zeilenrang und Spaltenrang beweisen, ohne uns die Finger mit Indizes zu beschmutzen.

Satz. Für jede Matrix $A \in M(m \times n; K)$ gilt

$$\text{Zeilenrang } A = \text{Spaltenrang } A.$$

Diese Zahl ist nach 2.2.1 gleich $\text{rang } A$.

Beweis. Wir betrachten $A: K^n \rightarrow K^m$ als lineare Abbildung und wählen in K^n und K^m Basen $\mathcal{A}$ und $\mathcal{B}$ entsprechend 2.4.3, d.h. mit

$$M_{\mathcal{B}}^{\mathcal{A}}(A) = B = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Für B ist offensichtlich

$$\text{Zeilenrang } B = r = \text{Spaltenrang } B.$$

Um zu zeigen, daß sich diese Gleichheit auf A überträgt, wählen wir entsprechend 2.6.5 invertierbare Matrizen S und T mit

$$B = SAT.$$

Es genügt also der Beweis von folgendem

Hilfssatz. Für $A \in M(m \times n; K)$, $S \in GL(m; K)$ und $T \in GL(n; K)$ gilt

- 1) Spaltenrang $SAT =$ Spaltenrang A ,
- 2) Zeilenrang $SAT =$ Zeilenrang A .

Beweis des Hilfssatzes. Zu den gegebenen Matrizen gehört ein kommutatives Diagramm linearer Abbildungen

$$\begin{array}{ccc} K^n & \xrightarrow{A} & K^m \\ \uparrow T & & \downarrow S \\ K^n & \xrightarrow{SAT} & K^m. \end{array}$$

Da S und T Isomorphismen sind, haben die linearen Abbildungen A und SAT gleichen Rang, d.h. es gilt 1). Daraus folgt 2) durch Transposition, denn

$$\text{Zeilenrang } A = \text{Spaltenrang } {}^t A \quad \text{und} \quad {}^t(SAT) = {}^t T \cdot {}^t A \cdot {}^t S.$$

Man beachte, daß bei Multiplikation von A von rechts sogar der Spaltenraum, bei Multiplikation von links nur seine Dimension erhalten bleibt. $\square$

Wie man die Matrizen S und T aus A berechnen kann, werden wir in 2.7.6 sehen.

2.6.7. Die Transformationsformel aus 2.6.5 ergibt in die Sprache der Matrizen übersetzt die folgende

Definition. Zwei Matrizen $A, B \in M(m \times n; K)$ heißen *äquivalent*, wenn es $S \in GL(m; K)$ und $T \in GL(n; K)$ gibt mit

$$B = SAT^{-1}.$$

Im Spezialfall $m = n$ nennen wir $A, B \in M(m \times m; K)$ *ähnlich*, wenn es ein $S \in GL(m; K)$ gibt mit

$$B = SAS^{-1}.$$

Aus 2.6.5 folgt sofort die

Bemerkung. Zwei Matrizen sind genau dann äquivalent, wenn sie bezüglich verschiedener Paare von Basen die gleiche lineare Abbildung beschreiben.

Zwei quadratische Matrizen sind genau dann ähnlich, wenn sie bezüglich verschiedener Basen den gleichen Endomorphismus beschreiben. $\square$

Daß dieser Begriff der Äquivalenz nichts Neues liefert, zeigt das

Lemma. Zwei Matrizen sind genau dann äquivalent, wenn sie den gleichen Rang haben. Insbesondere ist jede Matrix vom Rang r äquivalent zu

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Diese speziellen Matrizen repräsentieren die Äquivalenzklassen und heißen *Normalformen*.

Zum *Beweis* genügt es, die Argumente aus 2.6.6 zu wiederholen. Daß äquivalente Matrizen gleichen Rang haben, folgt aus dem dort bewiesenen Hilfssatz.

Ist A vom Rang r , so sieht man durch entsprechende Wahl von Basen in K^m und K^n , daß A äquivalent zu der obigen *Normalform* ist. Also ist A äquivalent zu jeder anderen Matrix B vom Rang r (benutze Aufgabe 4). $\square$

Viel schwieriger ist die Frage nach *Normalformen für Klassen ähnlicher Matrizen*. Das ist Thema von Kapitel 4.

Aufgaben zu 2.6

1. Gegeben sei ein endlichdimensionaler Vektorraum V mit Basen $\mathcal{A}$, $\mathcal{B}$ und $\mathcal{C}$. Beweisen Sie die „Kürzungsregel“

$$T_C^A = T_C^B \cdot T_B^A.$$

2. Im $\mathbb{R}^3$ seien die Basen

$$\mathcal{A} = ((1, -1, 2), (2, 3, 7), (2, 3, 6)) \text{ und } \mathcal{B} = ((1, 2, 2), (-1, 3, 3), (-2, 7, 6))$$

gegeben.

a) Berechnen Sie die Transformationsmatrix T_B^A .

b) Bestimmen Sie die Koordinaten des Vektors

$$v = 2 \cdot (1, -1, 2) + 9 \cdot (2, 3, 7) - 8 \cdot (2, 3, 6)$$

bezüglich der Basis $\mathcal{B}$.

3. V sei ein $\mathbb{R}$ -Vektorraum mit Basis $\mathcal{A} = (v_1, \dots, v_4)$, W sei ein $\mathbb{R}$ -Vektorraum mit Basis $\mathcal{B} = (w_1, \dots, w_5)$. $F: V \rightarrow W$ sei die lineare Abbildung, die gegeben ist durch

$$M_B^A(F) = \begin{pmatrix} 3 & 1 & -2 & 2 \\ -2 & -2 & 7 & -3 \\ 4 & 0 & 3 & 1 \\ 1 & 3 & 12 & 4 \\ 0 & 4 & -17 & 5 \end{pmatrix}.$$

Schließlich seien $\mathcal{A}' = (v'_1, \dots, v'_4)$ mit $v'_1 = v_1 + v_2$, $v'_2 = v_2 + v_3$, $v'_3 = v_3 + v_4$, $v'_4 = v_4$ und $\mathcal{B}' = (w'_1, \dots, w'_5)$ mit $w'_1 = w_1$, $w'_2 = w_1 + w_2$, $w'_3 = -w_1 + w_3$, $w'_4 = w_1 + w_4$, $w'_5 = w_1 + w_5$.

- Zeigen Sie, dass $\mathcal{A}'$ eine Basis von V und $\mathcal{B}'$ eine Basis von W ist.
- Berechnen Sie $M_{\mathcal{B}}^{\mathcal{A}'}(F)$, $M_{\mathcal{B}'}^{\mathcal{A}}(F)$ und $M_{\mathcal{B}'}^{\mathcal{A}'}(F)$.
- Bestimmen Sie $F^{-1}(\text{span}(w_1, w_2, w_3))$.

4. Zeigen Sie, dass durch

$$A \sim B \Leftrightarrow A \text{ und } B \text{ sind äquivalent}$$

(vgl. 2.6.7) tatsächlich eine Äquivalenzrelation auf der Menge $M(m \times n; K)$ gegeben ist und durch

$$A \sim B \Leftrightarrow A \text{ und } B \text{ sind ähnlich}$$

(vgl. 2.6.7) eine Äquivalenzrelation auf $M(m \times m; K)$ erklärt ist.

5. Zeigen Sie, dass für $A \in M(n \times n; \mathbb{R})$ gilt: $\text{rang } A = \text{rang } (A \cdot {}^t A)$. Gilt dies auch, falls $A \in M(n \times n; \mathbb{C})$?

2.7 Elementarmatrizen und Matrizenumformungen

Die Bedeutung von Matrizenumformungen sieht man schon am Eliminationsverfahren von GAUSS. Es gestattet die Lösung durch wiederholte Anwendung elementarer Schritte. Auf diese Weise kann man in der linearen Algebra viele Probleme lösen. In diesem Abschnitt geben wir weitere Beispiele dafür: die Inversion einer quadratischen Matrix und die explizite Bestimmung der Transformationsmatrizen, die eine gegebene Matrix auf Normalform bringen. Zur theoretischen Begründung ist es dabei sehr hilfreich, Matrizenumformungen zu interpretieren als Multiplikation mit speziellen invertierbaren Matrizen, die ihrerseits die gesamte allgemeine lineare Gruppe erzeugen. Diese sogenannten „Elementarmatrizen“ kann man als die Heizeilmännchen der linearen Algebra bezeichnen: sie erledigen geräuschlos und zuverlässig die kleinen Schmutzarbeiten.

2.7.1. Ist m eine beliebige natürliche Zahl, $1 \leq i, j \leq m$ mit $i \neq j$ und $\lambda \in K^*$, so nennt man die quadratischen Matrizen

$$\begin{array}{c}
 \begin{array}{cc}
 i\text{-te} & j\text{-te} \\
 \text{Spalte} & \text{Spalte} \\
 \downarrow & \downarrow
 \end{array} \\
 S_i(\lambda) := \left(\begin{array}{cccc|ccc}
 1 & & & & & & & \\
 & \ddots & & & & & & \\
 & & 1 & & & & & \\
 - & - & \lambda & - & - & 0 & - & - \\
 & & & 1 & & & & \\
 & & & & \ddots & & & \\
 & & & & & 1 & & \\
 - & - & 0 & - & - & 1 & - & - \\
 & & & & & & 1 & \\
 & & & & & & & \ddots \\
 & & & & & & & 1
 \end{array} \right) \begin{array}{l} \\ \\ \\ \leftarrow i\text{-te Zeile} \\ \\ \\ \leftarrow j\text{-te Zeile} \\ \\ \end{array}
 \end{array}$$

$$\begin{array}{c}
 Q_i^j := \left(\begin{array}{cccc|ccc}
 1 & & & & & & & \\
 & \ddots & & & & & & \\
 & & 1 & & & & & \\
 - & - & 1 & - & - & 1 & - & - \\
 & & & 1 & & & & \\
 & & & & \ddots & & & \\
 & & & & & 1 & & \\
 - & - & 0 & - & - & 1 & - & - \\
 & & & & & & 1 & \\
 & & & & & & & \ddots \\
 & & & & & & & 1
 \end{array} \right) \begin{array}{l} \\ \\ \\ \leftarrow i\text{-te Zeile} \\ \\ \\ \leftarrow j\text{-te Zeile} \\ \\ \end{array}
 \end{array}$$

$$Q_i^j(\lambda) := \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ - & - & 1 & - & \lambda & - & - \\ & & & 1 & & \\ & & & & \ddots & \\ - & - & 0 & - & 1 & - & - \\ & & & & & 1 & \\ & & & & & & \ddots & \\ & & & & & & & 1 \end{pmatrix} \begin{matrix} \\ \\ \\ \leftarrow i\text{-te Zeile} \\ \\ \\ \leftarrow j\text{-te Zeile} \\ \\ \end{matrix}$$

$$P_i^j := \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ - & - & 0 & - & 1 & - & - \\ & & & 1 & & \\ & & & & \ddots & \\ - & - & 1 & - & 0 & - & - \\ & & & & & 1 & \\ & & & & & & \ddots & \\ & & & & & & & 1 \end{pmatrix} \begin{matrix} \\ \\ \\ \leftarrow i\text{-te Zeile} \\ \\ \\ \leftarrow j\text{-te Zeile} \\ \\ \end{matrix}$$

aus $M(m \times m; K)$ *Elementarmatrizen*. Außer den eingetragenen oder durch Punkte angedeuteten Komponenten sind dabei alle Komponenten gleich Null.

Sind E_i^j die in 1.5.1 definierten Matrizen und ist E die m -reihige Einheitsmatrix, so ist

$$Q_i^j = E + E_i^j, \quad Q_i^j(\lambda) = E + \lambda E_i^j \quad \text{und} \quad S_i(\lambda) = E + (\lambda - 1)E_i^i.$$

Weiter ist selbstverständlich

$$Q_i^j = Q_i^j(1) \quad \text{und} \quad P_i^j = P_i^j.$$

Grundlegend ist der Zusammenhang mit den elementaren Umformungen von Matrizen. Ist $A \in M(m \times n; K)$ und $\lambda \in K^*$ gegeben, so hatten wir in 1.5.7 umgeformte Matrizen betrachtet, die wie folgt aus A entstanden waren:

A_I durch Multiplikation der i -ten Zeile mit λ ,

A_{II} durch Addition der j -ten Zeile zur i -ten Zeile,

A_{III} durch Addition der λ -fachen j -ten Zeile zur i -ten Zeile,

A_{IV} durch Vertauschen der i -ten und der j -ten Zeile.

Wie man sofort sieht, gilt

$$\begin{aligned} A_I &= S_i(\lambda) \cdot A, & A_{II} &= Q_i^j \cdot A, \\ A_{III} &= Q_i^j(\lambda) \cdot A, & A_{IV} &= P_i^j \cdot A. \end{aligned}$$

Man sollte sich davon auch anhand von Beispielen überzeugen, um mehr Zutrauen zu den eigenartigen Elementarmatrizen zu gewinnen.

Ganz Entsprechendes gilt, wenn man anstatt Zeilen immer Spalten umformt. Wir wollen es sicherheitshalber notieren. Ist $A \in M(m \times n; K)$ und $\lambda \in K^*$, so betrachten wir wieder die wie folgt aus A entstandenen Matrizen:

A^I durch Multiplikation der i -ten Spalte mit λ ,

A^{II} durch Addition der i -ten Spalte zur j -ten Spalte,

A^{III} durch Addition der λ -fachen i -ten Spalte zur j -ten Spalte,

A^{IV} durch Vertauschen der j -ten und der i -ten Spalte.

Verwenden wir die entsprechenden n -reihigen Elementarmatrizen, so gilt

$$\begin{aligned} A^I &= A \cdot S_i(\lambda), & A^{II} &= A \cdot Q_i^j, \\ A^{III} &= A \cdot Q_i^j(\lambda), & A^{IV} &= A \cdot P_i^j. \end{aligned}$$

Kurz ausgedrückt: Multiplikation von *links* mit Elementarmatrizen bewirkt *Zeilen*umformungen, und Multiplikation von *rechts* mit Elementarmatrizen bewirkt *Spalten*umformungen. Man beachte dabei die Vertauschung von i und j beim Übergang von links nach rechts.

Bemerkung. Die Elementarmatrizen $Q_i^j(\lambda)$ und P_i^j sind Produkte von Elementarmatrizen vom Typ $S_i(\lambda)$ und Q_i^j . Genauer gilt

$$Q_i^j(\lambda) = S_j\left(\frac{1}{\lambda}\right) \cdot Q_i^j \cdot S_j(\lambda), \quad P_i^j = Q_i^j \cdot Q_i^j(-1) \cdot Q_j^i \cdot S_j(-1).$$

Dies entspricht dem in 1.5.7 bemerkten Sachverhalt, daß sich Matrizenumformungen vom Typ III und IV aus Umformungen vom Typ I und II kombinieren lassen.

Der *Beweis* der Bemerkung ist ganz einfach, wenn man die Multiplikation als Zeilen- oder Spaltenumformung interpretiert. Wir wollen dies dem Leser überlassen. □

2.7.2. Lemma. Die Elementarmatrizen sind invertierbar und ihre Inversen sind wieder Elementarmatrizen. Genauer gilt:

$$\begin{aligned} (S_i(\lambda))^{-1} &= S_i\left(\frac{1}{\lambda}\right), & (Q_i^j)^{-1} &= Q_i^j(-1), \\ (Q_i^j(\lambda))^{-1} &= Q_i^j(-\lambda), & (P_i^j)^{-1} &= P_i^j. \end{aligned}$$

Zum Beweis genügt es, die rechten Seiten der Gleichungen mit den linken zu multiplizieren und festzustellen, daß die Einheitsmatrix herauskommt. $\square$

2.7.3. Satz. Jede invertierbare Matrix $A \in M(n \times n; K)$ ist (endliches) Produkt von Elementarmatrizen.

Man sagt dafür auch, daß die Gruppe $GL(n; K)$ von den Elementarmatrizen erzeugt wird.

Beweis. Nach 2.5.6 ist der Zeilenrang von A gleich n . Wie in 0.4.7 ausgeführt ist, kann man A durch elementare Zeilenumformungen zu einer Matrix der Form

$$B = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ & \ddots & \vdots \\ 0 & & b_{nn} \end{pmatrix}$$

mit von Null verschiedenen Diagonalelementen $b_{11}, \dots, b_{nn}$ machen. Nach 2.7.1 gibt es Elementarmatrizen $B_1, \dots, B_r$, so daß

$$B = B_r \cdot \dots \cdot B_1 \cdot A.$$

Man kann nun B durch weitere Zeilenumformungen zur Einheitsmatrix E_n machen. Dazu beseitigt man zunächst $b_{1n}, \dots, b_{n-1,n}$ mit Hilfe der letzten Zeile, dann $b_{1,n-1}, \dots, b_{n-2,n-1}$ mit Hilfe der vorletzten Zeile, usw. Schließlich normiert man die Komponenten in der Diagonalen auf 1. Es gibt also nach 2.7.1 weitere Elementarmatrizen $B_{r+1}, \dots, B_s$, so daß

$$E_n = B_s \cdot \dots \cdot B_{r+1} B = B_s \cdot \dots \cdot B_1 \cdot A.$$

Daraus folgt

$$A^{-1} = B_s \cdot \dots \cdot B_1, \quad \text{also} \quad A = B_1^{-1} \cdot \dots \cdot B_s^{-1},$$

und die Behauptung folgt aus 2.7.2. $\square$

2.7.4. Der Beweis von Satz 2.7.3 gestattet nun, ein einfaches Rechenverfahren für die Bestimmung der inversen Matrix anzugeben. Es hat die angenehme zusätzliche Eigenschaft, daß man von der gegebenen quadratischen Matrix im voraus gar nicht zu wissen braucht, ob sie invertierbar ist. Das stellt sich im Laufe der Rechnung heraus.

Sei also $A \in M(n \times n; K)$ gegeben. Man schreibt die Matrizen A und E_n nebeneinander. Alle Umformungen, die im folgenden an A vorgenommen werden, führt man parallel an E_n durch.

Zunächst bringt man A durch Zeilenumformungen auf Zeilenstufenform. Dabei stellt sich heraus, ob

$$\text{Zeilenumrang } A = n,$$

d.h. ob A invertierbar ist (vgl. 2.5.6). Ist der Zeilenumrang von A kleiner als n , so kann man aufhören; die Umformungen mit E_n waren dann umsonst. Ist der Zeilenumrang von A gleich n , so führt man weitere Zeilenumformungen durch, bis aus A die Matrix E_n geworden ist. Schematisch sieht das so aus (die Umformungen sind als Multiplikation mit Elementarmatrizen beschrieben):

A	E_n
$B_1 \cdot A$	$B_1 \cdot E_n$
$\vdots$	$\vdots$
$B_s \cdot \dots \cdot B_1 \cdot A$	$B_s \cdot \dots \cdot B_1 \cdot E_n$

Ist nun links aus A die Einheitsmatrix E_n entstanden, so hat sich rechts aus E_n die inverse Matrix A^{-1} aufgebaut, denn aus

$$B_s \cdot \dots \cdot B_1 \cdot A = E_n$$

folgt

$$B_s \cdot \dots \cdot B_1 \cdot E_n = B_s \cdot \dots \cdot B_1 = A^{-1}.$$

Diese erste schöne Anwendung wird den Leser hoffentlich schon vom Wert der Elementarmatrizen überzeugen.

Vorsicht! Anstelle von Zeilenumformungen kann man auch *ausschließlich* Spaltenumformungen anwenden. Aber bei *abwechselnder* Anwendung von Zeilen- und Spaltenumformungen funktioniert das Verfahren im allgemeinen nicht.

2.7.5. Beispiele. a)

$$\begin{array}{l}
 A = \begin{array}{c} \begin{array}{ccc|ccc} 0 & 1 & -4 & 1 & 0 & 0 \\ 1 & 2 & -1 & 0 & 1 & 0 \\ 1 & 1 & 2 & 0 & 0 & 1 \end{array} \\ P_2^1 \\ \begin{array}{ccc|ccc} 1 & 2 & -1 & 0 & 1 & 0 \\ 0 & 1 & -4 & 1 & 0 & 0 \\ 1 & 1 & 2 & 0 & 0 & 1 \end{array} \\ Q_3^1(-1) \\ \begin{array}{ccc|ccc} 1 & 2 & -1 & 0 & 1 & 0 \\ 0 & 1 & -4 & 1 & 0 & 0 \\ 0 & -1 & 3 & 0 & -1 & 1 \end{array} \\ Q_3^2 \\ \begin{array}{ccc|ccc} 1 & 2 & -1 & 0 & 1 & 0 \\ 0 & 1 & -4 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & -1 & 1 \end{array} \\ S_3(-1) \\ \begin{array}{ccc|ccc} 1 & 2 & -1 & 0 & 1 & 0 \\ 0 & 1 & -4 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 & -1 \end{array} \\ Q_1^2(-2) \\ \begin{array}{ccc|ccc} 1 & 0 & 7 & -2 & 1 & 0 \\ 0 & 1 & -4 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 & -1 \end{array} \\ Q_1^3(-7) \\ \begin{array}{ccc|ccc} 1 & 0 & 0 & 5 & -6 & 7 \\ 0 & 1 & -4 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 & -1 \end{array} \\ Q_2^3(4) \\ \begin{array}{ccc|ccc} 1 & 0 & 0 & 5 & -6 & 7 \\ 0 & 1 & 0 & -3 & 4 & -4 \\ 0 & 0 & 1 & -1 & 1 & -1 \end{array} \end{array} = A^{-1}.
 \end{array}$$

Man berechne zur Kontrolle $A \cdot A^{-1}$!

b)

$$A = \begin{array}{c} \\ Q_3^1(-1) \\ \\ Q_3^2 \end{array} \begin{array}{|cc|cc|} \hline 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \\ \hline 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 & 0 & 1 \\ \hline 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 & 1 \\ \hline \end{array}$$

A ist nicht invertierbar, denn Zeilenrang $A = 2$.

2.7.6. Ist $A \in M(m \times n; K)$ und

$$A: K^n \rightarrow K^m, \quad x \mapsto Ax,$$

die zugehörige lineare Abbildung, so gibt es nach 2.6.5 Transformationsmatrizen $S \in GL(m; K)$ und $T \in GL(n; K)$ mit

$$SAT^{-1} = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix},$$

wobei $r = \text{rang } A$. Wir leiten nun ein *Rechenverfahren für die Bestimmung von S und T* ab. Dazu betrachten wir folgendes Schema:

E_m	A	
$B_1 \cdot E_m$	$B_1 \cdot A$	
$\vdots$	$\vdots$	
$B_k \cdot \dots \cdot B_1 \cdot E_m$	$B_k \cdot \dots \cdot B_1 \cdot A$	E_n
	$B_k \cdot \dots \cdot B_1 \cdot A \cdot C_1$	$E_n \cdot C_1$
	$\vdots$	$\vdots$
	$B_k \cdot \dots \cdot B_1 \cdot A \cdot C_1 \cdot \dots \cdot C_l$	$E_n \cdot C_1 \cdot \dots \cdot C_l$

Zunächst wird A durch Zeilenumformungen auf Zeilenstufenform gebracht. Die Zeilenumformungen entsprechen der Multiplikation von links mit m -reihigen Elementarmatrizen $B_1, \dots, B_k$. Diese Umformungen führt man parallel an

E_m durch. Da die Matrix

$$B_k \cdot \dots \cdot B_1 \cdot A$$

Zeilenstufenform hat, kann man sie durch Spaltenumformungen auf die Form

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

mit $r = \text{rang } A$ bringen. Dies entspricht Multiplikation von rechts mit n -reihigen Elementarmatrizen $C_1, \dots, C_l$. Diese Spaltenumformungen führt man parallel an E_n durch. Wegen

$$B_k \cdot \dots \cdot B_1 \cdot A \cdot C_1 \cdot \dots \cdot C_l = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

sind durch

$$S = B_k \cdot \dots \cdot B_1 = B_k \cdot \dots \cdot B_1 E_m \quad \text{und} \quad T^{-1} = C_1 \cdot \dots \cdot C_l = E_n C_1 \cdot \dots \cdot C_l$$

Transformationsmatrizen der gewünschten Art gefunden.

Beispiel. Sei $K = \mathbb{R}$ und $A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 2 & 1 \end{pmatrix}$.

$$S = \begin{array}{|c|c|c|c|c|} \hline 1 & 0 & 1 & 2 & 0 \\ 0 & 1 & 2 & 2 & 1 \\ \hline \end{array} = A$$

$$S = \begin{array}{|c|c|c|c|c|c|c|} \hline 1 & 0 & 1 & 2 & 0 & 1 & 0 & 0 \\ -2 & 1 & 0 & -2 & 1 & 0 & 1 & 0 \\ \hline & & & & & 0 & 0 & 1 \\ \hline & & 1 & 0 & 2 & 1 & 0 & 0 \\ & & 0 & 1 & -2 & 0 & 0 & 1 \\ & & & & & 0 & 1 & 0 \\ \hline & & 1 & 0 & 0 & 1 & 0 & -2 \\ & & 0 & 1 & -2 & 0 & 0 & 1 \\ & & & & & 0 & 1 & 0 \\ \hline & & 1 & 0 & 0 & 1 & 0 & -2 \\ & & 0 & 1 & 0 & 0 & 0 & 1 \\ & & & & & 0 & 1 & 2 \\ \hline \end{array} = T^{-1}.$$

Ist

$$D = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix},$$

so erhält man auch Basen $\mathcal{A}$ und $\mathcal{B}$ von K^n und K^m , bezüglich derer A durch D beschrieben wird. Dazu betrachten wir das Diagramm

$$\begin{array}{ccc} K^n & \xrightarrow{D} & K^m \\ \uparrow T & & \uparrow S \\ K^n & \xrightarrow{A} & K^m, \end{array}$$

das wegen $D = SAT^{-1}$ kommutativ ist. $\mathcal{A}$ und $\mathcal{B}$ sind die Bilder der kanonischen Basen $\mathcal{K}$ und $\mathcal{K}'$ von K^n und K^m unter den Isomorphismen T^{-1} und S^{-1} . Also erhält man $\mathcal{A}$ und $\mathcal{B}$ als Spaltenvektoren von T^{-1} und S^{-1} . Dazu muß man S noch invertieren.

In unserem Beispiel ist

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

also sind

$$((1, 0, 0), (0, 0, 1), (-2, 1, 2)) \quad \text{und} \quad ((1, 2), (0, 1))$$

Basen der gesuchten Art. Zur Kontrolle prüft man nach:

$$A \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad A \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{und} \quad A \cdot \begin{pmatrix} -2 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

2.7.7. Natürlich kann man auch das Gaußsche Eliminationsverfahren mit Hilfe von Elementarmatrizen beschreiben. Sei das System

$$A \cdot x = b$$

gegeben. Die elementaren Zeilenumformungen von A und (A, b) werden bewirkt durch Multiplikation von links mit Elementarmatrizen aus $GL(m; K)$. Ihr Produkt ergibt eine Matrix

$$S \in GL(m; K) \quad \text{mit} \quad (\tilde{A}, \tilde{b}) = S \cdot (A, b) = (SA, Sb),$$

wobei $(\tilde{A}, \tilde{b})$ die auf Zeilenstufenform gebrachte erweiterte Koeffizientenmatrix ist. Man beachte, daß S allein durch A bestimmt ist. Die Berechnung von S kann

man wieder schematisch durchführen, indem man die Zeilenumformungen von A parallel an E_m durchführt.

E_m	A
$B_1 \cdot E_m$	$B_1 \cdot A$
$\vdots$	$\vdots$
$B_s \cdot \dots \cdot B_1 \cdot E_m$	$B_s \cdot \dots \cdot B_1 \cdot A$

Ist nun aus A die Matrix $\tilde{A} = B_s \cdot \dots \cdot B_1 \cdot A$ in Zeilenstufenform entstanden, so hat sich links die Matrix

$$S = B_s \cdot \dots \cdot B_1$$

aufgebaut. Damit kann man sofort entscheiden, ob für ein $b \in K^m$ Lösungen von $Ax = b$ existieren. Man berechnet

$$S \cdot b = \tilde{b} = \begin{pmatrix} \tilde{b}_1 \\ \vdots \\ \tilde{b}_r \\ \tilde{b}_{r+1} \\ \vdots \\ \tilde{b}_m \end{pmatrix}$$

und sieht nach, ob $\tilde{b}_{r+1} = \dots = \tilde{b}_m = 0$, wobei $r = \text{rang } A = \text{rang } \tilde{A}$ (vgl. Aufgabe 5).

Aufgaben zu 2.7

1. Stellen Sie die folgende Matrix A als Produkt von Elementarmatrizen dar:

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix}.$$

2. Sind die folgenden Matrizen invertierbar? Wenn ja, dann geben die inverse Matrix an.

$$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \in M(4 \times 4; \mathbb{R}), \quad \begin{pmatrix} 6 & 3 & 4 & 5 \\ 1 & 2 & 2 & 1 \\ 2 & 4 & 3 & 2 \\ 3 & 3 & 4 & 2 \end{pmatrix} \in M(4 \times 4; \mathbb{R}),$$

$$\begin{pmatrix} 1 & 2 & 0 \\ 1 & 1 & 1 \\ 2 & 0 & 1 \end{pmatrix} \in M(3 \times 3; \mathbb{R}), \quad \begin{pmatrix} 1 & 2 & 0 \\ 1 & 1 & 1 \\ 2 & 0 & 1 \end{pmatrix} \in M(3 \times 3; \mathbb{Z}/3\mathbb{Z}).$$

3. Zeigen Sie:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M(2 \times 2; K) \text{ ist invertierbar} \Leftrightarrow ad - bc \neq 0.$$

Berechnen Sie in diesem Fall die Inverse von A .

4. Modifizieren Sie das Rechenverfahren aus 2.7.6 so, dass man statt S die inverse Matrix S^{-1} erhält (benutzen Sie dabei die Inversen der Elementarmatrizen aus 2.7.2).

5. Finden Sie für die Gleichungssysteme $Ax = b$ aus 0.3.5 sowie aus Aufgabe 2 in 0.4 jeweils eine Matrix S , so dass $\tilde{A} = SA$ in Zeilenstufenform ist, und berechnen Sie $\tilde{b} = Sb$.

6. Beweisen Sie:

a) Für $A \in M(n \times n; K)$ und $m \in \mathbb{N}$ gilt:

$$E_n - A^m = (E_n - A) \left(\sum_{i=0}^{m-1} A^i \right) = \left(\sum_{i=0}^{m-1} A^i \right) (E_n - A).$$

(Dabei sei $A^0 := E_n$.)

b) Ist $A \in M(n \times n; K)$ eine Matrix, für die ein $m \in \mathbb{N}$ existiert mit $A^m = 0$, so ist $E_n - A$ invertierbar. Wie sieht die inverse Matrix aus?

Kapitel 3

Determinanten

In den vorhergehenden Kapiteln wurde laufend mit Linearkombinationen gerechnet, das gilt als der „triviale“ Teil der linearen Algebra. Nun steigen wir eine Stufe höher zur Determinante, das ist eine Zahl, die man einer quadratischen Matrix zuordnet. LEIBNIZ gab schon um 1690 eine Formel zur Berechnung dieser Zahl an ([Kow 1], §1). WEIERSTRASS benutzte in seinen Vorlesungen eine andere Methode: Er führte die Determinante mit axiomatisch angegebenen Eigenschaften ein. Dadurch kann man die chronischen Vorzeichenprobleme erst einmal im Hintergrund halten und all das bereitstellen, was man zur praktischen Berechnung der Determinanten benötigt. Es zeigt sich, daß auch hier das Verfahren aus Kapitel 0 zur Überführung einer Matrix in Zeilenstufenform zum Ziel führt. Diesen Weg haben wir mit Rücksicht auf eilige Leser in 3.1 beschritten. Die Vorzeichenspiele werden systematisch in 3.2 vorgeführt.

Für genauere historische Hinweise seien die entsprechenden Exkurse in [B1], [Fr], [Koe] und [Kow1] zur Lektüre empfohlen.

3.1 Beispiele und Definitionen

3.1.1. Zunächst geben wir zwei charakteristische Beispiele für das Auftreten von Determinanten.

1) Gegeben sei ein lineares Gleichungssystem

$$a_{11}x_1 + a_{12}x_2 = b_1, \quad \text{I}$$

$$a_{21}x_1 + a_{22}x_2 = b_2. \quad \text{II}$$

Wir suchen eine Formel für die Lösung, die man bei beliebigen Werten der Koeffizienten anwenden kann. Umformungen ergeben die Gleichungen

$$(a_{11}a_{22} - a_{12}a_{21})x_1 = a_{22}b_1 - a_{12}b_2, \quad a_{22}\text{I} - a_{12}\text{II}$$

$$(a_{11}a_{22} - a_{12}a_{21})x_2 = a_{11}b_2 - a_{21}b_1. \quad a_{11}\text{II} - a_{21}\text{I}$$

Definieren wir allgemein für eine (2×2) -Matrix eine *Determinante* durch

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} := ad - bc,$$

so erhält man für die Lösungen des obigen Systems

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}, \quad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}.$$

Das ist der einfachste Fall der *CRAMERschen Regel*. Sie versagt, wenn

$$a_{11}a_{22} - a_{12}a_{21} = 0, \quad \text{also} \quad a_{22}(a_{11}, a_{12}) - a_{12}(a_{21}, a_{22}) = (0, 0).$$

Das bedeutet, daß der Rang der Koeffizientenmatrix kleiner als 2 ist. Entsprechend 2.3 gibt es hier keine eindeutige Lösung.

2) Die Fläche eines *Dreiecks* ist gegeben durch

$$\frac{1}{2} (\text{Grundlinie mal Höhe}).$$

Das zeigt man mit Hilfe des *CAVALIERIschen Prinzips*, indem man zuerst das Dreieck verdoppelt zu einem *Parallelogramm*, und dieses dann verschiebt zu einem Rechteck.

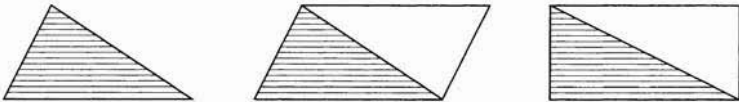


Bild 3.1

Zur Berechnung der Fläche eines Parallelogramms in der Ebene nehmen wir an, dieses sei durch zwei Vektoren

$$v = (a_1, a_2) \quad \text{und} \quad w = (b_1, b_2)$$

gegeben.

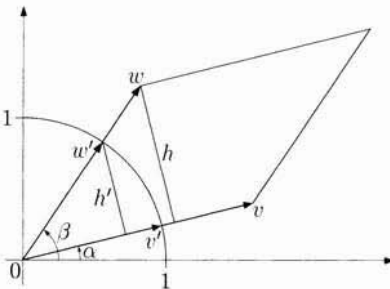


Bild 3.2

Wie in der Zeichnung angedeutet, ist

$$v = \varrho \cdot v' = \varrho \cdot (\cos \alpha, \sin \alpha), \quad w = \sigma \cdot w' = \sigma \cdot (\cos \beta, \sin \beta),$$

wobei $\varrho, \sigma > 0$. Bezeichnet F bzw. F' die Fläche des von v und w bzw. v' und w' aufgespannten Parallelogramms, so verläuft die Rechnung für $0 \leq \beta - \alpha \leq \pi$ wie folgt:

$$h' = \sin(\beta - \alpha) = \cos \alpha \sin \beta - \cos \beta \sin \alpha = \begin{vmatrix} \cos \alpha & \sin \alpha \\ \cos \beta & \sin \beta \end{vmatrix},$$

also

$$F = \varrho \cdot \sigma \cdot F' = \varrho \cdot \sigma \cdot h' = \begin{vmatrix} \varrho \cdot \cos \alpha & \varrho \cdot \sin \alpha \\ \sigma \cdot \cos \beta & \sigma \cdot \sin \beta \end{vmatrix} = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}.$$

Die Fläche ist also gleich der Determinante. Daher kann man einige Eigenschaften der Abbildung

$$\det: M(2 \times 2; \mathbb{R}) \rightarrow \mathbb{R}, \quad A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto ad - bc = \det A$$

geometrisch illustrieren. Dazu die folgende Notation: sind

$$v = (a, b) \quad \text{und} \quad w = (c, d)$$

die Zeilenvektoren von A , so schreiben wir $\det A = \det \begin{pmatrix} v \\ w \end{pmatrix}$.

a) Für $\lambda, \mu \in \mathbb{R}$ ist

$$\det \begin{pmatrix} \lambda v \\ w \end{pmatrix} = \lambda \cdot \det \begin{pmatrix} v \\ w \end{pmatrix}, \quad \det \begin{pmatrix} v \\ \mu w \end{pmatrix} = \mu \cdot \det \begin{pmatrix} v \\ w \end{pmatrix}.$$

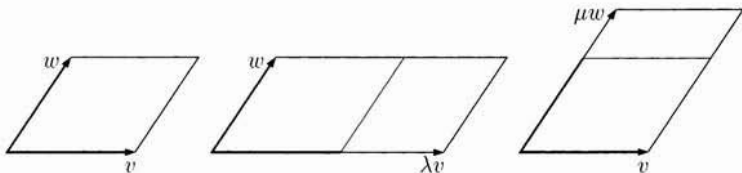


Bild 3.3

Das bedeutet, daß die Fläche so gestreckt wird wie einzelne Seiten.

b) Für $\lambda \in \mathbb{R}$ ist

$$\det \begin{pmatrix} v \\ w \end{pmatrix} = \det \begin{pmatrix} v \\ w + \lambda v \end{pmatrix},$$

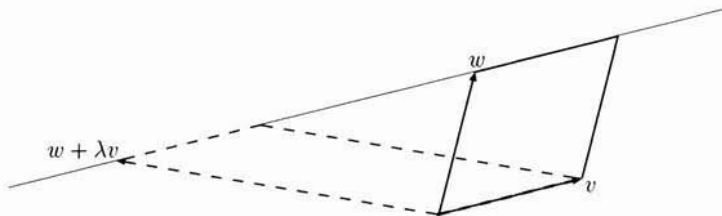


Bild 3.4

was die Invarianz der Fläche unter *Scherungen* nach dem Cavalierischen Prinzip bedeutet.

c)
$$\det \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = 1.$$

Das bedeutet, das Einheitsquadrat hat die Fläche 1.

d)
$$\det \begin{pmatrix} w \\ v \end{pmatrix} = -\det \begin{pmatrix} v \\ w \end{pmatrix}.$$

Daran sieht man, daß nicht nur der Betrag, sondern auch das Vorzeichen der Determinante eine geometrische Bedeutung hat. Es hängt von der Orientierung des Paares v, w ab, darauf kommen wir in 3.4 zurück.

e)
$$\det \begin{pmatrix} v \\ w \end{pmatrix} = 0$$

ist gleichbedeutend mit der linearen Abhängigkeit von v und w , d.h. das Parallelogramm hat die Fläche Null.

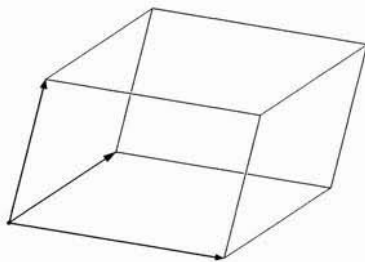


Bild 3.5

So wie zwei Vektoren im $\mathbb{R}^2$ ein Parallelogramm erklären, spannen drei Vektoren im $\mathbb{R}^3$ einen *Spat* und n Vektoren im $\mathbb{R}^n$ ein *Parallelotop* auf, und es entsteht das

Problem, den Inhalt zu berechnen. Wie man in der Analysis lernt, ist das die Determinante der $n \times n$ -Matrix mit den Vektoren als Zeilen ([Fo 3], §5).

3.1.2. Zur Erklärung der Determinante einer $n \times n$ -Matrix gibt es mehrere Möglichkeiten, zwei davon sind:

1) Eine Formel, in der die Einträge vorkommen, so wie das oben bei (2×2) -Matrizen angegeben war. Das hatte schon LEIBNIZ bei größeren Matrizen ausgeführt, das Ergebnis – die allgemeine Leibniz-Formel in 3.2.5 – ist leider ziemlich umständlich.

2) Eine Charakterisierung der Determinante durch Axiome, sie geht auf WEIERSTRASS zurück (vgl. [Fr]). Das ist nicht nur eleganter, sondern ergibt auch einfachere Methoden zur Berechnung als die Leibniz-Formel.

Nun zur axiomatischen Einführung der Determinante, wir benutzen dabei die folgende Notation:

Ist A eine n -reihige quadratische Matrix, so bezeichnen wir stets mit $a_1, \dots, a_n$ die Zeilenvektoren von A . Dann schreiben wir

$$A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

Definition. Sei K ein Körper und n eine von Null verschiedene natürliche Zahl. Eine Abbildung

$$\det: M(n \times n; K) \rightarrow K, \quad A \mapsto \det A,$$

heißt *Determinante*, falls folgendes gilt:

D1 $\det$ ist *linear in jeder Zeile*. Genauer heißt das folgendes. Für jeden Index $i \in \{1, \dots, n\}$ gilt:

a) Ist $a_i = a'_i + a''_i$, so ist

$$\det \begin{pmatrix} \vdots \\ a_i \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ a'_i \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ a''_i \\ \vdots \end{pmatrix}.$$

b) Ist $a_i = \lambda a'_i$, so ist

$$\det \begin{pmatrix} \vdots \\ a_i \\ \vdots \end{pmatrix} = \lambda \cdot \det \begin{pmatrix} \vdots \\ a'_i \\ \vdots \end{pmatrix}.$$

An den mit Punkten bezeichneten Stellen stehen dabei jeweils unverändert die Zeilenvektoren $a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_n$.

D2 $\det$ ist *alternierend*, d.h. hat A zwei gleiche Zeilen, so ist $\det A = 0$.

D3 $\det$ ist *normiert*, d.h. $\det E_n = 1$.

Eine andere Schreibweise für die Determinante ist

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} := \det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

Man beachte dabei, daß die senkrechten Striche nichts mit einem Absolutbetrag zu tun haben.

3.1.3. Diese Definition ist sehr einfach, aber es bleibt die Existenz und Eindeutigkeit zu zeigen, und das wird etwas Mühe machen. Zunächst spielen wir mit den Axiomen und leiten daraus weitere Regeln ab.

Satz. Eine Determinante $\det: M(n \times n; K) \rightarrow K$ hat die folgenden weiteren Eigenschaften:

D4 Für jedes $\lambda \in K$ ist $\det(\lambda \cdot A) = \lambda^n \cdot \det A$.

D5 Ist eine Zeile von A gleich Null, so ist $\det A = 0$.

D6 Entsteht B aus A durch eine Zeilenvertauschung, so ist

$$\det B = -\det A.$$

Die Determinante ändert also bei Zeilenumformungen vom Typ IV ihr Vorzeichen.

D7 Ist $\lambda \in K$, und entsteht B aus A durch Addition der λ -fachen j -ten Zeile zur i -ten Zeile ($i \neq j$), so ist

$$\det B = \det A.$$

Die Determinante bleibt also bei Zeilenumformungen vom Typ III unverändert.

D8 Ist A eine obere Dreiecksmatrix, also

$$A = \begin{pmatrix} \lambda_1 & \cdots & \\ & \ddots & \vdots \\ 0 & & \lambda_n \end{pmatrix},$$

so ist $\det A = \lambda_1 \cdot \dots \cdot \lambda_n$.

D9 Sei $n \geq 2$ und $A \in M(n \times n; K)$ von der Gestalt

$$A = \begin{pmatrix} A_1 & C \\ 0 & A_2 \end{pmatrix},$$

wobei A_1 und A_2 quadratisch sind. Dann gilt

$$\det A = (\det A_1) \cdot (\det A_2).$$

D10 $\det A = 0$ ist gleichbedeutend mit $\text{rang } A < n$.

D11 Es gilt der Determinanten-Multiplikationssatz

$$\det(A \cdot B) = \det A \cdot \det B$$

für alle $A, B \in M(n \times n; K)$. Insbesondere gilt für $A \in GL(n; K)$

$$\det A^{-1} = (\det A)^{-1}.$$

Vorsicht! Die „Regel“ $\det(A + B) = \det A + \det B$ ist für $n \geq 2$ falsch.

Beweis. D4 und D5 folgen sofort aus D1 b). Zum Beweis von D6 nehmen wir an, daß die Zeilen $i < j$ vertauscht werden. Dann ist wegen D1 a) und D2

$$\begin{aligned} \det A + \det B &= \det \begin{pmatrix} a_i \\ a_j \end{pmatrix} + \det \begin{pmatrix} a_j \\ a_i \end{pmatrix} \\ &= \det \begin{pmatrix} a_i \\ a_i \end{pmatrix} + \det \begin{pmatrix} a_i \\ a_j \end{pmatrix} + \det \begin{pmatrix} a_j \\ a_i \end{pmatrix} + \det \begin{pmatrix} a_j \\ a_j \end{pmatrix} \\ &= \det \begin{pmatrix} a_i + a_j \\ a_i + a_j \end{pmatrix} = 0. \end{aligned}$$

Dabei sind zur Vereinfachung der Schreibweise nur die Einträge der Zeilen i und j angegeben, in den restlichen Zeilen ändert sich nichts.

Es sei bemerkt, daß D2 aus D6 folgt, wenn $\text{char}(K) \neq 2$, denn hat A zwei gleiche Zeilen, so ist nach D6

$$\det A = -\det A, \quad \text{also} \quad 2 \det A = 0.$$

D7: Wegen D1 und D2 ist

$$\det B = \det \begin{pmatrix} a_i + \lambda a_j \\ a_j \end{pmatrix} = \det A + \lambda \det \begin{pmatrix} a_j \\ a_j \end{pmatrix} = \det A.$$

D8: Sind alle $\lambda_i \neq 0$, so folgt durch wiederholte Anwendung von D7

$$\det A = \det \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = \lambda_1 \cdot \dots \cdot \lambda_n \cdot \det E_n = \lambda_1 \cdot \dots \cdot \lambda_n.$$

Gibt es ein i mit $\lambda_i = 0$, so wählen wir i maximal, d.h. $\lambda_{i+1}, \dots, \lambda_n \neq 0$. Mit Hilfe von $\lambda_{i+1}, \dots, \lambda_n$ räumt man den Rest der i -ten Zeile aus, und mit D7 und D5 folgt $\det A = 0$.

D9: Durch Zeilenumformungen vom Typ III und IV an A mache man A_1 zu einer oberen Dreiecksmatrix B_1 . Dabei bleibt A_2 unverändert, aus C werde C' . Ist k die Anzahl der ausgeführten Zeilenvertauschungen, so ist

$$\det A_1 = (-1)^k \cdot \det B_1.$$

Dann mache man A_2 durch Zeilenumformungen vom Typ III und IV an A zu einer oberen Dreiecksmatrix. Dabei bleiben B_1 und C' unverändert. Ist l die Anzahl der ausgeführten Zeilenvertauschungen, so ist

$$\det A_2 = (-1)^l \cdot \det B_2.$$

Ist

$$B := \begin{pmatrix} B_1 & C' \\ 0 & B_2 \end{pmatrix},$$

so sind B , B_1 und B_2 obere Dreiecksmatrizen, es ist also nach D8 offensichtlich

$$\det B = (\det B_1) \cdot (\det B_2).$$

Wegen

$$\det A = (-1)^{k+l} \cdot \det B$$

folgt die Behauptung.

D10: Durch Zeilenumformungen vom Typ III und IV bringen wir A auf Zeilenstufenform B . Dann ist B obere Dreiecksmatrix, also

$$B = \begin{pmatrix} \lambda_1 & \cdots & \\ & \ddots & \vdots \\ 0 & & \lambda_n \end{pmatrix},$$

und nach D6 und D7 ist $\det B = \pm \det A$. Weiter ist $\text{rang } A = \text{rang } B$ und wegen D8

$$\text{rang } B = n \Leftrightarrow \det B = \lambda_1 \cdot \dots \cdot \lambda_n \neq 0.$$

D11: Ist $\text{rang } A < n$, so ist $\text{rang } (A \cdot B) < n$, und die Gleichung lautet $0 = 0$ nach D10.

Andernfalls können wir $A \in \text{GL}(n; K)$ annehmen. Nach 2.7.3 gibt es Elementarmatrizen $C_1, \dots, C_s$, so daß

$$A = C_1 \cdot \dots \cdot C_s.$$

Es genügt also zu zeigen, daß für jede Elementarmatrix C vom Typ $S_i(\lambda)$ oder Q_i^j (vgl. 2.7.1)

$$\det(C \cdot B) = \det C \cdot \det B$$

gilt. Nach Eigenschaft D8 (die natürlich auch für untere Dreiecksmatrizen gilt) ist

$$\det S_i(\lambda) = \lambda \quad \text{und} \quad \det Q_i^j = 1.$$

Multiplizieren von links mit $S_i(\lambda)$ multipliziert die i -te Zeile von B mit λ , also ist

$$\det(S_i(\lambda) \cdot B) = \lambda \cdot \det B$$

nach D1. Multiplizieren von links mit Q_i^j bewirkt die Addition der j -ten zur i -ten Zeile, also ist

$$\det(Q_i^j \cdot B) = 1 \cdot \det B. \quad \square$$

Dieser letzte Beweis wird den Leser hoffentlich in seiner Wertschätzung der Elementarmatrizen bestärken.

3.1.4. Für die Praxis der Berechnung von Determinanten hat man nun alle erforderlichen Hilfsmittel zur Verfügung. Man bringt A durch Zeilenumformungen vom Typ III und IV auf obere Dreiecksgestalt B . Ist k die Anzahl der dabei durchgeführten Zeilenvertauschungen, so gilt

$$\det A = (-1)^k \cdot \det B = (-1)^k \cdot \lambda_1 \cdot \dots \cdot \lambda_n.$$

Beispiele. a)

$$\begin{vmatrix} 0 & 1 & 2 \\ 3 & 2 & 1 \\ 1 & 1 & 0 \end{vmatrix} = - \begin{vmatrix} 1 & 1 & 0 \\ 3 & 2 & 1 \\ 0 & 1 & 2 \end{vmatrix} = - \begin{vmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & 2 \end{vmatrix} = - \begin{vmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 3 \end{vmatrix} = 3.$$

b) Die Berechnung von Determinanten wird interessanter, wenn man die Einträge a_{ij} der Matrix als *Unbestimmte* auffaßt, das sind Zahlen, für die man beliebige Elemente des Körpers einsetzen kann, und zwar unabhängig voneinander.

Es ist üblich, das dadurch anzudeuten, daß man statt a den Buchstaben x verwendet. Auf diese Weise berechnet man mit Hilfe von D7 und D8

$$\begin{vmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{vmatrix} = \begin{vmatrix} x_{11} & x_{12} \\ 0 & x_{22} - \frac{x_{21}}{x_{11}}x_{12} \end{vmatrix} = x_{11}x_{22} - x_{21}x_{12}. \quad (*)$$

Man beachte, daß x_{11} während der Rechnung vorübergehend im Nenner steht, nicht aber am Anfang und am Ende. Mit Hilfe von D6 kann man noch einmal direkt überprüfen, daß die Formel (*) auch für $x_{11} = 0$ gilt.

c) Eine Matrix $A = (a_{ij}) \in M(n \times n; K)$ heißt *schiefsymmetrisch*, wenn $a_{ij} = -a_{ji}$ und $a_{ii} = 0$ (im Fall $\text{char} K \neq 2$ folgt die zweite Bedingung aus der ersten). Die Berechnung solcher Determinanten ist besonders interessant, wir betrachten die Einträge wieder als Unbestimmte. Für $n = 2$ und 3 ist

$$\begin{vmatrix} 0 & x_{12} \\ -x_{12} & 0 \end{vmatrix} = - \begin{vmatrix} x_{12} & 0 \\ 0 & -x_{12} \end{vmatrix} = x_{12}^2,$$

$$\begin{vmatrix} 0 & x_{12} & x_{13} \\ -x_{12} & 0 & x_{23} \\ -x_{13} & -x_{23} & 0 \end{vmatrix} = \begin{vmatrix} 0 & x_{12} & x_{13} \\ -x_{12} & 0 & x_{23} \\ 0 & 0 & 0 \end{vmatrix} = 0.$$

Dabei wurde zu Zeile III die Kombination $\frac{x_{23}}{x_{12}}\text{I} - \frac{x_{13}}{x_{12}}\text{II}$ addiert. Das ist ungefährlich, denn für $x_{12} = 0$ ist das Ergebnis ohnehin klar. Nun zum Fall $n = 4$:

$$\det A = \begin{vmatrix} 0 & x_{12} & x_{13} & x_{14} \\ -x_{12} & 0 & x_{23} & x_{24} \\ -x_{13} & -x_{23} & 0 & x_{34} \\ -x_{14} & -x_{24} & -x_{34} & 0 \end{vmatrix} = \begin{vmatrix} 0 & x_{12} & x_{13} & x_{14} \\ -x_{12} & 0 & x_{23} & x_{24} \\ 0 & 0 & 0 & Q(x) \\ 0 & 0 & -Q(x) & 0 \end{vmatrix}$$

durch geeignete Umformungen der Zeilen III und IV, wobei

$$Q(x) = \frac{P(x)}{x_{12}} \quad \text{mit} \quad P(x) := x_{12}x_{34} - x_{13}x_{24} + x_{14}x_{23}.$$

Aus D9 folgt schließlich

$$\det A = (P(x))^2.$$

Man nennt $P(x)$ ein *PAFFSches Polynom* (vgl. Aufgabe 5 und Aufgabe 8 zu 3.2).

Aufgaben zu 3.1

1. Berechnen Sie die Determinanten von

$$\begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 1 \\ 2 & 7 & 9 \end{pmatrix}.$$

2. Zeigen Sie:

$$\det \begin{pmatrix} x & 1 & 1 \\ 1 & x & 1 \\ 1 & 1 & x \end{pmatrix} = (x-1)^2(x+2),$$

$$\det \begin{pmatrix} a^2+1 & ab & ac \\ ab & b^2+1 & bc \\ ac & bc & c^2+1 \end{pmatrix} = a^2 + b^2 + c^2 + 1.$$

3. Berechnen Sie:

$$\det \begin{pmatrix} \sin \alpha & \cos \alpha & a \sin \alpha & b \cos \alpha & ab \\ -\cos \alpha & \sin \alpha & -a^2 \sin \alpha & b^2 \cos \alpha & a^2 b^2 \\ 0 & 0 & 1 & a^2 & b^2 \\ 0 & 0 & 0 & a & b \\ 0 & 0 & 0 & -b & a \end{pmatrix}.$$

4. Zeigen Sie, dass für eine Matrix $A = (a_{ij}) \in M(n \times n; K)$ gilt:

$$\det(a_{ij}) = \det((-1)^{i+j} \cdot a_{ij}).$$

5. Sei K ein Körper mit $\text{char } K \neq 2$ und $A \in M(n \times n; K)$ alternierend (vgl. Aufgabe 3 zu 1.6). Zeigen Sie:

a) Ist n ungerade, so ist $\det A = 0$.

(Hinweis: Benutzen Sie Satz 3.2.6)

b) Ist n gerade, so ist $\det A$ Quadrat eines Polynoms in den Einträgen von A (vgl. Aufgabe 8 zu 3.2).

6. Sind $f = a_m t^m + \dots + a_0, g = b_n t^n + \dots + b_0 \in K[t]$ Polynome mit $\deg f = m, \deg g = n$, so ist die *Resultante* von f und g definiert durch

$$\operatorname{Res}_{f,g} := \det \left(\begin{array}{ccccccc} a_0 & & \cdots & & \cdots & & a_m \\ & \ddots & & & & & \\ & & a_0 & & \cdots & \cdots & a_m \\ b_0 & \cdots & \cdots & b_n & & & \\ & & \ddots & & \ddots & & \\ & & & & b_0 & \cdots & \cdots & b_n \end{array} \right) \left. \begin{array}{l} \\ \\ \\ \\ \\ \end{array} \right\} \begin{array}{l} n \text{ Zeilen} \\ \\ \\ m \text{ Zeilen} \end{array}$$

Zeigen Sie die Äquivalenz der folgenden Aussagen:

- i) $\operatorname{Res}_{f,g} = 0$.
- ii) $f, tf, \dots, t^{n-1}f, g, tg, \dots, t^{m-1}g$ sind linear abhängig.
- iii) Es existieren $p, q \in K[t]$, $p, q \neq 0$, mit $\deg p \leq n-1, \deg q \leq m-1$ und $pf = qg$.

Mit etwas Teilbarkeitstheorie von Polynomen kann man zeigen, dass i) bis iii) äquivalent sind zu

- iv) f und g haben einen gemeinsamen nichtkonstanten Teiler $h \in K[t]$.

Insbesondere ist also $\operatorname{Res}_{f,g} = 0$, falls f und g eine gemeinsame Nullstelle haben, und im Fall $K = \mathbb{C}$ gilt: $\operatorname{Res}_{f,g} = 0 \Leftrightarrow f$ und g haben eine gemeinsame Nullstelle.

3.2 Existenz und Eindeutigkeit

Bei der Berechnung der Determinante nach der Methode aus 3.1.4 bleibt eine kleine, wieder spitzfindig erscheinende Frage: die durchzuführenden Zeilenvertauschungen sind nicht eindeutig bestimmt, man hat Wahlmöglichkeiten. Aber das Ergebnis $(-1)^k \lambda_1 \dots \lambda_n$ muß unabhängig von allen Auswahlen sein, insbesondere muß klar sein, ob k gerade oder ungerade ist. Das wird besonders deutlich an dem charakteristischen Beispiel einer Einheitsmatrix mit veränderter Reihenfolge der Zeilen: Dazu kommen wir zurück auf die schon in 1.2.2 betrachteten Permutationen. Ist σ eine bijektive Abbildung von $\{1, \dots, n\}$ auf sich, und bezeichnen $e_1, \dots, e_n$ die kanonischen Basisvektoren, so betrachten wir die Matrix

$$E_\sigma := \begin{pmatrix} e_{\sigma(1)} \\ \vdots \\ e_{\sigma(n)} \end{pmatrix}$$

mit den Basisvektoren in permutierter Reihenfolge als Zeilen, und die Vorzeichen-Frage spitzt sich zu zur Alternative

$$\det E_\sigma = \pm 1 ?$$

Vorzeichen sind eine Art von Butterbroten: die haben zwei Möglichkeiten zu fallen ([E]). Zur Beantwortung der Vorzeichenfrage benötigen wir eine zuverlässige Methode, an der Permutation zu erkennen, auf welche Arten sie durch wiederholte Vertauschungen rückgängig gemacht werden kann.

3.2.1. Wie wir gerade gesehen haben, ist zunächst ein kleiner Exkurs über Permutationen unvermeidlich.

Wie in 1.2.2 bezeichnen wir für jede natürliche Zahl $n > 0$ mit S_n die *symmetrische Gruppe* von $\{1, \dots, n\}$, d.h. die Gruppe aller bijektiven Abbildungen

$$\sigma: \{1, \dots, n\} \rightarrow \{1, \dots, n\}.$$

Die Elemente von S_n nennen wir *Permutationen*. Das neutrale Element von S_n ist die identische Abbildung, die wir mit id bezeichnen. Wie üblich schreiben wir $\sigma \in S_n$ explizit in der Form

$$\sigma = \begin{bmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{bmatrix}.$$

Für $\sigma, \tau \in S_n$ ist dann

$$\begin{aligned} \tau \cdot \sigma &= \begin{bmatrix} 1 & \cdots & n \\ \tau(1) & \cdots & \tau(n) \end{bmatrix} \cdot \begin{bmatrix} 1 & \cdots & n \\ \sigma(1) & \cdots & \sigma(n) \end{bmatrix} \\ &= \begin{bmatrix} 1 & \cdots & n \\ \tau(\sigma(1)) & \cdots & \tau(\sigma(n)) \end{bmatrix}, \end{aligned}$$

zum Beispiel

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{bmatrix},$$

aber

$$\begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix}.$$

Man beachte dabei, daß die *rechts* stehende Permutation *zuerst* angewandt wird, wie das bei Abbildungen üblich ist.

Bemerkung. Die Gruppe S_n enthält

$$n! := n \cdot (n-1) \cdot \dots \cdot 2 \cdot 1$$

(sprich: *n*-Fakultät) Elemente. Für $n \geq 3$ ist S_n nicht abelsch.

Beweis. Wir überlegen, wie viele Möglichkeiten es gibt, Elemente $\sigma \in S_n$ aufzubauen. Zunächst hat man für die Auswahl von

$\sigma(1)$ genau n Möglichkeiten.

Da σ injektiv sein soll, muß $\sigma(2) \neq \sigma(1)$ sein. Es verbleiben für die Auswahl von

$\sigma(2)$ genau $n-1$ Möglichkeiten.

Sind schließlich $\sigma(1), \dots, \sigma(n-1)$ gewählt, so ist $\sigma(n)$ festgelegt, es gibt also für

$\sigma(n)$ nur eine Möglichkeit.

Insgesamt gibt es daher

$$n \cdot (n-1) \cdot \dots \cdot 2 \cdot 1 = n!$$

verschiedene Permutationen in S_n . Ist für $n \geq 3$

$$\sigma = \begin{bmatrix} 1 & 2 & 3 & 4 & \dots & n \\ 1 & 3 & 2 & 4 & \dots & n \end{bmatrix} \quad \text{und} \quad \tau = \begin{bmatrix} 1 & 2 & 3 & 4 & \dots & n \\ 2 & 3 & 1 & 4 & \dots & n \end{bmatrix},$$

so folgt wie oben $\tau \cdot \sigma \neq \sigma \cdot \tau$. Die Gruppen S_1 und S_2 sind abelsch, wie man sofort sieht. $\square$

Um sehen zu können, wie schnell $n!$ mit n wächst, benutzt man die Formel von STIRLING

$$n! \approx \sqrt{2\pi n} \left(\frac{n}{e}\right)^n,$$

wobei $\approx$ eine asymptotische Näherung bezeichnet ([Fo1], §20). Die Zahl $60!$ ist ungefähr gleich 10^{82} , das ist in etwa die geschätzte Zahl der Nukleonen des Universums. Eine (60×60) -Matrix ist dagegen in den Problemen der Anwendungen als klein anzusehen.

3.2.2. Um die Veränderung des Vorzeichens der Determinante bei Umordnung der Zeilen zu kontrollieren, vertauscht man mehrfach jeweils zwei Zeilen. Solche Permutationen haben einen eigenen Namen.

Eine Permutation $\tau \in S_n$ heißt *Transposition*, falls τ zwei Elemente aus $\{1, \dots, n\}$ vertauscht und alle übrigen fest läßt, d.h. wenn es $k, l \in \{1, \dots, n\}$ mit $k \neq l$ gibt, so daß gilt:

$$\begin{aligned}\tau(k) &= l, \\ \tau(l) &= k \quad \text{und} \\ \tau(i) &= i \quad \text{für } i \in \{1, \dots, n\} \setminus \{k, l\}.\end{aligned}$$

Offensichtlich gilt $\tau^{-1} = \tau$ für jede Transposition $\tau \in S_n$.

Daß man allein mit Vertauschungen von Zeilen auskommt, zeigt das

Lemma. Ist $n \geq 2$, so gibt es zu jedem $\sigma \in S_n$ (keineswegs eindeutig bestimmte) Transpositionen $\tau_1, \dots, \tau_k \in S_n$ mit

$$\sigma = \tau_1 \cdot \dots \cdot \tau_k.$$

Beweis. Ist $\sigma = \text{id}$ und $\tau \in S_n$ irgendeine Transposition, so ist

$$\text{id} = \tau \cdot \tau^{-1} = \tau \cdot \tau.$$

Andernfalls gibt es ein $i_1 \in \{1, \dots, n\}$ mit

$$\sigma(i) = i \quad \text{für } i = 1, \dots, i_1 - 1 \quad \text{und} \quad \sigma(i_1) \neq i_1, \quad \text{also sogar} \quad \sigma(i_1) > i_1.$$

Sei τ_1 die Transposition, die i_1 mit $\sigma(i_1)$ vertauscht, und $\sigma_1 := \tau_1 \cdot \sigma$. Dann ist

$$\sigma_1(i) = i \quad \text{für } i = 1, \dots, i_1.$$

Entweder ist nun $\sigma_1 = \text{id}$, oder es gibt ein i_2 mit $i_2 > i_1$ und

$$\sigma_1(i) = i \quad \text{für } i = 1, \dots, i_2 - 1 \quad \text{und} \quad \sigma_1(i_2) > i_2.$$

Analog erhält man τ_2 und σ_2 und schließlich ein $k \leq n$ sowie Transpositionen $\tau_1, \dots, \tau_k$ mit $\sigma_k = \tau_k \cdot \dots \cdot \tau_1 \cdot \sigma = \text{id}$. Daraus folgt

$$\sigma = (\tau_k \cdot \dots \cdot \tau_1)^{-1} = \tau_1^{-1} \cdot \dots \cdot \tau_k^{-1} = \tau_1 \cdot \dots \cdot \tau_k. \quad \square$$

Zur Vorsorge noch eine kleine technische

Bemerkung. Sei $n \geq 2$ und

$$\tau_0 := \begin{bmatrix} 1 & 2 & 3 & \cdots & n \\ 2 & 1 & 3 & \cdots & n \end{bmatrix} \in S_n$$

die Transposition, die 1 und 2 vertauscht. Dann gibt es zu jeder beliebigen Transposition $\tau \in S_n$ ein $\sigma \in S_n$ mit

$$\tau = \sigma \cdot \tau_0 \cdot \sigma^{-1}.$$

Beweis. Seien k und l die von τ vertauschten Elemente. Wir behaupten, daß jedes $\sigma \in S_n$ mit

$$\sigma(1) = k \quad \text{und} \quad \sigma(2) = l$$

die verlangte Eigenschaft hat. Sei $\tau' := \sigma \cdot \tau_0 \cdot \sigma^{-1}$. Wegen $\sigma^{-1}(k) = 1$ und $\sigma^{-1}(l) = 2$ ist

$$\tau'(k) = \sigma(\tau_0(1)) = \sigma(2) = l \quad \text{und} \quad \tau'(l) = \sigma(\tau_0(2)) = \sigma(1) = k.$$

Für $i \notin \{k, l\}$ ist $\sigma^{-1}(i) \notin \{1, 2\}$, also

$$\tau'(i) = \sigma(\tau_0(\sigma^{-1}(i))) = \sigma(\sigma^{-1}(i)) = i.$$

Daraus folgt $\tau' = \tau$. □

3.2.3. Die Zerlegung einer Permutation in Transpositionen ist nicht eindeutig. Wir müssen aber beweisen, daß die Anzahl der nötigen Transpositionen entweder immer gerade oder immer ungerade ist. Zu diesem Zweck ordnen wir jeder Permutation ein Vorzeichen zu. Elementar kann man es so beschreiben:

Ist $\sigma \in S_n$, so nennt man jedes Paar $i, j \in \{1, \dots, n\}$ mit

$$i < j, \quad \text{aber} \quad \sigma(i) > \sigma(j),$$

einen *Fehlstand* von σ . Zum Beispiel hat

$$\sigma = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}$$

insgesamt 2 Fehlstände, nämlich

$$1 < 3, \quad \text{aber} \quad 2 > 1, \quad \text{und} \quad 2 < 3, \quad \text{aber} \quad 3 > 1.$$

Wir definieren das *Signum* (d.h. „Vorzeichen“) von σ durch

$$\text{sign } \sigma := \begin{cases} +1, & \text{falls } \sigma \text{ eine gerade Anzahl von Fehlständen hat,} \\ -1, & \text{falls } \sigma \text{ eine ungerade Anzahl von Fehlständen hat.} \end{cases}$$

Man nennt $\sigma \in S_n$

gerade, falls $\text{sign } \sigma = +1$, und

ungerade, falls $\text{sign } \sigma = -1$.

Diese Definition ist recht gut geeignet, um das Signum durch systematisches Zählen zu berechnen, aber zu schwerfällig für theoretische Überlegungen. Daher

ist es hilfreich, das Zählen der Fehlstände und das Berechnen des Signums in einer Formel zusammenzufassen. In den folgenden Produkten sollen i und j stets die Menge $\{1, \dots, n\}$ durchlaufen, und zwar mit den unter dem Produktsymbol vermerkten Nebenbedingungen.

Lemma. Für jedes $\sigma \in S_n$ gilt

$$\text{sign } \sigma = \prod_{i < j} \frac{\sigma(j) - \sigma(i)}{j - i}.$$

Beweis. Man mache sich erst einmal klar, daß man das Produkt als einen langen Bruch schreiben kann, bei dem in Nenner und Zähler die gleichen Differenzen vorkommen, allerdings im Zähler im allgemeinen an anderer Stelle und – das ist der Kniff – im Fall eines Fehlstandes mit negativem Vorzeichen.

Diese Vorstellung übersetzt man in die folgende Rechnung, bei der m die Anzahl der Fehlstände bezeichnet:

$$\begin{aligned} \prod_{i < j} (\sigma(j) - \sigma(i)) &= \left(\prod_{\substack{i < j \\ \sigma(i) < \sigma(j)}} (\sigma(j) - \sigma(i)) \right) \cdot (-1)^m \prod_{\substack{i < j \\ \sigma(i) > \sigma(j)}} |\sigma(j) - \sigma(i)| \\ &= (-1)^m \prod_{i < j} |\sigma(j) - \sigma(i)| = (-1)^m \prod_{i < j} (j - i). \end{aligned}$$

Bei der letzten Gleichung wird verwendet, daß die beiden Produkte bis auf die Reihenfolge die gleichen Faktoren enthalten. Das folgt aus der Bijektivität der Abbildung σ . $\square$

Die entscheidende Eigenschaft des Signums ist, daß es mit der Hintereinanderschaltung von Permutationen verträglich ist.

