Studienbücherei

J. Flachsmeyer,
L. Prohaska

Algebra

VEB Deutscher Verlag der Wissenschaften

Mathematik für Lehrer
Band 3

Herausgegeben von:

W. Engel, S. Brehmer, M. Schneider, H. Wussing

Unter Mitarbeit von:

G. Asser, J. Böhm, J. Flachsmeyer, G. Geise, T. Glocke,
K. Härtig, G. Kasdorf, O. Krötenheerdt, H. Lugowski,
P. H. Müller, G. Porath

Studienbücherei

Algebra

J. Flachsmeyer
L. Prohaska

Mit 20 Abbildungen

Mit historischen Bemerkungen
von H. Wussing

Vierte Auflage

VEB Deutscher Verlag
der Wissenschaften
Berlin 1980

Verlagslektor: Dipl.-Math. K. Bratz
Umschlaggestaltung: R. Wendt

Printed in the German Democratic Republic
Lizenz-Nr. 206·435/91/80
Satz: VEB Druckhaus „Maxim Gorki", Altenburg
Offsetnachdruck und buchbinderische Verarbeitung: Volksdruckerei Zwickau
LSV 1024
Bestellnummer 570 195 1
DDR 14,80 M

Vorwort

In den Bänden 1 und 2 dieser Lehrbuchreihe wurde der Leser noch einmal vom systematischen Standpunkt mit der Arithmetik — dem Zahlenrechnen — vertraut gemacht. Dabei standen in gleicher Weise die Objekte, mit denen gerechnet wird, die Zahlen, wie auch die vollzogenen Verknüpfungen, die Rechenoperationen, im Brennpunkt des Interesses.

Der Algebra geht es in Abstraktion von individuellen Rechenobjekten um eine Untersuchung von Rechenoperationen im allgemeinen, worunter sich dann die Rechenoperationen der Arithmetik als Spezialfälle einreihen. In diesem Sinne ist die gebräuchliche pauschale Äußerung zu verstehen, daß die Algebra eine höhere Arithmetik sei. Vom mengentheoretischen Standpunkt ist Algebra die Theorie der Mengen mit Operationen. Natürlich werden aus der denkbaren Fülle solcher strukturtheoretischen Betrachtungen vorwiegend jene interessieren, die das Ordnen der Wirklichkeit möglichst unterstützen und uns zu „lebensnahen" Erkenntnissen führen. Mit anderen Worten, man wird solche Mengen mit Operationen gesondert betrachten, die als Widerspiegelung vielfältiger Sachverhalte in Erscheinung treten. Die lange Entwicklung der Mathematik hat hinsichtlich der Algebra einen gewissen klassischen Bestand an solchen Modellierungen herausgeschält. Dazu gehören beispielsweise die Begriffe der Gruppe, des Ringes, des Körpers, des Vektorraumes (und damit zusammenhängende Begriffe wie Linearformen, lineare Abbildungen, Dualraum) und auch der Begriff des Verbandes. Diese Begriffe spielen in der Schulmathematik keine explizite Rolle, aber sie stecken immer wieder in den schulmathematischen Dingen. Die Ausbildung der Lehrer zielt darauf ab, den Lehrer dafür sehend zu machen, damit er von einem mehr überschauenden Standpunkt die zu lehrende Thematik übersieht. Sachverhalte, die vorher nur lose zusammenhängend oder gar voneinander isoliert gesehen wurden, können in Wahrheit sehr eng verwandt sein. Das Darlegen solcher natürlichen Verwandtschaften ist Sinn und Zweck der Abstraktion, des Absehens von überflüssigem, konkretem Beiwerk. Abstraktion ist ein wesentlicher Erkenntnisschritt. Eine zu weit getriebene Abstraktion ist hingegen ermüdend und abstoßend. Der mit einer Abstraktion gewon-

nene Denkvorteil darf nicht solange auf sich warten lassen. Hierin besteht die Kunst guten Lehrens.

In der Algebra herrscht ein hoher Abstraktionsgrad. Das macht erfahrungsgemäß anfänglich Schwierigkeiten. Wir haben uns bemüht, durch Anordnung und Darstellung diese Schwierigkeiten so zu halten, daß sie zu überwinden sind. Die Behandlung der sogenannten linearen Algebra hätte aus systematischen und logischen Gründen erst nach Kenntnis allgemeiner algebraischer Strukturen, wie Gruppen, Ringe und Körper erfolgen sollen. Aber von den Erfordernissen eines Lernenden, möglichst oft an einem natürlichen Heranwachsen der Begriffe teilnehmen zu können, empfiehlt sich wohl eine Hinlenkung zur allgemeinen Algebra durch eine Beschäftigung mit der klassischen Grundaufgabe der linearen Algebra, nämlich dem Auflösen von linearen Gleichungen mit reellen Koeffizienten. Demgemäß ist hier zunächst alles um diese Aufgabenstellung gruppiert. Der Leser wird damit an die Algebra herangeführt. Ohne eigene (intensive!) Arbeit wird es nicht abgehen. Von geometrischen Schulkenntnissen haben wir zur Illustration Gebrauch gemacht. Vom logischen Betrachtungspunkt wäre natürlich ein Auskommen ohne jegliche geometrische Vorleistung erstrebenswert. Die Aneignung mathematischer Kenntnisse darf sich aber nicht darauf beschränken, einen logisch reinen Aufbau überprüfend nachzuvollziehen, sondern sie muß genügend von der Erkenntniskomponente „Kunst des Findens" enthalten. Zu einem späteren Zeitpunkt wird der Leser hinsichtlich der Geometrie noch ausreichend mit der „Kunst des logischen Absicherns" in Berührung kommen.

Der erstgenannte Autor hat die Abschnitte 2 bis 10 verfaßt, der zweitgenannte Autor die anschließenden Abschnitte. Herr Kollege Wussing hat dem Ganzen den Abschnitt 1 vorangestellt. Wir danken ihm herzlich dafür. In der Folgezeit wird noch ein Ergänzungsband (Spezialstudium) über Algebra erscheinen.

Dem VEB Deutscher Verlag der Wissenschaften und den Herausgebern, insbesondere Herrn Prof. Dr. W. Engel, danken wir für die Möglichkeit, mit diesem Band einen Beitrag zur Studienbücherei „Mathematik für Lehrer" leisten zu können. Unser Dank gilt aber auch Herausgebern und Kollegen, die mit ihren kritischen Bemerkungen zum Manuskript unser gemeinsames Anliegen gefördert haben. Besondere Anerkennung verdient schließlich die sorgfältige Arbeit des VEB Druckhaus „Maxim Gorki" in Altenburg.

Zur weiteren Gestaltung und zukünftigen Verbesserung des Bandes wünschen wir uns eine möglichst vielfältige Reaktion der Leserschaft, insbesondere der Studenten, die danach arbeiten, der danach Lehrenden und der Lehrer, die zum Zwecke des Auffrischens ihrer Kenntnisse das Buch zur Hand nehmen. Mögen uns viele Zuschriften erreichen. Eine auf offizielle Rezensionen beschränkte Äußerung wäre uns nicht ausreichend.

Greifswald und Rostock, August 1973

Jürgen Flachsmeyer
Ludwig Prohaska

Inhalt

1. Bemerkungen zur Geschichte der Algebra

HANS WUSSING

I

Innerhalb der Mathematik der Neuzeit stellte die Algebra ein außerordentlich umfangreiches, in ihren Zielen weitgespanntes und in Methode und Inhalt weitreichendes mathematisches Gebiet dar. Auch heute befindet sie sich in rascher Entwicklung.

Die Anfänge der Algebra reichen weit zurück. Der Begriffsinhalt und damit ihre Zielstellung unterlagen dabei einer langen historischen Entwicklung. Etwas vereinfacht kann man die Geschichte der Algebra folgendermaßen periodisieren: Erste implizite algebraische Denk- und Arbeitsweisen treten uns bereits in den frühen Klassengesellschaften Chinas, Indiens, Ägyptens und Mesopotamiens entgegen; einige Nachwirkungen davon sind in der hellenistischen Antike nachweisbar und gewannen dort insbesondere bei DIOPHANTOS VON ALEXANDRIA (um 250 u. Z.) ansatzweise explizite Gestalt. Eine zweite Periode könnte man als Periode der Entwicklung der Algebra zur selbständigen mathematischen Disziplin bezeichnen. Sie reicht von der rechnerisch-algebraisch orientierten islamischen Mathematik des 9. und 10. Jahrhunderts über die Rechenmeister und Cossisten der europäischen Renaissance bis hin zum Ausgang des 15. Jahrhunderts. Während einer dritten Periode gewann die Algebra im 17. und 18. Jahrhundert den hauptsächlichen Begriffsinhalt als Kunst, Gleichungen aufzulösen; als natürlicher Abschluß dieser Periode können die endgültigen, lückenlosen Beweise des Fundamentalsatzes der Algebra durch CARL FRIEDRICH GAUSS (1777–1855) angesehen werden.

Mit GAUSS, ABEL, GALOIS und anderen beginnt sich, zunächst allerdings nur in impliziter Gestalt, eine neue Form algebraischen Denkens herauszubilden, die auf das Studium algebraischer Strukturen abzielt. Am Ende des 19. und am Anfang des 20. Jahrhunderts vollzog sich auch äußerlich eine radikale Wendung von einer als Gleichungstheorie verstandenen Algebra zur Algebra als Disziplin, welche die Erforschung (algebraischer) Strukturen zum Gegenstand hat.

II

Die altägyptische Mathematik im 2. Jahrtausend v. u. Z. besaß insofern Ansatzpunkte zur Algebra, als sie imstande war, lineare Gleichungen mit einer Variablen exakt zu behandeln und für die Variable einen feststehenden Terminus besaß, nämlich das Schriftzeichen *hau*, das *Haufen* oder *Menge* bedeutet, als Symbol für die zu bestimmende Größe oder Menge.

Im engeren Sinne echt algebraische Ansätze finden sich in der hochentwickelten Rechentechnik der sogenannten babylonischen Mathematik. Die Analyse jener Rechengänge auf mesopotamischen Keilschrifttafeln aus dem 1. Jahrtausend v. u. Z. zeigt bei außerordentlich komplizierten Aufgaben eine erstaunliche Geschicklichkeit im Umformen von Gleichungen. Es werden u. a. zweckmäßige Hilfsgrößen eingeführt; wenn mehrere Variable auftreten, werden Größen eliminiert. Es zeigt sich weiter, daß die Rechnungen im Grunde so verlaufen, wie wir heute an solche Rechenaufgaben herangehen würden (vgl. [7, S. 46]).

In diesen Zusammenhängen vollzog sich der Übergang zu einer zwar unvollständigen, aber doch durch Konvention fixierten „Formel"-Schreibweise und die Ausbildung einer Art Fachterminologie. Zum Beispiel repräsentierten die Worte *tab* bzw. *lal* die Operationen der Addition bzw. Subtraktion, die zugleich die Rolle von „Vorzeichen" spielten. Es gab ein inneres Lückenzeichen nach Art einer Null im sexagesimalen Zahlensystem, und es gab in der Geometrie einen feststehenden Terminus als Zeichen der Gleichheit zweier Seiten. Schließlich verdichteten sich im Laufe der Zeit die Texte zu einer ideographischen Kurzschreibweise von algebraischem Charakter.

Die griechisch-hellenistische Mathematik nahm etwa seit dem Ende des 4. Jahrhunderts v. u. Z. den Charakter einer geometrischen Algebra an (vgl. dazu [1, 7]). In Anbetracht der unbewältigten Probleme des Umgangs mit dem Unendlichen und insbesondere der Problematik der inkommensurablen Größen (u. a. der irrationalen Zahlen) wurden algebraische Probleme mit Hilfe geometrischer Konstruktionen behandelt. Beispielsweise wurden quadratische Gleichungen mittels der Methode der Flächenanlegung gelöst. Erst gegen Ende der Antike traten auch Elemente einer „algebraischen Algebra" wieder hervor, insbesondere bei DIOPHANTOS VON ALEXANDRIA. Sein Hauptwerk, die *Arithmetik*, verwendete feste Symbole für die gesuchte Zahl (vermutlich einen abgewandelten Buchstaben des griechischen Alphabets) und für deren k-te Potenzen, $k = \pm 1, \pm 2, \ldots, \pm 6$. Für die Subtraktion hatte DIOPHANTOS ebenfalls ein festes Zeichen, eine Art umgekehrtes ψ; für die Gleichheit verwendete er den Buchstaben ι, den ersten des griechischen Wortes $\iota\sigma o\iota$, das *gleich* bedeutet.

III

Die Mathematik der Länder des Islam (d. h. die sogenannte arabische Mathematik) nahm wesentliche Teile der weiterentwickelten indischen sowie der hellenistischen Mathematik in sich auf und erreichte, insbesondere auch bei der Fortführung der rechnerisch-algebraischen Ansätze der indischen Mathematik, eine bedeutende Höhe.

Das Wort *Algebra* geht zurück auf den Titel eines Buches des aus Choresm (heute Chiwa) stammenden AL-HWÂRAZMÎ (780?–850?). Er lautet *Hisâb aljabr w'almuqâbalah* und bedeutet soviel wie „Buch von der Ergänzung (*aljabr*) und der Ausgleichung (*almuqâbalah*)". An Hand von Musterbeispielen erläutert der Autor dort die Verfahren zur Auflösung von Gleichungen. Beispielsweise wird in einer Gleichung, die wir als $13x - 5 = 7x + 4$ schreiben würden, die linke Seite um 5 „ergänzt", weil dort ein negativer Term vorkommt; diese 5 muß dann rechts ebenfalls „ergänzt" werden. Dann folgt ein zweiter Schritt, die „Ausgleichung" der mit x behafteten Glieder in der Gleichung $13x = 7x + 9$. Es ergibt sich $6x = 9$ und damit die Lösung

$$x = \frac{3}{2}.$$

Dieses schrittmachende Buch von AL-HWÂRAZMÎ erlangte eine große Verbreitung und wurde vom 12. Jahrhundert an auch in Europa bekannt. Der Buchtitel, insbesondere das erste Wort *aljabr*, wurde allmählich zum Synonym für die dort verwendeten Methoden der Auflösung von Gleichungen; so entstand das Fachwort *Algebra* durch Latinisierung des arabischen Wortes. Jedoch schufen erst die Entwicklung des Frühkapitalismus und der damit verbundene Übergang von der Natural- zur Geldwirtschaft das allgemeine gesellschaftliche Bedürfnis zur umfassenden Anwendung von Rechenmethoden. Den Rechenmeistern des 15. und 16. Jahrhunderts — in Deutschland wurde ADAM RIES (1492–1559) am bekanntesten — dankt man u. a. die Einführung der indisch-arabischen Ziffern, das damals als sehr schwierig empfundene Rechnen mit Brüchen, die kalkülmäßige Durcharbeitung der vier Grundrechenarten und des Radizierens und die Einführung erster Symbole für die Rechenoperationen und Unbekannten. Beispielsweise verwendete LUCA PACIOLI (1445–1515) die Zeichen $\bar{p}$ und $\bar{m}$ für plus und minus. Im Rechenbuch des JOHANN WIDMANN (geb. um 1460) aus dem Jahre 1489 traten, vermutlich zum erstenmal, die Zeichen + und − im Druck auf. Der englische Arzt R. RECORDE (1510?–1558) schlug 1557 den Gebrauch des heutigen Gleichheitszeichens vor, es setzte sich jedoch erst im 17. Jahrhundert durch.

Unter der Bezeichnung *Coß* oder *cossische Kunst* — welche auf die italienische Benennung *cosa* (*Sache*) für die Variable in Gleichungen zurückgeht — erreichte die Verwendung von Symbolen und Abkürzungen schon im 16. Jahrhundert einen recht hohen Stand, in Deutschland durch das Wirken von CHR. RUDOLFF (1500?–1545?) und MICHAEL STIFEL (1486–1567), in Frankreich durch N. CHUQUET († um 1500),

in Italien durch RAFAEL BOMBELLI (16. Jahrhundert), HIERONIMO CARDANO (1501 bis 1576) und andere.

Die endgültige Herausbildung der Algebra vollzog sich erst am Ausgang des 16. und zu Anfang des 17. Jahrhunderts. Insbesondere schuf FRANÇOIS VIETA (1540 bis 1603) mit seiner — natürlich an Vorgängern orientierten — durchgebildeten Buchstabenalgebra, die er im Unterschied zum Rechnen mit konkreten Zahlen als *logistica speciosa* (etwa: prachtvolle Rechenkunst) bezeichnete, ein einheitliches algebraisches Bezeichnungssystem. Durchgehend hat VIETA die Variablen (Unbekannten) durch die Vokale A, E, I, O, U, Y und die bekannten Größen durch die Konsonanten B, C, D, ... bezeichnet; er verwendete u. a. die Zeichen + und −, geschweifte Klammern, den Bruchstrich als Zeichen der Division, das Wörtchen *in* als feststehendes Kurzzeichen der Multiplikation. Die Gleichheit zweier Terme drückte VIETA noch durch die Worte *aequibitur* oder *aequale* aus. (Die heute fast allgemein gewordene Verwendung der letzten Buchstaben x, y, z des Alphabets für die Variablen geht auf RENÉ DESCARTES (1596—1650) zurück.) Insgesamt trat durch VIETA die Algebra als neuer Zweig der Mathematik gleichberechtigt neben die Geometrie, die bis dahin weitgehend mit Mathematik überhaupt identisch gewesen war, hatte es bisher doch nur dort echte mathematische Sätze und Beweise gegeben.

Zu Beginn des 16. Jahrhunderts traten in Verbindung mit den sich rasch entfaltenden frühkapitalistischen Produktionsverhältnissen auch bedeutende mathematische Leistungen hervor, darunter insbesondere bei der Behandlung algebraischer Gleichungen. So fand um 1500 SCIPIO DEL FERRO (1465?—1526) die algorithmische Auflösung der kubischen Gleichung; doch blieb sie unpubliziert. Unabhängig davon gelangte der Rechenmeister NICCOLÒ TARTAGLIA (1500?—1557) im Jahre 1535 zu eben derselben Lösung und teilte sie 1539 unter dem Siegel der Verschwiegenheit dem Universitätsprofessor CARDANO aus Milano mit, der jedoch eidbrüchig TARTAGLIAS Methode — zusammen mit der von seinem Schüler LUDOVICO FERRARI (1522—1565) herrührenden Lösungsmethode für die Gleichung vierten Grades — in seiner *Ars magna* von 1545 veröffentlichte.

Für die allgemeinen algebraischen Gleichungen höheren als vierten Grades suchten die Mathematiker der nachfolgenden Generation mit außergewöhnlicher Anstrengung nach analogen Lösungsverfahren, d. h. nach Lösungen, die sich durch Wurzelschachtelungen darstellen lassen. Vorübergehend glaubte EHRENFRIED WALTER VON TSCHIRNHAUS (1651—1708) durch geeignete Variablensubstitution (Tschirnhaus-Transformation) die Gleichungen aller Grade so umformen zu können, daß eine Auflösung möglich sein müsse, doch diese und weitere Hoffnungen zerschlugen sich. Am Ende des 18. Jahrhunderts wurde es zweifelhaft, ob eine derartige Lösung des Problems mit den bisherigen Methoden überhaupt möglich sein werde; JOSEPH LOUIS LAGRANGE (1736—1813) sprach dies 1770/71 als erster aus. Der Italiener PAOLO RUFFINI (1765—1822) konnte von 1799 an Schritt für Schritt einen im wesentlichen vollständigen Beweis dafür erbringen, daß die allgemeine Gleichung fünften Grades nicht in Radikalen lösbar ist. Unabhängig von RUFFINI gelangte 1824 der

Norweger NIELS HENRIK ABEL (1802—1829) zum gleichen Ergebnis und konnte zwei Jahre später den Beweis liefern, daß die allgemeine Gleichung höheren als vierten Grades nicht durch Radikale lösbar ist.

Die Frage nach der Existenz der Wurzeln algebraischer Gleichungen war schon seit der Renaissance bewußt gestellt worden und hatte überdies in der Folgezeit philosophisch und mathematisch schwierige Fragen des Umgangs mit imaginären (wörtlich: eingebildeten!) und komplexen Zahlen aufgeworfen (vgl. [10, 13]). ALBERT GIRARD (1595—1632) formulierte 1629 unter Einbeziehung komplexer Zahlen den Satz, daß jede algebraische Gleichung n-ten Grades genau n Wurzeln besitzt. Doch erst GAUSS konnte, nachdem 1746 bereits JEAN BAPTISTE LE ROND D'ALEMBERT (1717—1783) Wesentliches beigetragen hatte, im Jahre 1796 seinen ersten lückenlosen Beweis des Fundamentalsatzes der Algebra geben.

Auch andere Elemente der klassischen Algebra reichen weit in die Vergangenheit zurück und fanden während der Neuzeit ihre Durchbildung. Beispielsweise war die Behandlung linearer Gleichungssysteme schon im mittelalterlichen China hochentwickelt; die in der sogenannten *fang-cheng-Methode* verwendeten Ideen stehen dem Gebrauch von Determinanten und Matrizen sehr nahe. Bei GOTTFRIED WILHELM LEIBNIZ (1646—1716) scheint sich zum erstenmal eine allgemeingültige Definition der Determinanten vorgefunden zu haben, die indessen in Vergessenheit geriet. So begann eine moderne Theorie der Determinanten 1750 mit dem Schweizer Mathematiker GABRIEL CRAMER (1704—1752); der Ausbau erfolgte u. a. durch AUGUSTIN-LOUIS CAUCHY (1789—1857) und CARL GUSTAV JACOB JACOBI (1804—1851). Die Präzision der Lösungsverhältnisse bei m linearen Gleichungen in n Variablen — wozu auch die entscheidende Begriffsbildung des Ranges einer Matrix gehört — verdankt man der britischen algebraischen Schule um ARTHUR CAYLEY (1821—1895) und JAMES JOSEPH SYLVESTER (1814—1897) sowie LEOPOLD KRONECKER (1823—1891) und GEORG FROBENIUS (1849—1917).

Mit der Entwicklung kapitalistischer Produktionsverhältnisse, insbesondere seit der industriellen Revolution, ergingen eine Vielzahl von direkten und indirekten Impulsen an die Entwicklung der Naturwissenschaften und der Mathematik. Unter anderem nahm auch die Algebra während des 19. Jahrhunderts einen raschen Aufschwung. Noch innerhalb einiger Bestandteile der klassischen Algebra, insbesondere der Auflösungstheorie algebraischer Gleichungen, entwickelten sich Elemente der künftigen Strukturalgebra. Beispielsweise ist die Theorie der sogenannten Gaußschen Perioden identisch mit der Untersuchung der Untergruppen der Galoisschen Gruppe der Kreisteilungsgleichung; hieraus entnahmen ABEL und EVARISTE GALOIS (1811 bis 1832) wesentliche Anregungen. Durch ABEL wurde der Begriff *Gruppe* zu einem mathematischen Fachausdruck; er untersuchte endliche kommutative Permutationsgruppen bei der Frage der Bestimmung aller in Radikalen auflösbaren algebraischen Gleichungen. GALOIS erkannte die Bedeutung der Normalteiler und konnte jeder algebraischen Gleichung eine (Permutations-)Gruppe zuordnen, aus deren Struktur auf die Lösungsverhältnisse der Gleichung geschlossen werden kann, insbesondere

wird entscheidbar, ob sie in Radikalen auflösbar ist. KRONECKER stellte 1870 ein erstes explizites Axiomensystem für eine endliche, kommutative Gruppe auf; im selben Jahr erschien die bedeutende Monographie *Traité des substitutions et des équations algébriques* von CAMILLE JORDAN (1838—1922). Am Ende des 19. Jahrhunderts hatte sich der Begriff der Gruppe — entstanden aus zunächst impliziten gruppentheoretischen Denkformen in Zahlentheorie, Geometrie und Auflösungstheorie algebraischer Gleichungen — als erster abstrakter algebraischer Strukturbegriff herausgebildet.

Ähnlich entwickelten sich, in Anlehnung an das methodische Vorbild der Gruppentheorie, die abstrakte Körpertheorie, die Idealtheorie und die Theorie der hyperkomplexen Systeme. Dabei wurden bei Studien über mögliche Axiomensysteme auch nichtkommutative Verknüpfungen innerhalb einer Menge mit doppelter Verknüpfung analysiert. Auf diesen Gebieten vollbrachten u. a. Sir WILLIAM ROWAN HAMILTON (1805—1865), ERNST EDUARD KUMMER (1810—1893), GEORGE BOOLE (1815—1864), KRONECKER und RICHARD DEDEKIND (1831—1916) bedeutende Forschungsleistungen. Die ausführliche, immer wieder nachgedruckte Lehrbuchdarstellung der klassischen Algebra durch HEINRICH WEBER (1842—1913) von 1895/96 stand geradezu symbolisch am Ende dieses Abschnittes der Entwicklung der Algebra.

Auf der Grundlage der mengentheoretischen Durchdringung der gesamten Mathematik und unter der Wirkung des sogenannten „Zahlberichtes" (1897) von DAVID HILBERT (1862—1943) sowie der Theorie der algebraischen Körper (1910) von ERNST STEINITZ (1871—1928) kam es zu Anfang der zwanziger Jahre unseres Jahrhunderts zu einem qualitativen Umschwung innerhalb der Algebra.

Die hochbedeutende Mathematikerin EMMY NOETHER (1882—1935), die 1933 von den Faschisten aus Göttingen vertrieben wurde, EMIL ARTIN (1898—1962) und deren Schüler wie HELMUT HASSE (* 1898), O. SCHREIER (1901—1929), W. KRULL (1899—1971) erhoben die Untersuchung der algebraischen Strukturen in ihrer abstrakten Form, also losgelöst von einer Repräsentation durch „konkrete" mathematische Objekte, zum Gegenstand der Algebra. Wegbereitend wirkte hier das Lehrbuch *Moderne Algebra* (1930/31) von BARTEL LEENDERT VAN DER WAERDEN (* 1903).

Obwohl die Wissenschaft der Gegenwart in den sozialistischen und kapitalistischen Staaten gänzlich anderen gesellschaftlichen Interessen dient, ist sie doch allgemein in rascher Entwicklung begriffen. Die Sowjetunion und die USA besitzen gegenwärtig die stärksten und erfolgreichsten mathematischen Zentren, darunter auch hervorragende algebraische Schulen. Hier, aber auch anderswo, entwickelten sich in Wechselwirkung mit den Anwendungsgebieten der Algebra (Topologie, Funktionalanalysis, algebraische Geometrie, theoretische Physik, Computertechnik) rasch neue Teilgebiete der Algebra, darunter die Theorie der Kategorien, die Verbandstheorie und die Theorie der Halbgruppen. Für diese umfassende, vom strukturellen Denken geprägte und auf die Herausarbeitung des Allgemeinen, des vom algebraischen Standpunkt Wesentlichen abzielende Algebra scheint sich in jüngster Zeit eine neue Bezeichnung, *Allgemeine Algebra*, einzubürgern.

Ausgewählte Literaturangaben

[1] Zeuthen, H. G.: Geschichte der Mathematik im Altertum und Mittelalter (dänisch), Kopenhagen 1896. (Deutsche Übersetzung: Kopenhagen 1896; russische Übersetzung: Moskau—Leningrad 1938).
[2] Tropfke, J.: Geschichte der Elementarmathematik, Band II, 2. Aufl., Berlin und Leipzig 1921.
[3] Cajori, F.: History of Mathematical Notations, Band I und II, Chicago 1928/29.
[4] Neugebauer, O.: Vorlesungen über Geschichte der antiken mathematischen Wissenschaften, 1. Band: Vorgriechische Mathematik, Berlin 1934.
[5] Kuroš, A. G.: Vorlesungen über allgemeine Algebra (Übersetzung aus dem Russischen), Leipzig 1964.
[6] Juschkewitsch, A. P.: Mathematik im Mittelalter, Leipzig 1964 (Übersetzung aus dem Russischen).
[7] Wussing, H.: Mathematik in der Antike, 2. Aufl., Leipzig 1965.
[8] Wussing, H.: Die Genesis des abstrakten Gruppenbegriffes, Berlin 1969.
[9] Struik, D. J.: Abriß der Geschichte der Mathematik, 6. Aufl., Berlin 1976 (Übersetzung aus dem Amerikanischen).
[10] Gericke, H.: Geschichte des Zahlbegriffes, Mannheim—Wien—Zürich 1970.
[11] Nový, L.: Origins of Modern Algebra, Prague 1973.
[12] Vieta, F.: Einführung in die neue Algebra, Übersetzt und erläutert von K. Reich und H. Gericke, München 1973.
[13] Wussing, H.: Zur Geschichte der Zahlzeichen und des Zahlbegriffes. In: Wisliceny, J.: Grundbegriffe der Mathematik II, Studienbücherei MfL, Band 2, 2. Aufl.. Berlin 1975.

2. Der n-dimensionale reelle Zahlenraum

2.1. Abstrakte Erklärung des $\mathbb{R}^n$

Definition 1 (n-dimensionaler reeller Zahlenraum, $\mathbb{R}^n$). Unter dem *n-dimensionalen reellen Zahlenraum* — dem $\mathbb{R}^n$ — versteht man das n-fache kartesische Mengenprodukt der Menge der reellen Zahlen $\mathbb{R}$ mit sich selbst:

$$\mathbb{R}^n = \underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{n\text{-mal}} = \{(x_1, x_2, \ldots, x_n) : x_i \in \mathbb{R},\ i = 1, \ldots, n\}.$$

Die Elemente von $\mathbb{R}^n$ (auch Punkte des $\mathbb{R}^n$ genannt) sind also (geordnete) n-Tupel reeller Zahlen

$$\boldsymbol{x} = (x_1, x_2, \ldots, x_n).$$

Die das n-Tupel $\boldsymbol{x}$ konstituierenden n reellen Zahlen $x_1, x_2, \ldots, x_n$ heißen die *Komponenten* oder auch *Koordinaten* von $\boldsymbol{x}$, und x_i ist die i-te Komponente bzw. i-te Koordinate von $\boldsymbol{x}$.

Unter der *Gleichheit von zwei n-Tupeln* $\boldsymbol{x}$ und $\boldsymbol{y}$ wird gemäß der Produktmengenerklärung die koordinatenweise Gleichheit verstanden. Zwei Elemente $\boldsymbol{x}$, $\boldsymbol{y}$ des $\mathbb{R}^n$ sind also genau dann voneinander verschieden, wenn sie sich wenigstens in einer Komponente unterscheiden.

2.2. Veranschaulichung des $\mathbb{R}^n$

1. Eine Art der Veranschaulichung des $\mathbb{R}^n$ für die Fälle $n = 1, 2, 3$ besteht in der Deutung der Elemente des $\mathbb{R}^n$ als Punkte einer „Zahlengeraden" ($\mathbb{R}^1 (= \mathbb{R})$), einer „Zahlenebene" ($\mathbb{R}^2$) bzw. eines „Zahlenraumes" ($\mathbb{R}^3$). Sie wird euklidische Veranschaulichung genannt. Abbildung 1 verdeutlicht die genannten Fälle.
2. Eine andere Art der Veranschaulichung des $\mathbb{R}^n$, n beliebige natürliche Zahl, besteht in der Deutung der Elemente des $\mathbb{R}^n$ als Abbildung einer n-elementigen Menge, etwa der Menge $\{1, 2, \ldots, n\}$, in die Menge $\mathbb{R}$ der reellen Zahlen. Abbildung 2

zeigt die Veranschaulichung von x als Funktion x: $\{1, 2, \ldots, n\} \to R$ $(x = (x_1, x_2, \ldots, x_n))$. Diese Veranschaulichung nennen wir die fasernweise oder schichtenweise Veranschaulichung des R^n. Der R^n wird aus n Schichten, jeweils bestehend aus R, aufgebaut gedacht. Ein Punkt des R^n, kann dann als ein „Faden" aufgefaßt werden, der sich durch die Schichten an den Stellen x_1 bzw. x_2 ... bzw. x_n zieht.

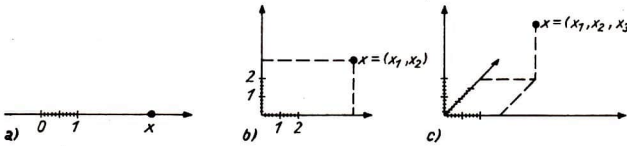


Abb. 1

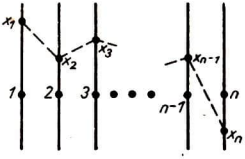


Abb. 2

2.3. Arithmetische Struktur des R^n

Mit den reellen Zahlen, mit den Elementen des R^1, kann man rechnen. Das Rechnen überträgt sich in bestimmter Weise auch auf die Elemente des R^n. Solch eine Situation ist für $n = 2$ schon im Zusammenhang mit den komplexen Zahlen aufgetaucht. Die dort interessierende Addition war die koordinatenweise Addition während die dortige Multiplikation sehr wohl von der koordinatenweisen Multiplikation verschieden ist. Die zu erklärenden Operationen hängen ganz von dem verfolgten Zweck ab! In der linearen Algebra benötigt man in erster Linie eine Addition von Elementen des R^n und eine Multiplikation eines Elementes des R^n mit einer reellen Zahl.

Definition 1 (Koordinatenweise Addition und Multiplikation mit einem Skalar im R^n). Es seien $x, y \in R^n$ mit $x = (x_1, x_2, \ldots, x_n)$, $y = (y_1, y_2, \ldots, y_n)$. Ferner sei $\alpha \in R$. Unter der ***koordinatenweisen Addition*** im R^n versteht man die folgende Operation:

$$x + y := (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n),$$

d. h., die i-ten Koordinaten von $\boldsymbol{x}$ und von $\boldsymbol{y}$ werden jeweils addiert. Unter der *koordinatenweisen Multiplikation* mit *einem Skalar* im $\mathbf{R}^n$ versteht man die folgende Operation:

$$\alpha \cdot \boldsymbol{x} := (\alpha x_1, \alpha x_2, \ldots, \alpha x_n),$$

d. h., die i-te Koordinate von $\boldsymbol{x}$ wird jeweils mit α multipliziert. Anstelle von $\alpha \cdot \boldsymbol{x}$ schreibt man vielfach auch nur $\alpha \boldsymbol{x}$.

Bemerkungen.

1. Die koordinatenweise Addition im $\mathbf{R}^n$ ist also eine Abbildung von $\mathbf{R}^n \times \mathbf{R}^n$ in den $\mathbf{R}^n$; jedem Paar von reellen n-Tupeln wird wieder ein reelles n-Tupel zugeordnet:

$$+ : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}^n \quad \text{vermöge} \quad (\boldsymbol{x}, \boldsymbol{y}) \to \boldsymbol{x} + \boldsymbol{y} \quad \text{für} \quad \boldsymbol{x}, \boldsymbol{y} \in \mathbf{R}^n.$$

2. Die koordinatenweise Multiplikation mit einem Skalar im $\mathbf{R}^n$ ist also eine Abbildung von $\mathbf{R} \times \mathbf{R}^n$ in den $\mathbf{R}^n$; jedem Paar aus einer reellen Zahl und einem n-Tupel reeller Zahlen wird wieder ein reelles n-Tupel zugeordnet:

$$\cdot : \mathbf{R} \times \mathbf{R}^n \to \mathbf{R}^n \quad \text{vermöge} \quad (\alpha, \boldsymbol{x}) \to \alpha \boldsymbol{x} \quad \text{für} \quad \alpha \in \mathbf{R}, \boldsymbol{x} \in \mathbf{R}^n.$$

Einem Sprachgebrauch der Physiker folgend, hat sich in diesem Zusammenhang anstelle der Sprechweise von der Multiplikation mit einer reellen Zahl die Bezeichnung von der Multiplikation mit einem Skalar eingebürgert.

3. Die beiden genannten Operationen verfolge man zweckmäßig unter Benutzung der Veranschaulichungen des $\mathbf{R}^n$. Die Addition wird in der ersten Veranschaulichung als Addition „nach dem Kräfteparallelogramm" wiedergegeben, im zweiten Falle als übliche „Funktionenaddition" oder „Superposition von Funktionen", wie man auch sagt. Die Multiplikation mit einem Skalar bedeutet im ersten Fall eine radiale Verschiebung des Punktes, im zweiten Fall eine Verzerrung der Funktion, hinzu kommt in beiden Fällen bei $\alpha < 0$ noch eine Spiegelung.

Die soeben betrachteten Operationen der Addition und der Multiplikation mit einem Skalar genügen gewissen Gesetzen. Die grundlegenden „Rechenregeln" für diese Operationen sind nachfolgend zusammengestellt.

Satz 1 (Grundeigenschaften der koordinatenweisen Addition und der Multiplikation mit einem Skalar im $\mathbf{R}^n$). *Die im $\mathbf{R}^n$ erklärte binäre Operation der koordinatenweisen Addition und die Multiplikation mit einem reellen Skalar haben die folgenden Grundeigenschaften:*

I. *Für die Addition $+$ gilt:*

1. $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{y} + \boldsymbol{x}$ *für alle* $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{R}^n$ (*Kommutativität der Addition*).
2. $(\boldsymbol{x} + \boldsymbol{y}) + \boldsymbol{z} = \boldsymbol{x} + (\boldsymbol{y} + \boldsymbol{z})$ *für alle* $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbf{R}^n$ (*Assoziativität der Addition*).
3. *Es gibt ein* (*eindeutig bestimmtes*) *Element* $\mathbf{0} \in \mathbf{R}^n$, *so daß* $\boldsymbol{x} + \mathbf{0} = \boldsymbol{x}$ *für alle* $\boldsymbol{x} \in \mathbf{R}^n$ *ist* (*Existenz und Einzigkeit des Nullelements*).

4. *Zu jedem* $\boldsymbol{x} \in \mathbb{R}^n$ *existiert ein (eindeutig bestimmtes) Element, bezeichnet mit* $-\boldsymbol{x}$, *für welches* $\boldsymbol{x} + (-\boldsymbol{x}) = \boldsymbol{0}$ *ist (Existenz und Einzigkeit des Inversen).*

II. *Für die Multiplikation mit einem reellen Skalar gilt:*

5. $1 \cdot \boldsymbol{x} = \boldsymbol{x}$ *für alle* $\boldsymbol{x} \in \mathbb{R}^n$.

6. $(\alpha \cdot \beta)\boldsymbol{x} = \alpha(\beta\boldsymbol{x})$ *für alle* $\alpha, \beta \in \mathbb{R}$ *und alle* $\boldsymbol{x} \in \mathbb{R}^n$ *(Assoziativität der Multiplikation mit einem Skalar).*

III. *Für das Zusammenspiel der koordinatenweisen Addition* $+$ *mit der Multiplikation mit einem reellen Skalar gilt:*

7. $(\alpha + \beta)\boldsymbol{x} = \alpha\boldsymbol{x} + \beta\boldsymbol{x}$ *für alle* $\alpha, \beta \in \mathbb{R}$ und alle $\boldsymbol{x} \in \mathbb{R}^n$ *(Distributivität der Multiplikation mit einem Skalar bezüglich der Addition von Skalaren).*

8. $\alpha(\boldsymbol{x} + \boldsymbol{y}) = \alpha\boldsymbol{x} + \alpha\boldsymbol{y}$ *für alle* $\alpha \in \mathbb{R}$ *und alle* $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ *(Distributivität der Multiplikation mit einem Skalar bezüglich der koordinatenweisen Addition).*

Bemerkungen.

1. Die in der Gruppe I zusammengefaßten Rechenregeln für die Operation der koordinatenweisen Addition sind uns schon mehrfach für andere Operationen begegnet. Und zwar gehorcht die Addition im Bereich der ganzen Zahlen, im Bereich der rationalen Zahlen, im Bereich der reellen Zahlen, im Bereich der komplexen Zahlen und die Multiplikation im Bereich der von Null verschiedenen rationalen Zahlen, im Bereich der von Null verschiedenen reellen Zahlen, im Bereich der von Null verschiedenen komplexen Zahlen formal den gleichen Grundregeln. Hier bietet sich also im Sinne einer Denkökonomie eine Heraushebung dieses Umstandes in Form eines eigenständigen Begriffes an, zumal wir noch weitere Bereiche kennenlernen werden, die gleichartig strukturiert sind. Es hat sich für den einschlägigen Begriff die Bezeichnungsweise *abelsche Gruppe* eingebürgert. Eine systematische Analyse dieses Begriffes erfolgt in der Gruppentheorie (vgl. hierzu Kap. 12).

Wir sagen kurz: Der $\mathbb{R}^n$ bildet bezüglich der koordinatenweisen Addition eine Gruppe.

2. Die aufgeführten Grundregeln sind natürlich nicht die einzigen geltenden Regeln hinsichtlich der Addition und der Multiplikation mit Skalaren im $\mathbb{R}^n$. Beispielsweise nennen wir noch folgende:

(i) Zu je zwei Elementen $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ gibt es ein $\boldsymbol{z} \in \mathbb{R}^n$ mit $\boldsymbol{x} + \boldsymbol{z} = \boldsymbol{y}$.

(ii) Für jedes Element $\boldsymbol{x} \in \mathbb{R}^n$ gilt stets $0 \cdot \boldsymbol{x} = \boldsymbol{0}$.

Noch viele andere mehr wären zu nennen. Aber diese Regeln kann man alle aus den ausgewählten Grundregeln ableiten, ohne auf die konkrete Definition zurückzugreifen, wie koordinatenweise addiert bzw. mit einem Skalar multipliziert wird. Diese Tatsachen haben sich im Laufe der Mathematikentwicklung herausgestellt und zu dem sogenannten axiomatischen Verfahren bzw. zum deduktiven Aufbau in der Mathematik geführt. Was dabei jeweils als Grundregel ausgewählt wird, unterliegt in einem bestimmten Maße dem Belieben. Konkrete Angaben hierzu findet man etwa in der Gruppentheorie.

3. Nach dem Assoziativgesetz der Addition im $\mathbf{R}^n$ ist bei drei Summanden die Art der Klammerung ohne Einfluß auf das Ergebnis, es kann daher eine Klammerung überhaupt unterbleiben, genauso wie das auch schon beim Rechnen mit Zahlen erfolgte. Entsprechend hierzu hat dann auch eine k-gliedrige Summe im $\mathbf{R}^n$ einen wohlbestimmten Sinn:

$$\sum_{i=1}^{k} \boldsymbol{x}_i, \qquad \boldsymbol{x}_i \in \mathbf{R}^n.$$

Die binäre Operation der koordinatenweisen Addition $+$ im $\mathbf{R}^n$ ist zu einer k-nären Operation im $\mathbf{R}^n$ ausgedehnt worden:

$$\sum_{i=1}^{k} : \underbrace{\mathbf{R}^n \times \mathbf{R}^n \times \cdots \times \mathbf{R}^n}_{k\text{-mal}} \to \mathbf{R}^n.$$

4. Anstelle von $\boldsymbol{x} + (-\boldsymbol{y})$ schreibt man kürzer $\boldsymbol{x} - \boldsymbol{y}$, genauso wie das auch für das Zahlenrechnen üblich ist.

Beweis des Satzes.

Zu I. 1. Es sei $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{R}^n$ mit $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$, $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$. Es ist

$$\boldsymbol{x} + \boldsymbol{y} = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$$

und

$$\boldsymbol{y} + \boldsymbol{x} = (y_1 + x_1, y_2 + x_2, \ldots, y_n + x_n).$$

Infolge der Kommutativität der Addition der reellen Zahlen haben wir $x_i + y_i = y_i + x_i$ für alle $i \in \{1, 2, \ldots, n\}$. Also stimmen $\boldsymbol{x} + \boldsymbol{y}$ und $\boldsymbol{y} + \boldsymbol{x}$ koordinatenweise überein, d. h. aber gerade $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{y} + \boldsymbol{x}$.

Die anderen Eigenschaften werden ganz analog bewiesen. Wir vermerken lediglich noch folgendes:

Zu I. 3. Es ist $\mathbf{0} = (0, 0, \ldots, 0)$ — das n-Tupel reeller Zahlen, wo jede Koordinate 0 ist.

Zu I. 4. Es ist $-\boldsymbol{x} = (-x_1, -x_2, \ldots, -x_n)$, sofern $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ ist.

2.4. Übungsaufgaben

1. Es bezeichne $\mathscr{P}$ die Menge aller reellen Polynome. Es handelt sich also um die folgende Menge:

$\mathscr{P} := \{P : P : \mathbf{R} \to \mathbf{R}$ mit einer Darstellung $P(x) = \sum_{i=0}^{n} a_i x^i$ für ein gewisses $n \in \mathbf{N}$ und gewisse $a_i \in \mathbf{R}\}$.

In $\mathscr{P}$ betrachte man die übliche Addition und Multiplikation mit reellen Skalaren:

$$+ : \mathscr{P} \times \mathscr{P} \to \mathscr{P} \qquad (P + Q)(x) := P(x) + Q(x), \qquad P, Q \in \mathscr{P};$$
$$\cdot : \mathbf{R} \times \mathscr{P} \to \mathscr{P} \qquad (\alpha P)(x) := \alpha P(x), \qquad \alpha \in \mathbf{R},\ P \in \mathscr{P}.$$

Man prüfe, welche der im Satz über die Grundeigenschaften der koordinatenweisen Addition und der Multiplikation mit einem reellen Skalar im $\mathbf{R}^n$ genannten Rechengesetze für den Bereich $\mathscr{P}$ hinsichtlich der zu betrachtenden Operationen gelten.

2. Entsprechend zur Aufgabe 1 führe man eine analoge Überprüfung für die folgenden Teilmengen von $\mathscr{P}$ durch, wo jedesmal dieselben vorherbetrachteten Operationen gemeint sind:

a) $\mathscr{P}_{n_0} := \{P : P \in \mathscr{P}, P(x) = \sum_{i=0}^{n} a_i x^i \text{ mit } n \leq n_0,\ n_0 \in \mathbf{N} \text{ fixiert}\}$. ($\mathscr{P}_{n_0}$ ist die Menge der reellen Polynome vom Grade höchstens n_0.)

b) $\mathscr{P} \setminus \mathscr{P}_{n_0}$.

c) $\mathscr{P} \setminus \bigcup \mathscr{P}_{n_0}\ (n_0 \neq k)$, k fixiert.

3. Für den Bereich $\mathbf{Q}$ der rationalen Zahlen seien folgende Operationen betrachtet:

$\oplus : \mathbf{Q} \times \mathbf{Q} \to \mathbf{Q}$ mit $r \oplus s := (-r) + (-s)$, $r, s \in \mathbf{Q}$

(rechts die üblichen Operationen in $\mathbf{Q}$);

$\odot : \mathbf{R} \times \mathbf{Q} \to \mathbf{Q}$ mit $\alpha \odot r := 0$, $\alpha \in \mathbf{R}$, $r \in \mathbf{Q}$.

Man überprüfe wieder analog zur Aufgabe 1 für die beiden soeben erklärten Operationen, welche der Grundregeln gültig sind.

4. Aus den Grundregeln der Addition und Multiplikation im $\mathbf{R}^n$ leite man die Gültigkeit der folgenden Gleichungen ab:

$-(-\boldsymbol{x}) = \boldsymbol{x}$,

$(-\alpha)\boldsymbol{x} = -\alpha\boldsymbol{x}$ für $\alpha \in \mathbf{R}$ und $\boldsymbol{x} \in \mathbf{R}^n$,

$\alpha \cdot \mathbf{0} = \mathbf{0}$ für $\alpha \in \mathbf{R}$ und das Nullelement $\mathbf{0}$ aus $\mathbf{R}^n$.

Weiterhin zeige man ebenso, lediglich unter Benutzung der Grundregeln, die Gültigkeit der Gleichungen

$$(\alpha - \beta)\boldsymbol{x} = \alpha\boldsymbol{x} - \beta\boldsymbol{x} \text{ für } \alpha, \beta \in \mathbf{R} \text{ und } \boldsymbol{x} \in \mathbf{R}^n,$$

$$\alpha(\boldsymbol{x} - \boldsymbol{y}) = \alpha\boldsymbol{x} - \alpha\boldsymbol{y} \text{ für } \alpha \in \mathbf{R} \text{ und } \boldsymbol{x}, \boldsymbol{y} \in \mathbf{R}^n.$$

(Bei Benutzung der Erklärung der Operationen im $\mathbf{R}^n$ als koordinatenweise Addition und Multiplikation mit einem Skalar werden die angegebenen Gleichungen unmittelbar einsichtlich. Es kommt aber in dieser Aufgabe darauf an, daß man an Hand der Herleitung aus den Grundregeln wirklich deren grundlegende Rolle mehr und mehr erkennt.)

3. Linearformen auf dem n-dimensionalen reellen Zahlenraum

3.1. Ein einführendes Beispiel. Erklärung der reellen Linearformen in n Variablen

Es seien $a_1, a_2, \ldots, a_n$ gegebene reelle Zahlen, wobei etwa die Größe a_i den durch die i-te Abteilung eines Betriebes erreichten Gewinn pro erzeugter Einheit angibt. Der Gesamtgewinn des Betriebes hängt natürlich von der von jeder Abteilung erzeugten Menge ab, er berechnet sich als

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n,$$

sofern x_i die von der i-ten Abteilung erzeugte Menge in Produktionseinheiten angibt.

Jedes Produktionsergebnis $(x_1, x_2, \ldots, x_n)$ ergibt einen Gewinn von $\sum_{i=1}^{n} a_ix_i$.

Die Art der funktionalen Abhängigkeit ist hierbei, wie man sagt, durch einen reellen linearen Ausdruck in n reellen Variablen $x_1, x_2, \ldots, x_n$ bestimmt.

Definition 1 (Reelle Linearformen in n Variablen). Unter einer *reellen Linearform in n reellen Variablen* $x_1, x_2, \ldots, x_n$ *versteht man eine reelle Funktion auf dem* $\mathbf{R}^n$,

$$f\colon \mathbf{R}^n \to \mathbf{R},$$

mit dem folgenden Verlauf:

$$f(x_1, x_2, \ldots, x_n) = a_1x_1 + a_2x_2 + \cdots + a_nx_n$$

bei fest vorgegebenen reellen Zahlen $a_i \in \mathbf{R}$, $i = 1, 2, \ldots, n$, mit $(x_1, x_2, \ldots, x_n) \in \mathbf{R}^n$. Die gegebenen reellen Zahlen $a_1, \ldots, a_n$ heißen die *Koeffizienten der Linearform.*

Es ist für f auch die Bezeichnung *lineares Funktional* auf dem $\mathbf{R}^n$ üblich.

Bemerkungen.

1. In der angeführten Definition ist hinsichtlich der Koeffizienten der betrachteten reellen Linearform keine Forderung bezüglich Positivität gemacht, wie das etwa in dem obigen Beispiel naturgemäß auftritt.

2. Im obigen Beispiel handelt es sich genauer gesagt also um eine reelle Linearform mit positiven Koeffizienten und einem *eingeschränkten* Definitionsbereich, in dem nämlich nur solche $(x_1, x_2, \ldots, x_n)$ betrachtet werden, für die $x_i \geqq 0$, $i = 1, 2, \ldots, n$, gilt.

3.2. Abstrakte Beschreibung und Veranschaulichung der reellen Linearformen

Die reellen Linearformen auf dem $\mathbf{R}^n$ können als Funktionen durch ihr Verhalten gegenüber der arithmetischen Struktur des $\mathbf{R}^n$ leicht wie folgt beschrieben werden.

Satz 1 (Reelle Linearformen als lineare Abbildungen). *Es sei* $f\colon \mathbf{R}^n \to \mathbf{R}$ *eine reelle Funktion auf dem* $\mathbf{R}^n$. *Dann gilt:* f *ist eine reelle Linearform, d. h., es gibt endlich viele reelle Zahlen* $a_1, a_2, \ldots, a_n$ *mit*

$$f(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} a_i x_i \qquad (*)$$

für alle $(x_1, \ldots, x_n) \in \mathbf{R}^n \Leftrightarrow f$ *ist eine lineare Abbildung von* $\mathbf{R}^n$ *in* $\mathbf{R}$, *d. h.,* f *hat bezüglich der koordinatenweisen Addition im* $\mathbf{R}^n$ *und der Multiplikation mit reellen Skalaren die folgenden Eigenschaften:*

1. $f(\boldsymbol{x} + \boldsymbol{y}) = f(\boldsymbol{x}) + f(\boldsymbol{y})$ *für alle* $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{R}^n$ (*Additivität der Funktion* f).
2. $f(\alpha \boldsymbol{x}) = \alpha f(\boldsymbol{x})$ *für alle* $\alpha \in \mathbf{R}$ *und* $\boldsymbol{x} \in \mathbf{R}^n$ (*Homogenität der Funktion* f).

Bemerkung. Die Eigenschaften 1 und 2, die Additivität und die Homogenität von f, kann man natürlich zu einer einzigen Eigenschaft (Linearität von f) zusammenziehen:

$$f(\alpha \boldsymbol{x} + \beta \boldsymbol{y}) = \alpha f(\boldsymbol{x}) + \beta f(\boldsymbol{y})$$

für alle $\alpha, \beta \in \mathbf{R}$ und alle $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{R}^n$. Per Induktion folgt dann auch

$$f\left(\sum_{i=1}^{k} \alpha_i \boldsymbol{x}_i\right) = \sum_{i=1}^{k} \alpha_i f(\boldsymbol{x}_i)$$

bei $\alpha_i \in \mathbf{R}$ für beliebiges $k \in \mathbf{N}$.

Beweis des Satzes. Der Aussagenteil „$\Rightarrow$" ist nach Erklärung der arithmetischen Struktur des $\mathbf{R}^n$ klar.

Zu „$\Leftarrow$": Es sind Zahlen $a_i \in \mathbf{R}$ zu finden, so daß f eine Darstellung $(*)$ hat. Bei einer gegebenen Linearform f haben die a_i folgende Bedeutung:

$$a_i = f(\boldsymbol{e}_i) \qquad \text{mit} \qquad \boldsymbol{e}_i = (0, \ldots, 1, \ldots, 0),$$

wobei in dem n-Tupel e_i lediglich an der i-ten Stelle eine 1 steht und sonst lauter Nullen vorkommen. Man wird also auch bei einer gegebenen linearen Abbildung f den Ausdruck

$$f(\boldsymbol{x}) = \sum_{i=1}^{n} f(\boldsymbol{e}_i)\, x_i \qquad \text{mit} \qquad \boldsymbol{x} = (x_1, x_2, \ldots, x_n)$$

zu bestätigen suchen. Das gilt auch wirklich wegen der möglichen Darstellung

$$x = \sum_{i=1}^{n} x_i e_i \qquad \text{mit} \qquad e_i = (0, \ldots, \underset{i\text{-te Stelle}}{1}, \ldots, 0).$$

Damit ist der Satz bewiesen.

Für die Fälle $n = 1$ und $n = 2$ lassen sich die Linearformen auf dem $\mathbf{R}^n$ durch ihre Funktionsgraphen gut veranschaulichen, wenn man dabei für den $\mathbf{R}^1 \times \mathbf{R}^1$ bzw. $\mathbf{R}^2 \times \mathbf{R}^1$ die euklidische Veranschaulichung wählt (vgl. die erste Art der Veranschaulichung des $\mathbf{R}^n$ in 2.2.).

Die Linearformen auf dem $\mathbf{R}^1$ entsprechen genau den sämtlichen Ursprungsgeraden im $\mathbf{R}^2$, ausgenommen diejenige, die senkrecht zur Grundgeraden steht. Die Linearformen auf dem $\mathbf{R}^2$ entsprechen genau den sämtlichen Ursprungsebenen im $\mathbf{R}^3$,

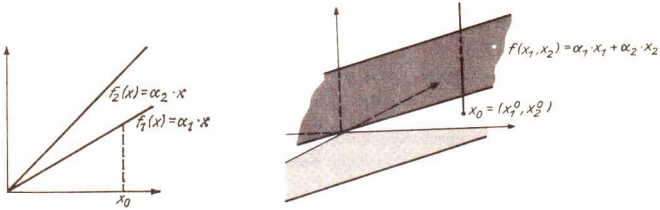


Abb. 3

ausgenommen diejenigen, die senkrecht zur Grundebene stehen. Abbildung 3 zeigt Funktionsgraphen der Linearformen auf dem $\mathbf{R}^1$ und dem $\mathbf{R}^2$. (Unter Heranziehung von Schulkenntnissen mache man sich die geometrische Bedeutung der Koeffizienten der Linearformen $f(x) = \alpha x$ bzw. $f(x_1, x_2) = \alpha_1 x_1 + \alpha_2 x_2$ klar.)

Fragt man nach einer Veranschaulichung der Linearformen auf dem $\mathbf{R}^n$ unter Benutzung der faserweisen Veranschaulichung, so wird man auf folgendes geführt. Die Linearformen des $\mathbf{R}^n$ entsprechen genau den sämtlichen Punkten des $\mathbf{R}^n$, und zwar die Linearform $f(x) = \alpha_1 x_1 + \cdots + \alpha_n x_n$ genau dem Punkt $(\alpha_1, \alpha_2, \ldots, \alpha_n)$. Eine Linearform auf dem $\mathbf{R}^n$ ist eindeutig durch das geordnete n-Tupel ihrer Koeffizienten bestimmt. Mit den Punkten des $\mathbf{R}^n$ können arithmetische Operationen ausgeführt werden. Vermöge der angegebenen Entsprechung zwischen den Linearformen und Punkten des $\mathbf{R}^n$ kann man also in natürlicher Weise auch für die Linearformen arithmetische Operationen erklären, was jetzt näher erörtert werden soll.

3.3. Arithmetische Struktur im Bereich der Linearformen des $\mathbf{R}^n$

Es bezeichne $\mathscr{L}(\mathbf{R}^n)$ die Menge aller reellen Linearformen auf dem $\mathbf{R}^n$. Die erwähnte Zuordnung zwischen den Punkten des $\mathbf{R}^n$ und den Linearformen auf dem $\mathbf{R}^n$ stellt eine eineindeutige surjektive Abbildung $\Phi\colon \mathbf{R}^n \to \mathscr{L}(\mathbf{R}^n)$ dar. Der Verlauf von Φ ist dabei bestimmt durch

$$\Phi(y) := f \quad \text{mit} \quad f(x) = y_1x_1 + y_2x_2 + \cdots + y_nx_n$$

für alle $x = (x_1, x_2, \ldots, x_n)$. Mittels dieser Abbildung Φ kommt eine Übertragung der arithmetischen Struktur des $\mathbf{R}^n$ in die Menge aller Linearformen auf dem $\mathbf{R}^n$ zustande, nämlich

$$\Phi(y) + \Phi(z) := \Phi(y + z),$$
$$\alpha\Phi(y) := \Phi(\alpha y).$$

Nun ist aber noch auf eine formal andere Weise eine Addition und eine Multiplikation mit reellen Skalaren für die Linearformen erklärt, und zwar die punktweise Addition (Superposition) der Linearformen und die punktweise Multiplikation mit einem reellen Skalar, denn es handelt sich ja bei den Linearformen um Abbildungen vom $\mathbf{R}^n$ in $\mathbf{R}$.