Satz. Für alle $\sigma, \tau \in S_n$ gilt $\text{sign}(\tau \cdot \sigma) = (\text{sign } \tau) \cdot (\text{sign } \sigma)$.

Insbesondere gilt $\text{sign } \sigma^{-1} = \text{sign } \sigma$ für jedes $\sigma \in S_n$.

Anders ausgedrückt bedeutet das, daß die Abbildung

$$\text{sign} : S_n \rightarrow \{+1, -1\}$$

ein Homomorphismus in die Gruppe von zwei Elementen ist.

Beweis. Es ist

$$\begin{aligned} \text{sign}(\tau \cdot \sigma) &= \prod_{i < j} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{j - i} \\ &= \prod_{i < j} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j) - \sigma(i)} \cdot \prod_{i < j} \frac{\sigma(j) - \sigma(i)}{j - i}. \end{aligned}$$

Da das zweite Produkt gleich $\text{sign } \sigma$ ist, genügt es zu zeigen, daß das erste Produkt gleich $\text{sign } \tau$ ist.

$$\begin{aligned}
 & \prod_{i < j} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j) - \sigma(i)} \\
 = & \prod_{\substack{i < j \\ \sigma(i) < \sigma(j)}} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j) - \sigma(i)} \cdot \prod_{\substack{i < j \\ \sigma(i) > \sigma(j)}} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j) - \sigma(i)} \\
 = & \prod_{\substack{i < j \\ \sigma(i) < \sigma(j)}} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j) - \sigma(i)} \cdot \prod_{\substack{i > j \\ \sigma(i) < \sigma(j)}} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j) - \sigma(i)} \\
 = & \prod_{\sigma(i) < \sigma(j)} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j) - \sigma(i)}.
 \end{aligned}$$

Da σ bijektiv ist, enthält dieses letzte Produkt bis auf die Reihenfolge die gleichen Faktoren wie

$$\prod_{i < j} \frac{\tau(j) - \tau(i)}{j - i} = \text{sign } \tau,$$

und der Satz ist bewiesen. $\square$

Damit ist die zu Beginn von 3.2.3 gestellte Frage über die Anzahl der Transpositionen beantwortet:

Korollar 1. Sei $n \geq 2$.

1) Für jede Transposition $\tau \in S_n$ gilt $\text{sign } \tau = -1$.

2) Ist $\sigma \in S_n$ und $\sigma = \tau_1 \cdot \dots \cdot \tau_k$ mit Transpositionen $\tau_1, \dots, \tau_k \in S_n$, so ist

$$\text{sign } \sigma = (-1)^k.$$

Beweis. Ist τ_0 die Transposition, die 1 und 2 vertauscht, so ist $\text{sign } \tau_0 = -1$, denn τ_0 hat genau einen Fehlstand. Nach der Bemerkung aus 3.2.2 gibt es ein $\sigma \in S_n$ mit $\tau = \sigma \cdot \tau_0 \cdot \sigma^{-1}$, also folgt aus obigem Satz

$$\text{sign } \tau = \text{sign } \sigma \cdot \text{sign } \tau_0 \cdot (\text{sign } \sigma)^{-1} = \text{sign } \tau_0 = -1.$$

2) folgt aus 1) wieder nach obigem Satz. $\square$

Als entscheidende Folgerung für Determinanten erhalten wir das

Korollar 2. Für jede Permutation $\sigma \in S_n$ ist

$$\det \begin{pmatrix} e_{\sigma(1)} \\ \vdots \\ e_{\sigma(n)} \end{pmatrix} = \text{sign } \sigma.$$

Beweis. Ist $\sigma = \tau_1 \cdot \dots \cdot \tau_k$, so kann man E_n durch k Zeilenvertauschungen in die obige Matrix überführen. $\square$

3.2.4. Die Gruppe S_n zerfällt in zwei Klassen, die der geraden und die der ungeraden Permutationen, und diese beiden Klassen sind nahezu gleichberechtigt. Zunächst zu den geraden:

$$A_n := \{\sigma \in S_n : \text{sign } \sigma = +1\} \subset S_n$$

ist nach dem Satz aus 3.2.3 eine Untergruppe, sie heißt die *alternierende Gruppe*. Für jedes $\tau \in S_n$ haben wir

$$A_n \tau := \{\sigma \cdot \tau : \sigma \in A_n\}.$$

Für $\text{sign } \tau = +1$ ist offenbar $A_n \tau = A_n$.

Bemerkung. Ist $\tau \in S_n$ mit $\text{sign } \tau = -1$ gegeben, so ist

$$S_n = A_n \cup A_n \tau \quad \text{und} \quad A_n \cap A_n \tau = \emptyset.$$

Insbesondere ist die Anzahl der Elemente von A_n gleich $\frac{1}{2}n!$.

Beweis. Sei $\sigma \in S_n$ mit $\text{sign } \sigma = -1$ gegeben. Nach 3.2.3 ist $\text{sign}(\sigma \cdot \tau^{-1}) = +1$, also ist $\sigma \in A_n \tau$, denn $\sigma = (\sigma \cdot \tau^{-1}) \cdot \tau$. Für jedes $\sigma \in A_n \tau$ ist $\text{sign } \sigma = -1$, also ist die Vereinigung auch disjunkt.

Nach 1.2.4 ist die Abbildung $A_n \rightarrow A_n \tau$, $\sigma \mapsto \sigma \cdot \tau$, bijektiv. Da S_n aus $n!$ Elementen besteht, enthalten A_n und $A_n \tau$ je $\frac{1}{2}n!$ Elemente. $\square$

3.2.5. Es gibt mehrere Möglichkeiten, die Existenz und Eindeutigkeit der mit den Axiomen von WEIERSTRASS charakterisierten Determinante zu beweisen. Wir benutzen dazu den Weg über die Formel von LEIBNIZ, die zwar wenig beliebt, aber dennoch klassisch und manchmal für die Theorie nützlich ist.

Theorem. Ist K ein Körper und $n \geq 1$, so gibt es genau eine Determinante

$$\det: M(n \times n; K) \rightarrow K,$$

und zwar ist für $A = (a_{ij}) \in M(n \times n; K)$

$$\det A = \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot a_{1\sigma(1)} \cdot \dots \cdot a_{n\sigma(n)}. \quad (*)$$

Die Formel (*) von LEIBNIZ hat für jede Permutation einen, also insgesamt $n!$ Summanden.

Beweis. Wir zeigen zunächst, daß die Formel (*) aus den Axiomen folgt, das beweist die *Eindeutigkeit*. Dazu zerlegen wir jeden Zeilenvektor a_i von A in

$$a_i = a_{i1}e_1 + \dots + a_{in}e_n$$

und wenden Zeile für Zeile von oben nach unten das Axiom D1 an, bis eine Summe mit n^n Summanden entstanden ist. Das nennt man eine *Entwicklung* von A nach Zeilen.

$$\begin{aligned} \det \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} &= \sum_{i_1=1}^n a_{1i_1} \cdot \det \begin{pmatrix} e_{i_1} \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \\ &= \sum_{i_1=1}^n a_{1i_1} \cdot \sum_{i_2=1}^n a_{2i_2} \cdot \det \begin{pmatrix} e_{i_1} \\ e_{i_2} \\ a_3 \\ \vdots \\ a_n \end{pmatrix} = \dots = \\ &= \sum_{i_1=1}^n \sum_{i_2=1}^n \dots \sum_{i_n=1}^n a_{1i_1} \cdot a_{2i_2} \cdot \dots \cdot a_{ni_n} \cdot \det \begin{pmatrix} e_{i_1} \\ \vdots \\ e_{i_n} \end{pmatrix}. \end{aligned}$$

Nun gilt nach Korollar 2 aus 3.2.3 und D2

$$\det \begin{pmatrix} e_{i_1} \\ \vdots \\ e_{i_n} \end{pmatrix} = \begin{cases} \text{sign } \sigma, & \text{falls es ein } \sigma \in S_n \text{ gibt mit } \sigma(v) = i_v \text{ für alle } v, \\ 0, & \text{sonst.} \end{cases}$$

Also bleiben „nur“ $n!$ der n^n Summanden übrig, und es gilt (*).

Leider ist der Beweis damit noch nicht zu Ende, denn allein aus der Tatsache, daß man die Axiome zur Ableitung der Formel benutzt hat, folgt noch nicht, daß die Formel alle diese Eigenschaften hat.

Für den *Existenzbeweis* definieren wir nun die Determinante durch die Formel von LEIBNIZ, und es sind die Axiome von WEIERSTRASS nachzuprüfen. Danach herrscht wieder Frieden zwischen den beiden alten Herren.

D1 sieht man direkt an der Summe

$$\begin{aligned}
 \det \begin{pmatrix} \vdots \\ a'_i + a''_i \\ \vdots \end{pmatrix} &= \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot a_{1\sigma(1)} \cdot \dots \cdot (a'_{i\sigma(i)} + a''_{i\sigma(i)}) \cdot \dots \cdot a_{n\sigma(n)} \\
 &= \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot a_{1\sigma(1)} \cdot \dots \cdot a'_{i\sigma(i)} \cdot \dots \cdot a_{n\sigma(n)} \\
 &\quad + \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot a_{1\sigma(1)} \cdot \dots \cdot a''_{i\sigma(i)} \cdot \dots \cdot a_{n\sigma(n)} \\
 &= \det \begin{pmatrix} \vdots \\ a'_i \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ a''_i \\ \vdots \end{pmatrix},
 \end{aligned}$$

also gilt D1 a). Für b) zeigt man, daß jeder Summand mit λ multipliziert wird.

D2 Angenommen, die k -te und die l -te Zeile von A seien gleich, wobei $k < l$. Ist τ die Transposition, die k und l vertauscht, so ist nach 3.2.4

$$S_n = A_n \cup A_n \tau,$$

und diese Vereinigung ist disjunkt. Ist $\sigma \in A_n$, so gilt

$$\text{sign} \sigma = +1 \quad \text{und} \quad \text{sign}(\sigma \cdot \tau) = -1.$$

Wenn σ die Gruppe A_n durchläuft, durchläuft $\sigma \cdot \tau$ die Menge $A_n \tau$. Also ist

$$\det A = \sum_{\sigma \in A_n} a_{1\sigma(1)} \cdot \dots \cdot a_{n\sigma(n)} - \sum_{\sigma \in A_n} a_{1\sigma(\tau(1))} \cdot \dots \cdot a_{n\sigma(\tau(n))}. \quad (**)$$

Da die k -te und die l -te Zeile von A gleich sind, gilt nach der Definition von τ

$$\begin{aligned}
 &a_{1\sigma(\tau(1))} \cdot \dots \cdot a_{k\sigma(\tau(k))} \cdot \dots \cdot a_{l\sigma(\tau(l))} \cdot \dots \cdot a_{n\sigma(\tau(n))} \\
 &= a_{1\sigma(1)} \cdot \dots \cdot a_{k\sigma(l)} \cdot \dots \cdot a_{l\sigma(k)} \cdot \dots \cdot a_{n\sigma(n)} \\
 &= a_{1\sigma(1)} \cdot \dots \cdot a_{k\sigma(k)} \cdot \dots \cdot a_{l\sigma(l)} \cdot \dots \cdot a_{n\sigma(n)} \\
 &= a_{1\sigma(1)} \cdot \dots \cdot a_{n\sigma(n)}.
 \end{aligned}$$

Also heben sich in (**) die Summanden gegenseitig auf, und es folgt

$$\det A = 0.$$

D3 Ist δ_{ij} das Kronecker-Symbol und $\sigma \in S_n$, so ist

$$\delta_{1\sigma(1)} \cdot \dots \cdot \delta_{n\sigma(n)} = \begin{cases} 0 & \text{für } \sigma \neq \text{id}, \\ 1 & \text{für } \sigma = \text{id}. \end{cases}$$

Also ist

$$\det E_n = \det(\delta_{ij}) = \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot \delta_{1\sigma(1)} \cdot \dots \cdot \delta_{n\sigma(n)} = \text{sign}(\text{id}) = 1. \quad \square$$

3.2.6. Durch den Existenz- und Eindeutigkeitssatz ist die theoretische Rechtfertigung dafür geliefert, daß man Determinanten so berechnen darf, wie wir es in 3.1.4 mit Hilfe von Zeilenumformungen getan hatten. Man kann aber für kleine n auch die Formel von LEIBNIZ verwenden:

Für $n = 2$ ist

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

Für $n = 3$ ist

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{aligned} & a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ & - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}. \end{aligned}$$

Die Summe hat $3! = 3 \cdot 2 \cdot 1 = 6$ Summanden. Man kann sich diese Formel leicht merken durch die *Regel von SARRUS*: Man schreibt den ersten und zweiten Spaltenvektor noch einmal hinter die Matrix:

$$\begin{array}{cccccc} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} & \\ & a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ & a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{array}$$

Bild 3.6

Die Koeffizienten längs der „Hauptdiagonale“ und ihrer Parallelen ergeben dann die Summanden mit positivem Vorzeichen, die Koeffizienten längs der „Nebendiagonalen“ und ihrer Parallelen ergeben die Summanden mit negativem Vorzeichen.

Für $n = 4$ erhält man eine Summe mit $4! = 24$ Summanden, was schon höchst unangenehm ist. Man beachte, daß ein Analogon zur Regel von SARRUS für $n \geq 4$ nicht gilt. Für $n = 4$ würde man auf diese Weise nur 8 von 24 Summanden erhalten, und die Vorzeichen würden im allgemeinen nicht stimmen (wovon man sich zur Übung ein für allemal überzeugen sollte).

Wegen des rasanten Wachstums von $n!$ sind auch Computer mit der Formel von LEIBNIZ überfordert (vgl. Aufgabe 5). Die Methode aus 3.1.4 erfordert wesentlich weniger Rechenaufwand.

Zur Rehabilitation der Formel von LEIBNIZ geben wir zwei theoretische Anwendungen. Die Einträge a_{ij} der Matrix kann man als insgesamt n^2 Unbestimmte ansehen, dann ist die Abbildung

$$\det: K^{n^2} \rightarrow K$$

ein Polynom, im Fall $K = \mathbb{R}$ oder $\mathbb{C}$ insbesondere differenzierbar und somit

stetig.

In den Axiomen der Determinante sind die Zeilen vor den Spalten ausgezeichnet. Das ist nur scheinbar so:

Satz. Für eine Matrix $A \in M(n \times n; K)$ gilt $\det {}^t A = \det A$.

Beweis. Ist $A = (a_{ij})$, so ist ${}^t A = (a'_{ij})$ mit $a'_{ij} = a_{ji}$. Nun gilt

$$\begin{aligned} \det {}^t A &= \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot a'_{1\sigma(1)} \cdot \dots \cdot a'_{n\sigma(n)} \\ &= \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot a_{\sigma(1)1} \cdot \dots \cdot a_{\sigma(n)n} \\ &= \sum_{\sigma \in S_n} \text{sign}(\sigma^{-1}) \cdot a_{1\sigma^{-1}(1)} \cdot \dots \cdot a_{n\sigma^{-1}(n)} \\ &= \det A. \end{aligned}$$

Bei der vorletzten Gleichung wurde benutzt, daß für jedes $\sigma \in S_n$

$$a_{\sigma(1)1} \cdot \dots \cdot a_{\sigma(n)n} = a_{1\sigma^{-1}(1)} \cdot \dots \cdot a_{n\sigma^{-1}(n)}$$

gilt, denn bis auf die Reihenfolge enthalten die beiden Produkte die gleichen Faktoren. Außerdem wurde verwendet, daß nach 3.2.3

$$\text{sign} \sigma = \text{sign} \sigma^{-1}$$

gilt. Für die letzte Gleichung haben wir benutzt, daß mit σ auch σ^{-1} „ganz S_n durchläuft“. Genauer gesagt bedeutet das, daß die Abbildung

$$S_n \rightarrow S_n, \quad \sigma \mapsto \sigma^{-1},$$

bijektiv ist. Dies folgt sofort aus der Eindeutigkeit des inversen Elementes (vgl. 1.2.3). $\square$

3.2.7. Wie nützlich es ist, wenn man abwechselnd mit Zeilen und Spalten operieren kann, zeigt das

Beispiel. Wir betrachten $x_1, \dots, x_n$ als Unbestimmte und definieren die **VANDERMONDE-Determinante**

$$\Delta_n := \det \begin{pmatrix} 1 & x_1 & \dots & x_1^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & x_n & \dots & x_n^{n-1} \end{pmatrix}.$$

Offensichtlich ist $\Delta_1 = 1$, $\Delta_2 = x_2 - x_1$. Für $n = 3$ formt man zuerst Spalten

um und zieht dann Faktoren aus den Zeilen:

$$\begin{aligned}
 \begin{vmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{vmatrix} &= \begin{vmatrix} 1 & x_1 & x_1^2 - x_1^2 \\ 1 & x_2 & x_2^2 - x_1 x_2 \\ 1 & x_3 & x_3^2 - x_1 x_3 \end{vmatrix} = \begin{vmatrix} 1 & x_1 - x_1 & 0 \\ 1 & x_2 - x_1 & x_2(x_2 - x_1) \\ 1 & x_3 - x_1 & x_3(x_3 - x_1) \end{vmatrix} \\
 &= \begin{vmatrix} x_2 - x_1 & x_2(x_2 - x_1) \\ x_3 - x_1 & x_3(x_3 - x_1) \end{vmatrix} = (x_2 - x_1)(x_3 - x_1) \begin{vmatrix} 1 & x_2 \\ 1 & x_3 \end{vmatrix} \\
 &= (x_2 - x_1)(x_3 - x_1)(x_3 - x_2).
 \end{aligned}$$

Nach diesem Muster erhält man durch Induktion über n (Aufgabe 2)

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

Daraus folgt, daß die Zeilen oder Spalten der obigen Matrix genau dann linear abhängig sind, wenn $x_i = x_j$ für mindestens ein Paar $i \neq j$.

3.2.8. Wie wir gesehen haben, ist es oft nützlich, in eine Matrix nicht nur Elemente eines Körpers, sondern allgemeinere Symbole – etwa Unbestimmte – einzutragen und damit rechnen zu dürfen. Damit das kein Ritt über den Bodensee bleibt, wird etwas theoretische Rechtfertigung dafür bereitgestellt.

Wir gehen aus von einem kommutativen Ring R mit Einselement 1 (s. 1.3.1). Mit

$$M(n \times n; R)$$

bezeichnen wir die quadratischen n -reihigen Matrizen mit Einträgen aus R . Addition und Multiplikation von Matrizen kann man wie bei einem Körper erklären (1.4.1 und 2.5.1), damit wird $M(n \times n; R)$ zu einem Ring mit Einselement E_n . Da im Beweis des Existenz- und Eindeutigkeitstheorems aus 3.2.5 nirgendwo durch Einträge der Matrix dividiert wird, kann man ihn wörtlich auf R übertragen. Man überzeuge sich, daß die Kommutativität verwendet wird! Daraus folgt, daß die *Determinante*

$$\det: M(n \times n; R) \rightarrow R, \quad A = (a_{ij}) \mapsto \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot a_{1\sigma(1)} \cdot \dots \cdot a_{n\sigma(n)}$$

die Eigenschaften D1, D2 und D3 aus 3.1.2 hat.

Man beachte, daß von allen Folgerungen D4 bis D11 nur diejenigen direkt übertragen werden können, bei deren Beweis nicht dividiert wurde (Aufgabe 7).

Aufgaben zu 3.2

1. Stellen Sie die Permutation

$$\sigma = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 2 & 1 \end{bmatrix}$$

als Produkt von Transpositionen dar.

2. Beweisen Sie mit Induktion nach n , dass für die Vandermonde-Determinante gilt:

$$\det \begin{pmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & x_n & \cdots & x_n^{n-1} \end{pmatrix} = \prod_{1 \leq i < j \leq n} (x_j - x_i).$$

3. Geben Sie eine unendliche Teilmenge des $\mathbb{R}^n$ an, in der jeweils n verschiedene Punkte linear unabhängig sind.

4. Zeigen Sie noch einmal

$$\det(a_{ij}) = \det((-1)^{i+j} \cdot a_{ij}),$$

(vgl. Aufgabe 4 zu 3.1), aber benutzen Sie nun zum Beweis die Formel von LEIBNIZ.

5. In dieser Aufgabe soll der Aufwand zum Berechnen der Determinante mit Hilfe der Leibniz-Formel bzw. des Gauß-Algorithmus verglichen werden.

- a) Bestimmen Sie die Anzahl der Additionen und Multiplikationen, die nötig sind, wenn man die Determinante von $A = (a_{ij}) \in M(n \times n; \mathbb{R})$
- mit der Leibniz-Formel,
 - durch Umformung der Matrix in Zeilenstufenform mit dem Gauß-Algorithmus und Aufmultiplizieren der Diagonalelemente berechnet.
- b) Es stehe ein Computer zur Verfügung, der Addition und Multiplikation in 0.2 Mikrosekunden durchführen kann. Schätzen Sie ab, für welche Größe von Matrizen man mit den Verfahren i) bzw. ii) in einer vorgegebenen Rechenzeit von höchstens 48 Stunden auf diesem Computer Determinanten berechnen kann.

6. Beweisen Sie die Regeln D4 bis D11 aus 3.1.3 mit Hilfe der Leibniz-Formel.

7. Welche der Eigenschaften D4 bis D11 gelten, falls man Determinanten von Matrizen aus $M(n \times n; R)$ für einen Ring R betrachtet (vgl. 3.2.8)?

8. (Fortsetzung von Aufgabe 5 zu 3.1.)

Sei K ein Körper mit $\text{char } K \neq 2$, $n \in \mathbb{N} \setminus \{0\}$ gerade, also $n = 2m$ für ein $m \in \mathbb{N}$ und $A \in M(n \times n; K)$ schiefsymmetrisch. Definiert man

$$P(x_{11}, \dots, x_{nn}) = \sum \text{sign}(\sigma) \cdot x_{\sigma(1)\sigma(2)} \cdot \dots \cdot x_{\sigma(2m-1)\sigma(2m)},$$

wobei über alle $\sigma \in S_n$ mit $\sigma(2i) > \sigma(2i-1)$ für $i = 1, \dots, m$ summiert wird, so gilt $\det A = (\frac{1}{m!} P(a_{11}, \dots, a_{nn}))^2$. Man nennt P ein *Pfaffsches Polynom*.

9. Seien v, w zwei verschiedene Punkte des K^2 und $L \subset K^2$ die Gerade durch v und w . Dann gilt:

$$L = \{(x_1, x_2) \in K^2 : \det \begin{pmatrix} 1 & v_1 & v_2 \\ 1 & w_1 & w_2 \\ 1 & x_1 & x_2 \end{pmatrix} = 0\}.$$

10.* Zeigen Sie, dass die Menge

$$\mathrm{SL}(2; \mathbb{Z}) := \{A \in \mathrm{M}(2 \times 2; \mathbb{Z}) : \det A = 1\}$$

eine Gruppe bzgl. der Multiplikation ist und erzeugt wird von den Matrizen

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

d.h. $\mathrm{SL}(2; \mathbb{Z}) = \mathrm{erz}(A, B)$ (vgl. Aufgabe 4 zu 1.2).

11. Gegeben sei ein offenes Intervall $I \subset \mathbb{R}$ und die $\mathbb{R}$ -Vektorräume

$$\mathcal{C} := \mathcal{C}(I; \mathbb{R}) = \{\alpha : I \rightarrow \mathbb{R} : \alpha \text{ stetig}\},$$

$$\mathcal{D} := \mathcal{D}(I; \mathbb{R}^n) = \{\varphi = {}^t(\varphi_1, \dots, \varphi_n) : I \rightarrow \mathbb{R}^n : \\ \varphi_i \text{ beliebig oft differenzierbar}\}.$$

Matrizen $A \in \mathrm{M}(n \times n; \mathcal{C})$ und $b \in \mathrm{M}(n \times 1; \mathcal{C})$ bestimmen das lineare Differentialgleichungssystem

$$y' = A \cdot y + b. \quad (*)$$

Für $b = 0$ heißt das System homogen. Die Lösungsräume sind erklärt durch

$$\mathcal{L} := \{\varphi \in \mathcal{D} : \varphi' = A \cdot \varphi + b\} \quad \text{und} \quad \mathcal{L}_0 := \{\varphi \in \mathcal{D} : \varphi' = A \cdot \varphi\}.$$

a) Zeigen Sie, dass $\mathcal{L}_0 \subset \mathcal{D}$ ein Untervektorraum und $\mathcal{L} \subset \mathcal{D}$ ein affiner Unterraum ist.

b) Zeigen Sie, dass für $\varphi^{(1)}, \dots, \varphi^{(n)} \in \mathcal{L}_0$ folgende Bedingungen äquivalent sind:

- $\varphi^{(1)}, \dots, \varphi^{(n)}$ sind über $\mathbb{R}$ linear unabhängig.
- Für ein $x_0 \in I$ sind $\varphi^{(1)}(x_0), \dots, \varphi^{(n)}(x_0) \in \mathbb{R}^n$ linear unabhängig.
- $\det(\varphi_i^{(j)}) \neq 0$. Diese Determinante heißt *WRONSKI-Determinante*.

c) Zeigen Sie, dass $\dim \mathcal{L} = n$ (unabhängig von A).

Hinweis: Man benutze die in der Analysis bewiesene Existenz- und Eindeutigkeitsaussage ([Fo 2], §12), wonach es bei gegebenem x_0 zu beliebigem Anfangswert $c \in \mathbb{R}^n$ genau eine Lösung φ von $(*)$ mit $\varphi(x_0) = c$ gibt.

12. Bestimmen Sie alle Lösungen der Differentialgleichung $y'' = -y$. Überführen Sie dazu die Differentialgleichung mit dem Ansatz $y_0 = y$, $y_1 = y'$ in ein lineares Differentialgleichungssystem wie in Aufgabe 11, und benutzen Sie, dass φ genau dann eine Lösung von $y'' = -y$ ist, wenn (φ, φ') eine Lösung des linearen Systems ist.

3.3 Minoren*

Wie wir gesehen haben, wird die Berechnung von Determinanten bei wachsendem n sehr viel schwieriger. Daher kann es manchmal helfen, in einer Matrix Zeilen und Spalten zu streichen, und zunächst die Determinante der kleineren Matrix zu berechnen.

3.3.1. Zunächst benötigen wir einige sehr technische Vorbereitungen. Ist $A = (a_{ij}) \in M(n \times n; K)$, so sei A_{ij} für festes i, j die Matrix, die aus A entsteht, indem man a_{ij} durch 1 und alle anderen Komponenten, die in der i -ten Zeile oder der j -ten Spalte stehen, durch 0 ersetzt. Ausgeschrieben ist

$$A_{ij} = \begin{pmatrix} a_{11} & \cdots & a_{1,j-1} & 0 & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & 0 & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & 0 & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & 0 & a_{n,j+1} & \cdots & a_{nn} \end{pmatrix}.$$

Die Matrix

$$A^\sharp = (a_{ij}^\sharp) \in M(n \times n; K) \quad \text{mit} \quad a_{ij}^\sharp := \det A_{ji}$$

heißt die zu A *komplementäre Matrix*. Man beachte dabei die Umkehrung der Reihenfolge der Indizes. Weiter bezeichnen wir mit

$$A'_{ij} := \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{pmatrix} \in M((n-1) \times (n-1); K)$$

die Matrix, die man durch *Streichen* der i -ten Zeile und der j -ten Spalte aus A erhält.

Bemerkung 1. Es gilt stets

$$\det A_{ij} = (-1)^{i+j} \det A'_{ij}.$$

Beweis. Durch $i - 1$ Vertauschungen benachbarter Zeilen und $j - 1$ Vertauschungen benachbarter Spalten kann man A_{ij} auf die Form

$$\begin{pmatrix} 1 & 0 \\ 0 & A'_{ij} \end{pmatrix}$$

bringen. Also folgt die Behauptung aus D6 und D9 wegen

$$(-1)^{(i-1)+(j-1)} = (-1)^{i+j}.$$

□

Ist $A = (a^1, \dots, a^n) \in M(n \times n; K)$, wobei $a^1, \dots, a^n$ die Spaltenvektoren von A sind, und ist

$$e^i = {}^t e_i = {}^t(0, \dots, 0, 1, 0, \dots, 0),$$

so ist

$$(a^1, \dots, a^{j-1}, e^i, a^{j+1}, \dots, a^n)$$

die Matrix, die aus A entsteht, indem man a_{ij} durch 1 und alle anderen Komponenten der j -ten Spalte durch 0 ersetzt. Im Gegensatz zu A_{ij} bleiben die restlichen Komponenten der i -ten Zeile unverändert.

Bemerkung 2. Es gilt stets

$$\det A_{ij} = \det(a^1, \dots, a^{j-1}, e^i, a^{j+1}, \dots, a^n).$$

Beweis. Durch Addition von Vielfachen der j -ten Spalte zu den anderen Spalten kann man $(a^1, \dots, a^{j-1}, e^i, a^{j+1}, \dots, a^n)$ in A_{ij} überführen. Also folgt die Behauptung aus D7. □

Satz. Ist $A \in M(n \times n; K)$ und $A^\sharp$ die zu A komplementäre Matrix, so ist

$$A^\sharp \cdot A = A \cdot A^\sharp = (\det A) \cdot E_n.$$

Beweis. Wir berechnen die Komponenten von $A^\sharp \cdot A$:

$$\begin{aligned} \sum_{j=1}^n a_{ij}^\sharp a_{jk} &= \sum_{j=1}^n a_{jk} \det A_{ji} \\ &= \sum_{j=1}^n a_{jk} \det(a^1, \dots, a^{i-1}, e^j, a^{i+1}, \dots, a^n) \quad \text{nach Bem. 2} \\ &= \det(a^1, \dots, a^{i-1}, \sum_{j=1}^n a_{jk} e^j, a^{i+1}, \dots, a^n) \quad \text{nach D1} \\ &= \det(a^1, \dots, a^{i-1}, a^k, a^{i+1}, \dots, a^n) \\ &= \delta_{ik} \cdot \det A. \quad \text{nach D2} \end{aligned}$$

Also ist $A^\sharp \cdot A = (\det A) \cdot E_n$. Analog berechnet man $A \cdot A^\sharp$. □

3.3.2. In den Definitionen und Beweisen von 3.3.1 kann man den Körper K durch einen kommutativen Ring R mit Einselement ersetzen (vgl. 3.2.8). Der Leser möge zur Übung die Einzelheiten überprüfen. Kritischer Punkt ist der Beweis von Bemerkung 2; aber da man die 1 als Pivot für die Umformungen hat, geht alles gut. Ergebnis ist das

Korollar. Ist R ein kommutativer Ring mit 1 und $A^\sharp \in M(n \times n; R)$ die zu $A \in M(n \times n; R)$ komplementäre Matrix, so gilt

$$A^\sharp \cdot A = A \cdot A^\sharp = (\det A) \cdot E_n. \quad \square$$

Dies wird in 4.5.3 sehr nützlich sein.

3.3.3. Die gerade bewiesene Eigenschaft der komplementären Matrix hat wichtige Konsequenzen. Wir beschränken uns wieder auf den Fall eines Körpers K .

Entwicklungssatz von LAPLACE. Ist $n \geq 2$ und $A \in M(n \times n; K)$, so gilt für jedes $i \in \{1, \dots, n\}$

$$\det A = \sum_{j=1}^n (-1)^{i+j} \cdot a_{ij} \cdot \det A'_{ij}$$

(Entwicklung nach der i -ten Zeile) und für jedes $j \in \{1, \dots, n\}$

$$\det A = \sum_{i=1}^n (-1)^{i+j} \cdot a_{ij} \cdot \det A'_{ij}$$

(Entwicklung nach der j -ten Spalte). Dabei bezeichnet A'_{ij} jeweils die in 3.3.1 definierte Streichungsmatrix.

Beweis. Nach Satz 3.3.1 ist $\det A$ für jedes i gleich der i -ten Komponente in der Diagonale der Matrix $A \cdot A^\sharp$, also

$$\det A = \sum_{j=1}^n a_{ij} a_{ji}^\sharp = \sum_{j=1}^n a_{ij} \cdot \det A_{ij} = \sum_{j=1}^n (-1)^{i+j} a_{ij} \cdot \det A'_{ij}$$

nach Bemerkung 1 aus 3.3.1. Indem man ebenso mit $A^\sharp \cdot A$ verfährt, erhält man die Formel für die Entwicklung nach der j -ten Spalte. $\square$

Genau genommen gibt der Entwicklungssatz von Laplace nur ein Verfahren an, die Summanden der Formel von LEIBNIZ in einer speziellen Reihenfolge aufzuschreiben. Das kann aber doch nützlich sein, etwa dann, wenn in einer Zeile oder Spalte viele Nullen stehen. Als Beispiel berechnen wir noch einmal

$$\begin{vmatrix} 0 & 1 & 2 \\ 3 & 2 & 1 \\ 1 & 1 & 0 \end{vmatrix} = 0 \cdot \begin{vmatrix} 2 & 1 \\ 1 & 0 \end{vmatrix} - 1 \cdot \begin{vmatrix} 3 & 1 \\ 1 & 0 \end{vmatrix} + 2 \cdot \begin{vmatrix} 3 & 2 \\ 1 & 1 \end{vmatrix} = 0 + 1 + 2 = 3.$$

Die durch den Faktor $(-1)^{i+j}$ bewirkte Vorzeichenverteilung kann man sich als „Schachbrettmuster“ vorstellen:

+	-	+	-	+	-	+	-
-	+	-	+	-	+	-	+
+	-	+	-	+	-	+	-
-	+	-	+	-	+	-	+
+	-	+	-	+	-	+	-
-	+	-	+	-	+	-	+
+	-	+	-	+	-	+	-
-	+	-	+	-	+	-	+

3.3.4. Aus Satz 3.3.1 sieht man sofort, daß die komplementäre Matrix bis auf den Faktor $\det A$ gleich der *inversen Matrix* ist. Das kann man nach Bemerkung 1 in 3.3.1 auch so ausdrücken:

Satz. Sei $A \in GL(n; K)$. Definiert man $C = (c_{ij}) \in M(n \times n; K)$ durch

$$c_{ij} := (-1)^{i+j} \cdot \det A'_{ij},$$

so ist

$$A^{-1} = \frac{1}{\det A} \cdot {}^t C. \quad \square$$

Im Spezialfall $n = 2$ erhält man

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc} \cdot {}^t \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} = \frac{1}{ad-bc} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Auch für $n = 3, 4$ ist dieses Verfahren noch nützlich. Man berechne zur Übung

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

und vergleiche das Ergebnis mit dem in 2.7.5 erhaltenen.

Für größere n ist der Rechenaufwand wieder zu groß, weil man n^2 Determinanten berechnen muß. Dafür eine schöne Anwendung in der Theorie. Wegen der Stetigkeit von

$$\det: M(n \times n; \mathbb{R}) \rightarrow \mathbb{R}$$

ist $GL(n; \mathbb{R}) \subset M(n \times n; \mathbb{R})$ eine offene Teilmenge, und aus den obigen Formeln folgt, daß die Abbildung

$$GL(n; \mathbb{R}) \rightarrow GL(n; \mathbb{R}), \quad A \mapsto A^{-1},$$

differenzierbar ist. Das ist nützlich in der Analysis.

3.3.5. Auch für die Lösung linearer Gleichungssysteme ist die komplementäre Matrix nützlich, und zwar im Fall

$$A \cdot x = b \quad \text{mit} \quad A \in GL(n; K).$$

Hier ist die entsprechend 2.3.6 eindeutige Lösung gegeben durch

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

Man kann x auch ohne explizite Berechnung der inversen Matrix bestimmen: Sind $a^1, \dots, a^n$ die Spaltenvektoren von A , so hat A^{-1} nach 3.3.1 in der i -ten Zeile und der j -ten Spalte die Komponente

$$\frac{\det A_{ji}}{\det A} = \frac{\det(a^1, \dots, a^{i-1}, e^j, a^{i+1}, \dots, a^n)}{\det A}.$$

Für die i -te Komponente von $x = A^{-1}b$ folgt nach D1 und Satz 3.2.6

$$x_i = \sum_{j=1}^n b_j \frac{\det A_{ji}}{\det A} = \frac{\det(a^1, \dots, a^{i-1}, b, a^{i+1}, \dots, a^n)}{\det A}.$$

Man kann also x_i berechnen aus der Determinante von A und der Determinante der Matrix, die aus A durch Austausch der i -ten Spalte gegen b entsteht. Wir fassen das Ergebnis noch einmal zusammen.

CRAMERSche Regel. Sei $A \in GL(n; K)$, $b \in K^n$ und $x = (x_1, \dots, x_n) \in K^n$ die eindeutig bestimmte Lösung des Gleichungssystems

$$A \cdot x = b.$$

Bezeichnen $a^1, \dots, a^n$ die Spaltenvektoren der Matrix A , so gilt für jedes $i \in \{1, \dots, n\}$

$$x_i = \frac{\det(a^1, \dots, a^{i-1}, b, a^{i+1}, \dots, a^n)}{\det A}. \quad \square$$

Für große n ist diese Regel zur Bestimmung der Lösung nicht praktisch, denn man muß dazu $n+1$ Determinanten berechnen. Für theoretische Untersuchungen ist die Cramersche Regel jedoch sehr wertvoll. Im Fall $K = \mathbb{R}$ kann man damit zum Beispiel leicht einsehen, daß die Lösung x des Gleichungssystems $A \cdot x = b$ stetig von den Koeffizienten von A und b abhängt.

Als Beispiel betrachten wir das Gleichungssystem

$$\begin{aligned}x_1 + x_2 &= 1, \\x_2 + x_3 &= 1, \\3x_1 + 2x_2 + x_3 &= 0,\end{aligned}$$

mit der Koeffizientenmatrix

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 3 & 2 & 1 \end{pmatrix}.$$

Durch elementare Umformungen erhält man das reduzierte Gleichungssystem

$$\begin{aligned}x_1 + x_2 &= 1, \\x_2 + x_3 &= 1, \\2x_3 &= -2,\end{aligned}$$

und daraus die Lösung

$$(x_1, x_2, x_3) = (-1, 2, -1).$$

Wegen $\det A = 2$ ergibt die Cramersche Regel

$$\begin{aligned}x_1 &= \frac{1}{2} \begin{vmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 2 & 1 \end{vmatrix} = -1, & x_2 &= \frac{1}{2} \begin{vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 3 & 0 & 1 \end{vmatrix} = 2, \\x_3 &= \frac{1}{2} \begin{vmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 3 & 2 & 0 \end{vmatrix} = -1.\end{aligned}$$

3.3.6. Nach D10 ist $\text{rang } A < n$ für $A \in M(n \times n; K)$ gleichbedeutend mit $\det A = 0$. Um zu sehen, wie weit der Rang absinkt, muß man weitere Determinanten berechnen. Das kann man sogar auf beliebige Matrizen ausdehnen.

Ist $A \in M(m \times n; K)$ und $k \leq \min\{m, n\}$, so heißt eine quadratische Matrix $A' \in M(k \times k; K)$ eine *k-reihige Teilmatrix* von A , wenn A durch Zeilen- und Spaltenvertauschungen auf die Form

$$\begin{pmatrix} A' & * \\ * & * \end{pmatrix}$$

gebracht werden kann (wobei an den mit * bezeichneten Stellen beliebige Matrizen stehen können). Das ist insbesondere dann der Fall, wenn A' aus A durch Streichen von $m - k$ Zeilen und $n - k$ Spalten entstanden ist. $\det A'$ heißt ein *k-reihiger Minor*.

Satz. Sei $A \in M(m \times n; K)$ und $r \in \mathbb{N}$ mit $0 < r \leq \min \{m, n\}$. Dann sind folgende Bedingungen gleichwertig:

- i) $r = \text{rang } A$.
- ii) Es gibt einen r -reihigen Minor $\neq 0$, und für $k > r$ ist jeder k -reihige Minor gleich Null.

Beweis. Es genügt zu zeigen, daß für jedes $k \in \mathbb{N}$ mit $0 < k \leq \min \{m, n\}$ folgende Bedingungen gleichwertig sind:

- a) $\text{rang } A \geq k$.
- b) Es gibt eine k -reihige Teilmatrix A' von A mit $\det A' \neq 0$.

$b) \Rightarrow a)$. Aus $\det A' \neq 0$ folgt $\text{rang } A' = k$ und somit $\text{rang } A \geq k$, weil sich der Rang einer Matrix bei Zeilen- und Spaltenvertauschungen nicht ändert.

$a) \Rightarrow b)$. Ist $\text{rang } A \geq k$, so gibt es k linear unabhängige Zeilenvektoren in A . Nach Zeilenvertauschungen können wir sie in die ersten k Zeilen bringen. Sei B die Matrix, die aus diesen Zeilen besteht. Wegen

$$\text{Zeilenrang } B = k = \text{Spaltenrang } B$$

gibt es k linear unabhängige Spaltenvektoren in B . Durch Spaltenvertauschungen können wir sie in die ersten k Spalten bringen. Sei $A' \in M(k \times k; K)$ die aus diesen Spalten bestehende Matrix. A' ist eine Teilmatrix von A , und wegen $\text{rang } A' = k$ gilt $\det A' \neq 0$, was zu zeigen war. $\square$

3.3.7. Die in 3.1.3 als Eigenschaft D11 bewiesene Multiplikativität der Determinante hat eine Verallgemeinerung, die zum Beispiel in der Analysis bei der Berechnung von Inhalten benutzt wird (vgl. [Fo3], §14). Sie betrifft rechteckige Matrizen. Dafür kann man zwar i.a. keine Determinante mehr erklären, aber Minoren. Sei $A = (a^1, \dots, a^n) \in M(m \times n; K)$, wobei die $a^j \in K^m$ die Spaltenvektoren bezeichnen. Ist $m \leq n$, so definieren wir für $1 \leq k_1 < \dots < k_m \leq n$ die Teilmatrix

$$A^{k_1, \dots, k_m} := (a^{k_1}, \dots, a^{k_m}) \in M(m \times m; K).$$

$\det(A^{k_1, \dots, k_m})$ heißt ein m -reihiger Minor von A , davon gibt es

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

Stück (vgl. [Fo1], §1). Hat B die gleiche Größe wie A , so ist $A \cdot {}^t B$ quadratisch, also gibt es davon eine Determinante. Wie man sie berechnen kann, sagt das

Determinanten-Multiplikationstheorem. Ist $m \leq n$, so gilt für alle Matrizen $A, B \in M(m \times n; K)$

$$\det(A \cdot {}^t B) = \sum_{1 \leq k_1 < \dots < k_m \leq n} (\det A^{k_1, \dots, k_m}) \cdot (\det B^{k_1, \dots, k_m}). \quad (*)$$

Der Fall $m > n$ ist langweilig (vgl. Aufgabe 2), für $m = 1$ ist die Aussage offensichtlich.

Beweis. Für sehr kleine m und n kann man die Formel $(*)$ durch direkte Rechnung mit der Leibniz-Formel beweisen (Aufgabe 3), aber im allgemeinen Fall ist das eine einzige Index-Schlacht (vgl. etwa [Kow 1], §34). Übersichtlicher ist eine geeignete Zerlegung der Rechnung in elementare Schritte. Bei quadratischem A ging das mit Hilfe einer Produktdarstellung durch Elementarmatrizen (D11 in 3.1.3). Im rechteckigen Fall wird A bei festem, aber beliebigem B zeilenweise aus besonders einfachen Matrizen aufgebaut, das kann man als Umkehrung der Entwicklung nach Zeilen aus 3.2.5 ansehen.

1) Wir zeigen die Gültigkeit von $(*)$ für

$$A = \begin{pmatrix} e_{j_1} \\ \vdots \\ e_{j_m} \end{pmatrix} \quad \text{für beliebige } j_1, \dots, j_m \in \{1, \dots, n\}.$$

Man beachte, daß die Zeilenvektoren e_j im K^n liegen. Für dieses spezielle A sind die Minoren ganz einfach zu berechnen: Ist $1 \leq k_1 < \dots < k_m \leq n$, so gilt

$$\det A^{k_1, \dots, k_m} = \begin{cases} \text{sign } \sigma, & \text{falls } j_i = k_{\sigma(i)} \text{ für ein } \sigma \in S_m, \\ 0, & \text{sonst.} \end{cases}$$

Nun betrachten wir die Wirkung der Multiplikation mit A : Ist

$$B = (b^1, \dots, b^n), \quad \text{so folgt} \quad A \cdot {}^t B = \begin{pmatrix} b^{j_1} \\ \vdots \\ b^{j_m} \end{pmatrix}.$$

Die Determinante von $A \cdot {}^t B$ ist also höchstens dann von Null verschieden, wenn $j_1, \dots, j_m$ paarweise verschieden sind, d.h. wenn es eine Permutation $\sigma \in S_m$ und $1 \leq k_1 < \dots < k_m \leq n$ gibt mit $j_i = k_{\sigma(i)}$, und es ist dann

$$\det(A \cdot {}^t B) = \text{sign } \sigma \cdot \det B^{k_1, \dots, k_m}.$$

Also gilt in diesem Fall $(*)$ mit einem einzigen Summanden auf der rechten Seite.

2) Gilt die Formel für A , und entsteht $\tilde{A}$ aus A durch Multiplikation der i -ten Zeile mit $\lambda \in K$, so gilt die Formel auch für $\tilde{A}$. In Zeichen

$$(*) \Rightarrow (\tilde{*}).$$

Ist $C := A \cdot {}^t B$ und $\tilde{C} := \tilde{A} \cdot {}^t B$, so entsteht $\tilde{C}$ aus C durch Multiplikation der i -ten Zeile mit λ . Nach D1 ist

$$\det \tilde{A}^{k_1, \dots, k_m} = \lambda \cdot \det A^{k_1, \dots, k_m} \quad \text{und} \quad \det \tilde{C} = \lambda \cdot \det C,$$

also folgt $(\tilde{*}) = \lambda \cdot (*)$.

3) Ist die i -te Zeile a_i von A eine Summe $a_i = \tilde{a}_i + \tilde{\tilde{a}}_i$, und gilt die Formel für

$$\tilde{A} = \begin{pmatrix} \vdots \\ \tilde{a}_i \\ \vdots \end{pmatrix} \quad \text{und} \quad \tilde{\tilde{A}} = \begin{pmatrix} \vdots \\ \tilde{\tilde{a}}_i \\ \vdots \end{pmatrix},$$

so gilt sie auch für A . In Zeichen

$$(\tilde{*}) \text{ und } (\tilde{\tilde{*}}) \Rightarrow (*).$$

Wir setzen $\tilde{C} := \tilde{A} \cdot {}^t B$, $\tilde{\tilde{C}} := \tilde{\tilde{A}} \cdot {}^t B$ und $C = A \cdot {}^t B$. Axiom D1 angewandt auf $A^{k_1, \dots, k_m}$ und C ergibt

$$\det A^{k_1, \dots, k_m} = \det \tilde{A}^{k_1, \dots, k_m} + \det \tilde{\tilde{A}}^{k_1, \dots, k_m} \quad \text{und} \quad \det C = \det \tilde{C} + \det \tilde{\tilde{C}},$$

also folgt $(*) = (\tilde{*}) + (\tilde{\tilde{*}})$.

Es bleibt zu zeigen, daß man mit Hilfe von 1), 2) und 3) jede beliebige Matrix

$$A = (a_{ij}) = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \quad \text{mit Zeilenvektoren } a_i \in K^n$$

schrittweise aufbauen kann: In der ersten Etappe wählt man die Indizes $j_2, \dots, j_m \in \{1, \dots, n\}$ beliebig, aber fest. Läßt man j_1 von 1 bis n laufen, und wählt man nacheinander $\lambda_{j_1} = a_{1j_1}$, so folgt die Gültigkeit von $(*)$ für

$$A_1 = \begin{pmatrix} a_1 \\ e_{j_2} \\ \vdots \\ e_{j_m} \end{pmatrix} \quad \text{für beliebige } j_2, \dots, j_m.$$

In der zweiten Etappe hält man a_1 sowie $j_3, \dots, j_m$ fest und läßt j_2 von 1 bis n

laufen. Das ergibt wie oben die Gültigkeit von (*) für

$$A_2 = \begin{pmatrix} a_1 \\ a_2 \\ e_{j_3} \\ \vdots \\ e_{j_m} \end{pmatrix} \quad \text{für beliebige } j_3, \dots, j_m.$$

In der m -ten Etappe erhält man schließlich (*) für $A_m = A$. $\square$

3.3.8. Im Spezialfall $m = n$ erhält man mit der oben angewandten Methode einen neuen Beweis des Determinantenmultiplikationssatzes D11. Für $B = A$ ergibt sich das

Korollar. Für jedes $A \in M(m \times n; K)$ mit $m \leq n$ gilt

$$\det(A \cdot {}^t A) = \sum_{1 \leq k_1 < \dots < k_m \leq n} \det(A^{k_1, \dots, k_m})^2. \quad \square$$

Man nennt $\det(A \cdot {}^t A)$ eine GRAMsche *Determinante*. Insbesondere ist sie für $K = \mathbb{R}$ nie negativ und genau dann Null, wenn $\text{rang } A < m$ (vgl. 3.3.6).

Aufgaben zu 3.3

1. In dieser Aufgabe geht es um weitere Eigenschaften der komplementären Matrix.

a) Ist die Abbildung $M(n \times n; K) \rightarrow M(n \times n; K)$, $A \mapsto A^\sharp$ linear?

b) Zeigen Sie: ${}^t(A^\sharp) = ({}^t A)^\sharp$, $(AB)^\sharp = B^\sharp A^\sharp$.

c) $\det A^\sharp = (\det A)^{n-1}$.

d) $(A^\sharp)^\sharp = (\det A)^{n-2} \cdot A$.

2. Sind $A, B \in M(m \times n; K)$ und ist $m > n$, so folgt $\det A \cdot {}^t B = 0$.

3. Beweisen Sie die Formel für $\det A \cdot {}^t B$ aus 3.3.7 durch direktes Ausrechnen, wenn $A, B \in M(2 \times 3; K)$ sind.

4. Beweisen Sie:

$$\det \begin{pmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{pmatrix} = (a^2 + b^2 + c^2 + d^2)^2.$$

5. Für $x = (x_1, \dots, x_n)$ und $y = (y_1, \dots, y_n)$ aus K^n sind äquivalent:

- i) x und y sind linear abhängig.
 ii) $\det \begin{pmatrix} x_i & y_i \\ x_j & y_j \end{pmatrix} = 0$ für alle i, j .

6. Ist $E = \text{span}(x, y) \subset K^n$ ein 2-dimensionaler Untervektorraum, so definieren wir

$$p_{ij} = \det \begin{pmatrix} x_i & y_i \\ x_j & y_j \end{pmatrix} \text{ für } 1 \leq i < j \leq n.$$

Man nennt $p(x, y) = (p_{ij})_{1 \leq i < j \leq n} \in K^{\binom{n}{2}}$ die (homogenen) *Plückerkoordinaten* von $E = \text{span}(x, y)$; nach Aufgabe 5 ist $p(x, y) \neq 0$.

- a) Zeigen Sie, dass die Plückerkoordinaten bis auf einen Faktor aus $K \setminus \{0\}$ nur von E abhängen: Ist $E = \text{span}(x, y) = \text{span}(x', y')$, so existiert ein $\lambda \in K \setminus \{0\}$ mit $p(x, y) = \lambda \cdot p(x', y')$. In diesem Sinne wollen wir auch einfach von den Plückerkoordinaten $p(E)$ von E reden, diese sind dann bis auf einen Faktor $\neq 0$ eindeutig bestimmt.
 b) Zeigen Sie: Sind $E_1, E_2 \subset K^n$ Untervektorräume der Dimension 2, so dass $p(E_1)$ und $p(E_2)$ linear abhängig sind, so folgt $E_1 = E_2$.
 c) Ist $E = \text{span}(x, y) \subset K^4$, so erfüllen die Plückerkoordinaten (p_{ij}) von E die Gleichung $p_{12}p_{34} - p_{13}p_{24} + p_{14}p_{23} = 0$. Ist umgekehrt $p = (p_{ij})_{1 \leq i < j \leq 4} \in K^6 \setminus 0$ gegeben mit $p_{12}p_{34} - p_{13}p_{24} + p_{14}p_{23} = 0$, so existiert ein 2-dimensionaler Untervektorraum $E = \text{span}(x, y) \subset K^4$ mit $p(E) = p$.
 d) Sind $E_1 = \text{span}(x, y), E_2 = \text{span}(x', y') \subset K^4$ zweidimensionale Untervektorräume mit Plückerkoordinaten $p(E_1) = (p_{ij}), p(E_2) = (q_{ij})$, so gilt:

$$E_1 \cap E_2 \neq \{0\} \Leftrightarrow \det \begin{pmatrix} x_1 & y_1 & x'_1 & y'_1 \\ x_2 & y_2 & x'_2 & y'_2 \\ x_3 & y_3 & x'_3 & y'_3 \\ x_4 & y_4 & x'_4 & y'_4 \end{pmatrix} = 0$$

$$\Leftrightarrow p_{12}q_{34} - p_{13}q_{24} + p_{14}q_{23} + p_{23}q_{14} - p_{24}q_{13} + p_{34}q_{12} = 0.$$

7. Zeigen Sie, dass $\det(x) = \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot x_{1\sigma(1)} \cdot \dots \cdot x_{n\sigma(n)} \in K[x_{11}, \dots, x_{nn}]$ ein irreduzibles Polynom ist, das heißt, dass aus $\det(x) = P \cdot Q$ mit Polynomen P und Q stets $P \in K$ oder $Q \in K$ folgt.

3.4 Determinante eines Endomorphismus und Orientierung*

3.4.1. Nach den vielen Rechnungen mit Determinanten kommen wir zurück zur Frage nach ihrer geometrischen Bedeutung. Eine Matrix $A \in M(n \times n; K)$ kann auch als Endomorphismus des K^n angesehen werden. Dabei wird die Standardbasis $e_1, \dots, e_n$ auf die Spaltenvektoren $a^1, \dots, a^n$ von A abgebildet. Im Fall $K = \mathbb{R}$ hat der von $e_1, \dots, e_n$ aufgespannte Würfel im $\mathbb{R}^n$ das Volumen 1, das von $a^1, \dots, a^n$ aufgespannte Parallelotop hat das Volumen

$$|\det A| = |\det A|$$

(vgl. [Fo3], §5). Der Betrag der Determinante ist also das Maß für die durch den Endomorphismus bewirkte Veränderung des Volumens.

Sind ein beliebiger K -Vektorraum V mit $\dim V < \infty$ und ein Endomorphismus

$$F: V \rightarrow V$$

gegeben, so kann man eine Basis B von V wählen und

$$\det F := \det M_B(F) \in K$$

setzen. Ist A eine andere Basis von V , so gibt es nach 2.6.5 ein $S \in GL(n; K)$ mit

$$M_A(F) = S \cdot M_B(F) \cdot S^{-1}, \quad \text{also ist} \quad \det M_A(F) = \det M_B(F)$$

nach D11 aus 3.1.3. Daher haben wir eine wohldefinierte Abbildung

$$\det: \text{End}(V) \rightarrow K, \quad F \mapsto \det F.$$

Im Gegensatz zum $\mathbb{R}^n$ hat man in einem allgemeinen V keine Volumenmessung mehr. Immerhin folgt aus D10 in 3.1.3 die

Bemerkung. Für $F \in \text{End}(V)$ sind folgende Bedingungen gleichwertig:

i) F ist surjektiv.

ii) $\det F \neq 0$. □

3.4.2. Für einen Endomorphismus des $\mathbb{R}^n$ ist der Betrag der Determinante ein Maß für die Veränderung der Volumina (vgl. 3.4.1). Das Vorzeichen der Determinante hat ebenfalls eine geometrische Bedeutung.

Beispiele. a) Sei

$$A = \begin{pmatrix} -1 & -\frac{1}{4} \\ 1 & -\frac{1}{2} \end{pmatrix} \quad \text{und} \quad A' = \begin{pmatrix} -\frac{1}{4} & 1 \\ 1 & \frac{1}{2} \end{pmatrix}.$$

Es ist $\det A = \frac{3}{4} > 0$ und $\det A' = -\frac{9}{8} < 0$. Die Wirkung von A auf den Buchstaben F ist in Bild 3.7 zu sehen: Das Bild unter A ist nach einer Drehung wieder als F zu erkennen, das Bild unter A' ist gespiegelt.

b) Bei einem Automorphismus A des $\mathbb{R}^3$ betrachtet man das Bild einer linken Hand. Ist $\det A > 0$, so sieht es wieder aus wie eine linke Hand, falls $\det A < 0$, wird daraus eine rechte Hand.

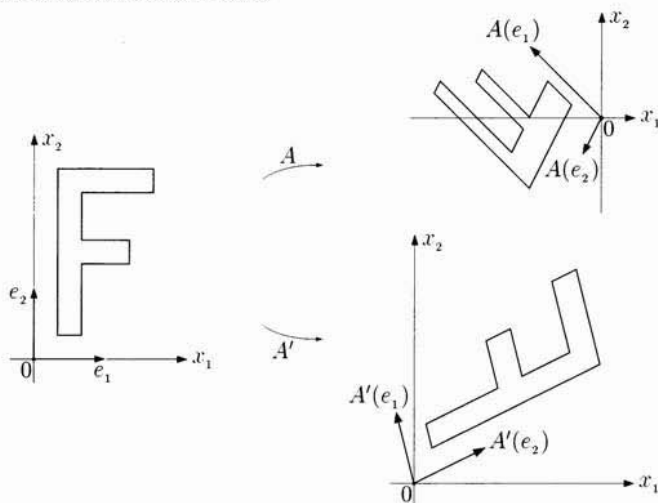


Bild 3.7

Das motiviert die

Definition. Ein Automorphismus eines $\mathbb{R}$ -Vektorraumes V mit $\dim V < \infty$ heißt
orientierungstreu, falls $\det F > 0$, und
orientierungsuntreu, falls $\det F < 0$.

3.4.3. Es ist bezeichnend für die Schwierigkeit der Definition, daß man „orientierungstreu“ erklären kann, bevor klar ist, was eine „Orientierung“ ist. Das wird nun nachgeholt:

Definition. Seien $\mathcal{A} = (v_1, \dots, v_n)$ und $\mathcal{B} = (w_1, \dots, w_n)$ Basen des $\mathbb{R}$ -Vektorraumes V und F der nach 2.4.1 eindeutig bestimmte Automorphismus von V mit

$$F(v_1) = w_1, \dots, F(v_n) = w_n.$$

Dann heißen $\mathcal{A}$ und $\mathcal{B}$ *gleichorientiert*, in Zeichen $\mathcal{A} \sim \mathcal{B}$, wenn $\det F > 0$.

Offensichtlich ist dadurch eine Äquivalenzrelation in der Menge X aller Basen von V erklärt, und X zerfällt in zwei Äquivalenzklassen

$$X = X^* \cup X_*,$$

wobei je zwei Basen aus derselben Klasse gleichorientiert sind. Man beachte, daß X^* und X_* völlig gleichberechtigt sind.

Unter einer *Orientierung* von V versteht man eine Äquivalenzklasse gleichorientierter Basen; dafür gibt es zwei Möglichkeiten.

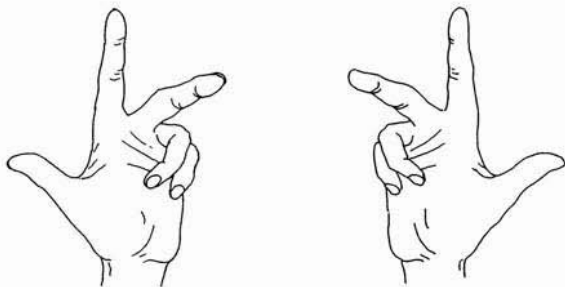


Bild 3.8 (aus [B 2])

Für $V = \mathbb{R}^n$ kann man die Zerlegung von X explizit beschreiben. Man hat eine kanonische bijektive Abbildung

$$M: X \rightarrow \mathrm{GL}(n; \mathbb{R}), \quad \mathcal{A} \mapsto A = M(\mathcal{A}),$$

wenn $M(\mathcal{A})$ die Matrix mit den Vektoren der Basis $\mathcal{A}$ als Spalten bezeichnet (vgl. Aufgabe 1), und es gilt

$$\mathcal{A} \sim \mathcal{B} \Leftrightarrow \det M(\mathcal{A}) \cdot \det M(\mathcal{B}) > 0.$$

Man beachte dabei, daß $\det M(\mathcal{B})$ und $\det M(\mathcal{B})^{-1}$ das gleiche Vorzeichen haben. Die Gruppe $\mathrm{GL}(n; \mathbb{R})$ hat eine disjunkte Zerlegung in

$$G_+ := \{A \in \mathrm{GL}(n; \mathbb{R}) : \det A > 0\} \quad \text{und}$$

$$G_- := \{A \in \mathrm{GL}(n; \mathbb{R}) : \det A < 0\}.$$

Im $\mathbb{R}^n$ gibt es die *kanonische* (d.h. nach DUDEN „den kirchlichen Bestimmungen gemäß“) Basis $\mathcal{K}$, sie ist in einer der beiden Klassen X^* oder X_* enthalten, und diese ist dadurch ausgezeichnet (also ist es hier schon wieder vorbei mit der Gleichberechtigung). Ist etwa $\mathcal{K} \in X^*$, so folgt

$$G_+ = M(X^*) \quad \text{und} \quad G_- = M(X_*).$$

Wir können also im Fall $V = \mathbb{R}^n$ die Zerlegung

$$X = X^* \cup X_* \quad \text{ersetzen durch} \quad \mathrm{GL}(n; \mathbb{R}) = G_+ \cup G_-.$$

Offensichtlich ist $G_+ \subset GL(n; \mathbb{R})$ eine Untergruppe, und es gilt die

Bemerkung. Ist $T \in GL(n; \mathbb{R})$ mit $\det T < 0$ gegeben, und ist

$$\tau: GL(n; \mathbb{R}) \rightarrow GL(n; \mathbb{R}), \quad A \mapsto A \cdot T,$$

die Rechtstranslation (vgl. 1.2.4), so gilt

$$\tau(G_+) = G_- \quad \text{und} \quad \tau(G_-) = G_+.$$

Der Beweis verläuft wie in 3.2.4. □

3.4.4. Nun wollen wir zeigen, daß zwei Basen des $\mathbb{R}^n$ genau dann gleichorientiert sind, wenn sie sich „stetig ineinander verformen“ lassen. Vor allem muß präzisiert werden, was das heißen soll. Dabei können wir entsprechend 3.4.3 die Menge aller Basen durch $GL(n; \mathbb{R})$ ersetzen.

Definition. Sind $A, B \in GL(n; \mathbb{R})$, so versteht man unter einem Weg von A nach B eine stetige Abbildung

$$\varphi: I \rightarrow GL(n; \mathbb{R}), \quad t \mapsto \varphi(t) = (\varphi_{ij}(t)),$$

wobei $I = [\alpha, \beta] \subset \mathbb{R}$ ein Intervall ist, mit $\varphi(\alpha) = A$ und $\varphi(\beta) = B$. Die Stetigkeit von φ bedeutet, daß die n^2 Funktionen φ_{ij} mit Werten in $\mathbb{R}$ stetig sind. Wesentlich dabei ist, daß die Matrix $\varphi(t)$ für jedes $t \in I$ invertierbar ist.

Die Matrizen $A, B \in GL(n; \mathbb{R})$ nennt man *verbindbar* (in Zeichen $A \sim B$), wenn es einen Weg von A nach B gibt.

Man betrachtet t als Zeitparameter, die Spalten von $\varphi(t)$ beschreiben eine Basis, die sich im Lauf der Zeit verformt.

Theorem. Für zwei Matrizen $A, B \in GL(n; \mathbb{R})$ sind folgende Bedingungen gleichwertig:

- i) A und B liegen in derselben Klasse G_+ oder G_- , d.h. $\det(A \cdot B) > 0$,
- ii) A und B sind in $GL(n; \mathbb{R})$ verbindbar.

In der Sprache der Topologie sagt man dafür, daß $GL(n, \mathbb{R})$ in die zwei Zusammenhangskomponenten G_+ und G_- zerfällt.

Korollar. Zwei Basen A und B des $\mathbb{R}^n$ sind genau dann gleichorientiert, wenn sie stetig ineinander verformbar sind.

Wesentlich dabei ist wieder, daß man zu jedem Zeitpunkt der Verformung eine Basis hat.

Beispiele. a) Im $\mathbb{R}^1$ besteht eine Basis aus einem einzigen Vektor $x \in \mathbb{R}^* = GL(1; \mathbb{R})$. Ist y ein weiterer, so gibt es genau dann einen Weg von x nach y ,

wenn beide gleiches Vorzeichen haben (Bild 3.9). Dabei benutzt man den Zwischenwertsatz ([Fo1], §11).

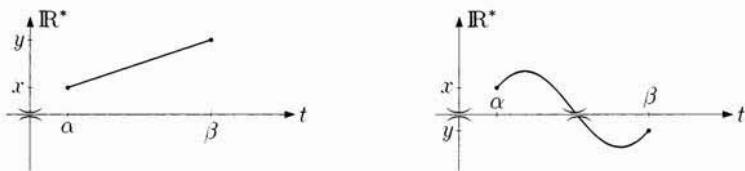


Bild 3.9

b) Die Basen $\mathcal{A} = (e_1, e_2)$ und $\mathcal{B} = (-e_1, e_2)$ sind nicht stetig ineinander verformbar. Man mache dazu ein Experiment mit einer nicht mehr benutzten oder gedachten Uhr: versuche, sie durch unabhängige Bewegung beider Zeiger von 3 Uhr auf 9 Uhr zu stellen, so daß die beiden Zeiger dabei nie auf einer Geraden liegen!

c) Daumen, Zeigefinger und Mittelfinger einer nicht ganz flach gestreckten linken und rechten Hand repräsentieren die beiden möglichen Orientierungen des $\mathbb{R}^3$. Dreht man so lange, bis Mittelfinger und Zeigefinger der beiden Hände aufeinander liegen, so zeigen die Daumen in verschiedene Richtungen.

Wir benutzen folgende Teilresultate:

Lemma 1. Die Verbindbarkeit ist eine Äquivalenzrelation in $GL(n; \mathbb{R})$.

Lemma 2. Ist $\varphi: I \rightarrow GL(n; \mathbb{R})$ ein Weg von A nach B , so haben $\det A$ und $\det B$ das gleiche Vorzeichen.

Lemma 3. Ist $A \in GL(n; \mathbb{R})$ mit $\det A > 0$, so gibt es einen Weg von A nach E_n .

Daraus ergibt sich leicht der *Beweis des Theorems*: Nach Lemma 2 gilt ii) $\Rightarrow$ i). Es bleibt also i) $\Rightarrow$ ii) zu zeigen:

Sind $A, B \in G_+$, so gibt es nach Lemma 3 Wege von A nach E_n und von B nach E_n , also nach Lemma 1 auch einen Weg von A nach B .

Sind $A, B \in G_-$, so setze man etwa

$$T := \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}, \quad \text{dann ist } T^{-1} = T.$$

Die Rechtstranslation $C \mapsto C \cdot T$ bewirkt eine Multiplikation der ersten Spalte mit -1 , das ist eine stetige Abbildung von $\mathrm{GL}(n; \mathbb{R})$ auf sich. Da $A \cdot T$ und $B \cdot T \in G_+$, gibt es einen Weg dazwischen, durch die Rechtstranslation mit T erhält man daraus einen Weg von A nach B . $\square$

Zum *Beweis von Lemma 1* verweisen wir auf Aufgabe 2. $\square$

Beweis von Lemma 2. Nach 3.2.6 ist die Determinante stetig, also ist

$$\det \circ \varphi: I \rightarrow \mathbb{R}, \quad t \mapsto \det(\varphi(t)),$$

eine stetige Funktion. Da $\det(\varphi(t)) \neq 0$ für alle t , hat $\det(\varphi(t))$ nach dem Zwischenwertsatz für alle t das gleiche Vorzeichen. $\square$

Beweis von Lemma 3. Durch eine Folge verschiedenartiger Umformungen wird A in E_n überführt. Der Kniff ist, diese Umformungen nicht auf einen Schlag, sondern langsam durchzuführen. Dadurch erhält man kleine Wegstückchen, die man schließlich zum ganzen Weg von A nach E_n zusammenfügt.

Zur Vorbereitung der *ersten Etappe* überlegt man sich, daß jede invertierbare Matrix durch Spaltenumformungen vom Typ III (Addition einer μ -fachen i -ten Spalte zu einer j -ten Spalte mit $j \neq i$) in eine Diagonalmatrix

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

überführt werden kann (Aufgabe 3). Es ist wichtig, Spaltenvertauschungen zu vermeiden, weil sie das Vorzeichen der Determinante verändern.

Seien also Elementarmatrizen $C_1, \dots, C_k$ vom Typ III gefunden, so daß

$$D = A \cdot C_1 \cdot \dots \cdot C_k.$$

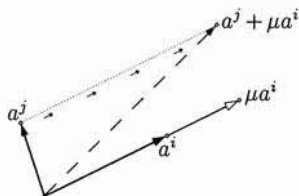
Ist $C = Q_i^j(\mu)$ eine davon, so zeigen wir, wie daraus ein Wegstückchen gemacht werden kann. Ist

$$A = (\dots, a^i, \dots, a^j, \dots), \quad \text{so ist} \quad A \cdot Q_i^j(\mu) = (\dots, a^i, \dots, a^j + \mu a^i, \dots).$$

Definieren wir eine stetige Abbildung

$$\psi: [0, 1] \rightarrow \mathrm{GL}(n; \mathbb{R}), \quad t \mapsto A \cdot Q_i^j(t \cdot \mu),$$

so ist $\psi(0) = A$ und $\psi(1) = A \cdot Q_i^j(\mu)$. Wie die stetige Ausführung dieser Scherung verläuft, ist in Bild 3.10 angedeutet.

**Bild 3.10**

Nach Lemma 1 kann man die so erhaltenen k Wegstückchen zusammensetzen zur ersten Etappe

$$\varphi_1: [\alpha_1, \beta_1] \rightarrow \text{GL}(n; \mathbb{R}) \quad \text{mit} \quad \varphi_1(\alpha_1) = A \quad \text{und} \quad \varphi_1(\beta_1) = D.$$

Aus D7 folgt $\det(\varphi_1(t)) = \det A$ für jedes $t \in [\alpha_1, \beta_1]$, insbesondere

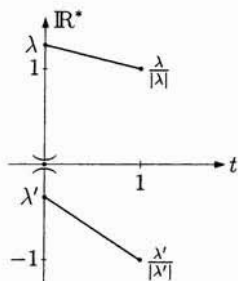
$$\det A = \det D = \lambda_1 \cdot \dots \cdot \lambda_n.$$

In der zweiten Etappe wird D verbunden mit

$$D' = \begin{pmatrix} \varepsilon_1 & & \\ & \ddots & \\ & & \varepsilon_n \end{pmatrix}, \quad \text{wobei} \quad \varepsilon_i = \frac{\lambda_i}{|\lambda_i|} = \pm 1.$$

Dazu benutzen wir, daß man in $\mathbb{R}^*$ jedes λ mit $\varepsilon = \frac{\lambda}{|\lambda|}$ geradlinig verbinden kann durch

$$\chi: [0, 1] \rightarrow \mathbb{R}^*, \quad t \mapsto \lambda + (\varepsilon - \lambda)t = \lambda \left(1 + \left(\frac{1}{|\lambda|} - 1 \right) t \right).$$

**Bild 3.11**

Die Multiplikation der Spalten mit Faktoren kann man auch durch Elementarmatrizen vom Typ $S_j(\mu)$ bewirken, also ist

$$D' = D \cdot S_1 \left(\frac{1}{|\lambda_1|} \right) \cdot \dots \cdot S_n \left(\frac{1}{|\lambda_n|} \right),$$

und jedes der n Wegstückchen von D nach D' ist beschrieben durch Multiplikation mit

$$S_j \left(1 + \left(\frac{1}{|\lambda_j|} - 1 \right) t \right)$$

für $t \in [0, 1]$. Diese n Stückchen zusammengefügt ergeben einen Weg

$$\varphi_2: [\alpha_2, \beta_2] \rightarrow \mathrm{GL}(n; \mathbb{R}) \quad \text{mit} \quad \varphi_2(\alpha_2) = D \quad \text{und} \quad \varphi_2(\beta_2) = D'.$$

Da $1 + \left(\frac{1}{|\lambda|} - 1 \right) t > 0$ für $0 \leq t \leq 1$, ist $\det D' = +1$, also ist die Anzahl der Einträge -1 in D' gerade.

In der *dritten Etappe* wird D' mit E_n verbunden, dazu nimmt man sich in jedem Schritt ein Pärchen von negativen Einsen vor. Wie sie gemeinsam ins Positive gewendet werden können, sieht man am besten im Spezialfall $n = 2$. Man nimmt die stetige Abbildung

$$\alpha: [-\pi, 0] \rightarrow \mathrm{GL}(2; \mathbb{R}), \quad t \mapsto \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}.$$

Für sie gilt

$$\alpha(-\pi) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \alpha(0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = E_2.$$

Falls die beiden Einträge -1 an beliebigen Stellen sitzen, verwendet man die Abbildung

$$t \mapsto \begin{pmatrix} \ddots & & & \\ & \cos t & \cdots & -\sin t \\ & \vdots & \ddots & \vdots \\ & \sin t & \cdots & \cos t \\ & & & \ddots \end{pmatrix} \in \mathrm{GL}(n; \mathbb{R}).$$

Macht man dies nacheinander für jedes Paar, so erhält man insgesamt das Wegstück

$$\varphi_3: [\alpha_3, \beta_3] \rightarrow \mathrm{GL}(n; \mathbb{R}) \quad \text{mit} \quad \varphi_3(\alpha_3) = D' \quad \text{und} \quad \varphi_3(\beta_3) = E_n.$$

Durch Zusammensetzung von φ_1 , φ_2 und φ_3 erhält man schließlich den gesuchten Weg von A nach E_n . $\square$

Wie wir gesehen haben, gibt es für einen reellen Vektorraum zwei mögliche Orientierungen. Viel schwieriger ist es, den Begriff der Orientierung für eine reelle Mannigfaltigkeit, etwa eine Fläche, zu erklären. Das einfachste Beispiel für eine nicht orientierbare Fläche ist das um 1850 entdeckte **MÖBIUSband**, das nicht nur Anselm Wüßteger und Sophie beschäftigt hat (mehr darüber in [P]):



Aufgaben zu 3.4

1. Sei V ein K -Vektorraum, X die Menge aller Basen von V und $B \in X$. Zeigen Sie, dass die Abbildung

$$\Phi: X \rightarrow \mathrm{GL}(n; K), \quad \mathcal{A} \mapsto T_{\mathcal{A}}^B = M_{\mathcal{A}}^B(\mathrm{id})$$

bijektiv ist. Wie hängt Φ im Fall $V = \mathbb{R}^n$ mit der in 3.4.3 definierten kanonischen Bijektion

$$M: X \rightarrow \mathrm{GL}(n; \mathbb{R})$$

zusammen?

2. Beweisen Sie, dass die Verbindbarkeit von Matrizen in $\mathrm{GL}(n; \mathbb{R})$ eine Äquivalenzrelation in $\mathrm{GL}(n; \mathbb{R})$ definiert.

3. Zeigen Sie, dass man eine invertierbare Matrix $A \in \mathrm{GL}(n; K)$ durch Spaltenumformungen vom Typ III auf Diagonalgestalt bringen kann.

4. Zeigen Sie, dass in $M(n \times n; \mathbb{R})$ je zwei Matrizen durch einen Weg verbindbar sind.

5. Beweisen Sie, dass $\mathrm{GL}(n; \mathbb{C})$ zusammenhängend ist, das heißt, dass je zwei Matrizen aus $\mathrm{GL}(n; \mathbb{C})$ durch einen Weg in $\mathrm{GL}(n; \mathbb{C})$ verbunden sind.

Kapitel 4

Eigenwerte

In Abschnitt 2.2.4 hatten wir für eine lineare Abbildung $F: V \rightarrow W$ ein Paar von Basen konstruiert, bezüglich derer F durch

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

mit $r = \text{rang } F$ dargestellt wird. Die nötigen Transformationsmatrizen sind ganz einfach explizit zu berechnen (vgl. 2.7.6).

Zur Beschreibung eines Endomorphismus benutzt man eine einzige Basis, und ihre Anpassung an eine lineare Abbildung ist weit schwieriger als wenn man zwei Basen variieren kann. In die Matrizenrechnung übersetzt bedeutet diese Frage, zu einer quadratischen Matrix A eine möglichst einfache ähnliche Matrix

$$B = SAS^{-1}$$

zu finden (vgl. 2.6.7). Insbesondere wird sich zeigen, daß man hierzu stärkere Hilfsmittel aus der Algebra, nämlich Polynome höheren Grades benötigt, obwohl in der *linearen* Algebra zunächst nur der Grad 1 interessiert. Die benutzten Tatsachen über Polynome sind in Abschnitt 1.3 zusammengestellt.

4.1 Beispiele und Definitionen

4.1.1. Wir beginnen mit dem einfachsten Fall, nämlich einem K -Vektorraum V mit $\dim V = 1$. Ist $F \in \text{End}(V)$ und $0 \neq v \in V$, so gibt es ein eindeutig bestimmtes $\lambda \in K$ mit

$$F(v) = \lambda \cdot v,$$

und dieses λ hängt nicht von der Wahl von v ab. Denn ist $w = \mu v$ ein anderer Vektor von V , so ist

$$F(w) = F(\mu v) = \mu \lambda v = \lambda \mu v = \lambda w.$$

Die Abbildung F ist also durch die Zahl λ festgelegt.

Auch in höherdimensionalen Räumen darf man auf Vektoren hoffen, die bei Anwendung von F nur mit einem Faktor multipliziert werden, weil dann die Wirkung von F wenigstens in bestimmten Richtungen besonders einfach zu verstehen ist. Das führt zu der

Definition. Sei F ein Endomorphismus des K -Vektorraums V . Ein $\lambda \in K$ heißt *Eigenwert* von F , wenn es ein $v \in V$ mit $v \neq 0$ gibt, so daß gilt

$$F(v) = \lambda v.$$

Jedes vom Nullvektor verschiedene $v \in V$ mit $F(v) = \lambda v$ heißt *Eigenvektor* von F (zum Eigenwert λ).

Vorsicht! Man beachte, daß natürlich $0 \in K$ ein Eigenwert sein kann, der Nullvektor $0 \in V$ jedoch nach Definition nie Eigenvektor ist.

Das zentrale Problem dieses Kapitels ist die Frage nach der Existenz und der Vielfalt von Eigenvektoren. Dazu zunächst drei

Beispiele. a) Es sei $V = \mathbb{R}^2$ und F eine Drehung um den Winkel α , die beschrieben wird durch die Matrix

$$A = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}.$$

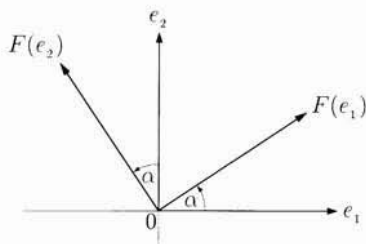


Bild 4.1

Anschaulich ist klar, daß es mit Ausnahme der Fälle $\alpha = 0$ und $\alpha = \pi$ keinen Eigenvektor geben kann.

b) Wir variieren das erste Beispiel, indem wir die Richtung von $F(e_2)$ umkehren. Dann ist die beschreibende Matrix

$$A' = \begin{pmatrix} \cos \alpha & \sin \alpha \\ \sin \alpha & -\cos \alpha \end{pmatrix}.$$

Wenn man die Gerade mit dem Winkel $\frac{\alpha}{2}$ einzeichnet, erkennt man, daß die Abbildung F eine Spiegelung an dieser Geraden ist. Damit hat man zwei Eigenvektoren gefunden:

$$\begin{aligned} v_1 &= \left(\cos \frac{\alpha}{2}, \sin \frac{\alpha}{2} \right) && \text{zum Eigenwert } \lambda_1 = 1 \text{ und} \\ v_2 &= \left(\cos \frac{\alpha+\pi}{2}, \sin \frac{\alpha+\pi}{2} \right) && \text{zum Eigenwert } \lambda_2 = -1, \end{aligned}$$

und diese beiden bilden eine Basis $\mathcal{B} = (v_1, v_2)$ von $\mathbb{R}^2$ mit

$$M_{\mathcal{B}}(F) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

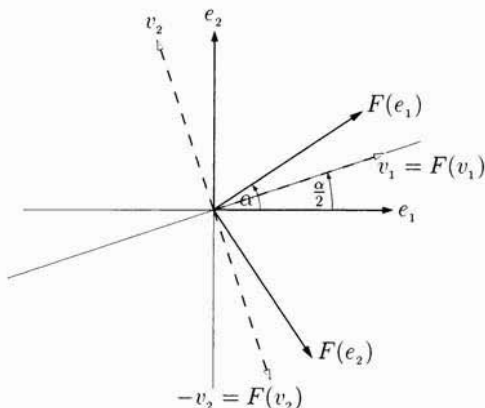


Bild 4.2

c) Sei $I \subset \mathbb{R}$ ein Intervall und $V = \mathcal{D}(I; \mathbb{R})$ der unendlichdimensionale $\mathbb{R}$ -Vektorraum der auf I beliebig oft differenzierbaren Funktionen. Ein Endomorphismus ist gegeben durch

$$F: V \rightarrow V, \quad \varphi \mapsto \varphi'.$$

Dieses F hat jedes beliebige $\lambda \in \mathbb{R}$ als Eigenwert, denn

$$\varphi(x) := ce^{\lambda x}$$

ist für jedes $c \in \mathbb{R}^*$ ein Eigenvektor zu λ . Das ist ein erster Hinweis auf die Bedeutung der Eigenwerttheorie für Differentialgleichungen (vgl. auch Aufgabe 3).

4.1.2. Endomorphismen wie im obigen Beispiel b) erhalten einen eigenen Namen:

Definition. Ein Endomorphismus heißt *diagonalisierbar*, wenn es eine Basis aus Eigenvektoren gibt.

Dieser Name erklärt sich aus der

Bemerkung. Ist $\dim V = n < \infty$, so ist $F \in \text{End}(V)$ genau dann diagonalisierbar, wenn es eine Basis $\mathcal{B} = (v_1, \dots, v_n)$ von V gibt, so daß $M_{\mathcal{B}}(F)$ eine Diagonalmatrix ist, d.h.

$$M_{\mathcal{B}}(F) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Das folgt unmittelbar aus der Definition der darstellenden Matrix. $\square$

Insbesondere nennt man eine Matrix $A \in M(n \times n; K)$ diagonalisierbar, wenn der durch A beschriebene Endomorphismus von K^n diagonalisierbar ist. Nach 2.6.7 ist das gleichwertig mit der Existenz einer Matrix $S \in GL(n; K)$ mit

$$SAS^{-1} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

d.h. A ist ähnlich zu einer Diagonalmatrix.

Es sei daran erinnert, daß die Spalten von S^{-1} eine Basis aus Eigenvektoren von A sind. Das sieht man zum Beispiel durch Multiplikation obiger Gleichung von links mit S^{-1} .

Vorsicht! Selbst wenn F diagonalisierbar ist, braucht nicht *jeder* Vektor ungleich Null aus V ein Eigenvektor zu sein. Man mache sich das an Beispielen klar (etwa Beispiel b) oben oder Bild 4.3)!

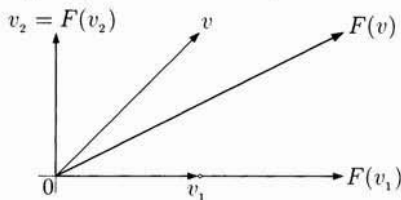


Bild 4.3

4.1.3. Bevor wir eine Methode angeben, Eigenwerte zu finden, beweisen wir den

Satz. Angenommen, $F \in \text{End}(V)$ hat paarweise verschiedene Eigenwerte $\lambda_1, \dots, \lambda_n$, wobei $n = \dim V$. Dann ist F diagonalisierbar.

Da es zu jedem Eigenwert λ_i mindestens einen Eigenvektor $v_i \neq 0$ gibt, folgt der Satz sofort aus dem

Lemma. Gegeben seien $F \in \text{End}(V)$ und Eigenvektoren $v_1, \dots, v_m$ zu paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_m$. Dann sind $v_1, \dots, v_m$ linear unabhängig. Insbesondere ist $m \leq \dim V$.

Beweis. Wir führen Induktion über m . Der Fall $m = 1$ ist wegen $v_1 \neq 0$ klar. Sei $m \geq 2$, und sei die Aussage für $m - 1$ bereits bewiesen. Wir betrachten die Bedingung

$$\alpha_1 v_1 + \dots + \alpha_m v_m = 0. \quad (*)$$

Darauf können wir einerseits F anwenden, andererseits mit λ_1 multiplizieren und die Ergebnisse voneinander subtrahieren. Das ergibt

$$\begin{aligned} \alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_2 v_2 + \dots + \alpha_m \lambda_m v_m &= 0, \\ \alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_1 v_2 + \dots + \alpha_m \lambda_1 v_m &= 0 \end{aligned}$$

und

$$\alpha_2 (\lambda_2 - \lambda_1) v_2 + \dots + \alpha_m (\lambda_m - \lambda_1) v_m = 0.$$

Nach Induktionsannahme sind $v_2, \dots, v_m$ linear unabhängig, also ist

$$\alpha_2 (\lambda_2 - \lambda_1) = \dots = \alpha_m (\lambda_m - \lambda_1) = 0,$$

und da die Eigenwerte verschieden sind, folgt

$$\alpha_2 = \dots = \alpha_m = 0.$$

Setzt man das in $(*)$ ein, so folgt $\alpha_1 v_1 = 0$ und wegen $v_1 \neq 0$ auch $\alpha_1 = 0$. $\square$

4.1.4. Wie wir gesehen haben, gibt es höchstens $n = \dim V$ Eigenwerte, dagegen im allgemeinen sehr viel mehr Eigenvektoren. Daher ist es nützlich, alle Eigenvektoren zu einem festen Eigenwert zusammenzufassen.

Definition. Ist F ein Endomorphismus von V und $\lambda \in K$, so nennen wir

$$\text{Eig}(F; \lambda) := \{v \in V : F(v) = \lambda v\}$$

den *Eigenraum* von F bezüglich λ .

Bemerkung.

- a) $\text{Eig}(F; \lambda) \subset V$ ist ein Untervektorraum.
- b) λ ist Eigenwert von $F \Leftrightarrow \text{Eig}(F; \lambda) \neq \{0\}$.
- c) $\text{Eig}(F; \lambda) \setminus \{0\}$ ist die Menge der zu λ gehörigen Eigenvektoren von F .

d) $\text{Eig}(F; \lambda) = \text{Ker}(F - \lambda \text{id}_V)$.

e) Sind $\lambda_1, \lambda_2 \in K$ verschieden, so ist $\text{Eig}(F; \lambda_1) \cap \text{Eig}(F; \lambda_2) = \{0\}$.

Beweis. a) bis d) ist ganz klar, e) folgt aus dem Lemma in 4.1.3. $\square$

Aufgaben zu 4.1

1. Zeigen Sie: Ein nilpotenter Endomorphismus hat null als einzigen Eigenwert.

2. Gegeben sei die lineare Abbildung $F: \mathcal{D}(I; \mathbb{R}) \rightarrow \mathcal{D}(I; \mathbb{R})$, $\varphi \mapsto \varphi'$, wobei $I \subset \mathbb{R}$ ein Intervall ist.

a) Bestimmen Sie die reellen Eigenwerte von F .

b) Bestimmen Sie eine Basis von $\text{Eig}(F, -1)$.

3. Sei $I \subset \mathbb{R}$ ein offenes Intervall. Durch eine Matrix $A \in M(n \times n; \mathbb{R})$ ist das homogene lineare Differentialgleichungssystem

$$y' = A \cdot y$$

bestimmt; nach Aufgabe 11 zu 3.2 hat der zugehörige Lösungsraum

$$\mathcal{L}_0 = \{\varphi \in \mathcal{D}(I; \mathbb{R}^n): \varphi' = A \cdot \varphi\} \subset \mathcal{D}(I; \mathbb{R}^n)$$

die Dimension n . Um Lösungen zu erhalten, kann man den Ansatz

$$\varphi(t) = e^{\lambda t} \cdot v$$

benutzen, wobei $\lambda \in \mathbb{R}$ und $v \in \mathbb{R}^n$. Zeigen Sie:

a) $\varphi(t) = e^{\lambda t} \cdot v$ ist eine Lösung $\neq 0$ von $y' = A \cdot y$ genau dann, wenn v Eigenvektor von A zum Eigenwert λ ist.

b) Lösungen $\varphi^{(1)}(t) = e^{\lambda_1 t} \cdot v_1, \dots, \varphi^{(k)}(t) = e^{\lambda_k t} \cdot v_k$ sind linear unabhängig genau dann, wenn $v_1, \dots, v_k$ linear unabhängig sind.

Insbesondere erhält man mit diesem Ansatz eine Basis des Lösungsraums, falls A diagonalisierbar ist.

4. Sei V ein K -Vektorraum und $F: V \rightarrow V$ linear. Zeigen Sie: Hat $F^2 + F$ den Eigenwert -1 , so hat F^3 den Eigenwert 1 .

5. Gegeben sei ein K -Vektorraum V und $F, G \in \text{End}(V)$. Beweisen Sie:

a) Ist $v \in V$ Eigenvektor von $F \circ G$ zum Eigenwert $\lambda \in K$, und ist $G(v) \neq 0$, so ist $G(v)$ Eigenvektor von $G \circ F$ zum Eigenwert λ .

b) Ist V endlichdimensional, so haben $F \circ G$ und $G \circ F$ dieselben Eigenwerte.

4.2 Das charakteristische Polynom

In diesem Abschnitt entwickeln wir eine systematische Methode zur Suche von Eigenwerten und zur Bestimmung der zugehörigen Eigenräume. Entscheidend dafür ist die Determinante eines Endomorphismus, von der wir in 3.4.1 gesehen hatten, daß sie wohldefiniert ist, wenn der zugrundeliegende Vektorraum *endlichdimensional* ist. Das muß in diesem ganzen Abschnitt stillschweigend vorausgesetzt werden.

4.2.1. Grundlegend ist die folgende einfache

Bemerkung. Für $F \in \text{End}(V)$ und $\lambda \in K$ sind folgende Bedingungen gleichwertig:

i) λ ist Eigenwert von F .

ii) $\det(F - \lambda \text{id}_V) = 0$.

Beweis. Für festes $\lambda \in K$ ist die Existenz eines $v \neq 0$ mit $F(v) = \lambda v$ gleichbedeutend mit

$$\begin{aligned} F(v) - \lambda v &= 0 \\ \Leftrightarrow (F - \lambda \text{id}_V)(v) &= 0 && \text{wegen der Linearität,} \\ \Leftrightarrow \text{Ker}(F - \lambda \text{id}_V) &\neq \{0\} && \text{nach der Definition des Kerns,} \\ \Leftrightarrow \text{Im}(F - \lambda \text{id}_V) &\neq V && \text{nach der Dimensionsformel aus 2.2.4,} \\ \Leftrightarrow \text{rang}(F - \lambda \text{id}_V) &< \dim V && \text{nach Definition des Ranges,} \\ \Leftrightarrow \det(F - \lambda \text{id}_V) &= 0 && \text{nach 3.4.1.} \quad \square \end{aligned}$$

4.2.2. Durch die obige Bemerkung ist die Suche nach Eigenwerten zurückgeführt auf die Suche nach Nullstellen der Abbildung

$$\tilde{P}_F: K \rightarrow K, \quad \lambda \mapsto \det(F - \lambda \text{id}_V).$$

Diese nennen wir die *charakteristische Funktion* von F . Wir zeigen, daß sie durch ein Polynom beschrieben wird.

Sei dazu $\mathcal{A}$ eine Basis von V und $A = M_{\mathcal{A}}(F)$. Ist t eine Unbestimmte, so definieren wir

$$P_A(t) := \det(A - t \cdot E_n) = \begin{vmatrix} a_{11} - t & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - t & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - t \end{vmatrix}.$$

Die Definition dieser Determinante ist etwas problematisch, weil in der Matrix die Unbestimmte t vorkommt. Man hat mehrere Möglichkeiten, damit umzugehen:

- 1) Man kann für t beliebige $\lambda \in K$ einsetzen und damit rechnen. Für unendlichen K ist das nach Korollar 2 aus 1.3.8 keine Einschränkung.
- 2) Man kann die Determinante formal nach der Formel von LEIBNIZ aus 3.2.5 ausrechnen. Eine Rechtfertigung dafür wurde in 3.2.8 gegeben.
- 3) Man kann die Einträge von $A - t \cdot E_n$ als Elemente des Körpers $K(t)$ der rationalen Funktionen ansehen (vgl. Aufgabe 10 in 1.3).

Das Ergebnis der Rechnung ist

$$P_A(t) = (a_{11} - t) \cdot \dots \cdot (a_{nn} - t) + Q,$$

wobei der erste Summand zur identischen Permutation gehört und Q die restliche Summe über $S_n \setminus \{\text{id}\}$ ist. Da in einem Summanden von Q als Faktoren höchstens $n - 2$ Diagonalkomponenten auftreten können, ist Q ein Polynom vom Grad höchstens $n - 2$. Nun ist

$$(a_{11} - t) \cdot \dots \cdot (a_{nn} - t) = (-1)^n t^n + (-1)^{n-1} (a_{11} + \dots + a_{nn}) t^{n-1} + Q_1,$$

wobei Q_1 ein Polynom vom Grad höchstens $n - 2$ ist. Also ist $P_A(t)$ ein Polynom vom Grad n mit Koeffizienten aus K , d.h. es gibt $\alpha_0, \dots, \alpha_n \in K$, so daß

$$P_A(t) = \alpha_n t^n + \alpha_{n-1} t^{n-1} + \dots + \alpha_1 t + \alpha_0.$$

Dabei ist

$$\begin{aligned} \alpha_n &= (-1)^n, \\ \alpha_{n-1} &= (-1)^{n-1} (a_{11} + \dots + a_{nn}) \quad \text{und} \\ \alpha_0 &= \det A. \end{aligned}$$

Man nennt $(a_{11} + \dots + a_{nn})$ auch die *Spur* von A . Die Koeffizienten $\alpha_1, \dots, \alpha_{n-2}$ sind nicht so leicht aufzuschreiben, sie haben auch keine speziellen Namen.

Diese Überlegung zeigt, daß $P_A(t)$ ein Element des Polynomrings $K[t]$ ist (siehe 1.3.6). Man nennt

$$P_A(t) = \det(A - t \cdot E_n) \in K[t]$$

das *charakteristische Polynom* der $(n \times n)$ -Matrix A .

Setzt man für die Unbekannte t ein $\lambda \in K$ ein, so erhält man eine Abbildung

$$\tilde{P}_A: K \rightarrow K, \quad \lambda \mapsto P_A(\lambda).$$

Nun zurück zu F . Für jedes $\lambda \in K$ ist

$$M_A(F - \lambda \cdot \text{id}_V) = A - \lambda \cdot E_n.$$

Also ist

$$\det(F - \lambda \cdot \text{id}_V) = \det(A - \lambda \cdot E_n) = P_A(\lambda),$$

d.h. die charakteristische Funktion von F ist beschrieben durch das charakteristische Polynom von A .

Ist nun $\mathcal{B}$ eine weitere Basis von V , so ist $B := M_{\mathcal{B}}(F)$ zu A ähnlich.

Lemma. *Ähnliche Matrizen haben das gleiche charakteristische Polynom.*

Beweis. Sei $B = SAS^{-1}$ mit $S \in \text{GL}(n; K)$. Eine formale Rechnung mit der Unbestimmten t (d.h. genauer eine Rechnung im Ring $M(n \times n; K[t])$, vgl. 3.2.8) ergibt

$$S \cdot t \cdot E_n \cdot S^{-1} = t \cdot E_n.$$

Also ist

$$B - t \cdot E_n = SAS^{-1} - S \cdot t \cdot E_n \cdot S^{-1} = S(A - t \cdot E_n)S^{-1}.$$

Anwendung der Determinante ergibt

$$\det(B - t \cdot E_n) = \det S \cdot \det(A - t \cdot E_n) \cdot (\det S)^{-1} = \det(A - t \cdot E_n)$$

und somit $P_B(t) = P_A(t)$. □

4.2.3. Damit ist gezeigt, daß das charakteristische Polynom der darstellenden Matrix nicht von der Wahl der Basis abhängt. Also ist folgende Definition sinnvoll:

Sei F ein Endomorphismus und $\mathcal{A}$ eine Basis von V . Ist $A = M_{\mathcal{A}}(F)$, so nennen wir

$$P_F(t) := P_A(t)$$

das *charakteristische Polynom* von F .

Insgesamt haben wir folgendes bewiesen:

Satz. *Sei V ein K -Vektorraum der Dimension $n < \infty$ und F ein Endomorphismus von V . Dann hat das charakteristische Polynom $P_F(t) \in K[t]$ die folgenden Eigenschaften:*

1) $\deg P_F = n$.

2) P_F beschreibt die Abbildung $K \rightarrow K, \lambda \mapsto \det(F - \lambda \text{id}_V)$.

3) Die Nullstellen von P_F sind die Eigenwerte von F .

4) Ist A eine Matrix, die F darstellt, so ist $P_F(t) = \det(A - t \cdot E_n)$. □

Damit ist das geometrische Problem, die Eigenwerte eines Endomorphismus zu finden, auf das algebraische Problem der Bestimmung der Nullstellen eines Polynoms zurückgeführt.

4.2.4. Hat man einen Eigenwert gefunden, so ist die Bestimmung des Eigenraumes ganz einfach. Wir können uns dabei auf den Fall $V = K^n$ beschränken. Aus 4.1.4 folgt sofort die

Bemerkung. Ist ein Endomorphismus $A: K^n \rightarrow K^n$ durch die Matrix $A \in M(n \times n; K)$ gegeben, so ist der Eigenraum $\text{Eig}(A; \lambda)$ für jedes $\lambda \in K$ der Lösungsraum des homogenen linearen Gleichungssystems

$$(A - \lambda E_n)x = 0. \quad \square$$

Nach der vielen Theorie ist es höchste Zeit für einige

Beispiele. a) Ist

$$A = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$$

die Matrix einer *Drehung* im $\mathbb{R}^2$ (vgl. 2.1.1), so ist

$$P_A(t) = t^2 - 2t \cos \alpha + 1.$$

Dieses quadratische Polynom hat nach 1.3.9 genau dann eine reelle Wurzel, wenn

$$4 \cos^2 \alpha - 4 \geq 0, \quad \text{d.h.} \quad \cos^2 \alpha = 1$$

gilt. Das ist nur für $\alpha = 0$ und $\alpha = \pi$ der Fall. Diese beiden Drehungen sind trivialerweise diagonalisierbar, alle anderen Drehungen besitzen keine Eigenvektoren. Damit ist noch einmal durch die allgemeine Theorie bestätigt, was wir in 4.1.1 anschaulich gesehen hatten.

b) Für eine Spiegelung

$$A = \begin{pmatrix} \cos \alpha & \sin \alpha \\ \sin \alpha & -\cos \alpha \end{pmatrix}$$

ist $P_A(t) = t^2 - 1 = (t+1)(t-1)$. Also ist A nach 4.1.3 diagonalisierbar, was wir in 4.1.1 schon direkt bewiesen hatten.

c) Für

$$A = \begin{pmatrix} -1 & 6 \\ -1 & 4 \end{pmatrix}$$

ist $P_A(t) = t^2 - 3t + 2 = (t-1)(t-2)$. Setzen wir $A_1 = A - E_2$, $A_2 = A - 2E_2$, so haben wir die linearen Gleichungssysteme

$$A_1 \cdot x = 0 \quad \text{und} \quad A_2 \cdot x = 0 \quad \text{mit Lösungen} \quad \begin{pmatrix} 3 \\ 1 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 2 \\ 1 \end{pmatrix},$$

das sind also Eigenvektoren zu den Eigenwerten 1 und 2. Zur Kontrolle berechnet

man

$$S \cdot A \cdot S^{-1} = \begin{pmatrix} 1 & -2 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} -1 & 6 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

d) Ist

$$\begin{aligned} A &= \begin{pmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix}, \quad \text{so ist} \quad P_A = \begin{vmatrix} -t & -1 & 1 \\ -3 & -2-t & 3 \\ -2 & -2 & 3-t \end{vmatrix} \\ &= -t \cdot \begin{vmatrix} -2-t & 3 \\ -2 & 3-t \end{vmatrix} + 3 \cdot \begin{vmatrix} -1 & 1 \\ -2 & 3-t \end{vmatrix} - 2 \cdot \begin{vmatrix} -1 & 1 \\ -2-t & 3 \end{vmatrix} \\ &= -t(t^2 - t) + 3(t-1) - 2(t-1) = -t^3 + t^2 + t - 1 \\ &= -(t-1)^2(t+1). \end{aligned}$$

Darauf kommen wir in 4.3.4 zurück.

Es darf nicht verschwiegen werden, daß die Beispiele c) und d) sehr künstlich sind, weil man die Nullstellen der charakteristischen Polynome mehr oder weniger schnell erkennen kann. Ist das nicht der Fall, so muß man die Methoden der Numerik verwenden, um die Nullstellen näherungsweise zu berechnen. Wie man ebenfalls in der Numerik lernt, gibt es Verfahren, Eigenwerte und Eigenvektoren in einem Aufwasch gemeinsam zu approximieren. Das hier angegebene Verfahren, zuerst die Eigenwerte und anschließend die Eigenvektoren zu suchen, hilft nur für die Theorie und in besonderen Glücksfällen, die bei den hier behandelten Beispielen inszeniert wurden. Dennoch haben derartige Beispiele einen Sinn: man kann erst einmal ohne Ablenkung durch größeren Rechenaufwand einen Lösungsweg deutlich erkennen.

Aufgaben zu 4.2

1. Berechnen Sie das charakteristische Polynom, die Eigenwerte und Eigenvektoren von

$$\begin{pmatrix} 2 & 2 & 3 \\ 1 & 2 & 1 \\ 2 & -2 & 1 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} -5 & 0 & 7 \\ 6 & 2 & -6 \\ -4 & 0 & 6 \end{pmatrix}.$$

2. Beweisen Sie: Ist $A \in M(2 \times 2; \mathbb{R})$ symmetrisch, so hat A reelle Eigenwerte.

3. Sei V ein endlichdimensionaler Vektorraum und $F \in \text{End}(V)$. Zeigen Sie, dass $P_F(0) \neq 0$ genau dann, wenn F ein Isomorphismus ist.

4. Zeigen Sie, dass die Matrix

$$A = \begin{pmatrix} 0 & \dots & 0 & -\alpha_0 \\ 1 & 0 & \dots & 0 & -\alpha_1 \\ & 1 & \ddots & 0 & \vdots \\ & & \ddots & 0 & \vdots \\ 0 & & & 1 & -\alpha_{n-1} \end{pmatrix}$$

das charakteristische Polynom $P_A(t) = (-1)^n(t^n + \alpha_{n-1}t^{n-1} + \dots + \alpha_1t + \alpha_0)$ besitzt.

5. Sei $A \in M(n \times n; K)$ und $\Phi: M(n \times n; K) \rightarrow M(n \times n; K)$ der Endomorphismus, der durch die Linksmultiplikation mit A gegeben ist, das heißt $\Phi(B) = AB$. Zeigen Sie, dass für die charakteristischen Polynome von A und Φ gilt: $P_\Phi = (P_A)^n$.

4.3 Diagonalisierung

4.3.1. Zunächst halten wir als Ergebnis der vorhergehenden Abschnitte folgendes fest:

Satz. Sei F ein Endomorphismus von V mit $n = \dim V$. Dann gilt:

- 1) Ist F diagonalisierbar, so ist $P_F = \pm(t - \lambda_1) \cdot \dots \cdot (t - \lambda_n)$, d.h. das charakteristische Polynom zerfällt in Linearfaktoren.
- 2) Ist $P_F = \pm(t - \lambda_1) \cdot \dots \cdot (t - \lambda_n)$ mit paarweise verschiedenen $\lambda_1, \dots, \lambda_n$, so ist F diagonalisierbar.

Beweis. 1) ist klar, weil man das charakteristische Polynom mit Hilfe einer Basis aus Eigenvektoren berechnen kann, und

$$P_F = \det \begin{pmatrix} \lambda_1 - t & & 0 \\ & \ddots & \\ 0 & & \lambda_n - t \end{pmatrix} = (\lambda_1 - t) \cdot \dots \cdot (\lambda_n - t).$$

2) folgt aus 4.1.3 und 4.2.3. □

4.3.2. Nach 4.3.1 bleibt also zu klären, wann F im Falle mehrfacher Nullstellen des charakteristischen Polynoms noch diagonalisierbar ist. Zu diesem Zweck fassen wir in P_F gleiche Linearfaktoren zusammen, d.h. wir schreiben

$$P_F = \pm(t - \lambda_1)^{r_1} \cdot \dots \cdot (t - \lambda_k)^{r_k},$$

wobei die $\lambda_1, \dots, \lambda_k$ paarweise verschieden sind, $1 \leq r_i \leq n$ für $i = 1, \dots, k$ und $r_1 + \dots + r_k = n$. Der Exponent r_i ist die Vielfachheit der Nullstelle λ_i von P_F , in der Notation von 1.3.8 ist $r_i = \mu(P_F; \lambda_i)$. Andererseits gehört zu λ_i der Eigenraum $\text{Eig}(F; \lambda_i)$.

Lemma. Ist λ Eigenwert von F , so gilt $1 \leq \dim \text{Eig}(F; \lambda) \leq \mu(P_F; \lambda)$.

Beweis. Sei $v_1, \dots, v_s$ eine Basis von $\text{Eig}(F; \lambda)$. Die Ungleichung $1 \leq s$ ist klar, da λ Eigenwert ist. Zum Beweis der zweiten Ungleichung ergänzen wir die Basis von $\text{Eig}(F; \lambda)$ zu einer Basis

$$\mathcal{A} = (v_1, \dots, v_s, v_{s+1}, \dots, v_n)$$

von V . Dann ist

$$A := M_{\mathcal{A}}(F) = \left(\begin{array}{ccc|ccc} \lambda & & 0 & & & \\ & \ddots & & & * & \\ 0 & & \lambda & & & \\ \hline & & & 0 & & A' \end{array} \right), \quad \left. \vphantom{\begin{pmatrix} \lambda & & 0 \\ & \ddots & \\ 0 & & \lambda \end{pmatrix}} \right\} s\text{-mal}$$

also $P_F = (t - \lambda)^s \cdot P_{A'}$ nach D8 und D9 in 3.1.3, und damit

$$\dim \operatorname{Eig}(F; \lambda) = s \leq \mu(P_F; \lambda).$$

□

4.3.3. Nun können wir zeigen, daß die obigen Ungleichungen ein leicht nachprüfbares Kriterium für die Diagonalisierbarkeit ergeben. Um den Beweis kürzer und die Aussage klarer zu machen, benutzen wir die in 1.6.4 bereitgestellten Tatsachen über direkte Summen.

Theorem. Sei V ein endlichdimensionaler K -Vektorraum und $F \in \operatorname{End}(V)$. Dann sind die folgenden Bedingungen äquivalent:

- i) F ist diagonalisierbar.
- ii) a) Das charakteristische Polynom P_F zerfällt in Linearfaktoren und
b) $\dim \operatorname{Eig}(F; \lambda) = \mu(P_F; \lambda)$ für alle Eigenwerte λ von F .
- iii) Sind $\lambda_1, \dots, \lambda_k$ die paarweise verschiedenen Eigenwerte von F , so ist

$$V = \operatorname{Eig}(F; \lambda_1) \oplus \dots \oplus \operatorname{Eig}(F; \lambda_k).$$

Beweis. i) $\Rightarrow$ ii). Ist F diagonalisierbar, so ordnen wir die Basis von V aus Eigenvektoren entsprechend den verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_k$, d.h. für $i = 1, \dots, k$ betrachten wir eine Basis

$$\left(v_1^{(i)}, \dots, v_{s_i}^{(i)} \right) \quad \text{von} \quad \operatorname{Eig}(F; \lambda_i).$$

Setzen wir $r_i := \mu(P_F, \lambda_i)$, so gilt

$$s_1 + \dots + s_k = n, \quad r_1 + \dots + r_k = n \quad \text{und} \quad s_i \leq r_i.$$

Das geht nur, wenn $s_i = r_i$ für alle i , das ist Bedingung ii) b).

ii) $\Rightarrow$ iii). Sei $W := \operatorname{Eig}(F; \lambda_1) + \dots + \operatorname{Eig}(F; \lambda_k)$. Nach Lemma 4.1.3 und der Definition der direkten Summe in 1.6.4 gilt

$$W = \operatorname{Eig}(F; \lambda_1) \oplus \dots \oplus \operatorname{Eig}(F; \lambda_k).$$

Aus ii) und Bedingung iii) in Satz 1.6.4 folgt $\dim W = r_1 + \dots + r_k = n$, also ist $W = V$.

iii) $\Rightarrow$ i). Für jedes $i \in \{1, \dots, k\}$ sei $\mathcal{B}_i = (v_1^{(i)}, \dots, v_{s_i}^{(i)})$ eine Basis von $\text{Eig}(F; \lambda_i)$. Nach 1.6.4 ist

$$\mathcal{B} := (v_1^{(1)}, \dots, v_{s_1}^{(1)}, \dots, v_1^{(k)}, \dots, v_{s_k}^{(k)})$$

eine Basis von V , und da sie nach Definition aus Eigenvektoren besteht, ist F diagonalisierbar. Insbesondere ist $s_i = r_i$ für alle i und

$$M_{\mathcal{B}}(F) = \left(\begin{array}{ccccccc} \lambda_1 & & & & & & \\ & \ddots & & & & & \\ & & \lambda_1 & & & & \\ & & & \ddots & & & \\ & & & & \ddots & & \\ & 0 & & & & \lambda_k & \\ & & & & & & \ddots \\ & & & & & & & \lambda_k \end{array} \right) \left. \begin{array}{l} \vdots \\ \vdots \\ \vdots \end{array} \right\} \begin{array}{l} r_1\text{-mal} \\ \vdots \\ r_k\text{-mal} \end{array} \quad \square$$

4.3.4. Mit Hilfe der gerade abgeschlossenen theoretischen Überlegungen kann man ein praktisches Verfahren für die Diagonalisierung eines Endomorphismus F von einem endlichdimensionalen K -Vektorraum V angeben:

1. Schritt: Mit Hilfe einer Basis $\mathcal{A}$ von V und der Matrix $A = M_{\mathcal{A}}(F)$ berechne man das charakteristische Polynom. Wenn man eine Zerlegung in Linearfaktoren angeben kann, ist das erste und für die Rechnung entscheidende Hindernis überwunden.

2. Schritt: Für jeden Eigenwert λ von F bestimme man durch Lösung eines linearen Gleichungssystems eine Basis von $\text{Eig}(F; \lambda)$. Damit kann man kontrollieren, ob

$$\dim \text{Eig}(F; \lambda) = \mu(P_F; \lambda)$$

gilt. Genau dann, wenn dies für alle λ der Fall ist, ist F diagonalisierbar. Entsprechend 4.3.3 bilden dann die Basen der Eigenräume zusammen eine Basis $\mathcal{B}$ von V .

Beispiel. Sei $F: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ gegeben durch

$$F(x, y, z) := (-y + z, -3x - 2y + 3z, -2x - 2y + 3z).$$

Bezeichnet $\mathcal{K}$ die kanonische Basis von $\mathbb{R}^3$, so ist

$$A := M_{\mathcal{K}}(F) = \begin{pmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix}$$

und $P_F = -t^3 + t^2 + t - 1 = -(t-1)^2(t+1)$. Also sind $\lambda_1 = 1$ und $\lambda_2 = -1$ die beiden einzigen Eigenwerte von F . $\text{Eig}(F; 1)$ ist der Lösungsraum von

$$\begin{pmatrix} -1 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3-1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

umgeformt $-x_1 - x_2 + x_3 = 0$. Es ist also $\mu(P_F; 1) = 2 = \dim \text{Eig}(F; 1)$, und $((1, 0, 1), (0, 1, 1))$ ist eine Basis von $\text{Eig}(F; 1)$.

Ebenso ist $\text{Eig}(F; -1)$ der Lösungsraum von

$$\begin{pmatrix} 1 & -1 & 1 \\ -3 & -2+1 & 3 \\ -2 & -2 & 3+1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

umgeformt

$$x_1 - x_2 + x_3 = 0, \quad -4x_2 + 6x_3 = 0.$$

Es ist also $\mu(P_F; -1) = 1 = \dim \text{Eig}(F; -1)$, und $(1, 3, 2)$ ist eine Basis von $\text{Eig}(F; -1)$. Somit ist

$$\mathcal{B} = ((1, 0, 1), (0, 1, 1), (1, 3, 2))$$

eine Basis von $\mathbb{R}^3$, bestehend aus Eigenvektoren von F . Für die Transformationsmatrix $S = T_{\mathcal{B}}^{\mathcal{K}}$ gilt

$$S^{-1} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 3 \\ 1 & 1 & 2 \end{pmatrix}, \quad \text{also} \quad S = \frac{1}{2} \begin{pmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ 1 & 1 & -1 \end{pmatrix}.$$

Für

$$D := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

muß somit $D = SAS^{-1}$ gelten, was man zur Kontrolle der Rechnung benutzen kann.

4.3.5. Ist eine Masse an einer Feder aufgehängt und zur Zeit $t = 0$ in senkrechter Richtung in die Position $y(0) = \alpha$ mit Geschwindigkeit $\dot{y}(0) = \beta$ ausgelenkt, so ist die weitere Bewegung bestimmt durch die Differentialgleichung der gedämpften Schwingung

$$\ddot{y} + 2\mu\dot{y} + \omega^2 y = 0, \quad y(0) = \alpha, \quad \dot{y}(0) = \beta.$$

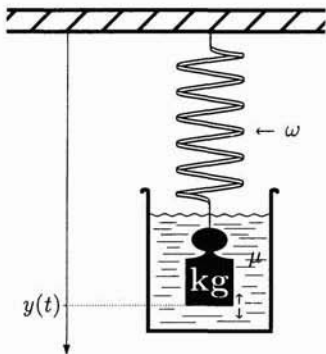


Bild 4.4

Dabei sind $\omega, \mu \in \mathbb{R}_+$ Konstanten, ω ist durch die Feder und μ durch die Reibung bestimmt. Wie üblich macht man daraus mit $y_0 = y$ und $y_1 = \dot{y}$ das lineare System erster Ordnung

$$\begin{aligned} \dot{y}_0 &= y_1, & y_0(0) &= \alpha, \\ \dot{y}_1 &= -\omega^2 y_0 - 2\mu y_1, & y_1(0) &= \beta. \end{aligned} \quad (*)$$

Das führt auf die Matrix mit von t unabhängigen Einträgen

$$A = \begin{pmatrix} 0 & 1 \\ -\omega^2 & -2\mu \end{pmatrix}.$$

Einer Diagonalisierung von A entspricht eine Entkoppelung des Systems (*), und wenn A zu einer oberen Dreiecksmatrix gemacht werden kann, ist das System von unten nach oben explizit lösbar (vgl. [Fo2], §14). Daher betrachten wir das charakteristische Polynom

$$P_A(\lambda) = \lambda^2 + 2\mu\lambda + \omega^2, \quad \text{also} \quad \lambda = -\mu \pm \sqrt{\mu^2 - \omega^2}.$$

Entscheidend für die Art der Bewegung ist die Diskriminante $\mu^2 - \omega^2$. Es sind drei Fälle möglich.

1) $0 \leq \mu < \omega$, d.h. $\mu^2 - \omega^2 < 0$, (schwache Dämpfung)

2) $\mu = \omega$, d.h. $\mu^2 - \omega^2 = 0$, (aperiodischer Grenzfall)

3) $\mu > \omega$, d.h. $\mu^2 - \omega^2 > 0$ (starke Dämpfung)

Im Fall 3) gibt es zwei verschiedene negative Eigenwerte

$$\lambda_1 = -\mu + \sqrt{\mu^2 - \omega^2}, \quad \lambda_2 = -\mu - \sqrt{\mu^2 - \omega^2},$$

also ist A diagonalisierbar.

Im Fall 2) ist $\lambda = -\mu = -\omega$ ein 2-facher Eigenwert, die Matrix

$$A - \lambda E_2 = \begin{pmatrix} \omega & 1 \\ -\omega^2 & -\omega \end{pmatrix}$$

hat den Rang 1, also ist $\dim \text{Eig}(A; -\omega) = 1$, und A ist nicht diagonalisierbar. In 4.4 werden wir sehen, daß A trigonalisierbar ist.

Im Fall 1) gibt es keine reellen Eigenwerte, dafür aber zwei verschiedene komplexe. Also ist A komplex diagonalisierbar, man kann also zunächst komplexe und daraus reelle Lösungen berechnen. Die explizite Durchführung der Lösung von (*) überlassen wir dem Leser (Aufgabe 4 und Aufgabe 4 zu 4.4).

4.3.6. Gelegentlich tritt das Problem auf, daß man zwei oder mehrere Endomorphismen mit einer gemeinsamen Basis diagonalisieren will (man nennt das *simultane Diagonalisierung*). Daß dies nicht immer gehen kann, sieht man am schnellsten mit Matrizen: Gibt es zu $A, B \in M(n \times n; K)$ ein $S \in GL(n; K)$ mit

$$SAS^{-1} = D \quad \text{und} \quad SBS^{-1} = \tilde{D},$$

wobei D und $\tilde{D}$ Diagonalmatrizen bezeichnen, so folgt aus $D \cdot \tilde{D} = \tilde{D} \cdot D$, daß

$$B \cdot A = S^{-1} \tilde{D} S \cdot S^{-1} D S = S^{-1} D S \cdot S^{-1} \tilde{D} S = A \cdot B.$$

Erfreulicherweise ist die Vertauschbarkeit auch hinreichend. Es gilt also:

Satz. Zwei diagonalisierbare Endomorphismen $F, G \in \text{End}_K(V)$ sind genau dann simultan diagonalisierbar, wenn $F \circ G = G \circ F$.

Beweis. Nach Voraussetzung hat man die Zerlegungen in Eigenräume

$$\begin{aligned} V &= \text{Eig}(F; \lambda_1) \oplus \dots \oplus \text{Eig}(F; \lambda_k) \\ &= \text{Eig}(G; \mu_1) \oplus \dots \oplus \text{Eig}(G; \mu_l), \end{aligned}$$

wobei λ_i bzw. μ_j die Eigenwerte von F bzw. G bezeichnen. Wir halten ein λ fest und betrachten

$$W := \text{Eig}(F; \lambda).$$

Ist $w \in W$, so ist $F(G(w)) = G(F(w)) = G(\lambda w) = \lambda G(w)$, also ist W auch G -invariant (vgl. 4.4.1). Wir definieren

$$W_j := W \cap \text{Eig}(G; \mu_j) \quad \text{für } j = 1, \dots, l,$$

und es genügt zu zeigen, daß $W = W_1 \oplus \dots \oplus W_l$, da dies dann für alle Eigenwerte von F gilt. Wegen Lemma 4.1.3 ist nur zu zeigen, daß

$$W = W_1 + \dots + W_l.$$

Zu $w \in W$ gibt es $w_j \in \text{Eig}(G; \mu_j)$, so daß $w = w_1 + \dots + w_l$. Daraus folgt

$$\begin{aligned} F(w) &= F(w_1) + \dots + F(w_l) \\ &= \lambda w = \lambda w_1 + \dots + \lambda w_l, \end{aligned}$$

daher ist wegen der Eindeutigkeit in der direkten Summe

$$F(w_j) = \lambda w_j, \quad \text{also } w_j \in W \text{ und somit } w_j \in W_j. \quad \square$$

Aufgaben zu 4.3

1. Beweisen Sie Teil 2) von Satz 4.3.1 mit Hilfe von Theorem 4.3.3.

2. Sind die folgenden Matrizen diagonalisierbar?

$$\begin{pmatrix} 1 & 2 & 0 & 4 \\ 0 & 2 & 3 & 1 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}, \quad \begin{pmatrix} -5 & 0 & 7 \\ 6 & 2 & -6 \\ -4 & 0 & 6 \end{pmatrix}, \quad \begin{pmatrix} 2 & 1 & 2 \\ -2 & -2 & -6 \\ 1 & 2 & 5 \end{pmatrix}.$$

3. Für welche $a, b \in \mathbb{R}$ ist die Matrix

$$\begin{pmatrix} -3 & 0 & 0 \\ 2a & b & a \\ 10 & 0 & 2 \end{pmatrix}$$

diagonalisierbar?

4. Wir betrachten das Differentialgleichungssystem mit Anfangswertbedingung

$$\dot{y} = A \cdot y, \quad y_0(0) = \alpha, \quad y_1(0) = \beta \quad (*)$$

für die gedämpfte Schwingung (siehe 4.3.5), wobei

$$A = \begin{pmatrix} 0 & 1 \\ -\omega^2 & -2\mu \end{pmatrix}.$$

- Im Fall $\mu > \omega$ ist A (reell) diagonalisierbar. Bestimmen Sie eine Basis des $\mathbb{R}^2$ aus Eigenvektoren von A und geben Sie eine Basis des Lösungsraums von $\dot{y} = A \cdot y$ an (vgl. Aufgabe 3 zu 4.1). Wie sieht die Lösung von $(*)$ aus?
- Im Fall $\mu < \omega$ ist $A \in M(2 \times 2; \mathbb{C})$ komplex diagonalisierbar. Bestimmen Sie die Eigenwerte von A und geben Sie eine Basis des $\mathbb{C}^2$ aus Eigenvektoren von A an. Ist

$\lambda \in \mathbb{C}$ Eigenwert von A zum Eigenvektor $v \in \mathbb{C}^2$, so ist $re^{i\lambda t} \cdot v$, im $e^{i\lambda t} \cdot v$ eine Basis des Lösungsraums von $\dot{y} = A \cdot y$ ([Fo2], §13). Bestimmen Sie auch in diesem Fall die Lösung von (*).

5. Diagonalisieren Sie die Matrizen

$$A = \begin{pmatrix} -5 & 1 & 6 & 6 \\ -12 & 2 & 12 & 12 \\ 1 & 1 & 0 & -2 \\ -4 & 0 & 4 & 6 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 0 & -1 & -4 \\ -3 & 1 & 3 & 0 \\ 2 & 0 & -1 & -2 \\ 1 & 0 & -1 & -3 \end{pmatrix}$$

aus $M(4 \times 4; \mathbb{R})$ simultan, d. h. bestimmen Sie eine Matrix $S \in GL(4; \mathbb{R})$, so dass SAS^{-1} und SBS^{-1} Diagonalmatrizen sind.

6. Seien $A, B \in M(n \times n; K)$ mit $AB = BA$ und alle Eigenwerte von A und B seien einfach. Dann gilt: A und B haben die gleichen Eigenvektoren.

7. Zeigen Sie, dass es für $\lambda \in K$ und natürliche Zahlen μ, n mit $1 \leq \mu \leq n$ stets eine Matrix $A \in M(n \times n; K)$ gibt mit $\mu(P_A; \lambda) = \mu$ und $\dim \text{Eig}(A; \lambda) = 1$.

8. Es sei K ein Körper mit $\text{char } K \neq 2$. Zeigen Sie, dass die Lösungen der Gleichung $A^2 = E_2$ in $M(2 \times 2; K)$ genau von der folgenden Gestalt sind:

$$A = E_2, \quad A = -E_2 \quad \text{oder} \quad A = SDS^{-1} \quad \text{mit} \quad D = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{und} \quad S \in GL(2; K).$$

9. Sei F ein diagonalisierbarer Endomorphismus eines endlichdimensionalen $\mathbb{R}$ -Vektorraums, für den gilt: Sind v und w Eigenvektoren von F , so ist $v + w$ ein Eigenvektor von F oder $v + w = 0$. Zeigen Sie, dass es ein $\lambda \in \mathbb{R}$ gibt mit $F = \lambda \cdot \text{id}$.

10. Seien $A, B \in M(3 \times 3; \mathbb{R})$ zwei Matrizen mit den charakteristischen Polynomen $P_A(t) = -t^3 + 2t^2 - t$ und $P_B(t) = -t^3 + 7t^2 - 9t + 3$. Zeigen Sie, dass der Kern von AB die Dimension 1 hat.

4.4 Trigonalisierung*

Wie wir in 4.3.3 gesehen hatten, gibt es zwei Bedingungen für die Diagonalisierbarkeit eines Endomorphismus:

- Das charakteristische Polynom muß in Linearfaktoren zerfallen, und
- die Dimensionen der Eigenräume müssen gleich den Vielfachheiten der Eigenwerte sein.

In diesem Abschnitt zeigen wir, daß ein Endomorphismus, der nur Bedingung a) erfüllt, wenigstens durch eine obere Dreiecksmatrix beschrieben werden kann.

4.4.1. Um die geometrischen Eigenschaften komplizierter Endomorphismen besser verstehen zu können, sind ein paar neue Begriffe nützlich.

Definition. Sei $F: V \rightarrow V$ ein Endomorphismus und $W \subset V$ ein Untervektorraum. W heißt *F-invariant*, wenn $F(W) \subset W$.

Es gibt stets die trivialen, aber nutzlosen Beispiele $W = \{0\}$ und $W = V$. Je mehr weitere invariante Unterräume existieren, desto übersichtlicher ist ein Endomorphismus. Ist etwa F diagonalisierbar und $(v_1, \dots, v_n)$ eine Basis aus Eigenvektoren, so ist

$$V = W_1 \oplus \dots \oplus W_n \quad \text{mit} \quad W_i = K \cdot v_i$$

eine Zerlegung in eindimensionale invariante Unterräume. Im Falle mehrfacher Eigenwerte (4.3.3) hat man eine weitere Zerlegung

$$V = \text{Eig}(F; \lambda_1) \oplus \dots \oplus \text{Eig}(F; \lambda_k)$$

in invariante Unterräume der Dimensionen $r_1, \dots, r_k$.

Die Beziehung zwischen invarianten Unterräumen und charakteristischem Polynom ist in einer Richtung ganz klar.

Bemerkung. Ist $W \subset V$ ein F -invarianter Unterraum, so ist $P_{F|W}$ ein Teiler von P_F .

Beweis. Wir ergänzen eine Basis $\mathcal{B}'$ von W zu einer Basis $\mathcal{B}$ von V . Dann ist

$$M_{\mathcal{B}}(F) = \left(\begin{array}{c|c} M_{\mathcal{B}'}(F|W) & * \\ \hline 0 & A \end{array} \right),$$

also $P_F = P_{F|W} \cdot P_A$ nach D9 in 3.1.3. □

Beispiel. In $\mathbb{Q}[t]$ hat das Polynom $t^n - 2$ für $n \geq 2$ keinen Teiler kleineren Grades (Aufgabe 1). Also hat der durch

$$A = \begin{pmatrix} 0 & \cdots & 0 & 2 \\ 1 & & 0 & 0 \\ & \ddots & & \vdots \\ 0 & & 1 & 0 \end{pmatrix}$$

beschriebene Endomorphismus des $\mathbb{Q}^n$ nur die trivialen invarianten Unterräume.

4.4.2. Sei $A \in M(n \times n; K)$ eine obere Dreiecksmatrix, d.h. $a_{ij} = 0$ für $i > j$. Unter dem Endomorphismus A des K^n sind die Untervektorräume

$$W_r := \text{span}(e_1, \dots, e_r)$$

für $1 \leq r \leq n$ invariant. Das motiviert die

Definition. Unter einer Fahne (V_r) in einem n -dimensionalen Vektorraum V versteht man eine Kette

$$\{0\} = V_0 \subset V_1 \subset \dots \subset V_n = V$$

von Untervektorräumen mit $\dim V_r = r$. Ist $F \in \text{End}(V)$, so heißt die Fahne F -invariant, wenn

$$F(V_r) \subset V_r \quad \text{für alle } r \in \{0, \dots, n\}.$$

Man kann sich V_0 als Befestigungspunkt, V_1 als Stange und V_2 als Tuch der Fahne vorstellen.

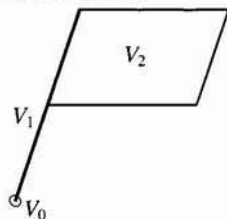


Bild 4.5

Mit Hilfe jeder Basis eines Vektorraumes kann man viele Fahnen konstruieren, aber aus der Existenz einer F -invarianten Fahne folgt insbesondere wegen $F(V_1) \subset V_1$, daß es einen Eigenvektor geben muß.

Bemerkung. Für $F \in \text{End}(V)$ sind folgende Bedingungen gleichwertig:

- i) Es gibt eine F -invariante Fahne in V .
- ii) Es gibt eine Basis $\mathcal{B}$ von V , so daß $M_{\mathcal{B}}(F)$ obere Dreiecksmatrix ist.

Ist das der Fall, so heißt F trigonalisierbar.

Beweis. Die Beziehung zwischen Fahne (V_r) und Basis $\mathcal{B} = (v_1, \dots, v_n)$ ist geregelt durch

$$V_r = \text{span}(v_1, \dots, v_r) \quad \text{für alle } r. \quad \square$$

Übersetzt in den Matrizenkalkül ergibt das die

Definition. $A \in M(n \times n; K)$ heißt trigonalisierbar, wenn es ein $S \in \text{GL}(n; K)$ gibt, so daß SAS^{-1} obere Dreiecksmatrix ist.

4.4.3. Ergebnis dieses Abschnittes ist der

Trigonalisierungssatz. Für einen Endomorphismus F eines n -dimensionalen K -Vektorraumes sind folgende Bedingungen äquivalent:

- i) F ist trigonalisierbar.
- ii) Das charakteristische Polynom P_F zerfällt in Linearfaktoren, d.h.

$$P_F = \pm(t - \lambda_1) \cdot \dots \cdot (t - \lambda_n) \quad \text{mit } \lambda_1, \dots, \lambda_n \in K.$$

Mit Hilfe des Fundamentalsatzes der Algebra (1.3.9) folgt das

Korollar. Jeder Endomorphismus eines endlich-dimensionalen komplexen Vektorraumes ist trigonalisierbar. $\square$

Eine weitere wichtige Anwendung betrifft die Lösung linearer Systeme von Differentialgleichungen ([Fo2], §14).

Beweis des Satzes. i) $\Rightarrow$ ii) ist klar, denn ist $A = (a_{ij}) = M_{\mathcal{B}}(F)$ eine obere Dreiecksmatrix, so folgt aus D8 in 3.1.3, daß

$$P_F = (a_{11} - t) \cdot \dots \cdot (a_{nn} - t).$$

ii) $\Rightarrow$ i) wird durch Induktion über $n = \dim V$ bewiesen. Für $n = 0, 1$ ist nichts zu zeigen. Ist $n \geq 2$, so wähle man einen Eigenvektor v_1 zu λ_1 und ergänze ihn zu einer Basis $\mathcal{B} = (v_1, w_2, \dots, w_n)$ von V . Dann ist

$$V = V_1 \oplus W \quad \text{mit } V_1 := \text{span}(v_1) \text{ und } W := \text{span}(w_2, \dots, w_n),$$

und

$$M_B(F) = \begin{pmatrix} \lambda_1 & a_{12} & \cdots & a_{1n} \\ 0 & \boxed{} & & \\ \vdots & & & \\ 0 & & & \end{pmatrix}.$$

V_1 ist F -invariant, W im allgemeinen nicht; das Hindernis dagegen sind die Einträge $a_{12}, \dots, a_{1n}$. Der Ausweg ist folgender: Durch

$$H(w_j) = a_{1j}v_1 \quad \text{und} \quad G(w_j) = a_{2j}w_2 + \dots + a_{nj}w_n$$

sind lineare Abbildungen $H: W \rightarrow V_1$ und $G: W \rightarrow W$ erklärt mit

$$F(w) = H(w) + G(w) \quad \text{für alle } w \in W.$$

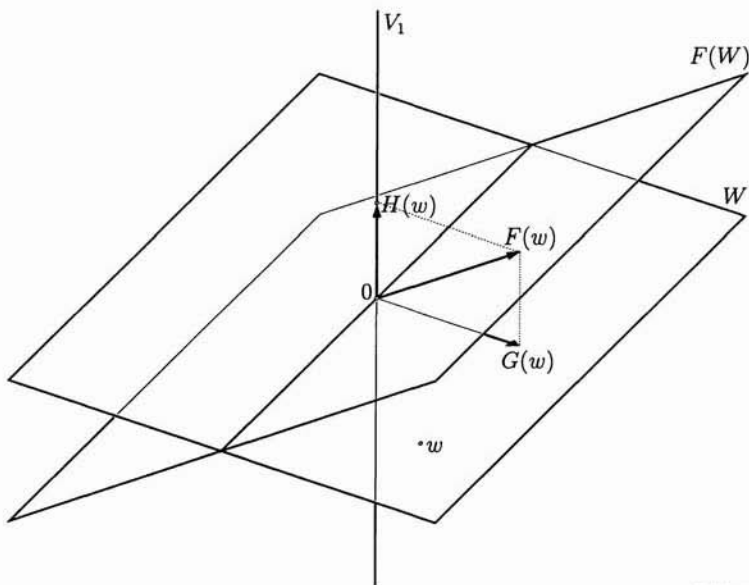


Bild 4.6

Für die charakteristischen Polynome gilt

$$P_F = (\lambda_1 - t) \cdot P_G, \quad \text{also} \quad P_G = (\lambda_2 - t) \cdot \dots \cdot (\lambda_n - t).$$

Die Induktionsvoraussetzung angewandt auf W ergibt eine G -invariante Fahne

$$\{0\} = W_0 \subset \dots \subset W_{n-1} = W,$$

und wir behaupten, daß durch $V_r := V_1 + W_{r-1}$ eine F -invariante Fahne gegeben ist. Das folgt aus

$$F(\mu v_1 + w) = \lambda_1 \mu v_1 + H(w) + G(w)$$

wegen $H(w) \in V_1$ und $G(w) \in W_{r-1}$ für $w \in W_{r-1}$. $\square$

4.4.4. Obwohl obiger Satz sehr nützlich ist, muß man doch bemerken, daß die Beweismethode nicht genügend sorgfältig auf die Geometrie der Abbildung achtet.

Beispiel. Ist $F \in \text{End}(\mathbb{R}^2)$ gegeben durch die Diagonalmatrix

$$D = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix},$$

so kann man mit $v_1 = e_1$ und $\lambda_1 = 1$ beginnen. Die Ergänzung durch w_2 ist beliebig: wählt man etwa $w_2 = (-1, 1)$, so wird

$$A = M_B(F) = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix},$$

und damit ist die schöne Diagonalmatrix versaut. Auch vom algebraischen Standpunkt hat A eine unangenehme Eigenschaft. Es ist

$$A = D + N \quad \text{mit} \quad N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Dabei ist N nilpotent (vgl. 4.5.7) mit $N^2 = 0$, aber $D \cdot N \neq N \cdot D$.

In 4.6 wird gezeigt, wie man den nilpotenten Anteil N auf eine der Abbildung angepaßte Normalform bringen kann.

4.4.5. Wir geben nun ein *Rechenverfahren zur Trigonalisierung eines Endomorphismus* an. Es genügt dazu, eine Matrix $A \in M(n \times n; K)$ zu betrachten, für die

$$P_A = \pm(t - \lambda_1) \cdot \dots \cdot (t - \lambda_n) \quad \text{mit} \quad \lambda_1, \dots, \lambda_n \in K.$$

Gesucht ist eine Matrix $S \in \text{GL}(n; K)$, so daß

$$D := S \cdot A \cdot S^{-1}$$

eine obere Dreiecksmatrix ist. Der Beweis des Trigonalisierungssatzes 4.4.3 ergibt das folgende Iterationsverfahren.

1. Schritt: Wir betrachten $W_1 = K^n$ mit der Basis $\mathcal{B}_1 = \mathcal{K}$ und dem Endomorphismus $A_1 = A$. Zu λ_1 berechnet man einen Eigenvektor $v_1 \in K^n$. Nach dem Austauschlemma 1.5.4 bestimmt man ein $j_1 \in \{1, \dots, n\}$, so daß

$$\mathcal{B}_2 := (v_1, e_1, \dots, \widehat{e}_{j_1}, \dots, e_n)$$

wieder eine Basis von K^n ist. Das Zeichen $\widehat{}$ bedeutet dabei, daß e_{j_1} ausgelassen wird. Wir betrachten die Transformationsmatrix

$$S_1^{-1} := T_{\mathcal{B}_1}^{\mathcal{B}_2}$$

mit der Basis $\mathcal{B}_2$ als Spalten. Dann ist

$$A_2 := S_1 \cdot A \cdot S_1^{-1} = \begin{pmatrix} \lambda_1 & * & \dots & * \\ 0 & \boxed{A'_2} \\ \vdots & & & \\ 0 & & & \end{pmatrix}.$$

2. Schritt: Wir betrachten W_2 mit der Basis $\mathcal{B}'_2 = (e_1, \dots, \widehat{e}_{j_1}, \dots, e_n)$ und dem Endomorphismus A'_2 . Es ist

$$P_{A'_2} = \pm(t - \lambda_2) \cdot \dots \cdot (t - \lambda_n).$$

Zu λ_2 berechnet man einen Eigenvektor $v_2 \in W_2$, und man wählt ein $j_2 \neq j_1$, so daß

$$\mathcal{B}'_3 = (v_2, e_1, \dots, \widehat{e}_{j_1}, \dots, \widehat{e}_{j_2}, \dots, e_n)$$

eine Basis von W_2 , also

$$\mathcal{B}_3 = (v_1, v_2, e_1, \dots, \widehat{e}_{j_1}, \dots, \widehat{e}_{j_2}, \dots, e_n)$$

eine Basis von K^n ist. Mit der Transformationsmatrix $S_2^{-1} = T_{\mathcal{B}_1}^{\mathcal{B}_3}$ erhält man

$$A_3 = S_2 \cdot A \cdot S_2^{-1} = \begin{pmatrix} \lambda_1 & * & \dots & \dots & * \\ 0 & \lambda_2 & * & \dots & * \\ \vdots & 0 & \boxed{A'_3} \\ \vdots & \vdots & & & \\ 0 & 0 & & & \end{pmatrix}.$$

Bei der Berechnung von S_2 kann man benutzen, daß

$$T_{\mathcal{B}_1}^{\mathcal{B}_3} = T_{\mathcal{B}_1}^{\mathcal{B}_2} \cdot T_{\mathcal{B}_2}^{\mathcal{B}_3} \quad \text{und} \quad T_{\mathcal{B}_2}^{\mathcal{B}_3} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \boxed{T_{\mathcal{B}'_2}^{\mathcal{B}'_3}} \\ \vdots & & & \\ 0 & & & \end{pmatrix}.$$

Spätestens im $(n-1)$ -ten Schritt erhält man eine obere Dreiecksmatrix A_n , denn A'_n ist eine (1×1) -Matrix. Also ist

$$D := A_n = S_{n-1} \cdot A \cdot S_{n-1}^{-1}$$

eine obere Dreiecksmatrix.