Definition 1 (Punktweise Addition (Superposition) von reellen Funktionen und punktweise Multiplikation mit Skalaren). Es sei X eine beliebige nichtleere Menge, f und g seien zwei beliebige reelle Funktionen auf X:

$$f\colon X \to \mathbf{R}, \qquad g\colon X \to \mathbf{R}.$$

Unter der *punktweisen Summe der Funktionen* f und g versteht man diejenige Funktion $h\colon X \to \mathbf{R}$, die den folgenden Verlauf hat:

$$h(x) = f(x) + g(x) \quad \text{für alle} \quad x \in X.$$

Man schreibt üblicherweise für diese Summe $f + g$.

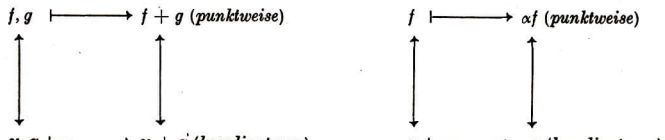
Unter dem *punktweisen Produkt der Funktion* $f\colon X \to \mathbf{R}$ mit einem reellen Skalar $\alpha \in \mathbf{R}$ versteht man diejenige Funktion $h\colon X \to \mathbf{R}$, die den folgenden Verlauf hat:

$$h(x) = \alpha f(x) \quad \text{für alle} \quad x \in X.$$

Man schreibt üblicherweise für dieses Produkt αf.

Bemerkung. Diese punktweise vollzogenen Operationen für reelle Funktionen sind uns schon in dem Spezialfall der koordinatenweisen Addition und Multiplikation mit Skalaren im $\mathbf{R}^n$ begegnet, wenn man die Elemente von $\mathbf{R}^n$ als Funktionen $x\colon \{1, 2, \ldots, n\} \to \mathbf{R}$ auffaßt. Hinsichtlich der vorhin erklärten Operationen für reelle Linearformen haben wir den folgenden

Satz 2 (Punktweise Operationen für Linearformen, widergespiegelt an ihren Koeffiziententupeln). *In der Menge $\mathscr{L}(\mathbf{R}^n)$ aller reellen Linearformen auf dem $\mathbf{R}^n$ stimmt die punktweise Addition von Linearformen mit der Addition überein, die aus der Addition der Koeffiziententupel hervorgeht. Ebenso stimmt die punktweise Multiplikation einer Linearform mit einem Skalar mit der Multiplikation überein, die aus der entsprechenden Multiplikation des Koeffiziententupels mit dem Skalar hervorgeht.*
In Diagrammform: $f, g \in \mathscr{L}(\mathbf{R}^n)$, $\boldsymbol{y}, \boldsymbol{z}$ ihre entsprechenden Koeffiziententupel



Beweis. Es sei $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$ und $\boldsymbol{z} = (z_1, z_2, \ldots, z_n)$. Dann gilt für jedes $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ aus dem $\mathbf{R}^n$

$$f(\boldsymbol{x}) = \sum_{i=1}^{n} y_i x_i, \quad g(\boldsymbol{x}) = \sum_{i=1}^{n} z_i x_i.$$

Die durch die koordinatenweise Addition der Koeffiziententupel hervorgehende Linearform ist

$$h(\boldsymbol{x}) = \sum_{i=1}^{n} (y_i + z_i)\, x_i.$$

Die durch die punktweise Addition hervorgehende Linearform ist $f + g$ mit

$$(f + g)(\boldsymbol{x}) = f(\boldsymbol{x}) + g(\boldsymbol{x}) = \sum_{i=1}^{n} y_i x_i + \sum_{i=1}^{n} z_i x_i.$$

Durch eine andere Anordnung ergibt sich

$$(f + g)(\boldsymbol{x}) = \sum_{i=1}^{n} (y_i + z_i)\, x_i.$$

Also gilt, wie behauptet, $h = f + g$. Entsprechendes erhält man für die Multiplikation mit einem Skalar.

Unter Verweis auf den im Band 1 der Reihe MfL erwähnten Isomorphiebegriff können wir sagen, daß die Abbildung Φ einen Isomorphismus zwischen der arithmetischen Struktur des $\mathbf{R}^n$ bezüglich der koordinatenweisen Operationen und der arithmetischen Struktur von $\mathscr{L}(\mathbf{R}^n)$ bezüglich der punktweisen Operationen darstellt.

Bemerkung. Insbesondere geht aus dem letzten Satz hervor, daß die punktweise Addition und die punktweise Multiplikation mit einem Skalar in $\mathscr{L}(\mathbf{R}^n)$ den gleichen Gesetzen genügen wie die entsprechenden koordinatenweisen Operationen im $\mathbf{R}^n$.

3.4. Lineare Gleichungssysteme und Linearformen. Lineare Teilräume

Linearformen auf dem $\mathbf{R}^n$ treten uns in linearen Gleichungssystemen mit n Unbekannten entgegen. Es seien uns etwa $m (m \in \mathbf{N})$ Linearformen $f_1, f_2, \ldots, f_m$ mit den Koeffiziententupeln

$$(a_{11}, a_{12}, \ldots, a_{1n}), \quad (a_{21}, a_{22}, \ldots, a_{2n}), \ldots, (a_{m1}, a_{m2}, \ldots, a_{mn})$$

gegeben. Außerdem seien m reelle Zahlen $b_1, b_2, \ldots, b_m$ vorgeschrieben. Wir suchen alle Elemente $x \in \mathbf{R}^n$ zu ermitteln, die den folgenden Beziehungen genügen:

$$\begin{aligned} f_1(x) &= b_1, \\ f_2(x) &= b_2, \\ &\vdots \\ f_m(x) &= b_m. \end{aligned}$$

Ausführlich geschrieben heißt das aber, daß man alle n-Tupel $(x_1, x_2, \ldots, x_n)$ ermitteln soll, die dem folgenden linearen Gleichungssystem genügen:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2, \\ &\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m. \end{aligned}$$

Zunächst verschaffen wir uns eine allgemeine Einsicht in die Struktur der Lösungsgesamtheit, um dann später ein Verfahren zur konkreten Bestimmung der Lösungsgesamtheit anzugeben.

Wir beginnen mit dem einfachsten Fall einer einzigen Gleichung, d. h. einer einzigen Linearform, wobei außerdem die gegebene rechte Seite b_1 den speziellen Wert 0 hat.

Definition 1 (Kern bzw. Nullraum einer Linearform). Es sei f eine reelle Linearform auf dem $\mathbf{R}^n$. Unter dem *Kern* bzw. dem *Nullraum dieser Linearform* versteht man die Lösungsgesamtheit der Gleichung

$$f(x) = 0.$$

Hiermit ist also die folgende Menge gemeint:

$$\{x : x \in \mathbf{R}^n, \quad f(x) = 0\}.$$

Man schreibt

$$\ker f = \{x : x \in \mathbf{R}^n, \quad f(x) = 0\}.$$

Wir illustrieren die Kerne von Linearformen an einigen Beispielen.

1. $f \in \mathscr{L}(\mathbf{R}^1)$ sei eine reelle Linearform auf dem $\mathbf{R}^1$. Dann gilt:

a) Die Linearform f ist identisch Null (die Nullform) $\Leftrightarrow \ker f = \mathbf{R}^1$.

b) Die Linearform f ist nicht ausgeartet (d. h. von der Nullform verschieden) $\Leftrightarrow \ker f = \{0\}$.

2. $f \in \mathscr{L}(\mathbf{R}^2)$ sei eine reelle Linearform auf dem $\mathbf{R}^2$. Dann gilt:

a) Die Linearform ist ausgeartet (die Nullform) $\Leftrightarrow \ker f = \mathbf{R}^2$.

b) Die Linearform ist nicht ausgeartet $\Leftrightarrow \ker f$ besteht aus allen Punkten einer Ursprungsgeraden im $\mathbf{R}^2$.

Die letzte Behauptung mache man sich mittels der euklidischen Veranschaulichung von Linearformen geometrisch klar.

Nun klären wir die Struktur der Kerne von Linearformen allgemein.

Satz 1 (Lineare Struktur des Kerns einer Linearform auf dem $\mathbf{R}^n$). *Es sei f eine Linearform auf dem $\mathbf{R}^n$. Dann hat der Kern dieser Linearform die folgenden Linearitätseigenschaften*:

1. $\boldsymbol{x}, \boldsymbol{y} \in \ker f \Rightarrow \boldsymbol{x} + \boldsymbol{y} \in \ker f$.

2. $\boldsymbol{x} \in \ker f, \quad \alpha \in \mathbf{R} \Rightarrow \alpha \boldsymbol{x} \in \ker f$.

Beweis. Eine Linearform ist eine lineare Abbildung von $\mathbf{R}^n$ in $\mathbf{R}$. Also gilt bei $\boldsymbol{x}, \boldsymbol{y} \in \ker f$ stets $f(\boldsymbol{x}) = 0$ und $f(\boldsymbol{y}) = 0$. Für beliebige reelle Zahlen $\alpha, \beta \in \mathbf{R}$ ist deshalb

$$0 = \alpha f(\boldsymbol{x}) + \beta f(\boldsymbol{y}) = f(\alpha \boldsymbol{x} + \beta \boldsymbol{y}),$$

d. h., $\alpha \boldsymbol{x} + \beta \boldsymbol{y} \in \ker f$.

Bemerkung. Wie schon beim Beweis geschehen, kann man beide Eigenschaften 1 und 2 zu der gleichwertigen Eigenschaft

$$\boldsymbol{x}, \boldsymbol{y} \in \ker f \quad \text{und} \quad \alpha, \beta \in \mathbf{R} \Rightarrow \alpha \boldsymbol{x} + \beta \boldsymbol{y} \in \ker f$$

zusammenfassen.

Teilmengen des $\mathbf{R}^n$ mit der genannten Linearitätseigenschaft spielen im weiteren eine beträchtliche Rolle. Sie werden deshalb herausgehoben.

Definition 2 (Lineare Teilräume des $\mathbf{R}^n$). Eine nichtleere Teilmenge L des $\mathbf{R}^n$ heißt ein *linearer Teilraum* des $\mathbf{R}^n$, wenn L die folgenden Linearitätseigenschaften hat:

1. $\boldsymbol{x}, \boldsymbol{y} \in L \Rightarrow \boldsymbol{x} + \boldsymbol{y} \in L$.

2. $\boldsymbol{x} \in L, \quad \alpha \in \mathbf{R} \Rightarrow \alpha \boldsymbol{x} \in L$.

Wir können den letzten Satz in Kurzform wie folgt formulieren:

Der Kern einer Linearform auf dem $\mathbf{R}^n$ ist ein linearer Teilraum des $\mathbf{R}^n$.

Unter Berufung auf Schulkenntnisse mache man sich klar, daß beispielsweise bei euklidischer Veranschaulichung des $\mathbf{R}^3$ seine sämtlichen linearen Teilräume folgendes geometrisches Aussehen haben:

1. Die einpunktige Menge $\{0\}$,
2. die Ursprungsgeraden,

3. die Ursprungsebenen,
4. der $\mathbb{R}^3$.

Es sei bemerkt, daß nicht alle linearen Teilräume des $\mathbb{R}^3$ als Kerne von Linearformen auftreten können. Wir wollen im späteren Verlauf entscheiden, welche linearen Teilräume gerade als Kerne von Linearformen in Frage kommen. Betrachtet man anstelle einer einzigen Linearform f auf dem $\mathbb{R}^n$ eine Schar von m Stück $f_1, f_2, \ldots, f_m \in \mathscr{L}(\mathbb{R}^n)$ und fragt man nach der Lösungsgesamtheit des Gleichungssystems

$$\begin{aligned} f_1(x) &= 0, \\ f_2(x) &= 0, \\ &\cdots\cdots \\ f_m(x) &= 0, \end{aligned}$$

so sieht man, daß die Lösungsgesamtheit gerade gleich dem Durchschnitt aller Kerne ist:

$$\{x : x \in \mathbb{R}^n \text{ mit } f_i(x) = 0 \text{ für alle } i = 1, 2, \ldots, m\} = \bigcap_{i=1}^{m} \ker f_i.$$

Diese Gesamtheit ist wieder ein linearer Teilraum des $\mathbb{R}^n$. Allgemein bestätigt man leicht den folgenden Satz:

Satz 2 (Durchschnitt von linearen Teilräumen). *Es sei $(L_i)_{i\in I}$ eine beliebige Familie von linearen Teilräumen des $\mathbb{R}^n$. Dann ist der Durchschnitt $\bigcap_{i\in I} L_i$ wieder ein linearer Teilraum des $\mathbb{R}^n$.*

Beweis. Es sei $L := \bigcap_{i\in I} L_i$. Man muß zweierlei zeigen: 1. $L \neq \emptyset$, 2. $x, y \in L$; $\alpha, \beta \in \mathbb{R} \Rightarrow \alpha x + \beta y \in L$.
Zu 1. genügt der Hinweis $0 \in L_i$ für alle $i \in I$.

Mit dem Durchschnittssatz gelangen wir zu dem wichtigen Begriff der linearen Hülle einer Teilmenge des $\mathbb{R}^n$.

Definition 3 (Von einer Teilmenge erzeugter linearer Teilraum, lineare Hülle). Es sei U eine beliebige Teilmenge des $\mathbb{R}^n$. Unter dem von dieser *Teilmenge erzeugten linearen Teilraum* $L(U)$ versteht man den kleinsten linearen Teilraum des $\mathbb{R}^n$, der U enthält, d. h., es wird definiert:

$$L(U) := \bigcap L \qquad (L \text{ linearer Teilraum des } \mathbb{R}^n \text{ mit } L \supseteq U).$$

Bezüglich der Inklusion ist $L(U)$ wirklich der kleinste lineare Teilraum, der U umfaßt. Man sagt auch, daß der lineare Teilraum $L(U)$ von der Menge U *aufgespannt* wird oder daß $L(U)$ die *lineare Hülle* der Menge U ist.

Man verdeutliche sich diesen Begriff etwa wieder durch einige geometrisch-anschauliche Überlegungen an der euklidischen Veranschaulichung des $\mathbb{R}^3$. Man findet:

1. $L(U) = \{\mathbf{0}\} \Leftrightarrow U = \{\mathbf{0}\}$ oder $U = \emptyset$.
2. Es sei $U = \{\boldsymbol{x}_0\}$ mit $\boldsymbol{x}_0 \neq \mathbf{0}$. Dann besteht $L(U)$ aus allen Punkten der Ursprungsgeraden, die durch $\boldsymbol{x}_0$ geht.
3. Es sei $U = \{\boldsymbol{x}_1, \boldsymbol{x}_2\}$, wobei $\boldsymbol{x}_1, \boldsymbol{x}_2$ zwei Punkte des $\mathbb{R}^3$ sind, die auf keiner gemeinsamen Ursprungsgeraden liegen. Dann besteht $L(U)$ aus allen Punkten der Ursprungsebene, die durch $\boldsymbol{x}_1$ und $\boldsymbol{x}_2$ geht.
4. Es sei $U = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3\}$, wobei $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3$ Punkte des $\mathbb{R}^3$ sind, die auf keiner gemeinsamen Ursprungsebene liegen. Dann besteht $L(U)$ gerade aus dem ganzen $\mathbb{R}^3$.

Man wird so zu der Vorstellung geführt, daß die linearen Teilräume des $\mathbb{R}^n$ sich immer schon aus endlichen Teilmengen aufspannen lassen. Die genaue Situation wird im nächsten Abschnitt erörtert. Zuvor vermerken wir noch als Gegenstück zur angegebenen äußeren Kennzeichnung der linearen Hülle eine innere Kennzeichnung von $L(U)$.

Definition 4 (Endliche Linearkombination von Elementen des $\mathbb{R}^n$). Es seien $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$ endlich viele beliebige Elemente aus $\mathbb{R}^n$. Diese brauchen nicht notwendig verschieden zu sein. Man sagt, daß ein Element $\boldsymbol{x} \in \mathbb{R}^n$ eine *(endliche) Linearkombination* dieser endlich vielen Elemente ist, wenn es gewisse reelle Zahlen $\alpha_1, \alpha_2, \ldots, \alpha_k$ gibt mit der Darstellung

$$\boldsymbol{x} = \sum_{i=1}^{k} \alpha_i \boldsymbol{x}_i .$$

$\boldsymbol{x}$ heißt auch *linear kombinierbar* aus den gegebenen Elementen $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$.

Satz 3 (Innere Beschreibung des von einer Teilmenge aufgespannten linearen Teilraumes). *Es sei U eine Teilmenge des $\mathbb{R}^n$, wobei diesmal $U \neq \emptyset$ vorausgesetzt werde. Dann gilt: Der von U aufgespannte lineare Teilraum $L(U)$ des $\mathbb{R}^n$ besteht gerade aus allen endlichen Linearkombinationen von Elementen aus U:*

$L(U) = \{\boldsymbol{x} : \boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{x}$ *läßt sich aus endlich vielen Elementen von U linear kombinieren*$\}$.

Beweis. Es sei K die Menge aller endlichen Linearkombinationen mittels Elementen aus U. Es ist $U \subseteq K$, man verweist einfach auf die Darstellung $\boldsymbol{x} = 1 \cdot \boldsymbol{x}$. K ist auch als linearer Teilraum von $\mathbb{R}^n$ zu erkennen:

$$\boldsymbol{x}, \boldsymbol{y} \in K \Rightarrow \boldsymbol{x} = \sum_{i=1}^{k} \alpha_i \boldsymbol{x}_i , \qquad \boldsymbol{y} = \sum_{j=1}^{l} \beta_j \boldsymbol{y}_j$$

mit gewissen $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k \in U$ und $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_l \in U$ und gewissen α_i und β_j. Also ist

$$\boldsymbol{x} + \boldsymbol{y} = \sum_{i=1}^{k} \alpha_i \boldsymbol{x}_i + \sum_{j=1}^{l} \beta_j \boldsymbol{y}_j ,$$

d. h. $\boldsymbol{x} + \boldsymbol{y} \in K$.

Folglich muß $K \supseteq L(U)$ sein. Andererseits enthält aber jeder lineare Teilraum von $\mathbf{R}^n$, der U umfaßt, auch jede endliche Linearkombination von Elementen aus U. Also haben wir auch $K \subseteq L(U)$ und damit $K = L(U)$.

3.5. Übungsaufgaben

1. Die Menge $\mathscr{P}$ aller reellen Polynome werde mit der punktweisen Addition und Multiplikation mit reellen Skalaren ausgestattet (vgl. 2.4., Aufgabe 1).
Welche der beiden folgenden reellen Abbildungen $f: \mathscr{P} \to \mathbf{R}$ und $g: \mathscr{P} \to \mathbf{R}$ ist ein lineares Funktional auf $\mathscr{P}$, hat also die Eigenschaft der Homogenität und Additivität:

$f: \mathscr{P} \to \mathbf{R}$ mit $f(P) = \sum_{i=0}^{n} a_i$, sofern $P(x) = \sum_{i=0}^{n} a_i x^i$,

$g: \mathscr{P} \to \mathbf{R}$ mit $g(P) = P(x_0)$, x_0 eine beliebige fixierte reelle Zahl?

2. Die Menge

$$M := \{(x, y) : x, y \in \mathbf{R}, x^2 + (y-1)^2 = 1,\ y \neq 2\}$$

(die Kreislinie, aus der der Nordpol herausgestochen ist) soll mit einer gewissen „Addition" und einer „Multiplikation mit reellen Skalaren" derart ausgestattet werden, daß man damit eine gewisse Veranschaulichung für den Raum der Linearformen auf dem $\mathbf{R}^1$ erhält. Wie hat das zu geschehen?

3. Man betrachte eine nichtausgeartete Linearform f auf dem $\mathbf{R}^2$, ihr Koeffiziententupel sei (a_1, a_2). Man versuche, eine geometrische Lagebeziehung zwischen dem Kern der Linearform (als Gerade in der euklidischen Veranschaulichung des $\mathbf{R}^2$ aufgefaßt) und der Ursprungsgeraden durch den Punkt (a_1, a_2) herauszufinden.

4. Die Vereinigung von zwei linearen Teilräumen des $\mathbf{R}^n$ ist nicht notwendig wieder ein linearer Teilraum des $\mathbf{R}^n$ (Beispiele!).
Welche gegenseitige Lagebedingung müssen zwei lineare Teilräume (sogar eine beliebige Familie linearer Teilräume) des $\mathbf{R}^n$ erfüllen, damit die Vereinigung ein linearer Teilraum wird?

5. Läßt sich das Element $(1, 1, 1)$ des $\mathbf{R}^3$ linear aus den Elementen $(1, -1, 0)$, $(2, 3, 1)$ kombinieren?

6. Im $\mathbf{R}^n$ seien die Elemente $x_0, x_1, \ldots, x_k$ vorgegeben. Welches rechnerische Anliegen hat man bei der Frage zu bewältigen, ob x_0 zu dem von den Elementen $x_1, \ldots, x_k$ aufgespannten linearen Teilraum $L(\{x_1, x_2, \ldots, x_k\})$ gehört?

7. Welche Inklusionsbeziehungen gelten zwischen $L(U_1) \cap L(U_2)$ und $L(U_1 \cap U_2)$ bzw. $L(U_1 \cup U_2)$ und $L(U_1) \cup L(U_2)$?

8. Kann der lineare Teilraum $L(\{x\})$, $x \in \mathbf{R}^3$, der also von dem einzigen Element x aufgespannt wird, Kern einer Linearform auf dem $\mathbf{R}^3$ sein?

4. Lineare Unabhängigkeit

4.1. Erklärung der linear unabhängigen Teilmengen des R^n

In unmittelbarer Fortsetzung zu den letzten Ausführungen versteht sich die folgende

Definition 1 (Lineare Abhängigkeit eines Elementes von einer Teilmenge des R^n). Es sei $x \in R^n$ und $U \subseteq R^n$. Das Element x heißt von der Teilmenge U *linear abhängig* genau dann, wenn $x \in L(U)$ gilt.

Lineare Abhängigkeit des Elementes x von U bedeutet also im Fall $U \neq \emptyset$, daß sich x aus U linear kombinieren läßt.

Definition 2 (Lineare Unabhängigkeit von Elementen des R^n). Es sei U eine nichtleere Teilmenge des R^n. Diese heißt *linear unabhängige Teilmenge* bzw. die Elemente von U heißen linear unabhängig genau dann, wenn kein Element von U von den übrigen linear abhängt, d. h., wenn stets $x \notin L(U \setminus \{x\})$ für $x \in U$ gilt.

Ist die Menge U nicht linear unabhängig, so nennt man diese Menge bzw. ihre Elemente *linear abhängig*. Es muß dann also wenigstens ein $x \in U$ geben, das von den übrigen linear abhängig ist.

Bemerkungen.

1. Das Nullelement des R^n ist von jeder beliebigen Teilmenge linear abhängig, denn es ist stets $0 \in L(U)$.

2. Es sei U eine einpunktige Teilmenge des R^n, $U = \{x\}$. Dann gilt: U linear unabhängig $\Leftrightarrow x \neq 0$.

3. Es sei U eine Teilmenge des R^1, dann gilt: U ist linear unabhängig $\Leftrightarrow$ U besteht aus genau einem Element x, und dieses ist ungleich 0.

4. Es sei U eine mehrpunktige Teilmenge des R^2. Die Menge U ist linear unabhängig $\Leftrightarrow$ U besteht aus genau zwei Punkten, und diese liegen nicht auf einer gemeinsamen Ursprungsgeraden.

Entsprechendes mache man sich an der euklidischen Veranschaulichung des R^3 klar.

Nun leiten wir ein wichtiges Kriterium für die lineare Abhängigkeit von endlich vielen Elementen aus dem R^n ab.

Satz 1 (Kennzeichnung der linearen Abhängigkeit bzw. Unabhängigkeit bei endlich vielen Elementen). *Es sei* $\{x_1, x_2, \ldots, x_r\}$ *eine beliebige endliche* (*r-elementige*) *nichtleere Teilmenge des* $\mathbb{R}^n$. *Dann gilt:*

1. *Die gegebenen Elemente* $x_1, x_2, \ldots, x_r$ $(r \geqq 1)$ *sind linear abhängig* $\Leftrightarrow$ *Es gibt r reelle Zahlen* $\alpha_1, \alpha_2, \ldots, \alpha_r$, *die nicht sämtlich gleich Null sind, so daß eine Darstellung* $\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_r x_r = 0$ *möglich ist.*

2. *Die gegebenen Elemente* $x_1, x_2, \ldots, x_r$ $(r \geqq 1)$ *sind linear unabhängig* $\Leftrightarrow$ *Aus einer Darstellung* $\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_r x_r = 0$ *folgt stets* $\alpha_1 = \alpha_2 = \cdots = \alpha_r = 0$.

Beweis. Zu 1. Es sei $\{x_1, x_2, \ldots, x_r\}$ eine linear abhängige Menge. Dann sei o. B. d. A.

$$x_1 \in L(\{x_1, x_2, \ldots, x_r\} \setminus \{x_1\})$$

angenommen, d. h., es gibt reelle Zahlen $\alpha_2, \alpha_3, \ldots, \alpha_r$ mit $x_1 = \sum_{i=2}^{r} \alpha_i x_i$. Dann hat man

$$1x_1 - \sum_{i=2}^{r} \alpha_i x_i = 0,$$

wobei der Koeffizient bei x_1 ungleich 0 ist.

Wenn es umgekehrt eine Darstellung $\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_r x_r = 0$ gibt, wobei nicht alle $\alpha_1, \alpha_2, \ldots, \alpha_r$ gleich Null sind, so ist etwa $\alpha_1 \neq 0$. Wir haben dann

$$x_1 = \sum_{i=2}^{r} \frac{-\alpha_i}{\alpha_1} x_i,$$

d. h. $x_1 \in L(\{x_1, x_2, \ldots, x_r\} \setminus \{x_1\})$.

Zu 2. Die lineare Unabhängigkeit einer nichtleeren Menge ist das Gegenteil der linearen Abhängigkeit. Also folgt alles aus 1.

Bemerkungen.

1. Folgende Sprechweisen für die nach dem Äquivalenzpfeil im Punkt 1 und Punkt 2 des Satzes aufgetretenen Situationen sind bequem.

a) Das Nullelement aus $\mathbb{R}^n$ läßt sich *nichttrivial linear kombinieren* aus den Elementen $x_1, x_2, \ldots, x_r$.

b) Das Nullelement des $\mathbb{R}^n$ läßt sich *nur trivial linear kombinieren* aus den Elementen $x_1, x_2, \ldots, x_r$.

2 (Lineare Abhängigkeit von Elementefamilien des $\mathbb{R}^n$). Die lineare Abhängigkeit einer Menge von r Elementen aus dem $\mathbb{R}^n$ läßt sich sinngemäß auf eine Familie von r Elementen des $\mathbb{R}^n$ ausdehnen. Die Elementefamilie $(x_i)_{i=1,2,\ldots,r}$ des $\mathbb{R}^n$ heißt *linear abhängig* genau dann, wenn sich das Nullelement des $\mathbb{R}^n$ durch eine gewisse Familie nicht sämtlich verschwindender reeller Zahlen $\alpha_1, \alpha_2, \ldots, \alpha_r$ (d. h., wenigstens eines der α_i ist von Null verschieden) in der Form $\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_r x_r = 0$ darstellen läßt.

Eine Familie $x_1, x_2, \ldots, x_r$ von Elementen des $\mathbf{R}^n$ (jetzt brauchen also nicht mehr je zwei voneinander verschieden zu sein) ist also gewiß linear abhängig, wenn in ihr zwei gleiche Elemente vorkommen.

Eine Familie von Elementen des $\mathbf{R}^n$ heißt *linear unabhängig* genau dann, wenn sie nicht linear abhängig ist. Notwendig für die lineare Unabhängigkeit der Familie $x_1, x_2, \ldots, x_r$ ist die paarweise Verschiedenheit der Elemente.

Der Begriff der linearen Abhängigkeit von Elementefamilien ist z. B. bei linearen Gleichungssystemen bzw. bei den Zeilen einer Matrix bequem.

3 (Lineare Unabhängigkeit bzw. Abhängigkeit im Bereich der Linearformen). Im Bereich $\mathscr{L}(\mathbf{R}^n)$ der Linearformen auf dem $\mathbf{R}^n$ herrscht die gleiche arithmetische Struktur wie im $\mathbf{R}^n$. Es ist daher in gleicher Weise die lineare Unabhängigkeit bzw. Abhängigkeit erklärt. Dieser Umstand wird uns später bei den linearen Gleichungssystemen nützlich sein.

4.2. Basen und Dimension von linearen Teilräumen des $\mathbf{R}^n$

Wenn man den von einer Menge $U \subseteq \mathbf{R}^n$ aufgespannten linearen Teilraum $L(U)$ betrachtet, könnte man also im Fall der linearen Abhängigkeit der Menge U sicher ein gewisses Element $x \in U$ finden mit

$$L(U) = L(U \setminus \{x\}).$$

Die den linearen Teilraum $L = L(U)$ aufspannende Menge U, oder anders gesagt, das *Erzeugersystem* U von L, hätte man um ein Element reduziert. Wie weit kann man diesen Reduktionsprozeß treiben? Uns wäre an einer möglichst weitgehenden Reduktion gelegen, da dann die Erzeugung von L an Übersichtlichkeit gewinnt. Um solch eine Übersichtlichkeit geht es uns allein schon wegen einer bequemen Beschreibung der Struktur von Lösungsgesamtheiten linearer Gleichungssysteme.

Definition 1 (Basis eines linearen Teilraumes des $\mathbf{R}^n$). Es sei L ein linearer Teilraum des $\mathbf{R}^n$, dieser sei vom Nullraum verschieden, d. h. $L \neq \{0\}$. Eine Teilmenge $\mathfrak{B} \subseteq \mathbf{R}^n$ heißt eine *Basis* für L genau dann, wenn gilt:

1. $\mathfrak{B}$ spannt L auf, d. h. $L = L(\mathfrak{B})$.

2. Keine echte Teilmenge von $\mathfrak{B}$ spannt den gegebenen linearen Teilraum auf.

Das pflegt man knapp so auszudrücken:

Eine Basis von L ist ein *minimales Erzeugendensystem* von L.

Satz 1 (Kennzeichnung der Basen durch lineare Unabhängigkeit). *Es sei L ein vom Nullraum verschiedener linearer Teilraum des $\mathbf{R}^n$ und $\mathfrak{B} \subseteq \mathbf{R}^n$. Dann sind folgende Aussagen paarweise äquivalent:*

1. *$\mathfrak{B}$ ist eine Basis von L.*

2. *$\mathfrak{B}$ spannt L auf, und $\mathfrak{B}$ ist eine linear unabhängige Menge.*

3. *$\mathfrak{B}$ ist eine linear unabhängige Teilmenge von L, und keine echte Obermenge von $\mathfrak{B}$, die zu L gehört, ist linear unabhängig.*

Bemerkung. Man kann dann also der obigen knappen Formulierung für eine Basis noch hinzufügen: Eine Basis von L ist ein linear unabhängiges Erzeugendensystem von L, sie ist ein maximales linear unabhängiges System in L.

Beweis des Satzes. Es ist zu zeigen $1. \Leftrightarrow 2.$, $2. \Leftrightarrow 3.$, $1. \Leftrightarrow 3.$ Diese sechs Teilaussagen kann man sich im Beweis abkürzen, indem man zyklisch beweist: $1. \Rightarrow 2. \Rightarrow 3. \Rightarrow 1.$

Zu $1. \Rightarrow 2.$: Es sei also $\mathfrak{B}$ eine Basis von L. Dann ist $L = L(\mathfrak{B})$. Es muß die lineare Unabhängigkeit von $\mathfrak{B}$ gezeigt werden. Angenommen, $\mathfrak{B}$ ist linear abhängig. Dann gibt es ein $\boldsymbol{x}_1 \in \mathfrak{B}$ mit $\boldsymbol{x}_1 \in L(\mathfrak{B} \setminus \{\boldsymbol{x}_1\})$. Also wäre $L(\mathfrak{B}) = L(\mathfrak{B} \setminus \{\boldsymbol{x}_1\})$. Also könnte $\mathfrak{B}$ entgegen der Voraussetzung nicht minimal sein.

Zu $2. \Rightarrow 3.$: Es sei $\overline{\mathfrak{B}} \subseteq L$ eine echte Obermenge von $\mathfrak{B}$. Wir wählen ein $\boldsymbol{x} \in \overline{\mathfrak{B}} \setminus \mathfrak{B}$. Wir haben wegen $\mathfrak{B} \subseteq \overline{\mathfrak{B}} \setminus \{\boldsymbol{x}\} \subset \overline{\mathfrak{B}} \subseteq L$ die Beziehung $L = L(\overline{\mathfrak{B}} \setminus \{\boldsymbol{x}\})$. Also ist $\boldsymbol{x} \in L(\overline{\mathfrak{B}} \setminus \{\boldsymbol{x}\})$, d. h., $\overline{\mathfrak{B}}$ ist linear abhängig.

Zu $3. \Rightarrow 1.$: Zunächst erzeugt $\mathfrak{B}$ ganz L. Denn wäre $L(\mathfrak{B}) \subset L$, so wäre $\mathfrak{B} \cup \{\boldsymbol{x}\}$ für jedes $\boldsymbol{x} \in L \setminus L(\mathfrak{B})$ linear unabhängig, was nach Voraussetzung von 3. nicht zutrifft. Könnte man L durch eine echte Teilmenge $U \subset \mathfrak{B}$ erzeugen, so ist für jedes $\boldsymbol{x} \in \mathfrak{B} \setminus U$

$$\boldsymbol{x} \in L(U) = L(\mathfrak{B} \setminus \{\boldsymbol{x}\}) = L(\mathfrak{B}) = L.$$

$\boldsymbol{x} \in L(\mathfrak{B} \setminus \{\boldsymbol{x}\})$ würde aber entgegen der Voraussetzung besagen, daß $\mathfrak{B}$ linear abhängig wäre.

Mit der folgenden Definition verweisen wir auf ein wichtiges Beispiel für Basen.

Definition 2 (Natürliche Basis im $\mathbb{R}^n$). Unter der *natürlichen Basis* des reellen n-dimensionalen Zahlenraumes $\mathbb{R}^n$ versteht man die folgende Basis $\mathfrak{B}$: $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$; wobei

$$\boldsymbol{e}_i = (0, \ldots, \underset{i\text{-te Stelle}}{1}, \ldots, 0)$$

dasjenige n-Tupel reeller Zahlen ist, das genau an der i-ten Stelle eine 1 aufweist und sonst lauter Nullen.

Bemerkung. Diese Definition bedarf natürlich einer Rechtfertigung, d. h., man muß sich von der Basiseigenschaft des Systems $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ überzeugen. Offenbar spannen diese n Elemente den $\mathbb{R}^n$ auf, denn es gilt für jedes $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$

$$\boldsymbol{x} = \sum_{i=1}^{n} x_i \boldsymbol{e}_i.$$

Zum anderen sind sie auch linear unabhängig, denn aus

$$\sum_{i=1}^{n} \alpha_i \boldsymbol{e}_i = \boldsymbol{0}$$

folgt $(\alpha_1, \alpha_2, \ldots, \alpha_n) = \boldsymbol{0}$, also $\alpha_i \doteq 0$ für alle $i = 1, 2, \ldots, n$.

Es gibt natürlich auch andere Basen als die soeben angegebene. Das macht man sich leicht am R^1, R^2 und R^3 klar. Spätere Ausführungen zeigen das noch zur Genüge.

Von grundlegender Bedeutung ist nun folgender Sachverhalt.

Satz 2 (Gleichmächtigkeit der Basen eines linearen Teilraumes des R^n). *Es sei L ein vom Nullraum verschiedener linearer Teilraum des R^n. Dann hat L eine endliche Basis und es bestehen alle Basen von L aus ein und derselben Anzahl von Elementen*:

$$\mathfrak{B}_1, \mathfrak{B}_2 \text{ Basen von } L \Rightarrow |\mathfrak{B}_1| = |\mathfrak{B}_2|.$$

Beweis. Wir setzen zunächst voraus, es sei schon bewiesen, daß L wenigstens eine endliche Basis $\mathfrak{B}$: $\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_m$ besitzt. Es sei nun $\tilde{\mathfrak{B}}$ eine weitere Basis von L. Wir zeigen folgende Austauschmöglichkeit.

1. Schritt: Es gibt ein Element $\boldsymbol{x}_1 \in \tilde{\mathfrak{B}}$, so daß $\tilde{\mathfrak{B}}_1 := (\tilde{\mathfrak{B}} \setminus \{\boldsymbol{x}_1\}) \cup \{\boldsymbol{b}_1\}$ eine Basis von L ist.

2. Schritt: Es gibt ein Element $\boldsymbol{x}_2 \in \tilde{\mathfrak{B}} \setminus \{\boldsymbol{x}_1\}$, so daß $\tilde{\mathfrak{B}}_2 := (\tilde{\mathfrak{B}} \setminus \{\boldsymbol{x}_1, \boldsymbol{x}_2\}) \cup \{\boldsymbol{b}_1, \boldsymbol{b}_2\}$ eine Basis von L ist.

Induktiv fortfahrend gelang man so zum abschließenden Schritt.

m-ter Schritt: Es gibt ein Element $\boldsymbol{x}_m \in \tilde{\mathfrak{B}} \setminus \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{m-1}\}$, so daß $\tilde{\mathfrak{B}}_m := (\tilde{\mathfrak{B}} \setminus \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_m\}) \cup \{\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_m\}$ eine Basis von L ist.

Dieses Verfahren ist als *Steinitzsches Austauschverfahren* bekannt. Nach Abschluß des Verfahrens hat man mit $\tilde{\mathfrak{B}}_m \supseteq \mathfrak{B}$ eine Basis gefunden, die $\mathfrak{B}$ umfaßt. Wegen der Maximalität des linear unabhängigen Systems $\mathfrak{B}$ muß $\mathfrak{B}_m = \mathfrak{B}$ sein, d. h., es war $\tilde{\mathfrak{B}} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_m\}$ eine Menge aus ebenfalls m Elementen.

Wir begründen den ersten Schritt des Austauschverfahrens. Wegen $\boldsymbol{b}_1 \in L(\tilde{\mathfrak{B}}) = L$ läßt sich $\boldsymbol{b}_1$ linear aus $\tilde{\mathfrak{B}}$ kombinieren. In dieser Kombination tritt wenigstens ein Element mit einem von Null verschiedenen Koeffizienten auf, es sei dies $\boldsymbol{x}_1$. Dann ist $\boldsymbol{x}_1 \in L((\tilde{\mathfrak{B}} \setminus \{\boldsymbol{x}_1\}) \cup \{\boldsymbol{b}_1\})$. Jedes Element $\boldsymbol{x} \in L$ läßt sich aus $\tilde{\mathfrak{B}}$ linear kombinieren. Das dabei eventuell auftretende Element $\boldsymbol{x}_1$ ersetze man durch eine Linearkombination mittels $(\tilde{\mathfrak{B}} \setminus \{\boldsymbol{x}_1\}) \cup \{\boldsymbol{b}_1\}$, also erzeugt $(\tilde{\mathfrak{B}} \setminus \{\boldsymbol{x}_1\}) \cup \{\boldsymbol{b}_1\}$ ganz L. Man zeigt auch noch die lineare Unabhängigkeit von $(\tilde{\mathfrak{B}} \setminus \{\boldsymbol{x}_1\}) \cup \{\boldsymbol{b}_1\}$, womit $\tilde{\mathfrak{B}}_1$ wirklich eine Basis von L ist. In gleicher Weise erledigen wir die anderen Schritte. Nun hatten wir aber vorausgesetzt, daß L überhaupt eine endliche Basis hat. Damit wissen wir also, daß alle Basen im R^n aus genau n Elementen bestehen. Ein linearer Teilraum L des R^n kann aber auch wirklich keine unendliche Basis $\mathfrak{B}$ haben. Denn wir können $\mathfrak{B}$ zu einer Basis von ganz R^n ergänzen, indem wir $\boldsymbol{e}_1$ im Fall $\boldsymbol{e}_1 \notin L(\mathfrak{B})$ zu $\mathfrak{B}$ hinzufügen. Dann ist $\mathfrak{B}_1 = \mathfrak{B} \cup \{\boldsymbol{e}_1\}$ linear unabhängig; ist $\boldsymbol{e}_1 \in L(\mathfrak{B})$, so sei $\mathfrak{B}_1 = \mathfrak{B}$. Ebenso fahren wir mit $\boldsymbol{e}_2$ fort und erhalten $\mathfrak{B}_2$. Schließlich gelangen wir zu einer linear unabhängigen Menge $\mathfrak{B}_n$, die ganz R^n erzeugt, denn die $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ tun es. Also ist $|\mathfrak{B}_n| = n$, d. h., wegen $\mathfrak{B} \subseteq \mathfrak{B}_n$ muß $|\mathfrak{B}| \leqq n$ sein.[1])

[1]) Eine Basis von L konstruiert man sich durch sukzessive Erweiterung einer linear unabhängigen Menge $U \subset L$. Man wähle ein beliebiges Element von $L \setminus L(U)$ und füge es zu U hinzu. Mit der entstehenden (linear unabhängigen!) Menge fahre man fort. Dieser Prozeß liefert nach endlich vielen Schritten eine maximale linear unabhängige Menge in L.

Definition 3 (Dimension von linearen Teilräumen des $\mathbf{R}^n$). Es sei L ein linearer Teilraum des $\mathbf{R}^n$. Ist L der Nullraum, so sagt man: L hat die *Dimension* 0. Ist L vom Nullraum verschieden, so besitzt er eine endliche Basis aus m Elementen (alle seine Basen haben die gleiche Anzahl von Elementen). Man sagt: L hat die *Dimension* m. Man schreibt $\dim L$ für Dimension von L. Demnach gilt für $L \subseteqq \mathbf{R}^n$:

$$\dim L = 0 \Leftrightarrow L = \{0\},$$

$$\dim L = m \Leftrightarrow L \text{ hat eine Basis aus } m \text{ Elementen.}$$

Bemerkungen.

1. Infolge dieser Begriffsbildung haben wir also $\dim \mathbf{R}^n = n$. Unsere neue Begriffsbildung stimmt demnach mit der schon laufend gebrauchten Bezeichnungsweise „n-dimensionaler reeller Zahlenraum" vollkommen überein.

2. Nach dem letzten Teil der Beweisausführungen zum vorstehenden Satz gilt für jeden linearen Teilraum L von $\mathbf{R}^n$ $\dim L \leqq n$. Etwas allgemeiner läßt sich die dortige Argumentation in gleicher Weise auf den Fall anwenden, daß man zwei lineare Teilräume L_1, L_2 von $\mathbf{R}^n$ hat mit $L_1 \subseteqq L_2$. Wir formulieren demzufolge das zusätzliche Ergebnis.

Satz 3 (Monotonie der Dimension für lineare Teilräume des $\mathbf{R}^n$). *Es seien L_1, L_2 zwei lineare Teilräume des $\mathbf{R}^n$. Dann gilt:*

$$L_1 \subseteqq L_2 \Rightarrow \dim L_1 \leqq \dim L_2.$$

Mittels des Dimensionsbegriffes können wir die vorhin offen gelassene Frage beantworten, welche linearen Teilräume des $\mathbf{R}^n$ gerade als Kerne von Linearformen auftreten.

Satz 4 (Beschreibung der Kerne von Linearformen auf dem $\mathbf{R}^n$). *Es sei f eine nichtausgeartete Linearform auf dem $\mathbf{R}^n$. Dann ist der Kern dieser Linearform ein $(n-1)$-dimensionaler linearer Teilraum von $\mathbf{R}^n$:*

$$\dim \ker f = n - 1.$$

Jeder $(n-1)$-dimensionale lineare Teilraum des $\mathbf{R}^n$ ist auch Kern einer gewissen nichtausgearteten Linearform auf dem $\mathbf{R}^n$. Zwei nichtausgeartete Linearformen haben genau dann denselben Kern, wenn sie linear abhängig sind.

Bemerkung. Der vorstehende Satz bedeutet eine erste Aussage über die vollständige Beschreibung von Lösungsgesamtheiten linearer Gleichungen. Eine solche Umformulierung ist wegen des Zusammenhangs zwischen Linearformen und linearen Gleichungen möglich. Die Lösungsgesamtheit der linearen Gleichung $a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$ in den n Unbekannten $x_1, x_2, \ldots, x_n$ und den gegebenen Koeffizienten $a_1, a_2, \ldots, a_n$ (nicht sämtlich gleich Null) ist ein $(n-1)$-dimensionaler linearer Teilraum des $\mathbf{R}^n$. Zu jedem $(n-1)$-dimensionalen linearen Teilraum L des $\mathbf{R}^n$ gibt es eine lineare Gleichung $a_1x_1 + \cdots + a_nx_n = 0$, deren Lösungsgesamtheit gerade der gegebene lineare Teilraum L ist. Zwei nichtentartete lineare Gleichungen $a_1x_1 + \cdots + a_nx_n = 0$ und $b_1x_1 + b_2x_2 + \cdots + b_nx_n = 0$ haben genau dann dieselbe

Lösungsgesamtheit L im R^n, wenn die Koeffiziententupel $(a_1, a_2, \ldots, a_n)$, $(b_1, b_2, \ldots, b_n)$ linear abhängig sind.

Beweis des Satzes. f sei nicht ausgeartet, also ist $\ker f \neq \mathsf{R}^n$. Demzufolge gibt es ein Element $\boldsymbol{x}_0 \notin \ker f$. Wir müssen zwei Fälle unterscheiden: dim $\ker f = 0$ und dim $\ker f > 0$.

Bei $\boldsymbol{x} \in \mathsf{R}^n$ gilt stets

$$f\left(\boldsymbol{x} - \frac{f(\boldsymbol{x})}{f(\boldsymbol{x}_0)}\boldsymbol{x}_0\right) = f(\boldsymbol{x}) - \frac{f(\boldsymbol{x})}{f(\boldsymbol{x}_0)} f(\boldsymbol{x}_0) = 0$$

(es ist $f(\boldsymbol{x}_0) \neq 0$).

Im ersten Fall muß also

$$\boldsymbol{x} - \frac{f(\boldsymbol{x})}{f(\boldsymbol{x}_0)}\boldsymbol{x}_0 = 0$$

sein, d. h., $\{\boldsymbol{x}_0\}$ ist eine Basis des R^n.

Im zweiten Fall wählen wir eine Basis $\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_m$ von $\ker f$. Es ist dann $\{\boldsymbol{x}_0, \boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_m\}$ linear unabhängig. Außerdem gilt wieder für beliebige $\boldsymbol{x} \in \mathsf{R}^n$

$$\boldsymbol{x} - \frac{f(\boldsymbol{x})}{f(\boldsymbol{x}_0)}\boldsymbol{x}_0 \in \ker f.$$

Daher läßt sich $\boldsymbol{x}$ stets aus $\boldsymbol{x}_0, \boldsymbol{b}_1, \ldots, \boldsymbol{b}_m$ linear kombinieren. $\{\boldsymbol{x}_0, \boldsymbol{b}_1, \ldots, \boldsymbol{b}_m\}$ ist also Basis von R^n, d. h. $m = n - 1$.

Es sei L ein $(n-1)$-dimensionaler Teilraum des R^n. Der Fall $n = 1$ ist klar. Im Fall $n > 1$ existiert eine Basis $\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_m$ von L. Wie aus dem Beweis des Satzes über die Gleichmächtigkeit der Basen ersichtlich, können wir $\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_m$ durch Hinzunahme eines gewissen Elementes $\boldsymbol{x}_0 \in \mathsf{R}^n$ zu einer Basis des R^n ergänzen. Es gehören genau die $\boldsymbol{x} = \alpha_0\boldsymbol{x}_0 + \alpha_1\boldsymbol{b}_1 + \cdots + \alpha_m\boldsymbol{b}_m$ zu L, für die $\alpha_0 = 0$ ist. Als Linearform mit dem Kern L haben wir daher f: $\mathsf{R}^n \to \mathsf{R}$ mit $f(\boldsymbol{x}) = \alpha_0$. (Man muß dabei natürlich bemerken, daß für jedes $\boldsymbol{x} \in \mathsf{R}^n$ der Wert α_0 und auch die Werte $\alpha_1, \ldots, \alpha_m$ eindeutig bestimmt sind, was aus der linearen Unabhängigkeit von $\boldsymbol{x}_0, \boldsymbol{b}_1, \ldots, \boldsymbol{b}_m$ folgt.)

Zum letzten Teil: Sind die beiden nichtausgearteten Linearformen $f, g \in \mathscr{L}(\mathsf{R}^n)$ linear abhängig, so gilt $f = \alpha g$ bei einem gewissen $\alpha \neq 0$ aus R. Also hat man $f(\boldsymbol{x}) = 0$ genau dann, wenn $g(\boldsymbol{x}) = 0$ ist, d. h. $\ker f = \ker g$. Umgekehrt sei $\ker f = \ker g$. Die Situation $n = 1$ ist wieder klar, da je zwei Linearformen im $\mathscr{L}(\mathsf{R}^1)$ linear abhängig sind. Im Fall $n > 1$ wählen wir wieder eine Basis $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_m$ $(m = n - 1)$ von $\ker f = \ker g$ und ergänzen sie durch ein $\boldsymbol{x}_0$ wie oben zu einer Basis des R^n. Dann haben wir für jedes $\boldsymbol{x} \in \mathsf{R}^n$ eine Darstellung $\boldsymbol{x} = \alpha_0\boldsymbol{x}_0 + \alpha_1\boldsymbol{b}_1 + \cdots + \alpha_m\boldsymbol{b}_m$. Dann ist

$$f(\boldsymbol{x}) = \alpha_0 f(\boldsymbol{x}_0) + f(\alpha_1\boldsymbol{b}_1 + \cdots + \alpha_m\boldsymbol{b}_m) = \alpha_0 f(\boldsymbol{x}_0),$$

$$g(\boldsymbol{x}) = \alpha_0 g(\boldsymbol{x}_0) + g(\alpha_1\boldsymbol{b}_1 + \cdots + \alpha_m\boldsymbol{b}_m) = \alpha_0 g(\boldsymbol{x}_0),$$

d. h.

$$f(x) = \frac{f(x_0)}{g(x_0)}\, g(x) \qquad (g(x_0) \neq 0).$$

Bemerkung. Der angegebene Satz machte eine allgemeine Strukturaussage über die Lösungsmannigfaltigkeit einer Gleichung $a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$, aber er liefert kein Verfahren, wie diese Lösungsgesamtheit gewonnen wird. Solch ein Lösungsverfahren ist aber auch unschwer zu erkennen. Man wähle einen beliebigen Koeffizienten $a_i \neq 0$, o. B. d. A. sei das etwa a_1. Dann ist

$$x_1 = -\frac{a_2}{a_1}\, x_2 - \frac{a_3}{a_1}\, x_3 - \cdots - \frac{a_n}{a_1}\, x_n. \tag{*}$$

Die Lösungsgesamtheit der betrachteten Gleichung erhält man also, indem man $x_2, x_3, \ldots, x_n$ beliebig wählt und x_1 dann gemäß der Beziehung (*) bestimmt. Die so gewonnenen $(x_1, x_2, \ldots, x_n)$ sind genau die sämtlichen Lösungen der Gleichung. Eine Basis des Lösungsraumes erhält man etwa, indem man die frei wählbaren $x_2, x_3, \ldots, x_n$ beispielsweise nacheinander jeweils als 1 und die restlichen dabei als 0 wählt. Also wäre eine Basis für den Lösungsraum:

$$\begin{aligned}
\boldsymbol{b}_1 &= \left(-\frac{a_2}{a_1},\ 1, 0, \ldots, 0\right),\\
\boldsymbol{b}_2 &= \left(-\frac{a_3}{a_1},\ 0, 1, \ldots, 0\right),\\
&\cdots\cdots\cdots\cdots\cdots\cdots\\
\boldsymbol{b}_{n-1} &= \left(-\frac{a_n}{a_1},\ 0, 0, \ldots, 1\right).
\end{aligned}$$

Dieses sind nämlich $n - 1$ linear unabhängige Lösungen der Gleichung. (Man hätte den ersten Teil des Struktursatzes auch so beweisen können, daß man von den eben angegebenen Elementen zeigt: Es sind $n - 1$ linear unabhängige Lösungen von

$$f(x) = a_1x_1 + \cdots + a_nx_n = 0,$$

und jede Lösung dieser Gleichung läßt sich aus ihnen linear kombinieren.) Der Wert des Struktursatzes besteht in seiner klärenden Funktion.

4.3. Koordinatendarstellungen in bezug auf Basen

Im Beweisverlauf des Struktursatzes über die Kerne der Linearformen haben wir schon einen Eindruck von der Nützlichkeit der Darstellung von Elementen an Hand geeigneter Basen bekommen. Eine solche Situation wird uns noch vielfach auch im

Zusammenhang mit anderen Fragen begegnen. Hier zunächst noch einige allgemeine Bemerkungen.

Satz 1 (Kennzeichnung der Basen durch eindeutige Darstellbarkeit von Elementen). *Es sei L ein vom Nullraum verschiedener linearer Teilraum des $\mathbb{R}^n$. Weiter sei $\mathfrak{B}: \boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_m$ ein endliches Erzeugendensystem von L. Dann gilt: $\mathfrak{B}$ ist eine Basis von $L \Leftrightarrow$ Jedes $\boldsymbol{x} \in L$ hat eine Darstellung*

$$\boldsymbol{x} = \xi_1 \boldsymbol{b}_1 + \xi_2 \boldsymbol{b}_2 + \cdots + \xi_m \boldsymbol{b}_m$$

mit eindeutig bestimmten Koeffizienten ξ_i.

Beweis. Es sei $\mathfrak{B}$ eine Basis von L und

$$\boldsymbol{x} = \xi_1 \boldsymbol{b}_1 + \xi_2 \boldsymbol{b}_2 + \cdots + \xi_m \boldsymbol{b}_m = \eta_1 \boldsymbol{b}_1 + \eta_2 \boldsymbol{b}_2 + \cdots + \eta_m \boldsymbol{b}_m.$$

Dann ist $(\xi_1 - \eta_1)\, \boldsymbol{b}_1 + \cdots + (\xi_m - \eta_m)\, \boldsymbol{b}_m = \boldsymbol{0}$. Wegen der linearen Unabhängigkeit von $\mathfrak{B}$ muß $\xi_1 - \eta_1 = 0, \ldots, \xi_m - \eta_m = 0$ gelten, d. h., die Darstellung für $\boldsymbol{x}$ ist eindeutig.

Die Darstellung sei für jedes $\boldsymbol{x}$ eindeutig. Also muß sich bei $\xi_1 \boldsymbol{b}_1 + \xi_2 \boldsymbol{b}_2 + \cdots + \xi_m \boldsymbol{b}_m = \boldsymbol{0}$ ergeben, daß $\xi_1 = \xi_2 = \cdots = \xi_m = 0$ ist, denn $0\boldsymbol{b}_1 + 0\boldsymbol{b}_2 + \cdots + 0\boldsymbol{b}_m$ ist eine Darstellung von $\boldsymbol{0}$. Also ist $\mathfrak{B}$ ein linear unabhängiges Erzeugendensystem von L.

Definition 1 (Koordinaten bezüglich einer Basis). Es sei L ein vom Nullraum verschiedener linearer Teilraum des $\mathbb{R}^n$ mit der Basis $\mathfrak{B}: \boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_m$ $(1 \leqq m \leqq n)$. Diese Basis werde in einer festen Anordnung betrachtet. Die eindeutig bestimmte Darstellung

$$\boldsymbol{x} = \xi_1 \boldsymbol{b}_1 + \xi_2 \boldsymbol{b}_2 + \cdots + \xi_m \boldsymbol{b}_m$$

für jedes $\boldsymbol{x} \in L$ mit $\xi_i \in \mathbb{R}$ nennt man *Koordinatendarstellung* von $\boldsymbol{x}$ bezüglich der (angeordneten) Basis $\mathfrak{B}: \boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_m$. $(\xi_1, \xi_2, \ldots, \xi_m)$ nennt man das *Koordinatentupel* von $\boldsymbol{x}$ bezüglich der (angeordneten) Basis $\mathfrak{B}$, ξ_1 die erste Koordinate von $\boldsymbol{x}$ bezüglich der Basis $\mathfrak{B}$, …, ξ_m die m-te Koordinate von $\boldsymbol{x}$ bezüglich der Basis $\mathfrak{B}$.

Bemerkung. Es sei $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ ein Element aus dem $\mathbb{R}^n$. Die x_i heißen die Koordinaten bzw. Komponenten von $\boldsymbol{x}$ ohne einen Zusatz hinsichtlich einer Basis. Oben ist die Koordinatensprechweise in bezug auf Basen eingeführt worden. Wir erkennen, daß die obigen x_i gerade die Koordinaten von $\boldsymbol{x}$ bezüglich der natürlichen Basis im $\mathbb{R}^n$ sind. (Dieser Sachverhalt hat gerade die Bezeichnungsweise natürliche Basis des $\mathbb{R}^n$ veranlaßt.)

Die arithmetischen Operationen — wie Addition und Multiplikation mit einem Skalar — für Elemente eines linearen Teilraumes können auch an den Koordinatentupeln bezüglich beliebiger Basen des Teilraumes ausgedrückt werden. Den genauen Sachverhalt beschreibt die folgende Aussage.

Satz 2 (Arithmetische Operationen in linearen Teilräumen beschrieben durch Koordinatendarstellungen). *Es sei L ein vom Nullraum verschiedener linearer Teil-*

raum des R^n *mit einer in einer bestimmten Anordnung fixierten Basis* $\mathfrak{B}: \boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_m$ $(1 \leqq m \leqq n)$. *Weiter seien* $\boldsymbol{x}$, $\boldsymbol{y}$ *beliebige Elemente von* L *mit den Koordinatentupeln* $(\xi_1, \xi_2, \ldots, \xi_m)$ *bzw.* $(\eta_1, \eta_2, \ldots, \eta_m)$ *in bezug auf* $\mathfrak{B}$. *Dann gilt:* $\boldsymbol{x} + \boldsymbol{y}$ *hat in bezug auf* $\mathfrak{B}$ *das Koordinatentupel* $(\xi_1 + \eta_1, \xi_2 + \eta_2, \ldots, \xi_m + \eta_m)$; $\alpha\boldsymbol{x}$, $\alpha \in \mathsf{R}$, *hat in bezug auf* $\mathfrak{B}$ *das Koordinatentupel* $(\alpha\xi_1, \alpha\xi_2, \ldots, \alpha\xi_m)$.