Vorsicht! Die im ersten Schritt berechnete erste Zeile von A_2 kann sich in jedem der folgenden Schritte wieder ändern, analog für die anderen Zeilen. Das ist nicht angenehm.

Beispiel. Sei $K = \mathbb{R}$, $n = 3$ und

$$A = \begin{pmatrix} 3 & 4 & 3 \\ -1 & 0 & -1 \\ 1 & 2 & 3 \end{pmatrix}.$$

Dann ist $P_A = -(t-2)^3$, also $\lambda_1 = \lambda_2 = \lambda_3 = 2$. Wegen $\text{rang}(A - 2E_3) = 2$ ist

$$\dim \text{Eig}(A; 2) = 1 < 3 = \mu(P_A; 2),$$

also ist A nicht diagonalisierbar. Die Iteration läuft wie folgt:

1. Schritt: $v_1 = (1, -1, 1)$, $j_1 = 1$, $B_2 = (v_1, e_2, e_3)$,

$$S_1^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \quad S_1 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix},$$

$$A_2 = S_1 \cdot A \cdot S_1^{-1} = \begin{pmatrix} 2 & 4 & 3 \\ 0 & \boxed{4} & 2 \\ 0 & -2 & 0 \end{pmatrix}.$$

2. Schritt: $W_2 = \text{span}(e_2, e_3)$, $v_2 = e_2 - e_3$, $j_2 = 2$,

$$S_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 2 & 1 & 3 \\ 0 & 2 & 2 \\ 0 & 0 & \boxed{2} \end{pmatrix} = D.$$

Also ist $S = S_2$.

Aufgaben zu 4.4

1. Zeigen Sie, dass das Polynom $t^n - 2 \in \mathbb{Q}[t]$ für $n \geq 2$ keinen Teiler $P \in \mathbb{Q}[t]$ mit $1 \leq \deg P \leq n-1$ besitzt.

2. Trigonalisieren Sie mit dem Verfahren aus 4.4.5 die Matrizen

$$\begin{pmatrix} 3 & 0 & -2 \\ -2 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} -1 & -3 & -4 \\ -1 & 0 & 3 \\ 1 & -2 & -5 \end{pmatrix}.$$

3. Zeigen Sie mit Induktion nach $n = \dim V$: Ist V ein endlichdimensionaler K -Vektorraum und $F: V \rightarrow V$ ein nilpotenter Endomorphismus, so existiert eine Basis $\mathcal{B}$ von V mit

$$M_{\mathcal{B}}(F) = \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix},$$

und es gilt $P_F(t) = \pm t^n$.

4. (Fortsetzung von Aufgabe 4 in 4.3.) Zeigen Sie, dass die Matrix

$$A = \begin{pmatrix} 0 & 1 \\ -\omega^2 & -2\mu \end{pmatrix}$$

im Fall $\mu = \omega$ trigonalisierbar ist, und bestimmen Sie eine Matrix $S \in GL(2; \mathbb{R})$, so dass $B = SAS^{-1}$ obere Dreiecksmatrix ist. Das System $\dot{y} = A \cdot y$ geht somit durch die Substitution $z = Sy$ über in $\dot{z} = B \cdot z$, und es reicht, das (einfachere) System $\dot{z} = B \cdot z$ zu lösen. Bestimmen Sie auf diese Weise eine Basis des Lösungsraums von $\dot{y} = A \cdot y$ und lösen (*) in 4.3.5 auch im aperiodischen Grenzfall.

4.5 Potenzen eines Endomorphismus*

Zur genauen Beschreibung eines Endomorphismus – insbesondere bei der Suche nach invarianten Unterräumen – muß man systematisch das Ergebnis der wiederholten Anwendung der Abbildung betrachten. Das legt die Weiterentwicklung der algebraischen Hilfsmittel nahe. Das wichtigste Ergebnis ist der Satz von CAYLEY-HAMILTON.

4.5.1. Die Eigenwertbedingung $F(v) = \lambda v$ kann man auch komplizierter ausdrücken: in dem linearen Polynom

$$L(t) = t - \lambda$$

setzt man F statt t und id_V statt 1 , das ergibt

$$L(F) := F - \lambda \text{id}_V,$$

und es ist $\text{Eig}(F; \lambda) = \text{Ker } L(F)$. Was hat man damit gewonnen? Ist

$$P(t) = \alpha_r t^r + \dots + \alpha_1 t + \alpha_0 \in K[t]$$

ein beliebiges Polynom, so kann man analog

$$P(F) := \alpha_r F^r + \dots + \alpha_1 F + \alpha_0 \text{id}_V \in \text{End}(V)$$

setzen. Ist $0 \neq v \in \text{Ker } P(F)$, so gilt

$$\alpha_r F^r(v) + \dots + \alpha_1 F(v) + \alpha_0 v = 0,$$

d.h. die Vektoren $v, F(v), \dots, F^r(v)$ sind linear abhängig, wenn $P \neq 0$. Ist $\alpha_r \neq 0$, so folgt, daß

$$W = \text{span}(v, F(v), \dots, F^{r-1}(v))$$

ein F -invarianter Unterraum von V mit $1 \leq \dim W \leq r$ ist. Ist $P(F) = 0$, so kann man diese Konstruktion mit jedem beliebigen v starten.

Beispiel. Sei F diagonalisierbar mit den paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_k$ und

$$Q(t) = (t - \lambda_1) \cdot \dots \cdot (t - \lambda_k).$$

Jedes $v \in V$ hat nach 4.3.3 eine Darstellung $v = v_1 + \dots + v_k$ mit $v_i \in \text{Eig}(F; \lambda_i)$. Benutzt man, daß für $\lambda \neq \mu$ und $0 \neq v \in \text{Eig}(F; \lambda)$

$$0 \neq (F - \mu \text{id}_V)(v) = (\lambda - \mu)v \in \text{Eig}(F; \lambda),$$

so folgt sofort $Q(F) = 0$ und $Q'(F) \neq 0$ für jeden echten Teiler Q' von Q . Weiter folgt $P_F(F) = 0$, wenn

$$P_F = (t - \lambda_1)^{r_1} \cdot \dots \cdot (t - \lambda_k)^{r_k}$$

das charakteristische Polynom bezeichnet.

4.5.2. Das Einsetzen eines Endomorphismus in Polynome ist beschrieben durch die Abbildung

$$\Phi_F: K[t] \rightarrow \text{End}(V), \quad P(t) \mapsto P(F).$$

Dies ist ein Homomorphismus von Ringen und auch von K -Vektorräumen. Das Bild

$$K[F] = \{P(F): P(t) \in K[t]\} \subset \text{End } V$$

ist ein kommutativer Unterring des (nicht kommutativen) Ringes $\text{End } V$, und der Kern

$$\mathcal{I}_F := \{P(t) \in K[t]: P(F) = 0\} \subset K[t]$$

heißt *Ideal von F* . Wegen $\dim K[t] = \infty$ und $\dim \text{End}(V) < \infty$ gibt es in $\mathcal{I}_F$ ein normiertes Polynom vom Grad $\leq n^2$ (vgl. Aufgabe 3). Ist F diagonalisierbar, so hatten wir in 4.5.1 sogar ein Polynom $Q \in \mathcal{I}_F$ vom Grad $k \leq n$ angegeben.

4.5.3. Wir zeigen nun, daß sich die Schranke der Grade in $\mathcal{I}_F$ für alle Endomorphismen F auf $n = \dim V$ verbessern läßt. Es gilt sogar noch mehr:

Satz von CAYLEY-HAMILTON. *Sei V ein endlichdimensionaler K -Vektorraum, $F \in \text{End}(V)$ und $P_F \in K[t]$ sein charakteristisches Polynom. Dann ist*

$$P_F(F) = 0 \in \text{End}(V).$$

Insbesondere gilt für jede Matrix $A \in M(n \times n; K)$

$$P_A(A) = 0 \in M(n \times n; K).$$

Vorsicht! Es ist entscheidend, daß die Null ein Endomorphismus oder eine Matrix, nicht die Null in K ist. Daher ist der naheliegende „Beweis“

$$P_A(A) = \det(A - A \cdot E_n) = \det(A - A) = \det(0) = 0$$

nicht richtig. Am Ende steht die $0 \in K$, der Fehler steht schon in der ersten Gleichung, weil ganz naiv bei einem Eintrag der Determinante $t = A$ gesetzt wurde, dann aber mit Einträgen aus K weitergerechnet wird.

Mit einem Trick von SERGE LANG [L] kann man diesen falschen Beweis zu einem richtigen aufmöbeln. Das Problem dabei ist folgendes: die Einträge einer Determinante müssen aus einem kommutativen Ring stammen (3.2.8), die Ringe quadratischer Matrizen sind i.a. nicht kommutativ. Der Ausweg ist, daß man sich auf den kommutativen Unterring (4.5.2)

$$K[A] \subset M(n \times n; K)$$

beschränken kann. Nach diesen Vorbemerkungen zum

Beweis. Man kann den Endomorphismus F von V mit Hilfe einer Basis durch eine Matrix $A \in M(n \times n; K)$ beschreiben. Also genügt es, den Satz dafür zu beweisen.

Für $n = 1$ ist die Aussage trivial, für $n = 2$ kann man sie noch durch direkte Rechnung mit Matrizen beweisen. Im allgemeinen Fall benutzen wir, wie oben erläutert, Matrizen mit Einträgen aus den kommutativen Ringen $K[t]$ und $K[A]$. Zunächst sei

$$B(t) := {}^t(A - tE_n) \in M(n \times n; K[t]).$$

Die Einträge von $B(t)$ außerhalb der Diagonalen sind Elemente von K , und in der Diagonalen stehen lineare Polynome, es ist

$$\det B(t) = P_A(t) \in K[t]. \quad (*)$$

Nun wird die Unbestimmte t durch die Matrix A und jeder Eintrag a_{ij} durch $a_{ij}E_n$ ersetzt. Das ergibt

$$B(A) = \begin{pmatrix} a_{11}E_n - A & a_{21}E_n & \cdots & a_{n1}E_n \\ \vdots & \vdots & & \vdots \\ a_{1n}E_n & a_{2n}E_n & \cdots & a_{nn}E_n - A \end{pmatrix} \in M(n \times n; K[A]).$$

Diese Matrix kann mit einer Spalte multipliziert werden, deren Einträge Spaltenvektoren des K^n sind. Insbesondere ist

$$B(A) \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} a_{11}e_1 - Ae_1 + a_{21}e_2 + \dots + a_{n1}e_n \\ \vdots \\ a_{1n}e_1 + a_{2n}e_2 + \dots + a_{nn}e_n - Ae_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}. \quad (**)$$

Nun sei $B^z(t) \in M(n \times n; K[t])$ die in 3.3.1 erklärte, zu $B(t)$ komplementäre Matrix. Ihre Einträge sind entsprechend der Definition Polynome vom Grad $\leq n - 1$, und es gilt wegen $(*)$

$$B^z(t) \cdot B(t) = P_A(t) \cdot E_n.$$

Setzt man auf beiden Seiten A für t ein, so folgt mit Hilfe von $(**)$

$$\begin{pmatrix} P_A(A) & & 0 \\ & \ddots & \\ 0 & & P_A(A) \end{pmatrix} \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} P_A(A)e_1 \\ \vdots \\ P_A(A)e_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Da $P_A(A)$ auf einer Basis die Werte $0 \in K^n$ hat, ist $P_A(A) = 0$ als Matrix. $\square$

Dieser Beweis ist extrem kurz und elegant, aber für einen Geometer mag er wie algebraischer Hokusokus aussehen. Daher noch ein anderer

Beweis, bei dem man vielleicht besser ein Gefühl dafür bekommen kann, warum die Kombination $P_F(F)$ von Potenzen von F alles in V zu Null macht.

Zunächst betrachten wir den Spezialfall $K = \mathbb{C}$. Nach 4.4.3 ist F trigonalisierbar, d.h. es gibt eine F -invariante Fahne

$$\{0\} = V_0 \subset V_1 \subset \dots \subset V_n = V.$$

Ist $\mathcal{B} = (v_1, \dots, v_n)$ eine Basis von V mit $V_i = \text{span}(v_1, \dots, v_i)$ für $i = 1, \dots, n$, so ist

$$M_{\mathcal{B}}(F) = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

wobei $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ die Eigenwerte von F sind. Da

$$P_F = (\lambda_1 - t) \cdot \dots \cdot (\lambda_n - t),$$

kann man $P_F(F)$ schrittweise aufbauen: Ist

$$G_i := (\lambda_1 \text{id}_V - F) \circ (\lambda_2 \text{id}_V - F) \circ \dots \circ (\lambda_i \text{id}_V - F) \in \text{End}(V)$$

für $i = 1, \dots, n$, so ist $G_n = P_F(F)$. Es genügt also durch Induktion zu zeigen, daß

$$G_i(V_i) = \{0\} \quad \text{für } i = 1, \dots, n.$$

Der Fall $i = 1$ ist klar, da $F(v_1) = \lambda_1 v_1$. Sei nun $i \geq 2$ und $v \in V_i$. Dann gibt es ein $w \in V_{i-1}$ und ein $\mu \in \mathbb{C}$, so daß $v = w + \mu v_i$. Da

$$\lambda_i w - F(w) \in V_{i-1} \quad \text{und} \quad \lambda_i v_i - F(v_i) \in V_{i-1},$$

ist nach Induktionsvoraussetzung

$$G_i(w) = (G_{i-1} \circ (\lambda_i \text{id}_V - F))(w) = G_{i-1}(\lambda_i w - F(w)) = 0$$

und

$$G_i(v_i) = (G_{i-1} \circ (\lambda_i \text{id}_V - F))(v_i) = G_{i-1}(\lambda_i v_i - F(v_i)) = 0.$$

Also ist $G_i(v) = G_i(w) + \mu G_i(v_i) = 0$.

Der Leser möchte sich überlegen, wie viel einfacher der Beweis wird, wenn F diagonalisierbar ist, d.h. wenn $F(v_i) = \lambda_i v_i$ (vgl. Aufgabe 5).

Den Fall $K = \mathbb{R}$ kann man durch *Komplexifizierung* auf den obigen Fall zurückführen, obwohl F nicht reell trigonalisierbar sein muß. Am einfachsten geht das mit Hilfe einer Matrix

$$A = M_{\mathcal{B}}(F),$$

wobei $\mathcal{B}$ irgend eine Basis von V ist (die „professionelle“ Methode der Komplexifizierung ist in 6.3.4 b) ausgeführt). A beschreibt einen Endomorphismus

$$A: \mathbb{C}^n \rightarrow \mathbb{C}^n,$$

also ist $P_A(A) = 0 \in \text{End } \mathbb{C}^n$. Nach 2.6.4 ist

$$M_B(P_F(F)) = P_F(M_B(F)) = P_A(A),$$

also auch $P_F(F) = 0 \in \text{End } V$.

Diese Methode ist nicht auf $\mathbb{R}$ und $\mathbb{C}$ beschränkt, denn zu jedem Körper K und jedem Polynom $f \in K[t]$ gibt es einen Oberkörper $K' \supset K$, so daß $f \in K'[t]$ in Linearfaktoren zerfällt: etwa den *Zerfällungskörper* von f oder den *algebraischen Abschluß* von K (siehe z.B. [F-S]).

4.5.4. Als Beispiel für eine Anwendung des Satzes von CAYLEY-HAMILTON wollen wir in Analogie zu Korollar 4.4.3 zeigen, daß sich jeder Endomorphismus eines reellen Vektorraumes beinahe trigonalisieren läßt.

Theorem. Ist F Endomorphismus eines reellen Vektorraumes V , so gibt es eine Basis B derart, daß

$$M_B(F) = \begin{pmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_r & & & \\ & & & \boxed{B_1} & & \\ & 0 & & & \ddots & \\ & & & & & \boxed{B_m} \end{pmatrix}$$

mit $\lambda_1, \dots, \lambda_r \in \mathbb{R}$ und $B_i = \begin{pmatrix} 0 & -c_i \\ 1 & -b_i \end{pmatrix}$ mit $b_i, c_i \in \mathbb{R}$ und $b_i^2 - 4c_i < 0$ für $i = 1, \dots, m$.

Lemma. Jeder Endomorphismus F eines reellen Vektorraumes V mit $\dim V \geq 1$ hat einen invarianten Unterraum W mit $1 \leq \dim W \leq 2$.

Beweis des Lemmas. Nach 1.3.10 gibt es eine Zerlegung

$$P_F = \pm(t - \lambda_1) \cdot \dots \cdot (t - \lambda_r) \cdot Q_1(t) \cdot \dots \cdot Q_m(t)$$

mit $Q_i(t) = t^2 + b_i t + c_i$ und $b_i^2 - 4c_i < 0$. Ist $r \geq 1$, so gibt es einen Eigenvektor v_1 zu λ_1 , und $W = \mathbb{R} \cdot v_1$ ist invariant.

Im Fall $r = 0$ starten wir mit einem beliebigen $0 \neq v \in V$. Nach CAYLEY-HAMILTON ist

$$P_F(F)(v) = (Q_1(F) \circ \dots \circ Q_m(F))(v) = 0.$$

Also gibt es ein eindeutig bestimmtes $i \in \{1, \dots, m\}$ mit

$$w := (Q_{i+1}(F) \circ \dots \circ Q_m(F))(v) \neq 0 \quad \text{und} \quad Q_i(F)(w) = 0,$$

wobei $w := v$ gesetzt wird, falls $i = m$, d.h. $Q_m(F)(v) = 0$. Wir setzen

$$W := \text{span}(w, F(w)).$$

Aus $Q_i(F)(w) = 0$ folgt

$$F(F(w)) + b_i F(w) + c_i w = 0,$$

also ist $F(W) \subset W$. Da F keinen Eigenwert hat, ist $\mathcal{B} = (w, F(w))$ eine Basis von W , und es gilt

$$M_{\mathcal{B}}(F|_W) = \begin{pmatrix} 0 & -c_i \\ 1 & -b_i \end{pmatrix},$$

also $P_{F|_W} = Q_i(t)$. □

Den *Beweis des Theorems* führt man schrittweise nach dem Muster des Induktionsbeweises von Satz 4.4.3. In der ersten Etappe baut man die Eigenräume ab, in der zweiten die zweidimensionalen invarianten Unterräume. Das ergibt eine Folge von Zerlegungen

$$\begin{aligned} V &= V_0 \oplus W_1 = V_1 \oplus W_2 = V_2 \oplus W_3 = \dots = V_r \oplus W_{r+1} \\ &= V_{r+2} \oplus W_{r+3} = V_{r+4} \oplus W_{r+5} = \dots = V_{r+2m} \oplus W_{r+2m+1}, \end{aligned}$$

wobei $W_1 = V$, $W_{r+2m+1} = \{0\}$ und

$$\{0\} = V_0 \subset V_1 \subset \dots \subset V_r \subset V_{r+2} \subset \dots \subset V_{r+2m} = V$$

mit $\dim V_j = j$ und $F(V_j) \subset V_j$. Dazu verwendet man die eindeutige Zerlegung von

$$F|_{W_i}: W_i \rightarrow V = V_{i-1} \oplus W_i \quad \text{in} \quad F|_{W_i}: H_i + F_i \quad \text{mit}$$

$$H_i: W_i \rightarrow V_{i-1} \quad \text{und} \quad F_i: W_i \rightarrow W_i.$$

Die Einzelheiten seien getrost dem Leser überlassen. □

4.5.5. Nun wollen wir das in 4.5.2 begonnene und durch den Satz von CAYLEY-HAMILTON unterbrochene Spiel mit den Polynomen noch etwas weiter treiben.

Definition. Eine Teilmenge $\mathcal{I}$ eines kommutativen Ringes R heißt *Ideal*, wenn gilt

$$\text{I 1 } P, Q \in \mathcal{I} \Rightarrow P - Q \in \mathcal{I},$$

$$\text{I 2 } P \in \mathcal{I}, Q \in R \Rightarrow Q \cdot P \in \mathcal{I}.$$

Ein Beispiel dafür ist unser Ideal $\mathcal{I}_F \subset K[t]$.

Satz. Zu jedem Ideal $\mathcal{I} \subset K[t]$ mit $\mathcal{I} \neq \{0\}$ gibt es ein eindeutiges Polynom M mit folgenden Eigenschaften:

- 1) M ist normiert, d.h. $M = t^d + \dots$, wobei $d = \deg M$.
- 2) Für jedes $P \in \mathcal{I}$ gibt es ein $Q \in K[t]$ mit $P = Q \cdot M$.

M heißt Minimalpolynom von $\mathcal{I}$, im Fall $\mathcal{I} = \mathcal{I}_F$ Minimalpolynom von F .

Beweis. Sei $d := \min \{r : \text{es gibt ein } P \in \mathcal{I} \text{ mit } P \neq 0 \text{ und } r = \deg P\} \in \mathbb{N}$. Wir nehmen ein normiertes $M \in \mathcal{I}$ mit $\deg M = d$. Ein beliebiges $P \in \mathcal{I}$ dividieren wir entsprechend 1.3.7 mit Rest durch M :

$$P = Q \cdot M + R,$$

wobei $R = 0$ oder $\deg R < d$. Es genügt, den Fall $R \neq 0$ auszuschließen. Dann wäre

$$R = P - Q \cdot M \in \mathcal{I}$$

im Widerspruch zur Minimalität von d . Die Eindeutigkeit ist klar. $\square$

4.5.6. Es ist in der Geometrie eines Endomorphismus begründet, wie weit sich das Minimalpolynom vom charakteristischen Polynom unterscheidet.

Beispiel. Sei

$$A = \begin{pmatrix} 0 & 1 & & & & \\ & \ddots & \ddots & & & \\ & & \ddots & 1 & & \\ & & & \ddots & \ddots & \\ & 0 & & & \ddots & 0 \\ & & & & & \ddots \\ & & & & & & 0 \end{pmatrix} \in M(n \times n; K),$$

wobei $d - 1$ mit $1 \leq d \leq n$ die Anzahl der Einsen in der Nebendiagonalen bezeichnet. Das bedeutet

$$A(e_1) = 0, \quad A(e_2) = e_1, \dots, \quad A(e_d) = e_{d-1} \quad \text{und} \quad A(e_i) = 0 \quad \text{sonst.}$$

Für die Potenzen A^l mit $1 \leq l \leq d$ gilt $A^l(e_d) = e_{d-l}$, also ist $A^d = 0$, aber $A^{d-1} \neq 0$. Daraus folgt

$$P_A = \pm t^n, \quad M_A = t^d.$$

Im Extremfall $d = 1$ ist $A = 0$ und $P_A = \pm M_A^n$.

Allgemein gilt folgende Beziehung zwischen charakteristischem Polynom P_F und Minimalpolynom M_F :

Satz. Sei V ein n -dimensionaler K -Vektorraum, $F \in \text{End}(V)$. Dann gilt:

1) M_F teilt P_F .

2) P_F teilt M_F^n .

Beweis. 1) folgt sofort aus dem Satz von CAYLEY-HAMILTON. Zum Beweis von 2) wollen wir voraussetzen, daß K Unterkörper von $\mathbb{C}$ ist. Im allgemeinen Fall kann man einen *algebraischen Abschluß* $\bar{K}$ verwenden, aber so weit wollen wir nicht in die Algebra ausschweifen (vgl. [F-S]).

Wir können F ersetzen durch $A \in M(n \times n; K)$. Die Polynome M_A und P_A zerfallen über $\mathbb{C}$ in Linearfaktoren (vgl. 1.3.9). Zunächst betrachten wir

$$P_A = \pm(t - \lambda_1)^{r_1} \cdot \dots \cdot (t - \lambda_k)^{r_k},$$

wobei $\lambda_1, \dots, \lambda_k \in \mathbb{C}$ paarweise verschieden sind. Aus Teil 1) folgt

$$M_A = (t - \lambda_1)^{s_1} \cdot \dots \cdot (t - \lambda_k)^{s_k} \quad \text{mit} \quad 0 \leq s_i \leq r_i \quad \text{für} \quad i = 1, \dots, k,$$

d.h. M_A hat keine weiteren Linearfaktoren. Wir zeigen, daß keiner fehlt, d.h. daß $s_i \geq 1$ ist für alle i . Angenommen $s_i = 0$ für ein i . Dann nehmen wir einen Eigenvektor $v_i \in V$ zu λ_i und wenden $M_A(A)$ darauf an. Mit der gleichen Rechnung wie im Beispiel 4.5.1 folgt

$$M_A(A)(v_i) \neq 0$$

im Widerspruch zu $M_A(A) = 0$. Nun definieren wir

$$Q := (t - \lambda_1)^{ns_1 - r_1} \cdot \dots \cdot (t - \lambda_k)^{ns_k - r_k} \in \mathbb{C}[t].$$

Nach Definition von Q gilt $M_A^n = \pm Q \cdot P_A$, und da $M_A^n, P_A \in K[t]$ folgt $Q \in K[t]$ (Aufgabe 5 zu 1.3). $\square$

4.5.7. Manchmal kann das Minimalpolynom helfen, Beweise etwas abstrakter und kürzer zu machen. Dafür geben wir ein naheliegendes Beispiel.

Definition. Man nennt $F \in \text{End}_K(V)$ *nilpotent*, wenn $F^k = 0$ für ein $k \geq 1$.

Satz. Ist $F \in \text{End}_K(V)$ und $n = \dim V$, so sind folgende Aussagen äquivalent:

i) F ist nilpotent.

ii) $F^d = 0$ für ein d mit $1 \leq d \leq n$.

iii) $P_F = \pm t^n$.

iv) Es gibt eine Basis B von V , so daß

$$M_B(F) = \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix}.$$

Beweis. i) bedeutet $t^k \in \mathcal{I}_F$. Also ist $M_F = t^d$ mit $1 \leq d \leq n$ und $P_F = \pm t^n$. Das zeigt i) $\Rightarrow$ ii) und iii). Der Trigonalisierungssatz 4.4.3 enthält iii) $\Rightarrow$ iv). Schließlich folgt iv) $\Rightarrow$ i) nach Aufgabe 2 zu 2.5. $\square$

Aufgabe 3 zu 4.4 zeigt, daß es auch ohne Minimalpolynom geht.

Aufgaben zu 4.5

1. Sei $F: V \rightarrow V$ linear und $P \in K[t]$. Zeigen Sie: Ist $\lambda \in K$ ein Eigenwert von F , so ist $P(\lambda)$ ein Eigenwert von $P(F)$.

2. Ist $F: V \rightarrow V$ linear und $P, Q \in K[t]$, so ist

$$P(F) \circ Q(F) = Q(F) \circ P(F) = (P \cdot Q)(F).$$

3. Sei $F: V \rightarrow V$ ein Endomorphismus des K -Vektorraums V . Dann gilt:

- Die Abbildung $\Phi_F: K[t] \rightarrow \text{End}(V)$, $P(t) \mapsto P(F)$ ist ein Homomorphismus von Ringen und von K -Vektorräumen.
- $K[F] = \{P(F) : P \in K[t]\}$ ist ein kommutativer Unterring von $\text{End}(V)$.
- Ist $\dim V = n < \infty$, so existiert ein normiertes Polynom $P \in K[t]$ vom Grad $\leq n^2$ mit $P(F) = 0$. (Hinweis: Betrachten Sie $\text{id}, F, F^2, \dots, F^{n^2}$.)

4. Beweisen Sie den Satz von CAYLEY-HAMILTON durch direkte Rechnung für Matrizen

$$A \in M(2 \times 2; K).$$

5. Beweisen Sie den Satz von CAYLEY-HAMILTON für einen diagonalisierbaren Endomorphismus.

6. Geben Sie noch einen anderen Beweis des Satzes von CAYLEY-HAMILTON durch Induktion nach $n = \dim V$ mit der folgenden Methode:

- Zeigen Sie, dass $F(W) \subset W$ und berechnen Sie $P_G(t)$ für $G := F|_W$ (siehe Aufgabe 4 in 4.2).
- Zeigen Sie $P_G(G) = 0 \in \text{End}(W)$.
- Folgern Sie daraus im Fall $k < n$ mit der Bemerkung aus 4.4.4 und der Induktionsannahme, dass $P_F(F) = 0$.

7. Seien $\lambda_1, \dots, \lambda_r \in K$ die paarweise verschiedenen Eigenwerte eines diagonalisierbaren Endomorphismus F über einem endlichdimensionalen Vektorraum. Zeigen Sie, dass $(t - \lambda_1) \cdots (t - \lambda_r) \in K[t]$ das Minimalpolynom von F ist.

4.6 Die Jordansche Normalform*

Gegenstand dieses Abschnittes sind Endomorphismen, deren charakteristisches Polynom in Linearfaktoren zerfällt. Bedingungen für die Diagonalisierbarkeit sind in 4.1.3 und 4.3.3 angegeben. Sind sie nicht erfüllt, so kann man entsprechend 4.4.3 wenigstens trigonalisieren. Unbefriedigend dabei ist, daß man gar nichts von den Einträgen über der Diagonale weiß. Das wird nun nachgeholt.

Wir benutzen dabei nur elementare Methoden der linearen Algebra und verzichten darauf, das Minimalpolynom und die Teilbarkeitstheorie im Polynomring einzusetzen. Damit geht manches etwas eleganter, aber insgesamt nicht kürzer (vgl. etwa [B2]).

4.6.1. Im folgenden sei stets K ein beliebiger Körper und V ein K -Vektorraum mit $\dim V \geq 1$. Ist $F \in \text{End}(V)$ und

$P_F = \pm(t - \lambda_1)^{r_1} \cdot \dots \cdot (t - \lambda_k)^{r_k}$ mit $\lambda_1, \dots, \lambda_k \in K$ paarweise verschieden, so ist V nach 4.3.3 genau dann direkte Summe der F -invarianten Eigenräume $\text{Eig}(F; \lambda_i)$, wenn

$$\dim \text{Eig}(F; \lambda_i) = \mu(P_F; \lambda_i) = r_i.$$

Ist die Dimension des Eigenraumes zu klein, so kann man ihn mit folgendem Trick passend vergrößern. Für einen Eigenwert λ der Vielfachheit $r \geq 1$ ist

$$\text{Eig}(F; \lambda) = \text{Ker}(F - \lambda \text{id}_V) \subset \text{Ker}(F - \lambda \text{id}_V)^r =: \text{Hau}(F; \lambda).$$

Man nennt $\text{Hau}(F; \lambda)$ den *Hauptraum* (oder *verallgemeinerten Eigenraum*) von F zum Eigenwert λ . Damit kann man schon das Ergebnis der ersten Etappe formulieren:

Satz über die Hauptraumzerlegung. Sei $F \in \text{End}_K(V)$ und

$$P_F = \pm(t - \lambda_1)^{r_1} \cdot \dots \cdot (t - \lambda_k)^{r_k}$$

mit paarweise verschiedenen $\lambda_1, \dots, \lambda_k \in K$. Es sei $V_i := \text{Hau}(F; \lambda_i) \subset V$ für jedes λ_i der Hauptraum. Dann gilt:

1) $F(V_i) \subset V_i$ und $\dim V_i = r_i$ für $i = 1, \dots, k$.

2) $V = V_1 \oplus \dots \oplus V_k$.

3) F hat eine Zerlegung $F = F_D + F_N$ mit

a) F_D diagonalisierbar,

b) F_N nilpotent,

c) $F_D \circ F_N = F_N \circ F_D$.

Übersetzt man das in die Sprache der Matrizen, und verwendet man noch, daß jede nilpotente Matrix in eine obere Dreiecksmatrix transformiert werden kann (4.5.7), so erhält man das

Korollar. Sei $A \in M(n \times n; K)$, so daß $P_A = \pm(t - \lambda_1)^{r_1} \cdot \dots \cdot (t - \lambda_k)^{r_k}$. Dann gibt es eine invertierbare Matrix $S \in GL(n; K)$ derart, daß

$$SAS^{-1} = \begin{pmatrix} \boxed{\lambda_1 E_{r_1} + N_1} & & 0 \\ & \ddots & \\ 0 & & \boxed{\lambda_k E_{r_k} + N_k} \end{pmatrix} =: \tilde{A},$$

wobei für $i = 1, \dots, k$

$$\lambda_i E_{r_i} + N_i = \begin{pmatrix} \lambda_i & & * \\ & \ddots & \\ 0 & & \lambda_i \end{pmatrix} \in M(r_i \times r_i; K).$$

Insbesondere ist $\tilde{A} = D + N$, wobei D Diagonalmatrix, N nilpotent ist, und es gilt

$$D \cdot N = N \cdot D. \quad \square$$

Folgerungen aus der letzten Eigenschaft sind in Aufgabe 4 und Aufgabe 5 enthalten.

4.6.2. Zur Vorbereitung der Hauptraumzerlegung betrachten wir einen Eigenwert λ von F und Potenzen von

$$G := F - \lambda \text{id}_V.$$

Das können wir auch für ein ganz beliebiges $G \in \text{End } V$ tun. Offensichtlich hat man die Ketten

$$\begin{aligned} \{0\} &\subset \text{Ker } G \subset \text{Ker } G^2 \subset \dots \subset \text{Ker } G^l, \\ V &\supset \text{Im } G \supset \text{Im } G^2 \supset \dots \supset \text{Im } G^l, \end{aligned}$$

und für jedes l ist $\dim \text{Ker } G^l + \dim \text{Im } G^l = \dim V$ nach 2.2.4. Im allgemeinen ist jedoch

$$\text{Ker } G^l \cap \text{Im } G^l \neq \{0\},$$

man hat also keine direkte Summe (vgl. Beispiel 4.5.6). Da V endlichdimensional ist, können die beiden Ketten nicht endlos auf- bzw. absteigen. Das ist der Schlüssel zum entscheidenden

Lemma von FITTING. Zu einem $G \in \text{End}_K(V)$ betrachten wir die beiden Zahlen

$$d := \min\{l \in \mathbb{N} : \text{Ker } G^l = \text{Ker } G^{l+1}\} \quad \text{und} \quad r := \mu(P_G; 0),$$

wobei $G^0 := \text{id}_V$. Dann gilt:

- 1) $d = \min\{l : \text{Im } G^l = \text{Im } G^{l+1}\}$.
- 2) $\text{Ker } G^{d+i} = \text{Ker } G^d$, $\text{Im } G^{d+i} = \text{Im } G^d$ für alle $i \in \mathbb{N}$.
- 3) Die Räume $U := \text{Ker } G^d$ und $W := \text{Im } G^d$ sind G -invariant.
- 4) $(G|U)^d = 0$ und $G|W : W \rightarrow W$ ist ein Isomorphismus.
- 5) Für das Minimalpolynom von $G|U$ gilt $M_{G|U} = t^d$.
- 6) $V = U \oplus W$, $\dim U = r \geq d$, $\dim W = n - r$.

Insbesondere gibt es eine Basis B von V , so daß

$$M_B(G) = \begin{pmatrix} N & 0 \\ 0 & C \end{pmatrix} \quad \text{mit} \quad N^d = 0 \quad \text{und} \quad C \in \text{GL}(n-r; K).$$

Beweis. Wir betrachten das Diagramm

$$\begin{array}{ccccc} \text{Ker } G^l & \subset & V & \xrightarrow{G^l} & \text{Im } G^l \\ & \cap & & \parallel & \cup \\ \text{Ker } G^{l+1} & \subset & V & \xrightarrow{G^{l+1}} & \text{Im } G^{l+1} \end{array}$$

Unter Benutzung der Dimensionsformel 2.2.4 für G^l und G^{l+1} folgt

$$\begin{aligned} \text{Im } G^{l+1} = \text{Im } G^l &\Leftrightarrow \dim \text{Im } G^{l+1} = \dim \text{Im } G^l \\ &\Leftrightarrow \dim \text{Ker } G^{l+1} = \dim \text{Ker } G^l \\ &\Leftrightarrow \text{Ker } G^{l+1} = \text{Ker } G^l. \end{aligned}$$

Dies ist weiter gleichwertig damit, daß $G|_{\text{Im } G^l} : \text{Im } G^l \rightarrow \text{Im } G^{l+1}$ ein Isomorphismus ist. Daraus folgt sofort 1) und $G|_W$ Isomorphismus, also auch 2). Der Rest von 3) und 4) ist klar. Zu 5) genügt es zu zeigen, daß aus $(G|U)^{d-1} = 0$ folgen würde, daß

$$\text{Ker } G^d = U \subset \text{Ker } G^{d-1}.$$

Ist $v \in U \cap W$, so ist $G^d(v) = 0$ und $v = G^d(w)$ für ein $w \in V$. Daraus folgt $G^{2d}(w) = 0$, also

$$w \in \text{Ker } G^{2d} = \text{Ker } G^d$$

nach 2), und somit $0 = G^d(w) = v$. Damit hat man $V = U \oplus W$, und es bleibt die Berechnung der Dimension von U . Nach Definition von U ist $\dim U \geq d$. Wegen

$$t^r \cdot Q = P_G = P_{G|U} \cdot P_{G|W} \quad \text{mit } Q(0) \neq 0$$

und $P_{G|U} = \pm t^m$ mit $m = \dim U$ folgt $m = r$, denn für den Isomorphismus $G|W$ ist $P_{G|W}(0) \neq 0$. $\square$

Insbesondere folgt aus den obigen Argumenten die

Bemerkung. Ist λ Eigenwert von F und $G := F - \lambda \text{id}_V$, so ist $d = 1$ gleichbedeutend mit

$$\text{Eig}(F; \lambda) = \text{Hau}(F; \lambda).$$

$\square$

4.6.3. Nun beweisen wir den Satz über die Hauptraumzerlegung durch Induktion über die Zahl $k \geq 1$ der verschiedenen Eigenwerte. Zu λ_1 definieren wir $G := F - \lambda_1 \cdot \text{id}_V$. Dann ist

$$P_G(t - \lambda_1) = P_F(t), \quad \text{also} \quad \mu(P_G; 0) = \mu(P_F; \lambda_1) = r_1.$$

Nach Lemma 4.6.2 ist $V = \text{Hau}(F; \lambda_1) \oplus W$, und die beiden Summanden sind wegen $F = G + \lambda_1 \text{id}_V$ auch F -invariant. Weiter gilt

$$P_{F|W} = \pm(t - \lambda_2)^{r_2} \cdot \dots \cdot (t - \lambda_k)^{r_k},$$

also können wir auf $F|W$ die Induktionsannahme anwenden, daraus folgen 1) und 2).

Die Zerlegung 3) zeigt man am einfachsten mit Matrizen, die Bezeichnungen seien wie im Korollar gewählt. Die Existenz erhält man mit

$$D := \begin{pmatrix} \boxed{\lambda_1 E_{r_1}} & & 0 \\ & \ddots & \\ 0 & & \boxed{\lambda_k E_{r_k}} \end{pmatrix} \quad \text{und} \quad N := \begin{pmatrix} \boxed{N_1} & & 0 \\ & \ddots & \\ 0 & & \boxed{N_k} \end{pmatrix},$$

eine einfache Rechnung zeigt

$$D \cdot N = \begin{pmatrix} \boxed{\lambda_1 N_1} & & 0 \\ & \ddots & \\ 0 & & \boxed{\lambda_k N_k} \end{pmatrix} = N \cdot D.$$

$\square$

Zusatz. Die Zerlegung $F = F_D + F_N$ ist sogar eindeutig, wenn man a), b) und c) verlangt.

Der Beweis benutzt 4.3.6 und etwas Teilbarkeitstheorie von Polynomen (vgl. etwa [B2], §11).

4.6.4. Aus den obigen Beweisen kann man ein Rechenverfahren ableiten, wir geben dafür ein

Beispiel. Gegeben sei die Matrix

$$A = \begin{pmatrix} 25 & 34 & 18 \\ -14 & -19 & -10 \\ -4 & -6 & -1 \end{pmatrix}$$

mit $P_A = -t^3 + 5t^2 - 7t + 3 = -(t-1)^2 \cdot (t-3)$. In der Notation von 4.6.3 ist $k=2$, $\lambda_1=1$, $r_1=2$, $\lambda_2=3$, $r_2=1$. Wir betrachten die Matrizen

$$B_1 = A - \lambda_1 E_3 = \begin{pmatrix} 24 & 34 & 18 \\ -14 & -20 & -10 \\ -4 & -6 & -2 \end{pmatrix}, \quad B_2 = A - \lambda_2 E_3 = \begin{pmatrix} 22 & 34 & 18 \\ -14 & -22 & -10 \\ -4 & -6 & -4 \end{pmatrix},$$

sowie

$$B_1^{r_1} = B_1^2 = \begin{pmatrix} 28 & 28 & 56 \\ -16 & -16 & -32 \\ -4 & -4 & -8 \end{pmatrix}, \quad B_2^{r_2} = B_2.$$

Es ist $\text{rang } B_1 = 2 > 1 = \text{rang } B_1^2$. Daraus folgt

$$\text{Eig}(A; \lambda_1) = \text{Ker } B_1 \subsetneq \text{Ker } B_1^2 = \text{Hau}(A; \lambda_1),$$

und somit ist A nicht diagonalisierbar. Indem man die zu B_1^2 und B_2 gehörenden homogenen Gleichungssysteme löst, erhält man Basen

$$({}'(2, 0, -1), {}'(0, 2, -1)) \quad \text{von } \text{Hau}(A; \lambda_1)$$

und

$${}'(-7, 4, 1) \quad \text{von } \text{Hau}(A; \lambda_2) = \text{Eig}(A; \lambda_2).$$

Trägt man die erhaltenen Basisvektoren als Spalten in eine Matrix ein, so erhält man

$$S^{-1} = \begin{pmatrix} 2 & 0 & -7 \\ 0 & 2 & 4 \\ -1 & -1 & 1 \end{pmatrix} \quad \text{und} \quad S = \frac{1}{2} \begin{pmatrix} -6 & -7 & -14 \\ 4 & 5 & 8 \\ -2 & -2 & -4 \end{pmatrix},$$

sowie

$$SAS^{-1} = \begin{pmatrix} 16 & 25 & 0 \\ -9 & -14 & 0 \\ 0 & 0 & 3 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 16 & 25 \\ -9 & -14 \end{pmatrix} = E_2 + N \quad \text{mit } N^2 = 0.$$

Nun muß man noch die Basis von $\text{Hau}(A; \lambda_1)$ so transformieren, daß ein Basisvektor v Eigenvektor zu λ_1 wird (vgl. 4.4.5). Er hat die Form

$$v = \alpha \cdot {}'(2, 0, -1) + \beta \cdot {}'(0, 2, -1).$$

Die Bedingung $Av = v$ ergibt $3\alpha + 5\beta = 0$, also kann man $v = {}'(5, -3, -1)$ wählen. Aus der neuen Basis von $\mathbb{R}^3$ erhält man die Transformationsmatrix

$$T^{-1} = \begin{pmatrix} 5 & 2 & -7 \\ -3 & 0 & 4 \\ -1 & -1 & 1 \end{pmatrix}, \quad T = \frac{1}{3} \begin{pmatrix} -4 & -5 & -8 \\ 1 & 2 & -1 \\ -3 & -3 & -6 \end{pmatrix},$$

sowie

$$TAT^{-1} = \begin{pmatrix} 1 & 6 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

4.6.5. Auf der zweiten Etappe des Weges zur JORDANSchen Normalform wird zu einem nilpotenten Endomorphismus eine Basis maßgeschneidert. Um das Ergebnis einfach formulieren zu können, definieren wir die *Jordanmatrix*

$$J_k := \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix} \in M(k \times k; K),$$

für die $J_k^k = 0$ gilt, und k die minimale Potenz mit dieser Eigenschaft ist. Man beachte, daß $J_1 = (0)$.

Theorem. Sei G ein nilpotenter Endomorphismus eines K -Vektorraumes V und $d := \min\{l : G^l = 0\}$. Dann gibt es eindeutig bestimmte Zahlen $s_1, \dots, s_d \in \mathbb{N}$ mit

$$d \cdot s_d + (d-1)s_{d-1} + \dots + s_1 = r = \dim V$$

und eine Basis $\mathcal{B}$ von V , so daß

$$M_B(G) = \begin{pmatrix} J_d & & & & \\ & \ddots & & & \\ & & J_d & & \\ & & & J_{d-1} & \\ & & & & \ddots & 0 \\ & & & & & \ddots & J_{d-1} \\ & & 0 & & & & \ddots & J_1 \\ & & & & & & & \ddots & J_1 \end{pmatrix} \begin{matrix} \left. \begin{matrix} \\ \\ \\ \\ \end{matrix} \right\} s_d\text{-mal} \\ \left. \begin{matrix} \\ \\ \end{matrix} \right\} s_{d-1}\text{-mal} \\ \vdots \\ \left. \begin{matrix} \\ \\ \end{matrix} \right\} s_1\text{-mal} \end{matrix}$$

Beweis. Wir definieren $U_l := \text{Ker } G^l$ und betrachten die Kette

$$\{0\} = U_0 \subset U_1 \subset \dots \subset U_{d-1} \subset U_d = V.$$

Da d minimal ist, sind alle Inklusionen echt. Wir vermerken zwei weitere Eigenschaften:

(1) Für $1 \leq l \leq d$ ist $G^{-1}(U_{l-1}) = U_l$, insbesondere $G(U_l) \subset U_{l-1}$.

Das ist klar, denn

$$v \in G^{-1}(U_{l-1}) \Leftrightarrow G(v) \in U_{l-1} \Leftrightarrow 0 = G^{l-1}(G(v)) = G^l(v) \Leftrightarrow v \in U_l.$$

Durch Anwendung von G auf die erste Gleichung folgt

$$G(U_l) = G(G^{-1}(U_{l-1})) \subset U_{l-1}.$$

Daß hier keine Gleichheit mehr gelten muß, sieht man etwa an der durch

$$G(e_1) = G(e_2) = 0, \quad G(e_3) = e_2$$

gegebenen Abbildung des K^3 mit $d = 2$.

(2) Ist $W \subset V$ ein Untervektorraum mit $W \cap U_l = \{0\}$ für irgend ein $l > 0$, so ist $G|_W$ injektiv.

Denn da $\text{Ker } G = U_1 \subset U_l$, gilt sogar $W \cap \text{Ker } G = \{0\}$.

Damit konstruieren wir nun schrittweise eine direkte Summenzerlegung von V . Zunächst wählen wir $W_d \subset V$ mit

$$V = U_d = U_{d-1} \oplus W_d.$$

Aus (1) folgt $G(W_d) \subset U_{d-1}$ und $G(W_d) \cap U_{d-2} = \{0\}$. Also gibt es eine Zerlegung

$$U_{d-1} = U_{d-2} \oplus W_{d-1} \quad \text{mit} \quad G(W_d) \subset W_{d-1}.$$

Die Iteration dieses Verfahrens kann man schematisch darstellen, wobei die Pfeile die Wirkung von G zeigen:

$$\begin{array}{ccccccc}
 U_d & & & & & & \\
 \downarrow & & & & & & \\
 U_{d-1} \oplus & W_d & & & & & \\
 \downarrow & & \downarrow & & & & \\
 U_{d-2} \oplus & W_{d-1} \oplus & W_d & & & & \\
 \downarrow & & \downarrow & & \downarrow & & \\
 \vdots & & \vdots & & \vdots & & \\
 U_1 \oplus & W_2 \oplus & W_3 \oplus & \dots \oplus & W_d & & \\
 \downarrow & & \downarrow & & \downarrow & & \downarrow \\
 U_0 \oplus & W_1 \oplus & W_2 \oplus & \dots \oplus & W_{d-1} \oplus & W_d &
 \end{array}$$

Jede Zeile ist dabei eine Zerlegung von V , wegen $U_0 = 0$ ist insbesondere

$$V = W_1 \oplus W_2 \oplus \dots \oplus W_d.$$

Da nach (2) all die Beschränkungen von G in der Kette

$$W_d \rightarrow W_{d-1} \rightarrow \dots \rightarrow W_1$$

injektiv sind, können wir nun bei W_d beginnend durch schrittweise Ergänzung Basen aller W_i und damit insgesamt eine Basis von V nach folgendem Schema erhalten:

$$\begin{array}{ccccccc}
 w_1^{(d)}, \dots, w_{s_d}^{(d)} & & & & & & \\
 G(w_1^{(d)}), \dots, G(w_{s_d}^{(d)}), & w_1^{(d-1)}, \dots, w_{s_{d-1}}^{(d-1)} & & & & & \\
 \vdots & \vdots & \vdots & & \vdots & & \\
 G^{d-1}(w_1^{(d)}), \dots, G^{d-1}(w_{s_d}^{(d)}), & G^{d-2}(w_1^{(d-1)}), \dots, G^{d-2}(w_{s_{d-1}}^{(d-1)}), & \dots, & w_1^{(1)}, \dots, w_{s_1}^{(1)} & & &
 \end{array}$$

Dabei ist die erste Zeile eine Basis von W_d , die zweite von W_{d-1} , und schließlich die letzte von

$$W_1 = U_1 = \text{Ker } G.$$

Die Matrix $M_{\mathcal{B}}(G)$ hat die versprochene Form, wenn man $\mathcal{B}$ in folgender Weise anordnet: Man läuft in jeder Spalte von unten nach oben und liest die Spalten von links nach rechts.

Um zu zeigen, daß die Zahlen $s_1, \dots, s_d$ allein durch G bestimmt sind, benutzen wir die Existenz von Zerlegungen

$$U_l = U_{l-1} \oplus W_l = U_{l-1} \oplus G(W_{l+1}) \oplus \tilde{W}_l$$

für $l = 1, \dots, d$ mit $W_{d+1} = 0$. Danach ist

$$s_l = \dim \tilde{W}_l = \dim U_l - \dim U_{l-1} - \dim W_{l+1},$$

also sind diese Zahlen rekursiv aus den Dimensionen der Kerne von G^l berechenbar (vgl. Aufgabe 6). $\square$

4.6.6. Wir berechnen ein einfaches

Beispiel. Gegeben sei die Matrix

$$B = \begin{pmatrix} 0 & 1 & 3 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{mit} \quad B^2 = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{und} \quad B^3 = 0.$$

Es ist $d = 3$, und für $U_l := \text{Ker } B^l$ gilt

$$\{0\} = U_0 \subset U_1 = \text{span } e_1 \subset U_2 = \text{span } (e_1, e_2) \subset U_3 = \mathbb{R}^3.$$

Aus der Bedingung

$$\mathbb{R}^3 = U_2 \oplus W_3$$

folgt, daß wir $W_3 = \text{span } e_3$ wählen können. Somit ist $s_3 = 1$. Aus

$$\mathbb{R}^3 = U_1 \oplus W_2 \oplus W_3$$

folgt $\dim W_2 = 1$, also $s_2 = 0$, und $B \cdot e_3 = {}^t(3, 2, 0)$ ist der richtige Basisvektor von W_2 . Schließlich ist

$$\mathbb{R}^3 = U_0 \oplus W_1 \oplus W_2 \oplus W_3.$$

Der gesuchte Basisvektor von W_1 ist $B^2 \cdot e_3 = B \cdot {}^t(3, 2, 0) = 2e_1$, somit ist auch $s_1 = 0$. Trägt man die gefundenen Basisvektoren in der richtigen Reihenfolge als Spalten in eine Matrix ein, so erhält man

$$T^{-1} = \begin{pmatrix} 2 & 3 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{und daraus} \quad T = \frac{1}{4} \begin{pmatrix} 2 & -3 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

Das Ergebnis ist die Jordanmatrix

$$TBT^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

mit d Einträgen λ_i und $d - 1$ Einsen heißt *Jordanblock* der Länge d zu λ_i . Der oberste und größte Jordanblock in $\lambda_i E_{r_i} + N_i$ hat die Größe d_i mit

$$1 \leq d_i = \min \{l: N_i^l = 0\} \leq r_i,$$

das ist die Vielfachheit der Nullstelle λ_i im Minimalpolynom von F (Aufgabe 8).

Für $1 \leq j \leq d_i$ seien $s_j^{(i)} \geq 0$ die Anzahlen der Jordanblöcke der Größe j zu λ_i in $\lambda_i E_{r_i} + N_i$. Es ist $s_{d_i}^{(i)} \geq 1$, und durch Abzählung der Längen folgt

$$d_i s_{d_i}^{(i)} + (d_i - 1) s_{d_i-1}^{(i)} + \dots + s_1^{(i)} = r_i.$$

Die Elemente $\lambda_i \in K$ und $r_i, d_i, s_j^{(i)} \in \mathbb{N}$ sind allein durch den Endomorphismus F bestimmt (für die $s_j^{(i)}$ folgt das aus Aufgabe 6). Sie heißen daher *Invarianten* von F . Die Eigenwerte λ_i kann man im Fall $K = \mathbb{R}$ oder $\mathbb{C}$ *kontinuierliche* Invarianten nennen. Die anderen sind *diskret* und durch die oben angegebenen Bezeichnungen, sowie

$$r_1 + \dots + r_k = n$$

aneinander gekoppelt. Alle Invarianten zusammen beschreiben in komprimierter Form die „Geometrie von F “. Umgekehrt bedeutet dies, daß es so viele Möglichkeiten gibt, wesentlich verschiedene Endomorphismen zu konstruieren, wie man „Sätze“ von solchen Invarianten zusammenstellen kann. Nur die Reihenfolge der Jordanblöcke ist unwesentlich, sie kann durch eine passende Permutation der Basisvektoren beliebig verändert werden. Das ist die Aussage dieses wohl schwierigsten Theorems der elementaren linearen Algebra.

Das Verfahren, die Normalform zu berechnen, hat praktische Anwendungen, etwa bei der Lösung von Systemen linearer Differentialgleichungen mit konstanten Koeffizienten. An sich genügt es, die Systeme nach verschiedenen Eigenwerten zu entkoppeln (das geht mit der Hauptraumzerlegung) und in jedem Hauptraum zu trigonalisieren (dann kann man die Teilsysteme von unten nach oben lösen, vgl. [Fo 2], §14). Aber der Rechenaufwand für die verbesserte Trigonalisierung mit der JORDAN-Methode ist kaum größer als bei der Methode aus 4.4.5.

Nach all den schon geleisteten Vorbereitungen ist der *Beweis* nunmehr ganz kurz: Für $i = 1, \dots, k$ setzen wir

$$V_i := \text{Hau}(F; \lambda_i) \quad \text{und} \quad G_i := (F - \lambda_i \text{id}_V)|_{V_i}.$$

Anwendung von 4.6.5 auf G_i ergibt eine Basis $\mathcal{B}_i$ von V_i , und aus $\mathcal{B}_1, \dots, \mathcal{B}_k$ baut man sich das gesuchte $\mathcal{B}$ auf (vgl. 1.6.4). $\square$

Für die Minimalpolynome von G_i und F gilt (Aufgabe 8)

$$M_{G_i} = t^{d_i} \quad \text{und} \quad M_F = (t - \lambda_1)^{d_1} \cdot \dots \cdot (t - \lambda_k)^{d_k}.$$

Ist speziell $d_1 = \dots = d_k = 1$, so gibt es keine Einsen neben der Diagonale, und es folgt zusammen mit 4.6.1 und 4.6.7 das

Korollar. Für ein $F \in \text{End}_K(V)$ sind folgende Bedingungen gleichwertig:

- i) F ist diagonalisierbar.
- ii) $M_F = (t - \lambda_1) \cdot \dots \cdot (t - \lambda_k)$, wobei $\lambda_1, \dots, \lambda_k$ die verschiedenen Eigenwerte von F bezeichnen.
- iii) Es gibt paarweise verschiedene $\lambda_1, \dots, \lambda_m \in K$, so daß

$$(F - \lambda_1 \text{id}_V) \circ \dots \circ (F - \lambda_m \text{id}_V) = 0 \in \text{End}_K(V).$$

□

4.6.8. Für die praktische Berechnung der Jordanschen Normalform verwenden wir wieder das

Beispiel.

$$A = \begin{pmatrix} 3 & 4 & 3 \\ -1 & 0 & -1 \\ 1 & 2 & 3 \end{pmatrix}$$

mit dem charakteristischen Polynom $P_A = -(t-2)^3$ (vgl. 4.4.5). Wir betrachten

$$B := A - 2E_3 = \begin{pmatrix} 1 & 4 & 3 \\ -1 & -2 & -1 \\ 1 & 2 & 1 \end{pmatrix} \quad \text{mit} \quad B^2 = \begin{pmatrix} 0 & 2 & 2 \\ 0 & -2 & -2 \\ 0 & 2 & 2 \end{pmatrix}$$

und $B^3 = 0$. Dann ist

$$U_1 := \text{Ker } B = \text{span}'(1, -1, 1)$$

$$U_2 := \text{Ker } B^2 = \text{span}'(1, -1, 1), '(0, -1, 1)).$$

Entsprechend 4.6.5 haben wir die Zerlegungen

$$\mathbb{R}^3 = U_2 \oplus W_3 = U_1 \oplus W_2 \oplus W_3 = U_0 \oplus W_1 \oplus W_2 \oplus W_3,$$

das ergibt die Basisvektoren

$$e_3 \in W_3, \quad B \cdot e_3 = '(3, -1, 1) \in W_2 \quad \text{und} \quad B^2 \cdot e_3 = '(2, -2, 2) \in W_1.$$

In der richtigen Reihenfolge als Spalten in eine Matrix eingetragen erhält man

$$T^{-1} = \begin{pmatrix} 2 & 3 & 0 \\ -2 & -1 & 0 \\ 2 & 1 & 1 \end{pmatrix} \quad \text{mit} \quad T = \frac{1}{4} \begin{pmatrix} -1 & -3 & 0 \\ 2 & 2 & 0 \\ 0 & 4 & 4 \end{pmatrix},$$

und als Ergebnis

$$TAT^{-1} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Das kann man vergleichen mit den Ergebnissen der Trigonalisierung in 4.4.5

$$SAS^{-1} = D = \begin{pmatrix} 2 & 1 & 3 \\ 0 & 2 & 2 \\ 0 & 0 & 2 \end{pmatrix} \quad \text{mit} \quad S = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix},$$

und der Transformation der nilpotenten Matrix $D - 2E_3$ zu

$$\tilde{T}(D - 2E_2)\tilde{T}^{-1} = J_3 \quad \text{mit} \quad \tilde{T} = \frac{1}{4} \begin{pmatrix} 2 & -3 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

Es ist $\tilde{T} \cdot S = T$, also kann man T auch durch Kombination der vorher durchgeführten Rechnungen erhalten. Man beachte, daß dabei nicht notwendig dieselbe Transformationsmatrix entstehen muß, weil sie nicht eindeutig bestimmt ist.

Hat A mehr als einen Eigenwert, so muß man zunächst die Zerlegung in Haupträume berechnen. Dann kann man in jedem Hauptraum wie oben vorgehen. Da der gesamte Raum nach 4.6.1 direkte Summe der Haupträume ist, ergeben die unabhängig voneinander konstruierten Basen der Haupträume zusammen eine Basis der gewünschten Art des gesamten Raumes.

Aufgaben zu 4.6

1. Bestimmen Sie die Haupträume der folgenden Matrizen:

$$\begin{pmatrix} 1 & 4 & 2 & 1 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} 2 & 3 & 3 & 1 & 8 \\ 0 & 2 & 7 & 2 & 8 \\ 0 & 0 & 2 & 5 & 4 \\ 0 & 0 & 0 & -1 & -4 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix}.$$

2. Bestimmen Sie Basen, bezüglich derer die folgenden nilpotenten Matrizen Jordansche Normalform haben, und geben Sie jeweils das Minimalpolynom an:

$$\begin{pmatrix} 0 & 2 & 2 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & -2 & 0 & -1 & 2 \\ 1 & -3 & -1 & 0 & 3 \\ 0 & 2 & 1 & -1 & -3 \\ 1 & 0 & 0 & -1 & -2 \\ 0 & -1 & 0 & 0 & 2 \end{pmatrix}.$$

3. Bestimmen Sie Basen, bezüglich derer die folgenden Matrizen Jordansche Normalform haben, und geben Sie jeweils das charakteristische und das Minimalpolynom an:

$$\begin{pmatrix} 3 & 4 & 3 \\ -1 & 0 & -1 \\ 1 & 2 & 3 \end{pmatrix}, \quad \begin{pmatrix} 2 & 1 & 1 & 0 & -2 \\ 1 & 1 & 1 & 0 & -1 \\ 1 & 0 & 2 & 0 & -1 \\ 1 & 0 & 1 & 2 & -2 \\ 1 & 0 & 1 & 0 & 0 \end{pmatrix}.$$

4. Mit Hilfe des Satzes über die Jordansche Normalform kann man recht einfach hohe Potenzen von Matrizen berechnen. Zeigen Sie:

- a) Ist $A \in M(n \times n; K)$, $S \in GL(n; K)$ und $m \in \mathbb{N}$, so gilt $(SAS^{-1})^m = SA^mS^{-1}$.
 b) Sind $A, B \in M(n \times n; K)$ mit $AB = BA$ und $m \in \mathbb{N}$, so gilt

$$(A + B)^m = \sum_{k=0}^m \binom{m}{k} A^k B^{m-k}.$$

- c) Bestimmen Sie für die Matrix

$$A = \begin{pmatrix} 3 & 4 & 3 \\ -1 & 0 & -1 \\ 1 & 2 & 3 \end{pmatrix}$$

eine Matrix $S \in GL(3; \mathbb{R})$, so dass $A = S(D + N)S^{-1}$, wobei D Diagonalmatrix, N nilpotent und $DN = ND$ ist. Berechnen Sie mit Hilfe von a) und b) (und ohne Computer) A^{50} .

5. Betrachten Sie die Verallgemeinerung der Exponentialfunktion für Matrizen; für jede Matrix $A \in M(n \times n; \mathbb{R})$ existiert

$$\exp(A) := \lim_{m \rightarrow \infty} \sum_{k=0}^m \frac{1}{k!} A^k.$$

- a) Bestimmen Sie $\exp(D)$ für eine Diagonalmatrix D .

- b) Ist $A \in M(n \times n; \mathbb{R})$ und $S \in GL(n; \mathbb{R})$, so folgt $\exp(SAS^{-1}) = S \cdot \exp(A) \cdot S^{-1}$.
c) Sind $A, B \in M(n \times n; \mathbb{R})$ mit $AB = BA$, so gilt $\exp(A + B) = \exp(A)\exp(B)$.
d) Bestimmen Sie für die Matrix

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

eine Matrix $S \in GL(3; \mathbb{R})$, so dass $A = S(D + N)S^{-1}$, wobei D Diagonalmatrix, N nilpotent und $DN = ND$ ist, und berechnen Sie $\exp(A)$.

6. Zeigen Sie, dass für die Zahlen $s_1, \dots, s_d$ in 4.6.5 gilt:

$$s_l = \dim(U_l/U_{l-1}) - \dim(U_{l+1}/U_l).$$

7. Gegeben sei ein endlichdimensionaler $\mathbb{C}$ -Vektorraum V . Zwei Endomorphismen F und G von V heißen *ähnlich*, wenn es einen Isomorphismus H von V gibt mit $G = H \circ F \circ H^{-1}$.

- a) Zeigen Sie, dass dadurch eine Äquivalenzrelation auf der Menge der Endomorphismen von V gegeben ist.
b) Für $F, G \in \text{End}(V)$ sind folgende Bedingungen gleichwertig:
i) F und G sind ähnlich.
ii) Für jede Basis $\mathcal{B}$ von V sind $M_{\mathcal{B}}(F)$ und $M_{\mathcal{B}}(G)$ ähnlich.
iii) Die Jordanschen Normalformen von F und G haben (bis auf die Reihenfolge) die gleichen Invarianten.

8. Sei $F \in \text{End}_K(V)$ ein Endomorphismus, dessen charakteristisches Polynom in Linearfaktoren zerfällt. Beweisen Sie, dass man das Minimalpolynom M_F aus den Invarianten von F berechnen kann: Mit den Bezeichnungen von 4.6.7 gilt

$$M_F(t) = (t - \lambda_1)^{d_1} \cdot \dots \cdot (t - \lambda_k)^{d_k}.$$

9. Sei V ein 6-dimensionaler $\mathbb{R}$ -Vektorraum und F ein Endomorphismus von V mit $P_F(t) = (t - 1)(t + 2)^5$, $M_F(t) = (t - 1)(t + 2)^3$. Bestimmen Sie alle möglichen Jordanschen Normalformen von F .

10. Sei V ein endlichdimensionaler $\mathbb{R}$ -Vektorraum und F ein Endomorphismus von V mit $F^3 = F$. Zeigen Sie, dass F diagonalisierbar ist.

Kapitel 5

Euklidische und unitäre Vektorräume

In den vorhergehenden Kapiteln wurden Vektorräume über einem beliebigen Körper behandelt. Über den reellen und komplexen Zahlen hat man ein zusätzliches wichtiges Hilfsmittel, das die Messung von Längen und Winkeln gestattet, nämlich ein Skalarprodukt. Damit ausgestattet nennt man reelle bzw. komplexe Vektorräume euklidisch bzw. unitär. Bevor wir die üblichen allgemeinen Begriffe einführen, geben wir die wichtigsten kanonischen Beispiele.

5.1 Das kanonische Skalarprodukt im $\mathbb{R}^n$

5.1.1. Zwei Vektoren $x = (x_1, \dots, x_n)$ und $y = (y_1, \dots, y_n)$ des $\mathbb{R}^n$ ordnet man die Zahl

$$\langle x, y \rangle = x_1 y_1 + \dots + x_n y_n$$

zu. Das ergibt eine Abbildung

$$(\cdot, \cdot): \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}, \quad (x, y) \mapsto \langle x, y \rangle,$$

sie heißt *kanonisches Skalarprodukt*. Schreibt man x und y als Spaltenvektoren, so ist

$$\langle x, y \rangle = {}^t x \cdot y = (x_1, \dots, x_n) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Die formalen Eigenschaften dieses Produktes rechnet man mühelos nach:

$$\begin{aligned} 1. \quad \langle x + x', y \rangle &= \langle x, y \rangle + \langle x', y \rangle, & \langle x, y + y' \rangle &= \langle x, y \rangle + \langle x, y' \rangle, \\ \langle \lambda x, y \rangle &= \lambda \langle x, y \rangle, & \langle x, \lambda y \rangle &= \lambda \langle x, y \rangle, \end{aligned}$$

$$2. \quad \langle x, y \rangle = \langle y, x \rangle,$$

$$3. \quad \langle x, x \rangle \geq 0 \quad \text{und} \quad \langle x, x \rangle = 0 \Leftrightarrow x = 0,$$

für $x, x', y, y' \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$. Man beachte, daß nur bei 3. eine spezielle Eigenschaft der reellen Zahlen benutzt wird:

$$\langle x, x \rangle = x_1^2 + \dots + x_n^2,$$

und eine Summe von Quadraten ist nicht negativ. Insbesondere kann man daraus die Wurzel ziehen, was eine Abbildung

$$\| \cdot \|: \mathbb{R}^n \rightarrow \mathbb{R}_+, \quad x \mapsto \|x\| := \sqrt{\langle x, x \rangle},$$

ergibt, die *Norm* heißt. Im $\mathbb{R}^1$ ist $\|x\| = |x|$, im $\mathbb{R}^2$ ist

$$\|x\| = \sqrt{x_1^2 + x_2^2}$$

nach dem klassischen Satz von PYTHAGORAS die Länge von x . Auch im $\mathbb{R}^n$ ist $\|x\|$ der *Abstand* vom Nullpunkt zum Punkt x .

Um Abstände zwischen beliebigen Punkten zu messen, betrachtet man die Abbildung

$$d: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_+, \quad d(x, y) := \|y - x\|.$$

Explizit berechnet man den *Abstand* (oder die *Metrik*) durch

$$d(x, y) := \sqrt{(y_1 - x_1)^2 + \dots + (y_n - x_n)^2}.$$

5.1.2. Die wichtigsten Eigenschaften von Norm und Abstand sind die folgenden

N1 $\|x\| = 0 \Leftrightarrow x = 0,$

N2 $\|\lambda x\| = |\lambda| \cdot \|x\|,$

N3 $\|x + y\| \leq \|x\| + \|y\|,$ *Dreiecksungleichung*

D1 $d(x, y) = 0 \Leftrightarrow x = y,$

D2 $d(y, x) = d(x, y),$ *Symmetrie*

D3 $d(x, z) \leq d(x, y) + d(y, z).$ *Dreiecksungleichung*

Die jeweiligen Eigenschaften 1 und 2 sind klar. D3 folgt aus N3, denn

$$x - z = x - y + y - z.$$

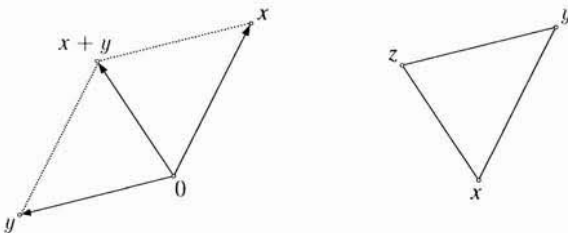


Bild 5.1

5.1.3. Man kann versuchen, die Ungleichung N3 direkt nachzurechnen. Aber es ist ökonomischer, sie aus der folgenden allgemeineren Ungleichung abzuleiten.

Ungleichung von CAUCHY-SCHWARZ. Für $x, y \in \mathbb{R}^n$ gilt

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\|$$

und

$$|\langle x, y \rangle| = \|x\| \cdot \|y\|$$

genau dann, wenn x und y linear abhängig sind.

Zunächst beweisen wir damit die *Dreiecksungleichung*:

$$\begin{aligned}\|x + y\|^2 &= \langle x + y, x + y \rangle = \langle x, x \rangle + 2\langle x, y \rangle + \langle y, y \rangle \\ &\leq \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2.\end{aligned}$$

Die Quadratwurzel ist monoton, also folgt N3. $\square$

Beweis der CAUCHY-SCHWARZschen Ungleichung. Wegen der Monotonie der Wurzel kann man beide Seiten quadrieren, es genügt also zu zeigen, daß

$$\langle x, x \rangle \cdot \langle y, y \rangle - \langle x, y \rangle^2 \geq 0. \quad (*)$$

Wie man das durch Ausrechnen zeigen kann, wird in 5.4.7 beschrieben. Wer gerne mit Kniffen rechnet, betrachte die Matrix

$$A = \begin{pmatrix} x_1 & \cdots & x_n \\ y_1 & \cdots & y_n \end{pmatrix}.$$

Dann ist

$$A \cdot {}^t A = \begin{pmatrix} \langle x, x \rangle & \langle x, y \rangle \\ \langle y, x \rangle & \langle y, y \rangle \end{pmatrix},$$

die abzuschätzende Differenz ist also gleich $\det(A \cdot {}^t A)$. Daher folgt (*) aus 3.3.7, denn

$$\det(A \cdot {}^t A) = \sum_{1 \leq i < j \leq n} \begin{vmatrix} x_i & x_j \\ y_i & y_j \end{vmatrix}^2 \geq 0.$$

Gleichheit bei Cauchy-Schwarz bedeutet, daß alle Minoren $x_i y_j - x_j y_i$ verschwinden, und das bedeutet $\text{rang } A < 2$ (vgl. 3.3.6). $\square$

5.1.4. Wie wir gesehen haben, hat die Cauchy-Schwarzsche Ungleichung eine Konsequenz für die *Längenmessung*, nämlich die Dreiecksungleichung. Man kann durch Betrachtung des Quotienten von linker und rechter Seite darüber hinaus eine sinnvolle *Winkelmessung* erklären:

Für zwei von Null verschiedene Vektoren $x, y \in \mathbb{R}^n$ ist nach der Ungleichung von CAUCHY-SCHWARZ

$$-1 \leq \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|} \leq +1.$$

Also ist dieser Quotient gleich $\cos \vartheta$ für genau ein $\vartheta \in [0, \pi]$. Dies nennt man den *Winkel zwischen x und y* , in Zeichen

$$\angle(x, y) := \arccos \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|}. \quad (*)$$

Um zu zeigen, daß diese Definition mit dem übereinstimmt, was man sich geometrisch unter einem (unorientierten) Winkel vorstellt, bemerken wir zunächst,

daß

$$\begin{aligned}\angle(x, y) &= \angle(y, x), \\ \angle(\lambda x, y) &= \angle(x, y) = \angle(x, \lambda y) \quad \text{für } \lambda > 0.\end{aligned}$$

Da die beiden Vektoren in einer Ebene liegen, behandeln wir nur den Fall $n = 2$ (daß dies keine echte Einschränkung ist, wird sich in 5.4.9 zeigen). Wir ersetzen x und y durch

$$x' = \frac{1}{\|x\|} \cdot x \quad \text{und} \quad y' = \frac{1}{\|y\|} \cdot y.$$

Dann ist $\|x'\| = \|y'\| = 1$ und $\angle(x, y) = \angle(x', y')$. Aus der Analysis weiß man, daß es $\alpha, \beta \in [0, 2\pi[$ gibt, so daß

$$x' = (\cos \alpha, \sin \alpha), \quad y' = (\cos \beta, \sin \beta).$$

Also gilt

$$\langle x', y' \rangle = \cos \alpha \cos \beta + \sin \alpha \sin \beta = \cos(\beta - \alpha),$$

nach einem „Additionstheorem“ für den Cosinus, wobei wir $\beta - \alpha \in [0, \pi]$ annehmen können. Insgesamt folgt

$$\angle(x, y) = \beta - \alpha,$$

was der Erwartung entspricht.

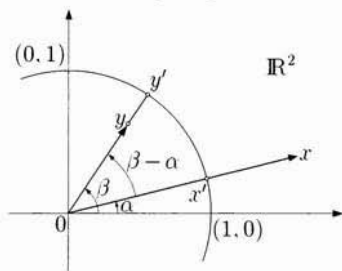


Bild 5.2

Man beachte, daß die Erklärung des Winkels mit Hilfe des Skalarprodukts formal einfacher ist als mit Hilfe eines Bogenmaßes in der Analysis. Anstatt eines Kurvenintegrals benötigt man lediglich die Potenzreihe des Arcussinus.

Offensichtlich ist $\angle(x, y) = \frac{\pi}{2}$ gleichbedeutend mit $\langle x, y \rangle = 0$. Das kann man ohne den Cosinus sehen:

Zwei Vektoren $x, y \in \mathbb{R}^n$ heißen *senkrecht*, wenn

$$\langle x, y \rangle = 0.$$

Dazu müssen die Vektoren nicht einmal von Null verschieden sein.

Die Definition (*) für den Winkel kann man auch in der Form

$$\langle x, y \rangle = \|x\| \cdot \|y\| \cdot \cos \angle(x, y) \quad (*)'$$

schreiben, im Gegensatz zu (*) ist dabei $x = 0$ oder $y = 0$ erlaubt. Die Formel (*) erlaubt eine geometrische Erklärung des Skalarproduktes als Produkt von $\|x\|$ und

$$\|y\| \cdot \cos \angle(x, y),$$

wobei der zweite Faktor gleich der senkrechten Projektion des Vektors y auf die durch x aufgespannte Gerade ist. Man beachte dabei das Vorzeichen: Ist $\vartheta = \angle(x, y)$, so gilt

$$\cos \vartheta \begin{cases} > 0 & \text{für } 0 \leq \vartheta < \frac{\pi}{2}, \\ < 0 & \text{für } \frac{\pi}{2} < \vartheta \leq \pi. \end{cases}$$

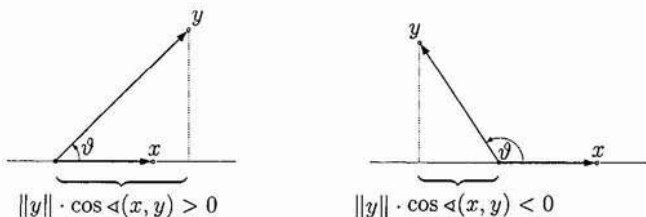


Bild 5.3

Aufgaben zu 5.1

1. Zeigen Sie, dass für alle $x, y \in \mathbb{R}^n$ gilt:

a) $\langle x + y, x - y \rangle = \|x\|^2 - \|y\|^2.$

b) $\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2\langle x, y \rangle = \|x\|^2 + \|y\|^2 - 2\|x\| \cdot \|y\| \cos \vartheta.$
(verallgemeinerter Satz von PYTHAGORAS oder Cosinussatz)

c) $\|x + y\|^2 - \|x - y\|^2 = 4\langle x, y \rangle.$

d) $\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2. \quad (\text{Parallelogramm-Gleichung})$

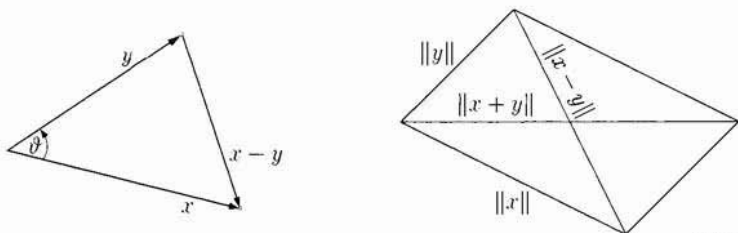


Bild 5.4

2. Beweisen Sie die Cauchy-Schwarzsche Ungleichung durch direkte Rechnung im Fall $n = 1, 2, 3$.

3. Mit Hilfe des Winkels zwischen Vektoren kann man auch den *Winkel zwischen Geraden* erklären. Sind $L = v + \mathbb{R}w$ und $L' = v' + \mathbb{R}w'$ Geraden im $\mathbb{R}^n$, so sei der Winkel zwischen L und L' erklärt durch

$$\angle(L, L') := \begin{cases} \angle(w, w') & \text{falls } \langle w, w' \rangle \geq 0, \\ \angle(-w, w') & \text{sonst.} \end{cases}$$

Zeigen Sie, dass diese Definition unabhängig von der Auswahl von w und w' ist, und dass $0 \leq \angle(L, L') \leq \frac{\pi}{2}$ gilt.

4. Zwei Vektoren $x, y \in \mathbb{R}^n$ heißen *orthogonal* (in Zeichen $x \perp y$), wenn $\langle x, y \rangle = 0$. Sind $x, y \neq 0$, so gilt offenbar

$$x \perp y \Leftrightarrow \angle(x, y) = \frac{\pi}{2}.$$

Ist $L = v + \mathbb{R}w \subset \mathbb{R}^n$ eine Gerade, so heißt $s \in \mathbb{R}^n$ orthogonal zu L , wenn $\langle s, x - y \rangle = 0$ für alle $x, y \in L$. Zeigen Sie:

a) Ist $L = v + \mathbb{R}w \subset \mathbb{R}^n$ eine Gerade und $s \in \mathbb{R}^n$, so gilt:

$$s \text{ ist orthogonal zu } L \Leftrightarrow s \perp w.$$

b) Ist $L = \{(x_1, x_2) \in \mathbb{R}^2 : a_1 x_1 + a_2 x_2 = b\}$ eine Gerade im $\mathbb{R}^2$, so ist (a_1, a_2) orthogonal zu L .

Zu einer Geraden orthogonale Vektoren kann man benutzen, um den *kürzesten Abstand zwischen einem Punkt und einer Geraden* zu bestimmen. Ist $L = v + \mathbb{R}w \subset \mathbb{R}^n$ eine Gerade und $u \in \mathbb{R}^n$, so ist der *Abstand zwischen u und L* definiert als

$$d(u, L) := \min\{\|x - u\| : x \in L\}.$$

Zeigen Sie, dass für den Abstand zwischen u und L gilt:

c) Es gibt ein eindeutig bestimmtes $x \in L$, so dass $(x - u)$ orthogonal zu L ist. Für x gilt $d(u, L) = \|x - u\|$ (d. h. der senkrechte Abstand ist der kürzeste).

Für Geraden im $\mathbb{R}^2$ kann man den Abstand von einem Punkt noch einfacher beschreiben.

Es gilt:

d) Ist $L \subset \mathbb{R}^2$ eine Gerade, $s \in \mathbb{R}^2 \setminus \{0\}$ orthogonal zu L und $v \in L$ beliebig, so ist

$$L = \{x \in \mathbb{R}^2 : \langle s, x - v \rangle = 0\}.$$

Ist $u \in \mathbb{R}^2$, so folgt aus c), dass

$$d(u, L) = \frac{|\langle s, u - v \rangle|}{\|s\|}.$$

Ist speziell $L = \{(x_1, x_2) \in \mathbb{R}^2 : a_1 x_1 + a_2 x_2 = b\}$ und $u = (u_1, u_2)$, so ergibt sich

$$d(u, L) = \frac{|a_1 u_1 + a_2 u_2 - b|}{\sqrt{a_1^2 + a_2^2}}.$$

Mit Hilfe von d) können wir nun für Gleichungen von Geraden im $\mathbb{R}^2$ die sogenannte *Hessesche Normalform* herleiten: Ist $s \in \mathbb{R}^2 \setminus \{0\}$ orthogonal zur Geraden $L \subset \mathbb{R}^2$, so sei $n := \frac{1}{\|s\|} \cdot s$. Dann ist $\|n\| = 1$. Man nennt n einen *Normalenvektor* zu L ; nach d) gilt für beliebiges $v \in L$, dass

$$L = \{x \in \mathbb{R}^2 : \langle n, x - v \rangle = 0\}.$$

Für jedes $u \in \mathbb{R}^2$ gilt dann $d(u, L) = |\langle n, u - v \rangle|$, die Funktion $\langle n, u - v \rangle$ misst also mit Vorzeichen den Abstand von u zu L .

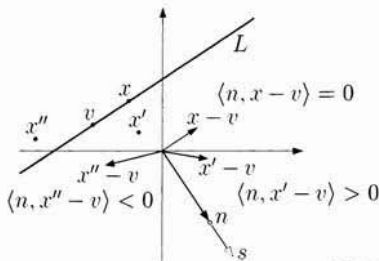
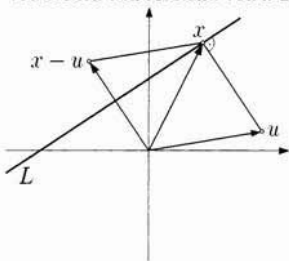


Bild 5.5

5. Aufgabe 4 lässt sich leicht verallgemeinern, um den Abstand zwischen einem Punkt und einer Hyperebene im $\mathbb{R}^n$ zu bestimmen; eine Teilmenge H des $\mathbb{R}^n$ heißt dabei *Hyperebene*, falls H ein affiner Unterraum der Dimension $(n-1)$ ist, d. h. es existiert ein $v \in \mathbb{R}^n$ und ein Untervektorraum $W \subset \mathbb{R}^n$ der Dimension $(n-1)$, so dass $H = v + W$. Ist $H = v + \text{span}(w_1, \dots, w_{n-1}) \subset \mathbb{R}^n$ eine Hyperebene, so heißt $s \in \mathbb{R}^n$ orthogonal zu H , wenn $\langle s, x - y \rangle = 0$ für alle $x, y \in H$. Zeigen Sie:

- s ist orthogonal zu $H \Leftrightarrow s \perp w_i$ für $i = 1, \dots, n-1$.
- Ist die Hyperebene gegeben durch $H = \{(x_1, \dots, x_n) \in \mathbb{R}^n : a_1 x_1 + \dots + a_n x_n = b\}$, so ist $(a_1, \dots, a_n)$ orthogonal zu H .

Ist die Hyperebene H also durch eine Gleichung gegeben, so findet man leicht einen zu H orthogonalen Vektor. Was man tun kann, falls die Ebene in Parameterdarstellung gegeben ist, wird in Aufgabe 6 zu 5.2 gezeigt.

Ist $H \subset \mathbb{R}^n$ eine Hyperebene und $u \in \mathbb{R}^n$, so ist der Abstand zwischen u und H erklärt durch

$$d(u, H) := \min\{\|x - u\| : x \in H\}.$$

Beweisen Sie:

- c) Es gibt ein eindeutig bestimmtes $x \in H$, so dass $(x - u)$ orthogonal zu H ist. Es gilt $d(u, H) = \|x - u\|$ (d. h. *der senkrechte Abstand ist der kürzeste*).
- d) Ist $H = \{(x_1, \dots, x_n) \in \mathbb{R}^n : a_1 x_1 + \dots + a_n x_n = b\}$ und $u = (u_1, \dots, u_n) \in \mathbb{R}^n$, so gilt

$$d(u, H) = \frac{|a_1 u_1 + \dots + a_n u_n - b|}{\sqrt{a_1^2 + \dots + a_n^2}}.$$

Ist N orthogonal zu H mit $\|N\| = 1$ und $v \in H$ beliebig, so leitet man wie in Aufgabe 4 die *Hessesche Normalform* der Gleichung der Hyperebene ab:

$$H = \{x \in \mathbb{R}^n : \langle N, x - v \rangle = 0\}.$$

6. Seien $\mathcal{N} \subset \mathcal{L}(\mathbb{R})$ wie in Beispiel 2b) aus 2.2.6. Betrachten Sie die Abbildungen

$$\|\cdot\| : \mathcal{L}(\mathbb{R}) \rightarrow \mathbb{R}, \quad f \mapsto \int_{\mathbb{R}} |f(t)| dt, \quad \text{und}$$

$$\|\cdot\|' : \mathcal{L}(\mathbb{R})/\mathcal{N} \rightarrow \mathbb{R}, \quad f + \mathcal{N} \mapsto \|f\|.$$

Welche davon ist eine Norm?

5.2 Das Vektorprodukt im $\mathbb{R}^3$

5.2.1. Eine ganz spezielle Multiplikation gibt es für $n = 3$:

$$\mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad (x, y) \mapsto x \times y, \quad \text{mit}$$

$$x \times y := (x_2 y_3 - x_3 y_2, x_3 y_1 - x_1 y_3, x_1 y_2 - x_2 y_1).$$

Diese Definition des *Vektorproduktes* (weil das Ergebnis wieder ein Vektor ist) kann man sich leichter merken durch die Regel

$$x \times y = \begin{vmatrix} e_1 & e_2 & e_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} = e_1 \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} - e_2 \begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix} + e_3 \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix},$$

wobei man die Determinante formal nach der ersten Zeile entwickelt. Daran sieht man auch, daß sich auf diese Weise für beliebiges n ein Produkt mit $n - 1$ Faktoren erklären läßt (Aufgabe 6).

Zunächst notieren wir wieder die formalen und ganz einfach zu beweisenden Rechenregeln. Für $x, x', y, y' \in \mathbb{R}^3$ und $\lambda \in \mathbb{R}$ gilt:

1. $(x + x') \times y = x \times y + x' \times y, \quad x \times (y + y') = x \times y + x \times y',$
 $\lambda x \times y = \lambda(x \times y), \quad x \times \lambda y = \lambda(x \times y),$
2. $y \times x = -x \times y, \quad \text{also} \quad x \times x = 0,$
3. $x \times y = 0 \Leftrightarrow x, y$ linear abhängig.

5.2.2. Den engen Zusammenhang zwischen Vektorprodukt und Skalarprodukt sieht man an den folgenden Regeln.

Bemerkung. Für $x, y, z \in \mathbb{R}^3$ gilt:

a)

$$\langle x \times y, z \rangle = \det \begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{pmatrix}, \quad \langle x \times y, x \rangle = \langle x \times y, y \rangle = 0.$$

b)

$$\|x \times y\|^2 = \|x\|^2 \|y\|^2 - \langle x, y \rangle^2, \quad \|x \times y\| = \|x\| \cdot \|y\| \cdot \sin \angle(x, y).$$

Beweis. a) folgt sofort aus den Definitionen. Eigenschaft b) ist eine Präzisierung der CAUCHY-SCHWARZschen Ungleichung und folgt sofort aus dem in 5.1.3

gegebenen Beweis für $n = 3$, da

$$\|x \times y\|^2 = \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix}^2 + \begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix}^2 + \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix}^2.$$

Die zweite Gleichung von b) folgt aus der Definition des Winkels $\vartheta = \angle(x, y)$ und

$$\|x \times y\|^2 = \|x\|^2 \|y\|^2 (1 - \cos^2 \vartheta) = \|x\|^2 \|y\|^2 \sin^2 \vartheta. \quad \square$$

Aus dieser Bemerkung kann man nun die geometrischen Eigenschaften des Vektorproduktes ablesen. Wir starten mit linear unabhängigen Vektoren $x, y \in \mathbb{R}^3$ (andernfalls ist $x \times y = 0$). Aus a) folgt, daß $x \times y$ senkrecht auf der von x und y aufgespannten Ebene steht. Also ist festgelegt, auf welcher Geraden $x \times y$ liegt.

Aus b) folgt, daß die Länge von $x \times y$ gleich der Fläche des von x und y aufgespannten Parallelogramms ist. Damit ist $x \times y$ bis auf das Vorzeichen festgelegt. Welches man zu wählen hat, ist eine Frage der Orientierung. Dazu betrachten wir die Basis

$$\mathcal{B} = (x, y, z = x \times y)$$

von $\mathbb{R}^3$. Nach Eigenschaft a) ist die Determinante gleich $\langle x \times y, x \times y \rangle$, also positiv, und somit sind $\mathcal{B}$ und die kanonische Basis gleichorientiert (vgl. 3.4.3). Also hat $x \times y$ zu x und y die gleiche Richtung wie e_3 zu e_1 und e_2 (vgl. Bild 5.6).

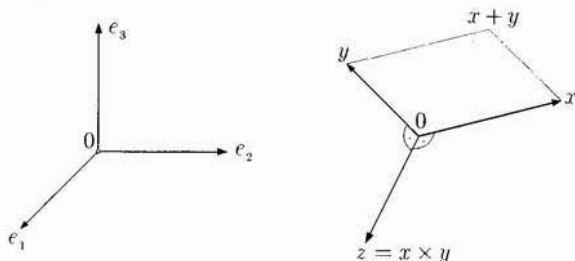


Bild 5.6

Außerdem folgt aus Teil b) der Bemerkung, daß ein Parallelogramm bei vorgegebenen Seitenlängen dann die größte Fläche hat, wenn es ein Rechteck ist. Eine Verallgemeinerung davon beweisen wir in 5.4.10.

Aufgaben zu 5.2

1. Zeigen Sie für $x, y, z \in \mathbb{R}^3$ die Grassmann-Identität

$$x \times (y \times z) = \langle x, z \rangle y - \langle x, y \rangle z$$

und folgern daraus die Jacobi-Identität

$$x \times (y \times z) + y \times (z \times x) + z \times (x \times y) = 0.$$

2. Für $x, x', y, y' \in \mathbb{R}^3$ gilt:

$$\text{a) } (x \times y) \times (x' \times y') = x' \cdot \det \begin{pmatrix} x_1 & y_1 & y'_1 \\ x_2 & y_2 & y'_2 \\ x_3 & y_3 & y'_3 \end{pmatrix} - y' \cdot \det \begin{pmatrix} x_1 & y_1 & x'_1 \\ x_2 & y_2 & x'_2 \\ x_3 & y_3 & x'_3 \end{pmatrix}.$$

$$\text{b) } \langle x \times y, x' \times y' \rangle = \langle x, x' \rangle \langle y, y' \rangle - \langle y, x' \rangle \langle x, y' \rangle.$$

3. Seien $x, y, z \in \mathbb{R}^3$. Dann gilt:

$$x, y, z \text{ sind linear unabhängig} \Leftrightarrow x \times y, y \times z, z \times x \text{ sind linear unabhängig}.$$

4. Gegeben sei eine Ebene $E = v + \mathbb{R}w_1 + \mathbb{R}w_2 \subset \mathbb{R}^3$. Zeigen Sie: Setzt man $a := w_1 \times w_2$ und $b := \langle v, a \rangle$, so gilt

$$E = \{x \in \mathbb{R}^3 : \langle x, a \rangle = b\}.$$

5. Wir wollen mit Hilfe des Vektorproduktes eine Parameterdarstellung der Schnittgeraden zweier nichtparalleler Ebenen im $\mathbb{R}^3$ bestimmen. Sind zwei Ebenen $E = v + \mathbb{R}w_1 + \mathbb{R}w_2$, $E' = v' + \mathbb{R}w'_1 + \mathbb{R}w'_2 \subset \mathbb{R}^3$ gegeben, so sei $W = \mathbb{R}w_1 + \mathbb{R}w_2$, $W' = \mathbb{R}w'_1 + \mathbb{R}w'_2$. Da die beiden Ebenen nicht parallel sind, ist $W \neq W'$, und damit hat $U = W \cap W'$ die Dimension 1. Zeigen Sie:

a) Ist $L = E \cap E'$ und $u \in L$, so ist $L = u + U$.

b) Seien $s = w_1 \times w_2$, $s' = w'_1 \times w'_2$ und $w = s \times s'$. Dann gilt $U = \mathbb{R}w$.

Bestimmen Sie nach diesem Verfahren eine Parameterdarstellung von $E \cap E'$, wobei

$$E = (0, 2, 3) + \mathbb{R}(3, 6, 5) + \mathbb{R}(1, 7, -1),$$

$$E' = (-1, 3, 2) + \mathbb{R}(8, 2, 3) + \mathbb{R}(2, -1, -2).$$

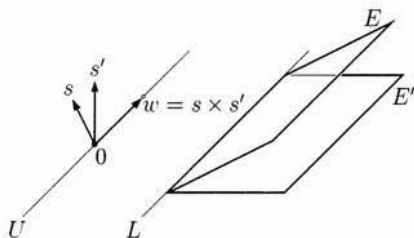


Bild 5.7

6. Das Vektorprodukt zweier Vektoren im $\mathbb{R}^3$ lässt sich für $n \geq 3$ folgendermaßen zu einem Produkt von $n - 1$ Vektoren im $\mathbb{R}^n$ verallgemeinern: Sind $x^{(1)}, \dots, x^{(n-1)} \in \mathbb{R}^n$, so sei

$$x^{(1)} \times \dots \times x^{(n-1)} := \sum_{i=1}^n (-1)^{i+1} (\det A_i) \cdot e_i,$$

wobei $A \in M((n-1) \times n; \mathbb{R})$ die Matrix ist, die aus den Zeilen $x^{(1)}, \dots, x^{(n-1)}$ besteht und A_i aus A durch Streichen der i -ten Spalte entsteht. Wie im Fall $n = 3$ entsteht $x^{(1)} \times \dots \times x^{(n-1)}$ also durch formales Entwickeln von

$$\det \begin{pmatrix} e_1 & e_2 & \dots & e_n \\ x_1^{(1)} & x_2^{(1)} & \dots & x_n^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{(n-1)} & x_2^{(n-1)} & \dots & x_n^{(n-1)} \end{pmatrix}$$

nach der ersten Zeile. Zeigen Sie, dass für das verallgemeinerte Vektorprodukt gilt:

$$\begin{aligned} \text{a) } & x^{(1)} \times \dots \times x^{(i-1)} \times (x + y) \times x^{(i+1)} \times \dots \times x^{(n-1)} = \\ & x^{(1)} \times \dots \times x^{(i-1)} \times x \times x^{(i+1)} \times \dots \times x^{(n-1)} + \\ & x^{(1)} \times \dots \times x^{(i-1)} \times y \times x^{(i+1)} \times \dots \times x^{(n-1)}, \\ & x^{(1)} \times \dots \times x^{(i-1)} \times (\lambda x) \times x^{(i+1)} \times \dots \times x^{(n-1)} = \\ & \lambda (x^{(1)} \times \dots \times x^{(i-1)} \times x \times x^{(i+1)} \times \dots \times x^{(n-1)}). \end{aligned}$$

$$\text{b) } x^{(1)} \times \dots \times x^{(n-1)} = 0 \Leftrightarrow x^{(1)}, \dots, x^{(n-1)} \text{ linear abhängig.}$$

$$\text{c) } \langle x^{(1)} \times \dots \times x^{(n-1)}, y \rangle = \det \begin{pmatrix} y_1 & y_2 & \dots & y_n \\ x_1^{(1)} & x_2^{(1)} & \dots & x_n^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{(n-1)} & x_2^{(n-1)} & \dots & x_n^{(n-1)} \end{pmatrix},$$

$$\text{d) } \langle x^{(1)} \times \dots \times x^{(n-1)}, x^{(i)} \rangle = 0, \text{ für } i = 1, \dots, n-1.$$

5.3 Das kanonische Skalarprodukt im $\mathbb{C}^n$

5.3.1. Zu einer komplexen Zahl $z = x + iy$ hatten wir in 1.3.4

$$\bar{z} = x - iy \quad \text{und} \quad |z| = \sqrt{z\bar{z}} = \sqrt{x^2 + y^2}$$

erklärt. Der Zahl $z \in \mathbb{C}$ entspricht der Vektor $(x, y) \in \mathbb{R}^2$, und mit der im $\mathbb{R}^2$ definierten Norm ist

$$|z| = \|(x, y)\|.$$

Für die Geometrie des $\mathbb{C}^n$ benutzt man die Abbildung

$$\mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{C}, \quad (z, w) \mapsto \langle z, w \rangle_c, \quad \text{mit}$$

$$\langle z, w \rangle_c := z_1 \bar{w}_1 + \dots + z_n \bar{w}_n \quad \text{für } z = (z_1, \dots, z_n), \quad w = (w_1, \dots, w_n).$$

Sie heißt *kanonisches Skalarprodukt*. Der Index c wird nur vorübergehend angehängt, um auf den Unterschied zum entsprechenden Skalarprodukt im $\mathbb{R}^n$ hinzuweisen. Ohne jede Mühe beweist man die folgenden Rechenregeln.

1. $\langle z + z', w \rangle_c = \langle z, w \rangle_c + \langle z', w \rangle_c, \quad \langle z, w + w' \rangle_c = \langle z, w \rangle_c + \langle z, w' \rangle_c,$
 $\langle \lambda z, w \rangle_c = \lambda \langle z, w \rangle_c, \quad \langle z, \lambda w \rangle_c = \bar{\lambda} \langle z, w \rangle_c,$
2. $\langle w, z \rangle_c = \overline{\langle z, w \rangle_c},$
3. $\langle z, z \rangle_c \in \mathbb{R}_+ \quad \text{und} \quad \langle z, z \rangle_c = 0 \Leftrightarrow z = 0$

für $z, z', w, w' \in \mathbb{C}^n$ und $\lambda \in \mathbb{C}$. Wegen der Eigenschaft 3. erhält man eine Norm

$$\mathbb{C}^n \rightarrow \mathbb{R}_+, \quad z \mapsto \|z\| := \sqrt{\langle z, z \rangle_c}.$$

5.3.2. Besonders wichtig ist die Beziehung zum kanonischen Skalarprodukt im $\mathbb{R}^n$. Bezüglich der natürlichen Inklusion

$$\mathbb{R} \subset \mathbb{C}, \quad x = x + i \cdot 0,$$

gibt es kein Problem: entsprechend ist $\mathbb{R}^n \subset \mathbb{C}^n$, und $\langle \cdot, \cdot \rangle_c$ ist eine Fortsetzung von $\langle \cdot, \cdot \rangle$. Anders ist es bei der Abbildung

$$\mathbb{R}^{2n} \rightarrow \mathbb{C}^n, \quad v = (x_1, y_1, \dots, x_n, y_n) \mapsto (x_1 + iy_1, \dots, x_n + iy_n) = z.$$

Ist $v' \in \mathbb{R}^{2n}$ ein anderer Vektor, dem $z' \in \mathbb{C}^n$ entspricht, so ist

$$\begin{aligned} \langle z, z' \rangle_c &= \sum_{v=1}^n z_v \bar{z}'_v = \sum_{v=1}^n (x_v x'_v + y_v y'_v) - i \sum_{v=1}^n \det \begin{pmatrix} x_v & y_v \\ x'_v & y'_v \end{pmatrix} \\ &= \langle v, v' \rangle - i \omega(v, v'), \end{aligned}$$

wobei $\langle \cdot, \cdot \rangle$ das Skalarprodukt im $\mathbb{R}^{2n}$ bezeichnet und

$$\omega: \mathbb{R}^{2n} \times \mathbb{R}^{2n} \rightarrow \mathbb{R}$$

durch die darüberstehende Summe erklärt ist. Offensichtlich ist

$$\omega(v', v) = -\omega(v, v').$$

Das alte Skalarprodukt im $\mathbb{R}^{2n}$ ist also der Realteil des Skalarproduktes im $\mathbb{C}^n$, dazu kommt ein alternierender Imaginärteil. Da er für $v = v'$ verschwindet, erhält man auf $\mathbb{R}^{2n}$ und $\mathbb{C}^n$ die gleiche Norm.

Aufgaben zu 5.3

1. Zeigen Sie, dass die schiefsymmetrische Bilinearform (vgl. 5.4.1) $\omega: \mathbb{R}^{2n} \times \mathbb{R}^{2n} \rightarrow \mathbb{R}$ aus 5.3.2 nicht-entartet ist, d. h.: Ist $\omega(v, w) = 0$ für alle $w \in \mathbb{R}^{2n}$, so ist $v = 0$.

2. Sei $J: \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n}$ der Endomorphismus, der gegeben ist durch

$$J(x_1, y_1, \dots, x_n, y_n) = (-y_1, x_1, \dots, -y_n, x_n).$$

(Identifiziert man $\mathbb{R}^{2n}$ mit $\mathbb{C}^n$ wie in 5.3.2, so ist J einfach die Multiplikation mit i .) Zeigen Sie, dass für das kanonische Skalarprodukt $\langle \cdot, \cdot \rangle$ im $\mathbb{R}^{2n}$, die Abbildung ω aus 5.3.2 und J der folgende Zusammenhang besteht:

$$\text{Für alle } v, w \in \mathbb{R}^{2n} \text{ ist } \langle v, w \rangle = \omega(v, J(w)).$$

3. Eine komplexe Struktur auf einem $\mathbb{R}$ -Vektorraum V ist ein Endomorphismus J von V mit $J^2 = -\text{id}$. Zeigen Sie:

- Mit der skalaren Multiplikation $(x + iy) \cdot v := xv + yJ(v)$ wird V zu einem $\mathbb{C}$ -Vektorraum.
- Ist V endlichdimensional, so ist $\dim_{\mathbb{R}} V$ gerade.

5.4 Bilinearformen und Sesquilinearformen

Nachdem wir die wichtigsten Beispiele beschrieben haben, werden nun die üblichen allgemeinen Begriffe eingeführt.

5.4.1. Zunächst darf K wieder ein allgemeiner Körper sein. Ist V ein Vektorraum darüber, so betrachten wir eine Abbildung

$$s: V \times V \rightarrow K, \quad (v, w) \mapsto s(v, w).$$

Sie heißt *Bilinearform auf V* , wenn folgendes gilt:

$$\mathbf{B1} \quad s(v + v', w) = s(v, w) + s(v', w), \quad s(\lambda v, w) = \lambda s(v, w),$$

$$\mathbf{B2} \quad s(v, w + w') = s(v, w) + s(v, w'), \quad s(v, \lambda w) = \lambda s(v, w).$$

Die Abbildung s heißt *symmetrisch*, falls

$$\mathbf{S} \quad s(v, w) = s(w, v),$$

und *alternierend* (oder *schiefsymmetrisch*), falls

$$\mathbf{A} \quad s(w, v) = -s(v, w).$$

Dabei ist jeweils $v, v', w, w' \in V$ und $\lambda \in K$.

Beispiel. Sei $K = \mathbb{R}$, $I = [a, b] \subset \mathbb{R}$ ein Intervall und $V = \mathcal{C}(I; \mathbb{R})$ der Vektorraum der darauf stetigen Funktionen. Da jede stetige Funktion integrierbar ist, hat man eine Abbildung

$$V \times V \rightarrow \mathbb{R}, \quad (f, g) \mapsto \int_a^b f(t)g(t) dt,$$

und diese ist nach den Rechenregeln für Integrale eine symmetrische Bilinearform.

5.4.2. Ist V im Gegensatz zu obigem Beispiel endlich-dimensional, so kann man Bilinearformen durch Matrizen beschreiben. Sei also s eine Bilinearform und $\mathcal{B} = (v_1, \dots, v_n)$ eine Basis von V . Wir betrachten die Matrix

$$M_{\mathcal{B}}(s) := (s(v_i, v_j))_{ij} \in M(n \times n; K).$$

Wie bei linearen Abbildungen nennt man sie die *darstellende Matrix* von s bezüglich $\mathcal{B}$. Durch sie ist s vollständig festgelegt. Genauer gilt:

Bemerkung. Sei s eine Bilinearform auf V mit Basis $\mathcal{B}$, $\Phi_{\mathcal{B}}: K^n \rightarrow V$ das zugehörige Koordinatensystem. Wir betrachten die Matrix $A = M_{\mathcal{B}}(s)$ und für

$v, w \in V$ die Koordinaten $x = \Phi_B^{-1}(v)$, $y = \Phi_B^{-1}(w)$. Dann ist

$$s(v, w) = {}^t x A y = (x_1, \dots, x_n) \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Beweis. Es ist

$$s(v, w) = s(x_1 v_1 + \dots + x_n v_n, y_1 v_1 + \dots + y_n v_n),$$

also folgt die Behauptung durch wiederholte Anwendung von B1 und B2, weil $a_{ij} = s(v_i, v_j)$. Das Ergebnis kann man auch als Doppelsumme mit n^2 Summanden schreiben:

$$s(v, w) = \sum_{i,j=1}^n a_{ij} x_i y_j. \quad \square$$

Ist umgekehrt auf V mit Basis B eine beliebige quadratische Matrix $A = (a_{ij})$ gegeben, so ist durch die obige Formel eine Bilinearform s mit

$$s(v_i, v_j) = a_{ij}$$

erklärt. Das folgt etwa aus den Rechenregeln für die Matrizenmultiplikation. Die Symmetrie von s ist gleichbedeutend mit der Symmetrie von A . Zusammenfassend gilt:

Satz. Sei V ein K -Vektorraum mit $n = \dim V < \infty$ und B eine Basis. Dann ist die Abbildung

$$s \mapsto M_B(s)$$

von den Bilinearformen auf V in $M(n \times n; K)$ bijektiv, und s ist genau dann symmetrisch, wenn $M_B(s)$ symmetrisch ist.

Daß zu verschiedenen Matrizen verschiedene Bilinearformen gehören, kann man in einer Art von Kürzungsregel ausdrücken:

Lemma. Gegeben seien $A, B \in M(n \times n; K)$ derart, daß

$${}^t x A y = {}^t x B y$$

für alle Spaltenvektoren $x, y \in K^n$. Dann ist $A = B$.

Beweis. Es genügt, für x und y die kanonische Basis einzusetzen und zu bemerken, daß

$${}^t e_i A e_j = a_{ij}. \quad \square$$

5.4.3. Die Beschreibung einer Bilinearform durch eine Matrix hat Ähnlichkeit mit der entsprechenden Beschreibung eines Endomorphismus (2.4.3). Ein wesentlicher Unterschied zeigt sich im Transformationsverhalten.

Transformationsformel. Sei V ein endlichdimensionaler Vektorraum mit Basen $\mathcal{A}$, $\mathcal{B}$, und sei $T_{\mathcal{A}}^{\mathcal{B}}$ die entsprechende Transformationsmatrix (2.6.2). Für jede Bilinearform s auf V gilt dann

$$M_{\mathcal{B}}(s) = {}^t T_{\mathcal{A}}^{\mathcal{B}} \cdot M_{\mathcal{A}}(s) \cdot T_{\mathcal{A}}^{\mathcal{B}}.$$

Man beachte, daß dagegen für einen Endomorphismus F von V nach 2.6.5 gilt:

$$M_{\mathcal{B}}(F) = T_{\mathcal{B}}^{\mathcal{A}} \cdot M_{\mathcal{A}}(F) \cdot T_{\mathcal{A}}^{\mathcal{B}}, \quad \text{wobei} \quad T_{\mathcal{B}}^{\mathcal{A}} = (T_{\mathcal{A}}^{\mathcal{B}})^{-1}.$$

Beweis. Seien $v, v' \in V$ mit $v = \Phi_{\mathcal{A}}(x) = \Phi_{\mathcal{B}}(y)$ und $v' = \Phi_{\mathcal{A}}(x') = \Phi_{\mathcal{B}}(y')$. Ist $T := T_{\mathcal{A}}^{\mathcal{B}}$, so ist $x = Ty$ und $x' = Ty'$. Setzen wir weiter $A := M_{\mathcal{A}}(s)$ und $B := M_{\mathcal{B}}(s)$, so folgt

$${}^t y B y' = s(v, v') = {}^t x A x' = {}^t (Ty) A (Ty') = {}^t y ({}^t T A T) y'.$$

Da dies für alle y und y' gilt, folgt die Behauptung $B = {}^t T A T$ aus Lemma 5.4.2. $\square$

Ist $S = T_{\mathcal{B}}^{\mathcal{A}} = T^{-1}$, so können wir die beiden Transformationsformeln auch in der folgenden Form vergleichen:

$$\begin{aligned} A &= {}^t S B S && \text{für eine Bilinearform und} \\ B &= S A S^{-1}, \quad \text{d.h.} \quad A = S^{-1} B S && \text{für einen Endomorphismus.} \end{aligned}$$

5.4.4. Ist $s: V \times V \rightarrow K$ eine symmetrische Bilinearform, so erhält man daraus eine Abbildung

$$q: V \rightarrow K, \quad v \mapsto q(v) := s(v, v).$$

Sie heißt die zu s gehörige *quadratische Form*. Da s bilinear ist, folgt

$$q(\lambda v) = \lambda^2 q(v)$$

für $\lambda \in K$. Ist insbesondere $V = K^n$ und s durch eine symmetrische Matrix $A = (a_{ij})$ gegeben, so ist

$$q(x_1, \dots, x_n) = \sum_{i,j=1}^n a_{ij} x_i x_j = \sum_{i=1}^n a_{ii} x_i^2 + \sum_{1 \leq i < j \leq n} 2a_{ij} x_i x_j,$$

das ist ein *homogenes quadratisches Polynom* in den Unbestimmten $x_1, \dots, x_n$.

Es ist nicht selbstverständlich, daß man die ursprüngliche Form s aus q rekonstruieren kann. Diesen Kniff nennt man

Polarisierung. Ist $\text{char}(K) \neq 2$, so gilt für jede symmetrische Bilinearform s und die zugehörige quadratische Form q auf V

$$s(v, w) = \frac{1}{2} (q(v+w) - q(v) - q(w)) .$$

Das rechnet man ohne Schwierigkeit nach. □

5.4.5. Für die Längenmessung im $\mathbb{C}^n$ benötigt man die komplexe Konjugation, das ist eine Abbildung

$$\mathbb{C} \rightarrow \mathbb{C}, \quad z \mapsto \bar{z},$$

die ein Körperhomomorphismus ist, und die zweimal angewandt die Identität ist.

Ist V ein komplexer Vektorraum, so heißt eine Abbildung $F: V \rightarrow V$ *semi-linear* (d.h. halb oder halbwegs linear), wenn

$$F(v+w) = F(v) + F(w) \quad \text{und} \quad F(\lambda v) = \bar{\lambda} F(v)$$

für $v, w \in V$ und $\lambda \in \mathbb{C}$. Entsprechend nennt man eine Abbildung

$$s: V \times V \rightarrow \mathbb{C}$$

sesquilinear (d.h. $1\frac{1}{2}$ -fach linear), wenn s im ersten Argument linear und im zweiten Argument semilinear ist, d.h.

$$\mathbf{B1} \quad s(v+v', w) = s(v, w) + s(v', w), \quad s(\lambda v, w) = \lambda s(v, w),$$

$$\mathbf{\bar{B}2} \quad s(v, w+w') = s(v, w) + s(v, w'), \quad s(v, \lambda w) = \bar{\lambda} s(v, w),$$

und *hermitesch*, wenn zusätzlich

$$\mathbf{H} \quad s(w, v) = \overline{s(v, w)}.$$

Dabei ist jeweils $v, v', w, w' \in V$ und $\lambda \in \mathbb{C}$. Standardbeispiel dafür ist das kanonische Skalarprodukt im $\mathbb{C}^n$.

Wie eine Bilinearform kann man auch eine Sesquilinearform durch eine Matrix beschreiben, indem man an der richtigen Stelle einen Querstrich einfügt. Wir notieren nur die Ergebnisse: Ist $\mathcal{A}$ Basis von V und

$$A = M_{\mathcal{A}}(s) = (s(v_i, v_j)), \quad v = \Phi_{\mathcal{A}}(x) \quad \text{und} \quad w = \Phi_{\mathcal{A}}(y),$$

so ist

$$s(v, w) = {}^t x A \bar{y}.$$

Ist $\mathcal{B}$ eine weitere Basis, $B = M_{\mathcal{B}}(s)$ und $T = T_{\mathcal{A}}^{\mathcal{B}}$, so hat man die *Transformationsformel*

$$B = {}^t T A \bar{T}.$$

Weiter ist die Sesquilinearform s genau dann hermitesch, wenn die Matrix $A = M_A(s)$ hermitesch ist, d.h.

$${}^t A = \bar{A}.$$

Schließlich hat man auch im Komplexen für Sesquilinearformen eine *Polarisierung*

$$s(v, w) = \frac{1}{4} (q(v+w) - q(v-w) + iq(v+iw) - iq(v-iw)).$$

5.4.6. Symmetrische Bilinearformen oder hermitesche Formen sind viel allgemeiner als die Standardbeispiele aus 5.1 und 5.3, denn zum Beispiel die Nullabbildung s (d.h. $s(v, w) = 0$ für alle v, w) erfüllt trivialerweise all die angegebenen Bedingungen. Zur Formulierung einer starken Qualitätsbedingung beschränken wir uns auf den reellen und komplexen Fall. Dazu ist es vorteilhaft, den Buchstaben $\mathbb{K}$ für

$$\mathbb{K} = \mathbb{R} \quad \text{oder} \quad \mathbb{K} = \mathbb{C}$$

zu verwenden. Sei also V ein $\mathbb{K}$ -Vektorraum und

$$s: V \times V \rightarrow \mathbb{K}$$

eine symmetrische Bilinearform (bzw. hermitesche Form). Dann heißt s *positiv definit*, wenn

$$s(v, v) > 0 \quad \text{für jedes } v \in V \text{ mit } v \neq 0.$$

Man beachte dabei, daß auch im hermiteschen Fall $s(v, v) \in \mathbb{R}$ ist.

Eine symmetrische (bzw. hermitesche) Matrix A heißt *positiv definit*, wenn

$${}^t x A \bar{x} > 0 \quad \text{für jeden Spaltenvektor } x \neq 0 \text{ aus } \mathbb{K}^n.$$

In 5.7.3 geben wir relativ leicht nachprüfbare Bedingungen dafür an.

Vorsicht! Es kann

$$s(v_i, v_i) > 0$$

für alle Vektoren v_i einer Basis sein, ohne daß s positiv definit ist. Man betrachte als Beispiel

$$s: \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (x_1, x_2, y_1, y_2) \mapsto x_1 y_1 - x_2 y_2,$$

und überlege sich, welches Vorzeichen $s(v, v)$ in Abhängigkeit von v hat.

Zur Abkürzung nennt man eine positiv definite symmetrische Bilinearform bzw. hermitesche Form ein *Skalarprodukt* und einen reellen bzw. komplexen Vektorraum zusammen mit einem Skalarprodukt einen *euklidischen* bzw. *unitären* Vektorraum.

In den folgenden Abschnitten wird sich zeigen, wie sich diese zusätzliche Struktur benutzen läßt, um neue Ergebnisse über die Diagonalisierung zu beweisen.

Die Standardräume $\mathbb{R}^n$ bzw. $\mathbb{C}^n$ haben neben der natürlichen Vektorraumstruktur auch eine natürliche euklidische bzw. unitäre Struktur in sich (vgl. 5.1 und 5.3). Man kann sich also fragen, was die vielen abstrakten Begriffe nützen. Eine Rechtfertigung besteht darin, daß die abstrakten Bedingungen oft einfacher zu handhaben und zu verstehen sind, als wenn man immer nur mit n -Tupeln von Zahlen rechnet. So wie im Wald, den man oft wegen der vielen Bäume nicht sieht.

5.4.7. Sei V ein euklidischer bzw. unitärer Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$. Wir definieren eine Norm

$$\|v\| = \sqrt{\langle v, v \rangle}.$$

Um zu zeigen, daß sie die Eigenschaften einer Norm aus 5.1.2 hat, benötigt man wieder die

Ungleichung von CAUCHY-SCHWARZ. Ist V ein euklidischer bzw. unitärer Vektorraum, so gilt für alle $v, w \in V$

$$|\langle v, w \rangle| \leq \|v\| \cdot \|w\|,$$

und die Gleichheit gilt genau dann, wenn v und w linear abhängig sind.

Beweis. Für alle $\lambda, \mu \in \mathbb{K}$ gilt

$$0 \leq \langle \lambda v + \mu w, \lambda v + \mu w \rangle = \lambda \bar{\lambda} \langle v, v \rangle + \mu \bar{\mu} \langle w, w \rangle + \lambda \bar{\mu} \langle v, w \rangle + \mu \bar{\lambda} \langle w, v \rangle.$$

Ist $w \neq 0$, so kann man $\lambda := \langle w, w \rangle > 0$ einsetzen und durch λ dividieren, also ist

$$0 \leq \langle v, v \rangle \langle w, w \rangle + \mu \bar{\mu} \langle v, w \rangle + \mu \langle w, v \rangle.$$

Setzt man weiter $\mu := -\langle v, w \rangle$, so folgt

$$0 \leq \langle v, v \rangle \langle w, w \rangle - \langle v, w \rangle \overline{\langle v, w \rangle} = \|v\|^2 \cdot \|w\|^2 - |\langle v, w \rangle|^2.$$

Durch Wurzelziehen folgt die Ungleichung, falls $w \neq 0$. Für $w = 0$ lautet sie $0 = 0$.

Um festzustellen, wann die Gleichung

$$|\langle v, w \rangle| = \|v\| \cdot \|w\| \tag{*}$$

gilt, können wir wieder $w \neq 0$ annehmen. Ist $v = \alpha \cdot w$, so ist

$$|\langle v, w \rangle| = |\alpha| \cdot \langle w, w \rangle = |\alpha| \cdot \|w\| \cdot \|w\| = \|\alpha \cdot w\| \cdot \|w\|,$$

also folgt (*). Ist (*) erfüllt, so folgt mit λ und μ wie oben

$$0 = \langle \lambda v + \mu w, \lambda v + \mu w \rangle, \quad \text{also } \lambda v + \mu w = 0. \quad \square$$

Der obige Beweis benutzt nur die abstrakten Eigenschaften des Skalarproduktes und ist einfacher als der in 5.1.3 gegebene. Dafür konnten wir dort die Differenz

der Quadrate der beiden Seiten von CAUCHY-SCHWARZ explizit als Summe von Quadraten darstellen, was zum Beispiel in 5.2.2 nützlich gewesen war.

Als Folgerung aus der CAUCHY-SCHWARZschen Ungleichung erhalten wir, daß in einem euklidischen oder unitären Vektorraum die oben definierte *Norm* und die daraus erhaltene *Metrik*

$$d(v, w) := \|w - v\|$$

die in 5.1.2 angegebenen Eigenschaften haben. Die Beweise verlaufen wie dort. Man beachte, daß umgekehrt auf einem $\mathbb{K}$ -Vektorraum nicht jede Metrik wie oben aus einer Norm und nicht jede Norm aus einem Skalarprodukt entsteht (vgl. Aufgabe 5).

5.4.8. In einem euklidischen Vektorraum kann man wie in 5.1.4 mit Hilfe der CAUCHY-SCHWARZschen Ungleichung eine Winkelmessung erklären. Im folgenden sind wir in erster Linie daran interessiert, wann zwei Vektoren senkrecht stehen. Das kann man auch im komplexen Fall erklären.

Definition. Sei V ein euklidischer bzw. unitärer Vektorraum.

a) Zwei Vektoren $v, w \in V$ heißen *orthogonal*, in Zeichen

$$v \perp w : \Leftrightarrow \langle v, w \rangle = 0.$$

b) Zwei Untervektorräume $U, W \subset V$ heißen *orthogonal*, in Zeichen

$$U \perp W : \Leftrightarrow u \perp w \quad \text{für alle } u \in U \text{ und alle } w \in W.$$

c) Ist $U \subset V$ ein Untervektorraum, so definiert man sein *orthogonales Komplement*

$$U^\perp := \{v \in V : v \perp u \quad \text{für alle } u \in U\}.$$

Das ist wieder ein Untervektorraum.

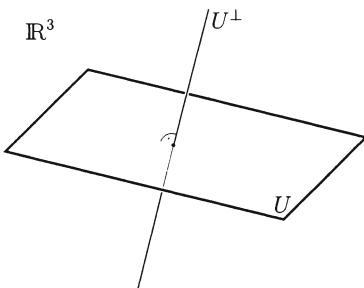


Bild 5.8

d) Eine Familie $(v_1, \dots, v_n)$ in V heißt *orthogonal*, wenn

$$v_i \perp v_j \quad \text{für alle } i \neq j.$$

Sie heit *orthonormal*, falls zustzlich

$$\|v_i\| = 1 \quad \text{fr alle } i,$$

und *Orthonormalbasis*, falls sie auch eine Basis ist, d.h. eine Basis mit

$$\langle v_i, v_j \rangle = \delta_{ij}.$$

e) Ist $V = V_1 \oplus \dots \oplus V_k$, so heit die direkte Summe *orthogonal*, in Zeichen

$$V = V_1 \oplus \dots \oplus V_k, \quad \text{falls } V_i \perp V_j \quad \text{fr alle } i \neq j.$$

Bemerkung 1. Ist $(v_1, \dots, v_n)$ eine orthogonale Familie in V und $v_i \neq 0$ fr alle i , so gilt:

a) Die Familie $(\alpha_1 v_1, \dots, \alpha_n v_n)$ mit $\alpha_i := \|v_i\|^{-1}$ ist orthonormal.

b) $(v_1, \dots, v_n)$ ist linear unabhngig.

Beweis. a) Aus

$$\langle \alpha_i v_i, \alpha_j v_j \rangle = \alpha_i \bar{\alpha}_j \langle v_i, v_j \rangle$$

folgt fr $i \neq j$, da die Vektoren orthogonal sind, und fr $i = j$ gilt wegen $\alpha_i \in \mathbb{R}$

$$\langle \alpha_i v_i, \alpha_i v_i \rangle = \alpha_i \bar{\alpha}_i \langle v_i, v_i \rangle = \alpha_i^2 \|v_i\|^2 = 1.$$

b) Bildet man das Skalarprodukt von beiden Seiten der Gleichung

$$\lambda_1 v_1 + \dots + \lambda_n v_n = 0$$

mit v_i , so folgt $\lambda_i \langle v_i, v_i \rangle = 0$, also $\lambda_i = 0$, da $\langle v_i, v_i \rangle \neq 0$. Das entspricht ganz der Vorstellung, da ein senkrecht stehender Vektor unabhngig ist. $\square$

Eine ganz einfache, aber wichtige Eigenschaft einer Orthonormalbasis ist die folgende:

Bemerkung 2. Sei $(v_1, \dots, v_n)$ eine Orthonormalbasis von V und $v \in V$ beliebig. Setzt man

$$\lambda_i := \langle v, v_i \rangle, \quad \text{so ist} \quad v = \lambda_1 v_1 + \dots + \lambda_n v_n.$$

Beweis. Die λ_i sind eindeutig bestimmt, also gengt es, die rechte Seite skalar mit v_i zu multiplizieren:

$$\langle v, v_i \rangle = \lambda_1 \langle v_1, v_i \rangle + \dots + \lambda_n \langle v_n, v_i \rangle = \lambda_i. \quad \square$$

5.4.9. Im $\mathbb{R}^n$ oder $\mathbb{C}^n$ mit dem kanonischen Skalarprodukt ist die kanonische Basis orthonormal. Wenn eine gewisse geometrische Situation vorgegeben ist,

kann es nützlich sein, eine Orthonormalbasis daran anzupassen. Das führt zu der Frage, wie viele Möglichkeiten es gibt, eine Orthonormalbasis zu finden. Eine konstruktive Antwort gibt der auf J. GRAM und E. SCHMIDT zurückgehende

Orthonormalisierungssatz. Sei V ein endlichdimensionaler euklidischer bzw. unitärer Vektorraum und $W \subset V$ ein Untervektorraum mit Orthonormalbasis $(w_1, \dots, w_m)$. Dann gibt es eine Ergänzung zu einer Orthonormalbasis

$$(w_1, \dots, w_m, w_{m+1}, \dots, w_n) \quad \text{von } V.$$

Da der Fall $W = 0$ erlaubt ist, folgt sofort das

Korollar 1. Jeder endlichdimensionale euklidische bzw. unitäre Vektorraum besitzt eine Orthonormalbasis. $\square$

Die nach der Aussage des Satzes existierenden Vektoren $w_{m+1}, \dots, w_n$ stehen senkrecht auf W , also ist

$$W' := \text{span}(w_{m+1}, \dots, w_n) \subset W^\perp.$$

Ist umgekehrt

$$w = \lambda_1 w_1 + \dots + \lambda_m w_m + \lambda_{m+1} w_{m+1} + \dots + \lambda_n w_n \in W^\perp,$$

so folgt durch skalare Multiplikation dieser Gleichung von rechts mit w_i

$$0 = \langle w, w_i \rangle = \lambda_i \quad \text{für } i = 1, \dots, m,$$

also $w \in W'$, und insgesamt

$$W^\perp = \text{span}(w_{m+1}, \dots, w_n).$$

Insbesondere folgt mit Hilfe von 1.6.3 und der Notation aus 5.4.8 das

Korollar 2. Ist W Untervektorraum eines euklidischen bzw. unitären Vektorraumes V , so gilt

$$V = W \oplus W^\perp \quad \text{und} \quad \dim V = \dim W + \dim W^\perp. \quad \square$$

Beweis des Orthonormalisierungssatzes. Ist $W = V$, so ist nichts mehr zu tun. Andernfalls gibt es einen Vektor $v \in V \setminus W$, und wir definieren

$$\tilde{v} := \langle v, w_1 \rangle w_1 + \dots + \langle v, w_m \rangle w_m.$$

Dies nennt man die *senkrechte Projektion* von v auf W , denn setzt man

$$w := v - \tilde{v}, \quad \text{so ist} \quad w \perp W.$$

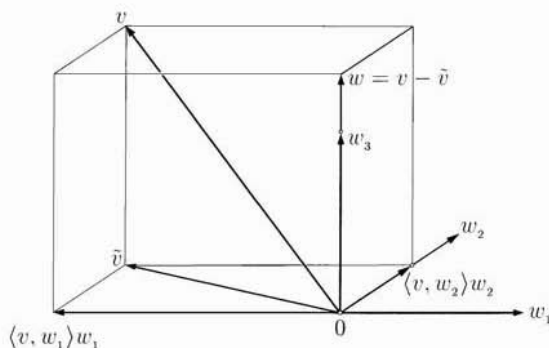


Bild 5.9

Dazu genügt es zu zeigen, daß $w \perp w_i$ für $i = 1, \dots, m$, und das folgt aus

$$\langle w, w_i \rangle = \langle v, w_i \rangle - \langle \tilde{v}, w_i \rangle = \langle v, w_i \rangle - \langle v, w_i \rangle = 0,$$

da $\langle w_i, w_j \rangle = \delta_{ij}$. Bis auf die Länge ist w schon für die Ergänzung der Basis geeignet, wir setzen

$$w_{m+1} := \frac{1}{\|w\|} \cdot w.$$

Nun hat $W' := \text{span}(w_1, \dots, w_{m+1})$ eine Orthonormalbasis $(w_1, \dots, w_{m+1})$, und es ist $\dim W' = m + 1$. Indem man das obige Verfahren so oft wie nötig wiederholt, erhält man schließlich die gewünschte Ergänzung. $\square$

Für die *praktische Rechnung* ergänzt man die Basis von W durch $v_{m+1}, \dots, v_n$ zu einer Basis von V und berechnet w_{m+1} durch Orthonormalisieren von v_{m+1} wie oben. Da aus der Konstruktion folgt, daß

$$v_{m+2} \notin \text{span}(w_1, \dots, w_m, w_{m+1}) = \text{span}(w_1, \dots, w_m, v_{m+1}),$$

kann man mit v_{m+2} fortfahren. Im letzten Schritt orthonormiert man v_n (vgl. Aufgabe 8).

5.4.10. Ein Skalarprodukt ist der abstrakte Hintergrund von Längenmessung. Damit kann man iterativ auch höherdimensionale Volumina messen: Die Fläche eines Parallelogramms ist gleich der Länge der Grundlinie und Höhe, das Volumen eines Spates gleich Grundfläche und Höhe, u.s.w. Zur Erklärung des allgemeinen m -dimensionalen Volumens verwenden wir eine Gramsche Determinante. Mit Hilfe des Orthogonalisierungsverfahrens aus 5.4.9 ergibt sich eine schöne geometrische Folgerung dafür.

In einem endlichdimensionalen euklidischen Vektorraum V seien beliebige Vektoren $v_1, \dots, v_m \in V$ gegeben, wobei $m \leq n$. Dann nennt man

$$G(v_1, \dots, v_m) := \det \begin{pmatrix} \langle v_1, v_1 \rangle & \cdots & \langle v_1, v_m \rangle \\ \vdots & & \vdots \\ \langle v_m, v_1 \rangle & \cdots & \langle v_m, v_m \rangle \end{pmatrix}$$

die *Gramsche Determinante* von $v_1, \dots, v_m$.

Bemerkung. Es gilt stets $G(v_1, \dots, v_m) \geq 0$ und

$$G(v_1, \dots, v_m) > 0 \iff (v_1, \dots, v_m) \text{ linear abhängig.}$$

Beweis. Nach 5.4.9 gibt es eine Orthonormalbasis $(w_1, \dots, w_n)$ von V . Ist

$$v_i = a_{i1}w_1 + \dots + a_{in}w_n$$

und $A := (a_{ij}) \in M(m \times n; \mathbb{R})$, so ist

$$G(v_1, \dots, v_m) = \det(A \cdot {}^t A).$$

Also folgt die Behauptung sofort aus 3.3.8. □

Nach dieser Bemerkung kann man aus der Gramschen Determinante die Wurzel ziehen, und durch

$$\text{Vol}(v_1, \dots, v_m) := \sqrt{G(v_1, \dots, v_m)}$$

das m -dimensionale *Volumen* des durch $v_1, \dots, v_m$ aufgespannten Spates erklären. Das ist die Grundlage der Definition von Verallgemeinerten Oberflächenintegralen in der Analysis (siehe etwa [Fo3]).

Im $\mathbb{R}^3$ mit dem kanonischen Skalarprodukt kann man eine beliebige Basis $(v_1, \dots, v_n)$ als Zeilen in eine Matrix A eintragen. Dann ist

$$\text{Vol}(v_1, \dots, v_n) = \sqrt{(\det A)^2} = |\det A|,$$

wie in 3.1.1, 2 e).

Nun entsteht die Frage, wie man mit m Vektoren gegebener Länge ein möglichst großes m -dimensionales Volumen aufspannen kann. Es ist höchst plausibel, daß die Vektoren dazu orthogonal sein müssen. Das besagt die

Ungleichung von HADAMARD. Seien $v_1, \dots, v_m$ beliebige Vektoren in einem n -dimensionalen Vektorraum V mit $m \leq n$. Dann ist

$$\text{Vol}(v_1, \dots, v_m) \leq \|v_1\| \cdots \|v_m\|,$$

und Gleichheit besteht genau dann, wenn die v_i paarweise orthogonal sind, d.h. wenn der aufgespannte Spat ein Quader ist.

Beweis. Es genügt, die Ungleichung in der quadratischen Form

$$G(v_1, \dots, v_m) \leq \langle v_1, v_1 \rangle \cdot \dots \cdot \langle v_m, v_m \rangle$$

zu beweisen, und dazu genügt es zu zeigen, daß für jedes $r \leq m$ folgendes gilt: Für die eindeutig bestimmte orthogonale Zerlegung (siehe 5.4.9)

$v_r = \tilde{v} + v'$ mit $\tilde{v} = \lambda_1 v_1 + \dots + \lambda_{r-1} v_{r-1}$ und $v' \perp \text{span}(v_1, \dots, v_{r-1})$ ist

$$G(v_1, \dots, v_r) = G(v_1, \dots, v_{r-1}) \cdot \langle v', v' \rangle \leq G(v_1, \dots, v_{r-1}) \cdot \langle v_r, v_r \rangle,$$

und die Ungleichung ist genau dann eine Gleichung, wenn

$$v_r \perp \text{span}(v_1, \dots, v_{r-1}).$$

Die Ungleichung und das Kriterium für die Gleichheit folgen unmittelbar aus der Definition. Zum Beweis der ersten Gleichung verwenden wir, daß

$$\langle v_i, v_r \rangle = \langle v_i, \tilde{v} \rangle \quad \text{für } i = 1, \dots, r-1 \quad \text{und} \quad \langle v_r, v_r \rangle = \langle v', v' \rangle + \langle \tilde{v}, \tilde{v} \rangle.$$

Daher ist

$$G(v_1, \dots, v_r) = \det \begin{pmatrix} \langle v_1, v_1 \rangle & \dots & \langle v_1, v_{r-1} \rangle & \langle v_1, \tilde{v} \rangle \\ \vdots & & \vdots & \vdots \\ \langle v_{r-1}, v_1 \rangle & \dots & \langle v_{r-1}, v_{r-1} \rangle & \langle v_{r-1}, \tilde{v} \rangle \\ \langle \tilde{v}, v_1 \rangle & \dots & \langle \tilde{v}, v_{r-1} \rangle & \langle v', v' \rangle + \langle \tilde{v}, \tilde{v} \rangle \end{pmatrix}.$$

Indem man von der letzten Spalte die Spalten 1 bis $r-1$, jeweils multipliziert mit $\lambda_1, \dots, \lambda_{r-1}$, abzieht, wird daraus die Spalte

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \\ \langle v', v' \rangle \end{pmatrix},$$

und die Behauptung folgt durch Entwicklung der Determinante nach der neuen letzten Spalte. $\square$

Aufgaben zu 5.4

1. Sei K ein Körper mit $\text{char } K \neq 2$ und V ein K -Vektorraum. Zeigen Sie, dass sich jede Bilinearform auf V in eindeutiger Weise als Summe einer symmetrischen und einer schiefsymmetrischen Bilinearform darstellen lässt.

2. Sei V ein 3-dimensionaler $\mathbb{R}$ -Vektorraum, $\mathcal{A} = (v_1, v_2, v_3)$ eine Basis von V und s eine Bilinearform auf V mit

$$M_{\mathcal{A}}(s) = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

Zeigen Sie, dass $\mathcal{B} = (v_1 + v_2, v_2 + v_3, v_2)$ eine Basis von V ist, und berechnen Sie $M_{\mathcal{B}}(s)$.

3. Gegeben seien $F, G \in \text{Hom}(\mathbb{R}^3, \mathbb{R})$ mit $F(x_1, x_2, x_3) = a_1x_1 + a_2x_2 + a_3x_3$, $G(x_1, x_2, x_3) = b_1x_1 + b_2x_2 + b_3x_3$. Zeigen Sie, dass $s: \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$, $(x, y) \mapsto F(x) \cdot G(y)$ eine Bilinearform ist, und bestimmen Sie die Matrix von s bezüglich der kanonischen Basis des $\mathbb{R}^3$.

4. Zeigen Sie, dass für einen $\mathbb{R}$ -Vektorraum V der folgende Zusammenhang zwischen Normen und Skalarprodukten gilt:

a) Ist $\langle \cdot, \cdot \rangle$ ein Skalarprodukt auf V mit zugehöriger Norm $\|v\| = \sqrt{\langle v, v \rangle}$, so gilt die *Parallelogramm-Gleichung*

$$\|v + w\|^2 + \|v - w\|^2 = 2\|v\|^2 + 2\|w\|^2.$$

b)* Ist umgekehrt $\|\cdot\|$ eine Norm auf V , die die Parallelogrammgleichung erfüllt, so existiert ein Skalarprodukt $\langle \cdot, \cdot \rangle$ auf V mit $\|v\| = \sqrt{\langle v, v \rangle}$.

5. Wir wollen zeigen, dass auf einem $\mathbb{R}$ -Vektorraum nicht jede Metrik aus einer Norm und nicht jede Norm aus einem Skalarprodukt entsteht. (Zur Erinnerung: Eine Norm auf einem $\mathbb{R}$ -Vektorraum V ist eine Abbildung $V \rightarrow \mathbb{R}_+$ mit den Eigenschaften N1, N2, N3 aus 5.1.2, eine Metrik auf V ist eine Abbildung $V \times V \rightarrow \mathbb{R}_+$ mit den Eigenschaften D1, D2, D3 aus 5.1.2.)

a) Zeigen Sie, dass für $n \geq 2$ auf dem $\mathbb{R}^n$ durch $\|x\| := \max\{|x_i| : 1 \leq i \leq n\}$ eine Norm definiert ist, für die kein Skalarprodukt $\langle \cdot, \cdot \rangle$ auf $\mathbb{R}^n$ existiert mit $\|x\| = \sqrt{\langle x, x \rangle}$.

b) Sei $V = \mathcal{C}(\mathbb{R}; \mathbb{R})$ der Vektorraum der stetigen Funktionen, und für $k \in \mathbb{N}$, $f \in V$ sei $\|f\|_k := \max\{|f(x)| : x \in [-k, k]\}$. Zeigen Sie, dass durch

$$d(f, g) := \sum_{k=0}^{\infty} 2^{-k} \frac{\|f - g\|_k}{1 + \|f - g\|_k}$$

eine Metrik auf V definiert ist, für die keine Norm $\|\cdot\|: V \rightarrow \mathbb{R}_+$ existiert, so dass $\|f - g\| = d(f, g)$.

6. Sei V ein endlichdimensionaler Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$ und $(u_1, \dots, u_r)$ eine orthonormale Familie in V . Beweisen Sie, dass die folgenden Bedingungen äquivalent sind:

- i) $(v_1, \dots, v_r)$ ist eine Basis von V .
- ii) Ist $v \in V$, so folgt aus $\langle v, v_i \rangle = 0$ für alle i , dass $v = 0$ ist.
- iii) Ist $v \in V$, so gilt: $v = \sum_{i=1}^r \langle v, v_i \rangle \cdot v_i$.
- iv) Für alle $v, w \in V$ gilt: $\langle v, w \rangle = \sum_{i=1}^r \langle v, v_i \rangle \cdot \langle v_i, w \rangle$.
- v) Für alle $v \in V$ gilt: $\|v\|^2 = \sum_{i=1}^r |\langle v, v_i \rangle|^2$.