Beweis. Es ist vorausgesetzt

$$\boldsymbol{x} = \sum_{i=1}^{m} \xi_i \boldsymbol{b}_i, \qquad \boldsymbol{y} = \sum_{i=1}^{m} \eta_i \boldsymbol{b}_i.$$

Dann ergibt sich

$$\boldsymbol{x} + \boldsymbol{y} = \sum_{i=1}^{m} \xi_i \boldsymbol{b}_i + \sum_{i=1}^{m} \eta_i \boldsymbol{b}_i = \sum_{i=1}^{m} (\xi_i + \eta_i)\, \boldsymbol{b}_i.$$

Infolge der Eindeutigkeit der Koordinatendarstellung ist $\xi_i + \eta_i$ die i-te Koordinate von $\boldsymbol{x} + \boldsymbol{y}$ in bezug auf die Basis $\mathfrak{B}$. Der andere Teil erledigt sich entsprechend.

Bemerkung. Der Satz macht darauf aufmerksam, daß die arithmetische Struktur des m-dimensionalen linearen Teilraumes L eigentlich die gleiche ist wie die arithmetische Struktur des Raumes R^m. Auf diesen Sachverhalt waren wir schon bei unseren geometrisch-anschaulichen Betrachtungen der linearen Teilräume des R^2 bzw. R^3 an Hand seiner euklidischen Veranschaulichung hingelenkt worden. Der nächste Abschnitt befaßt sich hiermit etwas prägnanter.

4.4. Isomorphie linearer Teilräume

Die im abschließenden Satz des letzten Abschnittes dargelegten Verhältnisse werden deutlicher, wenn wir sie uns noch einmal abbildungstheoretisch vor Augen führen.

L war ein m-dimensionaler Teilraum des R^n mit einer gewissen angeordneten Basis $\mathfrak{B}$. Diese Basis $\mathfrak{B}$ bestimmte über die Koordinatendarstellung der Elemente von L bezüglich $\mathfrak{B}$ eine Abbildung $\Phi: L \to \mathsf{R}^m$ mit dem Verlauf $\boldsymbol{x} \to (\xi_1, \xi_2, \ldots, \xi_m)$, sofern $\boldsymbol{x} = \sum_{i=1}^{m} \xi_i \boldsymbol{b}_i$ $(\mathfrak{B}: \boldsymbol{b}_1, \ldots, \boldsymbol{b}_m)$ ist. Diese Abbildung hat die folgenden Eigenschaften:

1. Φ ist eineindeutig, d. h. $\boldsymbol{x} \neq \boldsymbol{y} \Rightarrow \Phi(\boldsymbol{x}) \neq \Phi(\boldsymbol{y})$.
2. Φ ist surjektiv, d. h. $\Phi(L) = \mathsf{R}^m$ (jedes Element von R^m tritt als Bild unter Φ auf).
3. Φ ist linear, d. h., für $\boldsymbol{x}, \boldsymbol{y} \in L$, $\alpha, \beta \in \mathsf{R}$ gilt

$$\Phi(\alpha\boldsymbol{x} + \beta\boldsymbol{y}) = \alpha\Phi(\boldsymbol{x}) + \beta\Phi(\boldsymbol{y}).$$

In Diagrammform: $x, y \in L$, $\Phi(x), \Phi(y) \in \mathsf{R}^m$

Definition 1 (Isomorphismus linearer Teilräume). Es sei L ein linearer Teilraum des R^n und L' ein linearer Teilraum des R^m. Eine eineindeutige surjektive Abbildung $\Phi: L \to L'$ von L auf L' heißt ein *Isomorphismus*[1]) von L auf L' (oder zwischen L und L') genau dann, wenn Φ eine lineare Abbildung von L auf L' ist, d. h., wenn also gilt:

$$\Phi(\alpha x + \beta y) = \alpha\Phi(x) + \beta\Phi(y)$$

für alle $x, y \in L$ und $\alpha, \beta \in \mathsf{R}$. Besteht ein Isomorphismus zwischen L und L', so heißen diese linearen Teilräume zueinander *isomorph*.

Bemerkung. Der Isomorphiebegriff für lineare Teilräume bringt also zum Ausdruck, daß die betrachteten Teilräume dieselbe arithmetische Gestalt haben. In ihrer arithmetischen Struktur unterscheiden sie sich nicht, wenn man sie unter geeignetem Blickwinkel (d. h. unter gegenseitiger Bezugnahme mittels der Abbildung Φ betrachtet). Der Isomorphiebegriff ist nicht nur für lineare Teilräume des R^n bedeutungsvoll, sondern allgemein in der Mathematik für irgendwelche Strukturen von Wichtigkeit. Das kommt in der allgemeinen Algebra noch verschiedentlich ausführlicher zur Sprache. Für uns ergibt sich hier rückblickend auf die Linearformen, daß wir eine Isomorphiebeziehung auch zwischen $\mathscr{L}(\mathsf{R}^n)$ und R^n vorliegen haben, ohne daß es sich dabei im Fall des $\mathscr{L}(\mathsf{R}^n)$ von vornherein um einen Teilraum eines R^n handelt (vgl. den Satz über die punktweisen Operationen für Linearformen, widergespiegelt an ihren Koeffiziententupeln).

Wir formulieren unsere gewonnenen Einsichten.

Satz 1 (Isomorphie der linearen Teilräume des R^n zu gewissen R^m). *Es sei L ein linearer Teilraum des R^n mit positiver Dimension $m \geqq 1$. Dann wird durch jede angeordnete Basis $\mathfrak{B}$ von L in kanonischer Weise eine Isomorphie zwischen L und dem m-dimensionalen arithmetischen Zahlenraum R^m hergestellt*:

$$L \overset{\Phi_{\mathfrak{B}}}{\leftrightarrow} \mathsf{R}^m.$$

[1]) Isomorphismus: Iso (gleich) morph (Gestalt). Man vergleiche hierzu Isomorphiebetrachtungen in Band 1.

Satz 2 (Isomorphiekriterium mittels Dimension). *Es seien* $L \subseteq \mathbb{R}^n$, $L' \subseteq \mathbb{R}^m$ *lineare Teilräume des* $\mathbb{R}^n$ *bzw.* $\mathbb{R}^m$. *Dann gilt: Es gibt einen Isomorphismus zwischen* L *und* $L' \Leftrightarrow \dim L = \dim L'$.

4.5. Übungsaufgaben

1. Warum ist eine Übertragung des für den $\mathbb{R}^n$ erklärten Begriffs der linearen Unabhängigkeit auf den Bereich $\mathcal{P}$ aller reellen Polynome, ausgestattet mit der üblichen Funktionenaddition und Multiplikation mit reellen Skalaren, sinnvoll (vgl. hierzu Übungsaufgabe 1 zu 2.4.)? Man weise nach, daß die Menge aller Elementarpolynome $P_n(x) = x^n$, $n \geqq 0$, linear unabhängig in $\mathcal{P}$ ist!
2. Zwei Elemente $\boldsymbol{x} = (x_1, x_2)$, $\boldsymbol{y} = (y_1, y_2)$ des $\mathbb{R}^2$ sind genau dann linear abhängig, wenn $x_1 y_2 = x_2 y_1$ gilt.
3. Folgt im $\mathbb{R}^3$ für die drei Elemente $\boldsymbol{x} = (x_1, x_2, x_3)$, $\boldsymbol{y} = (y_1, y_2, y_3)$ und $\boldsymbol{z} = (z_1, z_2, z_3)$ stets aus dem Bestehen der Beziehung $x_1 y_1 z_3 = x_2 y_3 z_1 = x_3 y_1 z_2$ die lineare Abhängigkeit von $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$?
4. Wie müssen drei allgemein gegebene Elemente $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$ des $\mathbb{R}^3$ beschaffen sein, damit die aus den Elementen gebildete Menge linear unabhängig wird (notwendige und hinreichende Bedingungen)?
5. Man weise die Elemente $\boldsymbol{x} = (1, 2, 3)$, $\boldsymbol{y} = (4, 5, 6)$ des $\mathbb{R}^3$ als linear unabhängig nach und ergänze sie durch ein weiteres Element zu einer Basis des $\mathbb{R}^3$.
6. L_1 und L_2 seien zwei lineare Teilräume des $\mathbb{R}^n$, wobei $\mathfrak{B}_1$ eine Basis für L_1 sei und $\mathfrak{B}_2$ eine Basis für L_2. Man gebe Bedingungen an Hand von $L_1 \cap L_2$ und dim L_1, dim L_2 dafür an, daß $\mathfrak{B}_1 \cup \mathfrak{B}_2$ eine Basis für $\mathbb{R}^n$ wird.
7. L_1 und L_2 seien zwei lineare Teilräume des $\mathbb{R}^n$. Was für eine Beziehung läßt sich zwischen den Werten dim $(L(L_1 \cup L_2))$, dim L_1, dim L_2 und dim $(L_1 \cap L_2)$ ableiten?
8. Es sei L ein echter linearer Teilraum des $\mathbb{R}^n$. Gibt es einen Isomorphismus von L auf den $\mathbb{R}^n$?
9. In $\mathbb{R}^3$ werde die folgende Abbildung betrachtet: $A: \mathbb{R}^3 \to \mathbb{R}^3$ mit

$$(x_1, x_2, x_3) \xrightarrow[A]{} (-x_3, 2x_2, x_1)$$

für alle $(x_1, x_2, x_3) \in \mathbb{R}^3$. Ist diese Abbildung ein Isomorphismus des $\mathbb{R}^3$ auf sich?

10. Die Elemente

$$\boldsymbol{x} = (\cos\alpha, \sin\alpha), \quad \boldsymbol{y} = (-\sin\alpha, \cos\alpha)$$

des $\mathbb{R}^2$ weise man für beliebiges $\alpha \in \mathbb{R}$ als linear unabhängig nach.
Man bestimme zu den Elementen (1, 0), (0, 1), (1, 1) jeweils die Koordinatendarstellungen in bezug auf die Basis $\mathfrak{B}: \boldsymbol{x}, \boldsymbol{y}$.

5. Lösungsmannigfaltigkeiten linearer Gleichungssysteme

5.1. Die verschiedenen Typen linearer Gleichungssysteme

Es sei

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}\right\} \qquad (*)$$

oder in Funktionalform

$$\begin{aligned} f_1(x_1, x_2, \ldots, x_n) &= b_1, \\ f_2(x_1, x_2, \ldots, x_n) &= b_2, \\ \cdots\cdots\cdots\cdots & \\ f_m(x_1, x_2, \ldots, x_n) &= b_m \end{aligned}$$

ein lineares Gleichungssystem in den n Unbekannten $x_1, x_2, \ldots, x_n$ mit den gegebenen Koeffiziententupeln

$$\begin{aligned} &(a_{11}, a_{12}, \ldots, a_{1n}) && \text{(Koeffiziententupel für die Linearform } f_1), \\ &(a_{21}, a_{22}, \ldots, a_{2n}) && \text{(Koeffiziententupel für die Linearform } f_2), \\ &\cdots\cdots\cdots\cdots & & \\ &(a_{m1}, a_{m2}, \ldots, a_{mn}) && \text{(Koeffiziententupel für die Linearform } f_m) \end{aligned}$$

und den gegebenen rechten Seiten $b_1, \ldots, b_m$ (auch freie Glieder des Gleichungssystems genannt).

Definition 1 (Homogenes, inhomogenes, lösbares und unlösbares lineares Gleichungssystem). Ein lineares Gleichungssystem (*) heißt *homogenes lineares Gleichungssystem* genau dann, wenn die gegebenen rechten Seiten sämtlich gleich Null sind. Ein lineares Gleichungssystem (*) heißt *inhomogenes lineares Gleichungs-*

system genau dann, wenn die gegebenen rechten Seiten nicht sämtlich gleich Null sind. Ein lineares Gleichungssystem (*) heißt *lösbar* genau dann, wenn es wenigstens eine Lösung hat, d. h., wenn es wenigstens ein n-Tupel $\boldsymbol{x} = (x_1, \ldots, x_n)$ gibt, welches in die linken Seiten eingesetzt die rechten ergibt. Ein lineares Gleichungssystem (*) heißt *unlösbar* genau dann, wenn es keine Lösung hat (wenn es nicht lösbar ist).

Bemerkungen.

1. Ein homogenes lineares Gleichungssystem aus m Gleichungen mit n Unbekannten ist stets lösbar, denn es hat zumindest die Lösung $\boldsymbol{x} = (0, 0, \ldots, 0)$ — die sogenannte *triviale Lösung*.

2. Es gibt auch wirklich unlösbare Gleichungssysteme, was beispielsweise die folgenden einfachen Systeme zeigen.

$$x_1 = 1, \qquad x_1 = 2 \tag{1}$$

oder

$$\begin{aligned} x_1 + x_2 &= 1, \\ 2x_1 + 2x_2 &= 3. \end{aligned} \tag{2}$$

Definition 2 (Bestimmte und unbestimmte lineare Gleichungssysteme). Ein lineares Gleichungssystem (*) heißt ein *bestimmtes lineares Gleichungssystem* genau dann, wenn es eine einzige Lösung hat, wenn die Lösungsmenge aus genau einem n-Tupel $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ besteht. Besteht die Lösungsmenge von (*) aus mehreren n-Tupeln, so heißt das Gleichungssystem (*) *unbestimmt*.

Definition 3 (Das zu einem linearen Gleichungssystem gehörende homogene lineare Gleichungssystem). Das *zu dem linearen Gleichungssystem (*) gehörende homogene lineare Gleichungssystem* ist dasjenige, in dem die rechten Seiten alle durch Null ersetzt werden.

Definition 4 (Äquivalenz von linearen Gleichungssystemen). Es seien zwei lineare Gleichungssysteme in n Unbekannten gegeben, dabei bestehe das eine System aus m Gleichungen und das andere aus l Gleichungen.

$$\left.\begin{aligned} f_1(x_1, x_2, \ldots, x_n) &= b_1, \\ &\cdots\cdots\cdots\cdots \\ f_m(x_1, x_2, \ldots, x_n) &= b_m, \end{aligned}\right\} \tag{*}$$

$$\left.\begin{aligned} g_1(x_1, x_2, \ldots, x_n) &= c_1, \\ &\cdots\cdots\cdots\cdots \\ g_l(x_1, x_2, \ldots, x_n) &= c_l. \end{aligned}\right\} \tag{**}$$

Diese beiden linearen Gleichungssysteme (*) und (**) heißen *äquivalent* genau dann, wenn die Lösungsmengen (Teilmengen des $\mathbb{R}^n$) beider Systeme übereinstimmen.

Bemerkung. Im Fall der einfachen Systeme

$$f_1(x_1, x_2, \ldots, x_n) = 0 \tag{*}$$
$$g_1(x_1, x_2, \ldots, x_n) = 0, \tag{**}$$

wobei die Linearformen f_1 und g_1 beide nicht entartet waren, hatten wir schon im Satz über die Beschreibung der Kerne von Linearformen herausgefunden, daß beide Systeme genau dann äquivalent sind, wenn f_1, g_1 linear abhängig sind.

5.2. Lösungsmannigfaltigkeit homogener linearer Gleichungssysteme. Der Rang eines linearen Gleichungssystems

Über homogene lineare Gleichungssysteme in n Unbekannten im allgemeinen wissen wir bisher lediglich, daß die Lösungsmenge ein linearer Teilraum des $\mathbf{R}^n$ ist (vgl. S. 29 und S. 31). Unsere nächste Aufgabe wird in der Bestimmung der Dimension dieses Lösungsraumes bestehen.

Hilfssatz 1 (Trennbarkeit von linearen Teilräumen und Punkten durch Linearformen). *Es sei L ein linearer Teilraum des $\mathbf{R}^n$ und $\boldsymbol{x}$ ein Punkt des $\mathbf{R}^n$, der nicht zu L gehört. Dann gibt es eine Linearform $f \in \mathscr{L}(\mathbf{R}^n)$ mit $f(\boldsymbol{x}) \neq 0$ und $f(\boldsymbol{y}) = 0$ für alle $\boldsymbol{y} \in L$.*

Beweis. Es können die Fälle $\dim L = 0$ oder $\dim L > 0$ vorliegen. Im ersten Fall wähle man eine beliebige Basis $\mathfrak{B}$: $\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_n$ von $\mathbf{R}^n$ und betrachte die Koordinatendarstellung von $\boldsymbol{x}$ bezüglich $\mathfrak{B}$

$$\boldsymbol{x} = \xi_1 \boldsymbol{b}_1 + \xi_2 \boldsymbol{b}_2 + \cdots + \xi_n \boldsymbol{b}_n .$$

Wenigstens ein ξ_i ist von Null verschieden, es sei dies o. B. d. A. ξ_1. Dann nehme man als $f \in \mathscr{L}(\mathbf{R}^n)$ die Abbildung $f\colon \mathbf{R}^n \to \mathbf{R}$ mit $\boldsymbol{z} \mapsto$ erste Koordinate von $\boldsymbol{z}$ bezüglich der Basis $\mathfrak{B}$. Im Fall $\dim L > 0$ wähle man eine Basis $\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_m$ von L und ergänze diese zu einer Basis $\mathfrak{B}$: $\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_m, \ldots, \boldsymbol{b}_n$ von ganz $\mathbf{R}^n$. Dann ist

$$\boldsymbol{x} = \xi_1 \boldsymbol{b}_1 + \cdots + \xi_m \boldsymbol{b}_m + \cdots + \xi_n \boldsymbol{b}_n .$$

Es muß jetzt $\xi_i \neq 0$ für ein $i > m$ sein, weil sonst $\boldsymbol{x} \in L$ ist. Es sei o. B. d. A. $\xi_n \neq 0$. Dann nehme man als $f \in \mathscr{L}(\mathbf{R}^n)$ die Abbildung $f\colon \mathbf{R}^n \to \mathbf{R}$ mit $\boldsymbol{z} \mapsto n$-te Koordinate von $\boldsymbol{z}$ bezüglich $\mathfrak{B}$.

Mit diesem Hilfssatz gelangen wir zu der folgenden wichtigen Einsicht.

Satz 1 (Äquivalenz zweier homogener linearer Gleichungssysteme). *Es seien zwei homogene lineare Gleichungssysteme in n Unbekannten gegeben:*

$$\left.\begin{aligned} f_1(x_1, x_2, \ldots, x_n) &= 0, \\ &\cdots\cdots\cdots\cdots \\ f_m(x_1, x_2, \ldots, x_n) &= 0, \end{aligned}\right\} \quad (*)$$

$$\left.\begin{aligned} g_1(x_1, x_2, \ldots, x_n) &= 0, \\ &\cdots\cdots\cdots\cdots \\ g_l(x_1, x_2, \ldots, x_n) &= 0. \end{aligned}\right\} \quad (**)$$

Dann gilt: Diese beiden Systeme sind äquivalent ⇔ Die das Gleichungssystem (*) bestimmenden Linearformen $f_1, \ldots, f_m$ und die das Gleichungssystem (**) bestimmenden Linearformen $g_1, \ldots, g_l$ erzeugen im Raum $\mathscr{L}(\mathsf{R}^n)$ aller Linearformen auf dem R^n ein und denselben linearen Teilraum.

Beweis. Es sei L_1 der Lösungsraum von (*). Wir zeigen für eine Linearform $f \in \mathscr{L}(\mathsf{R}^n)$:

$$f \in L(\{f_1, f_2, \ldots, f_m\}) \Leftrightarrow f(\boldsymbol{x}) = 0 \quad \text{für alle } \boldsymbol{x} \in L_1.$$

Wenn $f \in L(\{f_1, \ldots, f_m\})$ ist, hat man eine Darstellung

$$f = \alpha_1 f_1 + \cdots + \alpha_m f_m$$

mit $\alpha_i \in \mathsf{R}$. Also ist für alle $\boldsymbol{x} \in \mathsf{R}^n$

$$f(\boldsymbol{x}) = \alpha_1 f_1(\boldsymbol{x}) + \cdots + \alpha_m f_m(\boldsymbol{x}),$$

d. h., für alle $\boldsymbol{x} \in L_1$ muß $f(\boldsymbol{x}) = 0$ sein.

Für die umgekehrte Richtung machen wir von dem Hilfssatz Gebrauch. Wir nehmen an, es sei $f \notin L(\{f_1, \ldots, f_m\})$. Um den Hilfssatz unmittelbar anwenden zu können, gehen wir von der im $\mathscr{L}(\mathsf{R}^n)$ vorliegenden Situation durch den kanonischen Isomorphismus zwischen $\mathscr{L}(\mathsf{R}^n)$ und R^n (wo also jedes $f \in \mathscr{L}(\mathsf{R}^n)$ in sein Koeffiziententupel übergeführt wird) zu einer Situation im R^n über. Wir haben dann einen Teilraum $L \subseteq \mathsf{R}^n$, der dem $L(\{f_1, \ldots, f_m\})$ entspricht, und einen Punkt $\boldsymbol{y} \in \mathsf{R}^n$, der dem f entspricht, mit $\boldsymbol{y} \notin L$. Es gibt nach dem Hilfssatz eine Linearform $h \in \mathscr{L}(\mathsf{R}^n)$ mit $h(\boldsymbol{y}) \neq 0$ und $h \equiv 0$ auf L. Nun kehren wir wieder zur Situation in $\mathscr{L}(\mathsf{R}^n)$ zurück. Der Linearform h entspricht dann ein Punkt $\boldsymbol{z} \in \mathsf{R}^n$. Wir haben folgende Entsprechung:

$$L(\{f_1, \ldots, f_m\}) \leftrightarrow L, \qquad f \leftrightarrow \boldsymbol{y}, \qquad \boldsymbol{z} \leftrightarrow h;$$

$f \notin L(\{f_1, \ldots, f_m\}) \Leftrightarrow \boldsymbol{y} \notin L$, $h(\boldsymbol{y}) \neq 0$ und $h \equiv 0$ auf $L \Leftrightarrow f(\boldsymbol{z}) \neq 0$ und $g(\boldsymbol{z}) = 0$ für alle $g \in L(\{f_1, \ldots, f_m\})$. Also wäre für den ermittelten Punkt $\boldsymbol{z}$ erwiesen, daß $\boldsymbol{z} \in L_1$ ist. Dabei sollte aber dann auch $f(\boldsymbol{z}) = 0$ gelten. Die Annahme $f \notin L(\{f_1, \ldots, f_m\})$ besteht folglich zu unrecht.

Die Gesamtaussage des Satzes ist nun leicht zu folgern.

Satz 2 (Struktur der Lösungsmenge eines homogenen linearen Gleichungssystems). *Es sei ein homogenes lineares Gleichungssystem in n Unbekannten gegeben:*

$$\left.\begin{aligned} f_1(x_1, x_2, \ldots, x_n) &= 0, \\ &\cdots\cdots\cdots\cdots \\ f_m(x_1, x_2, \ldots, x_n) &= 0. \end{aligned}\right\} \qquad (*)$$

Die Lösungsmenge dieses linearen homogenen Gleichungssystems ist dann ein linearer Teilraum L des R^n. Die Dimension von L beträgt $n - \dim L(\{f_1, \ldots, f_m\})$.

Umgekehrt gibt es zu jedem linearen Teilraum L des $\mathbf{R}^n$ von einer Dimension $0 \leqq k < n$ ein homogenes lineares Gleichungssystem von m Gleichungen, dessen Lösungsraum gleich dem gegebenen L ist. Die das Gleichungssystem bestimmenden Linearformen $f_1, \ldots, f_m$ sind genau bei $m = n - k$ linear unabhängig.

Beweis. Daß die Lösungsmenge eines homogenen linearen Gleichungssystems ein linearer Teilraum L_0 des $\mathbf{R}^n$ ist, hatten wir schon früher festgestellt. Wir müssen die Dimension von L_0 bestimmen. Nach dem Beweis des vorhergehenden Satzes gilt:

$$f \in L(\{f_1, \ldots, f_m\}) \Leftrightarrow f(\boldsymbol{x}) = 0 \quad \text{für alle} \quad \boldsymbol{x} \in L_0.$$

Also haben wir $\dim L(\{f_1, \ldots, f_m\}) = n \Leftrightarrow L_0 = \{0\}$, denn zu jedem $\boldsymbol{x} \neq \mathbf{0}$ gibt es wenigstens eine Linearform, die in $\boldsymbol{x}$ einen von Null verschiedenen Wert annimmt.

Es sei jetzt $\dim L_0 = r > 0$. Dann wählen wir eine Basis $\mathfrak{B}: \boldsymbol{b}_1, \ldots, \boldsymbol{b}_r$ von L_0. Diese können wir zu einer Basis von ganz $\mathbf{R}^n$ ergänzen: $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_r, \ldots, \boldsymbol{b}_n$. Wir betrachten im Fall $r < n$ (bei $r = n$ muß jedes $f_1, \ldots, f_m$ die Nullform sein, da der Lösungsraum einer einzigen nichtausgearteten Linearform schon $(n-1)$-dimensional ist) die Funktionale $g_{r+1}, \ldots, g_n$ mit $g_i(\boldsymbol{x}) = \alpha_i$, wobei α_i die i-te Koordinate von $\boldsymbol{x}$ in der Koordinatendarstellung bezüglich der Basis $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_r, \boldsymbol{b}_{r+1}, \ldots, \boldsymbol{b}_n$ ist:

$$\boldsymbol{x} = \sum_{i=1}^{n} \alpha_i \boldsymbol{b}_i.$$

Die Linearformen $g_{r+1}, \ldots, g_n$ sind linear unabhängig. Außerdem gehören sie alle zu $L(\{f_1, \ldots, f_m\})$, denn für $\boldsymbol{x} \in L_0$ ist jedes g_i, $i \geqq r$, gleich Null. Diese g_i erzeugen auch ganz $L(\{f_1, \ldots, f_m\})$, weil für $f \in L(\{f_1, \ldots, f_m\})$ gilt

$$\begin{aligned} f(\boldsymbol{x}) &= f\left(\sum_{i=1}^{r} \alpha_i \boldsymbol{b}_i\right) + \sum_{i=r+1}^{n} \alpha_i f(\boldsymbol{b}_i) \\ &= \sum_{i=r+1}^{n} \alpha_i f(\boldsymbol{b}_i) \\ &= \sum_{i=r+1}^{n} f(\boldsymbol{b}_i)\, g_i(\boldsymbol{x}), \end{aligned}$$

d. h.

$$f = \sum_{i=r+1}^{n} f(\boldsymbol{b}_i)\, g_i.$$

Demnach ist $\dim L(\{f_1, \ldots, f_m\}) = n - r$.

Zur umgekehrten Situation: Es sei ein linearer Teilraum $L \subseteq \mathbf{R}^n$ mit $0 \leqq \dim L < n$ gegeben. Bei $\dim L = 0$ ist L der Lösungsraum eines Gleichungssystems mit beliebigen n linear unabhängigen Linearformen. Im Fall $\dim L > 0$ wählen wir wie vorher eine Basis von L, ergänzen diese zu einer vollen Basis von $\mathbf{R}^n$ und verfahren genauso wie im ersten Beweisteil, d. h., unsere gewünschten Linearformen erhalten wir als $f_i(\boldsymbol{x}) = \alpha_i$, $\dim L < i \leqq n$, bei $\boldsymbol{x} = \sum_{i=1}^{n} \alpha_i \boldsymbol{b}_i$.

Für die Dimension des von den Linearformen eines Gleichungssystems im $\mathscr{L}(\mathbf{R}^n)$ aufgespannten Teilraumes hat sich die folgende Sprechweise eingebürgert, die auch im Fall von inhomogenen linearen Gleichungssystemen benutzt wird.

Definition 1 (Rang eines linearen Gleichungssystems). Es sei ein beliebiges lineares Gleichungssystem in n Unbekannten $x_1, \ldots, x_n$, bestehend aus m Gleichungen gegeben:

$$f_1(x_1, \ldots, x_n) = b_1,$$
$$\ldots\ldots\ldots\ldots$$
$$f_m(x_1, \ldots, x_n) = b_m.$$

Unter dem *Rang* dieses Gleichungssystems versteht man die Dimension des von den Linearformen $f_1, \ldots, f_m$ im Raum $\mathscr{L}(\mathbf{R}^n)$ aller Linearformen über $\mathbf{R}^n$ aufgespannten linearen Teilraumes. Haben die Linearformen $f_1, \ldots, f_m$ die Koeffiziententupel $(a_{11}, \ldots, a_{1n}), \ldots, (a_{m1}, \ldots, a_{mn})$, so ist der *Rang des Gleichungssystems* also gleich der Maximalzahl der in der Menge der $(a_{11}, \ldots, a_{1n}), \ldots, (a_{m1}, \ldots, a_{mn})$ linear unabhängigen Elemente.

Bemerkung. Für ein homogenes lineares Gleichungssystem in n Unbekannten gilt also in der neuen Sprechweise: Dimension des Lösungsraumes des Gleichungssystems + Rang des Gleichungssystems $= n$.

Im übernächsten Abschnitt werden wir ein geeignetes Verfahren zur Rangberechnung und zur effektiven Angabe des Lösungsraumes von vorgelegten linearen Gleichungssystemen abhandeln. Unsere im Augenblick angestellten Erörterungen haben die wichtige Funktion der allgemeinen Strukturklärung. Wir können durch sie von einem überschauenden Standpunkt, der geometrisch-anschauliche Züge trägt, verfolgen, was eigentlich beim Lösen linearer Gleichungssysteme passiert, während das rein Rechnerische beim Lösungsprozeß solche Einsichten nicht mehr bietet. Im nächsten Abschnitt soll nun auseinandergesetzt werden, daß die geometrisch-anschauliche Übersicht über die Lösungsmengen von linearen Gleichungssystemen auch für nichthomogene Gleichungssysteme bestehen bleibt.

5.3. Lösungsmannigfaltigkeiten beliebiger linearer Gleichungssysteme. Lineare Mannigfaltigkeiten im $\mathbf{R}^n$

Die Lösungsmannigfaltigkeiten linearer homogener Gleichungssysteme in n Unbekannten waren lineare Teilräume des $\mathbf{R}^n$. Das ist für nichthomogene lineare Gleichungssysteme nicht mehr richtig, da dann die Summe von zwei Lösungen nicht wieder eine Lösung ist und das skalare Vielfache einer Lösung nur für den Wert $\alpha = 1$ wieder eine Lösung ist. Man sieht aber unmittelbar, daß die Differenz zweier Lö-

sungen eines nichthomogenen linearen Gleichungssystems eine Lösung des zugehörigen homogenen linearen Gleichungssystems ist. Bei Verträglichkeit des inhomogenen linearen Gleichungssystems entstehen auf diese Weise auch alle Lösungen des zugehörigen homogenen linearen Gleichungssystems.

Wir fixieren diese Sachverhalte durch die folgenden Definitionen und Sätze.

Definition 1 (Komplexsumme von Teilmengen des $\mathbb{R}^n$). Es seien A, B zwei beliebige nichtleere Teilmengen des $\mathbb{R}^n$. Unter der *Komplexsumme* dieser Teilmengen versteht man die Menge

$$A + B := \{\boldsymbol{a} + \boldsymbol{b} : \boldsymbol{a} \in A, \boldsymbol{b} \in B\}.$$

Entsprechend versteht man unter der *Komplexdifferenz* dieser Teilmengen die Menge

$$A - B := \{\boldsymbol{a} - \boldsymbol{b} : \boldsymbol{a} \in A, \boldsymbol{b} \in B\}.$$

Bemerkungen.

1. Diese Komplexsumme ist nicht mit der Mengenvereinigung zu verwechseln und ebenso die Komplexdifferenz nicht mit der mengentheoretischen Differenz.

2. Man vergleiche hierzu die Ausführungen der allgemeinen Algebra hinsichtlich der „Komplexprodukte". Das Wort Komplex soll darauf hindeuten, daß hier die Summe bzw. Differenz bzw. Verknüpfung im allgemeinen für ganze Komplexe (Mengen) von Elementen betrachtet wird.

3. In Abb. 4 sind dazu einige Fälle an Hand der euklidischen Veranschaulichung des $\mathbb{R}^2$ gezeigt.

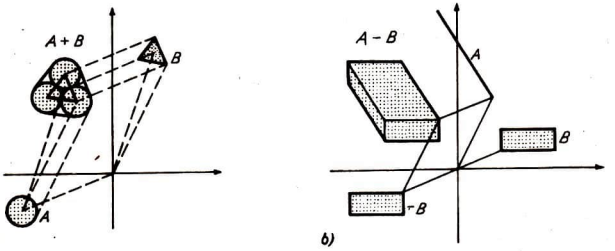


Abb. 4

Die Lösungsmenge M eines verträglichen inhomogenen linearen Gleichungssystems in n Unbekannten hat, wie vermerkt wurde, die Eigenschaft, daß die Komplexdifferenz von M mit sich selbst ein linearer Teilraum des $\mathbb{R}^n$ ist. Die Umkehrung hierzu gilt jedoch nicht, wie etwa im $\mathbb{R}^1$ das Beispiel $M = \{x : x \in \mathbb{R}^1, x \geqq 0\}$ zeigt.

Denn es ist $M - M = \mathsf{R}^1$ (nicht etwa $\{0\}$!), aber die Lösungsmengen von verträglichen linearen Gleichungssystemen mit einer Unbekannten sind einpunktig oder der ganze R^1. Für uns werden jetzt die Komplexsummen von Bedeutung sein, wo der eine Summand ein linearer Teilraum des R^n ist und der andere eine einpunktige Menge darstellt.

Wir schreiben anstelle der Komplexsumme $\{x\} + A$ kürzer und ebenso unmißverständlich $x + A$.

Definition 2 (Lineare Mannigfaltigkeiten im R^n). Unter einer *linearen Mannigfaltigkeit* M im R^n versteht man eine Teilmenge der Form $M = x_0 + L$, wobei x_0 ein festes Element des R^n und L ein linearer Teilraum des R^n ist.

Bemerkung. Man macht sich den Begriff der linearen Mannigfaltigkeit etwa wieder an der euklidischen Veranschaulichung des R^2 bzw. R^3 in seiner anschaulich-geometrischen Bedeutung klar. Die Teilmenge $M = x_0 + L$ ist aus L entstanden, indem man ganz L parallel verschoben hat, bis es durch x_0 geht. Demzufolge sind die linearen Mannigfaltigkeiten in R^2:

1. die einpunktigen Teilmengen des R^2 (L ist hier der Nullraum $\{0\}$),
2. die Geraden im R^2 (L ist hierbei eine Ursprungsgerade),
3. der ganze R^2.

Die linearen Mannigfaltigkeiten im R^3 sind entsprechend:

1. die einpunktigen Teilmengen des R^3 (L ist hier der Nullraum $\{0\}$),
2. die Geraden im R^3 (L ist hierbei eine Ursprungsgerade),
3. die Ebenen in R^3 (L ist hierbei eine Ursprungsebene),
4. der ganze R^3.

Satz 1 (Gleichheit von linearen Mannigfaltigkeiten). *Es seien* $M_1 = x_0 + L_1$, $M_2 = y_0 + L_2$ *zwei lineare Mannigfaltigkeiten im* R^n. *Dann gilt:* $M_1 = M_2 \Leftrightarrow L_1 = L_2$ *und* $x_0 - y_0 \in L_1$.

Beweis. Es ist $M_1 = \{x_0 + x : x \in L_1\}$, $M_2 = \{y_0 + y : y \in L_2\}$. Also gibt es bei $M_1 = M_2$ zu jedem $x \in L_1$ ein $y \in L_2$ mit $x_0 + x = y_0 + y$, und umgekehrt gibt es zu jedem $y \in L_2$ ein $x \in L_1$ mit $x_0 + x = y_0 + y$. Für $x = 0$ bzw. $y = 0$ ergibt sich somit $x_0 - y_0 \in L_1$ und $x_0 - y_0 \in L_2$. Dann ist also jedes $x \in L_1$ Summe zweier Elemente von L_2 und jedes $y \in L_2$ Summe zweier Elemente von L_1, d. h. $L_1 = L_2$.

Die andere Implikation ist noch einfacher $x_0 - y_0 \in L_1 (= L_2)$ impliziert für jedes $x \in L_1$

$$x + (x_0 - y_0) = y \in L_1.$$

Jedes $y \in L_1$ tritt auch stets als gewisses $x + (x_0 - y_0)$ auf. Also ist $x_0 + L_1 = y_0 + L_2$.

Die soeben bewiesene Aussage rechtfertigt wegen der eindeutigen Erzeugung einer linearen Mannigfaltigkeit im R^n aus einem wohlbestimmten linearen Teilraum die folgende Begriffsbildung.

Definition 3 (Dimension einer linearen Mannigfaltigkeit). Es sei M eine lineare Mannigfaltigkeit im R^n. Unter der *Dimension* (dim M) dieser linearen Mannigfaltigkeit versteht man dann die Dimension des diese lineare Mannigfaltigkeit hervorbringenden linearen Teilraumes:

$$M = \boldsymbol{x}_0 + L, \quad \dim M := \dim L.$$

Die linearen Mannigfaltigkeiten im R^n von der Dimension $n-1$ haben noch eine besondere Bezeichnung, die sich aus den für den R^3 in euklidischer Veranschaulichung geltenden Beziehungen erklärt.

Definition 4 (Hyperebenen im R^n). Unter einer *Hyperebene* im R^n versteht man eine lineare Mannigfaltigkeit M des R^n von der Dimension $n-1$ (dim $M = n-1$).

So wie die $(n-1)$-dimensionalen linearen Teilräume mit den Linearformen in Beziehung stehen, so gibt es auch eine Beziehung zwischen den Hyperebenen und Linearformen.

Satz 2 (Beschreibung der Hyperebenen durch Linearformen).

1. *Es sei f eine nichtausgeartete Linearform auf dem R^n. Dann ist für jedes gegebene reelle $\alpha \in \mathsf{R}$ die Lösungsmenge der Gleichung $f(\boldsymbol{x}) = \alpha$ eine Hyperebene im R^n.*

2. *Jede Hyperebene H im R^n ist Lösungsmenge einer Gleichung $f(\boldsymbol{x}) = \alpha$ mit einer gewissen nichtausgearteten Linearform f auf dem R^n und einem gewissen $\alpha \in \mathsf{R}$.*

3. *Zwei nichtausgeartete Linearformen f, g auf dem R^n bestimmen durch die Gleichungen $f(\boldsymbol{x}) = \alpha$ bzw. $g(\boldsymbol{x}) = \beta$ ein und dieselbe Hyperebene H im R^n genau dann, wenn es eine reelle Zahl $\lambda \neq 0$ gibt mit $f = \lambda g$ und $\alpha = \lambda\beta$.*

Bemerkung. Mit diesem Satz ist dann also auch die Struktur der Lösungsmannigfaltigkeit von einer inhomogenen linearen Gleichung mit n Unbekannten aufgeklärt.

Beweis des Satzes.

Zu 1. Die Lösungsmannigfaltigkeit von $f(\boldsymbol{x}) = \alpha$ sei H. Die Lösungsmenge der Gleichung $f(\boldsymbol{x}) = 0$ ist ein $(n-1)$-dimensionaler linearer Teilraum des R^n, der Kern (ker f) dieser Linearform. Es sei $\boldsymbol{x}_0$ eine gewisse Lösung von $f(\boldsymbol{x}) = \alpha$. Eine solche können wir uns beispielsweise verschaffen, wenn wir in dem Ausdruck $a_1x_1 + a_2x_2 + \cdots + a_nx_n = \alpha$ (wobei $(a_1, a_2, \ldots, a_n)$ das Koeffiziententupel von f bezeichnet) etwa $x_2, \ldots, x_n$ beliebig wählen und x_1 folgendermaßen berechnen:

$$x_1 = \frac{\alpha}{a_1} - \frac{a_2}{a_1}x_2 - \cdots - \frac{a_n}{a_1}x_n.$$

(Wir haben o. B. d. A. $a_1 \neq 0$ angenommen). $\boldsymbol{x}_0$ sei das n-Tupel $(x_1, x_2, \ldots, x_n)$ mit den soeben festgelegten Koordinaten.[1])

[1]) Die Existenz eines $\boldsymbol{x}_0$ mit $f(\boldsymbol{x}_0) = \alpha$ wäre auch so zu erschließen: Für eine nichtausgeartete Linearform $f\colon \mathsf{R}^n \to \mathsf{R}$ gilt $f(\mathsf{R}^n) = \mathsf{R}$, was sofort aus der Homogenität von f folgt! Also gibt es für jedes $\alpha \in \mathsf{R}$ wenigstens ein $\boldsymbol{x}_0 \in \mathsf{R}^n$ mit $f(\boldsymbol{x}_0) = \alpha$.

Dann gilt für H

$$H = \boldsymbol{x}_0 + \ker f,$$

denn wir haben die Beziehung:

$$f(\boldsymbol{x}) = \alpha \Leftrightarrow \boldsymbol{x}_0 - \boldsymbol{x} \in \ker f.$$

Zu 2. Es sei $H = \boldsymbol{x}_0 + L$ mit einem linearen Teilraum L des $\mathbf{R}^n$ der Dimension $n-1$. Dann gibt es eine Linearform f auf dem $\mathbf{R}^n$, so daß

$$L = \ker f$$

ist. Für diese Linearform erhalten wir H als Lösungsmenge der Gleichung

$$f(\boldsymbol{x}) = f(\boldsymbol{x}_0).$$

Zu 3. Wenn es eine reelle Zahl λ der genannten Eigenschaft gibt, haben wir

$$f(\boldsymbol{x}) = \alpha \Leftrightarrow \lambda g(\boldsymbol{x}) = \lambda\beta \Leftrightarrow g(\boldsymbol{x}) = \beta.$$

Also stimmen die durch die erste Gleichung und zweite Gleichung bestimmten Hyperebenen überein. Es sei jetzt

$$H_1 := \{\boldsymbol{x} : f(\boldsymbol{x}) = \alpha\}, \qquad H_2 := \{\boldsymbol{x} : g(\boldsymbol{x}) = \beta\}.$$

Vorausgesetzt wird $H_1 = H_2$. Wir haben demnach $\boldsymbol{x}_0 + \ker f = \boldsymbol{y}_0 + \ker g$ für $\boldsymbol{x}_0, \boldsymbol{y}_0 \in \mathbf{R}^n$ mit $f(\boldsymbol{x}_0) = \alpha, g(\boldsymbol{y}_0) = \beta$. Zunächst ist $\ker f = \ker g$ nach dem Satz über die Gleichheit von linearen Mannigfaltigkeiten. Dann muß aber wiederum $f = \lambda g$ für ein gewisses $\lambda \neq 0$ gelten. Weil noch $\boldsymbol{x}_0 - \boldsymbol{y}_0 \in \ker f$ ist, ist $f(\boldsymbol{x}_0) = f(\boldsymbol{y}_0) = \lambda g(\boldsymbol{y}_0)$, d. h. $\alpha = \lambda\beta$.

Wir erkennen damit, daß die Lösungsmannigfaltigkeit eines inhomogenen linearen Gleichungssystems in n Unbekannten der Durchschnitt von Hyperebenen des $\mathbf{R}^n$ ist.

Satz 3 (Durchschnitt von linearen Mannigfaltigkeiten des $\mathbf{R}^n$). *Es sei $(M_i)_{i\in I}$ eine beliebige Familie von linearen Mannigfaltigkeiten des $\mathbf{R}^n$. Dann ist der Durchschnitt $\bigcap_{i\in I} M_i$ entweder leer oder er ist wieder eine lineare Mannigfaltigkeit des $\mathbf{R}^n$. Im letzten Fall gilt:*

$$\bigcap_{i\in I} M_i = \boldsymbol{x}_0 + \bigcap_{i\in I} L_i,$$

sofern $M_i = \boldsymbol{x}_0 + L_i$ ist (L_i linearer Teilraum des $\mathbf{R}^n$).

Beweis. Es sei $\bigcap M_i \neq \emptyset$. Wir wählen einen Punkt $\boldsymbol{x}_0 \in \bigcap M_i$. Dann haben wir die Darstellung $M_i = \boldsymbol{x}_0 + L_i$. Für $\boldsymbol{x} \in \bigcap M_i$ ist also $\boldsymbol{x} = \boldsymbol{x}_0 + \boldsymbol{y}_i$ bei $\boldsymbol{y}_i \in L_i$, d. h. $\boldsymbol{x} - \boldsymbol{x}_0 \in L_i$, für alle $i \in I$. Demnach liegt $\boldsymbol{x}$ in $\boldsymbol{x}_0 + \bigcap L_i$.

Wenn umgekehrt $\boldsymbol{x} \in \boldsymbol{x}_0 + \bigcap L_i$ gilt, so haben wir $\boldsymbol{x} = \boldsymbol{x}_0 + \boldsymbol{y}$ mit $\boldsymbol{y} \in \bigcap L_i$, d. h. $\boldsymbol{x} \in M_i$ für alle $i \in I$.

Bemerkung. Die Lösungsgesamtheit eines inhomogenen linearen Gleichungssystems ist also entweder leer oder aber eine lineare Mannigfaltigkeit im $\mathbf{R}^n$. Die

nachfolgende Aussage macht noch eine zusätzliche Dimensionsangabe dieser Mannigfaltigkeit.

Satz 4 (Struktur der Lösungsmannigfaltigkeit eines lösbaren beliebigen linearen Gleichungssystems). *Es sei ein beliebiges (d. h. homogenes oder inhomogenes) lineares Gleichungssystem in n Unbekannten gegeben:*

$$\left.\begin{aligned} f_1(x_1, x_2, \ldots, x_n) &= b_1, \\ &\cdots\cdots\cdots\cdots \\ f_m(x_1, x_2, \ldots, x_n) &= b_m. \end{aligned}\right\} \qquad (*)$$

Dieses Gleichungssystem sei lösbar. Dann ist die Lösungsmannigfaltigkeit eine lineare Mannigfaltigkeit im R^n. *Ihre Dimension beträgt* n — *Rang des Gleichungssystems* (*).

Umgekehrt gibt es zu jeder linearen Mannigfaltigkeit M des R^n *von einer Dimension* k, $0 \leqq k < n$, *ein lösbares lineares Gleichungssystem vom Rang* $n - k$, *das die gegebene lineare Mannigfaltigkeit als Lösungsmenge besitzt.*

Beweis. Wir brauchen für den ersten Teil des Satzes lediglich noch die Dimensionsangabe zu beweisen. Es ist die Lösungsmannigfaltigkeit gleich $x_0 + \bigcap_{i=1}^{m} \ker f_i$. Nach dem Satz über die Struktur der Lösungsmenge eines homogenen linearen Gleichungssystems hat aber der lineare Teilraum $\bigcap_{i=1}^{m} \ker f_i$ gerade die Dimension n — Rang von (*).

Zur umgekehrten Situation: Es sei $M = x_0 + L$ mit einem linearen Teilraum L der Dimension k. L kann nach dem Satz über die Struktur der Lösungsmenge eines homogenen linearen Gleichungssystems beschrieben werden als Lösungsmenge von

$$\begin{aligned} f_1(x) &= 0, \\ &\cdots\cdots\cdots \\ f_m(x) &= 0 \end{aligned}$$

mit $m = n - k$ linear unabhängigen Linearformen. Dann ist M Lösungsmenge von

$$\begin{aligned} f_1(x) &= f_1(x_0), \\ &\cdots\cdots\cdots\cdots \\ f_m(x) &= f_m(x_0). \end{aligned}$$

5.4. Der Gaußsche Algorithmus

Nachdem wir mit den vorhergehenden Erörterungen eine vollständige Aufklärung der Struktur der Lösungsmannigfaltigkeiten von allgemeinen linearen Gleichungssystemen erzielt haben, bleibt jetzt die Frage nach einem effektiven Verfahren der

konkreten Bestimmung der Lösungsmannigfaltigkeiten zu beantworten. Ein solches Verfahren ist der sogenannte *Gaußsche Algorithmus*, er besteht in einer fortschreitenden Elimination der Unbekannten. Die Begründung für die Schlüssigkeit des in Rede stehenden Verfahrens kann nach unseren bisherigen Einsichten leicht erbracht werden. Es handelt sich eigentlich nur noch um eine zweckmäßige Anordnung der zuvor gewonnenen theoretischen Resultate.[1])

Das vorgelegte lineare Gleichungssystem

$$\begin{aligned} f_1(\boldsymbol{x}) &= b_1, \\ &\cdots\cdots\cdots \\ f_m(\boldsymbol{x}) &= b_m, \end{aligned} \qquad (*)$$

bestehe aus m Gleichungen mit n Unbekannten $x_1, x_2, \ldots, x_n$. Die Lösungsmenge M dieses Gleichungssystems ist der Durchschnitt der Hyperebenen $H_i := \{\boldsymbol{x} : f_i(\boldsymbol{x}) = b_i\}$, $i = 1, \ldots, m$. Nach dem Satz über die Beschreibung der Hyperebenen durch Linearformen erhalten wir dieselbe Hyperebene H_i, wenn wir die anfängliche Gleichung $f_i(\boldsymbol{x}) = b_i$ durch ein skalares Vielfaches $\lambda_i f_i(\boldsymbol{x}) = \lambda_i b_i$ mit einem von Null verschiedenen Skalar λ_i ersetzen. Wenn wir zwei Gleichungen $f_i(\boldsymbol{x}) = b_i$ und $f_j(\boldsymbol{x}) = b_j$ und ihre Lösungsmannigfaltigkeit $H_i \cap H_j$ betrachten, ist diese Menge $H_i \cap H_j$ auch Lösungsmannigfaltigkeit des Systems mit den Gleichungen $f_i(\boldsymbol{x}) = b_i$ und $f_i(\boldsymbol{x}) + f_j(\boldsymbol{x}) = b_i + b_j$, denn aus dem Bestehen der ersten beiden Gleichungen folgt das Bestehen der letzten beiden Gleichungen und umgekehrt.

Wir haben demnach das folgende Ergebnis.

Satz 1 (Elementare Umformungen von linearen Gleichungssystemen). *Es sei ein lineares Gleichungssystem mit m Gleichungen in n Unbekannten gegeben. Durch wiederholte Anwendung der folgenden elementaren Gleichungsumformungen gelangt man stets zu einem Gleichungssystem, welches zum Ausgangssystem äquivalent ist:*

1. *Vertauschen zweier Gleichungen.*

2. *Ersetzen einer Gleichung durch ein skalares Vielfaches mit einem von Null verschiedenen Skalar.*

3. *Ersetzen einer Gleichung durch die Summe dieser Gleichung und einer beliebigen anderen Gleichung des Systems.*

Dieser Satz bildet den Hintergrund für den Gaußschen Algorithmus, mit dem man ein vorgelegtes lineares Gleichungssystem durch geeignete Elementarumformungen auf ein möglichst übersichtliches äquivalentes Gleichungssystem bringt, an dem die konkreten Lösungen leicht abzulesen sind. Durch den Gaußschen Algorithmus bringt man das anfängliche Gleichungssystem auf *Trapezform*. Wir führen den Algorithmus

[1]) Wegen der großen praktischen Wichtigkeit der Lösung linearer Gleichungssysteme erfolgt eine weitere Behandlung in der Numerik (MfL, Bd. 9).

an Beispielen vor, die getätigten Zeilenumformungen sind jeweils in der ersten Spalte markiert.

1. Vorgelegtes Gleichungssystem:

$$\begin{aligned} x_1 + 2x_2 + 5x_3 &= -9, \\ x_1 - x_2 + 3x_3 &= 2, \\ 3x_1 - 6x_2 - x_3 &= 25. \end{aligned}$$

Umformung auf Trapezform

	Koeffizienten bei			absolutes Glied
	x_1	x_2	x_3	
(1)	1	2	5	−9
(2)	1	−1	3	2
(3)	3	−6	−1	25

	Koeffizienten bei			absolutes Glied
	x_1	x_2	x_3	
(1)	1	2	5	−9
(2) − (1)		−3	−2	11
(3) − 3(1)		−12	−16	52

	Koeffizienten bei			absolutes Glied
	x_1	x_2	x_3	
(1)	1	2	5	−9
(2)		−3	−2	11
(3) −4(2)			−8	8

Das Verfahren ist beendet. Die letzte Zeile ergibt $-8x_3 = 8$, also $x_3 = -1$. Dies in die vorletzte Zeile eingesetzt, liefert $x_2 = -3$. Woraus schließlich durch Einsetzen in die erste Zeile $x_1 = 2$ hervorgeht.

Das vorgelegte Gleichungssystem hat demnach genau eine Lösung, nämlich das Element

$$(x_1, x_2, x_3) = (2, -3, -1).$$

2. Vorgelegtes Gleichungssystem:

$$\begin{aligned} x_1 + 2x_2 &= 4, \\ 3x_1 - x_2 &= 2, \\ 2x_1 + x_2 &= 1. \end{aligned}$$

Umformung auf Trapezform:

	Koeffizienten bei x_1	x_2	absolutes Glied
(1)	1	2	4
(2)	3	−1	2
(3)	2	1	1

	Koeffizienten bei x_1	x_2	absolutes Glied
(1)	1	2	4
(2) −3(1)		−7	−10
(3) −2(1)		−3	−7

	Koeffizienten bei x_1	x_2	absolutes Glied
(1)	1	2	4
(2)		−7	−10
(3) $-\frac{3}{7}$(2)		0	$-7+\frac{30}{7}$

Das Verfahren ist beendet. Die letzte Zeile ergibt, daß die Nullform einen von Null verschiedenen Wert annehmen müßte, was jedoch nicht stattfinden kann.

Das vorgelegte Gleichungssystem ist demnach unverträglich, es besitzt keine Lösung.

3. Vorgelegtes Gleichungssystem:

$$\begin{aligned} 2x_1 - x_2 - x_3 + 3x_4 &= 1,\\ 4x_1 - 2x_2 - x_3 + x_4 &= 5,\\ 6x_1 - 3x_2 - x_3 - x_4 &= 9,\\ 2x_1 - x_2 + 2x_3 - 12x_4 &= 10. \end{aligned}$$

Umformung auf Trapezform:

	Koeffizienten bei x_1	x_2	x_3	x_4	absolutes Glied
(1)	2	−1	−1	3	1
(2)	4	−2	−1	1	5
(3)	6	−3	−1	−1	9
(4)	2	−1	2	−12	10

	Koeffizienten bei				absolutes Glied
	x_1	x_2	x_3	x_4	
(1)	2	−1	−1	3	1
(2) − 2(1)			1	−5	3
(3) − 3(1)			2	−10	6
(4) − (1)			3	−15	9

	Koeffizienten bei				absolutes Glied
	x_1	x_2	x_3	x_4	
(1)	2	−1	−1	3	1
(2)			1	−5	3
(3) − 2(2)			0	0	0
(4) − 3(2)			0	0	0

Das Verfahren ist in erster Etappe beendet. Das anfängliche Gleichungssystem ist in ein neues übergeführt worden, welches aus zwei Gleichungen mit vier Unbekannten besteht. Wir wählen zwei Unbekannte beliebig, etwa $x_1 = \alpha$, $x_2 = \beta$. Dann haben wir ein neues Gleichungssystem vor uns:

$$-x_3 + 3x_4 = 1 - 2\alpha + \beta$$
$$x_3 - 5x_4 = 3.$$

Die entsprechende Behandlung bringt folgendes:

	Koeffizienten bei		absolutes Glied
	x_3	x_4	
(1)	−1	3	$1 - 2\alpha + \beta$
(2)	1	−5	3

	Koeffizienten bei		absolutes Glied
	x_3	x_4	
(1)	−1	3	$1 - 2\alpha + \beta$
(2) + (1)		−2	$4 - 2\alpha + \beta$

Daraus ergibt sich:

$$x_4 = \alpha - \frac{\beta}{2} - 2,$$

$$x_3 = 5\alpha - \frac{5}{2}\beta - 7.$$

Das vorgelegte Gleichungssystem ist also unbestimmt. Die Lösungsmannigfaltigkeit besteht aus den Elementen

$$\left(\alpha, \beta, 5\alpha - \frac{5}{2}\beta - 7, \alpha - \frac{\beta}{2} - 2\right)$$

für beliebiges $\alpha, \beta \in \mathbf{R}$. Wegen der Darstellung

$$\left(\alpha, \beta, 5\alpha - \frac{5}{2}\beta - 7, \alpha - \frac{\beta}{2} - 2\right) = \alpha(1, 0, 5, 1) + \beta\left(0, 1, -\frac{5}{2}, -\frac{1}{2}\right) + (0, 0, -7, -2)$$

ist also die Lösungsmannigfaltigkeit

$$M = (0, 0, -7, -2) + L\left(\left\{(1, 0, 5, 1), \left(0, 1, -\frac{5}{2}, -\frac{1}{2}\right)\right\}\right).$$

Bemerkung (Rangberechnung von linearen Gleichungssystemen). Auf das zu einem gegebenen linearen Gleichungssystem gehörige homogene Gleichungssystem wendet man den Gaußschen Algorithmus an. Der Gaußsche Algorithmus führt zu einer Entscheidung, wie viele Unbekannte frei wählbar sind. Diese Anzahl der frei wählbaren Unbekannten gibt die Dimension des Lösungsraumes an. Der Rang beträgt dann: Anzahl der Unbekannten minus Dimension des Lösungsraumes. Gleichbedeutend dazu ist, daß der Rang gerade durch die Anzahl der nichttrivialen Zeilen gegeben wird, die sich nach Abschluß des Gaußschen Algorithmus in der Trapezform noch vorfinden. Man überzeugt sich nämlich davon, daß der Gaußsche Algorithmus zu linear unabhängigen Zeilen führt.

Für unsere vorherigen Beispiele heißt dies: Das erste homogene Gleichungssystem hat den Rang 3, das zweite homogene Gleichungssystem hat den Rang 2, das dritte homogene Gleichungssystem hat den Rang 2.

5.5. Übungsaufgaben

1. Im $\mathbf{R}^2$ bestimmt jede Gleichung $f(x) = \alpha$ mit nichtausgeartetem $f \in \mathscr{L}(\mathbf{R}^2)$ und $\alpha \in \mathbf{R}$ eine Gerade. (Bei Zugrundelegung der euklidischen Veranschaulichung des $\mathbf{R}^2$ macht die Lösungsgesamtheit der genannten Gleichung eine Gerade aus). Man ermittle Bedingungen dafür, daß die Gerade parallel zur x-Achse verläuft! (Ausgedrückt durch das Koeffiziententupel von f und den Wert α).
2. Wann und nur wann haben zwei Geraden im $\mathbf{R}^2$, die durch die Gleichungen $f(x) = \alpha$, $g(x) = \beta$ mit $f, g \in \mathscr{L}(\mathbf{R}^2)$, $\alpha, \beta \in \mathbf{R}$, beschrieben werden, genau einen Schnittpunkt? (Vgl. Aufgabe 1.)
3. Welche geometrische Deutung in $\mathbf{R}^2$ und $\mathbf{R}^3$ kann man dem Hilfssatz über die Trennbarkeit von linearen Teilräumen und Punkten durch Linearformen geben? (Existenz von Gerade bzw. Ebene, die nicht durch den gegebenen Punkt verlaufen und sich noch geeignet zu dem gegebenen linearen Teilraum verhalten.)

4. Man bestimme alle Elemente $x = (x_1, x_2, x_3)$ des $\mathbb{R}^3$, die auf der Hyperebene liegen, welche die Elemente $(1, 0, 0)$, $(0, 1, 0)$ $(0, 0, 1)$ enthält! Man gebe eine Hyperebenengleichung für die gewünschte Hyperebene an.

5. Das folgende Gleichungssystem ist zu lösen:
$$\begin{aligned} x_1 + x_2 + x_3 + x_4 &= 5, \\ 2x_1 + 3x_2 + x_3 + x_4 &= 2, \\ x_2 + 2x_3 + 3x_4 &= 18, \\ 5x_1 + 4x_2 + 3x_3 + 2x_4 &= 12. \end{aligned}$$

6. Das folgende Gleichungssystem ist in Abhängigkeit von λ zu lösen:
$$\begin{aligned} \lambda x + y + z &= 1, \\ -x + \lambda y + z &= \lambda, \\ -x - y + \lambda z &= -\lambda. \end{aligned}$$

7. Man bestimme eine Basis des Lösungsraumes des folgenden Gleichungssystems:
$$\begin{aligned} x_1 + x_2 - x_3 + x_4 &= 0, \\ x_1 - x_2 + x_3 + x_4 &= 0, \\ 3x_1 + x_2 - x_3 + 3x_4 &= 0, \\ x_2 - x_3 &= 0. \end{aligned}$$

8. Sind die folgenden Elemente des $\mathbb{R}^3$ linear unabhängig:
$x = (1, -2, 1)$, $y = (3, -1, 2)$, $z = (2, 1, 2)$?

9. Man ermittle drei Hyperebenengleichungen im $\mathbb{R}^3$, so daß der Durchschnitt der Hyperebenen gerade nur aus dem Element $(1, 1, 1)$ besteht.

10. Was für eine Figur entsteht im $\mathbb{R}^2$ (bei euklidischer Veranschaulichung), wenn man die Komplexsumme der Menge $\{(x_1, x_2): x_1^2 + x_2^2 = 1\}$ mit sich selbst betrachtet?

6. Lineare Abbildungen des n-dimensionalen reellen Zahlenraumes. Matrizen

6.1. Erklärung der linearen Abbildungen des n-dimensionalen reellen Zahlenraumes. Beispiele

Die mit einem gegebenen linearen Gleichungssystem verbundene Aufgabenstellung wollen wir jetzt von einem neuen abbildungstheoretischen Standpunkt betrachten. Es sei nun ein allgemeines lineares Gleichungssystem in n Unbekannten mit m Gleichungen vorgelegt:

$$\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1,\\
&a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2,\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m.
\end{aligned}$$

Wir lassen zunächst die vorgeschriebenen rechten Seiten außer Betracht und nehmen uns lediglich die gegebenen Linearformen $l_1, l_2, \ldots, l_m$ mit den angeschriebenen Koeffiziententupeln $(a_{11}, a_{12}, \ldots, a_{1n}), \ldots, (a_{m1}, \ldots, a_{mn})$ vor. Zu jedem $\boldsymbol{x} \in \mathsf{R}^n$ gehört dann ein Punkt

$$\boldsymbol{y} = \big(l_1(\boldsymbol{x}), l_2(\boldsymbol{x}), \ldots, l_m(\boldsymbol{x})\big)$$

des R^m. Wir haben damit eine Abbildung $A: \mathsf{R}^n \to \mathsf{R}^m$; $\boldsymbol{x} \to \big(l_1(\boldsymbol{x}), \ldots, l_m(\boldsymbol{x})\big)$. Das Lösen des gegebenen Gleichungssystems bedeutet dann, alle $\boldsymbol{x} \in \mathsf{R}^n$ ausfindig zu machen, für welche $A(\boldsymbol{x}) = \boldsymbol{b}$ ist mit $\boldsymbol{b} = (b_1, b_2, \ldots, b_m)$. Das Lösen von linearen Gleichungssystemen ist also in der Natur seiner Aufgabenstellung etwa vollkommen dem Lösen von Gleichungen n-ten Grades verwandt, nur mit dem Unterschied, daß die gegebenen Abbildungen verschiedener Natur sind. Über die Natur der bei linearen Gleichungssystemen auftretenden Abbildungen A kann man sich rasch einen weiteren Einblick verschaffen. Die Abbildung A hat hinsichtlich der arithmetischen Struktur im R^n die gleichen Eigenschaften wie eine Linearform, d. h., A ist additiv und homogen.

Definition 1 (Lineare Abbildungen des R^n in den R^m). Eine Abbildung $A: \mathsf{R}^n \to \mathsf{R}^m$ des n-dimensionalen reellen Zahlenraumes in den m-dimensionalen reellen Zahlen-

raum heißt *lineare Abbildung* genau dann, wenn hinsichtlich der koordinatenweisen Addition im R^n, R^m und der Multiplikation mit Skalaren folgendes gilt:

1. $A(\boldsymbol{x}+\boldsymbol{y}) = A(\boldsymbol{x}) + A(\boldsymbol{y})$ für alle $\boldsymbol{x}, \boldsymbol{y} \in R^n$.
2. $A(\alpha\boldsymbol{x}) = \alpha A(\boldsymbol{x})$ für alle $\boldsymbol{x} \in R^n$ und $\alpha \in R$.

Bemerkungen.

1. Reelle Linearformen in n Unbekannten sind ein Spezialfall der linearen Abbildungen. Bei ihnen handelt es sich um lineare Abbildungen des R^n in den R^1.

2. Ein anderer Typ von linearen Abbildungen ist uns bei den Isomorphismen vom R^n in den R^m ($m \geqq n$) begegnet. Hier tritt zur Linearität der Abbildung noch die Eineindeutigkeit.

Einige weitere Beispiele für lineare Abbildungen

1. Es sei λ eine beliebige fixierte reelle Zahl. Dann ist im R^n durch die Festsetzung $A(\boldsymbol{x}) := \lambda\boldsymbol{x}$, $\boldsymbol{x} \in R^n$, eine lineare Abbildung erklärt. Sie heißt *Homothetie* mit dem Koeffizienten λ. Man verfolge etwa die Wirkung von A an der geometrisch-euklidischen Veranschaulichung des R^2 bzw. R^3. Im Fall $0 \leqq \lambda \leqq 1$ wirkt die Abbildung stauchend und im Fall $\lambda > 1$ streckend. $\lambda = -1$ bedeutet eine zentrale Spiegelung, so daß also bei negativem λ eine Zusammensetzung der erstgenannten Wirkung mit einer zentralen Spiegelung zustande kommt.

2. Lineare Abbildungen sind insbesondere für die Geometrie wichtig. Das wird dort im einzelnen besprochen. Hier illustrieren wir lediglich noch etwas mehr den Begriff der linearen Abbildung, wobei wir uns auf Schulkenntnisse beziehen.

Wir betrachten im R^2 folgende anschaulich beschriebene Abbildung A. In der geometrisch-euklidischen Veranschaulichung des R^2 als Ebene wählen wir eine beliebige Gerade l durch den Ursprung. Zum anderen sei eine davon verschiedene Gerade g vorgegeben. Jetzt werde jeder Punkt der Ebene durch Parallelprojektion auf die Gerade l längs der vorgegebenen Richtung g projiziert (vgl. Abb. 5 und 6).

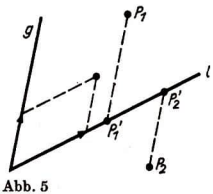


Abb. 5

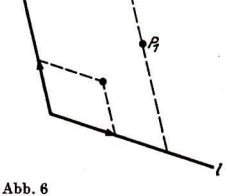


Abb. 6

Durch diese Vorschrift wird eine Abbildung $A: R^2 \to R^2$ festgelegt. Als Bildmenge tritt hierbei der eindimensionale Teilraum des R^2 auf, der durch die Ursprungs-

gerade l repräsentiert wird. An Hand von Schulkenntnissen mache man sich klar, daß die Abbildung linear ist.

3. Entsprechend zum vorstehenden Beispiel erörtere man, daß eine axiale Streckung im $\mathbf{R}^2$ in bezug auf eine gegebene Ursprungsgerade l längs einer gegebenen (davon verschiedenen) Geraden g mit einem Streckungsfaktor λ eine lineare Abbildung im $\mathbf{R}^2$ beschreibt. Der Fall $\lambda = -1$ bedeutet hierbei die Spiegelung an l längs der durch g gegebenen Richtung.

4. Es sei $A: \mathbf{R}^n \to \mathbf{R}^m$ eine beliebige lineare Abbildung des $\mathbf{R}^n$ in den $\mathbf{R}^m$. Wir betrachten die kanonischen Koordinatenprojektionen $p_1, p_2, \ldots, p_m$ in den $\mathbf{R}^m$. Jedes $p_i: \mathbf{R}^m \to \mathbf{R}^1$ ist, wie wir wissen, eine Linearform auf dem $\mathbf{R}^m$, die jedem Element $\boldsymbol{y} \in \mathbf{R}^m$ gerade als Funktionalwert seine i-te Koordinate zuordnet. Die Zusammensetzung von A mit den p_i, $i = 1, \ldots, m$, liefert uns m Linearformen $f_i := p_i \circ A$ auf dem $\mathbf{R}^n$. Die Abbildung A ist vollständig durch diese m Linearformen bestimmt. Es ist

$$A(\boldsymbol{x}) = \boldsymbol{y} = \big(f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \ldots, f_m(\boldsymbol{x})\big), \quad \boldsymbol{x} \in \mathbf{R}^n.$$

Die zu den f_i gehörenden Koeffiziententupel seien $(a_{i1}, a_{i2}, \ldots, a_{in})$. Dann ist also

$$\begin{aligned}
y_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n, \\
y_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n, \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
y_m &= a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n.
\end{aligned}$$

Der Verlauf der linearen Abbildung $A: \mathbf{R}^n \to \mathbf{R}^m$, wo also jedem $\boldsymbol{x} = (x_1, \ldots, x_n)$ des $\mathbf{R}^n$ das Element $A(\boldsymbol{x}) = \boldsymbol{y} = (y_1, y_2, \ldots, y_m)$ zugeordnet wird, ist demnach in der gleichen Weise beschrieben wie in der Situation, die uns zum Begriff der linearen Abbildung geführt hat. Ein lineares Funktional $f: \mathbf{R}^n \to \mathbf{R}^1$ wird durch ein Koeffiziententupel $(a_1, a_2, \ldots, a_n)$ beschrieben. Eine lineare Abbildung $A: \mathbf{R}^n \to \mathbf{R}^m$ wird — wie wir jetzt herausgefunden haben — durch ein Koeffizientenschema beschrieben, das aus m Zeilen und n Spalten besteht:

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

Mit solchen Koeffizientenschemata werden wir uns nun etwas ausführlicher beschäftigen.

6.2. Matrizen und die durch sie beschriebenen linearen Abbildungen

Definition 1 (Reelle Matrix vom Typ $m \times n$). Unter einer (reellen) *Matrix vom Typ* $m \times n$; m, n natürliche Zahlen; versteht man eine Abbildung $\boldsymbol{A}$ der Menge $\{1, 2, \ldots, m\} \times \{1, 2, \ldots, n\}$ in $\mathbf{R}$. Diese Abbildung

$$\boldsymbol{A}: \{1, 2, \ldots, m\} \times \{1, 2, \ldots, n\} \to \mathbf{R},$$

definiert auf der endlichen Menge $\{1, 2, \ldots, m\} \times \{1, 2, \ldots, n\}$, gibt man zweckmäßig durch ihren Verlauf $(i, j) \mapsto a_{ij}$ in einer rechteckigen Tabelle wie folgt an:

$$\boldsymbol{A}: \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

oder kürzer $(a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$. a_{ij} nennt man *das allgemeine Element der Matrix* $\boldsymbol{A}$. Der erste Index i des Elementes a_{ij} gibt den Zeilenstand und der zweite Index den Spaltenstand des Elementes a_{ij} an.

Man sagt auch: Eine Matrix vom Typ $m \times n$ (gelesen „m Kreuz n") ist ein rechteckiges Zahlenschema von m Zeilen und n Spalten. Man meint damit natürlich den ausgesprochenen Sachverhalt der Abbildung. Die Zweckmäßigkeit der Angabe der hier interessierenden Abbildungen durch die rechteckige Anordnung der auftretenden Bildelemente geht fürs erste schon aus dem oben vermerkten Zusammenhang mit linearen Gleichungssystemen hervor. Weitere Vorteile dieser Schreibweise werden sich bald zeigen, vor allem bezieht sich die noch zu erklärende Matrizenmultiplikation wesentlich auf die Zeilen-Spalten-Struktur einer Matrix.

Definition 2 (Gleichheit von Matrizen). Es seien $\boldsymbol{A}$, $\boldsymbol{B}$ zwei (reelle) Matrizen,

$$\boldsymbol{A} = (a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}, \qquad \boldsymbol{B} = (b_{kl})_{\substack{k=1,\ldots,r \\ l=1,\ldots,s}}.$$

Diese heißen gleich genau dann, wenn sie als Abbildung gleich sind. Für die Gleichheit von Matrizen ist also notwendig und hinreichend, daß ihr Typ übereinstimmt und sie außerdem Element für Element übereinstimmen:

$$\begin{matrix} m = r \\ n = s \end{matrix} \quad \text{und} \quad a_{ij} = b_{kl} \quad \text{für alle} \quad \begin{matrix} i = k(1, \ldots, m) \\ j = l(1, \ldots, n). \end{matrix}$$

Im letzten Beispiel des vorhergehenden Abschnitts haben wir bemerkt, daß jede lineare Abbildung $\boldsymbol{A}: \mathbf{R}^n \to \mathbf{R}^m$ durch eine gewisse Matrix vom Typ $m \times n$ beschrieben werden kann. Diesen Sachverhalt fassen wir jetzt allgemeiner, indem wir von gewissen Basen im $\mathbf{R}^n$ und $\mathbf{R}^m$ ausgehen. Der ursprüngliche Fall bedeutet dann gerade die Bezugnahme auf die natürlichen Basen im $\mathbf{R}^n$ und $\mathbf{R}^m$.

Definition 3 (Die einer linearen Abbildung assoziierte Matrix). Es sei im $\mathbb{R}^n$ eine beliebige Basis $\mathfrak{B}$ in einer bestimmten Anordnung fixiert: $\mathfrak{B}: \boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$. Es sei im $\mathbb{R}^m$ eine beliebige Basis $\mathfrak{C}$ in einer bestimmten Anordnung fixiert: $\mathfrak{C}: \boldsymbol{c}_1, \ldots, \boldsymbol{c}_m$. Jeder linearen Abbildung $A: \mathbb{R}^n \to \mathbb{R}^m$ läßt sich dann eine von den Basen $\mathfrak{B}$, $\mathfrak{C}$ abhängende Matrix $\boldsymbol{A}$ vom Typ $m \times n$ in der folgenden Weise zuordnen:

$$\boldsymbol{A} = (a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}},$$

wobei a_{ij} die i-te Koordinate von $A(\boldsymbol{b}_j)$ in bezug auf die Basis $\mathfrak{C}$ ist. In der j-ten Spalte von $\boldsymbol{A}$ steht also das Koordinatentupel des Elementes $A(\boldsymbol{b}_j)$ bezüglich der Basis $\mathfrak{C}$.

Die Matrix $\boldsymbol{A}$ heißt die der linearen Abbildung $A: \mathbb{R}^n \to \mathbb{R}^m$ bezüglich der angeordneten Basen $\mathfrak{B}$ und $\mathfrak{C}$ *assoziierte Matrix* oder auch die die lineare Abbildung $A: \mathbb{R}^n \to \mathbb{R}^m$ bezüglich der Basen $\mathfrak{B}$ und $\mathfrak{C}$ *beschreibende Matrix.*

Wir werden sogleich auseinandersetzen, in welchem Sinne die Matrix die lineare Abbildung bei gegebenen Basen beschreibt. Zuvor die folgenden

Bemerkungen.

1. Bei fixierten Basen gehören zu verschiedenen linearen Abbildungen auch stets verschiedene beschreibende Matrizen.

2. Jede Matrix vom Typ $m \times n$ ist auch stets beschreibende Matrix einer gewissen linearen Abbildung vom $\mathbb{R}^n$ in $\mathbb{R}^m$ bezüglich der fixierten Basen (Nachweis).

3. In dem Spezialfall einer linearen Abbildung $A: \mathbb{R}^n \to \mathbb{R}^n$ des $\mathbb{R}^n$ in sich tritt häufig der Umstand ein, daß eine beschreibende Matrix von A bezüglich eines Basispaares $\mathfrak{B}$, $\mathfrak{B}$ mit ein und demselben ersten und zweiten Glied zu betrachten ist. Man spricht in diesem Fall von der beschreibenden Matrix von A bezüglich der Basis $\mathfrak{B}$.

Satz 1 (Beschreibung einer linearen Abbildung durch die assoziierte Matrix). *In den Räumen $\mathbb{R}^n$ und $\mathbb{R}^m$ sei je eine geordnete Basis $\mathfrak{B}: \boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$ und $\mathfrak{C}: \boldsymbol{c}_1, \ldots, \boldsymbol{c}_m$ fixiert. Es sei eine lineare Abbildung $A: \mathbb{R}^n \to \mathbb{R}^m$ gegeben. Es bezeichne*

$$\boldsymbol{A} = (a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$$

die zu A bezüglich der Basen $\mathfrak{B}$ und $\mathfrak{C}$ gehörige beschreibende Matrix. Die Matrix beschreibt dann in der folgenden Weise den Verlauf der Abbildung A: Es seien $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{y} \in \mathbb{R}^m$ zwei gegebene Elemente. Für sie gilt dann: $A(\boldsymbol{x}) = \boldsymbol{y} \Leftrightarrow$ Für die Koordinatentupel $(\xi_1, \ldots, \xi_n)$ von $\boldsymbol{x}$ bezüglich $\mathfrak{B}$ und $(\eta_1, \ldots, \eta_m)$ von $\boldsymbol{y}$ bezüglich $\mathfrak{C}$ gilt

$$\eta_1 = a_{11}\xi_1 + a_{12}\xi_2 + \cdots + a_{1n}\xi_n,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\eta_m = a_{m1}\xi_1 + a_{m2}\xi_2 + \cdots + a_{mn}\xi_n.$$

Beweis. Unter den gemachten Voraussetzungen ist

$$\boldsymbol{x} = \xi_1 \boldsymbol{b}_1 + \cdots + \xi_n \boldsymbol{b}_n, \quad \boldsymbol{y} = \eta_1 \boldsymbol{c}_1 + \cdots + \eta_m \boldsymbol{c}_m$$

und

$$A(\boldsymbol{b}_j) = a_{1j}\boldsymbol{c}_1 + a_{2j}\boldsymbol{c}_2 + \cdots + a_{mj}\boldsymbol{c}_m.$$

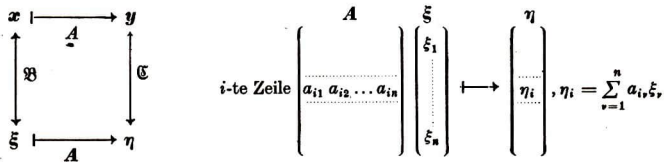
Daraus folgt wegen der Linearität von A

$$A(\boldsymbol{x}) = A\left(\sum_{\nu=1}^{n} \xi_\nu \boldsymbol{b}_\nu\right) = \sum_{\nu=1}^{n} \xi_\nu A(\boldsymbol{b}_\nu) = \sum_{\nu=1}^{n} \xi_\nu \sum_{\mu=1}^{m} a_{\mu\nu}\boldsymbol{c}_\mu = \sum_{\mu=1}^{m} \left(\sum_{\nu=1}^{n} a_{\mu\nu}\,\xi_\nu\right) \boldsymbol{c}_\mu.$$

Also sind die Koordinaten von $A(\boldsymbol{x})$ bezüglich der Basis $\mathfrak{C}$ gerade

$$\sum_{\nu=1}^{n} a_{1\nu}\xi_\nu, \sum_{\nu=1}^{n} a_{2\nu}\xi_\nu, \ldots, \sum_{\nu=1}^{n} a_{m\nu}\xi_\nu.$$

Wir notieren zwecks besserer Einprägung das folgende *Diagramm für den Zusammenhang zwischen A und $\boldsymbol{A}$* (es bezeichne A eine lineare Abbildung, $\boldsymbol{A}$ ihre beschreibende Matrix bezüglich der angeordneten Basen $\mathfrak{B}$ und $\mathfrak{C}$ im $\mathbb{R}^n$ bzw. $\mathbb{R}^m$, ξ das Koordinatentupel von $\boldsymbol{x} \in \mathbb{R}^n$ bezüglich $\mathfrak{B}$, η das Koordinatentupel von $\boldsymbol{y} \in \mathbb{R}^m$ bezüglich $\mathfrak{C}$):



Einige Beispiele für Matrixdarstellungen von linearen Abbildungen

1. Im $\mathbb{R}^n$ betrachten wir eine beliebige angeordnete Basis $\mathfrak{B}: \boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$. Wie lautet die assoziierte Matrix der identischen Abbildung $I: \mathbb{R}^n \to \mathbb{R}^n$, d. h. $I(\boldsymbol{x}) = \boldsymbol{x}$ für alle $\boldsymbol{x} \in \mathbb{R}^n$, bezüglich des Basispaares $\mathfrak{B}, \mathfrak{B}$? Die assoziierte Matrix $\boldsymbol{I}$ ist quadratisch vom Typ $n \times n$. In ihrer j-ten Spalte steht das Koordinatentupel von $\boldsymbol{b}_j$ bezüglich $\mathfrak{B}$, d. h.

$$\boldsymbol{I} = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & & \ldots & 1 \end{pmatrix}$$

Es treten also nur in der von links oben nach rechts unten laufenden Diagonalen Einsen auf und sonst Nullen. Diese Matrix heißt nach einem im nächsten Abschnitt

ersichtlichen Grunde die *Einheitsmatrix* vom Typ $n \times n$. Für gewöhnlich wird sie mit der folgenden Bezeichnungsweise abgekürzt:

$$\boldsymbol{I} = (\delta_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,n}}.$$

Hierbei heißt δ_{ij} das *Kronecker-Symbol*, es ist definiert durch

$$\delta_{ij} = \begin{cases} 1 & \text{für } i = j, \quad i,j = 1,\dots,n, \\ 0 & \text{für } i \neq j, \quad i,j = 1,\dots,n. \end{cases}$$

2. Wir nehmen uns das im vorhergehenden Abschnitt angedeutete Beispiel der Parallelprojektion $\boldsymbol{P}$ des $\mathbb{R}^2$ auf eine Ursprungsgerade längs einer vorgeschriebenen Richtung vor. Wählt man die Basis geeignet, so wird die Projektion $\boldsymbol{P}$ durch die Matrix $\boldsymbol{P}$ bezüglich des Basispaares $\mathfrak{B}$, $\mathfrak{B}$ wie folgt beschrieben werden:

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

Wir sehen an diesem Beispiel die Zweckmäßigkeit der geeigneten Basisauswahl. Man wird die Basen jeweils so wählen, daß übersichtliche Matrixdarstellungen zustande kommen.

3. Für das ebenfalls im vorhergehenden Abschnitt angedeutete Beispiel der axialen Streckung zeige man durch geeignete Basisauswahl eine Matrixdarstellung der Form

$$\boldsymbol{S} = \begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix}.$$

6.3. Algebraische Operationen für lineare Abbildungen. Matrizenkalkül

Für lineare Abbildungen des $\mathbb{R}^n$ in den $\mathbb{R}^m$ kann man auf eine uns schon geläufige Art eine Addition und Multiplikation mit reellen Skalaren erklären. Wir werden nämlich ganz entsprechend wie mit den Linearformen zu verfahren haben.

Definition 1 (**Punktweise Addition und Multiplikation mit einem Skalar für lineare Abbildungen**). Es bezeichne $\mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)$ das System aller linearen Abbildungen des $\mathbb{R}^n$ in den $\mathbb{R}^m$. Unter der *punktweisen Summe zweier linearer Abbildungen* $A, B \in \mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)$ versteht man die Abbildung $A + B\colon \mathbb{R}^n \to \mathbb{R}^m$ mit dem Verlauf $(A + B)(\boldsymbol{x}) := A(\boldsymbol{x}) + B(\boldsymbol{x})$ für alle $\boldsymbol{x} \in \mathbb{R}^n$ (hierbei meint natürlich $A(\boldsymbol{x}) + B(\boldsymbol{x})$ die koordinatenweise Addition im $\mathbb{R}^m$).

Unter der *punktweisen Multiplikation einer linearen Abbildung* $A \in \mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)$ *mit einem Skalar* $\alpha \in \mathbb{R}$ versteht man die Abbildung $\alpha A\colon \mathbb{R}^n \to \mathbb{R}^m$ mit dem Verlauf

$(\alpha A)(\boldsymbol{x}) := \alpha A(\boldsymbol{x})$ für alle $\boldsymbol{x} \in \mathbf{R}^n$ (hierbei meint natürlich $\alpha A(\boldsymbol{x})$ die koordinatenweise Multiplikation von α mit $A(\boldsymbol{x})$ im $\mathbf{R}^m$).

Satz 1 (Arithmetische Struktur des Systems aller linearen Abbildungen des $\mathbf{R}^n$ in den $\mathbf{R}^m$ hinsichtlich der punktweisen Addition und Multiplikation mit einem Skalar). *Die in $\mathscr{L}(\mathbf{R}^n, \mathbf{R}^m)$ erklärte binäre Operation der punktweisen Addition von linearen Abbildungen und die punktweise Multiplikation mit reellen Skalaren haben die folgenden Grundeigenschaften:*

I. *Für die Addition $+$ gilt:*

1. *$A + B = B + A$ für alle $A, B \in \mathscr{L}(\mathbf{R}^n, \mathbf{R}^m)$ (Kommutativität der Addition).*

2. *$(A + B) + C = A + (B + C)$ für alle $A, B, C \in \mathscr{L}(\mathbf{R}^n, \mathbf{R}^m)$ (Assoziativität der Addition).*

3. *Es gibt ein eindeutig bestimmtes Element $0 \in \mathscr{L}(\mathbf{R}^n, \mathbf{R}^m)$, so daß $A + 0 = A$ für alle $A \in \mathscr{L}(\mathbf{R}^n, \mathbf{R}^m)$ ist (Existenz und Einzigkeit des Nullelementes).*

4. *Zu jedem $A \in \mathscr{L}(\mathbf{R}^n, \mathbf{R}^m)$ existiert ein (eindeutig bestimmtes) Element, bezeichnet durch $-A$, für welches $A + (-A) = 0$ ist (Existenz und Einzigkeit des Inversen).*

II. *Für die Multiplikation mit einem reellen Skalar gilt:*

5. *$1A = A$ für alle $A \in \mathscr{L}(\mathbf{R}^n, \mathbf{R}^m)$.*

6. *$(\alpha\beta) A = \alpha(\beta A)$ für alle $\alpha, \beta \in \mathbf{R}$ und alle $A \in \mathscr{L}(\mathbf{R}^n, \mathbf{R}^m)$ (Assoziativität der Multiplikation mit einem Skalar).*

III. *Für das Zusammenspiel der punktweisen Addition mit der Multiplikation mit einem reellen Skalar gilt:*

7. *$(\alpha + \beta) A = \alpha A + \beta A$ für alle $\alpha, \beta \in \mathbf{R}$ und alle $A \in \mathscr{L}(\mathbf{R}^n, \mathbf{R}^m)$ (Distributivität der Multiplikation mit einem Skalar bezüglich der Addition von Skalaren).*

8. *$\alpha(A + B) = \alpha A + \alpha B$ für alle $\alpha \in \mathbf{R}$ und alle $A, B \in \mathscr{L}(\mathbf{R}^n, \mathbf{R}^m)$ (Distributivität der Multiplikation mit einem Skalar bezüglich der punktweisen Addition).*

Bemerkung. Der Spezialfall $\mathscr{L}(\mathbf{R}^n, \mathbf{R}^1)$ ist uns schon als System $\mathscr{L}(\mathbf{R}^n)$ der sämtlichen Linearformen auf dem $\mathbf{R}^n$ vertraut. Wir werden daher den ohne neue Ideen ablaufenden Beweis in der jetzigen Situation nicht reproduzieren. Die Details mache sich der Leser als nützliche Wiederholung selber klar.

Wir weisen nun auf einen wichtigen Umstand hin: Lineare Abbildungen können unter Bezugnahme auf Basen durch Matrizen beschrieben werden. Die algebraischen Operationen für lineare Abbildungen müssen sich folglich in gewissen Operationen für die Matrizen widerspiegeln. Wir zielen also auf ein gewisses Rechnen mit Matrizen ab. Ein solches Rechnen mit Matrizen — der sogenannte Matrizenkalkül — hat mit dem englischen Mathematiker A. Cayley im Jahre 1858 begonnen.

Definition 2 (Matrizenaddition und Multiplikation mit einem Skalar). Es bezeichne $\mathscr{M}(m \times n)$ die Menge aller reellen Matrizen vom Typ $m \times n$. Für Matrizen $\boldsymbol{A}, \boldsymbol{B} \in \mathscr{M}(m \times n)$ ist die *Matrixsumme* $\boldsymbol{A} + \boldsymbol{B}$ als Matrix vom Typ $m \times n$ mit dem allgemeinen Element $a_{ik} + b_{ik}$ erklärt, sofern

$$\boldsymbol{A} = (a_{ik})_{\substack{i=1,\ldots,m \\ k=1,\ldots,n}}, \qquad \boldsymbol{B} = (b_{ik})_{\substack{i=1,\ldots,m \\ k=1,\ldots,n}}$$

ist.

Für die Matrix $\boldsymbol{A} \in \mathcal{M}(m \times n)$ ist das *Produkt mit einem Skalar* $\alpha \in \mathsf{R}$ als Matrix $\alpha \boldsymbol{A}$ mit dem allgemeinen Element αa_{ik} erklärt. Wir erkennen unmittelbar den folgenden

Satz 2 (Summe und skalares Vielfaches von linearen Abbildungen und ihre Beschreibung durch Matrizen). *Es seien A, B lineare Abbildungen des R^n in den R^m und α sei ein reeller Skalar, $\alpha \in \mathsf{R}$. Sind $\boldsymbol{A}$ und $\boldsymbol{B}$ die assoziierten Matrizen von A, B bezüglich gewisser fixierter Basen $\mathfrak{C}$ und $\mathfrak{D}$ im R^n bzw. R^m, so gilt folgendes:*

1. *Der punktweisen Summe $A + B$ der linearen Abbildungen A, B kommt bezüglich des Basispaares $\mathfrak{C}$, $\mathfrak{D}$ in R^n, R^m als beschreibende Matrix die Matrix $\boldsymbol{A} + \boldsymbol{B}$ zu.*

In Diagrammform:

$$\begin{array}{ccc} \underbrace{A, B} \in \mathcal{L}(\mathsf{R}^n, \mathsf{R}^m) & \longmapsto & A + B \in \mathcal{L}(\mathsf{R}^n, \mathsf{R}^m) \\ \Big\uparrow \mathfrak{C}, \mathfrak{D} & & \Big\uparrow \mathfrak{C}, \mathfrak{D} \\ \widehat{\boldsymbol{A}, \boldsymbol{B}} \in \mathcal{M}(m \times n) & \longmapsto & \boldsymbol{A} + \boldsymbol{B} \in \mathcal{M}(m \times n) \end{array}$$

2. *Der punktweisen Multiplikation αA der linearen Abbildung A mit dem Skalar α kommt bezüglich des Basispaares B, C in R^n, R^m als beschreibende Matrix die Matrix $\alpha \boldsymbol{A}$ zu.*

In Diagrammform:

$$\begin{array}{ccc} A \in \mathcal{L}(\mathsf{R}^n, \mathsf{R}^m) & \longmapsto & \alpha A \in \mathcal{L}(\mathsf{R}^n, \mathsf{R}^m) \\ \Big\uparrow \mathfrak{C}, \mathfrak{D} & & \Big\uparrow \mathfrak{C}, \mathfrak{D} \\ \boldsymbol{A} \in \mathcal{M}(m \times n) & \longmapsto & \alpha \boldsymbol{A} \in \mathcal{M}(m \times n) \end{array}$$

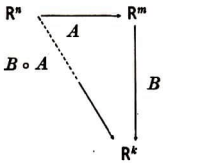
Bemerkung. Bei fixierten Basispaaren $(\mathfrak{C}, \mathfrak{D})$ für R^n, R^m besteht also eine eineindeutige Entsprechung zwischen $\mathcal{L}(\mathsf{R}^n, \mathsf{R}^m)$ und $\mathcal{M}(m \times n)$, die sogar die arithmetische Struktur invariant läßt. Wir können also sagen, daß das System der Matrizen $\mathcal{M}(m \times n)$ hinsichtlich der Matrizenaddition und Multiplikation mit Skalaren isomorph ist zu $\mathcal{L}(\mathsf{R}^n, \mathsf{R}^m)$ hinsichtlich der punktweisen Addition und Multiplikation mit Skalaren. Bei den Matrizen aus $\mathcal{M}(m \times n)$ handelt es sich nun um alle Abbildungen von $\{1, 2, \ldots, m\} \times \{1, 2, \ldots, n\}$ in R. Wir sehen somit, daß das System $\mathcal{M}(m \times n)$ bezüglich der Matrizenaddition und Multiplikation mit Skalaren seinerseits isomorph ist zu dem arithmetischen $m \cdot n$-dimensionalen reellen Zahlenraum R^{mn}.

Eine weitere Operation für lineare Abbildungen ist bedeutungsvoll: die Hintereinanderschaltung oder Zusammensetzung.

Definition 3 (Hintereinanderschaltung von linearen Abbildungen). Es sei $A: \mathbb{R}^n \to \mathbb{R}^m$ eine lineare Abbildung des $\mathbb{R}^n$ in den $\mathbb{R}^m$ und $B: \mathbb{R}^m \to \mathbb{R}^k$ eine lineare Abbildung des $\mathbb{R}^m$ in den $\mathbb{R}^k$. Die zusammengesetzte Abbildung $B \circ A: \mathbb{R}^n \to \mathbb{R}^k$ ist dann wie folgt erklärt: $(B \circ A)(x) := B(A(x))$ für alle $x \in \mathbb{R}^n$. Diese *Hintereinanderschaltung* von linearen Abbildungen ergibt wieder eine lineare Abbildung (Bestätigung!).

Wir machen extra noch einmal darauf aufmerksam, daß zur Hintereinanderschaltung von linearen Abbildungen A, B in der Form $B \circ A$ zuerst A ausgeführt wird und dann B und daß dabei der Wertebereich der zuerst auszuführenden Abbildung mit dem Definitionsbereich der danach auszuführenden Abbildung übereinstimmen muß.

In Diagrammform:



$$\left.\begin{array}{l} A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \\ \\ B \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^k) \end{array}\right\} \Rightarrow B \circ A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^k).$$

Bemerkung. Die Hintereinanderschaltung von Abbildungen nennt man bisweilen auch das *Produkt von Abbildungen*.

Jetzt wird die wichtige Frage nach der Widerspiegelung der Zusammensetzung von linearen Abbildungen an Hand von assoziierten Matrizen erörtert.

In $\mathbb{R}^n$, $\mathbb{R}^m$ und $\mathbb{R}^k$ seien angeordnete Basen

$$\mathfrak{B}: \boldsymbol{b}_1, \ldots, \boldsymbol{b}_n; \qquad \mathfrak{C}: \boldsymbol{c}_1, \ldots, \boldsymbol{c}_m; \qquad \mathfrak{D}: \boldsymbol{d}_1, \ldots, \boldsymbol{d}_k$$

fixiert. Für gegebene lineare Abbildungen $A: \mathbb{R}^n \to \mathbb{R}^m$ und $B: \mathbb{R}^m \to \mathbb{R}^k$ mögen $\boldsymbol{A} \in \mathcal{M}(m \times n)$, $\boldsymbol{B} \in \mathcal{M}(k \times m)$ die zu den Basispaaren $\mathfrak{B}, \mathfrak{C}$ bzw. $\mathfrak{C}, \mathfrak{D}$ gehörigen beschreibenden Matrizen von A und B sein. Welche Matrix vom Typ $(k \times n)$ gehört zu $B \circ A$ bezüglich des Basispaares $\mathfrak{B}, \mathfrak{D}$? Die gesuchte Matrix $\boldsymbol{C}$ hat in ihrer ersten Spalte das Koordinatentupel von $(B \circ A)(\boldsymbol{b}_1)$ bezüglich $\mathfrak{D}$, in ihrer zweiten Spalte das Koordinatentupel von $(B \circ A)(\boldsymbol{b}_2)$ bezüglich $\mathfrak{D}, \ldots$, in ihrer n-ten Spalte das Koordinatentupel von $(B \circ A)(\boldsymbol{b}_n)$ bezüglich $\mathfrak{D}$.

Unter Benutzung von

$$\boldsymbol{A} = (a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}, \qquad \boldsymbol{B} = (b_{\mu\nu})_{\substack{\mu=1,\ldots,k \\ \nu=1,\ldots,m}}$$

erhalten wir

$$A(\boldsymbol{b}_j) = \sum_{i=1}^{m} a_{ij}\boldsymbol{c}_i, \qquad (B \circ A)(\boldsymbol{b}_j) = \sum_{i=1}^{m} a_{ij} B(\boldsymbol{c}_i), \qquad B(\boldsymbol{c}_i) = \sum_{\mu=1}^{k} b_{\mu i}\boldsymbol{d}_\mu,$$

also insgesamt

$$(B \circ A)(\boldsymbol{b}_j) = \sum_{i=1}^{m} a_{ij} \sum_{\mu=1}^{k} b_{\mu i} \boldsymbol{d}_\mu = \sum_{\mu=1}^{k} \left(\sum_{i=1}^{m} b_{\mu i} a_{ij} \right) \boldsymbol{d}_\mu .$$

Das allgemeine Element von $\boldsymbol{C}$ ist demnach

$$c_{\mu j} = \sum_{i=1}^{m} b_{\mu i} a_{ij} .$$

Definition 4 (Matrizenmultiplikation). Es seien

$$\boldsymbol{B} = (b_{\mu\nu})_{\substack{\mu=1,\dots,k \\ \nu=1,\dots,m}} \qquad \boldsymbol{A} = (a_{ij})_{\substack{i=1,\dots,l \\ j=1,\dots,n}}$$

zwei Matrizen des Typs $k \times m$ bzw. $l \times n$. Diese heißen *verkettet* genau dann, wenn $m = l$ ist. Für zwei verkettete Matrizen $\boldsymbol{B}, \boldsymbol{A}$ wird ihr *Matrizenprodukt* $\boldsymbol{B} \cdot \boldsymbol{A}$ als die folgende Matrix erklärt: $\boldsymbol{B} \cdot \boldsymbol{A}$ ist vom Typ $k \times n$ und hat das allgemeine Element

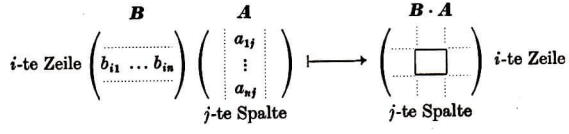
$$c_{\mu j} = \sum_{\nu=1}^{m} b_{\mu\nu} a_{\nu j} . \qquad (\mu = 1, \dots, k; \quad j = 1, \dots, n).$$

Merkregel zur Matrizenmultiplikation

1. Verkettungsbedingung

Für zwei Matrizen $\boldsymbol{B}$ und $\boldsymbol{A}$ ist ihr Produkt $\boldsymbol{B} \cdot \boldsymbol{A}$ nur dann erklärt, wenn der linksstehende Faktor genau so viele Spalten besitzt, wie der rechtsstehende Faktor Zeilen aufweist.

2. Aufbau der Produktmatrix



Das Element in der i-ten Zeile und j-ten Spalte der Matrix $\boldsymbol{B} \cdot \boldsymbol{A}$ hat die Gestalt

$$b_{i1}a_{1j} + b_{i2}a_{2j} + \dots + b_{in}a_{nj} .$$

Dieses Element in der Kreuzung der i-ten Zeile und j-ten Spalte von $\boldsymbol{B} \cdot \boldsymbol{A}$ erhält man als Summe der fortlaufend gebildeten Produkte der Elemente der i-ten Zeile von $\boldsymbol{B}$ mit den Elementen der j-ten Spalte von $\boldsymbol{A}$.

Wir notieren uns den folgenden

Satz 2 (Produkt von linearen Abbildungen und seine Beschreibung durch Matrizen). *Es seien* $A: \mathsf{R}^n \to \mathsf{R}^m$, $B: \mathsf{R}^m \to \mathsf{R}^k$ *lineare Abbildungen. Sind* $\boldsymbol{A}$ *und* $\boldsymbol{B}$ *die assoziierten Matrizen von* A, B *bezüglich gewisser fixierter Basen* $\mathfrak{B}$, $\mathfrak{C}$, $\mathfrak{D}$ *im* R^n, R^m, R^k,

so gilt folgendes: Dem Produkt $B \circ A : R^n \to R^k$ *der linearen Abbildungen* A, B *kommt bezüglich des Basispaares* $\mathfrak{B}$, $\mathfrak{D}$ *in* R^n, R^k *als beschreibende Matrix die Matrix* $B \cdot A$ *zu. In Diagrammform:*

$$\begin{array}{ccccc} A \in \mathcal{L}(R^n, R^m), & & B \in \mathcal{L}(R^m, R^k) & \longmapsto & B \circ A \in \mathcal{L}(R^n, R^k) \\ \Big\uparrow \mathfrak{B}, \mathfrak{C} & & \Big\uparrow \mathfrak{C}, \mathfrak{D} & & \Big\uparrow \mathfrak{B}, \mathfrak{D} \\ A \in \mathcal{M}(m \times n), & & B \in \mathcal{M}(k \times m) & \longmapsto & B \cdot A \in \mathcal{M}(k \times n) \end{array}$$

Bei der Hintereinanderschaltung von linearen Abbildungen und der entsprechenden Operation für die Matrizen war die Bezeichnungsweise Produkt gebraucht worden. Eine Rechtfertigung findet diese Bezeichnungsweise, wenn man sich das Zusammenspiel dieser Produktbildung mit der Addition von Abbildungen bzw. Matrizen vorhält, wie es jetzt geschehen soll.

Satz 3 (Grundeigenschaften der Matrizenmultiplikation). *Es seien* A, B, C *Matrizen. Dann gilt folgendes:*

1. *Sind* B *und* C *vom gleichen Typ und ist* A *mit* B *und mit* C *verkettet, so besteht die Beziehung* $A \cdot (B + C) = A \cdot B + A \cdot C$ (*Distributivgesetz der Multiplikation in bezug auf die Addition*). (Entsprechend mit Multiplikation von rechts, sofern die auftretenden Operationen ausführbar sind).

2. *Es bezeichne* O *die Nullmatrix eines beliebigen Typs. Dann bestehen für eine beliebige Matrix* A *bei entsprechender Verkettung mit* O *die Beziehungen*

$$A \cdot O = O,$$

$$O \cdot A = O.$$

3. *Es bezeichne* I *die (quadratische) Einheitsmatrix eines beliebigen Typs. Dann bestehen für eine beliebige Matrix* A *bei entsprechender Verkettung mit* I *die Beziehungen*

$$A \cdot I = A,$$

$$I \cdot A = A$$ (*Rechtfertigung der Bezeichnungsweise Einheitsmatrix*).

4. *Es sei die Matrix* A *mit der Matrix* B *verkettet und es sei* B *mit* A *verkettet. Dann gilt im allgemeinen*

$$A \cdot B \neq B \cdot A$$ (*Nichtkommutativität der Matrizenmultiplikation*).

5. *Es gibt von der Nullmatrix verschiedene Matrizen* A, B, *die verkettet sind, mit der Eigenschaft*

$$A \cdot B = O$$ (*Existenz von Nullteilern bei der Matrizenmultiplikation*).

6. *Es seien Matrizen* $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ *in der angeschriebenen Reihenfolge verkettet, dann besteht die Beziehung*

$$\boldsymbol{A} \cdot (\boldsymbol{B} \cdot \boldsymbol{C}) = (\boldsymbol{A} \cdot \boldsymbol{B}) \cdot \boldsymbol{C} \quad (\textit{Assoziativität der Matrizenmultiplikation}).$$

Beweis. Die Bestätigung ist einfach zu erbringen. Man kann zwei verschiedene Wege einschlagen. Der eine besteht in der direkten Anwendung der Matrizenmultiplikation. Der andere interpretiert die Matrizen als Darstellungen von linearen Abbildungen und läuft auf eine Bestätigung der analogen Beziehungen für die linearen Abbildungen hinaus. Die vorhergehenden Sätze garantieren dann, daß auch die entsprechenden Matrizenbeziehungen gelten. Der zweite Weg ist sogar weniger schreibaufwendig, was sich besonders für die Assoziativität zeigt. Die Details führe der Leser selber aus. Wir machen uns lediglich die Aussagen 4 und 5 klar.

Zu 4. Es sei

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad \boldsymbol{B} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Dann haben wir

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

also $\boldsymbol{A} \cdot \boldsymbol{B} \neq \boldsymbol{B} \cdot \boldsymbol{A}$.

Welche Abbildungen werden durch beide Matrizen im $\mathbb{R}^2$ hinsichtlich der natürlichen Basis beschrieben? Wir verfolgen diese Abbildungen in der euklidisch-geometrischen Veranschaulichung des $\mathbb{R}^2$. Es gilt für die der Matrix $\boldsymbol{A}$ entsprechende Abbildung $A: \boldsymbol{x} \mapsto A(\boldsymbol{x})$, wenn $\boldsymbol{x} = (x_1, x_2)$, so ist

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ -x_1 \end{pmatrix},$$

also

$$(x_1, x_2) \underset{A}{\mapsto} (x_2, -x_1).$$

Es ist für die der Matrix $\boldsymbol{B}$ entsprechende Abbildung $B: \boldsymbol{x} \mapsto B(\boldsymbol{x})$ bei $\boldsymbol{x} = (x_1, x_2)$

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix},$$

also

$$(x_1, x_2) \underset{B}{\mapsto} (x_1, -x_2).$$

Abb. 7 veranschaulicht die Wirkung der im Text beschriebenen Abbildungen. $\boldsymbol{A}$ beschreibt demzufolge eine Rechtsdrehung um einen rechten Winkel. $\boldsymbol{B}$ beschreibt demzufolge eine Spiegelung an der x-Achse. Die Hintereinanderschaltung $B \circ A$ der Rechtsdrehung um einen rechten Winkel mit der Spiegelung an der x-Achse ist von der Hintereinanderschaltung $A \circ B$ verschieden.

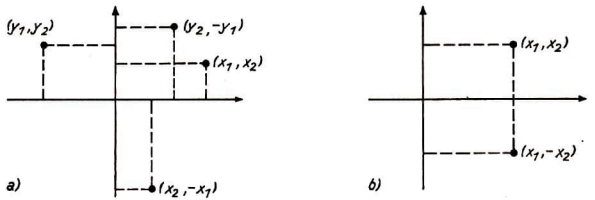


Abb. 7

Zu 5. Auch hier beziehen wir uns auf die euklidisch-geometrische Veranschaulichung des $\mathbb{R}^2$. Eine Orthogonalprojektion des $\mathbb{R}^2$ auf die x-Achse wird durch die Matrix

$$\boldsymbol{P}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

beschrieben.

Eine Orthogonalprojektion des $\mathbb{R}^2$ auf die y-Achse wird durch die Matrix

$$\boldsymbol{P}_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

beschrieben.

Die Hintereinanderschaltung der Orthogonalprojektion des $\mathbb{R}^2$ auf die x-Achse mit der Orthogonalprojektion auf die y-Achse ergibt aber die Nullabbildung, die jeden Punkt des $\mathbb{R}^2$ in den 0-Punkt abbildet. Demzufolge muß $\boldsymbol{P}_2 \cdot \boldsymbol{P}_1 = \boldsymbol{O}$ gelten (was man auch leicht durch direktes Ausrechnen bestätigt). Es ergibt sich auch $\boldsymbol{P}_1 \cdot \boldsymbol{P}_2 = \boldsymbol{O}$.

Bemerkungen.

1. Schauen wir zurück auf den durch eine assoziierte Matrix einer linearen Abbildung vermittelten Übergang der Koordinatentupel. Wir erkennen folgendes: Es sei A eine lineare Abbildung des $\mathbb{R}^n$ in den $\mathbb{R}^m$ und $\boldsymbol{A}$ die assoziierte Matrix von A bezüglich eines Basispaares $\mathfrak{B}$, $\mathfrak{C}$ in $\mathbb{R}^n$, $\mathbb{R}^m$. Es sei $\boldsymbol{x} \in \mathbb{R}^n$ und $\boldsymbol{\xi}$ sein Koordinatentupel bezüglich der Basis $\mathfrak{B}$. Dann hat das Element $A(\boldsymbol{x})$ bezüglich der Basis $\mathfrak{C}$ gerade das Koordinatentupel

$$\eta = \text{Matrizenprodukt von } \boldsymbol{A} \text{ mit der Matrix } \boldsymbol{\xi},$$

hierbei ist $\boldsymbol{\xi}$ als Matrix vom Typ $n \times 1$ und η als Matrix vom Typ $m \times 1$ aufgefaßt. In genauer Übereinstimmung mit der Matrixerklärung sollte eigentlich

$$\eta = \begin{pmatrix} \eta_{11} \\ \eta_{21} \\ \vdots \\ \eta_{m1} \end{pmatrix} \quad \text{und} \quad \xi = \begin{pmatrix} \xi_{11} \\ \xi_{21} \\ \vdots \\ \xi_{n1} \end{pmatrix}$$

geschrieben werden. Der zweite Index als Kennzeichnung der Spalte erübrigt sich aber, da nur eine einzige Spalte auftritt.

Bei der Matrizenmultiplikation ist also wesentlich, daß ein n-Tupel von Zahlen als eine Matrix vom Typ $n \times 1$ (eine Spalte) oder aber als eine Matrix vom Typ $1 \times n$ (eine Zeile) betrachtet werden kann. Es hat sich in diesem Zusammenhang folgende Bezeichnung eingebürgert: Ein Zahlentupel (mit n Komponenten) als eine Zeile aufgefaßt (Matrix vom Typ $1 \times n$) heißt ein *Zeilenvektor*, als eine Spalte aufgefaßt (Matrix vom Typ $n \times 1$) heißt es ein *Spaltenvektor*. Über die Bedeutung des Wortes „Vektor" vergleiche man Kapitel 9.

2. Es seien $\boldsymbol{x}$, $\boldsymbol{y}$ zwei Elemente des R^n, sie mögen die Komponenten x_i bzw. y_i, $i = 1, \ldots, n$, haben. Wird $\boldsymbol{x}$ als Zeilenvektor und $\boldsymbol{y}$ als Spaltenvektor aufgefaßt, so ist das Matrizenprodukt $\boldsymbol{x} \cdot \boldsymbol{y}$ erklärt, es ergibt sich eine Matrix vom Typ 1×1, das einzige Element ist

$$\sum_{i=1}^{n} a_i b_i .$$

Wir haben damit eine Abbildung von $\mathsf{R}^n \times \mathsf{R}^n$ in R.

Dieses spezielle Matrizenprodukt für Elemente aus dem R^n heißt das *innere Produkt* oder — weil das Ergebnis stets ein Skalar ist — das *Skalarprodukt* von $\boldsymbol{x}$ mit $\boldsymbol{y}$ (nicht zu verwechseln mit dem skalaren Vielfachen!). Dieser wichtige Spezialfall der Matrizenmultiplikation wird später noch weiter behandelt.

6.4. Kern und Bildraum linearer Abbildungen. Der Rang von Matrizen

Mit einer linearen Abbildung $A: \mathsf{R}^n \to \mathsf{R}^m$ gehen zwei lineare Teilräume, der Kern der Abbildung und der Bildraum der Abbildung einher, die einen gewissen Aufschluß über die Abbildung gestatten. Wir hatten die Nützlichkeit des Kerns von Linearformen schon früher feststellen können.

Definition 1 (Kern und Bildraum einer linearen Abbildung). Es sei $A: \mathsf{R}^n \to \mathsf{R}^m$ eine lineare Abbildung des R^n in den R^m. Die Menge $\{\boldsymbol{x}: \boldsymbol{x} \in \mathsf{R}^n, A(\boldsymbol{x}) = \boldsymbol{0}\}$ ist dann ein linearer Teilraum des R^n; er heißt der *Kern der Abbildung* A und wird bezeichnet mit $\ker A$. Die Menge $\{\boldsymbol{y}: \boldsymbol{y} \in \mathsf{R}^m$, es existiert ein $\boldsymbol{x} \in \mathsf{R}^n$ mit $A(\boldsymbol{x}) = \boldsymbol{y}\}$ ist dann ein linearer Teilraum des R^m, er heißt der *Bildraum der linearen Abbildung* A und wird bezeichnet mit $\operatorname{im} A$.

Der Leser bestätige die Teilraumeigenschaften!

Bemerkung. Der linearen Abbildung $A: \mathsf{R}^n \to \mathsf{R}^m$ entspricht bezüglich der natürlichen Basen in R^n, R^m eine Matrix $\boldsymbol{A} = (a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$. Der Kern von A ist dann gleichbedeutend mit dem Lösungsraum des homogenen linearen Gleichungs-

systems

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0.$$

(Die Anwendung des Matrizenproduktes gestattet die abkürzende Schreibweise als Matrizengleichung

$$A \cdot x = 0,$$

wobei $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ als Spaltenvektor erscheinen muß und 0 das Nullelement im $\mathbb{R}^m$ ist, das als Spaltenvektor aufgefaßt wird.)

Wenn die Räume $\ker A$ und $\operatorname{im} A$ für eine lineare Abbildung wirklich gewissen Einblick in die Wirkung von A ermöglichen, so werden gewiß die Dimensionszahlen $\dim \ker A$ und $\dim \operatorname{im} A$ bedeutungsvoll sein.

Definition 2 (Rang einer linearen Abbildung). Es sei $A: \mathbb{R}^n \to \mathbb{R}^m$ eine lineare Abbildung des $\mathbb{R}^n$ in den $\mathbb{R}^m$. Die Dimension des Bildraumes von A heißt dann der *Rang der linearen Abbildung.*

$$\operatorname{rang} A := \dim \operatorname{im} A.$$

Wir stellen zuerst die folgende Aussage fest:

Satz 1 (Beziehung zwischen den Dimensionen des Kerns und des Bildraumes einer linearen Abbildung). *Es sei $A: \mathbb{R}^n \to \mathbb{R}^m$ eine lineare Abbildung. Dann gilt die folgende Beziehung:*

$$\dim \ker A + \dim \operatorname{im} A = n.$$

Beweis.

1. **Fall:** $\dim \ker A = 0$. Es gilt: $\dim \ker A = 0 \Leftrightarrow A$ ist eine eineindeutige lineare Abbildung. Eine eineindeutige lineare Abbildung bedeutet aber einen Isomorphismus von $\mathbb{R}^n$ mit einem linearen Teilraum von $\mathbb{R}^m$. Bei einem Isomorphismus bleibt die Dimension invariant. Also ist $\dim \operatorname{im} A = n$.

2. **Fall:** $\dim \ker A = k > 0$. Es sei $b_1, \ldots, b_k$ eine Basis von $\ker A$. Diese kann dann zu einer vollen Basis $b_1, \ldots, b_k; b_{k+1}, \ldots, b_n$ des $\mathbb{R}^n$ ergänzt werden. Ist $k = n$, so haben wir $\operatorname{im} A = \{0\}$. Ist hingegen $k < n$, so spannen die Elemente $A(b_{k+1}), \ldots, A(b_n)$ den Teilraum $\operatorname{im} A$ auf. Diese Elemente sind notwendig linear unabhängig! Denn aus

$$\alpha_{k+1} A(b_{k+1}) + \cdots + \alpha_n A(b_n) = 0$$

folgt

$$A(\alpha_{k+1} b_{k+1} + \cdots + \alpha_n b_n) = 0.$$

Nun liegen aber nur die Linearkombinationen von $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_k$ in $\ker A$. Also muß gelten:

$$\alpha_{k+1}\boldsymbol{b}_{k+1} + \cdots + \alpha_n\boldsymbol{b}_n = 0,$$

was wegen der linearen Unabhängigkeit von $\boldsymbol{b}_{k+1}, \ldots, \boldsymbol{b}_n$ die Relation $\alpha_{k+1} = \cdots = \alpha_n = 0$ bedeutet. Demzufolge ist $\dim \operatorname{im} A = n - k$ bestätigt.

Bemerkung. Wir kennen schon den Begriff des Ranges eines linearen Gleichungssystems. Der vorstehende Satz führt uns mit dem Satz über die Struktur der Lösungsmenge eines homogenen linearen Gleichungssystems zu einem Vergleich des Rangbegriffes eines Gleichungssystems und des Rangbegriffes einer linearen Abbildung. Zuerst sprechen wir den Begriff des Ranges eines Gleichungssystems in formal leicht abgeänderter Form wie folgt aus.

Definition 3 (Zeilenrang einer Matrix). Es sei $\boldsymbol{A}$ eine Matrix vom Typ $m \times n$. Unter dem Zeilenrang dieser Matrix versteht man den Rang des durch $\boldsymbol{A}$ bestimmten homogenen Gleichungssystems $\boldsymbol{A} \cdot \boldsymbol{x} = 0$, d. h., der *Zeilenrang* von $\boldsymbol{A}$ ist gleich der Dimension des linearen Teilraumes von R^n, der durch die Zeilenvektoren von $\boldsymbol{A}$ aufgespannt wird.