7. Sei $B = (\frac{1}{2}\sqrt{2}, \cos x, \sin x, \cos 2x, \sin 2x, \dots)$ und

$$W = \text{span } B \subset \mathcal{C}([0, 2\pi]; \mathbb{R}) = V$$

(vgl. mit dem Vektorraum der trigonometrischen Polynome in Aufgabe 4 zu 1.4). Zeigen Sie:

- a) Durch $\langle f, g \rangle := \frac{1}{\pi} \int_0^{2\pi} f(x)g(x) dx$ ist ein Skalarprodukt auf V definiert.
- b) B ist eine Orthonormalbasis (bzgl. $\langle \cdot, \cdot \rangle$) von W .
- c) Ist $f(x) = \frac{a_0}{2}\sqrt{2} + \sum_{k=1}^n (a_k \cos kx + b_k \sin kx) \in W$, so gilt $a_k = \langle f, \cos kx \rangle$, $b_k = \langle f, \sin kx \rangle$. Für $f \in V$ heißen die Zahlen
 $a_k = \langle f, \cos kx \rangle$, $k \in \mathbb{N} \setminus \{0\}$, $b_l = \langle f, \sin lx \rangle$, $l \in \mathbb{N} \setminus \{0\}$,
 die *Fourierkoeffizienten* von f .
- d)* Ist $f \in V$ und sind a_k, b_k die Fourierkoeffizienten von f , so gilt die Ungleichung von Bessel:

$$\|f\|^2 \geq a_0^2 + \sum_{k=1}^{\infty} (a_k^2 + b_k^2).$$

- e)* Sind $f, g \in V$ stückweise stetig differenzierbar, und sind a_k, b_k die Fourierkoeffizienten von f und a'_k, b'_k die Fourierkoeffizienten von g , so gilt die Formel von Parseval:

$$\langle f, g \rangle = a_0 a'_0 + \sum_{k=1}^{\infty} (a_k a'_k + b_k b'_k).$$

8. Bestimmen Sie mit dem Schmidtschen Verfahren eine Orthonormalbasis des folgenden Untervektorraums des $\mathbb{R}^5$:

$$\text{span} \left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 0 \\ 2 \\ 3 \end{pmatrix} \right).$$

9. Gegeben sei auf $V = \text{span}(1, t, t^2, t^3) \subset \mathbb{R}[t]$ das Skalarprodukt

$$s(f, g) = \int_{-1}^1 f(t)g(t) dt.$$

- Bestimmen Sie die Matrix von s bezüglich der Basis $(1, t, t^2, t^3)$.
- Bestimmen Sie eine Orthonormalbasis von V .

10.* Ein *symplektischer Vektorraum* ist ein $\mathbb{R}$ -Vektorraum V mit einer schiefsymmetrischen Bilinearform ω , die nicht-entartet ist (d.h. dass aus $\omega(v, w) = 0$ für alle $w \in V$ stets $v = 0$ folgt). Eine Basis $(v_1, \dots, v_n, w_1, \dots, w_n)$ von V heißt *Darboux-Basis*, wenn gilt: $\omega(v_i, v_j) = \omega(w_i, w_j) = 0$ und $\omega(v_i, w_j) = \delta_{ij}$ für alle i, j . Zeigen Sie, dass jeder endlichdimensionale symplektische Vektorraum eine Darboux-Basis besitzt (und damit insbesondere gerade Dimension hat).

5.5 Orthogonale und unitäre Endomorphismen

Besonders wichtig sind die Abbildungen, die Abstände und Winkel erhalten. Es wird sich zeigen, daß sie geometrische Eigenschaften haben, die anschaulich gar nicht offensichtlich sind. Bei den grundlegenden Begriffen kann man den reellen und komplexen Fall gemeinsam behandeln, erst anschließend werden Unterschiede und interessante Beziehungen zwischen beiden Fällen sichtbar. Wie bisher setzen wir stets stillschweigend voraus, daß alle auftretenden Vektorräume endlichdimensional sind.

5.5.1. Für eine Abbildung, die Abstände erhalten soll, verlangen wir zunächst scheinbar etwas mehr.

Definition. Sei V ein euklidischer bzw. unitärer Vektorraum und F ein Endomorphismus von V . Dann heißt F *orthogonal* bzw. *unitär*, wenn

$$\langle F(v), F(w) \rangle = \langle v, w \rangle \quad \text{für alle } v, w \in V.$$

Bemerkung. Ein orthogonales bzw. unitäres $F \in \text{End}(V)$ hat folgende weiteren Eigenschaften:

- a) $\|F(v)\| = \|v\|$ für alle $v \in V$.
- b) $v \perp w \Rightarrow F(v) \perp F(w)$.
- c) F ist Isomorphismus und F^{-1} ist orthogonal bzw. unitär.
- d) Ist $\lambda \in \mathbb{K}$ Eigenwert von F , so ist $|\lambda| = 1$.

Beweis. a) und b) sind klar. Aus a) folgt, daß F injektiv ist, daher folgt c) aus 2.2.4. Ist v Eigenvektor zum Eigenwert λ , so ist

$$\|v\| = \|F(v)\| = \|\lambda v\| = |\lambda| \cdot \|v\|, \quad \text{also } |\lambda| = 1 \quad \text{wegen } \|v\| \neq 0. \quad \square$$

Man nennt orthogonale bzw. unitäre Abbildungen oft auch *Isometrien*, d.h. Abbildungen, die Abstände erhalten. Das ist gerechtfertigt durch das folgende

Lemma. Ist $F \in \text{End}(V)$ mit $\|F(v)\| = \|v\|$ für alle $v \in V$, so ist F orthogonal bzw. unitär.

Beweis. Aus der Invarianz der Norm folgt die Invarianz der zum Skalarprodukt gehörigen quadratischen Form. Aus den Polarisierungsgleichungen in 5.4.4 und 5.4.5 folgt daher die Invarianz des Skalarproduktes. $\square$

Vorsicht! Ist F orthogonal, so erhält F auch Winkel, insbesondere rechte Winkel. Aus der Definition des Winkels in 5.1.4 sieht man, daß für $0 \neq \varrho \in \mathbb{R}$ auch die Abbildung $\varrho \cdot F$ Winkel erhält. Für $|\varrho| \neq 1$ ist sie aber nicht mehr *orthogonal* im Sinne der Definition. Diese allgemein übliche Bezeichnungsweise ist also etwas irreführend.

5.5.2. Im $\mathbb{R}^n$ und $\mathbb{C}^n$ mit dem kanonischen Skalarprodukt sind Endomorphismen durch Matrizen beschrieben. Für die Abbildung $A: \mathbb{K}^n \rightarrow \mathbb{K}^n$ bedeutet orthogonal bzw. unitär dann

$${}^t x \bar{y} = {}^t (Ax) \overline{Ay} = {}^t x ({}^t A \bar{A}) \bar{y} \quad \text{für alle } x, y,$$

also ${}^t A \bar{A} = E_n$, d.h. $A^{-1} = {}^t \bar{A}$. Das erklärt die

Definition. Eine Matrix $A \in GL(n; \mathbb{R})$ heißt *orthogonal*, falls

$$A^{-1} = {}^t A,$$

und entsprechend heißt $A \in GL(n; \mathbb{C})$ *unitär*, falls

$$A^{-1} = {}^t \bar{A}.$$

Für eine unitäre Matrix A folgt aus $A \cdot {}^t \bar{A} = E_n$

$$|\det A|^2 = \det A \cdot \det \bar{A} = \det A \cdot \det {}^t \bar{A} = \det (A \cdot {}^t \bar{A}) = \det E_n = 1,$$

daß $|\det A| = 1$. Entsprechend ist für eine orthogonale Matrix $\det A = \pm 1$. Man nennt A *eigentlich orthogonal*, wenn $\det A = +1$. Das bedeutet nach 3.4.2, daß A orientierungstreu ist. Die Mengen

$$O(n) := \{A \in GL(n; \mathbb{R}): A^{-1} = {}^t A\}, \quad (\text{orthogonale Gruppe})$$

$$SO(n) := \{A \in O(n): \det A = 1\} \quad \text{und} \quad (\text{spezielle orthogonale Gruppe})$$

$$U(n) := \{A \in GL(n; \mathbb{C}): A^{-1} = {}^t \bar{A}\} \quad (\text{unitäre Gruppe})$$

der orthogonalen, eigentlich orthogonalen und unitären Matrizen sind Untergruppen von $GL(n; \mathbb{R})$ bzw. $GL(n; \mathbb{C})$. Wir zeigen das für $U(n)$. Sind A, B unitär, so ist

$$(AB)^{-1} = B^{-1} A^{-1} = {}^t \bar{B} {}^t \bar{A} = {}^t \overline{(AB)} \quad \text{und} \quad (A^{-1})^{-1} = A = {}^t \overline{(A^{-1})},$$

also sind AB und A^{-1} unitär.

Bemerkung. Für $A \in M(n \times n; \mathbb{K})$ sind folgende Bedingungen gleichwertig:

- i) A ist orthogonal bzw. unitär.
- ii) Die Spalten von A sind eine Orthonormalbasis von $\mathbb{K}^n$.
- iii) Die Zeilen von A sind eine Orthonormalbasis von $\mathbb{K}^n$.

Beweis. ii) bedeutet ${}^t A \bar{A} = E_n$, d.h. ${}^t \bar{A} A = E_n$ und iii) bedeutet $A \cdot {}^t \bar{A} = E_n$. $\square$

5.5.3. Der Übergang von den allgemeinen Räumen und Endomorphismen zu Standardräumen und Matrizen geschieht nach der folgenden Regel.

Satz. Sei V ein euklidischer bzw. unitärer Vektorraum mit einer Orthonormalbasis $\mathcal{B}$ und F ein Endomorphismus von V . Dann gilt:

F orthogonal (bzw. unitär) $\Leftrightarrow M_{\mathcal{B}}(F)$ orthogonal (bzw. unitär).

Beweis. Sei $A := M_{\mathcal{B}}(F) \in M(n \times n; \mathbb{K})$, und für $v, w \in V$ seien x, y die Koordinaten, d.h. $v = \Phi_{\mathcal{B}}(x)$ und $w = \Phi_{\mathcal{B}}(y)$. Da $\mathcal{B}$ orthonormal ist, folgt

$$\langle v, w \rangle = {}^t x \bar{y},$$

wobei x und y Spaltenvektoren im $\mathbb{K}^n$ sind. Daß F orthogonal bzw. unitär ist, bedeutet dann

$${}^t(Ax)\overline{Ay} = {}^t x \bar{y}, \quad \text{also} \quad {}^t A \bar{A} = E_n. \quad \square$$

5.5.4. Als interessante Beispiele wollen wir die orthogonalen Abbildungen des $\mathbb{R}^n$ für kleine n genauer ansehen.

a) Der Fall $n = 1$ ist unergiebig, es gibt nur die Möglichkeiten

$$F(x) = \pm x.$$

b) Für $n = 2$ genügt es, die orthogonalen (2×2) -Matrizen anzugeben.

Lemma. Ist $A \in O(2)$, so gibt es ein $\alpha \in [0, 2\pi[$, so daß

$$A = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \quad \text{oder} \quad A = \begin{pmatrix} \cos \alpha & \sin \alpha \\ \sin \alpha & -\cos \alpha \end{pmatrix}.$$

Im ersten Fall ist $A \in SO(2)$, die Abbildung ist eine *Drehung*. Im zweiten Fall ist $\det A = -1$, die Abbildung ist eine *Spiegelung* an einer Geraden (vgl. 4.1.1).

Vom Standpunkt der Topologie hat die Gruppe $O(2)$ zwei Zusammenhangskomponenten, jede ist homöomorph zu einer Kreislinie (vgl. 3.4.4).

Beweis. Ist $A \in O(2)$, so muß ${}^t A \cdot A = E_2$ sein. Ist

$$A = \begin{pmatrix} a & c \\ b & d \end{pmatrix},$$

so folgt

$$1. \ a^2 + b^2 = 1, \quad 2. \ c^2 + d^2 = 1 \quad \text{und} \quad 3. \ ac + bd = 0.$$

Wegen 1. und 2. gibt es $\alpha, \alpha' \in [0, 2\pi[$, so daß

$$a = \cos \alpha, \quad b = \sin \alpha, \quad c = \sin \alpha', \quad d = \cos \alpha'.$$

Nach 3. ist $0 = \cos \alpha \cdot \sin \alpha' + \sin \alpha \cdot \cos \alpha' = \sin(\alpha + \alpha')$. Also ist $\alpha + \alpha'$ entweder ein geradzahliges oder ein ungeradzahliges Vielfaches von π . Deshalb

ist entweder

$$\begin{aligned} c &= \sin \alpha' = -\sin \alpha \quad \text{und} \quad d = \cos \alpha' = \cos \alpha \quad \text{oder} \\ c &= \sin \alpha' = \sin \alpha \quad \text{und} \quad d = \cos \alpha' = -\cos \alpha. \end{aligned}$$

□

Wir erinnern daran, was schon in 4.2.4 über die Eigenwerte einer Matrix $A \in O(2)$ bewiesen wurde:

Ist $\det A = +1$, so gibt es nur im Fall $\alpha = 0$ oder $\alpha = \pi$ Eigenwerte.

Ist $\det A = -1$, so gibt es Eigenwerte $+1$ und -1 , und die zugehörigen Eigenvektoren stehen senkrecht aufeinander.

c) Ist $F: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ orthogonal, so betrachten wir das charakteristische Polynom P_F . Es hat den Grad 3, also nach dem Zwischenwertsatz der Analysis mindestens eine reelle Nullstelle. Also hat F einen Eigenwert λ_1 , und nach 5.5.1 ist $\lambda_1 = \pm 1$. Sei $w_1 \in \mathbb{R}^3$ ein Eigenvektor dazu, wir können $\|w_1\| = 1$ annehmen. Nach 5.4.9 können wir ihn zu einer Orthonormalbasis $\mathcal{B} = (w_1, w_2, w_3)$ ergänzen. Bezeichnet $W \subset \mathbb{R}^3$ die von w_2 und w_3 aufgespannte Ebene, so folgt aus der Bemerkung in 5.5.1, daß $F(W) = W$. Also ist

$$M_{\mathcal{B}}(F) = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \boxed{A'} \\ 0 & \end{pmatrix} =: A,$$

und aus 5.5.3 folgt $A' \in O(2)$. Weiter ist $\det A = \lambda_1 \cdot \det A'$. Nun sind Fallunterscheidungen nötig.

Sei $\det F = \det A = +1$. Ist $\lambda_1 = -1$, so muß $\det A' = -1$ sein. Daher kann man w_2 und w_3 als Eigenvektoren zu den Eigenwerten $\lambda_2 = +1$ und $\lambda_3 = -1$ wählen, d.h.

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

Ist $\lambda_1 = +1$, so muß auch $\det A' = +1$ sein, also gibt es ein $\alpha \in [0, 2\pi[$, so daß

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{pmatrix}.$$

Ist $\det F = -1$, so gibt es bei geeigneter Wahl von w_2 und w_3 für A die Möglich-

keiten

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{pmatrix}.$$

Man überlege sich, was das geometrisch bedeutet. Als Anwendung für das tägliche Leben notieren wir den

Satz vom Fußball. *Bei jedem Fußballspiel, in dem nur ein Ball benutzt wird, gibt es zwei Punkte auf der Oberfläche des Balles, die sich zu Beginn der ersten und der zweiten Halbzeit (wenn der Ball genau auf dem Anstoßpunkt liegt) an der gleichen Stelle im umgebenden Raum befinden.*

Beweis. Im Fall $\det F = +1$ gibt es stets einen Eigenwert $+1$. □

Diese Aussage ist zur Abwechslung leichter zu beweisen als anschaulich zu verstehen.

5.5.5. Bevor wir eine Normalform für beliebig große orthogonale Matrizen angeben, behandeln wir den unitären Fall, weil er einfacher ist.

Theorem. *Jeder unitäre Endomorphismus F eines unitären Vektorraumes besitzt eine Orthonormalbasis aus Eigenvektoren von F . Insbesondere ist er diagonalisierbar.*

Übersetzt in den Matrizenkalkül erhält man:

Korollar. *Zu $A \in U(n)$ gibt es ein $S \in U(n)$ mit*

$${}^t \bar{S} \cdot A \cdot S = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

wobei $\lambda_i \in \mathbb{C}$ mit $|\lambda_i| = 1$ für $i = 1, \dots, n$.

Beweis des Korollars. Als Spalten von S verwendet man eine Basis des $\mathbb{C}^n$, die aus Eigenvektoren von A besteht. □

Beweis des Theorems. Wir führen Induktion über $n = \dim V$. Für $n = 0$ ist nichts zu beweisen. Sei also $n \geq 1$ und

$$P_F = \pm(t - \lambda_1) \cdot \dots \cdot (t - \lambda_n) \quad \text{mit} \quad \lambda_1, \dots, \lambda_n \in \mathbb{C}$$

die nach dem Fundamentalsatz der Algebra (1.3.9) existierende Linearfaktorzerlegung des charakteristischen Polynoms von F . Zum Eigenwert λ_1 wählen wir

einen Eigenvektor v_1 mit $\|v_1\| = 1$. Wir betrachten das orthogonale Komplement zur Geraden $\mathbb{C}v_1$, d.h.

$$W := \{w \in V : \langle v_1, w \rangle = 0\}.$$

Die entscheidende Tatsache ist nun, daß $F(W) = W$ gilt, d.h. daß W ein *invarianter Unterraum* ist. Da F ein Isomorphismus ist, genügt es $F(W) \subset W$ zu beweisen. Wegen $|\lambda_1| = 1$ ist insbesondere $\lambda_1 \neq 0$, also folgt aus

$$\lambda_1 \langle v_1, F(w) \rangle = \langle \lambda_1 v_1, F(w) \rangle = \langle F(v_1), F(w) \rangle = \langle v_1, w \rangle = 0,$$

daß $\langle F(w), v_1 \rangle = 0$ und somit $F(W) \subset W$.

Nun betrachten wir den Endomorphismus $G := F|_W$ von W . Als Einschränkung von F ist er wieder unitär, und wegen $\dim W = n-1$ (vgl. 5.4.9) können wir auf G die Induktionsannahme anwenden. Danach gibt es eine Orthonormalbasis $(v_2, \dots, v_n)$ von W , bestehend aus Eigenvektoren von G und damit auch von F . Die Basis $(v_1, v_2, \dots, v_n)$ ist orthonormal und besteht aus Eigenvektoren von F . $\square$

Wie man eine derartige Orthonormalbasis konkret ausrechnen kann, wird in 5.6.3 erläutert.

5.5.6. Im Gegensatz zu der gerade im komplexen Fall bewiesenen Diagonalisierbarkeit unitärer Endomorphismen gibt es reell orthogonale Endomorphismen ohne Eigenwerte. Das einfachste Beispiel sind Drehungen der Ebene $\mathbb{R}^2$ (siehe 4.1.1). Wir zeigen nun, daß dieses Gegenbeispiel im wesentlichen das einzige ist: alle anderen lassen sich daraus aufbauen.

Theorem. Ist F ein orthogonaler Endomorphismus eines euklidischen Vektorraumes V , so gibt es in V eine Orthonormalbasis B derart, daß

$$M_B(F) = \begin{pmatrix} +1 & & & & & \\ & \ddots & & & & \\ & & +1 & & & \\ & & & -1 & & \\ & & & & \ddots & \\ & & & & & -1 \\ & 0 & & & & & \boxed{A_1} \\ & & & & & & & \ddots \\ & & & & & & & & \boxed{A_k} \end{pmatrix},$$

wobei für $j = 1, \dots, k$

$$A_j = \begin{pmatrix} \cos \vartheta_j & -\sin \vartheta_j \\ \sin \vartheta_j & \cos \vartheta_j \end{pmatrix} \in SO(2) \quad \text{mit} \quad \vartheta_j \in [0, 2\pi[, \text{ aber } \vartheta_j \neq 0, \pi.$$

F ist also charakterisiert durch die Anzahlen r und s der Eigenwerte $+1$ und -1 sowie der Winkel $\vartheta_1, \dots, \vartheta_k$, wobei $r + s + 2k = \dim V$. Die orthogonale Matrix heißt in *Normalform*.

Die Form der Matrix zeigt, dass V in ein- und zweidimensionale invariante Unterräume zerfällt. Das ist der springende Punkt. Man beachte auch den Unterschied zu 4.5.4: dort ist V nicht direkte Summe der invarianten Unterräume, also erhält man nicht nur Nullen oberhalb der Diagonale.

Lemma. *Zu einem orthogonalen Endomorphismus F eines euklidischen Vektorraumes V mit $\dim V \geq 1$ gibt es stets einen Untervektorraum $W \subset V$ mit*

$$F(W) \subset W \quad \text{und} \quad 1 \leq \dim W \leq 2.$$

Der *Beweis des Theorems* kann damit ganz einfach durch Induktion über $n = \dim V$ geführt werden.

Für $n = 0$ ist nichts zu beweisen, sei also $n \geq 1$. Nach dem Lemma gibt es einen Untervektorraum $W \subset V$ mit

$$1 \leq \dim W \leq 2 \quad \text{und} \quad F(W) = W,$$

denn eine orthogonale Abbildung ist injektiv. Insbesondere ist F^{-1} wieder orthogonal. Also ist für $w \in W$ und $v \in W^\perp$

$$\langle F(v), w \rangle = \langle F^{-1}(F(v)), F^{-1}(w) \rangle = \langle v, F^{-1}(w) \rangle = 0,$$

und es folgt $F(W^\perp) = W^\perp$. Damit haben wir F zerlegt in zwei orthogonale Abbildungen

$$G := F|_W: W \rightarrow W \quad \text{und} \quad H := F|_{W^\perp}: W^\perp \rightarrow W^\perp.$$

Da $\dim W^\perp < n$, können wir auf H die Induktionsvoraussetzung anwenden und erhalten eine Basis $\mathcal{B}'$ von $W^\perp$ der gewünschten Art.

Ist $\dim W = 1$, so gibt es einen Eigenvektor $v \in W$ mit $\|v\| = 1$ zu einem Eigenwert ± 1 . Ergänzt man $\mathcal{B}'$ an passender Stelle durch v zu $\mathcal{B}$, so hat diese Basis von V die gewünschten Eigenschaften.

Im Fall $\dim W = 2$ gibt es eine Orthonormalbasis (v_1, v_2) von W , bezüglich der G beschrieben wird durch eine Matrix der Form

$$\begin{pmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{pmatrix} \quad \text{oder} \quad \begin{pmatrix} \cos \vartheta & -\sin \vartheta \\ \sin \vartheta & \cos \vartheta \end{pmatrix} \quad \text{mit } \vartheta \neq 0, \pi.$$

Indem man v_1 und v_2 an passenden Stellen in $\mathcal{B}'$ einfügt, erhält man wieder die gewünschte Basis $\mathcal{B}$ von V .

Wie der Beweis zeigt, ist V die orthogonale Summe der invarianten Unterräume der Dimension ≤ 2 . □

Zum Beweis des Lemmas kann man verschiedene bereits bekannte Ergebnisse verwenden. Wir geben drei Alternativen.

1. Mit dem Satz von CAYLEY-HAMILTON. Das wurde schon in 4.5.4 ohne die Voraussetzung der Orthogonalität erledigt. Entscheidend dabei war die Zerlegung des charakteristischen Polynoms in lineare und quadratische Faktoren. $\square$

Bei den beiden folgenden Beweisen betrachten wir zur Vereinfachung den Spezialfall $V = \mathbb{R}^n$ mit dem kanonischen Skalarprodukt und eine orthogonale Matrix A .

2. Durch „Symmetrisierung“. Wir definieren

$${}^sA := A + {}^tA = A + A^{-1}.$$

Offenbar ist sA symmetrisch, also gibt es einen Eigenvektor von sA , d.h. $0 \neq v \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$ mit ${}^sAv = \lambda v$ (das wird zwar erst in 5.6.2 bewiesen, aber ohne Benutzung der Ergebnisse dieses Abschnitts). Wir behaupten, daß

$$W := \text{span}(v, Av)$$

die gewünschten Eigenschaften hat. Das folgt sofort durch Anwendung von A auf

$$Av + A^{-1}v = \lambda v \Rightarrow A^2v = -v + \lambda Av. \quad \square$$

Mit Hilfe der adjungierten Abbildung (6.2.4) kann man diesen Beweis auch für komplexes V durchführen.

3. Durch „Komplexifizierung“. Wir betrachten den Endomorphismus

$$A: \mathbb{C}^n \rightarrow \mathbb{C}^n, \quad z \mapsto Az.$$

Er ist unitär, denn A ist orthogonal. Also gibt es nach 5.5.5 einen komplexen Eigenvektor, d.h. $0 \neq z \in \mathbb{C}^n$ und $\lambda \in \mathbb{C}$, so daß $Az = \lambda z$. Da A reell ist, folgt

$$A\bar{z} = \overline{Az} = \overline{\lambda z} = \bar{\lambda}\bar{z}.$$

Also ist $\bar{z}$ Eigenvektor zu $\bar{\lambda}$. Wir definieren daraus reelle Vektoren

$$x := \frac{1}{2}(z + \bar{z}) \in \mathbb{R}^n \quad \text{und} \quad y = \frac{1}{2i}(z - \bar{z}) \in \mathbb{R}^n$$

und behaupten, daß $W := \text{span}(x, y) \subset \mathbb{R}^n$ unter A invariant ist. Dazu schreiben wir $\lambda = \alpha + i\beta$ mit $\alpha, \beta \in \mathbb{R}$. Dann folgt

$$\begin{aligned} Ax &= \frac{1}{2}(Az + A\bar{z}) = \frac{1}{2}(\lambda z + \bar{\lambda}\bar{z}) = \text{re } \lambda z = \alpha x - \beta y, \\ Ay &= \frac{1}{2i}(Az - A\bar{z}) = \frac{1}{2i}(\lambda z - \bar{\lambda}\bar{z}) = \text{im } \lambda z = \beta x + \alpha y. \end{aligned} \quad \square$$

Für allgemeines V kann man die Komplexifizierung mit Hilfe des Tensorprodukts (6.3.4 b) erklären.

Der erste Beweis ist sicher der eleganteste. Aber wenn man eine orthogonale Matrix explizit auf Normalform bringen will, geben die beiden anderen Beweise sofort Anleitungen, wie man das iterativ bewerkstelligen kann.

Aufgaben zu 5.5

1. Zeigen Sie, dass für $F \in O(3)$ gilt: $F(x) \times F(y) = (\det F) \cdot F(x \times y)$.

2. Ist V ein euklidischer Vektorraum und $F \in \text{End}(V)$, so heißt F *winkeltreu*, falls F injektiv ist und

$$\angle(v, w) = \angle(F(v), F(w)) \quad \text{für alle } v, w \in V \setminus \{0\}.$$

Zeigen Sie, dass F winkeltreu ist genau dann, wenn ein orthogonales $G \in \text{End}(V)$ und ein $\lambda \in \mathbb{R} \setminus \{0\}$ existieren mit $F = \lambda \cdot G$.

3. Sei $z = x + iy \in \mathbb{C}^n$, wobei $x, y \in \mathbb{R}^n$. Zeigen Sie:

x und y sind linear unabhängig über $\mathbb{R} \Leftrightarrow z$ und $\bar{z}$ sind linear unabhängig über $\mathbb{C}$.

4. Bestimmen Sie für die Matrix

$$A = \frac{1}{90} \begin{pmatrix} 66 & -18\sqrt{6} & 10\sqrt{18} \\ 6\sqrt{6} & 72 & 15\sqrt{12} \\ -14\sqrt{18} & -9\sqrt{12} & 60 \end{pmatrix}$$

eine Matrix $S \in U(3)$, so dass ${}^t \bar{S} A S$ Diagonalgestalt hat und eine Matrix $T \in O(3)$, so dass für ein $\alpha \in [0, 2\pi[$ gilt:

$${}^t T A T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{pmatrix}.$$

5. Sei $\pi \in S_n$ eine Permutation und die lineare Abbildung $f_\pi: \mathbb{R}^n \rightarrow \mathbb{R}^n$ definiert durch $f_\pi(x_1, \dots, x_n) = (x_{\pi(1)}, \dots, x_{\pi(n)})$. Bestimmen Sie die Eigenwerte von f_π .

6. Gegeben sei ein symplektischer Vektorraum V (vgl. Aufgabe 10 zu 5.4) und eine komplexe Struktur J auf V (vgl. Aufgabe 3 zu 5.3), so dass für alle $v, w \in V$ gilt: $\omega(v, w) = \omega(J(v), J(w))$.

a) Zeigen Sie, dass durch $\langle v, w \rangle := \omega(v, J(w))$ eine symmetrische Bilinearform auf V definiert wird, und dass J orthogonal bezüglich $\langle \cdot, \cdot \rangle$ ist, d. h. es gilt $\langle v, w \rangle = \langle J(v), J(w) \rangle$ für alle $v, w \in V$.

b) Die komplexe Struktur J heißt ω -kalibriert, wenn $\langle \cdot, \cdot \rangle$ positiv definit ist. Zeigen Sie, dass in diesem Fall durch $s(v, w) := \langle v, w \rangle - i\omega(v, w)$ eine positiv definite hermitesche Form auf dem von J induzierten $\mathbb{C}$ -Vektorraum V gegeben ist.

5.6 Selbstadjungierte Endomorphismen*

Dieser Abschnitt dient dem Studium von Endomorphismen, die durch symmetrische Matrizen beschrieben werden. Insbesondere werden wir zeigen, daß sie stets diagonalisierbar sind. Das hat zahlreiche Anwendungen in Geometrie und Analysis.

5.6.1. Mit V wird stets ein endlichdimensionaler euklidischer bzw. unitärer Vektorraum bezeichnet. Zu einem $F \in \text{End } V$ werden wir in 6.2.4 einen *adjungierten Endomorphismus* F^{ad} erklären, der durch die Bedingung

$$\langle F(v), w \rangle = \langle v, F^{\text{ad}}(w) \rangle \quad \text{für alle } v, w \in V$$

charakterisiert ist. Das erklärt die folgende

Definition. Ein Endomorphismus F eines euklidischen bzw. unitären Vektorraumes V heißt *selbstadjungiert*, wenn

$$\langle F(v), w \rangle = \langle v, F(w) \rangle \quad \text{für alle } v, w \in V.$$

In Matrizen übersetzt ist diese Beziehung gut bekannt:

Satz. Sei F ein Endomorphismus von V und $\mathcal{B}$ eine Orthonormalbasis. Dann gilt: F selbstadjungiert $\Leftrightarrow M_{\mathcal{B}}(F)$ symmetrisch bzw. hermitesch.

Beweis. Es genügt, den unitären Fall zu behandeln. Sind $x, y \in \mathbb{K}^n$ Spaltenvektoren, $v = \Phi_{\mathcal{B}}(x)$, $w = \Phi_{\mathcal{B}}(y)$ und $A = M_{\mathcal{B}}(F)$, so ist $\langle v, w \rangle = {}^t x \bar{y}$, da $\mathcal{B}$ orthonormal ist, also

$$\langle F(v), w \rangle = {}^t (Ax) \bar{y} = {}^t x {}^t A \bar{y} \quad \text{und} \quad \langle v, F(w) \rangle = {}^t x \overline{A y}.$$

Also ist F genau dann selbstadjungiert, wenn ${}^t A = \bar{A}$. □

5.6.2. Die Diagonale einer hermiteschen Matrix muß reell sein. Analog gilt das

Lemma. Ist F selbstadjungiert, so sind (auch im komplexen Fall) alle Eigenwerte reell. Insbesondere hat eine hermitesche Matrix nur reelle Eigenwerte.

Beweis. Ist $F(v) = \lambda v$ mit $v \neq 0$, so folgt

$$\lambda \langle v, v \rangle = \langle \lambda v, v \rangle = \langle F(v), v \rangle = \langle v, F(v) \rangle = \langle v, \lambda v \rangle = \bar{\lambda} \langle v, v \rangle,$$

also ist $\lambda = \bar{\lambda}$. □

Theorem. Ist F ein selbstadjungierter Endomorphismus eines euklidischen bzw. unitären Vektorraumes, so gibt es eine Orthonormalbasis von V , die aus Eigenvektoren von F besteht.

Korollar. Ist $A \in M(n \times n; \mathbb{K})$ eine symmetrische bzw. hermitesche Matrix, so gibt es eine orthogonale bzw. unitäre Matrix S , so daß

$${}^t\bar{S}AS = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \quad \text{mit } \lambda_1, \dots, \lambda_n \in \mathbb{R}.$$

□

Beweis des Theorems. Wir behandeln zunächst den Fall $\mathbb{K} = \mathbb{C}$. Da $\mathbb{C}$ algebraisch abgeschlossen ist, zerfällt das charakteristische Polynom in Linearfaktoren, und nach dem Lemma sind alle Eigenwerte reell, also ist

$$P_F = \pm(t - \lambda_1) \cdot \dots \cdot (t - \lambda_n) \quad \text{mit } \lambda_1, \dots, \lambda_n \in \mathbb{R}.$$

Nun führen wir Induktion über n . Der Fall $n = 0$ ist klar, für $n \geq 1$ gibt es einen Eigenvektor v_1 zu λ_1 mit $\|v_1\| = 1$ und wir definieren

$$W := \{w \in V : \langle v_1, w \rangle = 0\}.$$

W ist invariant unter F , d.h. $F(W) \subset W$, denn für $w \in W$ gilt

$$\langle v_1, F(w) \rangle = \langle F(v_1), w \rangle = \langle \lambda_1 v_1, w \rangle = \lambda_1 \langle v_1, w \rangle = 0.$$

Nach Induktionsannahme gibt es eine Orthonormalbasis von W aus Eigenvektoren von $F|_W$, durch v_1 wird sie zur gewünschten Orthonormalbasis von V ergänzt.

Der Fall $\mathbb{K} = \mathbb{R}$ kann genauso behandelt werden, wenn man weiß, daß das charakteristische Polynom P_F auch in diesem Fall in Linearfaktoren zerfällt. Dazu wählt man eine beliebige Orthonormalbasis $\mathcal{B}$ in V . Nach 5.6.1 ist $M_{\mathcal{B}}(F)$ symmetrisch, also auch hermitesch. Im Fall $\mathbb{K} = \mathbb{C}$ hatten wir gesehen, daß P_A in reelle Linearfaktoren zerfällt, und aus $P_F = P_A$ folgt die Behauptung.

Man beachte bei diesem Beweis, daß obiges Lemma nichts über die Existenz von Eigenwerten aussagt. Man kann es erst anwenden, wenn man sie hat, und dazu macht man den Ausflug ins Komplexe. □

5.6.3. Faßt man die Eigenvektoren der verschiedenen Eigenwerte zu Eigenräumen zusammen, so erhält man aus den Theoremen 5.5.5 und 5.6.2 das

Korollar. Sind $\lambda_1, \dots, \lambda_k$ die verschiedenen Eigenwerte eines selbstadjungierten oder unitären Endomorphismus F von V , so ist

$$V = \text{Eig}(F; \lambda_1) \oplus \dots \oplus \text{Eig}(F; \lambda_k).$$

Beweis. Ist λ ein Eigenwert der Vielfachheit r , so gehören dazu r Eigenvektoren aus der nach den Theoremen existierenden Orthonormalbasis, und sie bilden eine Basis von $\text{Eig}(F; \lambda)$. Daraus erhält man die Orthogonalität der Zerlegung.

Daß die Eigenräume zu verschiedenen Eigenwerten stets orthogonal sind, kann man auch noch einmal direkt nachrechnen: Sind v und w Eigenvektoren zu

$\lambda \neq \mu$, so ist im selbstadjungierten Fall wegen $\bar{\mu} = \mu$

$$\lambda \langle v, w \rangle = \langle \lambda v, w \rangle = \langle F(v), w \rangle = \langle v, F(w) \rangle = \langle v, \mu w \rangle = \mu \langle v, w \rangle,$$

also $\langle v, w \rangle = 0$, und im unitären Fall

$$\langle v, w \rangle = \langle F(v), F(w) \rangle = \langle \lambda v, \mu w \rangle = \lambda \bar{\mu} \langle v, w \rangle.$$

Aus $\langle v, w \rangle \neq 0$ folgt $\lambda \bar{\mu} = 1$. Wegen $|\mu| = 1$ folgt $\bar{\mu} = \mu^{-1}$, also $\lambda = \mu$. $\square$

Zur praktischen Berechnung einer Orthonormalbasis aus Eigenvektoren kann man wie folgt vorgehen:

- 1) Man bestimme die Linearfaktorzerlegung des charakteristischen Polynoms

$$P_F = \pm(t - \lambda_1)^{r_1} \cdot \dots \cdot (t - \lambda_k)^{r_k},$$

wobei $\lambda_1, \dots, \lambda_k$ paarweise verschieden sind.

- 2) Für jeden Eigenwert λ der Vielfachheit r bestimme man eine Basis von $\text{Eig}(F; \lambda)$ durch Lösen eines linearen Gleichungssystems.
- 3) Man orthonormalisiere die in 2) erhaltenen Basen nach dem Verfahren aus 5.4.9, und zwar unabhängig voneinander in den verschiedenen Eigenräumen.
- 4) Die k Basen der Eigenräume aus 3) bilden zusammen die gesuchte Basis aus Eigenvektoren von V .

Beispiel. Sei

$$A := \frac{1}{15} \begin{pmatrix} 10 & 5 & 10 \\ 5 & -14 & 2 \\ 10 & 2 & -11 \end{pmatrix} \in M(3 \times 3; \mathbb{R}).$$

Wie man leicht nachrechnet, bilden die Spaltenvektoren eine Orthonormalbasis von $\mathbb{R}^3$, und es ist $\det A = 1$. Also ist $A \in SO(3)$. Als charakteristisches Polynom erhält man

$$P_A = -t^3 - t^2 + t + 1 = -(t - 1)(t + 1)^2.$$

Zur Bestimmung von $\text{Eig}(A; 1)$ hat man das homogene lineare Gleichungssystem mit der Koeffizientenmatrix

$$\frac{1}{15} \begin{pmatrix} -5 & 5 & 10 \\ 5 & -29 & 2 \\ 10 & 2 & -26 \end{pmatrix}$$

zu lösen. Man findet $\text{Eig}(A; 1) = \mathbb{R} \cdot (5, 1, 2)$.

$\text{Eig}(A; -1)$ ist Lösungsraum des homogenen linearen Gleichungssystems mit der Koeffizientenmatrix

$$\frac{1}{15} \begin{pmatrix} 25 & 5 & 10 \\ 5 & 1 & 2 \\ 10 & 2 & 4 \end{pmatrix},$$

und man erhält $\text{Eig}(A; -1) = \text{Eig}(A; 1)^\perp$. Etwa

$$(0, -2, 1) \quad \text{und} \quad (1, -1, -2)$$

bilden eine orthogonale Basis von $\text{Eig}(A; -1)$. Also ist

$$\mathcal{B} := \left(\frac{1}{\sqrt{30}}(5, 1, 2), \frac{1}{\sqrt{5}}(0, -2, 1), \frac{1}{\sqrt{6}}(1, -1, -2) \right)$$

eine Orthonormalbasis des $\mathbb{R}^3$, bestehend aus Eigenvektoren von A . Setzen wir

$$T := \begin{pmatrix} \frac{5}{\sqrt{30}} & 0 & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{30}} & -\frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{30}} & \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{6}} \end{pmatrix} = T_{\mathcal{K}}^{\mathcal{B}},$$

so folgt

$${}^tTAT = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} =: D,$$

oder gleichbedeutend damit $T \cdot D \cdot {}^tT = A$, was man mit etwas weniger Rechenarbeit nachprüfen kann.

5.6.4. Wie in 5.6.2 bewiesen wurde, zerfällt das charakteristische Polynom einer reellen symmetrischen Matrix in reelle Linearfaktoren. Das ist eine höchst überraschende Tatsache, aus dem Beweis mit der Komplexifizierung kann man kaum den geometrischen Hintergrund erkennen. Daher geben wir noch einen ganz reellen Beweis. Nach der üblichen Methode, die Eigenwerte schrittweise abzubauen, genügt der Beweis von folgendem

Lemma. Jede symmetrische Matrix $A \in M(n \times n; \mathbb{R})$ hat einen reellen Eigenwert.

Beweis. Wir betrachten die zu A gehörige quadratische Form

$$q: \mathbb{R}^n \rightarrow \mathbb{R}, \quad x \mapsto {}^txAx.$$

Das ist ein homogenes Polynom vom Grad 2 (vgl. 5.4.4), insbesondere eine stetige Funktion. Die Sphäre

$$S := \{x \in \mathbb{R}^n: \|x\| = 1\}$$

ist kompakt, also nimmt q darauf ein Maximum an (vgl. [Fo2], §3). Es gibt also ein $v \in S$ mit

$${}^t v A v \geq {}^t x A x \quad \text{für alle } x \in S.$$

Wir behaupten nun, daß dieses v ein Eigenvektor ist. Dazu genügt es zu zeigen:

$$\text{Für } w \in S \text{ mit } v \perp w \text{ ist auch } Av \perp w. \quad (*)$$

Denn ist $W := (\mathbb{R}v)^\perp$, so folgt $Av \in W^\perp = \mathbb{R}v$.

Zum Nachweis von $(*)$ betrachten wir für $\tau \in]0, 1]$ und $\sigma := \sqrt{1 - \tau^2}$ den Vektor

$$x := \sigma v + \tau w.$$

Wegen $v \perp w$ ist $x \in S$. Da A symmetrisch ist, gilt ${}^t v A w = {}^t w A v$, also ist

$${}^t v A v \geq {}^t x A x = \sigma^2 \cdot {}^t v A v + 2\sigma\tau \cdot {}^t w A v + \tau^2 \cdot {}^t w A w.$$

Daraus folgt nach Division durch τ

$$2\sigma({}^t w A v) \leq \tau({}^t v A v - {}^t w A w). \quad (**)$$

Angenommen, Av wäre nicht orthogonal zu w . Indem wir eventuell w durch $-w$ ersetzen, können wir ${}^t w A v > 0$ annehmen. Da außerdem nach der Wahl von v

$${}^t v A v - {}^t w A w \geq 0$$

gilt, ergäbe sich aus $(**)$ ein Widerspruch, wenn man τ gegen 0 (und damit σ gegen 1) gehen läßt.

Wegen $\|v\| = 1$ folgt sofort, daß ${}^t v A v$ der Eigenwert zum Eigenvektor v ist. Da die weiteren Eigenwerte von A Werte der quadratischen Form q auf $W \cap S$ sind, folgt weiter, daß ${}^t v A v$ der größte Eigenwert von A ist (dabei ist wie oben $W = (\mathbb{R}v)^\perp$). $\square$

Bei dieser Beweismethode werden Eigenwert und Eigenvektor gleichzeitig gefunden, daraus kann auch leicht ein Approximationsverfahren zur numerischen Lösung gemacht werden.

Aufgaben zu 5.6

1. Sei $F: \mathbb{K}^n \rightarrow \mathbb{K}^n$ ein selbstadjungierter, nilpotenter Endomorphismus. Zeigen Sie, dass $F = 0$ ist.
2. Seien F und G zwei selbstadjungierte Endomorphismen auf einem endlichdimensionalen euklidischen bzw. unitären Vektorraum V . Zeigen Sie, dass $F \circ G$ selbstadjungiert ist genau dann, wenn $F \circ G = G \circ F$.
3. Bestimmen Sie für die Matrix

$$A = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

eine orthogonale Matrix $S \in O(3)$, so dass tSAS eine Diagonalmatrix ist.

5.7 Hauptachsentransformation*

Ist eine symmetrische Bilinearform positiv definit, so kann man dazu nach dem Verfahren von GRAM-SCHMIDT entsprechend 5.4.9 eine Orthonormalbasis konstruieren. Das wirft folgende Fragen auf:

- 1) Wie kann man einfach entscheiden, ob eine symmetrische Bilinearform positiv definit ist?
- 2) Gibt es einen Ersatz für eine Orthonormalbasis im nicht positiv definiten Fall?

Darauf werden in diesem Abschnitt Antworten gegeben.

5.7.1. Wir behandeln zunächst den reellen Fall und betrachten erst einmal eine symmetrische Bilinearform

$$s: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}.$$

Sie ist nach 5.4.2 bestimmt durch die symmetrische Matrix $A = (s(e_i, e_j))$. Für Spaltenvektoren $x, y \in \mathbb{R}^n$ gilt

$$s(x, y) = {}^t x A y = \langle Ax, y \rangle = \langle x, Ay \rangle,$$

wobei $\langle \cdot, \cdot \rangle$ das kanonische Skalarprodukt des $\mathbb{R}^n$ bezeichnet. Die Matrix A kann auch angesehen werden als ein selbstadjungierter Endomorphismus

$$A: \mathbb{R}^n \rightarrow \mathbb{R}^n,$$

der nach 5.6.2 diagonalisierbar ist. Es gibt also eine orthonormale Basis $\mathcal{B} = (w_1, \dots, w_n)$ des $\mathbb{R}^n$ aus Eigenvektoren, d.h. $\lambda_i \in \mathbb{R}$ mit

$$A w_i = \lambda_i w_i, \quad \text{also} \quad s(w_i, w_j) = \langle A w_i, w_j \rangle = \lambda_i \langle w_i, w_j \rangle = \lambda_i \cdot \delta_{ij}.$$

Sind die Eigenwerte $\lambda_1, \dots, \lambda_n$ so numeriert, daß

$$\lambda_1, \dots, \lambda_k > 0, \quad \lambda_{k+1}, \dots, \lambda_m < 0 \quad \text{und} \quad \lambda_{m+1} = \dots = \lambda_n = 0,$$

so setzen wir

$$w'_i := \begin{cases} |\lambda_i|^{-\frac{1}{2}} \cdot w_i & \text{für } i = 1, \dots, m, \\ w_i & \text{für } i = m+1, \dots, n. \end{cases}$$

Das ergibt eine bezüglich $\langle \cdot, \cdot \rangle$ orthogonale Basis $\mathcal{B}' = (w'_1, \dots, w'_n)$ des $\mathbb{R}^n$ mit

$$s(w'_i, w'_j) = \begin{cases} +1 & \text{für } 1 \leq i = j \leq k, \\ -1 & \text{für } k+1 \leq i = j \leq m, \\ 0 & \text{sonst.} \end{cases}$$

Mit Matrizen beschreibt man das so: Ist $S \in O(n)$ derart, daß

$$S A S^{-1} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} =: D,$$

so folgt aus $S^{-1} = {}^tS$ und der Transformationsformel 5.4.3 mit $T := S^{-1}$, daß

$$M_B(s) = {}^tTAT = D.$$

Weiter ist

$$M_{B'}(s) = \begin{pmatrix} E_k & & \\ & -E_l & \\ & & 0 \end{pmatrix} =: D'$$

mit $k + l = m$. Wir fassen das Ergebnis zusammen:

Hauptachsentransformation symmetrischer Matrizen. Sei $A \in M(n \times n; \mathbb{R})$ symmetrisch und s die durch A beschriebene Bilinearform auf $\mathbb{R}^n$. Dann gilt:

1) Ist $\mathcal{B} = (w_1, \dots, w_n)$ eine Orthonormalbasis des $\mathbb{R}^n$ bezüglich des Standardskalarproduktes bestehend aus Eigenvektoren des Endomorphismus A , so ist

$$M_B(s) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, \quad \text{d.h. } s(w_i, w_j) = \lambda_i \cdot \delta_{ij},$$

wobei $\lambda_1, \dots, \lambda_n$ die Eigenwerte von A sind.

2) Es gibt eine Basis $\mathcal{B}'$ des $\mathbb{R}^n$, so daß

$$M_{B'}(s) = \begin{pmatrix} E_k & & 0 \\ & -E_l & \\ 0 & & 0 \end{pmatrix} = D',$$

d.h. es gibt ein $T' \in GL(n; \mathbb{R})$ mit $D' = {}^tT'AT'$. □

Es sei bemerkt, daß man Teil 2) auch ohne Benutzung von 5.6.2 direkt durch simultane Zeilen- und Spaltenumformungen von A über jedem Körper K mit $\text{char}(K) \neq 2$ beweisen kann (siehe 5.7.5).

Vorsicht! Die Eigenwerte von A sind nur Invarianten des Endomorphismus, für eine Bilinearform ist der Begriff Eigenwert sinnlos.

In 5.7.4 werden wir zeigen, daß die Zahlen k und l Invarianten der Bilinearform sind.

5.7.2. Um den Namen *Hauptachsentransformation* zu erklären, geben wir folgendes

Beispiel. Sei

$$A = \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix} \in M(2 \times 2; \mathbb{R}).$$

Dann ist $s(x, y) = \alpha(x_1 y_1 + x_2 y_2) + \beta(x_1 y_2 + x_2 y_1)$. Aus

$$P_A(t) = t^2 - 2\alpha t + (\alpha^2 - \beta^2)$$

erhält man die Eigenwerte $\lambda_1 = \alpha + \beta$ und $\lambda_2 = \alpha - \beta$, sowie die Eigenvektoren

$$w_1 = \frac{1}{\sqrt{2}}(1, 1) \quad \text{und} \quad w_2 = \frac{1}{\sqrt{2}}(-1, 1).$$

Dafür gilt $s(w_1, w_2) = 0$, $s(w_1, w_1) = \alpha + \beta$ und $s(w_2, w_2) = \alpha - \beta$. Die zu s gehörige quadratische Form ist gegeben durch

$$q(x) = \alpha x_1^2 + 2\beta x_1 x_2 + \alpha x_2^2.$$

Sind z_1, z_2 die Koordinaten bezüglich der Basis (w_1, w_2) , so gilt

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix},$$

also

$$q(x(z)) = (\alpha + \beta)z_1^2 + (\alpha - \beta)z_2^2.$$

In den neuen z -Koordinaten ist also der gemischte Term mit $z_1 z_2$ verschwunden. Um das geometrisch zu interpretieren, betrachten wir die *Kurve*

$$C := \{x \in \mathbb{R}^2 : q(x) = 1\}.$$

Ist $\lambda_1 > 0$ und $\lambda_2 \neq 0$, so setzen wir

$$a := \frac{1}{\sqrt{\lambda_1}} \quad \text{und} \quad b := \frac{1}{\sqrt{|\lambda_2|}}.$$

Dann lautet die Gleichung von C in den z -Koordinaten

$$\frac{z_1^2}{a^2} \pm \frac{z_2^2}{b^2} = 1.$$

Im Fall $+$ ist das eine *Ellipse*, im Fall $-$ eine *Hyperbel*, a und b sind jeweils die *Hauptachsen*.

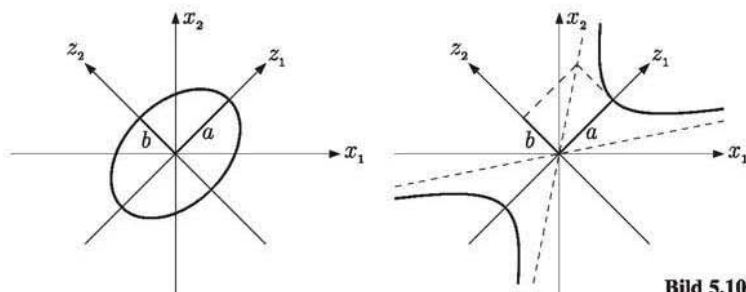


Bild 5.10

Dieser geometrische Aspekt der Bilinearformen wird ausführlich in [Fi] behandelt.

5.7.3. Nun geben wir einige Folgerungen aus der Hauptachsentransformation in 5.7.1 an

Korollar 1. Eine symmetrische Matrix $A \in M(n \times n; \mathbb{R})$ ist genau dann positiv definit, wenn ihre Eigenwerte $\lambda_1, \dots, \lambda_n$ positiv sind.

Beweis. Jeden Spaltenvektor $x \in \mathbb{R}^n$ stellen wir mit Hilfe einer Orthonormalbasis aus Eigenvektoren als

$$x = \mu_1 w_1 + \dots + \mu_n w_n$$

dar. Dann folgt unter Benutzung der Bilinearität

$${}^t x A x = \lambda_1 \mu_1^2 + \dots + \lambda_n \mu_n^2,$$

und daraus liest man sofort die Behauptung ab. $\square$

Um die positive Definitheit einer Matrix zu testen, muß man die Eigenwerte nicht ausrechnen; es reicht zu wissen, daß sie irgendwelche positiven Werte haben. Dazu genügt es, das charakteristische Polynom zu berechnen und die Vorzeichenregel aus 1.3.11 anzuwenden. Es folgt sofort

Korollar 2. Sei $A \in M(n \times n; \mathbb{R})$ eine symmetrische Matrix und

$$P_A = (-1)^n t^n + \alpha_{n-1} t^{n-1} + \dots + \alpha_1 t + \alpha_0$$

ihr charakteristisches Polynom. Dann gilt:

$$A \text{ positiv definit} \iff (-1)^j \alpha_j > 0 \text{ für } j = 0, \dots, n-1.$$

$\square$

5.7.4. Wir betrachten nun einen endlichdimensionalen $\mathbb{R}$ -Vektorraum V mit einer symmetrischen Bilinearform s und der zugehörigen quadratischen Form

$$q: V \rightarrow \mathbb{R}, \quad v \mapsto q(v) = s(v, v).$$

Das Ziel ist, q bezüglich einer geeigneten Basis von V möglichst einfach darzustellen. Dazu erklärt man zunächst

$$V_0 := \{v \in V: s(v, w) = 0 \text{ für alle } w \in V\} \subset V$$

als *Ausartungsraum* von s . Das ist offenbar ein Untervektorraum von V und

$$\text{rang}(s) := \dim V - \dim V_0$$

heißt der *Rang* von s (vgl. dazu Aufgabe 7).

Diagonalisierung einer quadratischen Form. Sei V ein $\mathbb{R}$ -Vektorraum mit $n := \dim V$ und s eine symmetrische Bilinearform mit der zugehörigen quadratischen Form q auf V . Dann gibt es eine Basis

$$\mathcal{B} = (v_1, \dots, v_k, v_{k+1}, \dots, v_r, v_{r+1}, \dots, v_n)$$

von V mit folgenden Eigenschaften: Es ist $r = \text{rang}(s)$ und für

$$v = \sum_{i=1}^n \alpha_i v_i \in V \quad \text{gilt} \quad q(v) = \sum_{i=1}^k \alpha_i^2 - \sum_{i=k+1}^r \alpha_i^2.$$

Insbesondere hat man eine Zerlegung

$$V = V_+ \oplus V_- \oplus V_0$$

mit $q(v) > 0$ für $0 \neq v \in V_+$, $q(v) < 0$ für $0 \neq v \in V_-$ und $q(v) = 0$ für $v \in V_0$.

Beweis. Wir wählen zunächst eine beliebige Basis $\mathcal{A}$ von V und betrachten die darstellende Matrix $A := M_{\mathcal{A}}(s)$. Mit Hilfe von Teil b) des Satzes über die Hauptachsentransformation aus 5.7.1 erhalten wir eine neue Basis $\mathcal{B} = (v_1, \dots, v_n)$ von V mit $s(v_i, v_j) = 0$ für $i \neq j$ und

$$\begin{aligned} q(v_i) &= +1 && \text{für } i = 1, \dots, k, \\ q(v_i) &= -1 && \text{für } i = k+1, \dots, r, \\ q(v_i) &= 0 && \text{für } i = r+1, \dots, n. \end{aligned}$$

Daraus folgt sofort die behauptete Darstellung von $q(v)$. Setzt man

$$V_+ := \text{span}(v_1, \dots, v_k) \quad \text{und} \quad V_- := \text{span}(v_{k+1}, \dots, v_r),$$

so erhält man die gewünschte direkte Summe. $\square$

Zur Untersuchung der Vorzeichen der Werte von q kann man die Mengen

$$C_+ := \{v \in V : q(v) > 0\} \cup \{0\}$$

$$C_- := \{v \in V : q(v) < 0\} \cup \{0\}$$

$$C_0 := \{v \in V : q(v) = 0\}$$

betrachten. Das sind jedoch fast nie Untervektorräume (Aufgabe 6). Es gilt

$$V_+ \subset C_+, \quad V_- \subset C_-, \quad V_0 \subset C_0,$$

aber diese Untervektorräume sind nicht eindeutig bestimmt. Ein klassisches Ergebnis besagt, dass wenigstens die Dimensionen eindeutig bestimmt sind.

Trägheitsgesetz von SYLVESTER. *Hat man für eine quadratische Form q auf einem $\mathbb{R}$ -Vektorraum V zwei Zerlegungen*

$$V = V_+ \oplus V_- \oplus V_0 = V'_+ \oplus V'_- \oplus V_0$$

mit $q(v) > 0$ für $0 \neq v \in V_+$ und $0 \neq v \in V'_+$,

sowie $q(v) < 0$ für $0 \neq v \in V_-$ und $0 \neq v \in V'_-$,

so folgt

$$\dim V'_+ = \dim V_+ \quad \text{und} \quad \dim V'_- = \dim V_-.$$

Die Zahlen $r_+(q) := \dim V_+$ und $r_-(q) := \dim V_-$ sind also neben dem Rang weitere Invarianten der quadratischen Form. Das Paar $(r_+(q), r_-(q))$ wird auch *Signatur* von q genannt.

Beweis. Es genügt folgendes zu zeigen: Ist $W \subset V$ ein Untervektorraum mit $q(w) > 0$ für $0 \neq w \in W$, so folgt

$$\dim W \leq \dim V_+. \quad (*)$$

Aufgrund der Symmetrie folgt dann $\dim V'_+ = \dim V_+$ und analog erhält man $\dim V'_- = \dim V_-$.

Angenommen $\dim W > \dim V_+ = \dim V - \dim(V_- \oplus V_0)$. Nach der Dimensionsformel aus 1.6.1 gibt es ein

$$0 \neq w \in W \cap (V_- \oplus V_0),$$

also $w = v_- + v_0$ mit $v_- \in V_-$ und $v_0 \in V_0$. Dann ist aber

$$q(w) = s(v_- + v_0, v_- + v_0) = s(v_-, v_-) + 0 \leq 0,$$

mit Widerspruch zu $q(w) > 0$. □

Für quadratische Matrizen weiß man, daß die Eigenwerte beim Übergang von A zu SAS^{-1} erhalten bleiben, beim Übergang zu tSAS dagegen im allgemeinen verändert werden. Aus dem Trägheitsgesetz folgt sofort

Korollar. Sei $A \in M(n \times n; \mathbb{R})$ symmetrisch und $S \in GL(n; \mathbb{R})$. Dann haben

$$A \quad \text{und} \quad {}^tSAS$$

mit Vielfachheit gezählt die gleichen Anzahlen positiver und negativer Eigenwerte.

5.7.5. Bisher haben wir Hauptachsentransformationen nur im reellen Fall behandelt, weil er besonders wichtig ist und im engen Zusammenhang zu den Skalarprodukten steht. Dabei bleibt ein Problem: Zur Kontrolle der Definitheit einer Matrix hat man entsprechend 5.7.3 das charakteristische Polynom, d.h. eine Determinante zu berechnen. Weil darin die Unbestimmte t auftritt, helfen Umformungen wenig und der Rechenaufwand ist für große Matrizen enorm. Daher ist es nützlich, eine einfachere Methode zu haben, bei der die Matrix schrittweise umgeformt wird. Überdies geht das für jeden Körper K , in dem $1 + 1 \neq 0$ ist. Zunächst geben wir als Verallgemeinerung des Orthonormalisierungssatzes in 5.4.9, den

Orthogonalisierungssatz. Sei V ein endlichdimensionaler Vektorraum über einem Körper K mit $\text{char}(K) \neq 2$ und

$$s: V \times V \rightarrow K$$

eine symmetrische Bilinearform. Dann gibt es eine Basis $\mathcal{B} = (v_1, \dots, v_n)$ von V mit

$$s(v_i, v_j) = 0 \quad \text{für } i \neq j.$$

Ist $q: V \rightarrow K$ die zugehörige quadratische Form und $\alpha_i := q(v_i)$, so folgt

$$q(v) = \alpha_1 x_1^2 + \dots + \alpha_n x_n^2 \quad \text{für } v = x_1 v_1 + \dots + x_n v_n.$$

Die quadratische Form ist also in diesen Koordinaten frei von „gemischten Termen“ $x_i x_j$ mit $i \neq j$.

Beweis. Wir führen Induktion über $n = \dim V$. Für $n = 1$ ist nichts zu beweisen. Ist $q(v) = 0$ für alle v , so folgt $s(v, w) = 0$ für alle $v, w \in V$ nach der Polarisierung in 5.4.4. Andernfalls gibt es ein $v_1 \in V$ mit $q(v_1) \neq 0$. Wir betrachten

$$W := \{w \in V : s(v_1, w) = 0\}$$

und behaupten $V = Kv_1 \oplus W$. Ist $v \in Kv_1 \cap W$, so folgt

$$v = \lambda v_1 \quad \text{und} \quad 0 = s(v_1, v) = \lambda s(v_1, v_1), \quad \text{also} \quad \lambda = 0 \quad \text{und} \quad v = 0.$$

Zum Nachweis von $V = Kv_1 + W$ definieren wir zu $v \in V$

$$v' := \frac{s(v_1, v)}{s(v_1, v_1)} v_1.$$

Dann ist $v = v' + (v - v')$, und aus $s(v_1, v) = s(v_1, v')$ folgt $v - v' \in W$.

Bezeichnet $\tilde{s}$ die Beschränkung von s auf W , so gibt es nach Induktionsannahme eine Basis $(v_2, \dots, v_n)$ von W mit

$$\tilde{s}(v_i, v_j) = 0 \quad \text{für } i, j \in \{2, \dots, n\} \text{ und } i \neq j.$$

Also kann man $B := (v_1, v_2, \dots, v_n)$ wählen. □

Mit Hilfe der Transformationsformel 5.4.3 folgt das

Korollar. Zu einer symmetrischen Matrix $A \in M(n \times n; K)$ gibt es ein $S \in GL(n; K)$, so daß

$${}^tSAS = \begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix}.$$

□

Es sei noch einmal bemerkt, daß die Zahlen α_i nicht eindeutig bestimmt sind. Im reellen Fall sind nach dem Trägheitssatz 5.7.4 jedoch die Vorzeichen festgelegt.

5.7.6. Zur Berechnung einer Transformationsmatrix S wie in obigem Korollar gibt es eine *symmetrische Umformungsmethode*. Dabei werden an A simultan Zeilen- und Spaltenumformungen durchgeführt, bis eine Diagonalmatrix entstanden ist. Die Zeilenumformungen werden ausgedehnt auf die danebenstehende Matrix E_n . Das kann man mit Hilfe von Elementarmatrizen C_i schematisch

so beschreiben:

A	E_n
${}^t C_1 \cdot A \cdot C_1$	$E_n \cdot C_1$
$\vdots$	$\vdots$
${}^t C_r \cdot \dots \cdot {}^t C_1 \cdot A \cdot C_1 \cdot \dots \cdot C_r$	$E_n \cdot C_1 \cdot \dots \cdot C_r$

Ist links eine Diagonalmatrix D entstanden, so hat man rechts S mit

$$D = {}^t S A S.$$

Wegen des Assoziativgesetzes der Matrizenmultiplikation ist es gleichgültig, ob man in jedem Schritt zuerst die Spalten oder die Zeilen umformt. Eine Strategie für die Umformungen wollen wir nicht allgemein aufschreiben, wir geben dafür zwei typische

Beispiele. a) Wir betrachten auf $\mathbb{R}^2$ die quadratische Form $q(x_1, x_2) = x_1 x_2$. Setzt man $x_1 = y_1 + y_2$ und $x_2 = y_1 - y_2$, so ist

$$q(x(y)) = (y_1 + y_2)(y_1 - y_2) = y_1^2 - y_2^2.$$

Mit Matrizen geht das so: zu q gehört die Bilinearform

$$s(x, \tilde{x}) = \frac{1}{2} x_1 \tilde{x}_2 + \frac{1}{2} \tilde{x}_1 x_2.$$

Ein mögliches Umformungsschema ist folgendes.

$$A = \begin{array}{cc|cc} 0 & \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & 0 & 0 & 1 \\ \hline 1 & \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & 0 & 1 & 1 \\ \hline 1 & 0 & 1 & -\frac{1}{2} \\ 0 & -\frac{1}{4} & 1 & \frac{1}{2} \\ \hline 1 & 0 & 1 & 1 \\ 0 & -1 & 1 & -1 \end{array} = E_2$$

$$D = \begin{array}{cc|cc} 1 & 0 & 1 & 1 \\ 0 & -1 & 1 & -1 \end{array} = S$$

Entsprechend 5.4.3 ist $x = Sy$.

b) Wir betrachten wieder auf $\mathbb{R}^2$ eine quadratische Form mit $a \neq 0$ und führen *quadratische Ergänzung* durch:

$$\begin{aligned}
q(x) &= ax_1^2 + 2bx_1x_2 + cx_2^2 \\
&= a \left(x_1^2 + 2\frac{b}{a}x_1x_2 + \frac{b^2}{a^2}x_2^2 \right) + cx_2^2 - \frac{b^2}{a}x_2^2 \\
&= a \left(x_1 + \frac{b}{a}x_2 \right)^2 + \left(c - \frac{b^2}{a} \right) x_2^2 \\
&= ay_1^2 + \left(c - \frac{b^2}{a} \right) y_2^2,
\end{aligned}$$

wenn $y_1 = x_1 + \frac{b}{a}x_2$ und $y_2 = x_2$.

In Matrizen geht das viel schneller:

$$\begin{aligned}
A &= \begin{pmatrix} a & b \\ b & c \end{pmatrix} = E_2 \\
D &= \begin{pmatrix} a & 0 \\ 0 & c - \frac{b^2}{a} \end{pmatrix} = S
\end{aligned}$$

Es ist wieder $x = Sy$. Weiter gilt

$$c - \frac{b^2}{a} = \frac{ac - b^2}{a} = \frac{\det A}{a}.$$

Also ist A genau dann positiv definit, wenn $a > 0$ und $\det A > 0$.

5.7.7. Nun kommen wir zu dem versprochenen Kriterium für die positive Definitheit einer Matrix $A \in M(n \times n; \mathbb{R})$. Hat man entsprechend 5.7.6 ein $S \in GL(n; \mathbb{R})$ gefunden, so daß

$${}^tSAS = \begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix},$$

so ist A nach dem Trägheitssatz 5.7.4 genau dann positiv definit, wenn

$$\alpha_i > 0 \quad \text{für } i = 1, \dots, n.$$

Das kann man noch etwas anders formulieren. Dazu bezeichnen wir für $k = 1, \dots, n$ mit

$$A_k \in M(k \times k; \mathbb{R})$$

die linke obere k -reihige und k -spaltige Teilmatrix von A . Ihre Determinante $\det A_k$ heißt *Hauptminor* von A . Schon bei JACOBI findet man das

Hauptminoren-Kriterium für Definitheit. Für eine symmetrische Matrix $A \in M(n \times n; \mathbb{R})$ gilt:

$$A \text{ positiv definit} \Leftrightarrow \det A_k > 0 \quad \text{für } k = 1, \dots, n.$$

Beweis. „ $\Rightarrow$ “: Jede positiv definite Matrix hat positive Determinante, denn ist

$${}^tSAS = \begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix} \quad \text{mit } \alpha_i > 0 \quad \text{für } i = 1, \dots, n,$$

so ist $\det A \cdot (\det S)^2 = \alpha_1 \cdot \dots \cdot \alpha_n$. Die Matrix A_k beschreibt die Beschränkung der zu A gehörigen Bilinearform auf

$$\{(x_1, \dots, x_n) \in \mathbb{R}^n : x_{k+1} = \dots = x_n = 0\},$$

also ist auch A_k positiv definit und somit $\det A_k > 0$.

„ $\Leftarrow$ “: Wir führen Induktion über n . Der Fall $n = 1$ ist klar. Nach Induktionsannahme ist A_{n-1} positiv definit, also gibt es ein $T \in \text{GL}(n-1; \mathbb{R})$ mit

$${}^tT \cdot A_{n-1} \cdot T = \begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_{n-1} \end{pmatrix} \quad \text{und } \alpha_i > 0 \quad \text{für } i = 1, \dots, n-1.$$

Wir definieren

$$T' := \left(\begin{array}{ccc|c} & & & 0 \\ & T & & \vdots \\ & & & 0 \\ \hline 0 & \dots & 0 & 1 \end{array} \right) \in \text{GL}(n; \mathbb{R}).$$

Dann ist

$${}^tT'AT' = \left(\begin{array}{ccc|c} \alpha_1 & & 0 & \beta_1 \\ & \ddots & & \vdots \\ 0 & & \alpha_{n-1} & \beta_{n-1} \\ \hline \beta_1 & \dots & \beta_{n-1} & \beta_n \end{array} \right) =: B.$$

Nach Voraussetzung ist $\det A = \det A_n > 0$, also auch $\det B > 0$. Wir setzen

$$S := \left(\begin{array}{ccc|c} & & & \gamma_1 \\ & E_{n-1} & & \vdots \\ & & & \gamma_{n-1} \\ \hline 0 & \dots & 0 & 1 \end{array} \right) \quad \text{mit } \gamma_i = -\frac{\beta_i}{\alpha_i},$$

daraus folgt

$${}^tSBS = \begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix}.$$

Wegen $0 < \det B = \frac{\alpha_1 \cdot \dots \cdot \alpha_n}{(\det S)^2}$ folgt auch $\alpha_n > 0$. $\square$

Dieses Kriterium ist mehr von theoretischem als von praktischem Interesse, weil dabei n Minoren auszurechnen sind. Schneller ist die Umformungsmethode aus 5.7.6.

Aufgaben zu 5.7

1. Sei s die symmetrische Bilinearform auf dem $\mathbb{R}^3$, die gegeben ist durch die Matrix

$$\begin{pmatrix} 3 & -2 & 0 \\ -2 & 2 & -2 \\ 0 & -2 & 1 \end{pmatrix}.$$

Bestimmen Sie eine Basis $\mathcal{A}$ des $\mathbb{R}^3$, so dass $M_{\mathcal{A}}(s)$ Diagonalgestalt hat und eine weitere Basis $\mathcal{B}$, so dass

$$M_{\mathcal{B}}(s) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

2. Sei $\mathcal{D} = \mathcal{D}(-1, 1[; \mathbb{R})$ der Vektorraum der auf $] -1, 1[$ differenzierbaren Funktionen.

a) Zeigen Sie, dass $d: \mathcal{D} \times \mathcal{D} \rightarrow \mathbb{R}$, $(f, g) \mapsto (fg)'(0)$ eine symmetrische Bilinearform ist.

b) Bestimmen Sie den Ausartungsraum $\mathcal{D}_0$ von d .