Dann können wir also sagen:

Es sei $\boldsymbol{A}$ eine Matrix vom Typ $m \times n$. Der Abbildungsrang der durch sie definierten Abbildung $\boldsymbol{x} \to \boldsymbol{A} \cdot \boldsymbol{x}$, $\boldsymbol{x} \in \mathsf{R}^n$ (als Spaltenvektor aufgefaßt), ist gleich dem Zeilenrang der Matrix $\boldsymbol{A}$.

Diese Einsicht können wir noch ein wenig weiter treiben.

Definition 4 (Spaltenrang einer Matrix). Es sei $\boldsymbol{A}$ eine Matrix vom Typ $m \times n$. Unter dem *Spaltenrang* dieser Matrix versteht man die Dimension des linearen Teilraumes von R^m, der durch die Spaltenvektoren von $\boldsymbol{A}$ aufgespannt wird.

Satz 2 (Gleichheit von Zeilen- und Spaltenrang einer Matrix). *Es sei $\boldsymbol{A}$ eine Matrix vom Typ $m \times n$. Dann stimmen Zeilenrang und Spaltenrang von $\boldsymbol{A}$ überein. Man spricht daher einfach von dem Rang einer Matrix. Es gilt* $0 \leqq \operatorname{rang} A \leqq \min\{m, n\}$.

Beweis. Die gegebene Matrix $\boldsymbol{A}$ definiert eine lineare Abbildung $A\colon \boldsymbol{x} \to \boldsymbol{A} \cdot \boldsymbol{x}$, $\boldsymbol{x} \in \mathsf{R}^n$ (als Spaltenvektor aufgefaßt) von R^n in R^m. Es sei der Zeilenrang von $\boldsymbol{A}$ gleich k. Wir sahen schon, daß die Beziehung $k = \operatorname{rang} A$ besteht. Nun ist der Rang von A als Dimension des Bildraumes von A erklärt. Die Spalten von $\boldsymbol{A}$ sind gerade die Koordinatentupel eines Erzeugendensystems von $\operatorname{im} A$ bezüglich der natürlichen Basis im R^m. Es ist der Spaltenrang von $\boldsymbol{A}$ gleich der Dimension des Bildraumes von A, weil der Übergang zu Koordinatentupeln bezüglich einer beliebigen Basis einen Isomorphismus darstellt. Die letzte Beziehung ist klar, da der Zeilenrang kleiner oder gleich der Zeilenzahl sein muß (entsprechend für den Spaltenrang).

Wir müssen nun die Frage beantworten, wie sich der Rang einer Abbildung durch eine beliebige beschreibende Matrix ausdrückt. Die Argumente dazu sind schon vollständig in den soeben erfolgten Erörterungen aufgetaucht.

Satz 3 (Rang einer Abbildung und Rang der assoziierten Matrizen). *Es sei* $A: R^n \to R^m$ *eine lineare Abbildung des* R^n *in den* R^m. *A sei die beschreibende Matrix von A bezüglich eines fixierten Basispaares* $\mathfrak{B}, \mathfrak{C}$ *in* R^n, R^m. *Dann gilt*

$$\operatorname{rang} A = \operatorname{rang} \boldsymbol{A}.$$

Beweis. Die Basis $\mathfrak{B}$ bestehe aus $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$. Die Elemente $A(\boldsymbol{b}_1), \ldots, A(\boldsymbol{b}_m)$ spannen im A auf. Es ist $\operatorname{rang} A = \dim L\big(\{A(\boldsymbol{b}_1), A(\boldsymbol{b}_2), \ldots, A(\boldsymbol{b}_n)\}\big)$. Die Spaltenvektoren von $\boldsymbol{A}$ seien $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n$. Die $\boldsymbol{s}_i$ sind die Koordinatentupel von $A(\boldsymbol{b}_i)$ bezüglich $\mathfrak{C}$. Die Zuordnung $\boldsymbol{y} \in \operatorname{im} A \mapsto$ Koordinatentupel von $\boldsymbol{y}$ bezüglich $\mathfrak{C}$ ist ein Isomorphismus von $\operatorname{im} A$ mit $L(\{\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_n\})$, daher ist $\dim \operatorname{im} A = \dim L(\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n\})$. Der letzte Wert ist aber der Spaltenrang (= Zeilenrang) von $\boldsymbol{A}$.

Nach dem letzten Satz können wir für eine lineare Abbildung bei Kenntnis einer beliebigen beschreibenden Matrix den Rang der Abbildung berechnen, da uns ein Berechnungsverfahren des Ranges eines Gleichungssystems (d. h. einer Matrix) schon geläufig ist. Eine Rangberechnung macht sich beispielsweise bei der Entscheidung nötig, ob eine gegebene lineare Abbildung eineindeutig ist oder nicht.

Satz 4 (Kennzeichnung der Eineindeutigkeit einer linearen Abbildung durch den Rang der Abbildung). *Es sei* $A: R^n \to R^m$ *eine lineare Abbildung des* R^n *in den* R^m. *A ist eineindeutig (d. h. ein Isomorphismus von* R^n *in* R^m*)* $\Leftrightarrow \operatorname{rang} A = n$.

Beweis. Wegen $n = \dim \ker A + \operatorname{rang} A$ ist $\operatorname{rang} A = n$ äquivalent mit $\dim \ker A = 0$. Letzteres ist aber mit der Eineindeutigkeit von A gleichbedeutend. Denn $A(\boldsymbol{x}) = A(\boldsymbol{y})$ ergibt $A(\boldsymbol{x} - \boldsymbol{y}) = \boldsymbol{0}$, d. h. $\boldsymbol{x} = \boldsymbol{y}$.

Abschließend sei noch das Verhalten des Ranges von Matrizen bzw. Abbildungen gegenüber der Produktbildung notiert.

Satz 5 (Abschätzung des Ranges einer Produktmatrix). *Es sei* $\boldsymbol{A}$ *eine Matrix vom Typ* $m \times n$ *und* $\boldsymbol{B}$ *eine Matrix vom Typ* $n \times k$. *Dann gilt für den Rang der Produktmatrix* $\boldsymbol{A} \cdot \boldsymbol{B}$ *die folgende Ungleichung:*

$$\operatorname{rang}(\boldsymbol{A} \cdot \boldsymbol{B}) \leqq \min\{\operatorname{rang} \boldsymbol{A}, \operatorname{rang} \boldsymbol{B}\}.$$

Beweis. $\boldsymbol{A}$ definiert eine Abbildung A von R^n in R^m und $\boldsymbol{B}$ eine Abbildung B von R^k in den R^n. Es ist $\dim \operatorname{im} A \circ B$ abzuschätzen!

Wir haben im $A \supseteq \operatorname{im} A \circ B$, d. h. $\dim \operatorname{im} A \circ B \leqq \dim \operatorname{im} A$. Zum anderen ist $\operatorname{im} A \circ B = A(\operatorname{im} B)$, also $\dim \operatorname{im} A \circ B \leqq \dim \operatorname{im} B$, denn bei einer linearen Abbildung hat der Bildraum eines linearen Teilraumes höchstens die Dimension des Urbildraumes, wie aus der Betrachtung von Erzeugendensystemen hervorgeht.

6.5. Lineare Abbildungen des $\mathbb{R}^n$ in sich. Invertierbare Matrizen

Je zwei lineare Abbildungen des $\mathbb{R}^n$ in sich können hintereinandergeschaltet werden, ebenso können je zwei quadratische Matrizen des Typs $n \times n$ (hierfür sagt man auch: quadratische Matrizen von der Ordnung n) miteinander multipliziert werden.

Die folgende Aussage stellt noch einmal die Grundeigenschaften der arithmetischen Struktur in $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ zusammen. Eine völlig analoge Aussage ist dann für die arithmetische Struktur im System $\mathcal{M}(n \times n)$ der Matrizen der Ordnung n gültig, weil beide Systeme zueinander isomorph sind.

Satz 1 (Arithmetische Struktur der linearen Abbildungen des $\mathbb{R}^n$ in sich bezüglich der Addition und Multiplikation). *Es bezeichne $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ das System aller linearen Abbildungen des $\mathbb{R}^n$ in sich. Man nennt die Elemente von $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ auch bisweilen die Endomorphismen bzw. Operatoren des $\mathbb{R}^n$. $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ hat dann hinsichtlich der punktweisen Addition und Multiplikation mit einem Skalar und hinsichtlich der Hintereinanderschaltung die folgenden Grundeigenschaften:*

I. *Für die Addition gilt:*

1. $A + B = B + A$ *für alle* $A, B \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ (*Kommutativität der Addition*).
2. $(A + B) + C = A + (B + C)$ *für alle* $A, B, C \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ (*Assoziativität der Addition*).
3. *Es gibt ein eindeutig bestimmtes Element* $0 \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$, *so daß* $A + 0 = A$ *für alle* $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ *ist* (*Existenz und Einzigkeit des Nullelementes*).
4. *Zu jedem* $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ *existiert ein* (*eindeutig bestimmtes*) *Element, bezeichnet durch* $-A$, *für welches* $A + (-A) = 0$ *ist* (*Existenz und Einzigkeit der Inversen*).

II. *Für die Multiplikation mit einem reellen Skalar gilt:*

5. $1 \cdot A = A$ *für alle* $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$.
6. $(\alpha\beta)A = \alpha(\beta A)$ *für alle* $\alpha, \beta \in \mathbb{R}$ *und alle* $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ (*Assoziativität der Multiplikation mit einem Skalar*).

III. *Für das Zusammenspiel der punktweisen Addition mit der Multiplikation mit einem reellen Skalar gilt:*

7. $(\alpha + \beta)A = \alpha A + \beta A$ *für alle* $\alpha, \beta \in \mathbb{R}$ *und alle* $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ (*Distributivität der Multiplikation mit einem Skalar bezüglich der Addition von Skalaren*).
8. $\alpha(A + B) = \alpha A + \alpha B$ *für alle* $\alpha \in \mathbb{R}$ *und alle* $A, B \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ (*Distributivität der Multiplikation mit einem Skalar bezüglich der punktweisen Addition*).

IV. *Für die Hintereinanderschaltung gilt:*

9. *Es gibt ein eindeutig bestimmtes Element* $I \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$, *so daß* $A \circ I = I \circ A = A$ *für alle* $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ *ist* (*Existenz und Einzigkeit des Einselementes*).
10. $(A \circ B) \circ C = A \circ (B \circ C)$ *für alle* $A, B, C \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ (*Assoziativität der Hintereinanderschaltung*).
11. $A \circ (B + C) = A \circ B + A \circ C$ *und* $(A + B) \circ C = A \circ C + B \circ C$ *für alle* $A, B, C \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ (*Distributivität der Hintereinanderschaltung bezüglich der Addition*).

Man sagt, $\mathscr{L}(\mathsf{R}^n, \mathsf{R}^n)$ bildet bezüglich der betrachteten Operationen eine Algebra über R (oder eine reelle Algebra). Diese Algebra ist bei $n \geqq 2$ nicht kommutativ, sie besitzt ein Einselement und bei $n \geqq 2$ Nullteiler.

Ein Teilsystem von $\mathscr{L}(\mathsf{R}^n, \mathsf{R}^n)$ ist von Bedeutung. Dazu die folgenden Betrachtungen: Eine lineare Abbildung des R^n in sich, die eineindeutig ist, muß zugleich auch surjektiv sein, denn bei eineindeutiger linearer Abbildung A des R^n in sich ist dim im$A = n$, also besteht die Beziehung im$A = \mathsf{R}^n$. Dann existiert aber zu A die inverse Abbildung $A^{-1}\colon \mathsf{R}^n \to \mathsf{R}^n$, deren Verlauf wie folgt erklärt ist:

$$x \in \mathsf{R}^n \underset{A^{-1}}{\mapsto} y \in \mathsf{R}^n : \Leftrightarrow A(y) = x.$$

A^{-1} gehört wiederum zu $\mathscr{L}(\mathsf{R}^n, \mathsf{R}^n)$.

Satz 2 (Rein arithmetische Kennzeichnung der Invertierbarkeit von linearen Abbildungen). *Es sei A eine lineare Abbildung des R^n in sich. A ist invertierbar (es existiert die inverse Abbildung, d. h., A ist eineindeutig (und surjektiv) $\Leftrightarrow$ Es besteht eine der folgenden äquivalenten Bedingungen:*

1. *Es gibt eine Abbildung $B \in \mathscr{L}(\mathsf{R}^n, \mathsf{R}^n)$ mit $B \circ A = E$ (identische Abbildung des R^n).*

2. *Es gibt eine Abbildung $C \in \mathscr{L}(\mathsf{R}^n, \mathsf{R}^n)$ mit $A \circ C = E$ (identische Abbildung des R^n).*

Bei Invertierbarkeit von A sind dann B und C eindeutig bestimmt, es gilt $B = A^{-1}$, $C = A^{-1}$.

Beweis. Im Fall der Invertierbarkeit von A hat man

$$A \circ A^{-1} = E = A^{-1} \circ A,$$

denn es ist bei der Existenz von A^{-1}

$$x \underset{A}{\mapsto} y \Leftrightarrow y \underset{A^{-1}}{\mapsto} x.$$

Wenn nun für ein gewisses $B \in \mathscr{L}(\mathsf{R}^n, \mathsf{R}^n)$ die Beziehung $B \circ A = E$ besteht (entsprechend für $A \circ C = E$), so gilt für jedes $x \in \mathsf{R}^n$

$$x = B\big(A(x)\big).$$

Demnach folgt aus $A(x_1) = A(x_2)$ die Gleichung $x_1 = x_2$, d. h., A ist eineindeutig. Dann existiert also A^{-1}, und $x = B\big(A(x)\big)$ bedeutet, daß $B = A^{-1}$ ist.

Definition 1 (Invertierbare Matrix). Eine quadratische Matrix $\boldsymbol{A}$ der Ordnung n heißt *invertierbar* genau dann, wenn eine der folgenden äquivalenten Bedingungen gilt:

1. Es existiert eine quadratische Matrix $\boldsymbol{B}$ der Ordnung n, so daß das Produkt $\boldsymbol{B} \cdot \boldsymbol{A}$ gleich der Einheitsmatrix der Ordnung n ist.

2. Es existiert eine quadratische Matrix $\boldsymbol{C}$ der Ordnung n, so daß das Produkt $\boldsymbol{A} \cdot \boldsymbol{C}$ gleich der Einheitsmatrix der Ordnung n ist.

Bemerkung. Wir sehen nach den voraufgegangenen Erörterungen, daß eine zu einer linearen Abbildung $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ assoziierte Matrix $\boldsymbol{A}$ (bei beliebigem Basispaar $\mathfrak{B}, \mathfrak{B}$) genau dann invertierbar ist, wenn die lineare Abbildung A invertierbar ist. Die Matrizen $\boldsymbol{B}$ und $\boldsymbol{C}$ sind dann eindeutig bestimmt, es handelt sich gerade um die assoziierte Matrix von A^{-1} bezüglich des Basispaares $\mathfrak{B}, \mathfrak{B}$ (ein Basispaar $\mathfrak{B}, \mathfrak{C}$ mit $\mathfrak{C} \neq \mathfrak{B}$ kommt bei diesen Betrachtungen nicht in Frage, weil sonst die identische Abbildung E: $\mathbb{R}^n \to \mathbb{R}^n$ mit $\boldsymbol{x} \overset{E}{\mapsto} \boldsymbol{x}$ nicht durch die Einheitsmatrix beschrieben wird). Man nennt diese zu der invertierbaren Matrix $\boldsymbol{A}$ eindeutig bestimmte Matrix, deren Produkt mit $\boldsymbol{A}$ (von links und auch von rechts) die Einheitsmatrix ergibt, die zu $\boldsymbol{A}$ *inverse Matrix*, sie wird mit $\boldsymbol{A}^{-1}$ bezeichnet. Invertierbare Matrizen heißen auch oftmals *reguläre Matrizen*. $\boldsymbol{A}^{-1}$ ist selbst wieder invertierbar, ihre inverse Matrix ist $\boldsymbol{A}$. Ein Kriterium für die Invertierbarkeit (Regularität) ist schon bekannt (Folgerung aus dem Satz über die Kennzeichnung der Eineindeutigkeit von linearen Abbildungen durch den Rang).

Satz 3 (Die Kennzeichnung der Invertierbarkeit einer Matrix durch den Rang). *Eine quadratische Matrix der Ordnung n ist genau dann invertierbar (regulär), wenn die Matrix den Rang n besitzt.*

Unser Matrizenkalkül ist bisher noch unvollständig, da wir kein Verfahren erörtert haben, die inverse Matrix $\boldsymbol{A}^{-1}$ einer invertierbaren Matrix $\boldsymbol{A}$ anzugeben, obgleich wir durch die Rangberechnung die Invertierbarkeit entscheiden können. Der Algorithmus zur Rangberechnung läßt sich aber leicht so ausgestalten, daß er im Falle der Invertierbarkeit gleich die inverse Matrix mitliefert. Das wird weiter unten besprochen. Zuvor noch einige Betrachtungen über das System der invertierbaren Matrizen.

Satz 4 (Struktur des Systems der invertierbaren Matrizen der Ordnung n hinsichtlich der Matrizenmultiplikation). *Das System aller invertierbaren quadratischen Matrizen der Ordnung n (ein Teilsystem von $\mathcal{M}(n \times n)$) hat folgende Eigenschaften in bezug auf die Matrizenmultiplikation:*

1. *Das Produkt zweier invertierbarer quadratischer Matrizen $\boldsymbol{A}$, $\boldsymbol{B}$ der Ordnung n ist wieder eine invertierbare quadratische Matrix der Ordnung n. Es besteht die Beziehung $(\boldsymbol{A} \cdot \boldsymbol{B})^{-1} = \boldsymbol{B}^{-1} \cdot \boldsymbol{A}^{-1}$.*

2. *Es gibt eine (eindeutig bestimmte) invertierbare quadratische Matrix $\boldsymbol{E}$ der Ordnung n, so daß für jede andere invertierbare quadratische Matrix $\boldsymbol{A}$ der Ordnung n $\boldsymbol{A} \cdot \boldsymbol{E} = \boldsymbol{E} \cdot \boldsymbol{A} = \boldsymbol{A}$ gilt.*

3. *Zu jeder invertierbaren quadratischen Matrix $\boldsymbol{A}$ der Ordnung n gibt es eine (eindeutig bestimmte) invertierbare quadratische Matrix $\tilde{\boldsymbol{A}}$ der Ordnung n, so daß $\boldsymbol{A} \cdot \tilde{\boldsymbol{A}} = \tilde{\boldsymbol{A}} \cdot \boldsymbol{A} = \boldsymbol{E}$ gilt.*

Das System aller invertierbaren quadratischen Matrizen der Ordnung n bildet hinsichtlich der Matrizenmultiplikation — wie man sagt — eine Gruppe. Man bezeichnet sie mit $GL(n)$ und nennt sie die generelle lineare Gruppe der Ordnung n.

Beweis. Die ausgesprochenen Eigenschaften erkennt man leicht für die Hintereinanderschaltung der eineindeutigen linearen Abbildungen des R^n in sich. Wegen der Isomorphie zu dem System der invertierbaren Matrizen hinsichtlich der Matrizenmultiplikation hat man die Bestätigung.

Nun erfolgt die Besprechung eines Verfahrens zur Ermittlung der inversen Matrix. Wir suchen für eine quadratische Matrix $\boldsymbol{A}$ eine quadratische Matrix $\boldsymbol{X}$ mit $\boldsymbol{A} \cdot \boldsymbol{X} = \boldsymbol{E}$. Sind die Spaltenvektoren von $\boldsymbol{X}$ nacheinander $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$, so sollen also $\boldsymbol{x}_i \in \mathsf{R}^n$ bestimmt werden mit

$$\boldsymbol{A}\boldsymbol{x}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \boldsymbol{A}\boldsymbol{x}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \boldsymbol{A}\boldsymbol{x}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Das ist aber eine Aufgabenstellung zur Lösung von n Gleichungssystemen in n Unbekannten. Diese Aufgabe läßt sich etwa nach dem Gaußschen Algorithmus erledigen. Wir formulieren diese Einsicht als den folgenden

Satz 5 (Verfahren zur Ermittlung der inversen Matrix). *Es sei* $\boldsymbol{A} = (a_{ij})_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}}$ *eine quadratische Matrix der Ordnung n. Die Frage nach der inversen Matrix* $\boldsymbol{A}^{-1}$, *d. h., die Frage nach deren Existenz und deren konkreter Gestalt, ist gleichbedeutend mit der Auflösung der folgenden n inhomogenen linearen Gleichungssysteme:*

$$\boldsymbol{A}\begin{pmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \boldsymbol{A}\begin{pmatrix} x_{12} \\ x_{22} \\ \vdots \\ x_{n2} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ \ldots, \ \boldsymbol{A}\begin{pmatrix} x_{1n} \\ x_{2n} \\ \vdots \\ x_{nn} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Es tritt dabei genau eine der folgenden Möglichkeiten ein:

1. *Nicht alle n Gleichungssysteme sind lösbar.*
2. *Jedes der n Gleichungssysteme hat genau eine Lösung.*

Im ersten Fall ist $\boldsymbol{A}$ *nicht invertierbar. Im zweiten Fall ist* $\boldsymbol{A}$ *invertierbar, und die zu* $\boldsymbol{A}$ *inverse Matrix lautet*

$$\boldsymbol{A}^{-1} = \begin{pmatrix} x_{11} x_{12} \ldots x_{1n} \\ x_{21} x_{22} \ldots x_{2n} \\ \ldots\ldots\ldots \\ x_{n1} x_{n2} \ldots x_{nn} \end{pmatrix}.$$

Die Entscheidung, welcher Fall vorliegt, erledigt man beispielsweise durch den Gaußschen Algorithmus.

Beispiele zur Invertierbarkeit von Matrizen

Der Gaußsche Algorithmus für die n Gleichungssysteme, die zur inversen Matrix führen, wird gleichzeitig für alle n Systeme ausgeführt.

1. Hat die quadratische Matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

eine Inverse?

Lösung.

Koeffizienten bei			absolut. Gl. bei		
x_{1j}	x_{2j}	x_{3j}	$j=1$	$j=2$	$j=3$
1	2	3	1	0	0
4	5	6	0	1	0
7	8	9	0	0	1

$\longmapsto$

1	2	3	1	0	0
0	−3	−6	−4	1	0
0	−6	−12	−7	0	−6

$\longmapsto$

1	2	3	1	0	0
0	−3	−6	−4	1	0
0	0	0	1	·	·

Also müßte für $j = 1$ die Nullform den Wert 1 annehmen, was nicht möglich ist. Folglich sind die Gleichungssysteme nicht alle lösbar. Demzufolge hat A keine Inverse! (Wir sehen außerdem, daß rang $A = 2$ ist.)

2. Hat die quadratische Matrix

$$A = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

eine Inverse?

Lösung.

Koeffiz. b.			absol. Gl. b.		
x_{1j}	x_{2j}	x_{3j}	$j=1$	$j=2$	$j=3$
1	1	1	1	0	0
−1	1	1	0	1	0
0	0	1	0	0	1

$\longmapsto$

1	1	1	1	0	0
0	2	2	1	1	0
0	0	1	0	0	1

1	0	0	$\frac{1}{2}$	$-\frac{1}{2}$	0
0	1	0	$\frac{1}{2}$	$\frac{1}{2}$	-1
0	0	1	0	0	1

Die erste Umformung bringt die Entscheidung, daß A eine Inverse besitzt, die weitere Umformung liefert die Inverse als

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

Man überzeugt sich durch die Produktbildung $A \cdot A^{-1}$ von der Richtigkeit der Rechnung.

Eine geometrische Interpretation der gegebenen Matrix A als eine beschreibende Abbildungsmatrix hätte uns auch leicht ohne Rechnung die Invertierbarkeit entscheiden lassen. A beschreibt im R^3 die Abbildung, durch welche $(1, 0, 0)$ in $(1, -1, 0)$; $(0, 1, 0)$ in $(1, 1, 0)$ und $(0, 0, 1)$ in $(1, 1, 1)$ übergeht. Die Elemente $(1, -1, 0)$, $(1, 1, 0)$, $(1, 1, 1)$ spannen aber ganz R^3 auf, da die beiden ersten die x, y-Ebene aufspannen und das dritte Element aus dieser Ebene herausweist.

3. Wann hat die quadratische Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

eine Inverse?

Zwei Fälle sind möglich, $a_{11} \neq 0$ oder $a_{12} \neq 0$. Denn bei $a_{11} = 0$ und $a_{12} = 0$ liegt keine Invertierbarkeit vor.

1. Fall: $a_{11} \neq 0$

Koeffiz. b.		absol. Gl. b.	
x_{1j}	x_{2j}	$j = 1$	$j = 2$
a_{11}	a_{12}	1	0
a_{21}	a_{22}	0	1

$\longmapsto$

a_{11}	a_{12}	1	0
0	$a_{22} - \frac{a_{21}}{a_{11}} a_{12}$	$-\frac{a_{21}}{a_{11}}$	1

2. Fall: $a_{12} \neq 0$

Koeffiz. b. x_{2j}	x_{1j}	absol. Gl. b. $j=1$	$j=2$
a_{12}	a_{11}	1	0
a_{22}	a_{21}	0	1

$\longmapsto$

a_{12}	a_{11}	1	0
0	$a_{21} - \frac{a_{22}}{a_{12}} a_{11}$	$-\frac{a_{22}}{a_{12}}$	1

Wir sehen also: Ist $a_{11} \neq 0$, so besteht Invertierbarkeit von $\boldsymbol{A}$ genau dann, wenn

$$a_{22} - \frac{a_{21}}{a_{11}} a_{12} \neq 0.$$

Ist $a_{12} \neq 0$, so besteht Invertierbarkeit von $\boldsymbol{A}$ genau dann, wenn

$$a_{21} - \frac{a_{22}}{a_{12}} a_{11} \neq 0.$$

Beide Bedingungen zusammenfassend kann man also sagen:

$$\boldsymbol{A} \text{ invertierbar} \Leftrightarrow a_{11}a_{22} - a_{12}a_{21} \neq 0.$$

Die Inverse von $\boldsymbol{A}$ lautet dann, wie man durch Weiterführung des obigen Verfahrens erhält:

$$\boldsymbol{A}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}.$$

6.6. Basiswechsel und Koordinatentransformationen

Jetzt untersuchen wir den Übergang von einer Basis des $\mathbb{R}^n$ zu einer neuen Basis und interessieren uns dabei vor allem für die Beschreibung durch Matrizen. Aus dem Bisherigen ersieht man direkt die Gültigkeit der folgenden Aussage.

Satz 1 (Basiswechsel als lineare Abbildung). *Es sei* $\mathfrak{B}: \boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$ *eine (angeordnete) Basis des* $\mathbb{R}^n$ *und* $A: \mathbb{R}^n \to \mathbb{R}^n$ *eine lineare Abbildung des* $\mathbb{R}^n$ *in sich.*

Die Abbildung A *überführt die Basis* $\mathfrak{B}$ *genau dann in eine Basis* $\mathfrak{C}: \boldsymbol{c}_1, \ldots, \boldsymbol{c}_n$ *mit* $A(\boldsymbol{b}_i) = \boldsymbol{c}_i$, $i = 1, \ldots, n$, *wenn* A *den Rang* n *hat. Zu zwei vorgegebenen angeordneten Basen* $\mathfrak{B}: \boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$ *und* $\mathfrak{C}: \boldsymbol{c}_1, \ldots, \boldsymbol{c}_n$ *des* $\mathbb{R}^n$ *gibt es genau eine lineare Abbildung* $A: \mathbb{R}^n \to \mathbb{R}^n$ *des* $\mathbb{R}^n$ *in sich, die die Basis* $\mathfrak{B}$ *in die Basis* $\mathfrak{C}$ *überführt:*

$$A(\boldsymbol{b}_i) = \boldsymbol{c}_i, \quad i = 1, \ldots, n.$$

Ziehen wir die Matrizendarstellung einer linearen Abbildung heran, so kommen wir also zu einer Beschreibung eines Basiswechsels mittels einer Matrix.

Definition 1 (Matrix für einen Basiswechsel). Es seien $\mathfrak{B}$: $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$ und $\mathfrak{C}$: $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n$ zwei (angeordnete) Basen des $\mathbb{R}^n$. Unter der *Matrix für den Basiswechsel* $\mathfrak{B} \mapsto \mathfrak{C}$ soll die beschreibende Matrix der eindeutig bestimmten Übergangsabbildung $A: \mathbb{R}^n \to \mathbb{R}^n$ mit $A(\boldsymbol{b}_i) = \boldsymbol{c}_i$ hinsichtlich der ersten Basis $\mathfrak{B}$ verstanden werden.

Demnach ist die Matrix für den Basiswechsel $\mathfrak{B} \mapsto \mathfrak{C}$

$$\begin{pmatrix} t_{11} & t_{12} & \ldots & t_{1n} \\ t_{21} & t_{22} & \ldots & t_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ t_{n1} & t_{n2} & \ldots & t_{nn} \end{pmatrix}, \text{ wobei } \begin{pmatrix} t_{1i} \\ t_{2i} \\ \vdots \\ t_{ni} \end{pmatrix}$$

das Koordinatentupel von $\boldsymbol{c}_i$ $\left(= A(\boldsymbol{b}_i)\right)$ bezüglich der Basis $\mathfrak{B}$ bedeutet.

Wie berechnen sich nun die Koordinaten eines Elementes des $\mathbb{R}^n$ bezüglich einer gegebenen Basis bei Kenntnis der Matrix des Basiswechsels?

Satz 2 (Koordinatentransformation bei Basiswechsel). *Es seien $\mathfrak{B}$ und $\mathfrak{C}$ zwei (angeordnete) Basen des $\mathbb{R}^n$. $T = (t_{ij})_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}}$ bezeichne die Matrix des Basiswechsels $\mathfrak{B} \mapsto \mathfrak{C}$. Ist*

$$\xi = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$$

das Koordinatentupel von $\boldsymbol{x} \in \mathbb{R}^n$ bezüglich der Basis $\mathfrak{B}$ (als Spaltenvektor aufgefaßt) und

$$\eta = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix}$$

das Koordinatentupel desselben Punktes $\boldsymbol{x} \in \mathbb{R}^n$ bezüglich der Basis $\mathfrak{C}$, so besteht folgender Zusammenhang zwischen ξ und η:

$$\eta = T^{-1}\xi \quad (\text{oder gleichbedeutend: } \xi = T\eta).$$

Bemerkung. Die Koordinatentupel eines Elements des $\mathbb{R}^n$ transformieren sich also bei einem Basiswechsel nicht im gleichen Sinne wie die Basiselemente. Man sagt dazu: Die Koordinaten transformieren sich *kontragredient* zu den Basiselementen.

Beweis des Satzes. Nach den Voraussetzungen des Satzes gilt einerseits

$$\boldsymbol{c}_j = \sum_{i=1}^{n} t_{ij}\boldsymbol{b}_i$$

und andererseits

$$\boldsymbol{x} = \sum_{i=1}^{n} \xi_i \boldsymbol{b}_i = \sum_{j=1}^{n} \eta_j \boldsymbol{c}_j.$$

Durch Einsetzen erhält man daraus

$$\boldsymbol{x} = \sum_{i=1}^{n} \xi_i \boldsymbol{b}_i = \sum_{j=1}^{n} \eta_j \sum_{i=1}^{n} t_{ij} \boldsymbol{b}_i = \sum_{j=1}^{n} \sum_{i=1}^{n} \eta_j t_{ij} \boldsymbol{b}_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} t_{ij} \eta_j \right) \boldsymbol{b}_i .$$

Wegen der Eindeutigkeit der Koordinatendarstellung ergibt sich somit

$$\xi_i = \sum_{j=1}^{n} t_{ij} \eta_j , \qquad i = 1, \ldots, n .$$

Das bedeutet aber gerade

$$\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = \boldsymbol{T} \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix} .$$

Nun ist die den Basiswechsel vermittelnde lineare Abbildung vom Rang n, also ist auch ihre beschreibende Matrix $\boldsymbol{T}$ vom Rang n. $\boldsymbol{T}$ besitzt demzufolge eine Inverse $\boldsymbol{T}^{-1}$. Durch Multiplikation ergibt sich

$$\boldsymbol{T}^{-1} \boldsymbol{\xi} = \eta .$$

Der soeben bewiesene Satz beantwortet die Frage nach der Berechnung der neuen Koordinaten bei Kenntnis der alten Koordinaten und der Matrix $\boldsymbol{T}$ des Basiswechsels. Dazu hat man zuvor die Inverse $\boldsymbol{T}^{-1}$ von $\boldsymbol{T}$ zu bestimmen. Sind hingegen die beiden Basen $\mathfrak{B}: \boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$ und $\mathfrak{C}: \boldsymbol{c}_1, \ldots, \boldsymbol{c}_n$ gegeben und die Matrix des Basiswechsels noch nicht direkt vorgelegt, so wird man zur Berechnung der neuen Koordinaten aus den alten nicht etwa zuerst $\boldsymbol{T}$ ermitteln und darauf die Inverse $\boldsymbol{T}^{-1}$ bestimmen, sondern man wird sogleich $\boldsymbol{T}^{-1}$ errechnen. Dazu notieren wir uns den folgenden

Satz 3 (Matrix eines iterierten Basiswechsels). *Es seien* $\mathfrak{B}$, $\mathfrak{C}$ *und* $\mathfrak{D}$ *gewisse angeordnete Basen des* $\mathbb{R}^n$. *Die Matrix* $\boldsymbol{T}$ *beschreibe den Basiswechsel* $\mathfrak{B} \mapsto \mathfrak{C}$. *Die Matrix* $\boldsymbol{S}$ *beschreibe den Basiswechsel* $\mathfrak{C} \mapsto \mathfrak{D}$. *Dann beschreibt die Matrix* $\boldsymbol{TS}$ *den Basiswechsel* $\mathfrak{B} \mapsto \mathfrak{D}$ (Reihenfolge!).

Beweis. Es sei $\mathfrak{B}: \boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$; $\mathfrak{C}: \boldsymbol{c}_1, \ldots, \boldsymbol{c}_n$ und $\mathfrak{D}: \boldsymbol{d}_1, \ldots, \boldsymbol{d}_n$. Dann gilt

$$\boldsymbol{c}_j = \sum_{i=1}^{n} t_{ij} \boldsymbol{b}_i , \qquad \boldsymbol{d}_l = \sum_{k=1}^{n} s_{kl} \boldsymbol{c}_k ,$$

wobei $\boldsymbol{T} = (t_{ij})_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}}$ und $\boldsymbol{S} = (s_{kl})_{\substack{k=1,\ldots,n \\ l=1,\ldots,n}}$ die jeweiligen Matrizen für den Basiswechsel $\mathfrak{B} \mapsto \mathfrak{C}$ und $\mathfrak{C} \mapsto \mathfrak{D}$ sind. Damit ergibt sich durch Einsetzen

$$\boldsymbol{d}_l = \sum_{k=1}^{n} s_{kl} \sum_{i=1}^{n} t_{ik} \boldsymbol{b}_i = \sum_{k=1}^{n} \sum_{i=1}^{n} t_{ik} s_{kl} \boldsymbol{b}_i = \sum_{i=1}^{n} \left(\sum_{k=1}^{n} t_{ik} s_{kl} \right) \boldsymbol{b}_i .$$

Demnach wird also der Basiswechsel $\mathfrak{B} \mapsto \mathfrak{D}$ durch die Matrix mit dem allgemeinen

Element

$$u_{il} = \sum_{k=1}^{n} t_{ik} s_{kl}$$

beschrieben, d. h., die Matrix für den iterierten Basiswechsel ist das Produkt TS, wobei der erste Faktor die Matrix des ersten Basiswechsels ist.

Als Spezialfall hiervon erhält man das folgende Ergebnis.

Satz 4 (Matrix für den inversen Basiswechsel). *Es seien $\mathfrak{B}$ und $\mathfrak{C}$ zwei (angeordnete) Basen des $\mathbb{R}^n$. Den Basiswechsel $\mathfrak{B} \mapsto \mathfrak{C}$ beschreibe die Matrix T. Dann beschreibt die Matrix T^{-1} den Basiswechsel $\mathfrak{C} \mapsto \mathfrak{B}$.*

Beweis. Es sei S die Matrix für den inversen Basiswechsel $\mathfrak{C} \mapsto \mathfrak{B}$ zu dem Basiswechsel $\mathfrak{B} \mapsto \mathfrak{C}$. Dann ist für den Basiswechsel $\mathfrak{B} \mapsto \mathfrak{B}$ nach dem vorhergehenden Satz die Produktmatrix TS beschreibende Matrix. Es ist aber außerdem die Einheitsmatrix beschreibende Matrix, also erhält man $TS = E$, d. h. $S = T^{-1}$.

Eine lineare Abbildung des $\mathbb{R}^n$ in sich kann bezüglich jeder beliebigen Basis des $\mathbb{R}^n$ durch eine Matrix beschrieben werden. Wie hängen zwei beschreibende Matrizen ein und derselben linearen Abbildung hinsichtlich verschiedener Basen miteinander zusammen?

Satz 5 (Basiswechsel in Auswirkung auf eine Matrixdarstellung linearer Abbildungen). *Es sei $A: \mathbb{R}^n \to \mathbb{R}^n$ eine lineare Abbildung des $\mathbb{R}^n$ in sich. Ferner seien zwei (angeordnete) Basen $\mathfrak{B}$ und $\mathfrak{C}$ des $\mathbb{R}^n$ vorgegeben. Die beschreibende Matrix von A bezüglich des Basispaares $\mathfrak{B}, \mathfrak{B}$ sei B, die beschreibende Matrix von A bezüglich des Basispaares $\mathfrak{C}, \mathfrak{C}$ sei C. Wenn der Basiswechsel $\mathfrak{B} \mapsto \mathfrak{C}$ durch die Matrix T beschrieben wird, besteht der folgende Zusammenhang zwischen B und C:*

$$C = T^{-1}BT.$$

Beweis. Die folgenden Daten sind für die beschriebene Situation bekannt:

$$\mathfrak{B}: \boldsymbol{b}_1, \ldots, \boldsymbol{b}_n, \quad \mathfrak{C}: \boldsymbol{c}_1, \ldots, \boldsymbol{c}_n, \quad \boldsymbol{c}_j = \sum_{i=1}^{n} t_{ij}\boldsymbol{b}_i, \quad A(\boldsymbol{b}_i) = \sum_{k=1}^{n} b_{ki}\boldsymbol{b}_k, \quad A(\boldsymbol{c}_j) = \sum_{l=1}^{n} c_{lj}\boldsymbol{c}_l.$$

Dann erhält man

$$A(\boldsymbol{c}_j) = \sum_{i=1}^{n} t_{ij} A(\boldsymbol{b}_i) = \sum_{i=1}^{n} t_{ij} \sum_{k=1}^{n} b_{ki}\boldsymbol{b}_k = \sum_{k=1}^{n} \sum_{i=1}^{n} b_{ki} t_{ij} \boldsymbol{b}_k$$

und zum anderen

$$A(\boldsymbol{c}_j) = \sum_{l=1}^{n} c_{lj} \sum_{i=1}^{n} t_{il}\boldsymbol{b}_i = \sum_{i=1}^{n} \sum_{l=1}^{n} t_{il} c_{lj} \boldsymbol{b}_i.$$

Aus dem Vergleich beider Darstellungen folgt also wegen der Eindeutigkeit der Koordinatendarstellung

$$\sum_{i=1}^{n} b_{ki} t_{ij} = \sum_{l=1}^{n} t_{kl} c_{lj} \quad \text{für alle} \quad k, j = 1, \ldots, n.$$

Die letzte Beziehung ist gleichbedeutend mit der Matrixgleichung $\boldsymbol{B} \cdot \boldsymbol{T} = \boldsymbol{T} \cdot \boldsymbol{C}$, woraus sich wegen der Invertierbarkeit von $\boldsymbol{T}$

$$\boldsymbol{C} = \boldsymbol{T}^{-1}\boldsymbol{B}\boldsymbol{T}$$

ergibt.

6.7. Übungsaufgaben

1. a) In der Einheitsmatrix $\boldsymbol{I}$ der Ordnung m vertausche man die i-te Zeile mit der j-ten Zeile. Die entstehende Matrix $\boldsymbol{T}$ der Ordnung m ist invertierbar. Welcher Spaltenvektor entsteht aus dem Spaltenvektor $\boldsymbol{b}$ mit den m Komponenten $b_1, b_2, \ldots, b_m$ durch Multiplikation mit $\boldsymbol{T}$? (Man vergleiche also $\boldsymbol{b}$ mit $\boldsymbol{Tb}$.)
   b) Die beim Gaußschen Algorithmus vorkommenden elementaren Zeilenumformungen eines gegebenen linearen Gleichungssystems

$$\begin{aligned} &a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ &a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{aligned}$$

   (in Matrixform $\boldsymbol{Ax} = \boldsymbol{b}$), die bekanntlich in der Vertauschung zweier beliebiger Zeilen, der Multiplikation mit einem Skalar oder auch in der Addition einer Zeile zu einer anderen bestehen, beschreibe man jeweils durch Multiplikation mit einer Matrix $\boldsymbol{T}$ als Übergang zu dem Gleichungssystem $\boldsymbol{TAx} = \boldsymbol{Tb}$!

2. Gibt es quadratische Matrizen $\boldsymbol{A}$ der Ordnung n ($n \geqq 2$) mit den folgenden Eigenschaften:
   $\boldsymbol{A}^2 (= \boldsymbol{A} \cdot \boldsymbol{A}) = \boldsymbol{I}$ und $\boldsymbol{A} \neq \boldsymbol{I}$?
   Man interpretiere diese Aufgabenstellung für lineare Abbildungen!

3. Für eine beliebige quadratische Matrix $\begin{pmatrix} x & y \\ u & v \end{pmatrix}$ gelte stets für die fixierte Matrix $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$

$$\begin{pmatrix} x & y \\ u & v \end{pmatrix} \cdot \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \cdot \begin{pmatrix} x & y \\ u & v \end{pmatrix}.$$

   Man leite daraus her:

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

4. Es werde durch folgende Festsetzung „Positivität" von quadratischen Matrizen der Ordnung $n \geqq 2$ erklärt: Gegeben sei $\boldsymbol{A} = (a_{ij})_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}}$.
   $\boldsymbol{A}$ heißt „positiv" ($\boldsymbol{A} \geqq \boldsymbol{0}$): $\Leftrightarrow a_{ij} \geqq 0$ für alle $i, j = 1, \ldots, n$.
   Damit sei dann in $\mathfrak{M}(n \times n)$ ein Vergleich von Matrizen erklärt:

$$\boldsymbol{A} \geqq \boldsymbol{B} : \Leftrightarrow \boldsymbol{A} - \boldsymbol{B} \geqq \boldsymbol{0}.$$

   Welche der folgenden für die reellen Zahlen und ihre natürliche Ordnung „$\geqq$" bekannten Grundregeln gelten dann in $\mathfrak{M}(n \times n)$ hinsichtlich „$\geqq$"?
   a) (Reflexivität) $\boldsymbol{A} \geqq \boldsymbol{A}$ für alle $\boldsymbol{A} \in \mathfrak{M}(n \times n)$?
   b) (Antisymmetrie) $\boldsymbol{A} \geqq \boldsymbol{B}$ und $\boldsymbol{B} \geqq \boldsymbol{A} \Leftrightarrow \boldsymbol{A} = \boldsymbol{B}$ für alle $\boldsymbol{A}, \boldsymbol{B} \in \mathfrak{M}(n \times n)$?
   c) (Transitivität) $\boldsymbol{A} \geqq \boldsymbol{B}$ und $\boldsymbol{B} \geqq \boldsymbol{C} \Rightarrow \boldsymbol{A} \geqq \boldsymbol{C}$ für alle $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C} \in \mathfrak{M}(n \times n)$?
   d) (Monotonie der Addition) $\boldsymbol{A} \geqq \boldsymbol{B} \Rightarrow \boldsymbol{A} + \boldsymbol{C} \geqq \boldsymbol{B} + \boldsymbol{C}$ für alle $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C} \in \mathfrak{M}(n \times n)$?

e) (Monotonie der Multiplikation) $A \geqq O, B \geqq O \Rightarrow AB \geqq O$ für alle $A, B \in \mathcal{M}(n \times n)$?
f) (Positivität der Quadrate) $A^2 \geqq O$ für alle $A \in \mathcal{M}(n \times n)$?

5. Im folgenden handle es sich um quadratische Matrizen, für diese zeige man:
a) $A = T^{-1}BT \Rightarrow A^n = T^{-1}B^nT$.
b) $AB = BA \Rightarrow (AB)^n = A^nB^n$.
c) $AB = BA \Rightarrow (A+B)^n = \sum_{k=0}^{n} \binom{n}{k} A^{n-k}B^k$, wobei $A^0 = E = B^0$ bedeuten soll.

6. Von den beiden folgenden Matrizen entscheide man die Invertierbarkeit und berechne gegebenenfalls die Inverse:

$$\begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{pmatrix}, \quad \begin{pmatrix} 3 & -2 & 1 \\ 2 & \frac{1}{2} & -1 \\ \frac{1}{3} & 3 & 7 \end{pmatrix}.$$

7. Von der Matrix

$$A = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$$

berechne man A^2 und A^3 und ermittle sodann einen Ausdruck von A^n.

8. Wie groß ist der Rang der in Aufgabe 7 angeschriebenen Matrix A (in Abhängigkeit von λ)? Man gebe den Kern und den Bildraum der durch A bezüglich des natürlichen Koordinatensystems im $\mathbb{R}^3$ beschriebenen linearen Abbildung an.

9. Welche Bedingungen sind an die natürlichen Zahlen n und m zu stellen, damit folgendes gilt?
a) Zu gegebenem linearen Teilraum L des $\mathbb{R}^n$ läßt sich eine lineare Abbildung $A: \mathbb{R}^n \to \mathbb{R}^m$ finden mit $\ker A = L$.
b) Zu gegebenem linearen Teilraum L des $\mathbb{R}^m$ läßt sich eine lineare Abbildung $A: \mathbb{R}^n \to \mathbb{R}^m$ finden mit $\operatorname{im} A = L$.

10. Für zwei vorgeschriebene lineare Abbildungen $A: \mathbb{R}^n \to \mathbb{R}^m$ und $B: \mathbb{R}^n \to \mathbb{R}^m$ gelte
$\ker A = \ker B$.
Folgt daraus im allgemeinen die lineare Abhängigkeit der beiden Elemente A, B des $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$?

11. Das System $\mathcal{M}(n \times m)$ aller Matrizen des Typs $n \times m$ ist — wie wir wissen — hinsichtlich der Matrizenaddition und der Multiplikation mit Skalaren isomorph zum $\mathbb{R}^{nm}$ mit der koordinatenweisen Addition und Multiplikation mit Skalaren.
Man ermittle zwei verschiedene Basen in $\mathcal{M}(n \times m)$ und entscheide die Frage, ob es eine Basis in $\mathcal{M}(n \times m)$ gibt, so daß die Rangzahlen der Basiselemente alle größer als Eins sind.

12. Man beweise, daß ein allgemeines lineares Gleichungssystem von n Gleichungen mit n Unbekannten genau dann eindeutig lösbar ist, wenn die Koeffizientenmatrix invertierbar ist.

13. Man löse folgende Matrizengleichungen:

a) $\begin{pmatrix} 1 & \frac{1}{2} \\ -1 & 0 \end{pmatrix} \cdot \begin{pmatrix} x & y \\ u & v \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 6 & 5 \end{pmatrix}$,

b) $\begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix} \cdot \begin{pmatrix} x & y \\ u & v \end{pmatrix} = \begin{pmatrix} 1 & \frac{1}{2} \\ 1 & 0 \end{pmatrix}$.

14. Es sei A eine lineare Abbildung des $\mathbb{R}^n$ in den $\mathbb{R}^m$ und 0 die Nullabbildung des $\mathbb{R}^n$ in den $\mathbb{R}^k$ (d. h. $\operatorname{im} 0 = \{0\}$).

Die Lösungsgesamtheit $\{X: X \in \mathscr{L}(\mathbb{R}^m, \mathbb{R}^k), X \circ A = 0\}$ ist als linearer Teilraum von $\mathscr{L}(\mathbb{R}^m, \mathbb{R}^k)$ nachzuweisen!
Man bestimme außerdem für den Fall $n = m = k = 2$ die Dimension dieses linearen Teilraumes in Abhängigkeit vom Rang der Abbildung A.

15. Im $\mathbb{R}^2$ betrachte man die folgenden Basen

$$\mathfrak{B}: (1, 0), (0, 1), \qquad \mathfrak{C}: (1, 1), (0, 1), \qquad \mathfrak{D}: (-1, 0), (-1, -1)$$

und ermittle die Matrizen $\boldsymbol{T}, \boldsymbol{S}$ für den Basiswechsel $\mathfrak{B} \overset{\boldsymbol{T}}{\mapsto} \mathfrak{C}$ und $\mathfrak{C} \overset{\boldsymbol{S}}{\mapsto} \mathfrak{D}$.
Man überzeuge sich davon, daß die Matrix für den Basiswechsel $\mathfrak{B} \mapsto \mathfrak{D}$ von der Produktmatrix $\boldsymbol{ST}$ verschieden ist!
Warum ist folgende Argumentation nicht stichhaltig?
Dem Basiswechsel $\mathfrak{B} \mapsto \mathfrak{C}$ entspricht eine lineare Abbildung $A_1: \mathbb{R}^2 \to \mathbb{R}^2$, die $\mathfrak{B}$ in $\mathfrak{C}$ überführt. Dem Basiswechsel $\mathfrak{C} \mapsto \mathfrak{D}$ entspricht eine lineare Abbildung $A_2: \mathbb{R}^2 \to \mathbb{R}^2$, die $\mathfrak{C}$ in $\mathfrak{D}$ überführt. Die beschreibende Matrix von A_1 ist $\boldsymbol{T}$, die beschreibende Matrix von A_2 ist $\boldsymbol{S}$. Sodann muß nach dem Satz über das Produkt von linearen Abbildungen und seine Beschreibung durch Matrizen für die dem Basiswechsel $\mathfrak{B} \mapsto \mathfrak{D}$ entsprechende lineare Abbildung $A_2 \circ A_1$ auch $\boldsymbol{ST}$ die beschreibende Matrix sein.

7. Das Skalarprodukt auf dem n-dimensionalen reellen Zahlenraum

7.1. Erklärung des Skalarproduktes auf dem $\mathbb{R}^n$ und seine abstrakte Beschreibung als Bilinearform

Es war schon darauf aufmerksam gemacht worden, daß man ausgehend von dem Matrizenprodukt eine spezielle Abbildung von $\mathbb{R}^n \times \mathbb{R}^n$ in $\mathbb{R}$ erhält, wenn man jedem Paar von Elementen $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ das Matrizenprodukt $\boldsymbol{x}\boldsymbol{y}$ zuordnet, sofern man $\boldsymbol{x}$ als Zeilenvektor und $\boldsymbol{y}$ als Spaltenvektor auffaßt. Diese Abbildung sollte das Skalarprodukt oder das innere Produkt der Elemente $\boldsymbol{x}, \boldsymbol{y}$ heißen. Sehr nützliche geometrische Zusammenhänge sind mit diesem Produkt verbunden. Wir heben noch einmal explizit seine Erklärung hervor.

Definition 1 (Skalarprodukt auf dem $\mathbb{R}^n$). Unter dem *Skalarprodukt* auf dem $\mathbb{R}^n$ versteht man die wie folgt erklärte Abbildung:

$$\langle ., . \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$$

mit dem Verlauf

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i=1}^{n} x_i\, y_i$$

für alle $\boldsymbol{x} = (x_1, \ldots, x_n)$ und $\boldsymbol{y} = (y_1, \ldots, y_n)$ des $\mathbb{R}^n$.

Aus der eingangs erwähnten Tatsache, daß es sich um ein spezielles Matrizenprodukt handelt, entnimmt man sogleich folgende Eigenschaften:

1. Das Skalarprodukt $\langle ., . \rangle$ ist *in jedem seiner Argumente linear*, d. h., es gilt

$$\langle \alpha\boldsymbol{x} + \beta\boldsymbol{y}, \boldsymbol{z} \rangle = \alpha\langle \boldsymbol{x}, \boldsymbol{z} \rangle + \beta\langle \boldsymbol{y}, \boldsymbol{z} \rangle \text{ für alle } \alpha, \beta \in \mathbb{R} \text{ und } \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^n,$$
$$\langle \boldsymbol{z}, \alpha\boldsymbol{x} + \beta\boldsymbol{y} \rangle = \alpha\langle \boldsymbol{z}, \boldsymbol{x} \rangle + \beta\langle \boldsymbol{z}, \boldsymbol{y} \rangle \text{ für alle } \alpha, \beta \in \mathbb{R} \text{ und } \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^n.$$

Zum anderen erkennt man noch

2. Das Skalarprodukt $\langle ., . \rangle$ ist *symmetrisch* (oder auch kommutativ), d. h., es gilt

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{y}, \boldsymbol{x} \rangle \quad \text{für alle } \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n.$$

3. Das Skalarprodukt $\langle ., . \rangle$ ist *positiv definit*, d. h., es gilt $\langle \boldsymbol{x}, \boldsymbol{x} \rangle \geqq 0$ für alle $\boldsymbol{x} \in \mathbb{R}^n$ und $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0$ genau dann, wenn $\boldsymbol{x} = \boldsymbol{0}$.

4. Das Skalarprodukt $\langle .,. \rangle$ ist hinsichtlich der *natürlichen Basis* $e_1, e_2, \ldots, e_n$ *von* $\mathbb{R}^n$ *normiert*, d. h., es gilt

$$\langle e_i, e_j \rangle = \delta_{ij} \quad \text{für alle} \quad i, j = 1, \ldots, n.$$

Diese Eigenschaften von $\langle .,. \rangle$ sind auch charakteristisch für das Skalarprodukt. Was dies im einzelnen besagen soll, bringen die folgenden Ausführungen zum Ausdruck.

Definition 2 (Bilinearform auf dem $\mathbb{R}^n$). Es sei $B\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$. Dazu sagt man auch, daß B eine reelle Abbildung zweier Argumente auf dem $\mathbb{R}^n$ ist. Diese Abbildung heißt eine *Bilinearform* auf dem $\mathbb{R}^n$ genau dann, wenn folgendes gilt:

1. Die Abbildung $B(\boldsymbol{x}, .)\colon \mathbb{R}^n \to \mathbb{R}$ mit dem Verlauf $\boldsymbol{y} \mapsto B(\boldsymbol{x}, \boldsymbol{y})$ ist für beliebig fixiertes $\boldsymbol{x} \in \mathbb{R}^n$ linear, d. h., man hat stets

$$B(\boldsymbol{x}, \alpha\boldsymbol{y} + \beta\boldsymbol{z}) = \alpha B(\boldsymbol{x}, \boldsymbol{y}) + \beta B(\boldsymbol{x}, \boldsymbol{z})$$

für alle $\alpha, \beta \in \mathbb{R}$ und $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^n$.

2. Die Abbildung $B(., \boldsymbol{y})\colon \mathbb{R}^n \to \mathbb{R}$ mit dem Verlauf $\boldsymbol{x} \mapsto B(\boldsymbol{x}, \boldsymbol{y})$ ist für beliebig fixiertes $\boldsymbol{y} \in \mathbb{R}^n$ linear, d. h., man hat stets

$$B(\alpha\boldsymbol{x} + \beta\boldsymbol{z}, \boldsymbol{y}) = \alpha B(\boldsymbol{x}, \boldsymbol{y}) + \beta B(\boldsymbol{z}, \boldsymbol{y})$$

für alle $\alpha, \beta \in \mathbb{R}$ und $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^n$.

Eine Bilinearform B auf dem $\mathbb{R}^n$ heißt *symmetrisch* genau dann, wenn stets gilt:

$$B(\boldsymbol{x}, \boldsymbol{y}) = B(\boldsymbol{y}, \boldsymbol{x}) \text{ für alle } \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n.$$

Eine Bilinearform B auf dem $\mathbb{R}^n$, heißt *positiv-definit* genau dann, wenn stets

$$B(\boldsymbol{x}, \boldsymbol{x}) \geqq 0 \text{ und dabei nur für } \boldsymbol{x} = 0 \text{ gilt: } B(\boldsymbol{x}, \boldsymbol{x}) = 0.$$

Satz 1 (Skalarprodukt als Bilinearform). *Das Skalarprodukt auf dem* $\mathbb{R}^n$ *ist die einzige symmetrische Bilinearform* B *auf dem* $\mathbb{R}^n$ *mit der zusätzlichen Eigenschaft* $B(e_i, e_j) = \delta_{ij}$ *für die natürliche Basis* $e_1, e_2, \ldots, e_n$ *des* $\mathbb{R}^n$. (Diese Bilinearform ist überdies positiv definit.)

Beweis. Das Skalarprodukt $\langle .,. \rangle\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ ist, wie vermerkt, eine Bilinearform mit den im Satz angegebenen Eigenschaften. Es sei nun $B\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ eine Bilinearform mit den genannten Eigenschaften. Es muß für alle $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ gezeigt werden:

$$B(\boldsymbol{x}, \boldsymbol{y}) = \langle \boldsymbol{x}, \boldsymbol{y} \rangle.$$

Dazu sei $\boldsymbol{x} = (x_1, \ldots, x_n)$ und $\boldsymbol{y} = (y_1, \ldots, y_n)$. Dann gilt

$$\boldsymbol{x} = \sum_{i=1}^{n} x_i e_i, \quad \boldsymbol{y} = \sum_{j=1}^{n} y_j e_j.$$

Durch Einsetzen erhält man

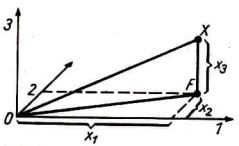
$$B(x, y) = B\left(\sum_{i=1}^{n} x_i e_i, \sum_{j=1}^{n} y_j e_j\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j B(e_i, e_j)$$

(Induktion!). Nun ist $B(e_i, e_j) = \delta_{ij}$, also verbleibt höchstens für $i = j$ ein von Null verschiedener Summand

$$B(x, y) = \sum_{i=1}^{n} x_i y_i = \langle x, y \rangle.$$

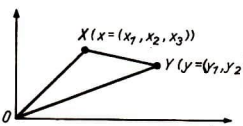
7.2. Geometrische Bedeutung des Skalarproduktes im Falle des R^2 und R^3. Norm im R^n

Die jetzigen Darlegungen sind im Sinne einer Motivierung des Interesses an dem Skalarprodukt im R^n im allgemeinen und an den noch zu entwickelnden „geometrischen" Betrachtungen wie Abstand und Orthogonalität im besonderen zu verstehen.



$$|OX|^2 = |OF|^2 + |FX|^2 = x_1^2 + x_2^2 + x_3^2$$

Abb. 8



$$|OX|^2 = x_1^2 + x_2^2 + x_3^2$$
$$|OY|^2 = y_1^2 + y_2^2 + y_3^2$$
$$|XY|^2 = (x_1 - y_1)^2 + (x_2 - y_1)^2 + (x_3 - y_3)^2$$

Abb. 9

Wir stützen uns hierbei auf einige Schulkenntnisse und verweisen auf die Ausführungen in MfL, Bde. 6 und 7.

In der euklidischen Veranschaulichung des R^2 bzw. R^3 versteht sich nach dem Pythagoras die Festsetzung, als Abstand des Punktes $x = (x_1, x_2, x_3)$ vom Ursprungspunkt die Zahl $\sqrt{x_1^2 + x_2^2 + x_3^2}$ ansehen zu wollen (vgl. Abb. 8).

Weiterhin hat man nach dem Cosinussatz für den Winkel γ zwischen den Ursprungsgeraden durch die Punkte $\boldsymbol{x} = (x_1, x_2, x_3)$ und $\boldsymbol{y} = (y_1, y_2, y_3)$

$$|OX|^2 + |OY|^2 - 2|OX| \cdot |OY| \cos\gamma = |XY|^2$$

(vgl. Abb. 9). Eine einfache Umformung führt schließlich auf

$$|OX| \cdot |OY| \cos\gamma = \sum_{i=1}^{3} x_i y_i .$$

Wir sehen somit, daß sowohl bei der Abstandsbestimmung als auch bei der Winkelbestimmung in $\mathbb{R}^2$ bzw. $\mathbb{R}^3$ eine Größe der gleichen Bauart eine entscheidende Rolle spielt, nämlich das Skalarprodukt $\langle .,. \rangle$. Für den Ursprungsabstand eines beliebigen Punktes $\boldsymbol{x}$ erhielten wir den Wert $\sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}$.

Für den Winkel zwischen den Ursprungsgeraden durch die Punkte $\boldsymbol{x}$ und $\boldsymbol{y}$ erhielten wir den Ausdruck

$$\sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle} \cdot \sqrt{\langle \boldsymbol{y}, \boldsymbol{y} \rangle} \cos\gamma = \langle \boldsymbol{x}, \boldsymbol{y} \rangle .$$

Daraus folgt insbesondere, daß für zwei von Null verschiedene Elemente $\boldsymbol{x}, \boldsymbol{y}$ des $\mathbb{R}^2$ bzw. $\mathbb{R}^3$ die durch $\boldsymbol{x}$ und $\boldsymbol{y}$ gehenden Ursprungsgeraden genau dann senkrecht aufeinanderstehen, wenn das Skalarprodukt $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ den Wert Null hat. Diese Verhältnisse geben Veranlassung, allgemein im $\mathbb{R}^n$ analoge Begriffsbildungen wie Abstand und Orthogonalität mittels des Skalarproduktes zu erklären.

Definition 1 (Norm der Elemente des $\mathbb{R}^n$). Im $\mathbb{R}^n$ soll für beliebiges $\boldsymbol{x} \in \mathbb{R}^n$ unter der *Norm* des Elementes $\boldsymbol{x}$ der folgende Zahlenwert verstanden werden:

$$\|\boldsymbol{x}\| := \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle} .$$

Wir können die Norm von $\boldsymbol{x}$ als eine Maßzahl des „Abstandes" des Elementes $\boldsymbol{x}$ vom Element $\boldsymbol{0}$ ansehen.

Satz 1 (Grundeigenschaften der aus dem Skalarprodukt abgeleiteten Norm im $\mathbb{R}^n$). *Die im $\mathbb{R}^n$ mit dem Skalarprodukt erklärte Norm der Elemente von $\mathbb{R}^n$ ist eine Abbildung $\|.\| : \mathbb{R}^n \to \mathbb{R}$ mit den folgenden Eigenschaften:*

1. $\|\boldsymbol{x}\| \geqq 0$ *für alle* $\boldsymbol{x} \in \mathbb{R}^n$ *und* $\|\boldsymbol{x}\| = 0$ *nur für* $\boldsymbol{x} = \boldsymbol{0}$.
2. $\|\alpha \cdot \boldsymbol{x}\| = |\alpha| \cdot \|\boldsymbol{x}\|$ *für alle* $\alpha \in \mathbb{R}$ *und* $\boldsymbol{x} \in \mathbb{R}^n$.
3. $\|\boldsymbol{x} + \boldsymbol{y}\| \leqq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$ *für alle* $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ (*sogenannte Dreiecksungleichung*).

Außerdem gilt noch die folgende Cauchy-Schwarz-Bunjakowski-Ungleichung:

4. $|\langle \boldsymbol{x}, \boldsymbol{y} \rangle| \leqq \|\boldsymbol{x}\| \cdot \|\boldsymbol{y}\|$ *für alle* $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$.

Bemerkung. Der Leser mache sich mittels der euklidischen Veranschaulichung des $\mathbb{R}^3$ klar, warum die Ungleichung $\|\boldsymbol{x} + \boldsymbol{y}\| \leqq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$ den Namen Dreiecksungleichung verdient, wenn man damit auf den Satz anspielt, daß in einem Dreieck

die Summe der Längen zweier Seiten stets größer oder gleich der Länge der restlichen Seite ist.

Beweis. Die Aussagen 1. und 2. sind wegen $\|x\| = \sqrt{\langle x, x\rangle}$ klar. Mittels der Cauchy-Schwarz-Bunjakowski-Ungleichung folgt auch leicht die Dreiecksungleichung. Denn es ist

$$\begin{aligned}\|x + y\|^2 &= \langle x + y, x + y\rangle \\ &= \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle \\ &= \|x\|^2 + 2\langle x, y\rangle + \|y\|^2.\end{aligned}$$

Wegen $|\langle x, y\rangle| \leqq \|x\|\,\|y\|$ ist dann

$$\|x + y\|^2 \leqq \|x\|^2 + 2\|x\|\,\|y\| + \|y\|^2,$$

d. h.

$$\|x + y\|^2 \leqq (\|x\| + \|y\|)^2,$$

oder gleichbedeutend

$$\|x + y\| \leqq \|x\| + \|y\|.$$

Verbleibt noch der Nachweis der Cauchy-Schwarz-Bunjakowski-Ungleichung. Im $\mathbb{R}^3$ ist diese klar, denn wir hatten gefunden

$$\|x\|\,\|y\| \cos\gamma = \langle x, y\rangle.$$

Nun beachtet man $-1 \leqq \cos\gamma \leqq 1$ und erhält

$$|\langle x, y\rangle| \leqq \|x\|\,\|y\|.$$

Jetzt soll diese Ungleichung aber unabhängig von jeglichen geometrischen Erörterungen allgemein im $\mathbb{R}^n$ abgeleitet werden. Dabei wird lediglich von den Grundeigenschaften des Skalarproduktes Gebrauch gemacht. Ein Zusammenhang zwischen $\langle x, y\rangle$ und $\|x\|$, $\|y\|$ besteht mittels $\langle x - y, x - y\rangle \geqq 0$ als

$$\|x\|^2 - 2\langle x, y\rangle + \|y\|^2 \geqq 0.$$

Damit gilt auch für beliebiges $\alpha \neq 0$

$$\alpha^2\|x\|^2 - 2\langle x, y\rangle + \frac{1}{\alpha^2}\|y\|^2 \geqq 0,$$

indem man von

$$\left\langle \alpha x - \frac{1}{\alpha} y, \alpha x - \frac{1}{\alpha} y \right\rangle \geqq 0$$

ausgeht.

Nun gibt es bei $\|x\| \neq 0$ und $\|y\| \neq 0$ ein $\alpha_0 \neq 0$ mit $\alpha_0\|x\| = \frac{1}{\alpha_0}\|y\|$, also ist $2\alpha_0^2\|x\|^2 \geqq 2\langle x, y\rangle$ und deshalb

$$\|x\|\,\|y\| \geqq \langle x, y\rangle.$$

Ebenso folgt aus

$$\left\langle \alpha \boldsymbol{x} + \frac{1}{\alpha}\boldsymbol{y}, \alpha \boldsymbol{x} + \frac{1}{\alpha}\boldsymbol{y} \right\rangle \geqq 0$$

die Beziehung

$$\|\boldsymbol{x}\|\|\boldsymbol{y}\| \geqq -\langle \boldsymbol{x}, \boldsymbol{y} \rangle$$

bei $\|\boldsymbol{x}\| \neq 0$, $\|\boldsymbol{y}\| \neq 0$.

Insgesamt ist dann $\|\boldsymbol{x}\|\|\boldsymbol{y}\| \geqq |\langle \boldsymbol{x}, \boldsymbol{y} \rangle|$, weil auch für $\|\boldsymbol{x}\| = 0$ oder $\|\boldsymbol{y}\| = 0$ die Ungleichung wegen $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$ richtig wird.

Bemerkung. In allen Argumentationen des vorstehenden Beweises ist von dem Skalarprodukt $\langle ., . \rangle$ lediglich benutzt worden, daß es sich um eine positiv-definite symmetrische Bilinearform handelt. Wir können daher in gleicher Weise für solche Bilinearformen eine Norm erklären und dieselben Grundeigenschaften der Norm ableiten.

Definition 2 (Normierung von $\mathbf{R}^n$ mittels einer positiv-definiten symmetrischen Bilinearform). Auf dem $\mathbf{R}^n$ sei eine positiv-definite symmetrische Bilinearform $B\colon \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ gegeben. Dann heißt der folgende Zahlenwert für ein beliebiges Element $\boldsymbol{x} \in \mathbf{R}^n$ die *Norm* von $\boldsymbol{x}$ bezüglich B:

$$\|\boldsymbol{x}\|_B := \sqrt{B(\boldsymbol{x}, \boldsymbol{x})}.$$

Satz 2 (Grundeigenschaften der aus einer positiv-definiten symmetrischen Bilinearform abgeleiteten Norm). *Die im $\mathbf{R}^n$ mit einer gegebenen positiv-definiten symmetrischen Bilinearform B erklärte Norm der Elemente von $\mathbf{R}^n$ ist eine Abbildung $\|.\|_B\colon \mathbf{R}^n \to \mathbf{R}$ mit den folgenden Eigenschaften:*

1. $\|\boldsymbol{x}\|_B \geqq 0$ *für alle* $\boldsymbol{x} \in \mathbf{R}^n$ *und* $\|\boldsymbol{x}\|_B = 0$ *nur für* $\boldsymbol{x} = \boldsymbol{0}$.
2. $\|\alpha \boldsymbol{x}\|_B = |\alpha| \cdot \|\boldsymbol{x}\|_B$ *für alle* $\alpha \in \mathbf{R}$ *und* $\boldsymbol{x} \in \mathbf{R}^n$.
3. $\|\boldsymbol{x} + \boldsymbol{y}\|_B \leqq \|\boldsymbol{x}\|_B + \|\boldsymbol{y}\|_B$ *für alle* $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{R}^n$.

Außerdem gilt noch die folgende Cauchy-Schwarz-Bunjakowski-Ungleichung:

4. $|B(\boldsymbol{x}, \boldsymbol{y})| \leqq \|\boldsymbol{x}\|_B \|\boldsymbol{y}\|_B$ *für alle* $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{R}^n$.

7.3. Orthogonalität im $\mathbf{R}^n$

Die folgende Begriffsbildung ist nach den Ausführungen des letzten Paragraphen sinnvoll, ihre Zweckmäßigkeit wird sich sodann unmittelbar zeigen.

Definition 1 (Orthogonalität im $\mathbf{R}^n$ hinsichtlich positiv-definiter symmetrischer Bilinearform). Im $\mathbf{R}^n$ sei eine positiv-definite symmetrische Bilinearform $B\colon \mathbf{R}^n \times \mathbf{R}^n$

$\to \mathbf{R}$ gegeben. Zwei Elemente $x, y \in \mathbf{R}^n$ heißen bezüglich B *orthogonal* genau dann, wenn $B(x, y) = 0$ gilt. Orthogonalität von $x, y \in \mathbf{R}^n$ bezüglich B wird durch $x \perp_B y$ bezeichnet. Orthogonalität ohne den Zusatz einer Bilinearform B meint dann einfach die Orthogonalität in bezug auf das Skalarprodukt im $\mathbf{R}^n$.

Definition 2 (Orthonormalsysteme im $\mathbf{R}^n$). Eine nichtleere Teilmenge $S \subseteq \mathbf{R}^n$ heißt ein *Orthonormalsystem* hinsichtlich einer gegebenen positiv-definiten symmetrischen Bilinearform $B\colon \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ genau dann, wenn gilt:

1. $B(x, x) = 1$ für alle $x \in S$.

2. $B(x, y) = 0$ für alle $x, y \in S$ mit $x \neq y$.

Bemerkung. Die erste Forderung besagt, daß jedes Element von S die Norm 1 hinsichtlich $\|.\|_B$ hat. Die zweite Forderung bedeutet, daß je zwei Elemente von S bezüglich B orthogonal sind. Damit erklärt sich auch die Bezeichnung als Zusammenziehung von „orthogonal" und „normiert".

Satz 1 (Orthonormalsysteme und lineare Unabhängigkeit). *Es sei im $\mathbf{R}^n$ eine positiv-definite symmetrische Bilinearform $B\colon \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ gegeben. $S \subseteq \mathbf{R}^n$ sei ein Orthonormalsystem im $\mathbf{R}^n$. Dann gilt: S ist eine linear unabhängige Menge.*

Beweis. Angenommen, es gibt ein Element $x \in S$, das sich aus $S \setminus \{x\}$ linear kombinieren läßt:

$$x = \sum_{i=1}^{k} \alpha_i x_i \qquad (x_i \in S \setminus \{x\}).$$

Nun sollte $B(x, x) = 1$ sein. Zum anderen ist aber

$$B\left(x, \sum_{i=1}^{k} \alpha_i x_i\right) = \sum_{i=1}^{k} \alpha_i B(x, x_i) = 0,$$

weil $B(y, z) = 0$ ist für alle $y, z \in S$ mit $y \neq z$. Demnach muß S wirklich linear unabhängig sein.

Wir ersehen aus dem Satz, daß der $\mathbf{R}^n$ bestenfalls Orthonormalsysteme aus höchstens n Elementen besitzen kann. Im Fall des Skalarproduktes ist die natürliche Basis ein solches maximales Orthonormalsystem. Hat nun aber auch stets der $\mathbf{R}^n$ bezüglich einer beliebigen positiv-definiten symmetrischen Bilinearform eine orthonormale Basis? Diese Frage wird durch das folgende Ergebnis beantwortet.

Satz 2 (Umwandlung eines linear unabhängigen Systems in ein Orthonormalsystem). *Es sei $\{a_1, \ldots, a_k\}$ eine linear unabhängige Teilmenge des $\mathbf{R}^n$, und im $\mathbf{R}^n$ sei eine positiv-definite, symmetrische Bilinearform $B\colon \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ gegeben. Dann gibt*

es ein Orthonormalsystem $\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_k\}$ *bezüglich* B *mit der folgenden Eigenschaft:*

$$L(\{\boldsymbol{a}_1\}) = L(\{\boldsymbol{b}_1\}),$$
$$L(\{\boldsymbol{a}_1, \boldsymbol{a}_2\}) = L(\{\boldsymbol{b}_1, \boldsymbol{b}_2\}),$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$L(\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_{k-1}, \boldsymbol{a}_k\}) = L(\{\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_{k-1}, \boldsymbol{b}_k\}).$$

Bemerkung. Das im anschließenden Beweis zur Anwendung kommende Konstruktionsverfahren von $\{\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_k\}$ heißt das *Orthonormalisierungsverfahren von Gram-E. Schmidt.*

Beweis. Die Konstruktion des Orthonormalsystems $\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_k\}$ geschieht schrittweise.

1. Schritt: Es wird

$$\boldsymbol{b}_1 := \frac{\boldsymbol{a}_1}{\|\boldsymbol{a}_1\|_B}$$

gebildet, wegen $\boldsymbol{a}_1 \neq \boldsymbol{0}$ ist $\|\boldsymbol{a}_1\|_B \neq 0$ und $\boldsymbol{b}_1$ wird in bezug auf B ein Einheitselement, d. h., seine Norm $\|.\|_B$ ist 1. Weil $\boldsymbol{b}_1$ ein skalares Vielfaches von $\boldsymbol{a}_1$ ist, versteht sich die Beziehung $L(\{\boldsymbol{a}_1\}) = L(\{\boldsymbol{b}_1\})$.

2. Schritt: Es wird

$$\boldsymbol{b}_2' := \boldsymbol{a}_2 + \gamma_1 \cdot \boldsymbol{b}_1$$

gesetzt, wobei der Koeffizient γ_1 so gewählt wird, daß $\boldsymbol{b}_2' \perp_B \boldsymbol{b}_1$, d. h., es muß

$$B(\boldsymbol{b}_2', \boldsymbol{b}_1) = B(\boldsymbol{a}_2, \boldsymbol{b}_1) + \gamma_1 \cdot B(\boldsymbol{b}_1, \boldsymbol{b}_1) = 0$$

gelten, woraus folgt, daß $\gamma_1 := -B(\boldsymbol{a}_2, \boldsymbol{b}_1)$ zu wählen ist.

Es ist $\boldsymbol{b}_2' \neq \boldsymbol{0}$, weil sonst $\boldsymbol{a}_2$ und $\boldsymbol{b}_1$ und damit $\boldsymbol{a}_2$ und $\boldsymbol{a}_1$ linear abhängig wären, was der Voraussetzung widerspräche. Es wird

$$\boldsymbol{b}_2 := \frac{\boldsymbol{b}_2'}{\|\boldsymbol{b}_2'\|_B}$$

gebildet. $\boldsymbol{b}_2$ ist in bezug auf B ein Einheitselement, d. h., seine Norm ist 1. $\boldsymbol{b}_1$ und $\boldsymbol{b}_2$ sind nach Konstruktion zueinander orthogonal, und es ist auch $L(\{\boldsymbol{a}_1, \boldsymbol{a}_2\}) = L(\{\boldsymbol{b}_1, \boldsymbol{b}_2\})$ erfüllt, denn wir haben $\boldsymbol{b}_1 \in L(\{\boldsymbol{a}_1, \boldsymbol{a}_2\})$, $\boldsymbol{b}_2 \in L(\{\boldsymbol{a}_1, \boldsymbol{a}_2\})$. Also ist $\boldsymbol{b}_1$, $\boldsymbol{b}_2$ ein orthonormiertes und damit linear unabhängiges System in $L(\{\boldsymbol{a}_1, \boldsymbol{a}_2\})$, es muß demzufolge eine Basis von $L(\{\boldsymbol{a}_1, \boldsymbol{a}_2\})$ sein.

Auf diese Weise fährt man induktiv in der Konstruktion fort.

Wir beschreiben dazu den allgemeinen Schritt:

ν-ter Schritt: Es sei für $\nu - 1$ schon ein orthonormiertes System mit den im Satz angegebenen Eigenschaften konstruiert.

Es wird

$$\boldsymbol{b}_\nu' := \boldsymbol{a}_\nu + \varepsilon_1 \boldsymbol{b}_1 + \varepsilon_2 \boldsymbol{b}_2 + \cdots + \varepsilon_{\nu-1} \boldsymbol{b}_{\nu-1}$$

gesetzt. Die Koeffizienten ε_i werden dabei so gewählt, daß $\boldsymbol{b}_\nu'$ zu jedem $\boldsymbol{b}_i$, $i = 1, \ldots, \nu - 1$, orthogonal wird, d. h., es muß

$$B(\boldsymbol{b}_\nu', \boldsymbol{b}_i) = B(\boldsymbol{a}_\nu, \boldsymbol{b}_i) + \varepsilon_i = 0$$

gelten, woraus folgt, daß

$$\varepsilon_i = -B(\boldsymbol{a}_\nu, \boldsymbol{b}_i)$$

zu wählen ist.

Es ist $\boldsymbol{b}_\nu' \neq 0$, weil sonst $\boldsymbol{a}_\nu \in L(\{\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_{\nu-1}\}) = L(\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{\nu-1}\})$ wäre, was der Voraussetzung widerspräche. Es wird

$$\boldsymbol{b}_\nu := \frac{\boldsymbol{b}_\nu'}{\|\boldsymbol{b}_\nu'\|_B}$$

gebildet. Damit ist $\{\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_{\nu-1}, \boldsymbol{b}_\nu\}$ ein Orthonormalsystem. Wegen $\boldsymbol{b}_i \in L(\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_\nu\})$ für $i = 1, \ldots, \nu$ ist $\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_\nu\}$ als Orthonormalsystem in $L(\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_\nu\})$ linear unabhängig, d. h., es muß eine Basis von $L(\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_\nu\})$ bilden (eine Basis besteht aus ν Elementen).

Als Folgerungen aus dem Orthonormierungsverfahren ergeben sich die beiden folgenden Sachverhalte:

Satz 3 (Maximale Orthonormalsysteme als Basen). *Es sei $S \subseteq \mathbf{R}^n$ ein Orthonormalsystem im $\mathbf{R}^n$ bezüglich einer beliebig vorgegebenen positiv-definiten symmetrischen Bilinearform B: $\mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$. Dann gilt:*

S ist eine Basis von $\mathbf{R}^n \Leftrightarrow S$ ist ein maximales Orthonormalsystem im $\mathbf{R}^n$, d. h., es gibt kein hinsichtlich B normiertes Element (Norm 1), welches zu allen Elementen von S orthogonal ist.

Beweis. „$\Rightarrow$": Angenommen, es gibt ein Element $\boldsymbol{x} \in \mathbf{R}^n$ mit $\|\boldsymbol{x}\|_B = 1$ und $\boldsymbol{x} \perp_B \boldsymbol{y}$ für alle $\boldsymbol{y} \in S$. Dann ist das System $S \cup \{\boldsymbol{x}\}$ ein Orthonormalsystem, also linear unabhängig. Das kann aber nicht eintreten, da S als Basis ein maximales linear unabhängiges System ist.

„$\Leftarrow$": S ist als Orthonormalsystem linear unabhängig. Wäre $L(S) \neq \mathbf{R}^n$, so könnte man S um weitere Elemente zu einer vollen Basis von $\mathbf{R}^n$ ergänzen. Auf diese das Orthonormalisierungsverfahren in geeigneter Weise zur Anwendung gebracht, würde man ein S umfassendes Orthonormalsystem erhalten, was nach Voraussetzung nicht geht. Also muß $L(S) = \mathbf{R}^n$ gelten, d. h., S ist Basis von $\mathbf{R}^n$.

Satz 4 (Positiv-definite symmetrische Bilinearformen als Skalarprodukt hinsichtlich Orthonormalbasen). *Es sei B: $\mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ eine positiv-definite symmetrische Bilinearform auf dem $\mathbf{R}^n$. $\mathfrak{B}$: $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$ sei eine beliebige Orthonormalbasis hinsichtlich B für den $\mathbf{R}^n$* (nach dem Orthonormierungsverfahren gibt es auch stets solche).

Dann drückt sich der Wert der Bilinearform für die Elemente $\boldsymbol{x}$ und $\boldsymbol{y}$ mit den Koordinatentupeln $(\xi_1, \ldots, \xi_n)$ und $(\eta_1, \ldots, \eta_n)$ bezüglich der Basis $\mathfrak{B}$ wie folgt aus:

$$B(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i=1}^{n} \xi_i \eta_i .$$

Beweis. Es ist nach Voraussetzung

$$\boldsymbol{x} = \sum_{i=1}^{n} \xi_i \boldsymbol{b}_i, \quad \boldsymbol{y} = \sum_{j=1}^{n} \eta_j \boldsymbol{b}_j$$

und $B(\boldsymbol{b}_i, \boldsymbol{b}_j) = \delta_{ij}$. Daraus folgt

$$B(\boldsymbol{x}, \boldsymbol{y}) = B\left(\sum_{i=1}^{n} \xi_i \boldsymbol{b}_i, \sum_{j=1}^{n} \eta_j \boldsymbol{b}_j\right) = \sum_{i,j=1}^{n} \xi_i \eta_j B(\boldsymbol{b}_i, \boldsymbol{b}_j) = \sum_{i=1}^{n} \xi_i \eta_i .$$

Den letzten Satz können wir so interpretieren, daß wir bei Betrachtungen von positiv-definiten symmetrischen Bilinearformen auf dem $\mathbf{R}^n$ eigentlich immer das Skalarprodukt zugrunde legen können, weil man ja durch geeigneten Basiswechsel stets diesen Fall herzustellen vermag. Wir erkennen hieran einmal mehr, wie nützlich die Bezugnahme einer Problematik der linearen Algebra des $\mathbf{R}^n$ auf eine zweckmäßige Basis ist.

7.4. Orthogonale lineare Abbildungen und orthogonale Matrizen

Der folgende Typ von linearen Abbildungen des $\mathbf{R}^n$ ist — besonders im Hinblick auf die Kongruenzgeometrie — von Bedeutung.

Definition 1 (Orthogonale lineare Abbildungen des $\mathbf{R}^n$ in sich). Es sei $A: \mathbf{R}^n \to \mathbf{R}^n$ eine lineare Abbildung des $\mathbf{R}^n$ in sich. Diese heißt eine *orthogonale Abbildung* genau dann, wenn durch A jedes Orthonormalsystem von $\mathbf{R}^n$ in ein Orthonormalsystem übergeführt wird.

Zunächst kümmern wir uns um eine andere Charakterisierung der orthogonalen Abbildungen.

Satz 1 (Kennzeichnung der orthogonalen linearen Abbildungen durch Wirkung auf das Skalarprodukt). *Es sei $A: \mathbf{R}^n \to \mathbf{R}^n$ eine lineare Abbildung des $\mathbf{R}^n$ in sich. Dann sind die folgenden Aussagen paarweise äquivalent:*

1. *A ist eine orthogonale Abbildung.*
2. *A läßt das Skalarprodukt invariant, d. h., es besteht für je zwei Elemente $\boldsymbol{x}, \boldsymbol{y}$ des $\mathbf{R}^n$ die Beziehung*

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle A(\boldsymbol{x}), A(\boldsymbol{y}) \rangle .$$

3. *A läßt die Norm der Elemente von $\mathbb{R}^n$ invariant, d. h., es besteht für jedes $x \in \mathbb{R}^n$ die Beziehung*

$$\|x\| = \|A(x)\|.$$

Beweis. Wir müssen die Äquivalenzen 1. $\Leftrightarrow$ 2., 2. $\Leftrightarrow$ 3. und 1. $\Leftrightarrow$ 3. zeigen. Dafür wird eine zyklische Beweisanordnung 1. $\Rightarrow$ 3. $\Rightarrow$ 2. $\Rightarrow$ 1. durchgeführt.

Zu 1. $\Rightarrow$ 3. Für $x = 0$ ist $A(x) = 0$, also $\|x\| = \|A(x)\|$. Wenn $x \neq 0$ vorausgesetzt wird, so ist $\left\{\frac{x}{\|x\|}\right\}$ ein Orthonormalsystem, also muß auch $A\left(\frac{x}{\|x\|}\right)$ Orthonormalsystem sein, d. h. $\left\|A\left(\frac{x}{\|x\|}\right)\right\| = 1$ oder gleichbedeutend

$$\frac{\|A(x)\|}{\|x\|} = 1.$$

Zu 3. $\Rightarrow$ 2. Es ist

$$\langle x, y\rangle = -\frac{1}{2}\left(\|x - y\|^2 - \|x\|^2 - \|y\|^2\right),$$

wie aus $\langle x - y, x - y\rangle = \|x - y\|^2$ hervorgeht. Nun hat man

$$\|x - y\| = \|A(x - y)\| = \|A(x) - A(y)\|$$

und

$$\|x\| = \|A(x)\|, \quad \|y\| = \|A(y)\|,$$

woraus

$$\begin{aligned}\langle x, y\rangle &= -\frac{1}{2}\left(\|x - y\|^2 - \|x\|^2 - \|y\|^2\right)\\ &= -\frac{1}{2}\left(\|A(x) - A(y)\|^2 - \|A(x)\|^2 - \|A(y)\|^2\right)\\ &= \langle A(x), A(y)\rangle\end{aligned}$$

hervorgeht.

Zu 2. $\Rightarrow$ 1. Es sei S ein Orthonormalsystem im $\mathbb{R}^n$. Dann gilt

$$\langle x, y\rangle = \begin{cases} 1 \text{ für } x, y \in S \text{ und } x = y,\\ 0 \text{ für } x, y \in S \text{ und } x \neq y.\end{cases}$$

Also ist

$$\langle A(x), A(y)\rangle = \begin{cases} 1 \text{ für } A(x), A(y) \in A(S) \text{ und } A(x) = A(y),\\ 0 \text{ für } A(x), A(y) \in A(S) \text{ und } A(x) \neq A(y),\end{cases}$$

d. h., bei $A(S)$ handelt es sich um ein Orthonormalsystem.

Bemerkung. Insbesondere enthält der vorstehende Satz die Tatsache, daß jede orthogonale lineare Abbildung des $\mathbb{R}^n$ eine eineindeutige Abbildung des $\mathbb{R}^n$ auf sich ist. Denn aus $\boldsymbol{x} \neq \boldsymbol{y}$ folgt $\|\boldsymbol{x} - \boldsymbol{y}\| \neq 0$ und $\|A(\boldsymbol{x}) - A(\boldsymbol{y})\| \neq 0$, d. h. $A(\boldsymbol{x}) \neq A(\boldsymbol{y})$. Die inverse Abbildung $A^{-1}: \mathbb{R}^n \to \mathbb{R}^n$ läßt dann auch die Norm invariant, ist also selbst wieder eine orthogonale Abbildung.

Durch welche Matrizen werden nun orthogonale lineare Abbildungen beschrieben? Um diese Frage zu entscheiden, betrachten wir im $\mathbb{R}^n$ eine lineare orthogonale Abbildung $A: \mathbb{R}^n \to \mathbb{R}^n$ und ihre beschreibende Matrix $\boldsymbol{A}$ bezüglich der natürlichen Basis $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$. Dann ist $A(\boldsymbol{e}_1), A(\boldsymbol{e}_2), \ldots, A(\boldsymbol{e}_n)$ ein Orthonormalsystem. Wir erkennen also, daß die Spalten von $\boldsymbol{A}$ die folgende Eigenschaft haben: Das Skalarprodukt einer jeden Spalte von $\boldsymbol{A}$ mit sich selbst ist 1, das Skalarprodukt zweier verschiedener Spalten von $\boldsymbol{A}$ ist Null.

Definition 2 (Orthogonale Matrizen). Es sei $\boldsymbol{A}$ eine quadratische Matrix der Ordnung n. Diese Matrix heißt genau dann eine *orthogonale Matrix*, wenn das Skalarprodukt einer jeden Spalte von $\boldsymbol{A}$ mit sich selbst 1 ist und das Skalarprodukt zweier verschiedener Spalten von $\boldsymbol{A}$ gleich 0 ist.

Wir haben dann in leichter Verallgemeinerung der vorherigen Feststellung den folgenden

Satz 2 (Kennzeichnung der orthogonalen linearen Abbildungen durch beschreibende Matrizen). *Es sei* $\mathfrak{B}: \boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$ *eine orthonormale Basis im* $\mathbb{R}^n$. $A: \mathbb{R}^n \to \mathbb{R}^n$ *sei eine lineare Abbildung des* $\mathbb{R}^n$ *in sich. Diese lineare Abbildung ist orthogonal* $\Leftrightarrow$ *Die beschreibende Matrix* $\boldsymbol{A}$ *von* A *bezüglich der Basis* $\mathfrak{B}$ *ist orthogonal.*

Beweis. Die Implikation „$\Rightarrow$" ist im wesentlichen schon vorher erörtert worden. Der Leser gehe noch einmal alle Punkte sorgfältig durch.

Zu „$\Leftarrow$": Wir zeigen $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle A(\boldsymbol{x}), A(\boldsymbol{y}) \rangle$ für alle $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$. Es ist

$$\boldsymbol{x} = \sum_{i=1}^{n} \xi_i \boldsymbol{b}_i, \qquad \boldsymbol{y} = \sum_{i=1}^{n} \eta_j \boldsymbol{b}_j,$$

und wegen Orthonormalität der Basis $\mathfrak{B}$ folgt $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i=1}^{n} \xi_i \eta_i$. Für alle $\langle A(\boldsymbol{x}), A(\boldsymbol{y}) \rangle$ ergibt sich unter Benutzung von

$$A(\boldsymbol{x}) = \sum_{i=1}^{n} \xi_i A(\boldsymbol{b}_i) = \sum_{i=1}^{n} \xi_i \sum_{k=1}^{n} a_{ki} \boldsymbol{b}_k = \sum_{k=1}^{n} \left(\sum_{i=1}^{n} a_{ki} \xi_i \right) \boldsymbol{b}_k$$

und

$$A(\boldsymbol{y}) = \sum_{j=1}^{n} \eta_j A(\boldsymbol{b}_j) = \sum_{j=1}^{n} \eta_j \sum_{k=1}^{n} a_{kj} \boldsymbol{b}_k = \sum_{k=1}^{n} \left(\sum_{j=1}^{n} a_{kj} \eta_j \right) \boldsymbol{b}_k,$$

$$\langle A(\boldsymbol{x}), A(\boldsymbol{y}) \rangle = \sum_{k=1}^{n} \left(\sum_{i=1}^{n} a_{ki} \xi_i \sum_{j=1}^{n} a_{kj} \eta_j \right) = \sum_{k=1}^{n} \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ki} a_{kj} \xi_i \eta_j$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\sum_{k=1}^{n} a_{ki} a_{kj} \right) \xi_i \eta_j.$$

Bei Orthogonalität von $\boldsymbol{A}$ ist aber für $i \neq j$

$$\sum_{k=1}^{n} a_{ki}a_{kj} = 0$$

und bei $i = j$

$$\sum_{k=1}^{n} a_{ki}a_{kj} = 1,$$

womit

$$\langle A(\boldsymbol{x}), A(\boldsymbol{y})\rangle = \sum_{i=1}^{n} \xi_i\eta_i = \langle \boldsymbol{x}, \boldsymbol{y}\rangle$$

verbleibt.

Die vorhin vermerkte Tatsache, daß jede orthogonale lineare Abbildung des $\mathbf{R}^n$ in sich invertierbar ist und als Inverse wieder eine orthogonale lineare Abbildung besitzt, kann zur folgenden Charakterisierung der Orthogonalität von Matrizen benutzt werden.

Satz 3 (Orthogonale Matrizen und ihre Inversen). *Es sei $\boldsymbol{A}$ eine quadratische Matrix der Ordnung n. Dann sind folgende Aussagen paarweise äquivalent:*

1. *$\boldsymbol{A}$ ist eine orthogonale Matrix.*

2. *$\boldsymbol{A}$ ist invertierbar und die Inverse $\boldsymbol{A}^{-1}$ ist orthogonal.*

3. *Die Zeilen von $\boldsymbol{A}$ bilden ein Orthonormalsystem im $\mathbf{R}^n$.*

4. *Es gilt $\boldsymbol{A} \cdot \boldsymbol{A}^{\mathsf{T}} = \boldsymbol{A}^{\mathsf{T}} \cdot \boldsymbol{A} = \boldsymbol{E}$, wobei $\boldsymbol{A}^{\mathsf{T}}$ die sogenannte transponierte Matrix zu $\boldsymbol{A}$ ist, diese hat die Spalten von $\boldsymbol{A}$ als Zeilen:*

$$\boldsymbol{A} = (a_{ij})_{\substack{i=1,\ldots,n\\ j=1,\ldots,n}}, \qquad \boldsymbol{A}^{\mathsf{T}} = (a_{ji})_{\substack{i=1,\ldots,n\\ j=1,\ldots,n}}.$$

Beweis. Wir zeigen einen Implikationszyklus $1. \Rightarrow 2. \Rightarrow 3. \Rightarrow 4. \Rightarrow 1.$

Zu $1. \Rightarrow 2.$ Wenn $\boldsymbol{A}$ orthogonal ist, ist die lineare Abbildung $A\colon \mathbf{R}^n \to \mathbf{R}^n$, die bezüglich der natürlichen Basis des $\mathbf{R}^n$ als beschreibende Matrix gerade $\boldsymbol{A}$ hat, orthogonal. Dann existiert aber die inverse Abbildung, und diese ist orthogonal. Die inverse Abbildung hat aber die inverse Matrix $\boldsymbol{A}^{-1}$ als beschreibende Matrix, also ist $\boldsymbol{A}^{-1}$ orthogonal.

Zu $2. \Rightarrow 3.$ Hinsichtlich der i-ten Zeile von $\boldsymbol{A}$ gilt: Das Skalarprodukt mit der j-ten Spalte von $\boldsymbol{A}^{-1}$ ist δ_{ij}. Die Spalten von $\boldsymbol{A}^{-1}$ bilden aber eine Orthonormalbasis $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n$ von $\mathbf{R}^n$, wenn man sie als Zeilenvektoren auffaßt. Demnach hat man für die i-te Zeile $\boldsymbol{z}_i$ von $\boldsymbol{A}$ die Beziehung $\boldsymbol{z}_i = \sum_{k=1}^{n} \sigma_k \boldsymbol{s}_k$. Woraus wegen $\langle \boldsymbol{z}_i, \boldsymbol{s}_j\rangle = \delta_{ij}$ folgt:

$$\sigma_k = \begin{cases} 1 \text{ für } k = i, \\ 0 \text{ für } k \neq i, \end{cases}$$

d. h., es ist $\boldsymbol{z}_i = \boldsymbol{s}_i$. Es ist also schon bewiesen worden, daß die Zeilen von $\boldsymbol{A}$ gerade mit den Spalten von $\boldsymbol{A}^{-1}$ übereinstimmen, sofern $\boldsymbol{A}^{-1}$ orthogonal ist.

Wir haben also aus 2. sowohl 3. als auch 4. gefolgert.

Zu 3. $\Rightarrow$ 4. Wenn die Zeilen von $\boldsymbol{A}$ ein Orthonormalsystem von $\mathbb{R}^n$ bilden, schließe man analog wie zuvor.

Zu 4. $\Rightarrow$ 1. Wenn

$$\boldsymbol{A}^{\mathsf{T}} \cdot \boldsymbol{A} = \boldsymbol{E} = \boldsymbol{A} \cdot \boldsymbol{A}^{\mathsf{T}}$$

gilt, dann heißt das gerade, daß die Spalten von $\boldsymbol{A}$ (die Zeilen von $\boldsymbol{A}$) ein Orthogonalsystem bilden.

Bemerkung. Für eine orthogonale Matrix $\boldsymbol{A}$ erhält man ihre Inverse also ganz einfach durch Transponieren: $\boldsymbol{A}^{-1} = \boldsymbol{A}^{\mathsf{T}}$. Nun fragen wir uns noch, welche Matrizen für den Basiswechsel einer Orthonormalbasis zu einer anderen Orthonormalbasis in Frage kommen.

Satz 4 (*Übergangsmatrix von Orthonormalbasen*). *Es sei* $\mathfrak{B}$ *eine Orthonormalbasis des* $\mathbb{R}^n$. *Die Matrix* $\boldsymbol{T}$ *beschreibe den Basiswechsel* $\mathfrak{B} \mapsto \mathfrak{C}$ *zu einer anderen Basis von* $\mathbb{R}^n$. *Dann gilt: Die Matrix* $\boldsymbol{T}$ *des Basiswechsels* $\mathfrak{B} \mapsto \mathfrak{C}$ *ist orthogonal* $\Leftrightarrow$ *Die neue Basis* $\mathfrak{C}$ *ist eine Orthonormalbasis.*

Beweis. „$\Rightarrow$“: Es gilt für die Basis $\mathfrak{C}$: $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n$

$$\boldsymbol{c}_j = \sum_{i=1}^{n} t_{ij}\, \boldsymbol{b}_i,$$

wenn $\boldsymbol{T} = (t_{ij})_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}}$ und $\mathfrak{B}$: $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$ ist. Für $\langle \boldsymbol{c}_j, \boldsymbol{c}_k \rangle$ erhält man

$$\langle \boldsymbol{c}_j, \boldsymbol{c}_k \rangle = \left\langle \sum_{i=1}^{n} t_{ij}\boldsymbol{b}_i, \sum_{l=1}^{n} t_{lk}\boldsymbol{b}_l \right\rangle = \sum_{i=1}^{n} \sum_{l=1}^{n} t_{ij} t_{lk} \langle \boldsymbol{b}_i, \boldsymbol{b}_l \rangle = \sum_{i=1}^{n} t_{ij}\, t_{ik} = \delta_{jk},$$

denn der vorletzte Summenausdruck ist das Skalarprodukt der j-ten Spalte mit der k-ten Spalte von $\boldsymbol{T}$.

„$\Leftarrow$“: Wie vorher hat man

$$\langle \boldsymbol{c}_j, \boldsymbol{c}_k \rangle = \sum_{i=1}^{n} t_{ij} t_{ik}.$$

Diesmal weiß man auf Grund der Orthonormalität von $\mathfrak{C}$, daß $\langle \boldsymbol{c}_j, \boldsymbol{c}_k \rangle = \delta_{jk}$ ist. Demnach bilden wegen

$$\sum_{i=1}^{n} t_{ij} t_{ik} = \delta_{jk}$$

die Spalten von $\boldsymbol{T}$ ein Orthonormalsystem.

Satz 5 (Struktur des Systems der orthogonalen Matrizen der Ordnung n hinsichtlich der Matrizenmultiplikation). *Das System aller orthogonalen Matrizen der Ordnung n (ein Teilsystem von $\mathcal{M}(n \times n)$) hat folgende Eigenschaften in bezug auf die Matrizenmultiplikation:*

1. *Das Produkt zweier orthogonaler Matrizen der Ordnung n ist wieder eine orthogonale Matrix der Ordnung n.*

2. *Es gibt eine (eindeutig bestimmte) orthogonale Matrix $\boldsymbol{I}$ der Ordnung n, so daß für jede andere orthogonale Matrix $\boldsymbol{A}$ der Ordnung n gilt:*

$$\boldsymbol{A} \cdot \boldsymbol{I} = \boldsymbol{I} \cdot \boldsymbol{A} = \boldsymbol{A}.$$

3. *Zu jeder orthogonalen Matrix $\boldsymbol{A}$ der Ordnung n gibt es eine (eindeutig bestimmte) orthogonale Matrix $\tilde{\boldsymbol{A}}$ der Ordnung n, so daß gilt:*

$$\boldsymbol{A} \cdot \tilde{\boldsymbol{A}} = \tilde{\boldsymbol{A}} \cdot \boldsymbol{A} = \boldsymbol{I}.$$

Das System aller orthogonalen Matrizen der Ordnung n bildet hinsichtlich der Matrizenmultiplikation — wie man sagt — eine Gruppe. Man bezeichnet sie mit $OL(n)$ und nennt sie die orthogonale lineare Gruppe der Ordnung n.

Beweis. Wir brauchen uns nur um eine Bestätigung der Eigenschaft 1. zu kümmern, da zu 2. der Hinweis auf die Einheitsmatrix genügt und die in 3. auftretende Matrix $\tilde{\boldsymbol{A}}$ die Inverse zu $\boldsymbol{A}$ ist. Die Aussage 1. folgt aber leicht aus der Hintereinanderschaltung von orthogonalen linearen Abbildungen und ihrer Matrixdarstellung in bezug auf die natürliche Basis (Bestätigung der Details!).

7.5. Die Struktur der orthogonalen linearen Abbildungen des $\mathbb{R}^2$

Zum Abschluß dieses Kapitels verschaffen wir uns zwecks Illustration der allgemein anstehenden Problematik der Strukturaufklärung der linearen Abbildungen wenigstens einen Einblick der Behandlung solch einer Problematik im einfachen zweidimensionalen Fall der orthogonalen linearen Abbildungen.

Es sei uns eine orthogonale lineare Abbildung $A: \mathbb{R}^2 \to \mathbb{R}^2$ des $\mathbb{R}^2$ in sich vorgelegt. Ihre Matrix bezüglich der natürlichen Basis im $\mathbb{R}^2$ sei

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Die Orthogonalität von A hat die Orthogonalität von $\boldsymbol{A}$ zur Folge. Das bedeutet

$$\begin{aligned} a_{11}^2 + a_{12}^2 &= 1, \\ a_{21}^2 + a_{22}^2 &= 1, \\ a_{11}a_{21} + a_{12}a_{22} &= 0. \end{aligned}$$

Die erste und zweite Zeile werden unter Bezugnahme auf trigonometrische Kenntnisse umgeschrieben: Es gibt genau eine Winkelgröße $\vartheta \in [\![0,2\pi[\![$ mit

$$a_{11} = \cos\vartheta, \quad a_{12} = \sin\vartheta.$$

Es gibt genau einen Winkel $\eta \in [\![0,2\pi[\![$ mit

$$a_{21} = \sin\eta, \quad a_{22} = \cos\eta.$$

Die dritte Zeile liefert dann

$$\sin\eta\cos\vartheta + \cos\eta\sin\vartheta = 0,$$

was wiederum gleichwertig mit $\sin(\eta + \vartheta) = 0$ ist. Folglich muß entweder $\eta + \vartheta = 0$ oder $\eta + \vartheta = 2\pi$ oder $\eta + \vartheta = \pi$ gelten.

Im ersten Fall ist A die Einheitsmatrix, d. h., der Operator A ist der identische Operator.

Im zweiten Fall sieht die Matrix wie folgt aus:

$$A = \begin{pmatrix} \cos\vartheta & \sin\vartheta \\ -\sin\vartheta & \cos\vartheta \end{pmatrix}.$$

Welche geometrische Bedeutung hat die zugehörige orthogonale Abbildung A?

Für einen Punkt

$$(x_1, x_2) = (r\cos\alpha, r\sin\alpha)$$

des $\mathbb{R}^2$ ergibt sich als Bildpunkt (y_1, y_2) mit der Koordinatenbeziehung

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \cos\vartheta & \sin\vartheta \\ -\sin\vartheta & \cos\vartheta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1\cos\vartheta + x_2\sin\vartheta \\ -x_1\sin\vartheta + x_2\cos\vartheta \end{pmatrix}$$

$$= r\begin{pmatrix} \cos\alpha\cdot\cos\vartheta + \sin\alpha\cdot\sin\vartheta \\ -\cos\alpha\cdot\sin\vartheta + \sin\alpha\cdot\cos\vartheta \end{pmatrix} = r\begin{pmatrix} \cos(\alpha - \vartheta) \\ \sin(\alpha - \vartheta) \end{pmatrix}.$$

Die Relation

$$(y_1, y_2) = \big(r\cos(\alpha - \vartheta),\ r\sin(\alpha - \vartheta)\big)$$

kann man aber leicht interpretieren. Es handelt sich bei $(x_1, x_2) \mapsto (y_1, y_2)$ um eine Rechtsdrehung um den Ursprung mit dem Drehwinkel ϑ.

Im dritten Fall sieht die Matrix wie folgt aus:

$$A = \begin{pmatrix} \cos\vartheta & \sin\vartheta \\ \sin\vartheta & -\cos\vartheta \end{pmatrix}.$$

Welche geometrische Bedeutung hat die zugehörige orthogonale Abbildung A?

Für einen Punkt

$$(x_1, x_2) = (r\cos\alpha,\ r\sin\alpha)$$

des $\mathbb{R}^2$ ergibt sich als Bildpunkt (y_1, y_2) mit der Koordinatenbeziehung

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \cos\vartheta & \sin\vartheta \\ \sin\vartheta & -\cos\vartheta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1\cos\vartheta + x_2\sin\vartheta \\ x_1\sin\vartheta - x_2\cos\vartheta \end{pmatrix}$$

$$= r \begin{pmatrix} \cos\alpha\cos\vartheta + \sin\alpha\sin\vartheta \\ \cos\alpha\sin\vartheta - \sin\alpha\cos\vartheta \end{pmatrix} = r \begin{pmatrix} \cos(\alpha - \vartheta) \\ -\sin(\alpha - \vartheta) \end{pmatrix}.$$

Die Relation

$$(y_1, y_2) = \big(r\cos(\alpha - \vartheta),\ -r\sin(\alpha - \vartheta)\big)$$

kann man leicht interpretieren. Es handelt sich bei $(x_1, x_2) \mapsto (y_1, y_2)$ um die Hintereinanderschaltung zweier Abbildungen $(x_1, x_2) \mapsto (y_1 - y_2) \mapsto (y_1, y_2)$. Der erste Übergang ist nach den vorherigen Ausführungen eine Rechtsdrehung um den Winkel ϑ.

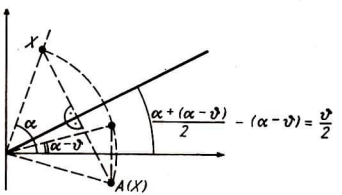


Abb. 10

Der zweite Übergang ist dann eine Spiegelung an der x-Achse. Das kann man nun wiederum weiter interpretieren als eine Spiegelung an der Ursprungsgeraden mit dem Neigungswinkel $\frac{\vartheta}{2}$ (vgl. Abb. 10).

Damit sind also sämtliche Typen der orthogonalen linearen Abbildungen im $\mathbb{R}^2$ aufgefunden, es sind dies in euklidisch-geometrischer Veranschaulichung die Drehungen um den Ursprung und die Spiegelungen an Ursprungsgeraden. Zu einer ähnlich geometrisch vollkommen durchsichtigen Klassifikation der orthogonalen linearen Abbildungen des $\mathbb{R}^3$ kann man ebenfalls gelangen. Die Durchführung des Weges macht von der Eigenwerttheorie linearer Operatoren Gebrauch, die uns hier nicht zur Verfügung steht. Die Eigenwerttheorie behandelt dabei die wichtige Frage, wie zu einem gegebenen linearen Operator des $\mathbb{R}^n$ eine Basis zu ermitteln ist, so daß die beschreibende Matrix des Operators in bezug auf diese Basis möglichst einfach wird. Die denkbar einfachste Gestalt wäre Diagonalform der Matrix, wo also höchstens in der Hauptdiagonalen von Null verschiedene Werte auftreten. Eine Reihe von Operatoren läßt für ihre Matrizen solche Normalformen zu.

7.6. Übungsaufgaben

1. Die Operation des Transponierens für Matrizen war im Text bisher nur im Spezialfall der quadratischen Matrizen vorgekommen. Man kann sie natürlich für allgemeine Matrizen des Typs $m \times n$ erklären. Es sei $\boldsymbol{A}$ eine Matrix des Typs $m \times n$. Dann versteht man unter der *transponierten Matrix* $\boldsymbol{A}^{\mathsf{T}}$ die Matrix vom Typ $n \times m$, deren i-te Zeile gerade die i-te Spalte von $\boldsymbol{A}$ ist, $i = 1, \ldots, n$. Es ist dann also für ein $\boldsymbol{x} \in \mathbb{R}^n$ unter $\boldsymbol{x}^{\mathsf{T}}$ der „Spaltenvektor $\boldsymbol{x}$" zu verstehen!
   a) Man zeige: Sind die Matrizen $\boldsymbol{A}$, $\boldsymbol{B}$ verkettet, so sind die Matrizen $\boldsymbol{B}^{\mathsf{T}}$, $\boldsymbol{A}^{\mathsf{T}}$ verkettet, und es gilt $(\boldsymbol{A} \cdot \boldsymbol{B})^{\mathsf{T}} = \boldsymbol{B}^{\mathsf{T}} \cdot \boldsymbol{A}^{\mathsf{T}}$.
   b) Was bedeutet für $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ das Matrizenprodukt $\boldsymbol{x}\boldsymbol{y}^{\mathsf{T}}$?
2. a) Es sei $\boldsymbol{B}$ eine beliebige quadratische Matrix der Ordnung n.
   Man zeige, daß die folgende Abbildung $B\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ mit dem Verlauf $B(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x}\boldsymbol{B}\boldsymbol{y}^{\mathsf{T}}$ eine Bilinearform auf $\mathbb{R}^n$ ist.
   b) Bezüglich einer geeignet gewählten Basis im $\mathbb{R}^n$ kann man eine vorgegebene Bilinearform des $\mathbb{R}^n$ auch immer in dem Sinne des Teiles a) beschreiben. Man gebe eine genaue Formulierung für diese Aussage und bestätige sie!
   c) Ist die Bilinearform $B\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ mit dem Verlauf $B(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x}\begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix}\boldsymbol{y}^{\mathsf{T}}$ für $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^2$ positiv-definit und symmetrisch?
   d) Was bedeutet die Symmetrie einer Bilinearform $B\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ mit dem Verlauf $B(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x}\boldsymbol{B}\boldsymbol{y}^{\mathsf{T}}$ hinsichtlich der Matrix $\boldsymbol{B}$?
3. Für die folgende Basis des $\mathbb{R}^3$ bringe man das Gram-Schmidtsche Orthonormierungsverfahren hinsichtlich des Skalarprodukts zur Anwendung. Welche Basis erhält man? Man verfolge den Prozeß in der euklidisch-geometrischen Veranschaulichung (Betrachtung gewisser Ebenen!)

$$\boldsymbol{a}_1 = (1, 1, 2),\ \boldsymbol{a}_2 = \left(-1, 3, -\frac{1}{2}\right),\ \boldsymbol{a}_3 = (-4, -1, 0).$$

4. Man gebe eine geometrische Interpretation der Bedingung $\|\boldsymbol{x} + \boldsymbol{y}\| = \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$ (etwa im $\mathbb{R}^3$!).
5. Die im letzten Abschnitt 7.5. aufgetretene orthogonale lineare Abbildung mit der Matrix $\begin{pmatrix} \cos\vartheta & \sin\vartheta \\ \sin\vartheta & -\cos\vartheta \end{pmatrix}$ hinsichtlich der natürlichen Basis erweise man als Spiegelung an der Ursprungsgeraden mit einem Neigungswinkel $\frac{\vartheta}{2}$, indem man die natürliche Basis in eine geeignete neue Lage dreht, die Spiegelung dann hinsichtlich der neuen Basis auf einfachste Weise durch eine Matrix beschreibt und außerdem die Übergangsmatrix von der ersten zur zweiten Basis benutzt!

8. Determinanten

8.1. Vorbemerkungen, insbesondere über Permutationen

Bei der Behandlung der Invertierbarkeit der quadratischen Matrix

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

hatte sich ergeben, daß Invertierbarkeit genau dann vorliegt, wenn der Ausdruck $a_{11}a_{22} - a_{12}a_{21}$ von Null verschieden ist. Untersucht man entsprechend hierzu den Fall der dreireihigen quadratischen Matrix, so würde man erhalten, daß die Matrix

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

genau dann invertierbar ist, wenn der folgende Ausdruck

$$a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$$

von Null verschieden ist.

Für die inverse Matrix tritt in jedem ihrer Glieder der obige Ausdruck als Nenner auf. Ebenso bemerkt man auch bei den vierreihigen quadratischen Matrizen, daß für die Invertierbarkeit ein in bestimmter Weise aus den Elementen der Matrix aufgebauter Ausdruck charakteristisch ist.

Der deutsche Philosoph und Mathematiker G. W. Leibniz (1646—1716) ist diesen Gesetzmäßigkeiten auf die Spur gekommen. Die von ihm entdeckten Determinanten klären die Sachlage; sie waren sodann für die Auflösungstheorie linearer Gleichungssysteme lange Zeit das beherrschende Element. Heutzutage hat sich die Situation mehr zugunsten einer determinantenfreien Behandlung der Grunddinge der linearen Algebra verschoben. Einerseits ist dafür wohl die moderne (generelle) Forderung nach bequemer algorithmischer Behandlung, andererseits auch die Vorleistung auf Fragestellungen der Funktionalanalysis (unendlichdimensionale Vektorräume) mit ausschlaggebend.

Die Leibnizsche Definition der Determinante macht von Kenntnissen über Permutationen Gebrauch. Wir verweisen dazu auf die entsprechenden Ausführungen in MfL, Bd. 1. Hier folgen noch einige Ergänzungen.

Eine Permutation einer nichtleeren Menge X war eine eineindeutige Abbildung von X auf sich. Im Fall der Endlichkeit von X ist die Forderung der Eineindeutigkeit

Abb. 11

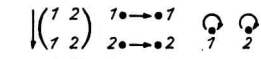


Abb. 12

mit der Forderung der Surjektivität gleichbedeutend. Besteht die Menge X aus den natürlichen Zahlen $1, 2, \ldots, n$ — den ersten natürlichen Zahlen —, so sagt man für eine Permutation der Menge X auch Permutation vom Grade n. Die übliche Schreibweise

$$\begin{pmatrix} 1 & 2 & \ldots & n \\ \pi(1) & \pi(2) & \ldots & \pi(n) \end{pmatrix}$$

für die Permutation π, indem man also unter die Zahl i das Bild von i hinsichtlich der permutierenden Abbildung π schreibt, darf nicht zu Fehldeutungen mit Matrizen verleiten. Zur prägnanteren Unterscheidung würde sich für die Permutation π etwa eine Bezeichnung

$$\downarrow \begin{pmatrix} 1 & 2 & \ldots & n \\ \pi(1) & \pi(2) & \ldots & \pi(n) \end{pmatrix}$$

empfehlen. Das ist jedoch nicht eingebürgert.

Wir lassen jetzt einige Beispiele für Permutationen folgen und machen dabei auf die beiden gebräuchlichen Darstellungen von Permutationen in Diagrammen aufmerksam.

In den Abbildungen 11, 12 und 13 werden Permutationen vom Grade $n = 1$, $n = 2$ und $n = 3$ im Pfeil- oder Leiterdiagramm und im Zyklendiagramm dargestellt.

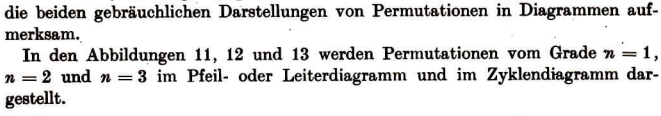


Abb. 13

Bei dem Pfeil- oder Leiterdiagramm werden also zwei Exemplare der zu permutierenden Menge $\{1, 2, \ldots, n\}$ gegenübergestellt und die durch die Permutation veranlaßte Zuordnung mittels Pfeilen von der ersten Reihe zur zweiten Reihe nach Art von verbindenden Sprossen angegeben. Bei dem Zyklendiagramm wird nur ein Exemplar der zu permutierenden Menge $\{1, 2, \ldots, n\}$ verwendet und die durch die Permutation veranlaßte Zuordnung mittels gerichteter Bögen innerhalb dieser Menge ausgedrückt, das gibt das Bild von (mehreren) geschlossenen Bahnkurven (Zyklen). Die erste Art der Darstellung eignet sich besonders bei dem anschaulichen Verfolgen der Hintereinanderschaltung von Permutationen desselben Grades. Eine solche Hintereinanderschaltung ist, wie aus MfL, Bd. 1 bekannt, gemäß der üblichen Zusammensetzung von Abbildungen zu verstehen.