3. Diagonalisieren Sie die Matrizen

$$\begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 4 \\ 2 & 4 & 4 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 2 \\ 1 & 1 & 0 & 0 \\ 0 & 2 & 0 & 2 \end{pmatrix}$$

mit der symmetrischen Umformungsmethode aus 5.7.6.

4. Eine symmetrische Matrix $A \in M(n \times n; \mathbb{R})$ heißt *negativ definit*, wenn

$${}^t x A x < 0$$

für jedes $0 \neq x \in \mathbb{R}^n$. Zeigen Sie: A negativ definit $\Leftrightarrow -A$ positiv definit.

5. Überprüfen Sie die folgenden Matrizen auf Definitheit:

$$\begin{pmatrix} 1 & 2 & -2 \\ 2 & 2 & 0 \\ -2 & 0 & -4 \end{pmatrix}, \quad \begin{pmatrix} -3 & 1 & -3 \\ 1 & -2 & 0 \\ -3 & 0 & -4 \end{pmatrix}, \quad \begin{pmatrix} 7 & 0 & -8 \\ 0 & 1 & 2 \\ -8 & 2 & 17 \end{pmatrix}.$$

6. Eine Teilmenge C eines K -Vektorraumes V heißt *Kegel*, wenn für $v \in C$ und $\lambda \in K$ auch $\lambda \cdot v \in C$ gilt.

Man zeige, dass die in 5.7.4 erklärten Mengen C_0 , C_+ und C_- im Fall $K = \mathbb{R}$ Kegel sind, und man bestimme diese Kegel sowie V_0 explizit für $V = \mathbb{R}^2$, wenn s durch die Matrix

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

expliziert ist.

7. Sei s eine symmetrische Bilinearform auf V , $n := \dim V$, $\mathcal{B}$ eine Basis von V und $A := M_{\mathcal{B}}(s)$. Ist

$$L := \{x \in \mathbb{R}^n : A \cdot x = 0\}$$

Lösungsraum des homogenen Gleichungssystems und $\Phi_{\mathcal{B}}: \mathbb{R}^n \rightarrow V$ das Koordinatensystem zu $\mathcal{B}$, so folgt

$$V_0 = \Phi_{\mathcal{B}}(L).$$

Insbesondere gilt $\text{rang}(s) = \text{rang}(A)$ und $\dim V_0 = n - \text{rang } A$.

Kapitel 6

Dualität und Tensorprodukte*

In diesem letzten Kapitel werden noch einige Dinge angefügt, die höchstens dadurch Schwierigkeiten bereiten, daß sie relativ abstrakt sind. Das ist vergleichbar mit den Atembeschwerden, die sich im Hochgebirge wegen der immer dünner werdenden Luft einstellen können.

6.1 Dualräume

Vektoren des Standardraumes K^n hatten wir als Zeilen oder Spalten geschrieben, je nachdem, was gerade günstiger war. Daraus kann man eine schöne Theorie machen, die eine neue Sicht auf lineare Gleichungssysteme und quadratische Matrizen eröffnet. Was hier zunächst als reine Spielerei erscheinen mag, ist ein wichtiges Hilfsmittel der theoretischen Physik.

6.1.1. Die linke Seite einer linearen Gleichung hat die Form

$$a_1x_1 + \dots + a_nx_n.$$

Schreibt man $x = {}^t(x_1, \dots, x_n)$ als Spalte und $a = (a_1, \dots, a_n)$ als Zeile, so kann man x als Element des K^n und a als eine lineare Abbildung

$$a: K^n \rightarrow K, \quad x \mapsto a \cdot x = a_1x_1 + \dots + a_nx_n,$$

d.h. als Element von $\text{Hom}(K^n, K)$ betrachten. Das ist nach 2.1.3 wieder ein Vektorraum.

Definition. Ist V ein K -Vektorraum, so heißt

$$V^* := \text{Hom}_K(V, K) = \{\varphi: V \rightarrow K: \varphi \text{ linear}\}$$

der *Dualraum* von V . Die Elemente von V^* heißen *Linearformen* auf V .

Im folgenden wird sich zeigen, daß der Name Dualraum gerechtfertigt ist. Besonders interessant ist der Fall, daß V ein unendlichdimensionaler Funktionenraum ist.

Beispiel. Sei $I = [a, b] \subset \mathbb{R}$ ein Intervall und $V := \mathcal{D}(I; \mathbb{R})$ der Vektorraum der auf I differenzierbaren Funktionen. Dann sind

$$\sigma: V \rightarrow \mathbb{R}, \quad f \mapsto \int_a^b f(x) dx$$

und

$$\delta: V \rightarrow \mathbb{R}, \quad f \mapsto \frac{df}{dx}(c),$$

wobei $a < c < b$, zwei Linearformen auf V .

Dieses Beispiel mag als Andeutung für den Nutzen der Dualräume in der Analysis dienen, das ist Gegenstand der Funktionalanalysis.

6.1.2. In der linearen Algebra kann man die Dualitätstheorie nur für endlichdimensionale Vektorräume entwickeln. Dies sei nun stets vorausgesetzt. Ist $\mathcal{B} = (v_1, \dots, v_n)$ eine Basis von V , so gibt es zu jedem $i \in \{1, \dots, n\}$ nach 2.4.1 genau eine lineare Abbildung

$$v_i^*: V \rightarrow K \quad \text{mit} \quad v_i^*(v_j) = \delta_{ij}. \quad (*)$$

Vorsicht! Die Linearform v_i^* hängt nicht nur vom Vektor v_i , sondern auch von den anderen Basisvektoren ab!

Bemerkung. Für jede Basis $\mathcal{B} = (v_1, \dots, v_n)$ von V ist $\mathcal{B}^* = (v_1^*, \dots, v_n^*)$ eine Basis von V^* .

Man nennt $\mathcal{B}^*$ die zu $\mathcal{B}$ *duale Basis*.

Beweis. Dies folgt aus 2.4.2, wir wollen es noch einmal direkt nachrechnen. Zu jedem $\varphi \in V^*$ sind $\lambda_1, \dots, \lambda_n$ gesucht, so daß

$$\varphi = \lambda_1 v_1^* + \dots + \lambda_n v_n^*.$$

Setzt man v_i in φ ein, so ergibt sich aus (*), daß $\lambda_i = \varphi(v_i)$ sein muß. Daraus folgt alles. $\square$

Da man jeden Vektor $v \neq 0$ zu einer Basis ergänzen kann, folgt das

Korollar 1. Zu jedem $v \in V$ mit $v \neq 0$ gibt es ein $\varphi \in V^*$ mit $\varphi(v) \neq 0$. $\square$

Mit Hilfe von 2.4.1 erhält man

Korollar 2. Zu jeder Basis $\mathcal{B} = (v_1, \dots, v_n)$ von V gibt es einen Isomorphismus

$$\Psi_{\mathcal{B}}: V \rightarrow V^* \quad \text{mit} \quad \Psi_{\mathcal{B}}(v_i) = v_i^*. \quad \square$$

Vorsicht! Dieser Isomorphismus hängt von der Auswahl der Basis ab, ebenso ist das φ aus Korollar 1 abhängig von der Ergänzung.

Beispiele. a) Im K^n hat man zur kanonischen Basis $\mathcal{K} = (e_1, \dots, e_n)$ die *kanonische Basis*

$$\mathcal{K}^* = (e_1^*, \dots, e_n^*) \quad \text{von } V^*.$$

Mit der Konvention, Vektoren als Spalten und Linearformen als Zeilen zu schreiben, ist

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

mit der 1 an der i -ten Stelle.

b) Im K^2 betrachten wir neben der kanonischen Basis die Basis $\mathcal{B} = (v_1, v_2)$ mit $v_1 = e_1$ und $v_2 = {}^t(1, 1)$. Aus $e_1 = v_1$ und $e_2 = v_2 - v_1$ folgt

$$v_1^*(e_1) = 1, \quad v_1^*(e_2) = -1, \quad v_2^*(e_1) = 0, \quad v_2^*(e_2) = 1,$$

$$v_1^* = e_1^* - e_2^*, \quad v_2^* = e_2^*, \quad \text{also} \quad \Psi_{\mathcal{B}}(e_1) = e_1^* - e_2^*, \quad \Psi_{\mathcal{B}}(e_2) = -e_1^* + 2e_2^*.$$

Vorsicht! Man beachte die Tücke der Notation: Es ist $v_1 = e_1$, aber $v_1^* \neq e_1^*$, weil die beiden Sternchen verschiedene Bedeutung haben.

6.1.3. In 0.2.4 hatten wir für eine parametrisierte Gerade in der Ebene eine Gleichung bestimmt. Diese Frage kann man nun noch einmal abstrakter betrachten. Ist $0 \neq x = {}^t(x_1, x_2) \in \mathbb{R}^2$ und $L = \mathbb{R} \cdot x \subset \mathbb{R}^2$ die Gerade durch 0 und x , so sind die $a = (a_1, a_2) \in (\mathbb{R}^2)^*$ gesucht mit

$$a_1x_1 + a_2x_2 = 0.$$

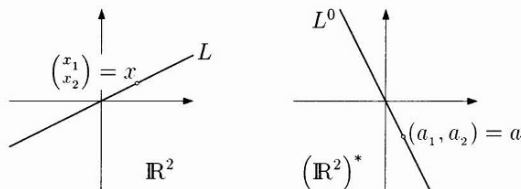


Bild 6.1

Sie liegen auf einer Geraden L^0 in $(\mathbb{R}^2)^*$, die senkrecht auf L steht, wenn man $\mathbb{R}^2$ und $(\mathbb{R}^2)^*$ nicht unterscheidet. Allgemeiner kann man für einen Untervektorraum alle „beschreibenden Gleichungen“ zusammenfassen.

Definition. Ist V ein K -Vektorraum und $U \subset V$ ein Untervektorraum, so heißt

$$U^0 := \{\varphi \in V^*: \varphi(u) = 0 \text{ für alle } u \in U\} \subset V^*$$

der zu U orthogonale Raum (oder der Annulator von U). Das ist offensichtlich ein Untervektorraum.

Man beachte den Unterschied zur Orthogonalität in Räumen mit Skalarprodukt: dort liegt das orthogonale Komplement $U^\perp$ in V , hier liegt U^0 in V^* . Dennoch gibt es starke Beziehungen:

Satz. Für jeden Untervektorraum $U \subset V$ gilt

$$\dim U^0 = \dim V - \dim U.$$

Genauer gilt: Ist $(u_1, \dots, u_k)$ Basis von U und $\mathcal{B} = (u_1, \dots, u_k, v_1, \dots, v_r)$ Basis von V , so bilden die Linearformen $v_1^*, \dots, v_r^*$ aus $\mathcal{B}^*$ eine Basis von U^0 .

Beweis. Da $v_1^*, \dots, v_r^*$ einer Basis entnommen wurden, sind sie linear unabhängig. Also bleibt

$$U^0 = \text{span}(v_1^*, \dots, v_r^*)$$

zu zeigen. „ $\supset$ “ ist klar, da $v_j^*(u_i) = 0$. Zum Nachweis von „ $\subset$ “ sei $\varphi \in U^0$ und

$$\varphi = \mu_1 u_1^* + \dots + \mu_k u_k^* + \lambda_1 v_1^* + \dots + \lambda_r v_r^*.$$

Setzt man u_i ein, so wird $0 = \varphi(u_i) = \mu_i$. □

6.1.4. Nun zeigen wir, daß man nicht nur Vektorräume, sondern auch lineare Abbildungen dualisieren kann. Dazu betrachten wir das Diagramm

$$\begin{array}{ccc} V & \xrightarrow{F} & W \\ & \searrow \psi \circ F & \downarrow \psi \\ & & K \end{array}$$

mit K -Vektorräumen V und W , sowie linearen Abbildungen F und ψ . Dann ist $\psi \in W^*$, und mit Hilfe von 2.1.3 folgt $\psi \circ F \in V^*$; also können wir eine *duale Abbildung*

$$F^*: W^* \rightarrow V^*, \quad \psi \mapsto F^*(\psi) := \psi \circ F,$$

erklären. Aus

$$\begin{aligned} F^*(\lambda_1 \psi_1 + \lambda_2 \psi_2) &= (\lambda_1 \psi_1 + \lambda_2 \psi_2) \circ F = \lambda_1 (\psi_1 \circ F) + \lambda_2 (\psi_2 \circ F) \\ &= \lambda_1 F^*(\psi_1) + \lambda_2 F^*(\psi_2) \end{aligned}$$

folgt die Linearität von F^* . Also hat man noch abstrakter eine Abbildung

$$\text{Hom}_K(V, W) \rightarrow \text{Hom}_K(W^*, V^*), \quad F \mapsto F^*,$$

die ein Vektorraumisomorphismus ist (Aufgabe 3). Das kann man auch mit Hilfe von Matrizen sehen:

Satz. Gegeben seien K -Vektorräume V und W mit Basen $\mathcal{A}$ und $\mathcal{B}$, sowie eine lineare Abbildung $F: V \rightarrow W$. Dann gilt:

$$M_{\mathcal{A}^*}^{\mathcal{B}^*}(F^*) = {}^t(M_{\mathcal{B}}^{\mathcal{A}}(F)).$$

Kurz ausgedrückt: Die *duale Abbildung* wird bezüglich der dualen Basen durch die *transponierte Matrix* beschrieben.

Beweis. Sei $\mathcal{A} = (v_1, \dots, v_n)$ und $\mathcal{B} = (w_1, \dots, w_m)$. Entsprechend 2.4.2 bedeutet $M_{\mathcal{B}}^{\mathcal{A}}(F) = (a_{ij})$, daß

$$F(v_j) = \sum_{i=1}^m a_{ij} w_i, \quad \text{also} \quad a_{ij} = w_i^*(F(v_j)) = F^*(w_i^*)(v_j).$$

Ist $M_{\mathcal{A}^*}^{B^*}(F^*) = (b_{ji})$, so folgt

$$F^*(w_i^*) = \sum_{j=1}^n b_{ji} v_j^*, \quad \text{also} \quad b_{ji} = F^*(w_i^*)(v_j),$$

und insgesamt $a_{ij} = b_{ji}$. □

6.1.5. Zu einer linearen Abbildung $F: V \rightarrow W$ hat man ein Diagramm

$$\text{Ker } F \subset V \xrightarrow{F} \text{Im } F \subset W,$$

und entsprechend für die duale Abbildung

$$V^* \supset \text{Im } F^* \xleftarrow{F^*} W^* \supset \text{Ker } F^*.$$

Einen Zusammenhang zwischen diesen beiden Folgen erhält man mit Hilfe der in 6.1.3 erklärten Orthogonalität:

Satz. Ist $F: V \rightarrow W$ eine lineare Abbildung zwischen endlichdimensionalen Vektorräumen, so gilt

$$\text{Im } F^* = (\text{Ker } F)^0 \quad \text{und} \quad \text{Ker } F^* = (\text{Im } F)^0.$$

Korollar 1. Unter den obigen Voraussetzungen gilt

$$\text{rang } F^* = \text{rang } F.$$

Beweis von Korollar 1. Mit Hilfe von 2.2.4 und 6.1.3 folgt

$$\begin{aligned} \text{rang } F^* &= \dim \text{Im } F^* = \dim (\text{Ker } F)^0 = \dim V - \dim \text{Ker } F \\ &= \dim \text{Im } F = \text{rang } F. \end{aligned}$$
□

Im Spezialfall $V = K^n$ und $W = K^m$ ist F durch eine Matrix A und F^* nach 6.1.4 durch $'A$ beschrieben. Daher ergibt Korollar 1 einen neuen und sehr abstrakten Beweis von

Korollar 2. Für jede Matrix $A \in M(m \times n; K)$ gilt

$$\text{Zeilenrang } A = \text{Spaltenrang } A.$$
□

Beweis des Satzes. Zum Nachweis der ersten Gleichung sind zwei Inklusionen zu zeigen: Dazu betrachten wir das Diagramm

$$\begin{array}{ccc}
 V & \xrightarrow{F} & W \\
 & \searrow \varphi & \downarrow \psi \\
 & & K
 \end{array}$$

Ist $\varphi = \psi \circ F$, so folgt $\varphi|_{\text{Ker } F} = 0$, also gilt „ $\subset$ “.

Sei umgekehrt $\varphi \in V^*$ mit $\varphi|_{\text{Ker } F} = 0$ gegeben. Zur Konstruktion eines $\psi \in W^*$ mit $\varphi = \psi \circ F$ wählen wir entsprechend 2.2.4 Basen

$$\mathcal{A} = (u_1, \dots, u_r, v_1, \dots, v_k) \text{ von } V \text{ und}$$

$$\mathcal{B} = (w_1, \dots, w_r, w_{r+1}, \dots, w_m) \text{ von } W$$

mit $\text{Ker } F = \text{span}(v_1, \dots, v_k)$, $\text{Im } F = \text{span}(w_1, \dots, w_r)$ und $F(u_i) = w_i$ für $i = 1, \dots, r$. Nach 2.4.1 gibt es genau ein lineares ψ mit

$$\psi(w_i) := \begin{cases} \varphi(u_i) & \text{für } i = 1, \dots, r, \\ 0 & \text{sonst.} \end{cases}$$

Nach Konstruktion von ψ ist $\varphi = \psi \circ F$.

Die zweite Gleichung kann man ähnlich beweisen oder mit Hilfe von 6.1.7 aus der ersten folgern. $\square$

6.1.6. Den Dualraum kann man zu jedem Vektorraum bilden, also auch zu V^* . Auf diese Weise erhält man zu V den *Bidualraum*

$$V^{**} := (V^*)^* = \text{Hom}(V^*, K).$$

Die in 6.1.2 konstruierten Isomorphismen $V \rightarrow V^*$ waren von der Auswahl einer Basis abhängig. Dagegen hat man eine *kanonische* Abbildung

$$\iota: V \rightarrow V^{**}, \quad v \mapsto \iota_v, \quad \text{mit } \iota_v(\varphi) := \varphi(v).$$

Mit Hilfe der Korollare 1 und 2 aus 6.1.2 folgt durch einfache Rechnungen der

Satz. Für jeden endlichdimensionalen K -Vektorraum V ist die kanonische Abbildung

$$\iota: V \rightarrow V^{**}$$

ein Isomorphismus. $\square$

Man kann also V mit V^{**} identifizieren und in suggestiver Form

$$v(\varphi) = \varphi(v)$$

schreiben. In dieser Form ist die Dualität zwischen Vektoren und Linearformen besonders deutlich.

Korollar. Für jede lineare Abbildung $F: V \rightarrow W$ gilt $F^{**} = F$. $\square$

Im K^n dualisiert man durch Transponieren:

$$\begin{aligned} K^n &\rightarrow (K^n)^* \rightarrow (K^n)^{**} = K^n, \\ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} &\mapsto (x_1, \dots, x_n) \mapsto \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}. \end{aligned}$$

6.1.7. Die in 6.1.3 eingeführten orthogonalen Räume ermöglichen eine abstrakte Beschreibung linearer Gleichungssysteme. Zunächst eine allgemeine

Bemerkung. Für jeden Untervektorraum $W \subset V$ gilt

$$(W^0)^0 = W \subset V = V^{**}.$$

Beweis. Wegen der Dimensionsformel in 6.1.3 genügt es, $W \subset (W^0)^0$ zu zeigen. Das ist aber klar, denn für $w \in W$ und $\varphi \in W^0$ gilt $w(\varphi) = \varphi(w) = 0$. $\square$

Damit kann man lineare Gleichungssysteme wie folgt interpretieren. Gegeben sei das System

$$A \cdot x = 0 \quad \text{mit } A \in M(m \times n; K) \text{ und } W = \text{Lös}(A, 0) \subset K^n.$$

Die Zeilen $a_1, \dots, a_m$ von A sind Linearformen, wir definieren

$$U := \text{span}(a_1, \dots, a_m) \subset (K^n)^* \quad \text{mit} \quad \dim U = \text{rang } A.$$

Nach Definition des Lösungsraumes gilt

$$W = U^0 \quad \text{und} \quad n = \dim W + \dim U.$$

Das Gleichungssystem lösen bedeutet also, zu vorgegebenem $U \subset (K^n)^*$ eine Basis von $U^0 \subset K^n$ zu finden, und die Zeilenumformungen von A entsprechend 0.4.7 dienen dazu, eine Basis von U zu konstruieren.

Das duale Problem ist folgendes: Gegeben ein Untervektorraum $W \subset K^n$ und gesucht eine Matrix A , so daß

$$W = \text{Lös}(A, 0).$$

Die Zeilen der Matrix A erhält man durch ein Erzeugendensystem des Raumes $U \subset (K^n)^*$ mit

$$U = W^0, \quad \text{d.h.} \quad U^0 = W.$$

W sei gegeben durch eine Matrix $X \in M(n \times l; K)$ mit Spalten $w_1, \dots, w_l$, so daß $W = \text{span}(w_1, \dots, w_l)$. Dann ist

$$U = \{a \in (K^n)^* : a \cdot X = 0\},$$

d.h. man hat das Gleichungssystem

$${}^tX \cdot {}^t a = 0$$

zu lösen. Eine Basis des Lösungsraumes liefert die Zeilen von A . Ist

$$k := \dim W \leq l, \quad \text{so ist} \quad r := n - k = \text{rang } A,$$

und es gilt $A \cdot X = 0$ in $M(r \times l; K)$.

Damit hat man ein Rezept für die *Berechnung von Gleichungen für W* : Man nehme ein Erzeugendensystem von W , forme daraus tX und löse das entsprechende Gleichungssystem.

Beispiel. Seien $w_1 = {}^t(1, 1, 1)$ und $w_2 = {}^t(0, -1, 1) \in \mathbb{R}^3$. Eine Gleichung der davon aufgespannten Ebene W erhält man durch Lösen des Gleichungssystems mit der Matrix

$${}^tX = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 1 \end{pmatrix}.$$

Eine Fundamentallösung ist $(a_1, a_2, a_3) = (-2, 1, 1)$, also lautet eine Gleichung für W

$$-2x_1 + x_2 + x_3 = 0.$$

Das geht auch mit dem Vektorprodukt aus 5.2, denn $w_1 \times w_2 = (2, -1, -1)$.

Aufgaben zu 6.1

1. Gegeben sei ein endlichdimensionaler Vektorraum V mit Basen $\mathcal{A}$ und $\mathcal{B}$. Sind $\mathcal{A}^*$ und $\mathcal{B}^*$ die zugehörigen dualen Basen von V^* , so gilt für die Transformationsmatrizen

$$T_{\mathcal{B}^*}^{\mathcal{A}^*} = ({}^t T_{\mathcal{B}}^{\mathcal{A}})^{-1}.$$

2. Gegeben sei der Untervektorraum

$$U = \text{span} \left(\begin{pmatrix} 2 \\ 3 \\ 1 \\ 4 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 5 \\ 1 \\ -1 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 0 \\ 1 \\ 1 \\ -2 \end{pmatrix} \right) \subset \mathbb{R}^5.$$

Bestimmen Sie eine Basis von U^0 .

3. Zeigen Sie, dass für Vektorräume V, W durch $\text{Hom}(V, W) \rightarrow \text{Hom}(W^*, V^*)$, $F \mapsto F^*$, ein Isomorphismus von Vektorräumen gegeben ist.

4. Sei $F: V \rightarrow W$ ein Homomorphismus und $U \subset W$ ein Untervektorraum. Zeigen Sie: $F^*(U^0) = (F^{-1}(U))^0$.

5. Es seien $W_1, W_2 \subset V$ Untervektorräume. Zeigen Sie:

a) $(W_1 + W_2)^0 = W_1^0 \cap W_2^0$,

b) $(W_1 \cap W_2)^0 = W_1^0 + W_2^0$.

6.2 Dualität und Skalarprodukte

In diesem Abschnitt wollen wir zeigen, wie schön und kanonisch die Dualitätstheorie in euklidischen und unitären Vektorräumen wird.

6.2.1. Zunächst wollen wir den Begriff einer Bilinearform aus 5.4.1 noch leicht verallgemeinern. Gegeben seien K -Vektorräume V und W sowie eine Abbildung

$$b: V \times W \rightarrow K, \quad (v, w) \mapsto b(v, w).$$

Für festes $v \in V$ und $w \in W$ definieren wir

$$\begin{aligned} b_v: W &\rightarrow K, \quad w \mapsto b(v, w), \quad \text{und} \\ b_w: V &\rightarrow K, \quad v \mapsto b(v, w). \end{aligned}$$

Man nennt b *bilinear*, wenn b_v und b_w für alle $v \in V$ und $w \in W$ linear sind.

Ist das der Fall, so erhält man lineare Abbildungen

$$\begin{aligned} b': V &\rightarrow W^*, \quad v \mapsto b_v, \quad \text{und} \\ b'': W &\rightarrow V^*, \quad w \mapsto b_w. \end{aligned}$$

Eine Bilinearform b heißt *nicht ausgeartet*, wenn b' und b'' injektiv sind.

Beispiele dafür sind $V \times V^* \rightarrow K, \quad (v, \varphi) \mapsto \varphi(v),$

für jeden K -Vektorraum V und seinen Dualraum V^* , sowie ein Skalarprodukt

$$V \times V \rightarrow \mathbb{R}, \quad (v, v') \mapsto \langle v, v' \rangle,$$

für jeden $\mathbb{R}$ -Vektorraum V . Das folgt aus Korollar 1 in 6.1.2 und der Definition eines Skalarproduktes in 5.4.6.

Satz. Sind V und W endlichdimensional, und ist

$$b: V \times W \rightarrow K$$

eine nicht ausgeartete Bilinearform, so sind die Abbildungen

$$b': V \rightarrow W^* \quad \text{und} \quad b'': W \rightarrow V^*$$

Isomorphismen.

Beweis. Da b' und b'' injektiv sind, ist nach 6.1.2

$$\dim V \leq \dim W^* = \dim W \quad \text{und} \quad \dim W \leq \dim V^* = \dim V.$$

Also folgt $\dim V = \dim W = \dim V^* = \dim W^*$. □

6.2.2. Aus dem obigen Satz folgt sofort das

Korollar. In einem euklidischen Vektorraum V ist die Abbildung

$$\Psi: V \rightarrow V^*, \quad v \mapsto \langle v, \rangle,$$

ein Isomorphismus. □

Man beachte die suggestive Notation: Für $\varphi = \langle v, \cdot \rangle$ ist $\varphi(v') = \langle v, v' \rangle$. Wegen der Symmetrie des Skalarproduktes kann man Ψ ebenso gut durch $v \mapsto \langle \cdot, v \rangle$ erklären.

Im Gegensatz zu den Isomorphismen $\Psi_B: V \rightarrow V^*$ aus 6.1.2, die von der Wahl der Basis abhängen, ist der obige Isomorphismus kanonisch, wenn ein Skalarprodukt vorgegeben ist.

Im Spezialfall $V = \mathbb{R}^n$ hat man die kanonische Basis $\mathcal{K} = (e_1, \dots, e_n)$ und das kanonische Skalarprodukt $\langle \cdot, \cdot \rangle$ mit

$$\langle e_i, e_j \rangle = \delta_{ij}.$$

In diesem Fall stimmt der Isomorphismus

$$\Psi: \mathbb{R}^n \rightarrow (\mathbb{R}^n)^*, \quad v \mapsto \langle v, \cdot \rangle,$$

mit dem Isomorphismus

$$\Psi_{\mathcal{K}}: \mathbb{R}^n \rightarrow (\mathbb{R}^n)^*, \quad e_i \mapsto e_i^*,$$

aus 6.1.2 überein.

6.2.3. In dem Beispiel aus 6.1.3 kann man neben $L \subset \mathbb{R}^2$ und $L^0 \subset (\mathbb{R}^2)^*$ auch die zu L senkrechte Gerade

$$L^\perp = \{(a_1, a_2) \in \mathbb{R}^2: a_1x_1 + a_2x_2 = 0\}$$

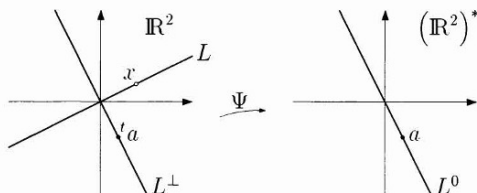


Bild 6.2

betrachten. Ihr Bild unter dem kanonischen Isomorphismus ist L^0 . Allgemein hat man die folgende Beziehung zwischen dem in 5.4.8 definierten orthogonalen Komplement und dem in 6.1.3 definierten Annulator:

Satz. Sei V ein euklidischer Vektorraum und $\Psi: V \rightarrow V^*$ der kanonische Isomorphismus. Dann gilt:

- 1) Für jeden Untervektorraum $U \subset V$ ist $\Psi(U^\perp) = U^0$.
- 2) Ist $\mathcal{B} = (v_1, \dots, v_n)$ eine Orthonormalbasis von V und $\mathcal{B}^* = (v_1^*, \dots, v_n^*)$ die duale Basis entsprechend 6.1.2, so ist $\Psi(v_i) = v_i^*$.

In einem euklidischen Vektorraum mit Orthonormalbasis treten also die in 6.1.2 beschriebenen Schwierigkeiten nicht auf!

Beweis. 1) Nach den Dimensionsformeln in 5.4.9 und 6.1.3 ist

$$\dim U^\perp = \dim V - \dim U = \dim U^0,$$

also genügt es, $\Psi(U^\perp) \subset U^0$ zu beweisen. Das ist aber klar, denn für $v \in U^\perp$ und $u \in U$ ist

$$\Psi(v)(u) = \langle v, u \rangle = 0.$$

2) folgt aus $\Psi(v_i)(v_j) = \langle v_i, v_j \rangle = \delta_{ij} = v_i^*(v_j)$. □

6.2.4. Wie in 5.6.1 versprochen, wird nun eine *adjungierte Abbildung* erklärt. Gegeben seien euklidische Vektorräume V und W , sowie eine lineare Abbildung

$$F: V \rightarrow W.$$

Dazu konstruieren wir eine lineare Abbildung

$$F^{\text{ad}}: W \rightarrow V \quad \text{mit} \quad \langle F(v), w \rangle = \langle v, F^{\text{ad}}(w) \rangle \quad (*)$$

für alle $v \in V$ und $w \in W$. Sind Φ und Ψ die kanonischen Isomorphismen, so bedeutet (*), daß das Diagramm

$$\begin{array}{ccc} V & \xleftarrow{F^{\text{ad}}} & W \\ \Phi \downarrow & & \downarrow \Psi \\ V^* & \xleftarrow{F^*} & W^* \end{array}$$

kommutiert, d.h. $F^{\text{ad}} = \Phi^{-1} \circ F^* \circ \Psi$. Das sieht man am einfachsten so: Für $w \in W$ ist

$$\Psi(w) = \langle \cdot, w \rangle, \quad \text{also} \quad F^*(\Psi(w)) = \langle F(\cdot), w \rangle.$$

Nach Definition von F^{ad} gilt

$$F^*(\Psi(w)) = \Phi(F^{\text{ad}}(w)) = \langle \cdot, F^{\text{ad}}(w) \rangle,$$

daraus folgt die Gleichung (*). Man erhält die adjungierte Abbildung also dadurch, daß man die duale Abbildung mit Hilfe der kanonischen Isomorphismen in die ursprünglichen Vektorräume zurückholt. Die Beschreibung durch Matrizen ist klar:

Bemerkung. Sind V und W euklidische Vektorräume mit Orthonormalbasen $\mathcal{A}$ und $\mathcal{B}$, so gilt für jede lineare Abbildung $F: V \rightarrow W$

$$M_{\mathcal{A}}^{\mathcal{B}}(F^{\text{ad}}) = {}^t(M_{\mathcal{B}}^{\mathcal{A}}(F)).$$

□

Aus Satz 6.1.5 folgt sofort der

Satz. Ist $F: V \rightarrow W$ eine lineare Abbildung zwischen euklidischen Vektorräumen, so gilt

$$\operatorname{Im} F^{\text{ad}} = (\operatorname{Ker} F)^{\perp} \quad \text{und} \quad \operatorname{Ker} F^{\text{ad}} = (\operatorname{Im} F)^{\perp}.$$

Inbesondere hat man im Fall $V = W$ orthogonale Zerlegungen

$$V = \operatorname{Ker} F \oplus \operatorname{Im} F^{\text{ad}} = \operatorname{Ker} F^{\text{ad}} \oplus \operatorname{Im} F.$$

Ist überdies F selbstadjungiert, d.h. $F = F^{\text{ad}}$, so gilt:

$$V = \operatorname{Ker} F \oplus \operatorname{Im} F. \quad \square$$

Diese letzte Zerlegung folgt schon aus der orthogonalen Diagonalisierbarkeit von F (5.6.3).

6.2.5. Was wir bisher für reelle euklidische Vektorräume beschrieben haben, kann man mit einer kleinen Modifikation auf komplexe unitäre Räume übertragen. Ist auf dem $\mathbb{C}$ -Vektorraum V eine sesquilineare Abbildung

$$s: V \times V \rightarrow \mathbb{C}$$

gegeben, so hat man dazu mit der Notation aus 6.2.1 wegen der Linearität im ersten Argument eine Abbildung

$$s'': V \rightarrow V^*, \quad v \mapsto s(\cdot, v).$$

Sie ist jedoch im Gegensatz zum bilinearen Fall nur semilinear. Im Fall eines Skalarproduktes erhält man einen kanonischen *Semi-Isomorphismus* (d.h. eine bijektive semilineare Abbildung)

$$\Psi: V \rightarrow V^*, \quad v \mapsto \langle \cdot, v \rangle.$$

Ist V ein unitärer Vektorraum und $F \in \operatorname{End}(V)$, so ist die *adjungierte Abbildung*

$$F^{\text{ad}} := \Psi^{-1} \circ F^* \circ \Psi$$

wieder $\mathbb{C}$ -linear, weil dabei zweimal konjugiert wird (durch Ψ und Ψ^{-1}), und man erhält insgesamt den folgenden

Satz. Sei F ein Endomorphismus eines unitären Vektorraumes V . Der dazu adjungierte Endomorphismus F^{ad} hat folgende Eigenschaften:

- 1) $\langle F(v), w \rangle = \langle v, F^{\text{ad}}(w) \rangle$ für alle $v, w \in V$.
- 2) $\operatorname{Im} F^{\text{ad}} = (\operatorname{Ker} F)^{\perp}$ und $\operatorname{Ker} F^{\text{ad}} = (\operatorname{Im} F)^{\perp}$.
- 3) Ist $\mathcal{B}$ eine Orthonormalbasis von V , so gilt $M_{\mathcal{B}}(F^{\text{ad}}) = {}^t \overline{M_{\mathcal{B}}(F)}$.

Beweis. 1) folgt wie in 6.2.4 aus der Definition der adjungierten Abbildung.

2) kann man aus Satz 6.1.5 folgern oder direkt nachrechnen. Wir tun das für die zweite Gleichung: $w \in (\operatorname{Im} F)^\perp$ gilt genau dann, wenn

$$0 = \langle F(v), w \rangle = \langle v, F^{\operatorname{ad}}(w) \rangle \quad \text{für alle } v \in V.$$

Da ein Skalarprodukt nicht ausgeartet ist, bedeutet das $w \in \operatorname{Ker} F^{\operatorname{ad}}$.

Zu 3) setzen wir für $B = (v_1, \dots, v_n)$

$$F(v_j) = \sum_{i=1}^n a_{ij} v_i \quad \text{und} \quad F^{\operatorname{ad}}(v_i) = \sum_{j=1}^n b_{ji} v_j.$$

Aus 1) folgt

$$a_{ij} = \langle F(v_j), v_i \rangle = \langle v_j, F^{\operatorname{ad}}(v_i) \rangle = \bar{b}_{ji}.$$

□

6.2.6. In Kapitel 5 hatten wir für folgende Arten von Endomorphismen F unitärer Vektorräume bewiesen, daß es eine Orthonormalbasis aus Eigenvektoren gibt:

1) Für unitäres F (5.5.5),

2) für selbstadjungiertes F (5.6.2).

Wir wollen nun zeigen, daß dies Spezialfälle eines allgemeineren Ergebnisses sind. Dafür benötigen wir die zunächst wenig motivierte

Definition. Ein Endomorphismus F eines unitären Vektorraumes V heißt *normal*, wenn

$$F \circ F^{\operatorname{ad}} = F^{\operatorname{ad}} \circ F.$$

Entsprechend heißt eine Matrix $A \in M(n \times n; \mathbb{C})$ *normal*, wenn $A \cdot {}^t\bar{A} = {}^t\bar{A} \cdot A$.

Beispiele. a) Jedes unitäre F ist normal, denn aus

$$\langle v, F^{-1}(w) \rangle = \langle F(v), w \rangle = \langle v, F^{\operatorname{ad}}(w) \rangle$$

folgt $F^{\operatorname{ad}} = F^{-1}$ und $F \circ F^{\operatorname{ad}} = \operatorname{id}_V = F^{\operatorname{ad}} \circ F$.

b) Jedes selbstadjungierte F ist normal, denn aus

$$\langle F(v), w \rangle = \langle v, F(w) \rangle = \langle v, F^{\operatorname{ad}}(w) \rangle$$

folgt $F = F^{\operatorname{ad}}$, also $F \circ F^{\operatorname{ad}} = F^2 = F^{\operatorname{ad}} \circ F$.

Entscheidend ist der folgende

Satz. Ist V unitär und $F \in \operatorname{End}(V)$ normal, so gilt

$$\operatorname{Ker} F^{\operatorname{ad}} = \operatorname{Ker} F \quad \text{und} \quad \operatorname{Im} F^{\operatorname{ad}} = \operatorname{Im} F.$$

Insbesondere hat man eine orthogonale Zerlegung $V = \text{Ker } F \oplus \text{Im } F$.

Beweis. Da das Skalarprodukt positiv definit ist, bedeutet $v \in \text{Ker } F$

$$\begin{aligned} 0 &= \langle F(v), F(v) \rangle = \langle v, F^{\text{ad}}(F(v)) \rangle = \langle v, F(F^{\text{ad}}(v)) \rangle \\ &= \overline{\langle F(F^{\text{ad}}(v)), v \rangle} = \overline{\langle F^{\text{ad}}(v), F^{\text{ad}}(v) \rangle} = \bar{0}. \end{aligned}$$

Das ist gleichwertig mit $v \in \text{Ker } F^{\text{ad}}$. Die zweite Gleichung folgt mit Hilfe von Satz 6.2.5. $\square$

Korollar. Ist F normal, so ist $\text{Eig}(F; \lambda) = \text{Eig}(F^{\text{ad}}; \bar{\lambda})$ für jedes $\lambda \in \mathbb{C}$.

Beweis. Für $G := F - \lambda \text{id}_V$ ist $G^{\text{ad}} = F^{\text{ad}} - \bar{\lambda} \text{id}_V$, und aus

$$\begin{aligned} G^{\text{ad}} \circ G &= F^{\text{ad}} \circ F + \lambda \bar{\lambda} \text{id}_V - \bar{\lambda} F - \lambda F^{\text{ad}} \\ &= F \circ F^{\text{ad}} + \lambda \bar{\lambda} \text{id}_V - \lambda F^{\text{ad}} - \bar{\lambda} F = G \circ G^{\text{ad}} \end{aligned}$$

folgt, daß auch G normal ist. Also ist nach obigem Satz

$$\text{Eig}(F; \lambda) = \text{Ker } G = \text{Ker } G^{\text{ad}} = \text{Eig}(F^{\text{ad}}; \bar{\lambda}). \quad \square$$

6.2.7. Nun können wir das abschließende Ergebnis beweisen:

Theorem. Für einen Endomorphismus F eines unitären $\mathbb{C}$ -Vektorraumes V sind folgende Bedingungen gleichwertig:

- i) Es gibt eine Orthonormalbasis von V bestehend aus Eigenvektoren von F .
- ii) F ist normal.

Beweis. i) $\Rightarrow$ ii): Ist $\mathcal{B} = (v_1, \dots, v_n)$ Orthonormalbasis und $F(v_i) = \lambda_i v_i$, so ist $\langle v_j, F^{\text{ad}}(v_i) \rangle = \langle F(v_j), v_i \rangle = \lambda_j \langle v_j, v_i \rangle = \delta_{ij} \lambda_j$. Also folgt $F^{\text{ad}}(v_i) = \bar{\lambda}_i v_i$ und

$$F(F^{\text{ad}}(v_i)) = F(\bar{\lambda}_i v_i) = \lambda_i \bar{\lambda}_i v_i = \bar{\lambda}_i \lambda_i v_i = F^{\text{ad}}(F(v_i)).$$

Da $F \circ F^{\text{ad}} = F^{\text{ad}} \circ F$ auf $\mathcal{B}$, folgt diese Gleichung allgemein.

ii) $\Rightarrow$ i): Wie bei den entsprechenden Beweisen in Kapitel 5 führen wir Induktion über $n = \dim V$. Das charakteristische Polynom zerfällt nach 1.3.9, also ist

$$P_F(t) = \pm(t - \lambda_1) \cdot \dots \cdot (t - \lambda_n) \quad \text{mit} \quad \lambda_1, \dots, \lambda_n \in \mathbb{C}.$$

Sei v_1 Eigenvektor zu λ_1 und $V_1 := \mathbb{C} \cdot v_1$. Wir definieren

$$W := \{w \in V : \langle w, v_1 \rangle = 0\}.$$

Dann ist $V = V_1 \oplus W$, und wegen $\dim W = n - 1$ und der Induktionsannahme genügt es, folgendes zu beweisen:

$$F(W) \subset W \quad \text{und} \quad F|_W \text{ ist normal.}$$

Die erste Behauptung folgt aus

$$\langle F(w), v_1 \rangle = \langle w, F^{\text{ad}}(v_1) \rangle = \langle w, \bar{\lambda}_1 v_1 \rangle = \lambda_1 \langle w, v_1 \rangle = 0.$$

Daß $F|W$ normal ist, folgt aus $F^{\text{ad}}(W) \subset W$, und letzteres gilt wegen

$$\langle v_1, F^{\text{ad}}(w) \rangle = \langle F(v_1), w \rangle = \langle \lambda_1 v_1, w \rangle = \lambda_1 \langle v_1, w \rangle = 0. \quad \square$$

Für Matrizen folgt das

Korollar. *Ein $A \in M(n \times n; \mathbb{C})$ ist genau dann normal, wenn es ein $S \in U(n)$ gibt, so daß SAS^{-1} eine Diagonalmatrix ist.* $\square$

6.2.8. Nach dem längeren Ausflug in die recht abstrakte Dualitätstheorie wollen wir wieder zurückkehren in die dreidimensionale Realität und mit all den zur Verfügung stehenden Werkzeugen eine Formel für den Abstand windschiefer Geraden herleiten.

Gegeben seien zwei Geraden im $\mathbb{R}^3$ in Parameterdarstellungen

$$L = v + \mathbb{R}w \quad \text{und} \quad L' = v' + \mathbb{R}w',$$

wobei $0 \neq w, w' \in \mathbb{R}^3$. Sie heißen *parallel*, wenn $\mathbb{R}w = \mathbb{R}w'$ und *windschief*, wenn sie sich nicht schneiden und nicht parallel sind. Nach Aufgabe 5 ist das gleichwertig damit, daß

$$(x := v' - v, w, w') \quad \text{linear unabhängig,}$$

also eine Basis von $\mathbb{R}^3$ sind. Wir definieren den Abstand (vgl. 5.1.1)

$$d(L, L') := \min \{d(u, u') = \|u' - u\| : u \in L, u' \in L'\}.$$

Es ist anschaulich plausibel, muß aber bewiesen werden, daß dieses Minimum tatsächlich angenommen wird. Vom Standpunkt der Analysis ist dazu die Funktion

$$\mathbb{R}^2 \rightarrow \mathbb{R}, \quad (\lambda, \lambda') \mapsto \|v' + \lambda'w' - v - \lambda w\|,$$

auf Extrema zu untersuchen. Einfacher ist eine direkte geometrische Methode.

Lemma. *Angenommen, $u \in L$ und $u' \in L'$ sind Punkte, so daß $y := u' - u$ auf w und w' senkrecht steht. Dann ist*

$$d(u, u') = d(L, L').$$

Das bedeutet, daß ein senkrechter Abstand minimal ist. Die Existenz und Eindeutigkeit von solchen *Fußpunkten* u und u' für ein gemeinsames Lot wird anschließend gezeigt.

Beweis. Da $v \in L$ und $v' \in L'$ beliebig gewählt werden können, genügt es,

$$d(u, u') \leq d(v, v') \quad (***)$$

zu zeigen. Wir halten $v' \in L'$ fest und suchen ein $\bar{v} \in L$, so daß $v' - \bar{v}$ auf L senkrecht steht. Man macht dazu den Ansatz

$$\bar{v} = v + \varrho w \quad \text{und} \quad 0 = \langle v' - \bar{v}, w \rangle = \langle x, w \rangle - \varrho \langle w, w \rangle.$$

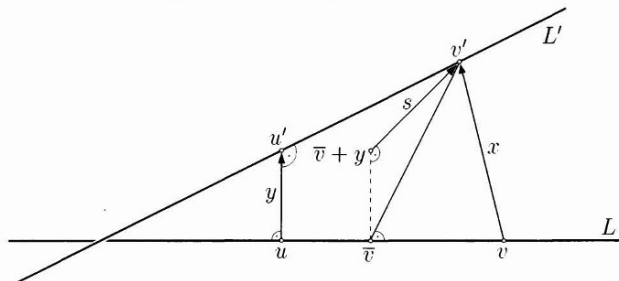


Bild 6.3

Daraus ist ϱ und somit auch $\bar{v}$ eindeutig bestimmt. Nach dem Satz von PYTHAGORAS im rechtwinkligen Dreieck mit den Ecken v, v' und $\bar{v}$ folgt

$$d(\bar{v}, v') \leq d(v, v'). \quad (*)$$

Für die nächste Abschätzung definieren wir $s := v' - \bar{v} - y$. Es gilt

$$\langle s, y \rangle = \langle v' - u' + u - \bar{v}, y \rangle = \langle v' - u', y \rangle + \langle u - \bar{v}, y \rangle = 0,$$

denn y steht auf L und L' senkrecht. Also steht s auf y senkrecht, und wir können im rechtwinkligen Dreieck mit den Ecken $\bar{v}, v'$ und $\bar{v} + y$ wieder den Satz von PYTHAGORAS anwenden, was

$$d(u, u') = \|y\| \leq d(\bar{v}, v') \quad (**)$$

ergibt. Aus (*) und (**) folgt (* * *). $\square$

Es bleibt also die Aufgabe, die Fußpunkte zu bestimmen. Wir leiten dafür zunächst notwendige Bedingungen her.

Das orthogonale Komplement von $\mathbb{R}w + \mathbb{R}w'$ wird aufgespannt vom Vektorprodukt $w \times w'$ (vgl. 5.2). Also sind $\lambda, \lambda', \beta \in \mathbb{R}$ gesucht mit $\beta \neq 0$ und

$$u = v + \lambda w, \quad u' = v' + \lambda' w', \quad y = u' - u = \beta(w \times w').$$

Das bedeutet

$$w \times w' = \alpha x + \mu w + \mu' w', \quad \text{wobei} \quad \alpha = \frac{1}{\beta}, \quad \mu = -\frac{\lambda}{\beta}, \quad \mu' = \frac{\lambda'}{\beta}. \quad (*)$$

Um α, μ, μ' aus dieser impliziten Formel zu extrahieren, verwenden wir, daß

$$w \times w', \quad w' \times x, \quad x \times w$$

wieder eine Basis von $\mathbb{R}^3$ ist (Aufgabe 3 zu 5.2). Daher ist

$$\langle \cdot, w \times w' \rangle, \quad \langle \cdot, w' \times x \rangle, \quad \langle \cdot, x \times w \rangle$$

eine Basis von $(\mathbb{R}^3)^*$. Wenden wir sie nacheinander auf Gleichung (*) an, so folgt nach den Beziehungen zwischen Vektorprodukt und Skalarprodukt (5.2.2), daß

$$\alpha = \frac{\|w \times w'\|^2}{\langle x, w \times w' \rangle}, \quad \mu = \frac{\langle w \times w', w' \times x \rangle}{\langle w, w' \times x \rangle}, \quad \mu' = \frac{\langle w \times w', x \times w \rangle}{\langle w', x \times w \rangle}$$

sein muß. Aus 5.2.2 folgt weiter, daß die drei Nenner gleich sind.

Definieren wir nun umgekehrt α , μ , μ' durch die oben berechneten Ausdrücke, und setzen wir

$$z := w \times w' - \alpha x - \mu w - \mu' w',$$

so folgt $z = 0$ aus

$$\langle z, w \times w' \rangle = \langle z, w' \times x \rangle = \langle z, x \times w \rangle = 0,$$

da diese drei Linearformen eine Basis von $(\mathbb{R}^3)^*$ bilden. Indem man aus α , μ , μ' noch β , λ , λ' ausrechnet, erhält man folgendes Ergebnis.

Satz. Für zwei windschiefe Geraden

$$L = v + \mathbb{R}w \quad \text{und} \quad L' = v' + \mathbb{R}w'$$

im $\mathbb{R}^3$ seien $u := v + \lambda w$ und $u' := v' + \lambda' w'$ mit

$$\lambda := \frac{\langle w \times w', (v' - v) \times w' \rangle}{\|w \times w'\|^2}, \quad \lambda' := \frac{\langle w \times w', (v' - v) \times w \rangle}{\|w \times w'\|^2}.$$

$$\text{Dann ist } d(L, L') = d(u, u') = \frac{|\langle w \times w', v' - v \rangle|}{\|w \times w'\|}. \quad \square$$

Für die Anwendung dieser Formel ist es vorteilhaft, die Vektoren w und w' zu normieren.

Aufgaben zu 6.2

1. Seien V, W euklidische Vektorräume, $F: V \rightarrow W$ linear und $U \subset W$ ein Untervektorraum. Dann gilt: $F^{\text{ad}}(U^\perp) = (F^{-1}(U))^\perp$.
2. Ist V ein euklidischer Vektorraum und $F: V \rightarrow V$ selbstadjungiert, so gilt $F(U^\perp) = (F^{-1}(U))^\perp$ für alle Untervektorräume $U \subset V$. Gilt die Umkehrung auch?
3. Zeigen Sie, dass für einen unitären Vektorraum V durch $\text{End}(V) \rightarrow \text{End}(V)$, $F \mapsto F^{\text{ad}}$ ein Semi-Isomorphismus gegeben ist.
4. Sei $A \in M(n \times n; \mathbb{C})$ antihermitesch, das heißt $-A = {}^t \bar{A}$. Zeigen Sie, dass A normal ist und alle Eigenwerte von A in $i\mathbb{R}$ liegen.

5. Seien $L = v + \mathbb{R}w$ und $L' = v' + \mathbb{R}w'$ zwei Geraden im $\mathbb{R}^n$ und $x := v' - v$. Zeigen Sie:

L und L' sind windschief $\Leftrightarrow x, w$ und w' sind linear unabhängig.

6. Gegeben seien zwei windschiefe Geraden $L = v + \mathbb{R}w$ und $L' = v' + \mathbb{R}w'$ im $\mathbb{R}^n$. Wir wollen zwei Methoden angeben, um den Abstand

$$d(L, L') = \min\{d(u, u') = \|u' - u\| : u \in L, u' \in L'\}$$

zu berechnen. Zur Vereinfachung nehmen wir $\|w\| = \|w'\| = 1$ an und definieren

$$\delta: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (\lambda, \lambda') \mapsto \|v' + \lambda'w' - v - \lambda w\|^2.$$

a) Untersuchen Sie die Funktion δ mit Hilfe der Differentialrechnung auf Extrema und bestimmen damit den Abstand $d(L, L')$.

b) Es gilt $\delta(\lambda, \lambda') = \lambda^2 + a\lambda\lambda' + \lambda'^2 + b\lambda + c\lambda' + d$. Setzen Sie $\mu := \lambda + \frac{a}{2}\lambda'$ und $\mu' = \frac{\sqrt{4-a^2}}{2}\lambda'$ und zeigen Sie, dass man auf diese Weise δ durch quadratische Ergänzung schreiben kann als $\delta(\lambda, \lambda') = (\mu - e)^2 + (\mu' - f)^2 + g$. Dann ist $g = d(L, L')$.

6.3 Tensorprodukte

„... what is the use of a book
without pictures or conversa-
tion?“

aus *Alice in Wonderland*

Der Begriff eines „Tensors“ wird in der theoretischen Physik seit langer Zeit benutzt; es hat etwas gedauert, bis es den Mathematikern gelungen ist, dafür einen befriedigenden formalen Rahmen zu zimmern. Er ist sehr präzise, aber höchst abstrakt und daher nicht allgemein beliebt.

Vom naiven Standpunkt startet man mit „skalaren“ Größen a : das sind die Elemente des Grundkörpers. Folgen davon, also $(a_1, a_2, \dots)$ sind „Vektoren“, sie haben einen einzigen Index. Folgen von Vektoren sind gegeben durch a_{ij} , dazu benötigt man zwei Indizes, und diese Objekte heißen „Tensoren“. Das wirft die Frage auf, warum man statt Tensor nicht einfach Matrix sagt. Dabei muß man bedenken, daß Matrizen Hilfsmittel sind, mit denen man verschiedenartige abstrakte Vorgänge beschreiben kann. Etwa eine quadratische Matrix kann sowohl einen Endomorphismus (2.4.4) als auch eine Bilinearform (5.4.2) beschreiben. Daher ist etwas Vorbereitung nötig, bis man sagen kann, was die charakteristischen Eigenschaften eines Tensors sind.

6.3.1. Der Begriff der Bilinearform aus 5.4.1 und 6.2.1 muß noch einmal verallgemeinert werden. Sind V , W und U Vektorräume über demselben Körper K , so heißt eine Abbildung

$$\xi: V \times W \rightarrow U$$

bilinear, wenn für jeweils festes $v \in V$ und $w \in W$ die Abbildungen

$$\xi_v: W \rightarrow U, \quad w \mapsto \xi(v, w), \quad \text{und}$$

$$\xi_w: V \rightarrow U, \quad v \mapsto \xi(v, w),$$

linear sind. Es ist leicht zu sehen, daß die Menge

$$\text{Bil}(V, W; U) := \{\xi: V \times W \rightarrow U: \xi \text{ ist bilinear}\}$$

wieder ein K -Vektorraum ist.

Beispiele. a) Seien $V = W = K[t]_d$ bzw. $U = K[t]_{2d}$ die Vektorräume der Polynome vom Grad $\leq d$ bzw. $\leq 2d$ mit Basen

$$1, t, \dots, t^d \quad \text{bzw.} \quad 1, t, \dots, t^{2d} \quad \text{und}$$

$$\xi: K[t]_d \times K[t]_d \rightarrow K[t]_{2d}, \quad (P, Q) \mapsto P \cdot Q,$$

die Multiplikationsabbildung, wobei das Produkt $P \cdot Q$ nach der in 1.3.6 angegebenen Formel erklärt ist. Daran sieht man sofort, daß ξ bilinear ist. Insbesondere ist $\xi(t^i, t^j) = t^{i+j}$, also liegt die Basis von $K[t]_{2d}$ im Bild $\text{Im } \xi \subset K[t]_{2d}$. Dies

ist jedoch im allgemeinen kein Untervektorraum. Sei etwa $d = 1$:

für $K = \mathbb{Q}$ ist $t^2 - 2 \notin \text{Im } \xi$, obwohl $t^2, -2 \in \text{Im } \xi$,

für $K = \mathbb{R}$ ist $t^2 + 1 \notin \text{Im } \xi$, obwohl $t^2, 1 \in \text{Im } \xi$,

denn diese quadratischen Polynome sind nicht Produkt von linearen Polynomen, weil $\sqrt{2} \notin \mathbb{Q}$ und $i \notin \mathbb{R}$.

b) Seien nun $V = W = K[t]$ bzw. $U = K[t_1, t_2]$ die unendlich dimensionalen Vektorräume der Polynome in einer Veränderlichen t bzw. in zwei Veränderlichen t_1, t_2 mit Basen

$$(t^i)_{i \in \mathbb{N}} \text{ von } K[t] \quad \text{bzw.} \quad (t_1^i \cdot t_2^j)_{(i,j) \in \mathbb{N} \times \mathbb{N}} \text{ von } K[t_1, t_2].$$

Auch hier hat man eine Multiplikation

$$\xi: K[t] \times K[t] \rightarrow K[t_1, t_2], \quad (P(t), Q(t)) \mapsto P(t_1) \cdot Q(t_2).$$

Das bedeutet, man ersetzt vor der Multiplikation die Variable t in P durch t_1 und in Q durch t_2 . Mit etwas Rechnung (vgl. Aufgabe 9 zu 1.3) sieht man, daß ξ wieder bilinear ist. Auch hier liegt die Basis von U im Bild von ξ , denn es ist

$$\xi(t^i, t^j) = t_1^i \cdot t_2^j,$$

aber $\text{Im } \xi \subset U$ ist kein Untervektorraum. Man überlege sich, daß etwa das Polynom $t_1 t_2 + 1$ für keinen Körper K in $\text{Im } \xi$ liegt.

6.3.2. In 2.4.1 hatten wir gesehen, daß eine lineare Abbildung durch die Bilder einer Basis eindeutig bestimmt ist. Für bilineare Abbildungen kann man das geeignet modifizieren. Man beachte dazu folgendes: Sind $(v_i)_{i \in I}$ und $(w_j)_{j \in J}$ Basen von V und W , so hat man eine durch die disjunkte Vereinigung von I und J indizierte Basis

$$((v_i, 0))_{i \in I}, \quad ((0, w_j))_{j \in J}$$

von $V \times W$ (vgl. Aufgabe 5 zu 1.5). Dagegen ist die durch $I \times J$ indizierte Familie

$$(v_i, w_j)$$

im allgemeinen weder ein Erzeugendensystem noch linear unabhängig in $V \times W$.

Grundlegend ist die folgende einfache

Bemerkung. Seien V bzw. W Vektorräume über K mit Basen $(v_i)_{i \in I}$ bzw. $(w_j)_{j \in J}$. Ist U ein weiterer K -Vektorraum, so gibt es zu einer beliebig vorgegebenen Familie $(u_{ij})_{(i,j) \in I \times J}$ in U genau eine bilineare Abbildung

$$\xi: V \times W \rightarrow U \quad \text{mit} \quad \xi(v_i, w_j) = u_{ij} \quad \text{für alle } (i, j) \in I \times J.$$

Vorsicht! Die Vektoren $(v_i, 0)$ und $(0, w_j)$ der Basis von $V \times W$ werden durch jede bilineare Abbildung auf Null abgebildet.

Beweis. Um ohne großen zusätzlichen Aufwand an Indizes auch unendliche Basen mit behandeln zu können, vereinbaren wir zunächst eine vereinfachte Schreibweise. Jeder Vektor $v \in V$ ist endliche Linearkombination der v_i . Das bedeutet, es gibt ein $m \in \mathbb{N}$, $i_1, \dots, i_m \in I$ und $\lambda_1, \dots, \lambda_m \in K$, so daß

$$v = \sum_{k=1}^m \lambda_k v_{i_k}.$$

Dafür schreiben wir einfacher

$$v = \sum_i' \lambda_i v_i,$$

wobei $\sum'$ andeuten soll, daß nur endlich viele der formal aufgeschriebenen Summanden $\neq 0$ sind. Das kann man etwa dadurch erreichen, daß man bis auf endlich viele Ausnahmen alle $\lambda_i = 0$ setzt. Ist I endlich, so ist die eingeschränkte Summe $\sum'$ gleich der vollen Summe $\sum$. Wer nur an diesem Fall interessiert ist, kann den Unterschied ignorieren.

Zur Definition von ξ betrachten wir ein beliebiges Paar $(v, w) \in V \times W$. Es sei $v = \sum' \lambda_i v_i$ und $w = \sum' \mu_j w_j$. Angenommen, es gibt ein bilineares ξ wie gesucht. Dann muß

$$\xi(v, w) = \xi\left(\sum_i' \lambda_i v_i, \sum_j' \mu_j w_j\right) = \sum_{i,j}' \lambda_i \mu_j \xi(v_i, w_j) = \sum_{i,j}' \lambda_i \mu_j u_{ij}$$

sein. Da die λ_i und μ_j durch (v, w) eindeutig bestimmt sind, gibt es höchstens eine solche Abbildung.

Umgekehrt ist zu zeigen, daß durch

$$\xi(v, w) := \sum_{i,j}' \lambda_i \mu_j u_{ij}$$

tatsächlich eine bilineare Abbildung erklärt ist. Dazu halten wir w fest. Dann ist

$$\xi_w: V \rightarrow U, \quad v = \sum_i' \lambda_i v_i \mapsto \sum_{i,j}' \lambda_i \mu_j u_{ij},$$

offensichtlich linear, mit

$$\xi_w(v_i) = \sum_j' \mu_j u_{ij}.$$

Analog können wir v festhalten und w variabel lassen. □

6.3.3. Der Vektorraum U und die Vektoren u_{ij} waren in obiger Bemerkung ganz beliebig gewählt gewesen. Nun wollen wir zeigen, daß unter all den möglichen Wahlen eine besonders ausgezeichnet ist, nämlich ein Vektorraum U , in dem die u_{ij} eine Basis sind.

Theorem. Seien V und W Vektorräume über K . Dann gibt es einen K -Vektorraum $V \otimes_K W$ zusammen mit einer bilinearen Abbildung

$$\eta: V \times W \rightarrow V \otimes_K W,$$

die folgende universelle Eigenschaft haben: Zu jedem K -Vektorraum U zusammen mit einer bilinearen Abbildung

$$\xi: V \times W \rightarrow U$$

gibt es genau eine lineare Abbildung

$$\xi_{\otimes}: V \otimes_K W \rightarrow U$$

mit $\xi = \xi_{\otimes} \circ \eta$. Das kann man durch ein kommutatives Diagramm illustrieren:

$$\begin{array}{ccc} V \times W & & \\ \eta \downarrow & \searrow \xi & \\ V \otimes_K W & \xrightarrow{\xi_{\otimes}} & U \end{array}$$

Weiter gilt: Falls $\dim V, \dim W < \infty$, so ist

$$\dim(V \otimes_K W) = (\dim V) \cdot (\dim W).$$

Falls klar ist, welches K Grundkörper ist, schreibt man nur $\otimes$ statt $\otimes_K$.

Man nennt $V \otimes_K W$ das *Tensorprodukt* von V und W über K . Die Elemente von $V \otimes_K W$ heißen *Tensoren*.

Beweis. Wir wählen Basen $(v_i)_{i \in I}$ von V und $(w_j)_{j \in J}$ von W (vgl. 1.5.3) und definieren $V \otimes W$ als den Vektorraum der endlichen Linearkombinationen von formalen Ausdrücken der Form $v_i \otimes w_j$. Präziser ausgedrückt betrachten wir den K -Vektorraum

$$\text{Abb}(I \times J, K) = \{\tau: I \times J \rightarrow K\}$$

(vgl. 1.4.1, Beispiel e) und seinen Untervektorraum

$$V \otimes W := \{\tau: I \times J \rightarrow K: \tau(i, j) \neq 0 \text{ für nur endlich viele } (i, j) \in I \times J\}.$$

Dann ist $v_i \otimes w_j \in V \otimes W$ die Abbildung, deren Wert an der einzigen Stelle (i, j) gleich 1 und sonst 0 ist. Offenbar ist $(v_i \otimes w_j)_{(i, j) \in I \times J}$ eine Basis von $V \otimes W$,

denn für $\tau \in V \otimes W$ gilt

$$\tau = \sum_{i,j} \tau(i, j)(v_i \otimes w_j),$$

wobei nur endlich viele Summanden $\neq 0$ sind. Also haben wir ein Erzeugendensystem. Ist

$$\tau := \sum_{i,j} \alpha_{ij}(v_i \otimes w_j) = 0,$$

so gilt $\tau(i, j) = \alpha_{ij} = 0$, weil die Nullabbildung überall den Wert Null hat. Zur Definition von η benutzen wir die Bemerkung 6.3.2: Es genügt,

$$\eta(v_i, w_j) := v_i \otimes w_j$$

zu setzen. Sind beliebige Vektoren

$$v = \sum_i' \lambda_i v_i \in V \quad \text{und} \quad w = \sum_j' \mu_j w_j \in W$$

gegeben, so ist wegen der Bilinearität von η

$$v \otimes w := \eta(v, w) = \eta\left(\sum_i' \lambda_i v_i, \sum_j' \mu_j w_j\right) = \sum_{i,j}' \lambda_i \mu_j (v_i \otimes w_j).$$

Nun zur universellen Eigenschaft: Ist $\xi: V \times W \rightarrow U$ gegeben, so betrachten wir die Vektoren $u_{ij} := \xi(v_i, w_j) \in U$. Wegen der Bedingung $\xi = \xi_{\otimes} \circ \eta$ muß

$$\xi_{\otimes}(v_i \otimes w_j) = u_{ij}$$

sein. Nach 2.4.1 gibt es genau eine lineare Abbildung $\xi_{\otimes}: V \otimes W \rightarrow U$ mit dieser Eigenschaft. Weiter ist nach 2.4.1

$$\xi_{\otimes}\left(\sum_{i,j}' \alpha_{ij}(v_i \otimes w_j)\right) = \sum_{i,j}' \alpha_{ij} u_{ij}.$$

Also ist $\xi_{\otimes}(v \otimes w) = \xi(v, w)$ für alle $(v, w) \in V \times W$, und die universelle Eigenschaft ist bewiesen.

Der Zusatz über die Dimensionen ist klar, denn besteht I aus m und J aus n Elementen, so besteht $I \times J$ aus $m \cdot n$ Elementen. $\square$

Aus der Bilinearität von η folgen sofort grundlegende

Rechenregeln für Tensoren. Ist $\eta: V \times W \rightarrow V \otimes W$ und $v \otimes w := \eta(v, w)$, so gilt für $v, v' \in V$, $w, w' \in W$ und $\lambda \in K$:

$$a) v \otimes w + v' \otimes w = (v + v') \otimes w, \quad v \otimes w + v \otimes w' = v \otimes (w + w'),$$

$$b) (\lambda \cdot v) \otimes w = v \otimes (\lambda \cdot w) = \lambda \cdot (v \otimes w).$$

$\square$

Man beachte, daß im Gegensatz zu a) eine beliebige Summe $v \otimes w + v' \otimes w'$ im allgemeinen nicht zu einem Produkt zusammengezogen werden kann. Die Regel

$$\lambda v \otimes_K w = v \otimes_K \lambda w$$

für $\lambda \in K$ kann man so lesen, daß der Index K unter dem Tensorzeichen $\otimes$ angibt, welche Faktoren zwischen den beiden Seiten ausgetauscht werden dürfen.

6.3.4. Wie schon gesagt, das Tensorprodukt ist ein sehr abstrakter Begriff, der bei der ersten Lektüre schwindeleerregend sein kann. Dem soll durch einige konkrete Beispiele entgegengewirkt werden. Der geometrische Hintergrund des eigenartigen Wechsels von bilinear zu linear wird in Aufgabe 3 beleuchtet.

Beispiele. a) Wie in 6.3.1 betrachten wir die Multiplikation von Polynomen

$$\xi: K[t] \times K[t] \rightarrow K[t_1, t_2].$$

Nach der universellen Eigenschaft des Tensorproduktes gibt es eine lineare Abbildung

$$\xi_{\otimes}: K[t] \otimes_K K[t] \rightarrow K[t_1, t_2], \quad t^i \otimes t^j \mapsto t_1^i \cdot t_2^j.$$

Da die *Monome* $t_1^i t_2^j$ eine Basis des K -Vektorraumes $K[t_1, t_2]$ bilden, ist $\xi_{\otimes}$ ein Isomorphismus. Man kann also die Identifikationen

$$K[t] \otimes_K K[t] = K[t_1, t_2], \quad t \otimes 1 = t_1 \text{ und } 1 \otimes t = t_2$$

vornehmen, was eine neue Methode ergibt, den Polynomring von zwei Veränderlichen aus dem einer Veränderlichen zu konstruieren. Indem man die Grade passend abschneidet, kann man daraus auch endlichdimensionale Beispiele erhalten.

b) Sei W ein beliebiger $\mathbb{R}$ -Vektorraum und $V = \mathbb{C}$ als $\mathbb{R}$ -Vektorraum. Wir betrachten die universelle bilineare Abbildung

$$\mathbb{C} \times W \rightarrow \mathbb{C} \otimes_{\mathbb{R}} W, \quad (\lambda, w) \mapsto \lambda \otimes w.$$

Ist $(w_j)_{j \in J}$ eine Basis von W , so erhalten wir zusammen mit der Basis 1, i von $\mathbb{C}$ die Basis

$$(1 \otimes w_j)_{j \in J} \cup (i \otimes w_j)_{j \in J} \quad \text{von } \mathbb{C} \otimes_{\mathbb{R}} W.$$

Also hat jedes $w^* \in \mathbb{C} \otimes_{\mathbb{R}} W$ eine eindeutige Darstellung

$$w^* = \sum_j' \alpha_j (1 \otimes w_j) + \sum_j' \beta_j (i \otimes w_j) = \sum_j' (\alpha_j + \beta_j i) \otimes w_j = \sum_j' \lambda_j \otimes w_j$$

mit $\alpha_j, \beta_j \in \mathbb{R}$ und $\lambda_j \in \mathbb{C}$. In dem $\mathbb{R}$ -Vektorraum $\mathbb{C} \otimes_{\mathbb{R}} W$ können wir eine Multiplikation mit komplexen Zahlen definieren durch

$$\mathbb{C} \times (\mathbb{C} \otimes_{\mathbb{R}} W) \rightarrow \mathbb{C} \otimes_{\mathbb{R}} W, \quad \left(\lambda, \sum_j' \lambda_j \otimes w_j \right) \mapsto \sum_j' \lambda \cdot \lambda_j \otimes w_j.$$

Ohne jede Schwierigkeit prüft man die Axiome aus 1.4.1 nach: $\mathbb{C} \otimes_{\mathbb{R}} W$ wird so zu einem komplexen Vektorraum mit Basis $(1 \otimes w_j)_{j \in J}$ über $\mathbb{C}$. Man nennt $\mathbb{C} \otimes_{\mathbb{R}} W$ die *Komplexifizierung* von W .