Die Hintereinanderschaltung von Permutationen gleichen Grades nennt man auch Produkt von Permutationen. Über diese Produktbildung war in MfL, Bd. 1 notiert worden, daß es sich um eine (im allgemeinen nicht kommutative) Gruppe handelt. Diese Gruppe heißt die *volle Permutationsgruppe* S_n *vom Grade* n (oder auch die *symmetrische Gruppe vom Grade* n).

Satz 1 (Elementezahl der vollen Permutationsgruppe). *Die volle Permutationsgruppe vom Grade* n *hat* $n!$ *verschiedene Elemente:* $|S_n| = n!$.

Beweis. Die Bestätigung ist einfach und sei dem Leser überlassen. Man wende vollständige Induktion nach n an.

Definition 1 (Signum einer Permutation). Es sei

$$\pi = \begin{pmatrix} 1 & 2 & \cdots & n \\ \pi(1) & \pi(2) & \cdots & \pi(n) \end{pmatrix}$$

eine Permutation der Ordnung n. Man sagt, daß in der Permutation π an den Stellen i, j im Falle $i < j$ eine *Inversion* vorliegt, wenn $\pi(i) > \pi(j)$ ist. Unter dem *Signum* von π versteht man die Zahl

$$\operatorname{sgn} \pi := (-1)^k,$$

wobei k die Anzahl der gesamten Inversionen von π ist.

Eine Permutation heißt *gerade* genau dann, wenn die Anzahl aller ihrer Inversionen gerade ist. Für die geraden Permutationen gilt also $\operatorname{sgn} \pi = +1$. Eine Permutation heißt *ungerade* genau dann, wenn die Anzahl aller ihrer Inversionen ungerade ist. Für die ungeraden Permutationen gilt also $\operatorname{sgn} \pi = -1$.

8.2. Die Leibnizsche Definition der Determinante. Determinanten als Multilinearform

Definition 1 (Determinante einer quadratischen n-reihigen Matrix nach LEIBNIZ). Es sei

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} \dots a_{1n} \\ a_{21} & a_{22} \dots a_{2n} \\ \dots\dots\dots\dots \\ a_{n1} & a_{n2} \dots a_{nn} \end{pmatrix} = (a_{ij})_{i,\,j=1,\dots,n}$$

eine quadratische n-reihige Matrix.

Unter der *Determinante* von $\boldsymbol{A}$ (in Zeichen $\det \boldsymbol{A}$ bzw. $|\boldsymbol{A}|$) versteht man den folgenden Zahlenausdruck:

$$\det \boldsymbol{A} = |\boldsymbol{A}| := \sum_{\pi = \begin{pmatrix} 1 & \dots & n \\ \pi(1) & \dots & \pi(n) \end{pmatrix}} \operatorname{sgn} \pi \; a_{1\pi(1)} a_{2\pi(2)} \cdots a_{n\pi(n)} .$$

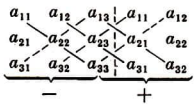
Die Summation wird dabei über alle Permutationen vom Grade n erstreckt, es treten damit genau $n!$ Summanden auf. Die Produkte $a_{1\pi(1)} a_{2\pi(2)} \cdots a_{n\pi(n)}$ werden also so gebildet, daß jedesmal aus jeder Matrixzeile und jeder Spalte genau ein Element entnommen wird.

Einfache Beispiele illustrieren sofort die Berechnung der Determinante:

1. $\boldsymbol{A} = (a_{11})$; $\qquad \det \boldsymbol{A} = a_{11}$.
2. $\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$; $\quad \det \boldsymbol{A} = a_{11}a_{22} - a_{12}a_{21}$.
3. $\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$;

$$\det \boldsymbol{A} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} .$$

Für den letzten Fall der Berechnung der Determinante einer dreireihigen quadratischen Matrix gibt es eine bequeme Merkregel — die *Sarrussche Regel:*



An die Matrix fügt man dazu noch einmal die ersten zwei Spalten an. Die von links oben nach rechts unten verlaufenden Diagonalen liefern dann die Produkte, die addiert werden (wobei sie das durch die Produktbildung entstandene individuelle Vorzeichen

behalten), während die von rechts oben nach links unten verlaufenden Diagonalen Produkte ergeben, die subtrahiert werden (wobei sie wieder das durch die Produktbildung entstandene individuelle Vorzeichen behalten).

Nun untersuchen wir die Haupteigenschaften der Determinanten, d. h. die Eigenschaften der Funktion

$$\det: \mathcal{M}(n \times n) \to \mathsf{R},$$

wo also jeder n-reihigen quadratischen Matrix ihre Determinante zugeordnet wird.

Satz 1 (Grundeigenschaften der Determinante). *Im Bereich $\mathcal{M}(n \times n)$ der quadratischen n-reihigen Matrizen hat die Determinante die folgenden Eigenschaften:*

1. *Besteht für die Matrix $\boldsymbol{A} \in \mathcal{M}(n \times n)$ eine Zeile aus lauter Nullen, so gilt* $\det \boldsymbol{A} = 0$.

2. *Multipliziert man in der Matrix $\boldsymbol{A}$ eine Zeile mit $\lambda \in \mathsf{R}$, so ist die Determinante gleich dem λ-fachen der Ausgangsdeterminante:*

$$\det \begin{pmatrix} \boldsymbol{a}_1 \\ \vdots \\ \lambda \boldsymbol{a}_i \\ \vdots \\ \boldsymbol{a}_n \end{pmatrix} = \lambda \cdot \det \begin{pmatrix} \boldsymbol{a}_1 \\ \vdots \\ \boldsymbol{a}_i \\ \vdots \\ \boldsymbol{a}_n \end{pmatrix}.$$

3. *Ist für die Matrix $\boldsymbol{A} \in \mathcal{M}(n \times n)$ eine Zeile die Summe zweier Zeilenvektoren, so ist die Determinante gleich der entsprechenden Summe der Determinanten:*

$$\det \begin{pmatrix} \boldsymbol{a}_1 \\ \vdots \\ \boldsymbol{a}_i + \boldsymbol{b}_i \\ \vdots \\ \boldsymbol{a}_n \end{pmatrix} = \det \begin{pmatrix} \boldsymbol{a}_1 \\ \vdots \\ \boldsymbol{a}_i \\ \vdots \\ \boldsymbol{a}_n \end{pmatrix} + \det \begin{pmatrix} \boldsymbol{a}_1 \\ \vdots \\ \boldsymbol{b}_i \\ \vdots \\ \boldsymbol{a}_n \end{pmatrix}.$$

4. *Ist für die Matrix $\boldsymbol{A} \in \mathcal{M}(n \times n)$ eine Zeile die Linearkombination von gewissen Zeilenvektoren, so läßt sich die Determinante als entsprechende Linearkombination schreiben.*

5. *Vertauscht man in der Matrix $\boldsymbol{A} \in \mathcal{M}(n \times n)$ zwei beliebige Zeilen miteinander, so ändert sich das Vorzeichen der Determinante.*

Beweis.

Zu 1. In den bei der Erklärung der Determinante auftretenden Produkten ist unter der gemachten Voraussetzung stets die Null Faktor. Die sämtlichen Produkte sind daher alle gleich Null.

Zu 2. Jedes Produkt in der Determinante der Matrix, die aus $\boldsymbol{A}$ entsteht, wenn man alle Elemente einer Zeile von $\boldsymbol{A}$ mit λ multipliziert, enthält den Faktor λ.

Zu 3. Die i-te Zeile der Matrix $\boldsymbol{A}$ sei die Summe der Zeilenvektoren $(a_{i1}, a_{i2}, \ldots, a_{in})$ und $(b_{i1}, b_{i2}, \ldots, b_{in})$. Dann sieht jedes Produkt in der Determinante von $\boldsymbol{A}$ wie folgt aus:

$$\begin{aligned} &\operatorname{sgn} \pi \, a_{1\pi(1)} \cdots (a_{i\pi(i)} + b_{i\pi(i)}) \cdots a_{n\pi(n)} \\ &= \operatorname{sgn} \pi \, a_{1\pi(1)} \cdots a_{i\pi(i)} \cdots a_{n\pi(n)} + \operatorname{sgn} \pi \, a_{1\pi(1)} \cdots b_{i\pi(i)} \cdots a_{n\pi(n)}. \end{aligned}$$

Also läßt sich $\det \boldsymbol{A}$ in der im Satz aufgeführten Weise als Summe zweier Determinanten schreiben.

Zu 4. Es handelt sich um eine Zusammenfassung der Eigenschaften 2. und 3. und eine Ausdehnung auf mehrere Summanden.

Zu 5. Es werde in der Matrix $\boldsymbol{A}$ die i-te Zeile mit der j-ten Zeile vertauscht. Die Produkte in der Determinante der Matrix $\boldsymbol{A}$ sind

$$\operatorname{sgn}\pi \, a_{1\pi(1)} \cdots a_{i\pi(i)} \cdots a_{j\pi(j)} \cdots a_{n\pi(n)},$$

in der Determinante der neuen Matrix lauten die Produkte entsprechend

$$\operatorname{sgn}\pi \, a_{1\pi(1)} \cdots b_{i\pi(i)} \cdots b_{j\pi(j)} \cdots a_{n\pi(n)}.$$

$b_{i\pi(i)}$ ist das $\pi(i)$-te Element der j-ten Zeile von $\boldsymbol{A}$, d. h.,

$$b_{i\pi(i)} = a_{j\pi(i)},$$

$b_{j\pi(j)}$ ist das $\pi(j)$-te Element der i-ten Zeile von $\boldsymbol{A}$, d. h. $b_{j\pi(j)} = a_{i\pi(j)}$. Also hat man

$$\begin{aligned}\operatorname{sgn}\pi \, a_{1\pi(1)} \cdots b_{i\pi(i)} \cdots b_{j\pi(j)} \cdots a_{n\pi(n)} &= \operatorname{sgn}\pi \, a_{1\pi(1)} \cdots a_{i\pi(j)} \cdots a_{j\pi(i)} \cdots a_{n\pi(n)} \\ &= (-1)\operatorname{sgn}\nu \, a_{1\nu(1)} \cdots a_{i\nu(i)} \cdots a_{j\nu(j)} \cdots a_{n\nu(n)},\end{aligned}$$

wobei die Permutationen ν und π wie folgt zusammenhängen:

$$\pi = \begin{pmatrix} 1 & \dots & i & \dots & j & \dots & n \\ \pi(1) & \dots & \pi(i) & \dots & \pi(j) & \dots & \pi(n) \end{pmatrix},$$

$$\nu = \begin{pmatrix} 1 & \dots & i & \dots & j & \dots & n \\ \pi(1) & \dots & \pi(j) & \dots & \pi(i) & \dots & \pi(n) \end{pmatrix}.$$

ν entsteht aus π durch Vertauschung der Elemente $\pi(i)$ und $\pi(j)$, damit ist $\operatorname{sgn}\pi = (-1)\operatorname{sgn}\nu$ (vgl. Übungsaufgabe 3).

Für eine zweckmäßige Fassung der Grundeigenschaften der Determinante verweisen wir auf die folgende Verallgemeinerung der Bilinearformen.

Definition 2 (Multilinearform auf dem $\mathbb{R}^n$). Unter einer *k-Multilinearform* ($k \in \mathbb{N}$, $k \geqq 1$) auf dem $\mathbb{R}^n$ versteht man eine Abbildung

$$\mu: \underbrace{\mathbb{R}^n \times \mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{k\text{-mal}} \to \mathbb{R},$$

die in jedem Argument linear ist, d. h., es gilt für jedes $i = 1, 2, \dots, k$

$$\mu(\boldsymbol{x}_1, \dots, \underbrace{\alpha\boldsymbol{x} + \beta\boldsymbol{y}}_{i\text{-tes Argument}}, \dots, \boldsymbol{x}_k) = \alpha\mu(\boldsymbol{x}_1, \dots, \boldsymbol{x}, \dots, \boldsymbol{x}_k) + \beta\mu(\boldsymbol{x}_1, \dots, \boldsymbol{y}, \dots, \boldsymbol{x}_k).$$

Eine k-Multilinearform μ auf dem $\mathbb{R}^n$ heißt *schiefsymmetrisch* genau dann, wenn μ bei der Vertauschung zweier Argumente das Vorzeichen wechselt:

$$\mu(\boldsymbol{x}_1, \dots, \boldsymbol{x}_i, \dots, \boldsymbol{x}_j, \dots, \boldsymbol{x}_k) = -\mu(\boldsymbol{x}_1, \dots, \boldsymbol{x}_j, \dots, \boldsymbol{x}_i, \dots, \boldsymbol{x}_k).$$

Bemerkung. Ist μ eine k-Multilinearform auf dem $\mathbf{R}^n$, so gilt bei Schiefsymmetrie allgemeiner

$$\mu(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k) = \operatorname{sgn}\pi\, \mu(\boldsymbol{x}_{\pi(1)}, \ldots, \boldsymbol{x}_{\pi(k)})$$

für jede Permutation

$$\pi = \begin{pmatrix} 1 & \ldots & k \\ \pi(1) & \ldots & \pi(k) \end{pmatrix}.$$

Überführt man nämlich in $\mu(\boldsymbol{x}_{\pi(1)}, \ldots, \boldsymbol{x}_{\pi(k)})$ das Argument $\boldsymbol{x}_1 \in \mathbf{R}^n$ in den ersten Platz, so muß man es sukzessive mit allen Vorgängern vertauschen. Dabei tritt jedesmal ein Vorzeichenwechsel auf. So mit allen Argumenten verfahren, wird das Vorzeichen in der Gesamtzahl der Inversionen von π gewechselt. Nach dem Satz über die Grundeigenschaften der Determinante ist die Determinante eine schiefsymmetrische Multilinearform auf dem $\mathbf{R}^n$. Es gilt sogar die folgende Kennzeichnung der Determinante.

Satz 2 (Weierstraßsche Determinantenkennzeichnung als Multilinearform). *Auf dem $\mathbf{R}^n$, $n \geqq 1$, gibt es genau eine schiefsymmetrische reelle n-Multilinearform mit der Eigenschaft, daß der geordneten natürlichen Basis der Wert 1 zugeordnet wird. Diese Multilinearform hat den folgenden Verlauf: Der Wert der Multilinearform für das n-Tupel $(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)$, $\boldsymbol{x}_i \in \mathbf{R}^n$, ist gleich*

$$\det \begin{pmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \vdots \\ \boldsymbol{x}_n \end{pmatrix}.$$

Beweis. $\det: \underbrace{\mathbf{R}^n \times \mathbf{R}^n \times \cdots \times \mathbf{R}^n}_{n\text{-mal}} \to \mathbf{R}$ ist eine Multilinearform mit den angegebenen Eigenschaften. Es sei jetzt μ eine n-Multilinearform auf $\mathbf{R}^n$, die schiefsymmetrisch und normiert ist. Für $\boldsymbol{x}_i \in \mathbf{R}^n$, $i = 1, 2, \ldots, n$, gelte

$$\boldsymbol{x}_i = \sum_{j=1}^{n} x_{ij} \boldsymbol{e}_j$$

hinsichtlich der natürlichen Basis $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ des $\mathbf{R}^n$. Dann ergibt sich

$$\begin{aligned} \mu(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n) &= \mu\left(\sum_{j=1}^{n} x_{1j}\boldsymbol{e}_j, \sum_{k=1}^{n} x_{2k}\boldsymbol{e}_k, \ldots, \sum_{s=1}^{n} x_{ns}\boldsymbol{e}_s\right) \\ &= \sum_{j=1}^{n} \sum_{k=1}^{n} \cdots \sum_{s=1}^{n} x_{1j}x_{2k} \cdots x_{ns}\mu(\boldsymbol{e}_j, \boldsymbol{e}_k, \ldots, \boldsymbol{e}_s). \end{aligned}$$

Nun ist jedes $\mu(\boldsymbol{e}_j, \boldsymbol{e}_k, \ldots, \boldsymbol{e}_s) = 0$ (Schiefsymmetrie), wenn zwei gleiche Basiselemente darin vorkommen. Also reduziert sich die Summe auf

$$\mu(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n) = \sum_{\pi} x_{1\pi(1)}x_{2\pi(2)} \cdots x_{n\pi(n)}\, \mu(\boldsymbol{e}_{\pi(1)}, \boldsymbol{e}_{\pi(2)}, \ldots, \boldsymbol{e}_{\pi(n)})$$

über alle Permutationen π vom Grade n.

Wegen $\mu(\boldsymbol{e}_{\pi(1)}, \boldsymbol{e}_{\pi(2)}, \ldots, \boldsymbol{e}_{\pi(n)}) = \operatorname{sgn}\pi \;\; \mu(\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n) = \operatorname{sgn}\pi$ folgt die Behauptung

$$\mu(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n) = \det \begin{pmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \vdots \\ \boldsymbol{x}_n \end{pmatrix}.$$

Nun beweisen wir noch einige weitere Determinantensätze.

Satz 3 (Kennzeichnung der Regularität einer quadratischen Matrix durch ihre Determinante). *Es sei*

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \\ \vdots \\ \boldsymbol{a}_n \end{pmatrix}$$

eine quadratische Matrix vom Typ $n \times n$. Dann gilt: $\boldsymbol{A}$ hat den Rang n (d. h., die Zeilen $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ bilden eine n-elementige linear unabhängige Menge) $\Leftrightarrow \det \boldsymbol{A} \neq 0$.

Beweis. „$\Rightarrow$“: Angenommen, es ist $\det \boldsymbol{A} = 0$ bei linearer Unabhängigkeit der Zeilenvektoren $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$. Für die Zeilenvektoren $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ der Einheitsmatrix gibt es eine Darstellung durch die Basis $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$:

$$\boldsymbol{e}_i = \sum_{j=1}^{n} e_{ij}\boldsymbol{a}_j, \qquad i = 1, \ldots, n.$$

Dann ist

$$\det \begin{pmatrix} \boldsymbol{e}_1 \\ \boldsymbol{e}_2 \\ \vdots \\ \boldsymbol{e}_n \end{pmatrix} = 1 = \sum_{\pi} e_{1\pi(1)} e_{2\pi(2)} \cdots e_{n\pi(n)} \det \begin{pmatrix} \boldsymbol{a}_{\pi(1)} \\ \vdots \\ \boldsymbol{a}_{\pi(n)} \end{pmatrix}$$

Nun war $\det \boldsymbol{A} = 0$, d. h., es ist auch

$$\det \begin{pmatrix} \boldsymbol{a}_{\pi(1)} \\ \vdots \\ \boldsymbol{a}_{\pi(n)} \end{pmatrix} = 0,$$

was in der Gleichungskette den Widerspruch $1 = 0$ ergibt.

„$\Leftarrow$“: Bei linearer Abhängigkeit der Zeilenvektoren $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ ist etwa $\boldsymbol{a}_1 = \sum_{i=2}^{n} \lambda_i \boldsymbol{a}_i$. Dann ergibt sich

$$\det \begin{pmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \\ \vdots \\ \boldsymbol{a}_n \end{pmatrix} = 0,$$

weil eine Determinante mit zwei gleichen Zeilen gleich Null ist.

Satz 4 (Multiplikationssatz für Determinanten). *Die Abbildung* $\det: \mathcal{M}(n \times n) \to \mathbb{R}$ *ist multiplikativ, d. h., sind* $\boldsymbol{A}, \boldsymbol{B}$ *n-reihige quadratische Matrizen, so gilt*

$$\det(\boldsymbol{AB}) = \det \boldsymbol{A} \cdot \det \boldsymbol{B}.$$

Beweis. Es sei

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \\ \vdots \\ \boldsymbol{a}_n \end{pmatrix}$$

mit $\boldsymbol{a}_i = (a_{i1}, a_{i2}, \ldots, a_{in})$ und

$$\boldsymbol{B} = \begin{pmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \\ \vdots \\ \boldsymbol{b}_n \end{pmatrix}$$

Die Produktmatrix $\boldsymbol{AB}$ hat als i-te Zeile $\sum_{j=1}^{n} a_{ij} \boldsymbol{b}_j$. Es ergibt sich dann

$$\det(\boldsymbol{AB}) = \sum_{j=1}^{n} \sum_{k=1}^{n} \cdots \sum_{s=1}^{n} a_{1j} a_{2k} \cdots a_{ns} \cdot \det \begin{pmatrix} \boldsymbol{b}_j \\ \boldsymbol{b}_k \\ \vdots \\ \boldsymbol{b}_s \end{pmatrix}.$$

Wenn in dem n-Tupel $(j, k, \ldots, s)$ nicht alle Elemente voneinander verschieden sind, ist die Determinante mit den entsprechenden $\boldsymbol{b}_j, \ldots, \boldsymbol{b}_s$ als Zeilenvektoren gleich Null. Damit reduziert sich die Darstellung auf

$$\det(\boldsymbol{AB}) = \sum_{\pi} a_{1\pi(1)} a_{2\pi(2)} \cdots a_{n\pi(n)} \cdot \det \begin{pmatrix} \boldsymbol{b}_{\pi(1)} \\ \boldsymbol{b}_{\pi(2)} \\ \vdots \\ \boldsymbol{b}_{\pi(n)} \end{pmatrix}$$

$$= \sum_{\pi} \operatorname{sgn} \pi \, a_{1\pi(1)} a_{2\pi(2)} \cdots a_{n\pi(n)} \det \boldsymbol{B} = \det \boldsymbol{A} \cdot \det \boldsymbol{B}.$$

Satz 5 (Determinante der transponierten Matrix). *Es sei* $\boldsymbol{A}$ *eine quadratische Matrix der Ordnung n. Die zu* $\boldsymbol{A}$ *transponierte Matrix* $\boldsymbol{A}^{\mathsf{T}}$ *hat die gleiche Determinante wie die Ausgangsmatrix:*

$$\det \boldsymbol{A} = \det \boldsymbol{A}^{\mathsf{T}}.$$

Bemerkung. Nach diesem Satz behalten alle Aussagen über Determinanten ihre Gültigkeit, wenn darin eventuell vorkommende Zeilenaussagen durch solche über Spalten ersetzt werden.

Beweis des Satzes. Es sei $A = (a_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,n}}$. Für die transponierte Matrix A^{T} gilt $A^{\mathsf{T}} = (b_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,n}}$ mit $b_{ij} = a_{ji}$.

Es errechnet sich $\det A^{\mathsf{T}}$ zu

$$\det A^{\mathsf{T}} = \sum_{\pi} \operatorname{sgn} \pi \, b_{1\pi(1)} b_{2\pi(2)} \cdots b_{n\pi(n)} = \sum_{\pi} \operatorname{sgn} \pi \, a_{\pi(1)1} \cdot a_{\pi(2)2} \cdots a_{\pi(n)n}.$$

Bezeichnen wir mit ν die zu π inverse Permutation, d. h., es ist $\pi(i) \overset{\nu}{\mapsto} i$, so kann man das Produkt $a_{\pi(1)1} a_{\pi(2)2} \cdots a_{\pi(n)n}$ auch als $a_{1\nu(1)} a_{2\nu(2)} \cdots a_{n\nu(n)}$ schreiben. Wegen $\operatorname{sgn} \pi = \operatorname{sgn} \nu$ erhalten wir die Bestätigung für die behauptete Determinantengleichung. Das folgende Ergebnis führen wir ohne Beweis an. Sämtliche dazu notwendigen Argumente stehen schon bereit. Der Leser führe zur Übung mindestens den Fall der quadratischen dreireihigen Matrix durch.

Satz 6 (Laplacescher Entwicklungssatz für Determinanten). *Es sei*

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

eine quadratische n-reihige Matrix. A_{ij} sei die aus der Matrix A durch Streichung der i-ten Zeile und j-ten Spalte entstehende quadratische Matrix der Ordnung $n - 1$. Dann gilt folgende Entwicklung der Determinante von A nach der i-ten Zeile:

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \cdot \det A_{ij}.$$

Bemerkung. Nach dem Entwicklungssatz kann man die Berechnung einer Determinante n-ter Ordnung auf die Berechnung von Determinanten $(n-1)$-ter Ordnung zurückführen. Der damit verbundene Rechenaufwand wird aber vergleichsweise groß, wenn man auf solche Art eine Reduktion bis zur zweiten Ordnung vornimmt. Eine direkte Berechnung nach der Leibnizschen Definition ist ebenso für $n \geqq 4$ wegen der $n!$ Summanden im allgemeinen ungeeignet. Hinreichend bequem ist die Determinantenberechnung nach dem Gaußschen Algorithmus der elementaren Zeilenumformung. Durch elementare Zeilenumformung bringt man die gegebene quadratische Matrix auf eine sogenannte Dreiecksform, wo unter der Hauptdiagonalen lauter Nullen stehen. Der gesuchte Determinantenwert ist gleich dem Produkt der Elemente der Hauptdiagonale in der erhaltenen Dreiecksmatrix, weil alle übrigen Produkte verschwinden.

8.3. Zur algebraischen und geometrischen Bedeutung der Determinante

Der Charakterisierungssatz der Regularität einer quadratischen Matrix mittels der Determinante führt uns zu der folgenden Umformulierung unserer Einsichten über die Lösbarkeit von linearen Gleichungssystemen.

Satz 1 (Eindeutige Lösbarkeit eines linearen Gleichungssystems mit quadratischer Koeffizientenmatrix). *Es sei ein lineares Gleichungssystem von n Gleichungen mit n Unbekannten gegeben:*

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{A} \in \mathcal{M}(n \times n), \boldsymbol{b} \in \mathbb{R}^n.$$

Dieses ist genau dann eindeutig lösbar, wenn die quadratische Koeffizientenmatrix A eine von Null verschiedene Determinante hat.

Beweis. Die eindeutige Lösbarkeit ergibt, daß die Lösungsgesamtheit eine nulldimensionale Mannigfaltigkeit ist. Demnach muß der Rang von $\boldsymbol{A}$ gleich n sein, d. h. $\det \boldsymbol{A} \neq 0$. Wenn $\det \boldsymbol{A} \neq 0$ gilt, ist $\boldsymbol{A}$ invertierbar. Also gibt es genau eine Lösung $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$.

Satz 2 (Cramersche Regel zur Bestimmung der Lösung eines linearen Gleichungssystems). *Es sei ein lineares Gleichungssystem von n Gleichungen mit n Unbekannten gegeben (quadratische Koeffizientenmatrix):*

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1, \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2, \\
&\cdots\cdots\cdots\cdots \\
a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n.
\end{aligned}$$

Bei eindeutiger Lösbarkeit gilt für die Lösung folgende Darstellung mittels Determinanten:

$$x_j = \frac{1}{\det \boldsymbol{A}} \cdot \det \begin{pmatrix} a_{11} & a_{12} \dots a_{1j-1} & b_1 & a_{1j+1} \dots a_{1n} \\ a_{21} & a_{22} \dots a_{2j-1} & b_2 & a_{2j+1} \dots a_{2n} \\ \vdots & \vdots & & \\ a_{n1} & a_{n2} \dots a_{nj-1} & b_n & a_{nj+1} \dots a_{nn} \end{pmatrix}$$

für $j = 1, 2, \dots, n$. Die rechtsstehende Matrix ergibt sich aus der Koeffizientenmatrix A mittels Ersetzen der j-ten Spalte durch den Spaltenvektor

$$\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

Beweis. Eindeutige Lösbarkeit des Gleichungssystems gilt genau für den Fall $\det \boldsymbol{A} \neq 0$. Entwickelt man die Determinante $\det \boldsymbol{A}$ nach der ersten Spalte, so ist

$$\det \boldsymbol{A} = \sum_{i=1}^{n} a_{i1}\alpha_{i1};$$

aus

$$\begin{aligned} &a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ &a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{aligned}$$

folgt durch zeilenweise Multiplikation mit $\alpha_{11}, \alpha_{21}, \ldots, \alpha_{n1}$ und Addition

$$\begin{aligned} &(a_{11}\alpha_{11} + a_{21}\alpha_{21} + \cdots + a_{n1}\alpha_{n1}) \cdot x_1 + \\ &(a_{12}\alpha_{11} + a_{22}\alpha_{21} + \cdots + a_{n2}\alpha_{n1}) \cdot x_2 + \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &(a_{1n}\alpha_{11} + a_{2n}\alpha_{21} + \cdots + a_{nn}\alpha_{n1}) \cdot x_n = b_1\alpha_{11} + b_2\alpha_{21} + \cdots + b_n\alpha_{n1}. \end{aligned}$$

In der ersten Klammer steht die Entwicklung der Determinante von $\boldsymbol{A}$ nach der ersten Spalte. Auf der rechten Seite steht die Entwicklung der Determinante derjenigen Matrix, die aus $\boldsymbol{A}$ durch Ersetzen der ersten Spalte durch $\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ entsteht. In den übrigen linken Klammern steht die Entwicklung der Determinante derjenigen Matrix, die aus $\boldsymbol{A}$ durch Ersetzen der ersten Spalte durch die zweite Spalte bzw. dritte Spalte, ..., bzw. n-te Spalte der Matrix $\boldsymbol{A}$ entsteht. Die letztgenannten Determinanten sind sämtlich gleich Null, da sie von Matrizen gebildet werden, die zwei gleiche Spalten aufweisen. Demnach ergibt sich für x_1 die behauptete Darstellung. Entsprechendes folgt für $x_2, \ldots, x_n$.

Der vorstehende Satz liefert uns kein eigentliches neues Lösungsverfahren von linearen Gleichungen mit quadratischer Koeffizientenmatrix. Er ist mehr als eine neue theoretische Einsicht über die geschlossene Darstellung der eindeutig bestimmten Lösung anzusehen.

Ähnlich hat man das folgende Ergebnis zu werten, das eine Umformulierung von Erkenntnissen über den Rang von Matrizen in die Sprache der Determinanten ist.

Satz 3 (Kennzeichnung des Matrizenranges durch Unterdeterminanten). *Es sei* $\boldsymbol{A} = (a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$ *eine Matrix vom Typ* $m \times n$. *Die Matrix* $\boldsymbol{A}$ *hat den Rang* $r \Leftrightarrow$ *Durch Streichen von Zeilen und Spalten kann man aus* $\boldsymbol{A}$ *eine quadratische Untermatrix der Ordnung* r *erhalten, die eine von Null verschiedene Determinante hat, aber es gibt keine quadratische Untermatrix höherer als* r*-ter Ordnung, deren Determinante ungleich Null ist.*

Beweis. Es sei $\boldsymbol{B}$ eine quadratische Untermatrix von $\boldsymbol{A}$ mit $\det \boldsymbol{B} \neq 0$. Dann sind die Zeilen von $\boldsymbol{B}$ linear unabhängig, weil $\boldsymbol{B}$ regulär ist. Die entsprechenden vollen

Zeilen von A sind dann ebenfalls linear unabhängig (Nachweis!). Also ist der Rang von A größer oder gleich der Maximalzahl der Ordnungen der quadratischen Untermatrizen von A, wo eine von Null verschiedene Determinante auftaucht. Wenn r der Rang von A ist, so gibt es r linear unabhängige Zeilen von A. Die auf diese r Zeilen reduzierte Matrix hat dann gewiß r linear unabhängige Spalten, weil der Zeilenrang gleich dem Spaltenrang ist. Reduziert man die Matrix nochmals auf diese r Spalten, so erhält man also eine quadratische Matrix B (Untermatrix von A), die r linear unabhängige Spalten hat. Es muß also $\det B \neq 0$ gelten. Daher ist der Rang von A gleich der im ersten Teil des Beweises betrachteten Maximalzahl.

Wir wollen zum Abschluß der Determinantenbetrachtungen noch Ausführungen machen, die insbesondere wieder bei geometrischen Anliegen aufgenommen werden.

Definition 1 (Das Vektorprodukt oder äußere Produkt im $\mathbb{R}^3$). Unter dem *Vektorprodukt* oder dem *äußeren Produkt* im $\mathbb{R}^3$ versteht man die folgende binäre Operation

$$\times : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$$

mit dem Verlauf

$$\boldsymbol{x} \times \boldsymbol{y} = \left(\det \begin{pmatrix} x_2 & x_3 \\ y_2 & y_3 \end{pmatrix}, \quad -\det \begin{pmatrix} x_1 & x_3 \\ y_1 & y_3 \end{pmatrix}, \quad \det \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} \right)$$

bei $\boldsymbol{x} = (x_1, x_2, x_3)$, $\boldsymbol{y} = (y_1, y_2, y_3)$.

Merkregel zur Bildung des äußeren Produktes

Das äußere Produkt $\boldsymbol{x} \times \boldsymbol{y}$ aus den Elementen $\boldsymbol{x} = (x_1, x_2, x_3)$, $\boldsymbol{y} = (y_1, y_2, y_3)$ des $\mathbb{R}^3$ erhält man durch „formale" Entwicklung der folgenden Determinante nach der ersten Zeile:

$$\det \begin{pmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix}.$$

Satz 4 (Die algebraischen Grundeigenschaften des äußeren Produktes im $\mathbb{R}^3$). *Das äußere Produkt im $\mathbb{R}^3$ hat die folgenden Grundeigenschaften:*

1. $\boldsymbol{x} \times \boldsymbol{y} = -(\boldsymbol{y} \times \boldsymbol{x})$ *für alle* $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^3$ (*Schiefsymmetrie*).
2. $\boldsymbol{x} \times (\boldsymbol{y} + \boldsymbol{z}) = \boldsymbol{x} \times \boldsymbol{y} + \boldsymbol{x} \times \boldsymbol{z}$ *für alle* $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^3$ (*Distributivität*).
3. $\alpha(\boldsymbol{x} \times \boldsymbol{y}) = (\alpha \boldsymbol{x}) \times \boldsymbol{y}$ *für alle* $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^3$ *und* $\alpha \in \mathbb{R}$ (*Assoziativität bei Multiplikation mit einem Skalar*).
4. *Das Assoziativgesetz für das äußere Produkt gilt im allgemeinen nicht.*

Bemerkungen.

1. Die Eigenschaften 2 und 3 ergeben, daß das äußere Produkt eine bilineare Abbildung von $\mathbb{R}^3 \times \mathbb{R}^3$ in den $\mathbb{R}^3$ ist.

2. Wegen der Gültigkeit des Distributivgesetzes heißt diese Operation Produkt. Der Zusatz „äußeres" ist zur Unterscheidung vom inneren Produkt gewählt worden; entsprechend erklärt sich die Bezeichnung Vektorprodukt zum Unterschied vom Skalarprodukt. Das Ergebnis beim Skalarprodukt ist ein Skalar, beim Vektorprodukt ein „Zeilen"vektor.

Beweis. Die Eigenschaften 1, 2 und 3 folgen sofort aus den Determinanteneigenschaften in direkter Rechnung. Zur letzten Aussage (4.) betrachten wir $(\boldsymbol{e}_1 \times \boldsymbol{e}_1) \times \boldsymbol{e}_2$ bzw. $\boldsymbol{e}_1 \times (\boldsymbol{e}_1 \times \boldsymbol{e}_2)$ mit $\boldsymbol{e}_1 = (1, 0, 0), \boldsymbol{e}_2 = (0, 1, 0)$. Es ergibt sich $(\boldsymbol{e}_1 \times \boldsymbol{e}_1) \times \boldsymbol{e}_2 = \boldsymbol{0}$ und $\boldsymbol{e}_1 \times (\boldsymbol{e}_1 \times \boldsymbol{e}_2) = -\boldsymbol{e}_2$.

Satz 5 (Die metrischen Grundeigenschaften des äußeren Produktes). *Über den Zusammenhang des äußeren Produktes mit dem inneren Produkt im* $\mathbb{R}^3$ *gilt:*

1. $\langle \boldsymbol{x}, \boldsymbol{x} \times \boldsymbol{y} \rangle = 0$ für $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^3$.

Das äußere Produkt $\boldsymbol{x} \times \boldsymbol{y}$ *ist orthogonal zu jedem seiner Faktoren.*

2. $\|\boldsymbol{x} \times \boldsymbol{y}\| = \sqrt{\det \begin{pmatrix} \langle \boldsymbol{x}, \boldsymbol{x} \rangle & \langle \boldsymbol{x}, \boldsymbol{y} \rangle \\ \langle \boldsymbol{y}, \boldsymbol{x} \rangle & \langle \boldsymbol{y}, \boldsymbol{y} \rangle \end{pmatrix}}$.

Die Norm des äußeren Produktes $\boldsymbol{x} \times \boldsymbol{y}$ *ist gleich der Wurzel aus der sogenannten Gramschen Determinante von* $\boldsymbol{x}, \boldsymbol{y}$.

Beweis.

1. Es ist für beliebige $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^3$ leicht zu bestätigen:

$$\langle \boldsymbol{x}, \boldsymbol{y} \times \boldsymbol{z} \rangle = \det \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \\ \boldsymbol{z} \end{pmatrix}.$$

Bei $\boldsymbol{x} = \boldsymbol{y}$ erscheinen aber in der Determinante zwei gleiche Zeilen.

2. Man berechnet $\|\boldsymbol{x} \times \boldsymbol{y}\|^2$ und $\det \begin{pmatrix} \langle \boldsymbol{x}, \boldsymbol{x} \rangle & \langle \boldsymbol{x}, \boldsymbol{y} \rangle \\ \langle \boldsymbol{y}, \boldsymbol{x} \rangle & \langle \boldsymbol{y}, \boldsymbol{y} \rangle \end{pmatrix} = \|\boldsymbol{x}\|^2 \|\boldsymbol{y}\|^2 - \langle \boldsymbol{x}, \boldsymbol{y} \rangle^2$ und findet Übereinstimmung.

Satz 6 (Kennzeichnung der linearen Abhängigkeit zweier Elemente des $\mathbb{R}^3$ durch das äußere Produkt). *Im* $\mathbb{R}^3$ *gilt für zwei Elemente* $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^3$: $\boldsymbol{x}, \boldsymbol{y}$ *sind linear abhängig* $\Leftrightarrow \boldsymbol{x} \times \boldsymbol{y} = \boldsymbol{0}$.

Beweis. Lineare Abhängigkeit von $\boldsymbol{x}, \boldsymbol{y}$ ist gleichbedeutend damit, daß der Rang der Matrix

$$\begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix}$$

höchstens gleich 1 ist. Nach dem Rangkennzeichnungssatz mittels Determinanten ist das gleichwertig damit, daß jede zweireihige Untermatrix eine Determinante gleich 0 hat. Das ist äquivalent zu $\boldsymbol{x} \times \boldsymbol{y} = \boldsymbol{0}$.

Unter Berufung auf Schulkenntnisse weisen wir jetzt noch auf mögliche geometrische Interpretationen des äußeren Produktes und dreireihiger Determinanten hin. Eine ausführliche Darstellung dieser Fragen erfolgt in der analytischen Geometrie. Dort wird dann insbesondere das äußere Produkt wie auch das innere Produkt zur Erlangung geometrischer Einsichten herangezogen. Die jetzigen Ausführungen sind nur im Sinne einer ersten Illustration und zur Hebung des Interesses zu verstehen (vgl. MfL, Bde. 6 und 7).

1. Für $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^3$ mißt $\|\boldsymbol{x} \times \boldsymbol{y}\|$ den Flächeninhalt des Parallelogramms, das durch die Strecken vom Ursprung nach $\boldsymbol{x}$ und vom Ursprung nach $\boldsymbol{y}$ aufgespannt wird. Die Ursprungsgerade durch $\boldsymbol{x} \times \boldsymbol{y}$ selber steht senkrecht auf diesem Parallelogramm. Es ist

$$\|\boldsymbol{x} \times \boldsymbol{y}\| = \sqrt{\|\boldsymbol{x}\|^2 \|\boldsymbol{y}\|^2 - \langle \boldsymbol{x}, \boldsymbol{y} \rangle^2}.$$

Nach der geometrischen Bedeutung des Skalarproduktes ersetze man $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ durch $\|\boldsymbol{x}\| \|\boldsymbol{y}\| \cos \sphericalangle (\boldsymbol{x}, \boldsymbol{y})$. Dann ergibt sich

$$\|\boldsymbol{x} \times \boldsymbol{y}\| = \|\boldsymbol{x}\| \|\boldsymbol{y}\| \sin \sphericalangle (\boldsymbol{x}, \boldsymbol{y}),$$

was den genannten Flächeninhalt bedeutet.

2. Für zwei Elemente $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^2$ mißt $\left| \det \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix} \right|$ den Flächeninhalt des durch $\boldsymbol{x}, \boldsymbol{y}$ aufgespannten Parallelogramms (vgl. hierzu die vorstehende Aussage). Anstelle von $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^2$ betrachte man

$$\boldsymbol{x} = (x_1, x_2, 0), \quad \boldsymbol{y} = (y_1, y_2, 0)$$

aus dem $\mathbb{R}^3$. Dann mißt $\|\boldsymbol{x} \times \boldsymbol{y}\|$ den Parallelogramminhalt, und es gilt

$$\|\boldsymbol{x} \times \boldsymbol{y}\| = \left| \det \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix} \right|.$$

3. Für drei Elemente $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^3$ mißt $\left| \det \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \\ \boldsymbol{z} \end{pmatrix} \right|$ den Rauminhalt des durch $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$ aufgespannten Parallelepipeds (auch Spat genannt). Es ist

$$\langle \boldsymbol{x}, \boldsymbol{y} \times \boldsymbol{z} \rangle = \det \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \\ \boldsymbol{z} \end{pmatrix}.$$

$\|\boldsymbol{y} \times \boldsymbol{z}\|$ mißt eine Grundfläche des Spats und $\left\langle \boldsymbol{x}, \dfrac{\boldsymbol{y} \times \boldsymbol{z}}{\|\boldsymbol{y} \times \boldsymbol{z}\|} \right\rangle$ bei $\boldsymbol{y} \times \boldsymbol{z} \neq \boldsymbol{0}$ die entsprechende Höhe (Zeichnung!).

8.4. Übungsaufgaben

1. Von den sämtlichen Permutationen vom Grade 4 bestimme man das Signum. Wie verhält sich dabei das Signum gegenüber der Produktbildung von Permutationen?
2. Unter einer Transposition versteht man eine Permutation, die alle bis auf zwei Elemente fest läßt. (Es werden nur zwei Elemente miteinander vertauscht!) Man zeige, daß jede Transposition eine ungerade Permutation ist.
3. Es sei π eine gegebene Permutation vom Grade $n \geqq 2$. In der Menge $\{1, \ldots, n\}$ werden zwei Elemente i, j fixiert. Die wie folgt erklärte Abbildung $\nu\colon \{1, \ldots, n\} \to \{1, \ldots, n\}$ mit dem Verlauf $\nu(k) := \pi(k)$ für $k \neq i, j$ und $\nu(i) := \pi(j)$, $\nu(j) := \pi(i)$ ist dann wieder eine Permutation vom Grade n. ν entsteht, indem man zuerst π ausführt und dann noch die Elemente $\pi(i)$ und $\pi(j)$ vertauscht. Es gilt dann $\operatorname{sgn} \pi = (-1)\operatorname{sgn} \nu$. (Man stelle eine Beziehung zur Aufgabe 2 her!)
4. Von den in der Aufgabe 6 von 6.7. genannten Matrizen berechne man die Determinanten.
5. Von den quadratischen Matrizen der Ordnung $n = 2, 3, 4, 5$, die in sukzessiver Aufeinanderfolge die ersten n^2 natürlichen Zahlen $1, 2, \ldots, n^2$ als Elemente enthalten, berechne man die Determinanten.
6. Von den orthogonalen quadratischen Matrizen berechnet sich die Determinante entweder zu $+1$ oder -1 (Produktsatz verwenden!).
7. Im Zusammenhang mit dem Laplaceschen Entwicklungssatz bestätige man, daß die Inverse einer regulären quadratischen Matrix $\boldsymbol{A}$ der Ordnung 3 sich wie folgt berechnet:

$$\boldsymbol{A}^{-1} = \frac{1}{\det \boldsymbol{A}} \begin{pmatrix} +\det A_{11} & -\det A_{21} & +\det A_{31} \\ -\det A_{12} & +\det A_{22} & -\det A_{32} \\ +\det A_{13} & -\det A_{23} & +\det A_{33} \end{pmatrix}.$$

   (Eine entsprechende Darstellung für $\boldsymbol{A}^{-1}$ beweise man für beliebiges n.)
8. Für die sogenannte *Vandermondesche Determinante*

$$\Delta(x_1, x_2, \ldots, x_n) = \det \begin{pmatrix} 1 & 1 & \ldots & 1 \\ x_1 & x_2 & \ldots & x_n \\ x_1^2 & x_2^2 & \ldots & x_n^2 \\ \ldots & \ldots & \ldots & \ldots \\ x_1^{n-1} & x_2^{n-1} & \ldots & x_n^{n-1} \end{pmatrix}$$

   bestätige man $\Delta = \prod_{i>j} (x_i - x_j)$ (Produkt über alle Binome $(x_i - x_j)$ mit $1 \leqq j < i \leqq n$ genommen).
9. Man bestimme eine Matrix

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$$

   mit vorgegebenen Werten

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = b, \quad \det \begin{pmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{pmatrix} = c, \quad \det \begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{pmatrix} = d.$$

9. Der Begriff des Vektorraumes

Die im n-dimensionalen reellen Zahlenraum $\mathbf{R}^n$, in der Menge der linearen Funktionale $\mathscr{L}(\mathbf{R}^n)$, in der Menge der linearen Abbildungen $\mathscr{L}(\mathbf{R}^n, \mathbf{R}^n)$ oder in der Menge $\mathbf{R}^X$ der reellen Funktionen auf X angestellten arithmetischen Betrachtungen haben einen gemeinsamen algebraischen Kern. Es war schon darauf hingewiesen worden, daß man eine der Gemeinsamkeiten mit dem algebraischen Begriff der Gruppe beschreibt, indem man sagt, daß die genannten Mengen hinsichtlich einer darin erklärten binären Operation (hier handelt es sich um die koordinatenweise bzw. um die punktweise Addition) jeweils eine Gruppe bilden. In diesen Gruppen war nun darüber hinaus noch eine weitere Operation erklärt, nämlich eine Multiplikation der Gruppenelemente mit reellen Zahlen. Die Verquickung der Gruppenaddition mit der Skalarmultiplikation unterlag dabei stets ein und denselben Rechengesetzen. Man sagt zu diesem Sachverhalt, daß es sich bei den betrachteten algebraischen Strukturen um lineare Räume (oder auch Vektorräume) handelt.

Definition 1 (Reeller linearer Raum oder reeller Vektorraum). Es sei E eine nichtleere Menge, die mit einer binären Operation „+“

$$+ : E \times E \to E, \quad x, y \in E \mapsto x + y \in E$$

sowie einer „Skalarmultiplikation“ „$\cdot$“

$$\cdot : \mathbf{R} \times E \to E, \ \alpha \in \mathbf{R}, \ x \in E \mapsto \alpha \cdot x \in E$$

ausgestattet ist.

Man sagt, daß E hinsichtlich dieser beiden Operationen einen *reellen linearen Raum* oder einen *reellen Vektorraum* bildet genau dann, wenn folgende Grundeigenschaften erfüllt sind:

I. Bezüglich der Operation „+“ gilt:

1. $x + y = y + x$ für alle $x, y \in E$ (Kommutativität der Operation $+$).
2. $(x + y) + z = x + (y + z)$ für alle $x, y, z \in E$ (Assoziativität der Operation $+$).
3. Es gibt ein (eindeutig bestimmtes) Element $0 \in E$, so daß $x + 0 = x$ für alle $x \in E$ ist (Existenz und Einzigkeit eines Nullelementes).

4. Zu jedem $x \in E$ existiert ein (eindeutig bestimmtes) Element, bezeichnet durch $-x$, für welches $x + (-x) = 0$ ist (Existenz und Einzigkeit der Inversen).

II. Für die Multiplikation mit einem reellen Skalar gilt:

5. $1 \cdot x = x$ für alle $x \in E$.

6. $(\alpha \cdot \beta)x = \alpha(\beta x)$ für alle $\alpha, \beta \in \mathbf{R}$ und alle $x \in E$ (Assoziativität der Multiplikation mit einem reellen Skalar).

III. Für das Zusammenspiel der Addition $+$ mit der Multiplikation mit einem reellen Skalar gilt:

7. $(\alpha + \beta)x = \alpha x + \beta x$ für alle $\alpha, \beta \in \mathbf{R}$ und alle $x \in E$ (Distributivität der Multiplikation mit einem Skalar bezüglich der Addition von Skalaren).

8. $\alpha(x + y) = \alpha x + \alpha y$ für alle $\alpha \in \mathbf{R}$ und alle $x, y \in E$ (Distributivität der Multiplikation mit einem Skalar bezüglich der Addition).

Bemerkung. Tritt an die Stelle des Skalarbereiches $\mathbf{R}$ der Skalarbereich $\mathbf{C}$ der komplexen Zahlen, so heißt die Struktur ein komplexer linearer Raum oder ein *komplexer Vektorraum*. Die eingangs gemachten Hinweise können als Verweise auf Beispiele zu den reellen linearen Räumen verstanden werden. Damit sind aber die Möglichkeiten für illustrierende Beispiele zu diesem Begriff noch längst nicht erschöpft. In der Analysis wird sich etwa noch ergeben, daß die stetigen reellen Funktionen, die differenzierbaren reellen Funktionen, die integrierbaren reellen Funktionen, die Folgen mit konvergenter Reihe usw. reelle lineare Räume (bei entsprechender Abänderung des Wertebereichs auch komplexe lineare Räume) bilden. Zum anderen erscheinen in der Geometrie ebenfalls lineare Räume, beispielsweise in Gestalt der Menge der Parallelverschiebungen in der Ebene bezüglich der Hintereinanderschaltung und der Multiplikation mit reellen Skalaren. Damit ist wohl hinreichend gerechtfertigt, daß man einen solchen Begriff wie den des linearen Raumes im Sinne einer Denkökonomie als abstrakten mathematischen Begriff hervorkehrt. Die Bezeichnung „linearer Raum" für diesen Begriff ist uns dabei erklärlich, jedoch bedarf es noch einiger Worte über die Bezeichnungsweise „Vektorraum".

Der deutsche Mathematiker Hermann Grassmann (1809—1877) hatte zum Zwecke einer geeigneten Behandlung geometrischer Probleme der Ebene und des Raumes die geometrische Vektormethode (das Operieren mit gerichteten Strecken) erfunden. Dabei verwirklichte Grassmann eigentlich eine Idee von Leibniz, der eine „characteristica geometrica" gefordert hatte, worunter eine Lehre zu verstehen wäre, „die die geometrischen Lagebeziehungen so zum Ausdruck bringt, wie die Algebra Größenbeziehungen". Außer Grassmann muß als Urheber einer Vektortheorie noch der irische Mathematiker W. R. Hamilton (1805—1865) genannt werden. Hamilton schuf mit den nach ihm benannten Quaternionen die meisten heute bekannten Einsichten der Vektortheorie. Von ihm stammt auch die Bezeichnung „Vektor" (vehere (lat.): fahren).

Heute hat die Vektortheorie außer den geometrischen Aspekten, die in Band 6 und 7 betrachtet werden, algebraische Aspekte. Die Bezeichnung Vektor steht dabei

allgemein für ein Element eines linearen Raumes (eines Vektorraumes). Dem „Vektor" kommt also keine absolute Bedeutung zu. Dieser Begriff versteht sich nur im Zusammenhang mit einer bestimmten algebraischen Struktur. Vielfach übliche, in der Geometrie zu präzisierende Äußerungen, daß Vektoren mathematische Größen bezeichnen, die außer durch einen Zahlenwert noch durch eine Richtung gekennzeichnet werden, sind in diesem Sinne unzutreffend, weil zu pauschal. Sie können höchstens als eine leichte Andeutung auf die Elemente des für die Elementargeometrie wichtigen „Vektorraumes" der Verschiebungen aufgefaßt werden.

10. Lineare Ungleichungen. Lineare Optimierung

Betrachtungen von linearen Ungleichungen werden zum Beispiel bei Fragen der linearen Optimierung nötig. Wir bringen ein sehr einfaches, aber instruktives Beispiel einer Aufgabe der linearen Optimierung.

Ein Betrieb stellt zwei Waren W_1 und W_2 her, wobei für jede Ware Rohstoffe dreierlei Art R_1, R_2, R_3 benötigt werden. Von dem Rohstoff R_i seien b_i Mengeneinheiten vorhanden. Zur Herstellung von W_1 möge pro Mengeneinheit benötigt werden:

a_{11} vom Rohstoff R_1,

a_{21} vom Rohstoff R_2,

a_{31} vom Rohstoff R_3.

Zur Herstellung von W_2 möge pro Mengeneinheit benötigt werden:

a_{12} vom Rohstoff R_1,

a_{22} vom Rohstoff R_2,

a_{32} vom Rohstoff R_3.

In den Koeffizienten a_{ij} gibt also der erste Index die Rohstoffart und der zweite Index die Warenart an. Die Ware W_i erbringe dem Betrieb pro Einheit einen gewissen Gewinn p_i, $i = 1, 2$.

Das interessierende Problem bestehe in der Frage: Wie muß der Betrieb seine Planung einrichten, damit der Gesamtgewinn möglichst groß wird? Es wird also die Zahl x_i der Wareneinheiten W_i, $i = 1, 2$, erfragt, die produziert werden müssen, um optimalen Gewinn zu erzielen. Mathematisch gesehen handelt es sich um die Bestimmung des Maximums der Funktion

$$\varphi(x_1, x_2) = p_1 x_1 + p_2 x_2 \quad (\textit{Zielfunktion}).$$

Hierbei sind den (x_1, x_2) folgende Bedingungen auferlegt:

$$x_1 \geqq 0, \; x_2 \geqq 0$$

(sogenannte Randbedingungen);

$$a_{11}x_1 + a_{12}x_2 \leqq b_1,$$
$$a_{21}x_1 + a_{22}x_2 \leqq b_2,$$
$$a_{31}x_1 + a_{32}x_2 \leqq b_3$$

(sogenannte Balance-Bedingungen).

Die Randbedingungen sind von sich aus klar. Die Balancebedingungen drücken die zu berücksichtigenden Stoffbeschränkungen aus. Die Zielfunktion φ ist ein lineares Funktional. Von diesem Funktional sind auf der Lösungsmenge der angeschriebenen fünf linearen Ungleichungen die Maximalstellen zu ermitteln!

In der Praxis treten solche linearen Optimierungsprobleme häufig auf. Zur rechnerischen Bewältigung gibt es verschiedene Verfahren (Algorithmen). Wir gehen auf die Lösungsverfahren nicht ein, sondern wir schaffen uns noch einen kleinen Einblick in die grundsätzliche Struktur der Problematik. Da wäre vor allem das Aussehen der Lösungsmenge von linearen Ungleichungen zu klären.

Im $\mathbf{R}^n$ sei eine lineare Ungleichung

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n \leqq b$$

mit den vorgeschriebenen Koeffizienten $a_1, a_2, \ldots, a_n$ und der vorgeschriebenen rechten Seite gegeben. Es ist eine Einsicht in die Lösungsmenge dieser Ungleichung gesucht. Um die Dinge geometrisch-anschaulich verfolgen zu können, beschränken wir uns zunächst auf den Fall $n \leqq 3$. Wird anstelle der Ungleichung die entsprechende Gleichung betrachtet, so ist die Lösungsmenge eine Ebene, Gerade oder aber ein Punkt (in Abhängigkeit von $n = 3$, $n = 2$ oder $n = 1$).

Für $n = 1$ stellt die Lösungsmenge einer linearen Ungleichung alle Punkte einer Halbgeraden dar, die durch den der Gleichung entsprechenden Punkt bestimmt wird. Für $n = 2$ stellt die Lösungsmenge einer linearen Ungleichung $f(\boldsymbol{x}) \leqq b$ alle Punkte einer Halbebene dar, die durch die Gerade bestimmt wird, welche der Gleichung $f(\boldsymbol{x}) = b$ entspricht. Analog stellt die Lösungsmenge einer linearen Ungleichung $f(\boldsymbol{x}) \leqq b$, $f \in \mathscr{L}(\mathbf{R}^3)$, alle Punkte eines Halbraumes dar, der durch die Ebene bestimmt wird, welche der linearen Gleichung $f(\boldsymbol{x}) = b$ entspricht. Zur Begründung der Aussagen betrachte man die Mengen

$$H_1 := \{\boldsymbol{x} : \boldsymbol{x} \in \mathbf{R}^n, \; f(\boldsymbol{x}) < b\}, \quad H := \{\boldsymbol{x} : \boldsymbol{x} \in \mathbf{R}^n, \; f(\boldsymbol{x}) = b\},$$
$$H_2 := \{\boldsymbol{x} : \boldsymbol{x} \in \mathbf{R}^n, \; f(\boldsymbol{x}) > b\}.$$

Diese drei Mengen stellen eine Zerlegung des $\mathbf{R}^n$ dar.

Nun läßt sich folgende Feststellung beweisen: Von zwei Elementen $\boldsymbol{y}, \boldsymbol{z} \in \mathbf{R}^n \setminus H$ gehört eins zu H_1, das andere zu H_2 genau dann, wenn die „Verbindungsstrecke" von

y, z die Menge H überschneidet, d. h., wenn

$$\{u: u \in \mathsf{R}^n,\ u = ty + (1-t)z,\ 0 \leqq t \leqq 1\} \cap H \neq \emptyset$$

gilt. Das bedeutet also, daß die Mengen H_1 und H_2 die beiden verschiedenen (offenen) Halbräume sind, die beim Herauslösen von H aus dem R^n entstehen. Die Lösungsmenge der betrachteten linearen Ungleichung $f(x) \leqq b$ ist damit ein „abgeschlossener" Halbraum, wo also die Punkte der begrenzenden Hyperebene H noch mitgerechnet werden. Hieraus folgt sodann, daß die Lösungsmenge eines Systems von m linearen Ungleichungen der Durchschnitt von (abgeschlossenen) Halbräumen ist. Es handelt sich um eine gewisse konvexe Menge. Beispielsweise ist im R^3 die Lösungsmenge des Ungleichungssystems

$$\begin{aligned} x_1 &\geqq 0, \\ x_2 &\geqq 0, \\ x_3 &\geqq 0, \\ x_1 + x_2 + x_3 &\leqq 1 \end{aligned}$$

das durch die Einheitspunkte auf den x_i-Achsen und dem Ursprungspunkt aufgespannte Tetraeder.