Die Einschränkung von $\mathbb{C} \times W \rightarrow \mathbb{C} \otimes_{\mathbb{R}} W$ auf $1 \times W$ ergibt eine $\mathbb{R}$ -lineare Abbildung

$$W \rightarrow \mathbb{C} \otimes_{\mathbb{R}} W, \quad w \mapsto 1 \otimes w.$$

Sie ist injektiv, also ist ihr Bild $1 \otimes W$ isomorph zu W , d.h. W kann als reeller Untervektorraum von $\mathbb{C} \otimes_{\mathbb{R}} W$ aufgefaßt werden.

Im Spezialfall $W = \mathbb{R}^n$ kann man das viel einfacher beschreiben: $\mathbb{R}^n$ hat die Basis

$$e_j = (0, \dots, 0, 1, 0, \dots, 0) \in \mathbb{R}^n, \quad j = 1, \dots, n.$$

Da die 1 an der Stelle j auch eine komplexe Zahl ist, kann man e_j ebenso als Element des $\mathbb{C}^n$ ansehen. Präziser kann man dafür $1 \otimes e_j \in \mathbb{C}^n$ schreiben. Dann ist die Abbildung

$$\mathbb{C} \otimes_{\mathbb{R}} \mathbb{R}^n \rightarrow \mathbb{C}^n, \quad \sum_{j=1}^n \lambda_j \otimes e_j \mapsto (\lambda_1, \dots, \lambda_n),$$

ein Isomorphismus von $\mathbb{C}$ -Vektorräumen, und $\mathbb{R}^n$ ist in kanonischer Weise ein reeller Untervektorraum des $\mathbb{C}^n$. Das hatten wir z.B. in 5.5.6 benutzt.

6.3.5. Die Beziehung des Tensorproduktes zu einem vertrauten Produkt sieht man schön am Polynomring (Beispiel a) in 6.3.3). Weitere Beispiele für solche Beziehungen erhält man mit Hilfe der Dualräume.

Sind V und W Vektorräume mit Dualräumen V^* und W^* , so betrachten wir folgende Konstruktionen:

a) Für zwei Linearformen $\varphi \in V^*$ und $\psi \in W^*$ ist das Produkt

$$\varphi \cdot \psi: V \times W \rightarrow K, \quad (v, w) \mapsto \varphi(v) \cdot \psi(w),$$

eine Bilinearform, die zugehörige Abbildung

$$(\varphi \cdot \psi)_{\otimes}: V \otimes W \rightarrow K, \quad v \otimes w \mapsto \varphi(v) \cdot \psi(w),$$

ist linear, also ist $(\varphi \cdot \psi)_{\otimes} \in (V \otimes W)^*$. Die so entstandene Abbildung

$$V^* \times W^* \rightarrow (V \otimes W)^*, \quad (\varphi, \psi) \mapsto (\varphi \cdot \psi)_{\otimes},$$

ist wiederum bilinear, d.h. sie wird zu einer linearen Abbildung

$$V^* \otimes W^* \rightarrow (V \otimes W)^*, \quad \varphi \otimes \psi \mapsto (\varphi \cdot \psi)_{\otimes}.$$

Daß $\varphi \otimes \psi$ eine Linearform auf $V \otimes W$ erklärt, kann man auch so sehen:

$$(\varphi \otimes \psi)(v \otimes w) := \varphi(v) \cdot \psi(w), \quad (\alpha)$$

wobei rechts das gewöhnliche Produkt in K steht.

b) Die durch Multiplikation mit Skalaren in W erhaltene Abbildung

$$V^* \times W \rightarrow \text{Hom}(V, W), \quad (\varphi, w) \mapsto \varphi(\cdot) \cdot w, \quad \text{wobei} \\ \varphi(\cdot) \cdot w: V \rightarrow W, \quad v \mapsto \varphi(v) \cdot w,$$

ist bilinear, dazu gehört die lineare Abbildung

$$V^* \otimes W \rightarrow \text{Hom}(V, W), \quad \varphi \otimes w \mapsto \varphi(\cdot) \cdot w.$$

Das kann man wieder kürzer ausdrücken durch

$$(\varphi \otimes w)(v) := \varphi(v) \cdot w. \quad (\beta)$$

Diese Spielereien gipfeln in dem

Satz. Sind V und W endlich-dimensionale Vektorräume, so sind die beiden oben erklärten kanonischen Abbildungen

$$\alpha: V^* \otimes W^* \rightarrow (V \otimes W)^* \quad \text{und} \quad \beta: V^* \otimes W \rightarrow \text{Hom}(V, W)$$

Isomorphismen.

Beweis. Wir benutzen Basen $(v_1, \dots, v_m)$ von V und $(w_1, \dots, w_n)$ von W , sowie die zugehörigen dualen Basen $(v_1^*, \dots, v_m^*)$ von V^* und $(w_1^*, \dots, w_n^*)$ von W^* (vgl. 6.1.2). Daraus erhält man Basen

$$(v_i^* \otimes w_j^*)_{i,j} \text{ von } V^* \otimes W^* \quad \text{und} \quad ((v_i \otimes w_j)^*)_{i,j} \text{ von } (V \otimes W)^*,$$

und es genügt $\alpha(v_i^* \otimes w_j^*) = (v_i \otimes w_j)^*$ zu zeigen. Das folgt aber aus

$$\alpha(v_i^* \otimes w_j^*)(v_k \otimes w_l) = v_i^*(v_k)w_j^*(w_l) = \delta_{ik}\delta_{jl} = (v_i \otimes w_j)^*(v_k \otimes w_l).$$

Um zu zeigen, daß β ein Isomorphismus ist, benutzen wir die Basen

$$(v_i^* \otimes w_j)_{i,j} \text{ von } V^* \otimes W \quad \text{und} \quad (F_i^j)_{i,j} \text{ von } \text{Hom}(V, W)$$

mit $F_i^j(v_k) = \delta_{ik}w_j$ wie in 2.4.2. Aus

$$\beta(v_i^* \otimes w_j)(v_k) = v_i^*(v_k)w_j = \delta_{ik}w_j$$

folgt, daß β diese Basen aufeinander abbildet. □

6.3.6. Nun kommen wir zurück auf die eingangs gestellte Frage nach den Beziehungen von Tensoren zu Matrizen. Es werde A als lineare Abbildung aufgefaßt; dann ist

$$A \in \text{Hom}(K^m, K^n) \cong (K^m)^* \otimes K^n.$$

Wirkt A als Bilinearform, so ist

$$A \in \text{Bil}(K^m, K^n; K) \cong \text{Hom}(K^m \otimes K^n, K) = (K^m \otimes K^n)^* \cong (K^m)^* \otimes (K^n)^*,$$

wobei all die Isomorphismen kanonisch sind (vgl. Aufgabe 2a). Allgemein ist

- eine lineare Abbildung von V nach W ein Tensor aus $V^* \otimes W$,
- eine Bilinearform auf $V \times W$ ein Tensor aus $V^* \otimes W^*$.

6.3.7. In einem Tensorprodukt $V \otimes V$ sind zwei Tensoren $v \otimes v'$ und $v' \otimes v$ im allgemeinen verschieden (vgl. etwa die Beispiele in 6.3.1). Man kann das Tensorprodukt so modifizieren, daß der Unterschied zwischen $v \otimes v'$ und $v' \otimes v$ genau kontrollierbar wird. Dazu die

Definition. Sind V und W Vektorräume über K , so heißt eine bilineare Abbildung

$$\xi: V \times V \rightarrow W$$

symmetrisch, wenn $\xi(v, v') = \xi(v', v)$ für alle $v, v' \in V$ und

alternierend, wenn $\xi(v, v) = 0$ für alle $v \in V$.

Bemerkung. Ist ξ alternierend, so gilt

$$\xi(v', v) = -\xi(v, v') \quad \text{für alle } v, v' \in V,$$

und im Fall $\text{char}(K) \neq 2$ gilt auch die Umkehrung.

Beweis. $\xi(v + v', v + v') = \xi(v, v) + \xi(v', v') + \xi(v, v') + \xi(v', v)$. □

Ob eine bilineare Abbildung ξ symmetrisch oder alternierend ist, kann man mit Hilfe der linearen Abbildung $\xi_{\otimes}$ entscheiden. Dazu betrachten wir die Untervektorräume

$$S(V) := \text{span}(v \otimes v' - v' \otimes v)_{v, v' \in V} \subset V \otimes V$$

und

$$A(V) := \text{span}(v \otimes v)_{v \in V} \subset V \otimes V.$$

Lemma. Für jedes bilineare $\xi: V \times V \rightarrow W$ gilt

$$\xi \text{ symmetrisch} \Leftrightarrow S(V) \subset \text{Ker } \xi_{\otimes},$$

$$\xi \text{ alternierend} \Leftrightarrow A(V) \subset \text{Ker } \xi_{\otimes}.$$

Beweis. Das folgt sofort aus

$$\xi(v, v') - \xi(v', v) = \xi_{\otimes}(v \otimes v') - \xi_{\otimes}(v' \otimes v) = \xi_{\otimes}(v \otimes v' - v' \otimes v),$$

$$\xi(v, v) = \xi_{\otimes}(v \otimes v).$$

□

6.3.8. Analog zum Tensorprodukt beweisen wir nun die Existenz eines äußeren Produktes.

Theorem. Für jeden K -Vektorraum V mit $\dim V \geq 2$ gibt es einen K -Vektorraum $V \wedge V$ zusammen mit einer alternierenden Abbildung

$$\wedge: V \times V \rightarrow V \wedge V,$$

die folgende universelle Eigenschaft haben: zu jedem K -Vektorraum W zusammen mit einer alternierenden Abbildung $\xi: V \times V \rightarrow W$ gibt es genau eine lineare Abbildung $\xi_{\wedge}$ derart, daß das Diagramm

$$\begin{array}{ccc} V \times V & & \\ \wedge \downarrow & \searrow \xi & \\ V \wedge V & \xrightarrow{\xi_{\wedge}} & W \end{array}$$

kommutiert. Ist $(v_1, \dots, v_n)$ eine Basis von V , so ist durch $v_i \wedge v_j := \wedge(v_i, v_j)$ mit $1 \leq i < j \leq n$ eine Basis von $V \wedge V$ gegeben. Insbesondere ist

$$\dim(V \wedge V) = \binom{n}{2} = \frac{n(n-1)}{2}.$$

Beweis. Wir erklären das äußere Produkt als Quotientenvektorraum des Tensorproduktes:

$$V \wedge V := (V \otimes V) / A(V),$$

wobei $A(V)$ der in 6.3.7 erklärte Untervektorraum ist. Bezeichnet

$$\varrho: V \otimes V \rightarrow V \wedge V$$

die kanonische Abbildung, so erklären wir $\wedge := \varrho \circ \eta$. Für $v, v' \in V$ ist also

$$v \wedge v' := \wedge(v, v') = \varrho(\eta(v, v')) = \varrho(v \otimes v').$$

Die Abbildung $\wedge$ ist bilinear und nach dem Lemma aus 6.3.7 auch alternierend.

Zum Beweis der universellen Eigenschaft betrachten wir das folgende Diagramm:

$$\begin{array}{ccccc} V \times V & & & & \\ & \searrow \eta & \searrow \xi & & \\ & & V \otimes V & \xrightarrow{\xi_{\otimes}} & W \\ & \downarrow \wedge & \swarrow \varrho & & \\ & & V \wedge V & \xrightarrow{\xi_{\wedge}} & W \end{array} \quad (*)$$

Zu ξ gibt es nach 6.3.3 ein eindeutiges lineares $\xi_{\otimes}$ und nach der universellen

Eigenschaft des Quotienten (2.2.7) wegen Lemma 6.3.7 ein eindeutiges lineares $\xi_\wedge$. Aus der Kommutativität des Diagramms (*) folgt

$$\xi_\wedge(v \wedge v') = \xi(v, v').$$

Es bleibt die Behauptung über die Basis von $V \wedge V$ zu zeigen, das ist die einzige Schwierigkeit. Da $V \otimes V$ von Tensoren der Form $v_i \otimes v_j$ erzeugt wird, erzeugen die Produkte $v_i \wedge v_j$ den Raum $V \wedge V$. Wegen $v_i \wedge v_i = 0$ und $v_j \wedge v_i = -v_i \wedge v_j$ sind schon die $\binom{n}{2}$ Produkte

$$v_i \wedge v_j \quad \text{mit } i < j$$

ein Erzeugendensystem, und es genügt, die lineare Unabhängigkeit zu beweisen.

Dazu betrachten wir den Vektorraum $W = K^N$ mit $N = \binom{n}{2}$, und wir bezeichnen die kanonische Basis von K^N mit

$$(e_{ij})_{1 \leq i < j \leq n}.$$

Eine alternierende Abbildung $\xi: V \times V \rightarrow K^N$ konstruieren wir wie folgt: sind

$$v = \sum \lambda_i v_i \quad \text{und} \quad v' = \sum \mu_i v_i$$

in V gegeben, so betrachten wir die Matrix

$$A = \begin{pmatrix} \lambda_1 & \cdots & \lambda_n \\ \mu_1 & \cdots & \mu_n \end{pmatrix}$$

und bezeichnen mit $a_{ij} := \lambda_i \mu_j - \lambda_j \mu_i$ den entsprechenden 2-Minor von A . Dann ist durch

$$\xi(v, v') := \sum_{i < j} a_{ij} e_{ij}$$

eine sehr gute alternierende Abbildung gegeben: wegen der universellen Eigenschaft muß

$$\xi_\wedge(v_i \wedge v_j) = \xi(v_i, v_j) = e_{ij}$$

sein, und aus der linearen Unabhängigkeit der e_{ij} in K^N folgt die Behauptung. Die so erhaltene Abbildung

$$\xi_\wedge: V \wedge V \rightarrow K^N$$

ist ein Isomorphismus. Er liefert eine etwas konkretere Beschreibung des äußeren Produktes. □

Die Tatsache, daß die Abbildung $\wedge$ alternierend ist, kann man auch ausdrücken in den folgenden

Rechenregeln für das äußere Produkt. Für $v, v', w, w' \in V$ und $\lambda \in K$ gilt:

$$a) (v + v') \wedge w = v \wedge w + v' \wedge w, \quad v \wedge (w + w') = v \wedge w + v \wedge w',$$

$$b) (\lambda v) \wedge w = v \wedge (\lambda w) = \lambda(v \wedge w),$$

$$c) v \wedge v = 0 \text{ und } v' \wedge v = -v \wedge v'.$$

□

Vorsicht! Ein Produkt $V \wedge W$ mit $V \neq W$ macht keinen Sinn! Daher nennt man das äußere Produkt manchmal auch *äußere Potenz*.

6.3.9. Da äußere Produkte in verschiedenartigen Verkleidungen auftreten können, wollen wir den Leser mit einigen kanonischen Isomorphismen wappnen. Zur Abkürzung verwenden wir die folgenden Bezeichnungen:

$$\text{Bil}(V; W) := \{\xi: V \times V \rightarrow W: \xi \text{ bilinear}\} \quad \text{und}$$

$$\text{Alt}^2(V; W) := \{\xi \in \text{Bil}(V, V; W): \xi \text{ alternierend}\}.$$

Die universelle Eigenschaft des Tensorproduktes ergibt den kanonischen Isomorphismus

$$\text{Bil}(V; K) \rightarrow (V \otimes V)^*, \quad \xi \mapsto \xi_{\otimes},$$

(Aufgabe 2). Aus der Quotientenabbildung

$$V \otimes V \rightarrow V \wedge V = (V \otimes V)/A(V)$$

erhalten wir den Isomorphismus

$$\pi: \{\psi \in (V \otimes V)^*: \psi|A(V) = 0\} \rightarrow (V \wedge V)^*, \quad \psi \mapsto \bar{\psi},$$

(Aufgabe 9 zu 2.2). Nach Lemma 6.3.7 ist die Abbildung

$$\text{Alt}^2(V; K) \rightarrow (V \wedge V)^*, \quad \xi \mapsto \bar{\xi}_{\otimes},$$

wohldefiniert und ein Isomorphismus. Wenn wir $(V \wedge V)^*$ mit Hilfe von π als Untervektorraum von $(V \otimes V)^*$ ansehen, ergibt sich das kommutative Diagramm von Inklusionen und Isomorphismen

$$\begin{array}{ccc} \text{Alt}^2(V; K) & \subset & \text{Bil}(V; K) \\ \downarrow & & \downarrow \\ (V \wedge V)^* & \subset & (V \otimes V)^*. \end{array}$$

Um eine Beziehung zwischen $(V \wedge V)^*$ und $V^* \wedge V^*$ zu erhalten, betrachten wir die Abbildung

$$\alpha: V^* \times V^* \rightarrow \text{Alt}^2(V; K), \quad (\varphi, \psi) \mapsto \alpha(\varphi, \psi),$$

mit

$$\alpha(\varphi, \psi)(v, w) := \det \begin{pmatrix} \varphi(v) & \varphi(w) \\ \psi(v) & \psi(w) \end{pmatrix}.$$

Aus den Eigenschaften der Determinante folgt, daß $\alpha(\varphi, \psi)$ alternierend ist. Also gehört dazu die lineare Abbildung

$$\alpha_{\wedge}: V^* \wedge V^* \rightarrow \text{Alt}^2(V; K).$$

Wir zeigen, daß sie für endlichdimensionales V ein Isomorphismus ist. Dazu wählen wir eine Basis $(v_1, \dots, v_n)$ von V und die duale Basis $(v_1^*, \dots, v_n^*)$ von V^* . Dann ist durch

$$v_i^* \wedge v_j^*, \quad 1 \leq i < j \leq n,$$

eine Basis von $V^* \wedge V^*$ gegeben. Nach Definition von α ist

$$\alpha_\wedge(v_i^* \wedge v_j^*)(v_k, v_l) = \begin{cases} 1 & \text{für } (i, j) = (k, l), \\ 0 & \text{sonst.} \end{cases}$$

Für ein beliebiges $\xi \in \text{Alt}^2(V; K)$ können wir daher ein eindeutiges Urbild unter $\alpha_\wedge$ finden, nämlich

$$\sum_{i < j} a_{ij}(v_i^* \wedge v_j^*) \quad \text{mit} \quad a_{ij} := \xi(v_i, v_j).$$

Also ist $\alpha_\wedge$ bijektiv. Insgesamt haben wir damit den

Satz. Für einen endlichdimensionalen Vektorraum V gibt es kanonische Isomorphismen

$$V^* \wedge V^* \rightarrow \text{Alt}^2(V; K) \rightarrow (V \wedge V)^*. \quad \square$$

Aufgaben zu 6.3

1. Es sei V ein Vektorraum über einen Körper K und $L \supset K$ ein Erweiterungskörper von L , d.h. L ist ein Körper und K ein Unterring von L (vgl. 1.3.2).

- a) Zeigen Sie, dass L eine Struktur als K -Vektorraum trägt.
- b) Für Elemente $\sum \lambda_i \otimes v_i \in L \otimes_K V$ und $\lambda \in L$ definieren wir eine skalare Multiplikation durch

$$\lambda \cdot \left(\sum \lambda_i \otimes v_i \right) := \sum \lambda \lambda_i \otimes v_i.$$

Zeigen Sie, dass $L \otimes_K V$ mit der üblichen Addition und dieser skalaren Multiplikation zu einem L -Vektorraum wird.

- c) Ist die Familie $(v_i)_{i \in I}$ eine Basis von V über K , so ist die Familie $(1 \otimes v_i)_{i \in I}$ eine Basis von $L \otimes_K V$ über L . Insbesondere gilt $\dim_K V = \dim_L (L \otimes_K V)$.
- d) Durch die Abbildung

$$\varphi: V \rightarrow L \otimes_K V, \quad v \mapsto 1 \otimes v,$$

ist ein Isomorphismus von K -Vektorräumen gegeben.

2. Es seien U, V, W Vektorräume über demselben Körper K .

- a) Zeigen Sie, dass die Menge $\text{Bil}(V, W; U)$ mit der Addition von Abbildungen und der üblichen Multiplikation mit Skalaren ein K -Vektorraum ist und dass die kanonische Abbildung

$$\begin{aligned}\text{Bil}(V, W; U) &\rightarrow \text{Hom}(V \otimes W, U), \\ \xi &\mapsto \xi_{\otimes},\end{aligned}$$

ein Vektorraumisomorphismus ist. Insbesondere erhält man für $V = W$ und $U = K$ einen Isomorphismus

$$\text{Bil}(V; K) \rightarrow (V \otimes V)^*, \quad \xi \mapsto \xi_{\otimes}.$$

- b) Zeigen Sie analog, dass die Menge $\text{Alt}^2(V; W)$ mit der Addition von Abbildungen und der üblichen Multiplikation von Skalaren ein K -Vektorraum ist, und dass die kanonische Abbildung

$$\begin{aligned}\text{Alt}^2(V; W) &\rightarrow \text{Hom}(V \wedge V, W), \\ \xi &\mapsto \xi_{\wedge},\end{aligned}$$

ein Vektorraumisomorphismus ist. Für $W = K$ erhält man einen Isomorphismus

$$\text{Alt}^2(V; K) \rightarrow V^* \wedge V^*, \quad \xi \mapsto \xi_{\wedge}.$$

3. In dieser Aufgabe betrachten wir die kanonische Abbildung

$$\eta: V \times W \rightarrow V \otimes W, \quad (v, w) \mapsto v \otimes w,$$

noch etwas genauer.

- a) Zeigen Sie, dass $Q := \eta(V \times W) \subset V \otimes W$ ein *Kegel* ist, d.h. für $u \in Q$ und $\lambda \in K$ ist $\lambda u \in Q$.
- b)* Für $V = K^m$ und $W = K^n$ gebe man Gleichungen für Q in $K^m \otimes K^n = K^{m \cdot n}$ an. (Hinweis: Beschreiben Sie η durch $z_{ij} := x_i y_j$.)
- c) Wann ist η injektiv/surjektiv/bijektiv?

4. Es seien V und W Vektorräume über einen Körper K und $(v_i)_{i \in I}$ bzw. $(w_j)_{j \in J}$ Familien linear unabhängiger Vektoren in V bzw. W .

- a) Die Familie

$$(v_i \otimes w_j)_{(i,j) \in I \times J}$$

ist linear unabhängig in $V \otimes_K W$.

- b) Für Vektoren $v \in V$ und $w \in W$ gilt:

$$v \otimes w = 0 \Rightarrow v = 0 \text{ oder } w = 0.$$

5. Für K -Vektorräume V, V', W, W' sowie Homomorphismen $F: V \rightarrow V'$ und $G: W \rightarrow W'$ definieren wir das Tensorprodukt von F und G durch

$$\begin{aligned}(F \otimes G): V \otimes W &\rightarrow V' \otimes W', \\ v \otimes w &\mapsto F(v) \otimes G(w).\end{aligned}$$

Zeigen Sie, dass hierdurch ein Vektorraum-Isomorphismus

$$\operatorname{Hom}_K(V, V') \otimes \operatorname{Hom}_K(W, W') \rightarrow \operatorname{Hom}_K(V \otimes W, V' \otimes W')$$

definiert wird.

6. Für Vektoren $v_1, v_2 \in V$ gilt:

$$v_1, v_2 \text{ sind linear abhängig} \Leftrightarrow v_1 \wedge v_2 = 0 \text{ in } V \wedge V.$$

7. Sei A ein K -Vektorraum. A heißt K -Algebra, falls es zusätzlich eine Multiplikationsabbildung

$$\mu: A \times A \rightarrow A, \quad (a, a') \mapsto \mu(a, a') =: a \cdot a',$$

mit folgenden Eigenschaften gibt:

1) μ ist K -bilinear.

2) A zusammen mit der Vektorraumaddition und der Multiplikation μ ist ein Ring.

a) Zeigen Sie, dass die folgenden K -Vektorräume auch K -Algebren sind:

i) der $\mathbb{R}$ -Vektorraum $\mathbb{C}$,

ii) der K -Vektorraum $M(n \times n; K)$ bzgl. der Matrizenmultiplikation,

iii) der K -Vektorraum $K[t_1, \dots, t_n]$ bzgl. der üblichen Multiplikation von Polynomen (vgl. Aufgabe 9 zu 1.3).

b) Sind K -Algebren A und B gegeben, so ist $A \otimes B$ als K -Vektorraum erklärt. Zeigen Sie, dass $A \otimes B$ auf eindeutige Weise so zu einer K -Algebra gemacht werden kann, dass für alle $a, a' \in A$ und $b, b' \in B$

$$(a \otimes b) \cdot (a' \otimes b') = (a \cdot a') \otimes (b \cdot b')$$

gilt.

c) Wird $K[t] \otimes K[t]$ wie in b) zu einer K -Algebra gemacht, so definiert der Vektorraum-Isomorphismus

$$K[t] \otimes K[t] \rightarrow K[t_1, t_2], \quad t^i \otimes t^j \mapsto t_1^i t_2^j,$$

aus Beispiel 6.3.4 a) einen Isomorphismus von Ringen mit $1_{K[t] \otimes K[t]} \mapsto 1_{K[t_1, t_2]}$.

8. Zeigen Sie in Analogie zu Theorem 6.3.8 die Existenz eines *symmetrischen Produktes*: Für jeden K -Vektorraum V gibt es einen K -Vektorraum $V \vee V$ zusammen mit einer symmetrischen Abbildung

$$\vee: V \times V \rightarrow V \vee V,$$

die folgende universelle Eigenschaft erfüllen: zu jedem K -Vektorraum W zusammen mit einer symmetrischen Abbildung $\xi: V \times V \rightarrow W$ gibt es genau eine lineare Abbildung $\xi_\vee$ derart, dass das Diagramm

$$\begin{array}{ccc}
 V \times V & & \\
 \downarrow & \searrow \xi & \\
 V \vee V & \xrightarrow{\xi_V} & W
 \end{array}$$

kommutiert. Ist $(v_1, \dots, v_n)$ eine Basis von V , so ist durch $v_i \vee v_j := \vee(v_i, v_j)$ mit $i \leq j$ eine Basis von $V \vee V$ gegeben. Insbesondere ist

$$\dim(V \vee V) = \binom{n+1}{2} = \frac{(n+1)n}{2}.$$

9. Beweisen Sie mit Hilfe der universellen Eigenschaften aus Theorem 6.3.3, Theorem 6.3.8 und Aufgabe 8 die Eindeutigkeit von Tensorprodukt, äußerem Produkt und symmetrischem Produkt, d.h.

- a) gibt es $\tilde{\eta}: V \times W \rightarrow V \tilde{\otimes} W$ mit denselben Eigenschaften, dann existiert ein Isomorphismus τ , so dass das Diagramm

$$\begin{array}{ccc}
 V \times W & & \\
 \eta \swarrow & & \searrow \tilde{\eta} \\
 V \otimes W & \xrightarrow{\tau} & V \tilde{\otimes} W
 \end{array}$$

kommutiert.

- b) gibt es $\tilde{\wedge}: V \times W \rightarrow V \tilde{\wedge} W$ mit denselben Eigenschaften, dann existiert ein Isomorphismus τ , so dass das Diagramm

$$\begin{array}{ccc}
 V \times W & & \\
 \wedge \swarrow & & \searrow \tilde{\wedge} \\
 V \wedge W & \xrightarrow{\tau} & V \tilde{\wedge} W
 \end{array}$$

kommutiert.

- c) gibt es $\tilde{\vee}: V \times W \rightarrow V \tilde{\vee} W$ mit denselben Eigenschaften, dann existiert ein Isomorphismus τ , so dass das Diagramm

$$\begin{array}{ccc}
 V \times W & & \\
 \vee \swarrow & & \searrow \tilde{\vee} \\
 V \vee W & \xrightarrow{\tau} & V \tilde{\vee} W
 \end{array}$$

kommutiert.

6.4 Multilineare Algebra

Die bisher behandelten Tensorprodukte haben zwei Faktoren, sie hängen zusammen mit bilinearen Abbildungen. Etwa in der Integrationstheorie ([Fo 3], §19) und der Differentialgeometrie ([K-N], Chapter I-2) benötigt man auch Produkte mit mehreren Faktoren. In diesem letzten Abschnitt soll noch kurz das Wichtigste darüber zusammengestellt werden.

6.4.1. Zunächst der grundlegende Begriff: Gegeben seien K -Vektorräume $V_1, \dots, V_k$ und W . Eine Abbildung

$$\xi: V_1 \times \dots \times V_k \rightarrow W$$

heißt *multilinear* (oder *k-fach linear*), wenn für jedes $i \in \{1, \dots, k\}$ und fest gewählte $v_j \in V_j$ ($j = 1, \dots, i-1, i+1, \dots, k$) die Abbildung

$$V_i \rightarrow W, \quad v \mapsto \xi(v_1, \dots, v_{i-1}, v, v_{i+1}, \dots, v_k),$$

K -linear ist. Kurz ausgedrückt: hält man alle bis auf eine Variable fest, so entsteht jeweils eine lineare Abbildung. Eine derartige Bedingung war schon bei der Definition der Determinante benutzt worden (D1 in 3.1.2).

Ganz analog zu 6.3.3 beweist man das

Theorem. Zu K -Vektorräumen $V_1, \dots, V_k$ gibt es einen K -Vektorraum $V_1 \otimes \dots \otimes V_k$ zusammen mit einer universellen multilinearen Abbildung

$$\eta: V_1 \times \dots \times V_k \rightarrow V_1 \otimes \dots \otimes V_k, \quad (v_1, \dots, v_k) \mapsto v_1 \otimes \dots \otimes v_k,$$

d.h. zu jeder multilinearen Abbildung

$$\xi: V_1 \times \dots \times V_k \rightarrow W$$

gibt es genau eine lineare Abbildung $\xi_{\otimes}$ derart, daß das Diagramm

$$\begin{array}{ccc} V_1 \times \dots \times V_k & & \\ \eta \downarrow & \searrow \xi & \\ V_1 \otimes \dots \otimes V_k & \xrightarrow{\xi_{\otimes}} & W \end{array}$$

kommutiert. Sind alle V_j endlichdimensional mit Basen

$$(v_1^{(j)}, \dots, v_{r_j}^{(j)}), \quad j = 1, \dots, k,$$

so ist eine Basis von $V_1 \otimes \dots \otimes V_k$ gegeben durch die Produkte

$$v_{i_1}^{(1)} \otimes \dots \otimes v_{i_k}^{(k)} \quad \text{mit } 1 \leq i_j \leq r_j.$$

Insbesondere ist $\dim(V_1 \otimes \dots \otimes V_k) = \dim V_1 \cdot \dots \cdot \dim V_k$. □

Ein wichtiger Spezialfall ist folgender: Für einen K -Vektorraum V erklärt man

$$T := \underbrace{V^* \otimes \dots \otimes V^*}_{p\text{-mal}} \otimes \underbrace{V \otimes \dots \otimes V}_{q\text{-mal}}.$$

Ein Element von T wird p -fach *kovarianter* und q -fach *kontravarianter* Tensor genannt. Für seine Darstellung sei $(v_1, \dots, v_n)$ eine Basis von V und (in der üblichen Bezeichnungsweise der Physik) $(v^1, \dots, v^n)$ die zugehörige duale Basis von V^* . Dann hat jedes Element von T eine eindeutige Darstellung

$$\sum_{i_1, \dots, i_p, j_1, \dots, j_q} a_{i_1 \dots i_p}^{j_1 \dots j_q} v^{i_1} \otimes \dots \otimes v^{i_p} \otimes v_{j_1} \otimes \dots \otimes v_{j_q}$$

mit $i_1, \dots, i_p, j_1, \dots, j_q \in \{1, \dots, n\}$ und $a_{i_1 \dots i_p}^{j_1 \dots j_q} \in K$. In dieser Form begegnet man Tensoren in der Physik.

6.4.2. Auch die Verallgemeinerung symmetrischer und äußerer Produkte von 2 auf k Faktoren bereitet keine grundsätzlichen Probleme, sie erfordert nur viele Buchstaben, die auf den ersten Blick recht verwirrend sein können. Zur Abkürzung schreiben wir dabei

$$V^k := \underbrace{V \times \dots \times V}_{k\text{-mal}}$$

für einen Vektorraum V und eine natürliche Zahl $k \geq 1$.

Definition. Sind V und W Vektorräume über K , so heißt eine k -fach lineare Abbildung

$$\xi: V^k \rightarrow W$$

symmetrisch, wenn $\xi(v_1, \dots, v_k) = \xi(v_{\sigma(1)}, \dots, v_{\sigma(k)})$
für jede Permutation $\sigma \in S_k$,

alternierend, wenn $\xi(v_1, \dots, v_k) = 0$,
falls $v_i = v_j$ für ein Paar (i, j) mit $i \neq j$.

Analog zu 3.2.3 beweist man die

Bemerkung. Ist ξ alternierend und $\sigma \in S_k$, so ist

$$\xi(v_{\sigma(1)}, \dots, v_{\sigma(k)}) = \text{sign}(\sigma) \cdot \xi(v_1, \dots, v_k).$$

In $\bigotimes^k V := \underbrace{V \otimes \dots \otimes V}_{k\text{-mal}}$ betrachten wir die folgenden Untervektorräume:

$$\begin{aligned} S^k(V) &:= \text{span}(v_1 \otimes \dots \otimes v_k - v_{\sigma(1)} \otimes \dots \otimes v_{\sigma(k)}), \\ &\quad \text{wobei } v_1 \otimes \dots \otimes v_k \in \bigotimes^k V \text{ und } \sigma \in S_k, \\ A^k(V) &:= \text{span}(v_1 \otimes \dots \otimes v_k), \\ &\quad \text{wobei } v_i = v_j \text{ für ein } (i, j) \text{ mit } i \neq j. \end{aligned}$$

Analog zu 6.3.7 folgt das

Lemma. Für jedes k -fach lineare $\xi: V^k \rightarrow W$ gilt:

$$\begin{aligned} \xi \text{ symmetrisch} &\Leftrightarrow S^k(V) \subset \text{Ker } \xi_{\otimes}, \\ \xi \text{ alternierend} &\Leftrightarrow A^k(V) \subset \text{Ker } \xi_{\otimes}. \end{aligned}$$

Theorem. Zu einem K -Vektorraum V und einer natürlichen Zahl $k \geq 1$ gibt es einen K -Vektorraum $\bigwedge^k V$ zusammen mit einer universellen alternierenden Abbildung

$$\wedge: V^k \rightarrow \bigwedge^k V,$$

d.h. zu jeder alternierenden Abbildung

$$\xi: V^k \rightarrow W$$

gibt es genau eine lineare Abbildung $\xi_{\wedge}$ derart, daß das Diagramm

$$\begin{array}{ccc} V^k & & \\ \downarrow \wedge & \searrow \xi & \\ \bigwedge^k V & \xrightarrow{\xi_{\wedge}} & W \end{array}$$

kommutiert. Ist $(v_1, \dots, v_n)$ eine Basis von V , so ist eine Basis von $\bigwedge^k V$ gegeben durch die Produkte

$$v_{i_1} \wedge \dots \wedge v_{i_k} \quad \text{mit } 1 \leq i_1 < \dots < i_k \leq n.$$

Insbesondere ist $\dim \bigwedge^k V = \binom{n}{k}$ für $1 \leq k \leq n = \dim V$. Für $k > n$ setzt man

$$\bigwedge^k V = 0.$$

Beweis. Man erklärt $\wedge$ als Komposition $V^k \rightarrow \bigotimes^k V \rightarrow \bigotimes^k V / A^k(V)$.

Zur Konstruktion der Basis benutze man $N = \binom{n}{k}$, den K^N mit Basis $e_{i_1 \dots i_k}$ und die alternierende Abbildung

$$\xi: V^k \rightarrow K^N, \quad (w_1, \dots, w_k) \mapsto \sum a_{i_1 \dots i_k} \cdot e_{i_1 \dots i_k}.$$

Dabei sei $w_i = \sum_{j=1}^n \lambda_{ij} v_j$, daraus erhält man eine $(k \times n)$ -Matrix $A = (\lambda_{ij})$, und $a_{i_1 \dots i_k}$ ist der zu den Spalten $i_1, \dots, i_k$ gehörende Minor der Matrix A . $\square$

Damit sollte der Leser gerüstet sein für das Studium der Integrationstheorie etwa in [Fo 3].

Aufgaben zu 6.4

1. Führen Sie die Einzelheiten des Beweises von Theorem 6.4.1 aus.

2. Zeigen Sie, dass für K -Vektorräume V_1, V_2 und V_3 die kanonischen Abbildungen, die gegeben sind durch

$$\begin{array}{ccccc} (V_1 \otimes V_2) \otimes V_3 & \longrightarrow & V_1 \otimes V_2 \otimes V_3 & \longleftarrow & V_1 \otimes (V_2 \otimes V_3) \\ (v_1 \otimes v_2) \otimes v_3 & \longmapsto & v_1 \otimes v_2 \otimes v_3 & \longleftarrow & v_1 \otimes (v_2 \otimes v_3), \end{array}$$

Isomorphismen sind. Folgern Sie daraus, dass für jeden K -Vektorraum W die Vektorräume

$$\text{Bil}((V_1 \otimes V_2), V_3; W), \quad \text{Bil}(V_1, (V_2 \otimes V_3); W)$$

(vgl. Aufgabe 2 zu 6.3) und

$$\text{Tril}(V_1, V_2, V_3; W) := \{\xi: V_1 \times V_2 \times V_3 \rightarrow W: \xi \text{ trilinear}\}$$

kanonisch isomorph sind.

3. Beweisen Sie Theorem 6.4.2.

4. Es sei V ein K -Vektorraum.

a) Für Vektoren $v_1, \dots, v_k \in V$ gilt:

$$v_1, \dots, v_k \text{ sind linear abhängig} \Leftrightarrow v_1 \wedge \dots \wedge v_k = 0 \text{ in } \bigwedge^k V.$$

b) Ist $\dim V = n$, so gilt $\bigwedge^k V = 0$ für $k > n$.

5. Beweisen Sie die folgende Verallgemeinerung von Aufgabe 7 zu Abschnitt 6.3:

Zu einem K -Vektorraum V und einer natürlichen Zahl $k \geq 1$ gibt es einen K -Vektorraum $\bigvee^k V$ zusammen mit einer universellen symmetrischen Abbildung

$$\vee: V^k \rightarrow \bigvee^k V,$$

d.h. zu jeder symmetrischen Abbildung

$$\xi: V^k \rightarrow W$$

gibt es genau eine lineare Abbildung $\xi_\vee$ derart, dass das Diagramm

$$\begin{array}{ccc} V^k & & \\ \downarrow \vee & \searrow \xi & \\ \bigvee^k V & \xrightarrow{\xi_\vee} & W \end{array}$$

kommutiert. Ist $(v_1, \dots, v_n)$ eine Basis von V , so ist eine Basis von $\bigvee^k V$ gegeben durch die Produkte

$$v_{i_1} \vee \dots \vee v_{i_k} \quad \text{mit } 1 \leq i_1 \leq \dots \leq i_k \leq n.$$

Insbesondere ist $\dim \bigvee^k V = \binom{n+k-1}{k}$.

Der Raum $\bigvee^k V$ heißt *symmetrisches Produkt der Ordnung k über V* .

6. Es sei $(e_1, \dots, e_n)$ eine Basis des Standardraumes K^n . Mit $K[t_1, \dots, t_n]_{(k)}$ bezeichnen wir den Vektorraum der homogenen Polynome vom Grad k in den Unbestimmten $t_1, \dots, t_n$ (vgl. Aufgabe 9 zu 1.3). Zeigen Sie, dass durch die Zuordnung

$$\bigvee^k K^n \rightarrow K[t_1, \dots, t_n]_{(k)}, \quad e_{i_1} \vee \dots \vee e_{i_k} \mapsto t_{i_1} \cdot \dots \cdot t_{i_k},$$

ein Isomorphismus von K -Vektorräumen definiert wird.

7. V sei ein endlichdimensionaler K -Vektorraum und $\alpha = (\alpha_1 \wedge \dots \wedge \alpha_k) \in \bigwedge^k V$ sowie $\beta = (\beta_1 \wedge \dots \wedge \beta_l) \in \bigwedge^l V$.

a) Zeigen Sie, dass eine bilineare Abbildung

$$\mu: \bigwedge^k V \times \bigwedge^l V \rightarrow \bigwedge^{k+l} V$$

mit

$$(\alpha, \beta) \mapsto \alpha_1 \wedge \dots \wedge \alpha_k \wedge \beta_1 \wedge \dots \wedge \beta_l$$

existiert. Das Element $\alpha \wedge \beta := \mu(\alpha, \beta)$ heißt *äußeres Produkt* von α und β .

b) Es gilt

$$\alpha \wedge \beta = (-1)^{k \cdot l} \beta \wedge \alpha.$$

8. Es sei V ein endlichdimensionaler K -Vektorraum mit $\dim V = n$.

a) Zeigen Sie, dass die bilinearen Abbildungen, die durch die folgenden Zuordnungen definiert werden, nicht ausgeartet sind (vgl. 6.2.1).

$$\text{i) } \bigwedge^k V \times \bigwedge^{n-k} V \rightarrow \bigwedge^n V \cong K, \quad (\alpha, \beta) \mapsto \alpha \wedge \beta.$$

Die Isomorphie von $\bigwedge^n V$ und K ergibt sich dabei aus Theorem 6.4.2.

ii) Als Verallgemeinerung des Beispiels aus 6.2.1

$$\bigwedge^k V^* \times \bigwedge^k V \rightarrow K, \quad (\varphi_1 \wedge \dots \wedge \varphi_k, v_1 \wedge \dots \wedge v_k) \mapsto \det \varphi(v),$$

wobei

$$\varphi(v) = \begin{pmatrix} \varphi_1(v_1) & \dots & \varphi_1(v_k) \\ \vdots & & \vdots \\ \varphi_k(v_1) & \dots & \varphi_k(v_k) \end{pmatrix}.$$

b) Folgern Sie aus Teil a), dass es kanonische Isomorphismen

$$\text{i) } \bigwedge^k V^* \rightarrow \left(\bigwedge^k V \right)^* \text{ und}$$

$$\text{ii) } \bigwedge^k V \rightarrow \bigwedge^{n-k} V^*$$

gibt.

9. V und W seien K -Vektorräume. Zeigen Sie, dass die Menge

$$\text{Alt}^k(V; W) := \{\xi: V^k \rightarrow W: \xi \text{ ist alternierend}\}$$

zusammen mit der Addition von Abbildungen und der üblichen Multiplikation mit Skalaren ein K -Vektorraum ist, und dass die kanonische Abbildung

$$\text{Alt}^k(V; W) \rightarrow \text{Hom}(\bigwedge^k V, W), \quad \xi \mapsto \xi_{\wedge},$$

ein Vektorraumisomorphismus ist. Insbesondere erhält man für $W = K$ einen kanonischen Isomorphismus

$$\text{Alt}^k(V; K) \rightarrow \bigwedge^k V^*.$$

Literaturverzeichnis

Lehrbücher der linearen Algebra

- [B 1] E. Brieskorn: *Lineare Algebra und Analytische Geometrie I*. Vieweg 1983.
- [B 2] E. Brieskorn: *Lineare Algebra und Analytische Geometrie II*. Vieweg 1985.
- [Bo] N. Bourbaki: *Éléments de mathématique*, speziell Serie VI. – Livre II. *Algèbre*, Ch. 2. *Algèbre linéaire*. Hermann 1962.
- [J] K. Jänich: *Lineare Algebra*. Springer 1979.
- [Koe] M. Koecher: *Lineare Algebra und analytische Geometrie*. Springer 1983.
- [Kow 1] G. Kowalewski: *Einführung in die Determinantentheorie*. Veit & Co 1909. 4. Auflage, W. de Gruyter 1953.
- [Kow 2] G. Kowalewski: *Einführung in die Analytische Geometrie*. Veit & Co 1910. 4. Auflage, W. de Gruyter 1953.
- [S-S] O. Schreier und E. Sperner: *Einführung in die Analytische Geometrie und Algebra*. Teubner 1931.
- [Str] G. Strang: *Linear Algebra and its Applications*. Academic Press 1976.

Ergänzende Literatur

- [C] H. Cartan: *Nicolas Bourbaki und die heutige Mathematik*. Arbeitsgemeinschaft für Forschung des Landes Nordrhein-Westfalen, Heft 76. Westdeutscher Verlag 1959.
- [D] L.E. Dickson: *New First Course in the Theory of Equations*. Wiley 1939.
- [DIN] DIN-Taschenbuch 202: *Formelzeichen, Formelsatz, Mathematische Zeichen und Begriffe*. Beuth 1994.
- [E] R. Ehrlich: *Why toast lands jelly-side down*. Princeton University Press 1997.
- [Fi] G. Fischer: *Analytische Geometrie*. Vieweg 1978.
- [F-L] W. Fischer und I. Lieb: *Funktionentheorie I*. Vieweg 1980.
- [Fo 1] O. Forster: *Analysis 1*. Vieweg 1976.
- [Fo 2] O. Forster: *Analysis 2*. Vieweg 1977.
- [Fo 3] O. Forster: *Analysis 3*. Vieweg 1981.
- [F-P] U. Friedrichsdorf und A. Prestel: *Mengenlehre für den Mathematiker*. Vieweg 1985.
- [Fr] F.G. Frobenius: *Zur Theorie der linearen Gleichungen*. Journal für die reine und angewandte Mathematik **129**, 175–180 (1905).
- [F-S] G. Fischer und R. Sacher: *Einführung in die Algebra*. Teubner 1974.
- [K-N] S. Kobayashi and K. Nomizu: *Foundations of Differential Geometry, Vol. 1*. Wiley-Interscience 1963.

- [L] S. Lang: *Algebra*. Addison-Wesley 1965.
- [Le] *Lexikon bedeutender Mathematiker*. Bibliographisches Institut Leipzig, 1990.
- [M-V] R. Meise und D. Vogt: *Einführung in die Funktionalanalysis*. Vieweg 1992.
- [O] G. Opfer: *Numerische Mathematik für Anfänger*. Vieweg 1992.
- [P] J.-P. Petit: *Das Geometrikon*. Vieweg 1995.
- [Sch] W. Scharlau: *Schulwissen Mathematik: Ein Überblick*. Vieweg 1994.
- [St] J. Stoer: *Einführung in die numerische Mathematik*. Springer 1972.
- [We] H. Weber: *Lehrbuch der Algebra, Band 1*. Vieweg 1895.
- [Wi] F.A. Willers: *Methoden der praktischen Analysis*, 3. Auflage. W. de Gruyter 1957.
- [Z] H.D. Ebbinghaus et al: *Zahlen*. Springer 1983.

Namensverzeichnis

- ABEL, NIELS HENRIK (1802–1829), 69
BESSEL, FRIEDRICH WILHELM (1784–1846), 286
BOURBAKI, NICOLAS (1934–1983 ?), VI
CAUCHY, AUGUSTIN (1789–1857), 275
CAVALIERI, BONAVENTURA (1598–1647), 175
CAYLEY, ARTHUR (1821–1895), VI, 20, 251
COURANT, RICHARD (1888–1972), V
CRAMER, GABRIEL (1704–1752), 175, 205
DESCARTES, RENÉ (1596–1650), 1, 72
DIRICHLET, PETER (1805–1859), 36
EUKLID (etwa 365–300 v.Chr.), 292
FITTING, HANS (1906–1938), 261
FONTENÉ, G. (?), 130
FROBENIUS, FERDINAND GEORG (1849–1917), 130
GAUSS, CARL FRIEDRICH (1777–1855), 22, 66
GOETHE, JOHANN WOLFGANG VON (1749–1832), 52
GRAM, JØRGEN (1850–1916), 210
HADAMARD, JAUQUES (1865–1963), 298
HAMILTON, SIR WILLIAM (1805–1865), 251
HILBERT, DAVID (1862–1943), VI
JACOBI, CARL GUSTAV (1804–1851), 145, 327
JORDAN, CAMILLE (1838–1922), 264
KOWALEWSKI, GERHARD (1876–1950), VI
KRONECKER, LEOPOLD (1823–1891), 93
LANG, SERGE (1927–), 251
LAPLACE, PIERRE SIMON (1749–1827), 203
LEIBNIZ, GOTTFRIED WILHELM (1646–1716), VI, 174, 192
MÖBIUS, AUGUST FERDINAND (1790–1868), 220
NOETHER, EMMY (1882–1935), VI
PEANO, GUISEPPE (1858–1939), 32
PFAFF, JOHANN FRIEDRICH (1765–1825), 183
PYTHAGORAS (etwa 580–500 v.Chr.), 347
ROUCHÉ, EUGENE (1832–1910), 130
SARRUS, PIERRE (1798–1861), 195
SCHMIDT, ERHARD (1876–1959), 296

- SCHREIER, OTTO (1901–1929), VI
SCHWARZ, HERMANN AMANDUS (1843–1921), 275
STEINITZ, ERNST (1871–1928), 89
STIRLING, JAMES (1692–1770), 187
SYLVESTER, JAMES (1814–1897), 322
VANDERMONDE, ALEXANDRE (1735–1796), 196
VIETA, FRANÇOIS (1540–1603), 70
WEBER, HEINRICH (1842–1913), VI
WEIERSTRASS, KARL (1815–1897), 174, 178, 192
WEYL, HERMANN (1885–1955), 40
ZORN, MAX (1906–1993), 88

Einzelheiten zu Leben und Werk findet man zum Beispiel in [Le].

Sachwortverzeichnis

- Abbildung, 34, 84
 - 2π -periodische, 84
 - adjungierte, 342
 - bilineare, 350
 - duale, 334
 - kanonische, 122, 357
 - lineare, 109
 - multilineare, 366
- abelsch, 43
- Ableitungshomomorphismus, 127, 141
- Absolutbetrag, 58
- Abstand, 275, 279, 349
 - Punkt von Gerade, 279
 - Punkt von Hyperebene, 280
 - windschiefer Geraden, 349
- abzählbar unendlich, 42
- Addition
 - von Abbildungen, 76, 84
 - von Matrizen, 75
 - von Polynomen, 62
 - von Vektoren, 2, 76
- Addition in Ringen, 54
- Additionstheorem, 146
- adjungiert, 342
- ähnlich, 160, 273
- äquivalent, 40, 120, 160
- Äquivalenzklasse, 40
- Äquivalenzrelation, 40
- äußeres Produkt, 358, 367
- affiner Unterraum, 30, 116
- Algebra, 364
- alternierend, 105, 179, 184, 192, 288, 358, 367, 371
- Annulator, 333, 341
- antihermitesch, 348
- Anzahl, 32
- arithmetisches Mittel, 43
- assoziativ, 37, 43, 54, 147
- aufgespannt, 80
- Ausartungsraum, 322
- Austauschlemma, 89
- Austauschsatz, 90
- Auswahlaxiom, 39
- Auswahlfunktion, 38
- Automorphismus, 109
- Basis, 86, 88
 - kanonische, 86
 - orthonormale, 295
 - verformbare, 215
- Basisauswahlsatz, 88
- Basisergänzungssatz, 91
- Beschränkung, 35
- Besselsche Ungleichung, 301
- Bidualraum, 336
- bijektiv, 35, 37
- Bild, 35, 114
- Bilinearform, 288
 - alternierende, 288
 - nicht ausgeartete, 340
 - symmetrische, 288
- Cauchy-Schwarzsche Ungleichung, 275, 293
- Cavalierisches Prinzip, 175
- Cayley-Hamilton
 - Satz von, 251
- Charakteristik, 61
- charakteristisches Polynom, 229, 230
- Cosinussatz, 278
- Cramersche Regel, 175, 205

- Darboux-Basis, 302
- darstellende Matrix, 288
- Determinante, 198
 - einer Matrix 174, 178, 192, 197
 - eines Endomorphismus, 210
 - Gramsche, 298
- Determinanten-Multiplikationssatz, 180
- Determinanten-Multiplikationstheorem, 208
- diagonalisierbar, 224, 225, 234, 235, 270, 307
 - simultan, 239
- Diagonalmatrix, 225
- Diagramm,
 - kommutatives, 157
- Diedergruppe, 53
- Differentialgleichung, 199, 200, 227, 240
- Differenzmenge, 34
- Dimension, 1, 91, 117
- Dimensionsformel,
 - für Kern und Bild, 117
 - für Summen, 100
- direkte Summe, 101
- direktes Produkt, 38
- Diskriminante, 69
- distributiv, 54, 56, 147
- Division mit Rest, 63
- Drehung, 106, 146, 223, 231, 305
- Dreieck, 175
- Dreiecksmatrix, 175, 243
- Dreiecksungleichung, 275
- duale Basis, 332
- Dualraum, 331
- Durchschnitt, 33, 79
- Ebene, 13
- Eigenraum, 226, 231, 259
- Eigenvektor, 223
- Eigenwert, 222
- Einheitsmatrix, 92, 147
- Einselement, 54
- Eintrag, 21
- Element, 32
 - inverses, 43, 45
 - negatives, 44
 - neutrales, 43, 44, 147
- Elementarmatrix, 164
- elementarsymmetrisch 70
- Elimination, 22, 30
- Eliminationsverfahren von Gauß, 20, 30, 130, 171
- Ellipse, 320
- endlich erzeugt, 86
- Endomorphismenring, 112
- Endomorphismus, 109, 112
 - adjunktiver, 312
 - ähnlicher, 273
 - diagonalisierbarer, 224
 - nilpotenter, 257
 - normaler, 344
 - orthogonaler, 303
 - selbstadjungierter, 312
 - trigonalisierbarer, 244
 - unitärer, 303
- Entwicklung, 203
- Erzeugendensystem, 86, 87
- erzeugt, 80, 166
- Erzeugung von linearen Abbildungen, 138
- euklidisch, 292
- Exponentialfunktion, 272
- Fahne
 - invariante, 243
- Faktorgruppe, 53
- Faktorisierungssatz, 119
- Fakultät, 187
- Familie, 38, 80
- Faser, 114
- Faserung, 115
- Fehlstand, 189
- Fitting
 - Lemma von, 261

- Fixpunkt, 112
Folge
 beschränkte, 85
 konvergente, 85
Fourierkoeffizient, 301
Fundamentalsatz der Algebra, 66
Fundamentalsystem, 133
Funktion
 differenzierbare, 78
 lineare, 106
 rationale, 74
 stetige, 78
Fußball
 Satz vom, 307

Gerade, 4, 8, 346
gleichmäßig, 42
Gleichung
 lineare, 2
Gleichungssystem
 lineares, 20, 129, 337
Grad, 61, 73
Gram-Schmidtscher Orthogonalisierungssatz, 296
Gramsche Determinante, 210, 298
Graph, 39, 40
Grassmann-Identität, 284
Gruppe, 43
 allgemeine lineare, 150
 alternierende, 192
 orthogonale, 304
 spezielle orthogonale, 304
 symmetrische, 44, 186
 unitäre, 304
 zyklische, 49, 53
Gruppentafel, 46

Hadamardsche Ungleichung, 298
Hauptachsentransformation, 319
Hauptminor, 327
Hauptminoren-Kriterium für Definitheit, 327
Hauptraum, 259
hermitesch, 291
Hesse-Matrix, 153
Hessesche Normalform, 280, 281
homogen, 73, 290
Homomorphismus
 von Gruppen, 48
 von Ringen, 56
 von Vektorräumen, 109
Hyperbel, 320
Hyperebene, 280

Ideal, 251, 255
imaginäre Einheit, 58
Imaginärteil, 58
Indexmenge, 33, 38
Induktion, 32
induziert, 48, 78
injektiv, 35, 37
invariant, 242
Isometrie, 303
Isomorphismus, 48, 109, 118, 149
 kanonischer, 336, 340

Jacobi-Identität, 284
Jakobi-Matrix, 145, 153
Jordanblock, 269
Jordanmatrix, 264
Jordansche Normalform, 268

kanonisch, 41, 86, 122, 140, 214, 274, 336
Kern, 114
Klasse, 41, 49
Koeffizient, 4, 61
Koeffizientenmatrix, 21, 129
 erweiterte, 21, 131
Körper, 56
kommutativ, 43, 54, 157

- Komplement, 34
 - orthogonales, 294, 341
- komplementär, 201
- komplexe Struktur, 287, 311
 - ω -kalibrierte, 311
- Komplexifizierung, 253, 310, 356
- Komponente, 1, 38
- Komposition, 36, 38
- kongruent, 50
- Konjugation, 291
- konjugiert, 58
- Koordinaten, 154
 - kartesische, 1
- Koordinatensystem, 138, 154
- Kronecker-Symbol, 93
- Kürzungsregel, 45, 161
- Länge einer Basis, 86, 90
- Längenmessung, 297
- Laplacescher Entwicklungssatz, 203
- Leibnizsche Formel, 192
- linear abhängig, 18, 81
- linear unabhängig, 18, 81
- lineare Optimierung, 2
- lineares Gleichungssystem, 20
- Linearfaktor, 66
- Linearform, 331
- Linearkombination, 79
 - triviale, 81
- Linkstranslation, 45
- Lösung, 4, 11, 21, 129
 - allgemeine, 130
 - eindeutige, 134
 - spezielle, 133
 - triviale, 26, 130, 134
 - universelle, 134
- Lösungsmenge, 21, 78
- Lösungsräume, 129
- Matrix, 20, 75
 - ähnliche, 160, 230
 - äquivalente, 160
 - antihermitesche, 348
 - darstellende, 139, 156, 288, 291
 - diagonalisierbare, 225
 - dünn besetzt, 151
 - hermitesche, 292, 312
 - inverse, 150, 166, 204
 - invertierbare, 149, 166
 - komplementäre, 201
 - normale, 344
 - orthogonale, 304
 - positiv definite, 292, 321, 327
 - Rang einer, 30
 - symmetrische, 104, 312, 315, 319, 325
 - transponierte, 334
 - trigonalisierbare, 244
 - unitäre, 304, 307
- Matrizenring, 148
- Menge, 32
 - endliche, 32
 - leere, 32
 - unendliche, 32
- Metrik, 275, 294, 300
- Minimalpolynom, 256
- Minor, 206, 207
- Modul, 92
- Möbiusband, 220
- Monom, 355
- multilinear, 366
- Multiplikation
 - mit Skalaren, 2, 76
 - von Endomorphismen, 112
 - von Matrizen, 144
 - von Polynomen, 62
- Multiplikation in Ringen, 54

- negativ definit, 330
- Negatives, 44, 76
- nicht ausgeartet, 370
- nilpotent, 227, 257, 264
- Norm, 274, 286, 293, 294, 300
- normal, 344
- Normalenvektor, 280
- Normalform, 161, 309
 - ähnliche Matrizen, 268
 - äquivalente Matrizen, 261
 - Jordansche, 268
 - nilpotenter Endomorphismus, 264
 - orthogonaler Endomorphismus, 308
 - selbstadjungierter Endomorphismus, 312
 - symmetrischer Bilinearformen, 324
 - unitärer Endomorphismen, 307
- normiert, 61, 179
- n -Tupel, 38
- Nullelement, 44, 54
- Nullpolynom, 61
- Nullstelle, 64
 - mehrfache, 66
 - Vielfachheit, 66
- nullstellenfrei, 57
- nullteilerfrei, 55
- Nullvektor, 3, 76
- Optimierung
 - lineare, 2
- Orientierung, 214
 - kanonische, 214
- orientierungstreu, 213
- orthogonal, 279, 294, 303, 304, 308, 333
- orthonormal, 295
- Orthonormalisierung, 296, 324
- paarweise verschieden, 32
- parallel, 346
- Parallelogramm, 175, 283
- Parallelogramm-Gleichung, 278, 300
- Parallelotop, 177
- Parallelprojektion, 119
- Parameter, 4, 24
- Parameterdarstellung, 11
- Parametrisierung, 4, 13, 22, 30, 131
- Parsevalsche Ungleichung, 301
- Peano-Axiome, 32
- Permutation, 44, 186
 - gerade, 189
 - ungerade, 189
- Pfaffsches Polynom, 183, 199
- Pivot, 23, 29
- Plückerkoordinaten, 211
- Polarisierung, 291, 292
- Polynom, 61, 73, 106, 350, 355
 - charakteristisches, 229, 230
 - homogenes, 73, 98, 370
 - trigonometrisches, 84, 301
- Polynomring, 62, 73
- positiv definit, 292, 321, 327
- Produkt
 - äußeres, 358, 370
 - direktes, 38, 84, 98, 104
 - symmetrisches, 364, 370
 - von Matrizen, 144
- Projektion, 38, 296
- Pythagoras,
 - Satz von, 278
- quadratische Ergänzung, 326
- quadratische Form, 290
- Quotientenmenge, 41
- Quotientenvektorraum, 123
- Rang, 114, 148, 159, 206, 335
- Rang einer Matrix, 30
- Raum
 - affiner, 30
- Realteil, 58
- Rechtstranslation, 45
- reflexiv, 40

Reihe

absolut konvergente, 85

Relation, 40

Repräsentant, 41, 50

Restklasse, 50

Restklassenring, 55

Resultante, 185

Ring, 54

Römerfläche, 51

Sarrussche Regel, 195

Scherung, 177

schiefsymmetrisch, 105, 183, 288

Schnitt, 119, 132

Schubladenprinzip, 36

Schwingung, 238

Schwingungsgleichung, 240

selbstadjungiert, 312

semilinear, 291

senkrecht, 277

sesquilinear, 291

Signum, 189

Skalarprodukt, 292, 300

kanonisches, 274, 286

Spaltenrang, 96, 99, 159

Spaltenraum, 96

Spaltenumformung, 165

Spat, 177, 298

spezielle lineare Gruppe, 199

Spiegelung, 223, 231, 305

Spur, 229

Standardbasis, 86

Standardebene, 13

Standardraum, 75

reeller, 1

stetig verformbar, 215

Stirlingsche Formel, 187

Sturmsche Ketten, 72

Summand

direkter, 102, 120

Summationsindex, 21

Summe

direkte, 101, 102

orthogonale, 295

von Vektorräumen, 100

Summenzeichen, 21

surjektiv, 35, 37

Sylvestersches Trägheitsgesetz, 322

symmetrisch, 40, 104, 232, 288, 358, 367, 369

symmetrische Gruppe, 44, 186

Symmetrisierung, 105, 310

symplektischer Vektorraum, 302, 311

Teiler, 62

Teilmatrix, 206

Teilmenge, 32, 33

Tensor, 350, 353

kontravarianter, 367

kovarianter, 367

Tensorprodukt, 353, 366

Trägheitssatz, 322

Transformationsformel, 290, 291

für Bilinearformen, 290

für Endomorphismen, 158

für lineare Abbildungen, 158

für Sesquilinearformen, 291

Transformationsmatrix, 154, 169, 325

transitiv, 40

Translation, 45

Transposition, 95, 188, 337

trigonalisierbar, 244

überabzählbar, 42

Umkehrabbildung, 35

Unbestimmte, 4, 61, 73, 182

unitär, 292, 303, 304, 307

universelle Eigenschaft, 369

des äußeren Produkts, 359, 368

des Quotientenvektorraums, 123

des Tensorprodukts, 353, 366

Untergruppe, 47

- Unterraum
 - affiner, 116
 - invarianter, 254
- Unterring, 56
- Untervektorraum, 77
 - invarianter, 242
- Urbild, 35
- Ursprung, 2
- Vandermonde-Determinante, 196, 198
- Variable
 - freie, 24
 - gebundene, 24
- Vektor, 1, 76
 - liegender, 20
 - stehender, 20
- Vektorprodukt, 282, 285
 - verallgemeinertes, 285
- Vektorraum, 76
- verbindbar, 221
- Vereinigung, 33
- Verknüpfung, 43
 - äußere, 75
 - innere, 75
- Verknüpfungstafel, 46
- Vielfachheit, 66
- Vietascher Wurzelsatz, 70
- Vorzeichenregel, 72, 321
- windschief, 11, 346
- Winkel, 276, 279
- winkeltreu, 311
- wohldefiniert, 50, 123
- Wronski-Determinante, 199
- Zahlen
 - ganze, 33
 - komplexe, 33, 57
 - natürliche, 32
 - rationale, 33, 57
 - reelle, 33, 57
- Zeichenwechsel, 71
- Zeilenrang, 96, 99, 159, 335
- Zeilenraum, 94
- Zeilenstufenform, 22, 28, 94, 131
- Zeilenumformung, 26, 93, 165
- zusammenhängend, 221

Symbolverzeichnis

$\square$	Ende eines Beweises oder Beweis klar	$(\cdot, \cdot)$	Skalarprodukt, 274
$a := b$	a ist definiert durch b , 32	$\ \cdot \ $	Norm, 274
$a \Rightarrow b$	aus a folgt b	$\angle$	Winkel, 276
$a \Leftrightarrow b$	a und b sind gleichwertig	d	Abstand, 275
$\{ \}$	Mengenklammern	$\mathbb{C}$	komplexe Zahlen, 33, 57
$\emptyset$	leere Menge, 32	$\mathbb{K}$	$\mathbb{R}$ oder $\mathbb{C}$, 292
$\in$	Element, 32	$\mathbb{N}$	natürliche Zahlen, 32
$\subset$	Teilmenge, 32	$\mathbb{Q}$	rationale Zahlen, 33
$\cup$	Vereinigung, 33	$\mathbb{R}$	reelle Zahlen, 1, 33
$\cap$	Durchschnitt, 33	$\mathbb{R}_+$	nicht-negative reelle Zahlen, 35
$\setminus$	Differenzmenge, 34	$\mathbb{R}_+^*$	positive reelle Zahlen, 43
$\times$	direktes Produkt, 38 oder Vektorprodukt, 269	$\mathbb{R}^n$	reeller Standardraum, 1
$\rightarrow, \mapsto$	Abbildungspfeile, 34	$\mathbb{Z}$	ganze Zahlen, 33
$\circ$	Komposition von Abbildungen, 36	$\mathbb{Z}/m\mathbb{Z}$	zyklische Gruppe, 50, 51
$ $	Beschränkung von Abbildungen, 35	K^*	Elemente ungleich Null, 56
f^{-1}	Umkehrabbildung von f , 35	V^*	dualer Vektorraum, 331
$\sim$	äquivalent, 40	K^n	Standardraum, 75
$(x_i)_{i \in I}$	Familie, 39	$K[t]$	Polynomring über K , 62
$\sum$	Summenzeichen, 21	$\mathcal{K}$	kanonische Basis, 86
$\sum'$	eingeschränkte Summe, 352	e_i	kanonischer Basisvektor, 81
$\prod$	Produktzeichen, 190	δ_{ij}	Kronecker-Symbol, 93
$+$	Summe, 102	$\mathcal{C}$	stetige Funktionen, 78
$\oplus$	direkte Summe, 101	$\mathcal{D}$	differenzierbare Funktionen, 78
$\ominus$	orthogonale Summe, 295	A_n	alternierende Gruppe, 192
$\otimes$	Tensorprodukt, 353	S_n	symmetrische Gruppe, 44
$\bigotimes^k$	k -faches Tensorprodukt, 368	$M(m \times n; K)$	Matrizenraum, 75
$\wedge$	äußeres Produkt, 359	$GL(n; K)$	allgemeine lineare Gruppe, 150
$\bigwedge^k$	k -faches äußeres Produkt von V , 368	$O(n)$	orthogonale Gruppe, 304
$\vee$	symmetrisches Produkt, 364	$SO(n)$	spezielle orthogonale Gruppe, 304
$\bigvee^k$	k -faches symmetrisches Produkt, 369	$U(n)$	unitäre Gruppe, 304

A^{-1}	inverse Matrix, 150	Abb	Abbildungen, 34
tA	transponierte Matrix, 95	Alt	alternierende Abbildungen, 361
$A^\sharp$	komplementäre Matrix, 201		371
E_i^j	Basismatrix, 86	Bil	bilineare Abbildungen, 350, 361
E_n	n -reihige Einheitsmatrix, 92	char	Charakteristik, 61
M_B^A	darstellende Matrix, 139	deg	Grad, 61
M_B	darstellende Matrix, 140, 158, 288	det	Determinante, 176, 178, 192, 212
T_B^A	Transformationsmatrix, 155	dim	Dimension, 91
Q_i^j	Elementarmatrix, 163	Eig	Eigenraum, 226
$Q_i^j(\lambda)$	Elementarmatrix, 164	Hau	Hauptraum, 259
$S_i(\lambda)$	Elementarmatrix, 163	End	Endomorphismen, 112
P_i^j	Elementarmatrix, 164	Hom	Homomorphismen, 111
		Im	Bild, 114
Φ_B	Koordinatensystem, 138	Ker	Kern, 114
F_i^j	Basishomomorphismen, 139	Lös	Lösungsmenge, 20
F^{ad}	adjungierte Abbildung, 342	rang	Rang, 114, 159
$\mathcal{I}_F$	Ideal von F , 251	sign	Signum, 190
		span	aufgespannter Vektorraum, 80