11. Algebraische Strukturen

11.1. Einleitung

In den Bemerkungen zur Geschichte der Algebra wurde bereits darauf aufmerksam gemacht, wie sich die Auffassungen vom Inhalt der Algebra während der historischen Entwicklung dieses Zweiges der Mathematik mehrmals geändert haben. Einer elementaren Lehre über die Auflösung von Gleichungen folgte die Beschäftigung mit der „Buchstabenrechnung", ehe das schwierige Problem der Auflösung einer Gleichung n-ten Grades in einer Unbekannten in den Mittelpunkt des Interesses rückte. GAUSS bewies, daß jedes Polynom positiven Grades, dessen Koeffizienten komplexe Zahlen sind, in der Menge der komplexen Zahlen mindestens eine Nullstelle besitzt. Entsprechend der zu jener Zeit herrschenden Auffassung vom Inhalt der Algebra wurde dieses Resultat als *Fundamentalsatz der Algebra* bezeichnet.

War auch der Hauptinhalt der „klassischen Algebra", wie sie von WEBER in seinem 1896 erschienenen *Lehrbuch der Algebra* dargestellt wurde, Gleichungstheorie, so erschienen doch in dieser Entwicklungsphase bereits Elemente einer „Strukturalgebra". In den Überlegungen von ABEL und GALOIS erwies sich der *Gruppenbegriff* als ein wesentliches Hilfsmittel zur Beherrschung des Problems der Gleichungsauflösung. Auch andere Strukturbegriffe wie *Modul*, *Körper*, *Ring* und *Verband* wurden bereits in dieser Zeit durch GAUSS, DEDEKIND, HILBERT und DEDEKIND eingeführt. Es stellte sich jedoch heraus, daß diese Hilfsbegriffe der klassischen Algebra weitaus größere Anwendungsmöglichkeiten besaßen, als man ursprünglich annehmen konnte, denn es setzte sich die Erkenntnis durch, daß die Beschaffenheit der Elemente einer solchen Struktur bei vielen Überlegungen unbeachtet bleiben darf. STEINITZ, der 1910 in seiner epochemachenden Arbeit *Algebraische Theorie der Körper* mit der Aufstellung des *Isomorphieprinzips* die Richtung der algebraischen Untersuchungen bestimmte, gab die präzise Formulierung dieser Auffassung.

Ausgehend vom Mengenbegriff entstand durch Hinzunahme von Operationen und Relationen der Begriff der *algebraischen Struktur*. Stellt man an eine solche noch bestimmte Forderungen (die i. allg. als ein Axiomensystem für die betrachtete Struktur formuliert werden), so erhält man Typen von Strukturen, wie sie unter den Namen Gruppe, Ring, Körper usw. geläufig sind. Ihre Erforschung ist die Aufgabe

der Algebra geworden. Diese „moderne Algebra" stellte VAN DER WAERDEN 1930/31 in seinem gleichnamigen, für die ganze Entwicklungsetappe grundlegenden Lehrbuch dar. Seit der vierten Auflage im Jahre 1955 gab ihm der Autor jedoch den Titel *Algebra*. In den letzten Jahrzehnten riefen nämlich die Eigenentwicklung der Algebra sowie die Ansprüche, welche die Anwendungsgebiete (Topologie, Funktionalanalysis, algebraische Geometrie u. a.) an sie stellten, einige neue und wichtige Zweige der Algebra (Verbandstheorie, Theorie der Halbgruppen, homologische Algebra, universelle Algebra u. a.) ins Leben. Einen Einblick in diese neue Entwicklung vermittelte KUROŠ in seinem Buch *Vorlesungen über allgemeine Algebra*, dessen deutsche Übersetzung 1964 erschien.

11.2. Der axiomatische Aufbau einer Theorie

In einer mathematischen Theorie betrachtet man Mengen von Objekten, zwischen denen gewisse Beziehungen erklärt sind, welche bestimmte Eigenschaften haben (vgl. MfL, Bd. 1, 2.7.). Der Inhalt der Theorie besteht nun darin, Begriffe und Beziehungen durch andere zu definieren und Eigenschaften von Beziehungen und Begriffen zu beweisen. Zum Beispiel kann man, von den natürlichen Zahlen ausgehend, die rationalen Zahlen definieren, kann auf der Menge der rationalen Zahlen eine zweistellige Operation erklären, die man Addition nennt, und von dieser etwa beweisen, daß sie kommutativ ist. Es ist unmittelbar einzusehen, daß man nicht alle Begriffe und Beziehungen definieren und sämtliche Eigenschaften allein mit Hilfe der Logik beweisen kann, da ja jede Definition den zu definierenden Begriff auf andere Begriffe zurückführt und beim Beweis jeder Eigenschaft andere Eigenschaften benutzt werden. Deshalb muß man bei jeder mathematischen Theorie gewisse Begriffe und Beziehungen ohne Definition an die Spitze stellen. Das sind die *Grundbegriffe* und *Grundbeziehungen* der betreffenden Theorie. Ebenso muß man gewisse Eigenschaften der Grundbegriffe und Grundbeziehungen voraussetzen. Diese sogenannten *Grundeigenschaften* werden in den *Axiomen* formuliert. Nach der Aufstellung der Grundbegriffe und Grundbeziehungen sowie der Axiome, die das Fundament der Theorie bilden, werden die abgeleiteten Begriffe und Sätze auf rein logischem Wege aus ihnen gefolgert.

Ein solcher *deduktiver* Aufbau ist für die Mathematik typisch. Es wäre jedoch falsch, wollte man *induktive* Betrachtungsweisen, die von speziellen Beobachtungen zu allgemeineren Aussagen gelangen, als bedeutungslos für die Mathematik ansehen. Die Mathematik ist wie jede andere Wissenschaft aus praktischen Erfahrungen entstanden und dient praktischen Bedürfnissen. Die Grundbegriffe und Axiome einer Theorie sind Widerspiegelungen der objektiven Realität mit einem häufig sehr hohen Abstraktionsgrad. Ehe der axiomatische Aufbau einer Theorie begonnen wird, sind oft bereits Kenntnisse vorhanden. So waren beispielweise die natürlichen Zahlen und

ihre Eigenschaften längst bekannt, ehe von dem *Peanoschen Axiomensystem* ausgehend ein axiomatischer Aufbau der Theorie der natürlichen Zahlen erfolgte. Auch waren bereits umfangreiche Kenntnisse der Geometrie vorhanden, als man einen axiomatischen Aufbau begann, der in den berühmten, im 3. Jahrhundert v. u. Z. in Alexandria geschriebenen „Elementen" des EUKLID dargestellt wurde. Es wäre aber auch verkehrt, wenn man die axiomatische Methode nur als ein Verfahren zur eleganten Darstellung einer Theorie ansehen wollte. Die Kritik am Axiomensystem des EUKLID hat die mathematische Forschung in zwei Jahrtausenden angeregt. Am Anfang des vorigen Jahrhunderts entdeckten NIKOLAI IWANOWITSCH LOBATSCHEWSKI (1793—1856), GAUSS und JANOS BOLYAI (1802—1860), ausgehend von einem Axiomensystem, die (hyperbolische) nichteuklidische Geometrie.

Wir haben Axiome als Grundaussagen ohne Beweis an die Spitze einer Theorie gestellt. Damit soll nicht ausgedrückt werden, daß die Axiome etwa deshalb keines Beweises bedürften, weil sie „offensichtlich" gelten. Ebensowenig vertreten wir den Standpunkt, daß die Grundbegriffe und Grundbeziehungen nicht definiert zu werden brauchten, weil ihre Bedeutung „klar" ist. Da die Grundbegriffe und Grundbeziehungen nicht aus der Theorie heraus erklärt werden, ergibt sich die Frage nach ihrer Bedeutung. Ist sie geklärt, so kann man auch sinnvoll nach der Wahrheit der Axiome fragen.

Man nennt eine vorgegebene Menge von Objekten, zwischen denen gewisse Beziehungen bestehen, ein *Modell* oder eine *Interpretation* eines Axiomensystems, wenn bei einer Belegung der Grundbegriffe und Grundbeziehungen des Axiomensystems durch konkrete Objekte und Beziehungen zwischen Objekten der gegebenen Menge die Aussagen des Axiomensystems zutreffen.

In diesem Fall kann dann die ganze Theorie auf die vorgegebene Menge und die zwischen ihren Elementen bestehenden Beziehungen angewendet werden. In einem festen Modell eines Axiomensystems haben die Grundbegriffe und Grundbeziehungen also eine wohlbestimmte Bedeutung, und die im Axiomensystem formulierten Aussagen sind für das Modell wahr, weil sie dort zutreffen.

Die Menge der Kardinalzahlen der endlichen Mengen liefert ein solches Modell für das Peanosche Axiomensystem (vgl. MfL, Bd. 1, 3.1.) und die Menge $\mathfrak{T}(M)$ aller 1-1-Abbildungen einer beliebigen nichtleeren Menge M auf sich bezüglich der Verkettungsoperation ein Modell für die Gruppenaxiome (vgl. MfL, Bd. 1, 2.4. (17a) bis (17c)). Wir werden im nächsten Abschnitt einige Axiomensysteme für algebraische Strukturen und zugehörige Modelle angeben.

Aus einem gegebenen Axiomensystem kann man weitere Aussagen ableiten. Erhält man dabei zwei Sätze, die einander widersprechen, d. h., von denen der eine die Negation des anderen ist, so heißt das Axiomensystem *widerspruchsvoll*. Grundlage für eine inhaltsreiche mathematische Theorie kann selbstverständlich nur ein Axiomensystem sein, das nicht widerspruchsvoll ist. Wenn aus einem Axiomensystem bis zu einem bestimmten Zeitpunkt keine Widersprüche entwickelt wurden, so ist das natürlich kein Beweis für seine *Widerspruchsfreiheit*. Es entsteht also die

Frage, wie man sich von der Widerspruchsfreiheit eines vorgelegten Axiomensystems überzeugen kann. KURT GÖDEL zeigte 1931, daß es unmöglich ist, innerhalb einer Theorie ihre Widerspruchsfreiheit nachzuweisen. Deshalb wird folgendes Verfahren benutzt: Hat man eine Theorie $\mathfrak{T}$, an deren Widerspruchsfreiheit man nicht zweifelt, sowie ein Axiomensystem $\mathfrak{A}$, dessen Widerspruchsfreiheit untersucht werden soll, und gelingt es, mit den Begriffen der Theorie $\mathfrak{T}$ ein Modell für das Axiomensystem $\mathfrak{A}$ zu konstruieren, so ist $\mathfrak{A}$ (*relativ* zur Theorie $\mathfrak{T}$) *widerspruchsfrei*.

Häufig ist es üblich, die Arithmetik der rationalen Zahlen als widerspruchsfreie Theorie $\mathfrak{T}$ zu akzeptieren und ein Axiomensystem $\mathfrak{A}$ sowie die darauf aufgebaute Theorie als widerspruchsfrei zu bezeichnen, wenn man mittels der Arithmetik der rationalen Zahlen ein Modell für $\mathfrak{A}$ konstruieren kann. Die im nächsten Abschnitt angegebenen Axiomensysteme für algebraische Strukturen werden sich in diesem Sinne als widerspruchsfrei erweisen.

Eine zweistellige Relation R in einer Menge M heißt *Äquivalenzrelation*, wenn sie den Axiomen

$$\bigwedge_{x \in M} xRx \qquad (R \text{ ist reflexiv}),$$

$$\bigwedge_{x,y,z \in M} (xRy \wedge yRz \Rightarrow xRz) \qquad (R \text{ ist transitiv}),$$

$$\bigwedge_{x,y \in M} (xRy \Rightarrow yRx) \qquad (R \text{ ist symmetrisch})$$

genügt (vgl. MfL, Bd. 1, 2.5. (12)).

Dieses System ist widerspruchsfrei, denn z. B. stellt die Menge der ganzen Zahlen mit der Gleichheit als Relation R ein Modell dar. Auch die Kongruenzrelation R_m in der Menge der natürlichen Zahlen (vgl. MfL, Bd. 1, 3.7. (63)) ist eine Äquivalenzrelation.

Ein Axiom heißt *unabhängig* von den übrigen Axiomen eines Systems, wenn es sich nicht aus diesen als Satz beweisen läßt. Beim axiomatischen Aufbau einer mathematischen Theorie bemüht man sich natürlich darum, möglichst nur unabhängige Axiome im Axiomensystem zu haben, d. h. das Axiomensystem „minimal" zu halten. Manchmal wird im Interesse der leichteren Handhabung einer axiomatischen Theorie ein mit abhängigen Axiomen angereichertes Axiomensystem verwendet, mit dessen Hilfe man oft schneller zu interessanten Folgerungen kommt. Die Unabhängigkeit eines Axioms A von den übrigen Axiomen eines gegebenen Systems $\mathfrak{S}$ kann man nachweisen durch Konstruktion eines Modells für dasjenige Axiomensystem $\mathfrak{S}'$, welches entsteht, wenn man in $\mathfrak{S}$ das Axiom A durch seine Negation ersetzt und alle anderen Axiome beibehält.

Die nachstehenden Relationen sind jeweils in der Menge $M = \{1, 2, 3\}$ erklärt:

$$R_1 = \{(x, y): x \in M \wedge y \in M \wedge x \leqq 2 \wedge y \leqq 2\}$$

ist nicht reflexiv, aber transitiv und symmetrisch,

$$R_2 = \{(x, y): x \in M \wedge y \in M \wedge |x - y| \leqq 1\}$$

ist nicht transitiv, aber reflexiv und symmetrisch.

$$R_3 = \{(x, y): x \in M \wedge y \in M \wedge x \leqq y\}$$

ist nicht symmetrisch, aber reflexiv und transitiv. Daher ist im Axiomensystem der Äquivalenzrelationen jedes Axiom von den beiden anderen unabhängig.

Die Modelle vieler Axiomensysteme sind Strukturen im Sinne des im ersten Band dargestellten allgemeinen Begriffes. So sind die eben angeführten Beispiele Strukturen über einer Trägermenge aus drei Elementen mit einer zweistelligen Grundrelation, in denen keine ausgezeichneten Elemente und keine Grundoperationen erklärt sind. Wir nennen zwei Modelle eines Axiomensystems *isomorph*, wenn sie als Strukturen isomorph sind (vgl. MfL, Bd. 1, 2.7.(2)), d. h.; wenn es eine 1-1-Abbildung von der Trägermenge des einen auf die Trägermenge des anderen Modells gibt, bei der sich die Grundelemente, -relationen und -operationen übertragen. Sind je zwei Modelle eines Axiomensystems isomorph, d. h. gibt es „bis auf Isomorphie" nur ein Modell, so heißt das Axiomensystem *kategorisch* (oder *monomorph*). Das Axiomensystem, welches die Äquivalenzrelationen beschreibt, ist nicht kategorisch, denn die Menge $M_1 = \{1, 2, 3\}$ mit der Relation

$$R_1 = \{(1, 1), (1, 2), (2, 1), (2, 2), (3, 3)\}$$

und die Menge $M_2 = \mathbf{Z}$ mit der Relation

$$R_2 = \{(x, y): x \in \mathbf{Z} \wedge y \in \mathbf{Z} \wedge 2 \mid |x - y|\}$$

sind Modelle dieses Systems, aber es gibt keine 1-1-Abbildung von M_1 auf M_2. Auch die Axiomensysteme der algebraischen Strukturen, die im nächsten Abschnitt angegeben werden, sind nicht monomorph.

Ohne Beweis merken wir an, daß das Peanosche Axiomensystem für die natürlichen Zahlen und das von HILBERT angegebene Axiomensystem für die euklidische Geometrie kategorisch sind.

11.3. Algebraische Strukturen

Unter einer *Struktur* oder *allgemeinen Algebra* versteht man ein $(k + m + n + 1)$-Tupel $(M, a_1, \ldots, a_k, R_1, \ldots, R_m, o_1, \ldots, o_n)$, wobei M eine Menge, die Trägermenge der Struktur, $a_1, \ldots, a_k$ ausgezeichnete Elemente aus M, $R_1, \ldots, R_m$ Relationen der Stellenzahlen $i_1, \ldots, i_m$ in M und $o_1, \ldots, o_n$ Operationen der Stellenzahlen $j_1, \ldots, j_n$ in M bezeichnen. Es sind auch die Fälle $k = 0$ und (oder) $m = 0$ und (oder) $n = 0$ zugelassen (vgl. MfL, Bd. 1, 2.6.). Von einer *algebraischen Struktur* spricht man, wenn M nicht die leere Menge bedeutet und wenigstens eine Operation in M erklärt ist. Gegenstand der Algebra ist die Untersuchung solcher algebraischen Strukturen. Dabei sind häufig noch Eigenschaften der Elemente von M, die durch $a_1, \ldots, a_k$,

$R_1, \ldots, R_m$ und $o_1, \ldots, o_n$ ausgedrückt werden können, in Axiomen für die Strukturen fixiert.

Wir wollen hier einige Beispiele für Typen von Strukturen angeben, die in der Algebra von Bedeutung sind. Dabei beschränken wir uns auf die Betrachtung zweistelliger Operationen. In der Entwicklung der Algebra ist aus diesen Strukturtypen durch weitere Verallgemeinerung der obige Strukturbegriff entstanden.

Den einfachsten Fall bilden offenbar die Strukturen $(M, \circ)$, die aus einer nichtleeren Menge M mit einer darin definierten zweistelligen Operation $\circ$ bestehen. Jede solche Struktur heißt *Gruppoid*.

$$(M, \circ) \textit{ heißt Gruppoid} :\Leftrightarrow M \neq \emptyset \wedge \circ \textit{ zweistellige Operation in } M. \tag{1}$$

Häufiger als diesen noch sehr umfassenden Begriff verwendet man denjenigen der *Halbgruppe*. Das ist ein Gruppoid, in welchem das *Assoziativgesetz*

$$(a \circ b) \circ c = a \circ (b \circ c)$$

für beliebige Elemente a, b, c aus M gilt. Hier tritt also erstmalig ein Axiom auf, in dem eine Eigenschaft einer Struktur fixiert ist.

$$(M, \circ) \textit{ heißt Halbgruppe} :\Leftrightarrow (M, \circ) \textit{ ist Gruppoid} \wedge \bigwedge_{a,b,c \in M} (a \circ b) \circ c = a \circ (b \circ c). \tag{2}$$

Beispiele für Halbgruppen sind die Menge der natürlichen Zahlen mit der Addition $(\mathbf{N}, +)$, die Menge der ganzen Zahlen mit der Multiplikation $(\mathbf{Z}, \cdot)$ und die Menge der Abbildungen von einer Menge M in sich mit der Hintereinanderausführung der Abbildungen als Operation.

Ist $(M, \circ)$ ein Gruppoid und $e \in M$, so heißt

$$e \textit{ neutrales Element} \text{ von } (M, \circ) :\Leftrightarrow \bigwedge_{a \in M} a \circ e = e \circ a = a. \tag{3}$$

$$\textit{Jedes Gruppoid } (M, \circ) \textit{ besitzt höchstens ein neutrales Element.} \tag{4}$$

Sind nämlich e und e' neutrale Elemente von $(M, \circ)$, so folgt aus (3)

$$e' = e' \circ e = e \circ e' = e.$$

Die obigen Beispiele haben der Reihe nach die neutralen Elemente 0, 1 und die identische Abbildung, die jedes Element von M auf sich abbildet, während die aus den positiven geraden Zahlen mit der Multiplikation als Operation bestehende Halbgruppe kein neutrales Element besitzt.

Ist $(M, \circ)$ ein Gruppoid mit dem neutralen Element e und sind a und $\bar{a}$ Elemente aus M, so heißt

$$\bar{a} \textit{ inverses Element zu } a :\Leftrightarrow a \circ \bar{a} = \bar{a} \circ a = e. \tag{5}$$

Offenbar ist dann a inverses Element zu $\bar{a}$.

$$\textit{Ist } (M, \circ) \textit{ eine Halbgruppe mit dem neutralen Element } e, \textit{ so besitzt jedes } a \in M \textit{ höchstens ein inverses Element.} \tag{6}$$

Sind nämlich $\bar{a}$ und $\bar{a}'$ inverse Elemente zu a, so ergibt sich aus (2), (3) und (5)

$$\bar{a}' = \bar{a}' \circ e = \bar{a}' \circ (a \circ \bar{a}) = (\bar{a}' \circ a) \circ \bar{a} = e \circ \bar{a} = \bar{a}.$$

In $(\mathbb{N}, +)$ besitzt nur 0 ein inverses Element, nämlich 0; in $(\mathbb{Z}, \cdot)$ besitzen 1 und -1 inverse Elemente, nämlich 1 bzw. -1. In der Menge der Abbildungen von einer Menge M in sich mit der Hintereinanderausführung der Abbildungen als Operation besitzen genau die 1-1-Abbildungen von M auf sich inverse Elemente.

Eine Halbgruppe $(M, \circ)$ wird *Gruppe* genannt, wenn zu je zwei Elementen $a, b \in M$ mindestens ein Element $x \in M$ mit der Eigenschaft $a \circ x = b$ und mindestens ein Element $y \in M$ mit der Eigenschaft $y \circ a = b$ existiert (Ausführbarkeit der links- und rechtsseitigen Division).

$$(M, \circ) \textit{ heißt Gruppe} :\Leftrightarrow (M, \circ) \textit{ ist Halbgruppe} \wedge \bigwedge_{a,b \in M} \left(\bigvee_{x \in M} a \circ x = b \wedge \bigvee_{y \in M} y \circ a = b \right). \tag{7}$$

Beispiele für Gruppen sind die Menge der positiven rationalen Zahlen mit der Multiplikation $(\mathbb{Q}_+^*, \cdot)$, die Menge der komplexen Zahlen vom absoluten Betrag 1 mit der Multiplikation als Operation und die Menge aller Permutationen einer endlichen Menge M mit der Nacheinanderausführung als Operation (vgl. MfL, Bd. 1, 2.4.).

In jeder Gruppe $(M, \circ)$ gibt es ein neutrales Element e. (8)

Zu einem festen Element $a \in M$ gibt es nach (7) in M jedenfalls zwei Elemente e, e' mit

$$a \circ e = a, \quad e' \circ a = a.$$

Mit a und b liegen auch solche Elemente x, y in M, daß

$$a \circ x = b, \quad y \circ a = b$$

gilt. Dann ist

$$e' \circ b = e' \circ (a \circ x) = (e' \circ a) \circ x = a \circ x = b$$

und

$$b \circ e = (y \circ a) \circ e = y \circ (a \circ e) = y \circ a = b.$$

Mit $b = e$ bzw. $b = e'$ ergibt sich daraus $e' \circ e = e$ bzw. $e' \circ e = e'$, also $e = e'$. Daher ist e neutrales Element von $(M, \circ)$.

Ist $(M, \circ)$ Gruppe, so besitzt jedes $a \in M$ ein inverses Element $\bar{a}$. (9)

Ist nämlich e das neutrale Element der Gruppe, so gibt es in M Elemente $\bar{a}$ und $\bar{a}'$, für die

$$a \circ \bar{a} = e, \quad \bar{a}' \circ a = e$$

gilt. Daher ist

$$\bar{a} = e \circ \bar{a} = (\bar{a}' \circ a) \circ \bar{a} = \bar{a}' \circ (a \circ \bar{a}) = \bar{a}' \circ e = \bar{a}'$$

ein inverses Element zu a, das nach (6) eindeutig bestimmt ist.

Ist $(M, \circ)$ eine Gruppe, so gibt es in M zu jedem Elementepaar a, b aus M genau ein x mit $a \circ x = b$ und genau ein y mit $y \circ a = b$. (10)

Die Existenz solcher Elemente x und y folgt aus (7). Bezeichnet e das neutrale Element der Gruppe und $\bar{a}$ das zu a inverse Element, so ist

$$x = e \circ x = (\bar{a} \circ a) \circ x = \bar{a} \circ (a \circ x) = \bar{a} \circ b$$

und

$$y = y \circ e = y \circ (a \circ \bar{a}) = (y \circ a) \circ \bar{a} = b \circ \bar{a},$$

es gibt also höchstens ein Element x und ein Element y in M, die den gegebenen Gleichungen genügen. Man rechnet sofort nach, daß $x = \bar{a} \circ b$ und $y = b \circ \bar{a}$ die geforderte Eigenschaft besitzen, denn es ist

$$a \circ x = a \circ (\bar{a} \circ b) = (a \circ \bar{a}) \circ b = e \circ b = b$$

und

$$y \circ a = (b \circ \bar{a}) \circ a = b \circ (\bar{a} \circ a) = b \circ e = b.$$

Eine Gruppe $(M, \circ)$ heißt *abelsch* (oder *kommutativ*), wenn für alle Elemente $a, b \in M$ das Kommutativgesetz $a \circ b = b \circ a$ gilt.

$$(M, \circ) \textit{ heißt abelsche Gruppe} :\Leftrightarrow (M, \circ) \textit{ ist Gruppe} \wedge \bigwedge_{a,b \in M} a \circ b = b \circ a. \quad (11)$$

Die Gruppe $(\mathbb{Q}_+^*, \cdot)$ ist abelsch, die Gruppe S_3 aller Permutationen einer Menge aus drei Elementen dagegen nicht (vgl. MfL, Bd. 1, 2.4.).

Meistens wird bei der Betrachtung dieser algebraischen Strukturen $(M, \circ)$ mit einer zweistelligen Operation die *multiplikative Schreibweise* verwendet, d. h., sind a und b Elemente von M, so schreibt man statt $a \circ b$ einfach $a \cdot b$ oder ab und nennt ab ohne Rücksicht auf die tatsächliche Bedeutung der vorliegenden Operation das *Produkt* der Elemente a und b. Das neutrale Element e der Gruppe nennt man dann auch das *Einselement* und bezeichnet es manchmal mit 1. Das zu a inverse Element wird mit a^{-1} bezeichnet.

Bei kommutativen Strukturen $(M, \circ)$ wird mitunter auch die sog. *additive Schreibweise* benutzt, d. h., statt $a \circ b$ wird $a + b$ geschrieben und von der *Summe* der Elemente a und b gesprochen. Das neutrale Element der Gruppe bezeichnet man dann häufig mit 0 und nennt es das *Nullelement* der Gruppe. Das zu a inverse Element wird mit $-a$ bezeichnet. Man setzt $b - a := b + (-a)$. Eine additiv geschriebene abelsche Gruppe $(M, +)$ wird *Modul* genannt.

$$(M, \circ) \textit{ heißt Modul} :\Leftrightarrow (M, \circ) \textit{ ist additiv geschriebene abelsche Gruppe.} \quad (12)$$

Die Menge aller ganzen Zahlen mit der Addition als Operation $(\mathbf{Z}, +)$ und die Menge aller Vektoren des dreidimensionalen euklidischen Raumes mit der Vektoraddition als Operation sind Beispiele für Moduln.

Eine algebraische Struktur (M, o_1, o_2) mit zwei binären Operationen wird *nichtassoziativer Ring* genannt, wenn sie hinsichtlich der Operation o_1 eine abelsche Gruppe, hinsichtlich o_2 ein Gruppoid ist und für alle Elemente a, b, c aus M die *Distributivgesetze*

$$a\, o_2\, (b\, o_1\, c) = (a\, o_2\, b)\, o_1\, (a\, o_2\, c)$$

und

$$(b\, o_1\, c)\, o_2\, a = (b\, o_2\, a)\, o_1\, (c\, o_2\, a)$$

gelten. Diese Erklärung schließt nicht aus, daß in (M, o_2) das Assoziativgesetz gilt, d. h., *„nichtassoziativ"* wird hier als Abkürzung für *„nicht notwendig assoziativ"* verwendet.

Es ist im allgemeinen üblich, o_1 durch die additive, o_2 durch die multiplikative Schreibweise auszudrücken. Man nennt dementsprechend $(M, o_1) = (M, +)$ die *additive Gruppe des Ringes* und $(M, o_2) = (M, \cdot)$ das *multiplikative Gruppoid des Ringes*. Für Elemente a, b, c aus M sei $ab + c := (ab) + c$, wie wir es vom Zahlenrechnen kennen.

$$(M, +, \cdot) \textit{ heißt nichtassoziativer Ring} :\Leftrightarrow (M, +) \textit{ ist Modul} \wedge (M, \cdot) \textit{ ist Gruppoid} \wedge \bigwedge_{a,b,c \in M} a(b + c) = ab + ac \wedge \bigwedge_{a,b,c \in M} (b + c)\, a = ba + ca. \tag{13}$$

Die Menge aller Vektoren des dreidimensionalen euklidischen Raumes mit der gewöhnlichen Vektoraddition und der vektoriellen Multiplikation als Operation ist ein Beispiel für einen nichtassoziativen Ring.

$$(M, +, \cdot) \textit{ nichtassoziativer Ring} \Rightarrow \bigwedge_{a,b,c \in M} a(b - c) = ab - ac$$
$$\wedge \bigwedge_{a,b,c \in M} (b - c)\, a = ba - ca. \tag{14}$$

Da $(M, +)$ Gruppe ist, gilt $(b - c) + c = b$, woraus nach Multiplikation mit a von links unter Ausnutzung des ersten Distributivgesetzes $a(b - c) + ac = ab$ folgt. Daraus ergibt sich

$$a(b - c) = ab - ac,$$

weil $(M, +)$ Gruppe ist. Die zweite Aussage kann analog bewiesen werden.

$$(M, +, \cdot) \textit{ nichtassoziativer Ring} \Rightarrow \bigwedge_{a \in M} a0 = 0a = 0. \tag{15}$$

Ist nämlich $x \in M$, so gilt nach (14)

$$a0 = a(x - x) = ax - ax = 0.$$

$$(M, +, \cdot) \textit{ nichtassoziativer Ring} \Rightarrow \bigwedge_{a,b \in M} (-a)\, b = a(-b) = -ab$$

$$\wedge \bigwedge_{a,b \in M} (-a)(-b) = ab. \tag{16}$$

Denn aus

$$0 = 0b = [a + (-a)]\, b = ab + (-a)\, b$$

ergibt sich $(-a)\, b = -ab$, da $(M, +)$ Gruppe ist. Ebenso beweist man $a(-b) = -ab$. Aus beiden Aussagen erhält man

$$(-a)\,(-b) = -[a(-b)] = -(-ab) = ab.$$

Ist das multiplikative Gruppoid eines nichtassoziativen Ringes $(M, +, \cdot)$ eine Halbgruppe, so heißt $(M, +, \cdot)$ *assoziativer Ring* oder einfach *Ring*.

$$(M, +, \cdot) \textit{ heißt (assoziativer) Ring} :\Leftrightarrow (M, +, \cdot) \textit{ ist nichtassoziativer Ring}$$

$$\wedge \bigwedge_{a,b,c \in M} a(bc) = (ab)c. \tag{17}$$

Beispiel. Die Menge aller n-reihigen quadratischen Matrizen mit reellen Elementen ($n \geqq 2$) mit der üblichen Matrizenaddition und Multiplikation als Operationen ist ein assoziativer Ring.

Ist die multiplikative Halbgruppe $(M, \cdot)$ des Ringes $(M, +, \cdot)$ kommutativ, so nennt man den Ring *kommutativ*.

$$(M, +, \cdot) \textit{ heißt kommutativer Ring} :\Leftrightarrow (M, +, \cdot) \textit{ ist Ring} \wedge \bigwedge_{a,b \in M} ab = ba. \tag{18}$$

Die Menge $\mathbf{Z}$ aller ganzen Zahlen bildet mit der Addition und Multiplikation als Operationen einen kommutativen Ring.

Ist $(M, +, \cdot)$ ein nichtassoziativer Ring, dessen Trägermenge nicht nur aus dem Nullelement 0 von $(M, +)$ besteht, so kann das multiplikative Gruppoid $(M, \cdot)$ keine Gruppe sein, denn wegen (15) kann 0 nicht neutrales Element von $(M, \cdot)$ sein und deshalb wegen (15) kein inverses Element in $(M, \cdot)$ besitzen. Bildet aber die Menge $M \setminus \{0\}$ mit der Multiplikation des Ringes eine Gruppe, so heißt der (dann notwendig assoziative) Ring $(M, +, \cdot)$ *Schiefkörper* und $(M \setminus \{0\}, \cdot)$ die *multiplikative Gruppe des Schiefkörpers*.

$$(M, +, \cdot) \textit{ heißt Schiefkörper} :\Leftrightarrow (M, +, \cdot) \textit{ ist Ring} \wedge (M \setminus \{0\}, \cdot) \textit{ ist Gruppe}. \tag{19}$$

$$(M, +, \cdot) \textit{ heißt Körper} :\Leftrightarrow (M, +, \cdot) \textit{ ist kommutativer Ring} \wedge (M \setminus \{0\}, \cdot) \textit{ ist Gruppe}. \tag{20}$$

Die Menge $\mathbf{Q}$ der rationalen Zahlen bildet ebenso wie die Menge der reellen Zahlen $\mathbf{R}$ und der komplexen Zahlen $\mathbf{C}$ mit der üblichen Addition und Multiplikation als Operationen einen Körper.

Vom Begriff des nichtassoziativen Ringes ausgehend kann man noch andere Strukturtypen erhalten, die für Anwendungen von Bedeutung sind. Beispielsweise heißt ein nichtassoziativer Ring $(M, +, \cdot)$, in dem für beliebige Elemente $a, b, c \in M$

$$aa = 0$$

und

$$(ab)\,c + (bc)\,a + (ca)\,b = 0$$

gilt, ein *Liescher Ring* (nach dem norwegischen Mathematiker Sophus Lie (1842—1899)).

Im Anschluß an (13) haben wir ein Beispiel für einen nichtassoziativen Ring angegeben, der auch ein Liescher Ring ist.

Weitere Beispiele für Liesche Ringe kann man sich leicht herstellen: Ist $(M, +, \cdot)$ ein assoziativer Ring und bezeichnen a, b beliebige Elemente aus M, so ist $a \circ b := ab - ba$ eine zweistellige Operation in M und $(M, +, \circ)$ ein Liescher Ring.

Wir zeigen zunächst die Gültigkeit eines Distributivgesetzes:

$$\begin{aligned} a \circ (b + c) &= a(b + c) - (b + c)a = ab + ac - (ba + ca) \\ &= ab + ac - ba - ca = (ab - ba) + (ac - ca) \\ &= a \circ b + a \circ c. \end{aligned}$$

Analog kann nachgewiesen werden, daß das zweite Distributivgesetz gilt. $(M, +, \circ)$ ist also ein nichtassoziativer Ring.

Für ein beliebiges Element $a \in M$ gilt $a \circ a = aa - aa = 0$, und für $a, b, c \in M$ ist

$$\begin{aligned} &(a \circ b) \circ c + (b \circ c) \circ a + (c \circ a) \circ b \\ &= (ab - ba)\,c - c(ab - ba) + (bc - cb)\,a - a(bc - cb) \\ &\quad + (ca - ac)\,b - b(ca - ac) = 0. \end{aligned}$$

Daher ist $(M, +, \circ)$ ein Liescher Ring.

Eine algebraische Struktur (M, o_1, o_2) mit zwei assoziativen und kommutativen binären Operationen heißt *Verband*, wenn außerdem noch die beiden *Absorptionsgesetze* (auch *Verschmelzungsgesetze* genannt)

$$a\,o_1\,(a\,o_2\,b) = a \quad \text{und} \quad a\,o_2\,(a\,o_1\,b) = a$$

für alle Elemente a und b aus M gelten.

Die Bildung des Durchschnitts $\cap$ und der Vereinigung $\cup$ je zweier Teilmengen einer gegebenen Menge M sind zwei Operationen in der Potenzmenge $\mathfrak{P}(M)$, die alle diese Bedingungen erfüllen. $(\mathfrak{P}(M), \cap, \cup)$ heißt *Verband aller Teilmengen* von M oder *voller Mengenverband*. In Anlehnung an dieses Beispiel verwenden wir für die Operationen in einem Verband die Bezeichnungen $\wedge$ und $\vee$.

$(M, \wedge, \vee)$ heißt Verband $:\Leftrightarrow M \neq \emptyset \wedge \wedge, \vee$ *zweistellige Operationen in* M

$$\wedge \bigwedge_{a,b,c \in M} a \wedge (b \wedge c) = (a \wedge b) \wedge c \wedge \bigwedge_{a,b,c \in M} a \vee (b \vee c) = (a \vee b) \vee c \tag{21.1}$$

$$\wedge \bigwedge_{a,b \in M} a \wedge b = b \wedge a \wedge \bigwedge_{a,b \in M} a \vee b = b \vee a \tag{21.2}$$

$$\wedge \bigwedge_{a,b \in M} a \wedge (a \vee b) = a \wedge \bigwedge_{a,b \in M} a \vee (a \wedge b) = a. \tag{21.3}$$

Die Menge $\mathbf{N}^*$ der natürlichen Zahlen $\neq 0$ mit der Bildung des größten gemeinsamen Teilers und des kleinsten gemeinsamen Vielfachen ist ein weiteres Beispiel für einen Verband $(\mathbf{N}^*, \sqcap, \sqcup)$.

Ist $(M, \wedge, \vee)$ ein Verband, so gelten für alle Elemente a, b aus M die Regeln

$$a \wedge a = a, \quad a \vee a = a \tag{22}$$

und

$$a \wedge b = a \Leftrightarrow a \vee b = b. \tag{23}$$

Denn nach (21.3) ist $a \vee (a \wedge a) = a$ und $a \wedge \big(a \vee (a \wedge a)\big) = a$, also $a \wedge a = a$. Analog folgt aus $a \wedge (a \vee a) = a$ und $a \vee \big(a \wedge (a \vee a)\big) = a$, daß $a \vee a = a$. Ferner ist $(a \wedge b) \vee b = b$. Daher folgt aus $a \wedge b = a$, daß $a \vee b = b$ ist. Analog ergibt sich wegen $a \wedge (a \vee b) = a$ aus $a \vee b = b$, daß $a \wedge b = a$ ist.

Es sei $M \neq \emptyset$ eine Menge und $\leqq$ bezeichne eine darin erklärte Ordnungsrelation, d. h., $\leqq$ ist eine transitive, reflexive und antisymmetrische binäre Relation in M (vgl. MfL, Bd. 1, 2.5.). Zu zwei Elementen a und b aus M definieren wir eine *größte untere Schranke* (*Infimum*, inf (a, b)) und eine *kleinste obere Schranke* (*Supremum*, sup (a, b)) durch folgende Eigenschaften:

$$x = \inf(a, b) :\Leftrightarrow x \in M \wedge x \leqq a \wedge x \leqq b \wedge \bigwedge_{\bar{x} \in M} (\bar{x} \leqq a \wedge \bar{x} \leqq b \Rightarrow \bar{x} \leqq x), \tag{24}$$

$$y = \sup(a, b) :\Leftrightarrow y \in M \wedge a \leqq y \wedge b \leqq y \wedge \bigwedge_{\bar{y} \in M} (a \leqq \bar{y} \wedge b \leqq \bar{y} \Rightarrow y \leqq \bar{y}). \tag{25}$$

Ist $x = \inf(a, b)$ und $x_1 = \inf(a, b)$, so folgt aus (24) $x_1 \leqq x$ und $x \leqq x_1$, wegen der Antisymmetrie von $\leqq$ gilt also $x = x_1$.

Ebenso zeigt man, daß es zu zwei Elementen a und b aus M höchstens eine kleinste obere Schranke gibt. Es brauchen aber sup (a, b) und inf (a, b) nicht immer zu existieren. Beispielsweise ist in der Menge $M = \{a, b, c, d, e\}$ durch die Relation

$$R = \{(a, a), (a, b), (a, c), (a, d), (a, e), (b, b), (b, d), (b, e), (c, c), (c, d), (c, e), (d, d), (e, e)\}$$

eine Ordnung erklärt $\big(x \leqq y :\Leftrightarrow (x, y) \in R\big)$, und es gilt inf $(b, c) = a$, während sup (b, c) nicht existiert.

In jedem Verband $(M, \wedge, \vee)$ *wird durch die Festlegung*

$$a \leqq b :\Leftrightarrow a \wedge b = a \qquad (a, b \in M) \tag{26.1}$$

eine Ordnungsrelation definiert, in der alle $\inf(a, b)$ *und* $\sup(a, b)$ *existieren. Es ist*

$$\inf(a, b) = a \wedge b \quad \textit{und} \quad \sup(a, b) = a \vee b. \tag{26.2}$$

Zum Beweis zeigen wir, daß durch (26.1) eine Ordnungsrelation in M definiert wird. Aus (22) ergibt sich $a \leqq a$ für alle $a \in M$. Da sich wegen (21.1) aus $a \wedge b = a$ und $b \wedge c = b$

$$a \wedge c = (a \wedge b) \wedge c = a \wedge (b \wedge c) = a \wedge b = a$$

ergibt, folgt aus $a \leqq b$ und $b \leqq c$ die Beziehung $a \leqq c$. Schließlich folgt aus $a \wedge b = a$ und $b \wedge a = b$ wegen (21.2) $a = b$, d. h., ist $a \leqq b$ und $b \leqq a$, so $a = b$.

Für beliebige Elemente a und b aus M gilt wegen (21) und (22)

$$(a \wedge b) \wedge a = a \wedge (b \wedge a) = a \wedge (a \wedge b) = (a \wedge a) \wedge b = a \wedge b,$$

d. h. wegen (26.1) $a \wedge b \leqq a$ und $(a \wedge b) \wedge b = a \wedge (b \wedge b) = a \wedge b$, d. h. $a \wedge b \leqq b$. Ist $x \in M$ und gilt $x \leqq a$ und $x \leqq b$, d. h. $x \wedge a = x$ und $x \wedge b = x$, so ist $x \wedge (a \wedge b) = (x \wedge a) \wedge b = x \wedge b = x$, was mit $x \leqq a \wedge b$ gleichbedeutend ist. Also ist $a \wedge b = \inf(a, b)$. Wir überlassen es dem Leser als Übungsaufgabe, unter Verwendung von (23) zu zeigen, daß $a \vee b = \sup(a, b)$ ist.

Umgekehrt gilt:

Ist die Menge $M \neq \emptyset$ *und* $\leqq$ *eine Ordnungsrelation in* M *derart, daß zu allen Elementen* $a, b \in M$ $\inf(a, b)$ *und* $\sup(a, b)$ *existieren, dann lassen sich in* M *auf genau eine Weise zwei Operationen* $\wedge, \vee$ *so definieren, daß* $(M, \wedge, \vee)$ *ein Verband ist und* $a \leqq b \Leftrightarrow a \wedge b = a$ $(a, b \in M)$ *gilt.* (27)

Wir gehen von einer geordneten Menge $(M, \leqq)$ aus, in der alle $\inf(a, b)$ und $\sup(a, b)$ $(a, b \in M)$ existieren und definieren

$$a \wedge b := \inf(a, b), \quad a \vee b := \sup(a, b).$$

Wegen $\inf(a, b) = \inf(b, a)$ und $\sup(a, b) = \sup(b, a)$ ist $a \wedge b = b \wedge a$ und $a \vee b = b \vee a$. Für beliebige Elemente a, b, c aus M gilt

$$\inf\big(\inf(a, b), c\big) = \inf\big(a, \inf(b, c)\big)$$

bzw.

$$\sup\big(\sup(a, b), c\big) = \sup\big(a, \sup(b, c)\big)$$

(Übungsaufgabe). Aus diesen Gleichungen folgt

$$(a \wedge b) \wedge c = a \wedge (b \wedge c) \quad \text{bzw.} \quad (a \vee b) \vee c = a \vee (b \vee c).$$

Wegen $\inf(a, \sup(a, b)) \leqq a$ ist $a \wedge (a \vee b) \leqq a$. Andererseits ist $a \leqq a$ und $a \leqq \sup(a, b)$, also $a \leqq \inf(a, \sup(a, b))$. Daher gilt $\inf(a, \sup(a, b)) = a$, d. h. $a \wedge (a \vee b) = a$.

Entsprechend weist man auch die Gültigkeit des zweiten Absorptionsgesetzes nach. Damit ist gezeigt, daß $(M, \wedge, \vee)$ ein Verband ist. In diesem gilt

$$a \wedge b = \inf(a, b) = a \Rightarrow a \leqq b$$

und

$$a \leqq b \Rightarrow a \leqq \inf(a, b) = a \wedge b \leqq a \Rightarrow a \wedge b = a.$$

Da wir bereits gezeigt haben, daß (26.2) aus (26.1) folgt, gibt es auch nur eine Möglichkeit, in $(M, \leqq)$ Operationen $\wedge$ und $\vee$ so zu definieren, daß $(M, \wedge, \vee)$ ein Verband ist, in dem (26.1) gilt.

In jedem Verband $(M, \wedge, \vee)$ ist durch eine Operation die andere eindeutig festgelegt. (28)

Denn nach (26.1) geht nur die Operation $\wedge$ in die Definition der Ordnungsrelation $\leqq$ ein, durch die dann nach (27) die Operation $\vee$ bestimmt ist. Mittels (23) erkennt man, daß auch durch die Operation $\vee$ die Ordnungsrelation $\leqq$ erklärt werden kann, durch die dann die Operation $\wedge$ festgelegt ist.

(26) und (27) besagen, daß die Verbände und diejenigen geordneten Mengen, in denen alle $\inf(a, b)$ und $\sup(a, b)$ existieren, identische Begriffe sind. Der Zusammenhang wird durch (26.1) vermittelt.

Die wichtigsten der vorgestellten Strukturen sind die Gruppen, Ringe und Körper. Entsprechend sind Gruppentheorie, Ringtheorie und Körpertheorie wichtige Teilgebiete der Algebra. Wir werden uns in den nächsten Abschnitten mit einer Einführung in diese Theorien beschäftigen. Dabei werden wir im allgemeinen die Trägermenge einer Struktur und die Struktur mit dem gleichen Buchstaben bezeichnen. Wir werden also beispielsweise von den Elementen einer Gruppe G oder eines Körpers K sprechen.

11.4. Übungsaufgaben

1. Man zeige, daß die angegebenen Axiome, welche ein Gruppoid $(M, \circ)$ erfüllen muß, um eine Gruppe zu sein, voneinander unabhängig sind.
   Dazu gebe man in einer dreielementigen Menge $M = \{a, b, c\}$ solche zweistelligen Operationen $\circ$ (z. B. in Form einer „Multiplikationstabelle" (vgl. 12.3.)) an, daß $(M, \circ)$ jeweils genau einem dieser Axiome nicht genügt.
   Wie viele Möglichkeiten gibt es, in $M = \{a, b, c\}$ eine zweistellige Operation $\circ$ so zu erklären, daß $(M, \circ)$ eine Gruppe ist?

2. Die Menge aller zweireihigen quadratischen Matrizen mit Elementen aus $\mathbf{N}$ bildet bezüglich der Matrizenmultiplikation eine Halbgruppe. Man zeige, daß es darin Matrizen $\boldsymbol{A}$ und $\boldsymbol{B}$ gibt, für die $\boldsymbol{AX} = \boldsymbol{B}$ unendlich viele Lösungen, $\boldsymbol{YA} = \boldsymbol{B}$ dagegen keine Lösung besitzt.

3. In der Menge M aller linearen Polynome $f(x) = ax + b$ mit Koeffizienten a $(+ 0)$, b aus $\mathbf{Q}$ wird durch $f(x) \circ g(x) := f(g(x))$ eine zweistellige Operation erklärt. Man beweise, daß $(M, \circ)$ eine nichtkommutative Gruppe ist, gebe ihr neutrales Element an und bestimme das zu $f(x)$ inverse Element.

4. In einer dreielementigen Menge $M = \{a, b, c\}$ gebe man solche zweistelligen Operationen $+, \cdot$ (z. B. durch Tabellen) an, daß $(M, +, \cdot)$ ein Ring ist.
Wie viele verschiedene Möglichkeiten gibt es?

5. Gilt in dem Ring $(M, +, \cdot)$ neben den Ringaxiomen auch noch

$$\bigvee_{e \in M} \bigwedge_{a \in M} ae = ea = a,$$

so heißt $(M, +, \cdot)$ *Ring mit Einselement* e. Im Axiomensystem für einen solchen Ring ist die Kommutativität der Addition eine Folgerung aus den übrigen Axiomen (sogar auch schon ohne das Axiom der Assoziativität der Multiplikation).

6. Wie viele Ordnungsrelationen kann man in der Menge $M = \{a, b, c\}$ so erklären, daß zu je zwei Elementen aus M das Infimum und das Supremum existieren?
In einem Fall gebe man die durch eine solche Ordnungsrelation festgelegten Operationen $\wedge$ und $\vee$, mit denen $(M, \wedge, \vee)$ nach 11.3. (27) ein Verband ist, in Tabellenform an.

12. Gruppen

12.1. Gruppenaxiome, Beispiele

12.1.1. Wir erinnern zunächst an die bereits im vorigen Abschnitt gegebene

Definition 1. Eine nichtleere Menge G mit einer (hier multiplikativ geschriebenen) zweistelligen Operation heißt *Gruppe*, wenn folgende Axiome erfüllt sind:

$$\bigwedge_{a,b,c\in G} (ab)c = a(bc) \quad \text{(Assoziativgesetz)}, \tag{1}$$

$$\bigwedge_{a,b\in G} \left(\bigvee_{x\in G} ax = b \wedge \bigvee_{y\in G} ya = b \right) \tag{2}$$

(Ausführbarkeit der links- und rechtsseitigen Division).

Diese Elemente x, y sind durch a und b eindeutig bestimmt (vgl. 11.3.(10)).

Die Gruppen können auch durch andere Aussagen charakterisiert werden:

Satz 1. *Eine nichtleere Menge G mit einer zweistelligen Operation ist dann und nur dann eine Gruppe, wenn folgendes gilt:*

$$\bigwedge_{a,b,c\in G} (ab)\,c = a(bc). \tag{1}$$

In G gibt es genau ein neutrales Element e mit der Eigenschaft

$$\bigwedge_{a\in G} ae = ea = a. \tag{3}$$

Zu jedem $a \in G$ gibt es in G genau ein inverses Element a^{-1} mit der Eigenschaft

$$aa^{-1} = a^{-1}a = e. \tag{4}$$

Beweis. Hat die Struktur G die Eigenschaften (1), (3), (4), so sind für beliebige Elemente a, b aus G auch $x = a^{-1}b$ sowie $y = ba^{-1}$ aus G, und es gilt

$$ax = a(a^{-1}b) = (aa^{-1})\,b = eb = b,$$
$$ya = (ba^{-1})\,a = b(a^{-1}a) = be = b.$$

Also sind (1) und (2) erfüllt, und G ist eine Gruppe.

Sind umgekehrt (1) und (2) erfüllt, so gilt (3) (vgl. 11.3.(3), (4), (8)) und (4) (vgl. 11.3.(5), (6), (9)).

In G sind zunächst nur Produkte aus zwei Elementen erklärt. Produkte aus mehr als zwei Elementen können durch mehrmalige Multiplikation von je zwei Elementen gebildet werden. Beispielsweise kann das Produkt der Elemente a, b, c einerseits bestimmt werden, indem zunächst $ab = p$ berechnet wird und dann $pc = (ab)\,c$. Andererseits kann man aber auch zuerst $bc = q$ und dann $aq = a(bc)$ bilden. Das Assoziativgesetz (1) besagt, daß sich in beiden Fällen das gleiche Ergebnis einstellt. Daher läßt man oft die Klammern weg, schreibt also $abc = (ab)c = a(bc)$ und spricht vom Produkt der drei Elemente a, b, c in der gegebenen Reihenfolge. Das Produkt aus n ($n \in \mathbb{N} \wedge n \geqq 3$) Gruppenelementen $a_1, \ldots, a_n$ ist erklärt, wenn es durch Beklammerung auf die Nacheinanderausführung von $n-1$ Produkten aus je zwei Gruppenelementen zurückgeführt wurde. Wir werden zeigen, daß jedes so gebildete Produkt bei fester Reihenfolge der Faktoren das gleiche Resultat liefert. Es kommt also auch bei einem Produkt aus n Faktoren nicht auf die Beklammerung an, und man schreibt daher einfach $a_1 a_2 \cdots a_n$.

Diese Aussage kann durch vollständige Induktion nach n bewiesen werden. Für $n = 3$ folgt die Richtigkeit aus (1). Wir nehmen an, die Behauptung sei für Produkte aus weniger als n Faktoren richtig. Für jede natürliche Zahl k mit $1 \leqq k < n$ sind also die Produkte $a_1 \cdots a_k = p_1$ und $a_{k+1} \cdots a_n = p_2$ bereits eindeutig durch die Angabe der Faktoren und ihrer Reihenfolge bestimmt. Das Gesamtprodukt der $a_1 \cdots a_n$ kann gebildet werden, indem man $(a_1 \cdots a_k)\,(a_{k+1} \cdots a_n) = p_1 p_2$ berechnet. Ist l eine andere natürliche Zahl mit $1 \leqq l < n$ und sind $q_1 = a_1 \cdots a_l$, $q_2 = a_{l+1} \cdots a_n$, so kann das Gesamtprodukt auch durch $(a_1 \cdots a_l)\,(a_{l+1} \cdots a_n) = q_1 q_2$ bestimmt werden. Zu zeigen ist $p_1 p_2 = q_1 q_2$. Sei $k < l$. Nach der Induktionsannahme ist $r = a_{k+1} \cdots a_l$ allein durch die Angabe der Faktoren und ihrer Reihenfolge festgelegt, und es gilt

$$p_1 r = q_1 \qquad \text{und} \qquad r q_2 = p_2 .$$

Aus (1) folgt dann

$$p_1 p_2 = p_1 (r q_2) = (p_1 r)\, q_2 = q_1 q_2 .$$

Für ein Produkt aus n Faktoren a ($n \in \mathbb{N}^*$) kann man wegen der Gültigkeit des assoziativen Gesetzes die Schreibweise

$$a^n := \underbrace{a a \cdots a}_{n \text{ Faktoren}}$$

einführen. Für die so erklärte n-te Potenz von a gelten die Regeln

$$a^m a^n = a^{m+n} = a^n a^m , \qquad (a^m)^n = a^{mn} = (a^n)^m . \tag{5}$$

Definiert man noch

$$a^0 := e , \qquad a^{-n} := (a^{-1})^n ,$$

so bestätigt man unter Verwendung von (4) leicht, daß die Regeln (5) für beliebige Exponenten $m, n \in \mathbf{Z}$ gelten.

In *abelschen Gruppen* (vgl. 11.3.(11)) ist für $n \in \mathbf{Z}$ und beliebige Elemente a, b außerdem

$$(ab)^n = a^n b^n. \tag{6}$$

In einer additiv geschriebenen Gruppe treten an die Stelle der Potenzen a^n die Vielfachen na. Die Regeln (5) heißen dann

$$ma + na = (m + n)\, a = na + ma, \qquad n(ma) = (nm)\, a = m(na).$$

Da $ab(b^{-1}a^{-1}) = a(bb^{-1})\, a^{-1} = aea^{-1} = e$ und das inverse Element eindeutig bestimmt ist, gilt

$$(ab)^{-1} = b^{-1}a^{-1}. \tag{7}$$

Definition 2. Als *Ordnung* $|G|$ einer Gruppe G bezeichnet man die Anzahl ihrer Elemente, wenn diese endlich ist, G wird dann *endliche Gruppe* genannt. Sonst heißt G *Gruppe unendlicher Ordnung* oder *unendliche Gruppe*.

12.1.2. Beispiele. Wollen wir von einer Menge $M \neq \emptyset$ und einer Korrespondenz, die geordneten Paaren von Elementen aus M gewisse Bilder zuordnet, nachweisen, daß durch sie eine Gruppe gegeben ist, so haben wir zu zeigen:

1. Diese Korrespondenz ist eine *Operation* in M (d. h. eine *eindeutige* Zuordnung *von* $M \times M$ *in* M).
2. Die Axiome (1) und (2) sind erfüllt. (Nach Satz 1 kann statt dessen auch die Gültigkeit von (1), (3) und (4) bewiesen werden.)

Bei den folgenden Beispielen werden wir auf ausführliche Nachweise häufig verzichten, da der Leser sie leicht ergänzen kann.

12.1.2.1. $D = \{a + b\sqrt{3} : a \in \mathbf{Z} \wedge b \in \mathbf{Z} \wedge a^2 - 3b^2 = 1\}$ bildet mit der üblichen Multiplikation eine unendliche abelsche Gruppe.

Aus $a^2 - 3b^2 = c^2 - 3d^2 = 1$ folgt nämlich, daß

$$(a + b\sqrt{3})\,(c + d\sqrt{3}) = (ac + 3bd) + (ad + bc)\sqrt{3}$$

Element von D ist, da folgendes gilt:

$$(ac + 3bd)^2 - 3(ad + bc)^2 = (a^2 - 3b^2)\,(c^2 - 3d^2) = 1.$$

Das Assoziativgesetz gilt bekanntlich in $\mathbf{R}$. $1 + 0\sqrt{3} = 1$ ist neutrales Element. In der Struktur gibt es kein weiteres neutrales Element (vgl. 11.3.(4)).

Zu $a + b\sqrt{3}$ ist $a - b\sqrt{3}$ invers, weil

$$\left(a + b\sqrt{3}\right)\left(a - b\sqrt{3}\right) = a^2 - 3b^2 = 1.$$

Wegen 11.3.(6) ist es das einzige Element mit dieser Eigenschaft.

Nach Satz 1 liegt also eine Gruppe vor.

12.1.2.2. Ist α eine feste Zahl aus $\mathbb{R}\setminus\{0, 1, -1\}$, dann bildet die Menge $Z_\infty = \{\alpha^k : k \in \mathbb{Z}\}$ bezüglich der üblichen Multiplikation eine unendliche Gruppe mit dem neutralen Element $\alpha^0 = 1$ und dem zu α^k inversen Element α^{-k}.

Eine solche Gruppe, deren Elemente die Potenzen eines einzigen Elementes sind, heißt *zyklisch*. Nach (5) ist jede zyklische Gruppe abelsch.

12.1.2.3. $(\mathbb{Z}, +)$ ist eine additiv geschriebene unendliche zyklische Gruppe, da sämtliche Elemente Vielfache von 1 sind. 0 ist das neutrale Element und zur ganzen Zahl a ist $-a$ invers.

$$(\mathbb{Q}, +),\ (\mathbb{R}, +),\ (\mathbb{C}, +),\ (\mathbb{Q}\setminus\{0\}, \cdot),\ (\mathbb{R}\setminus\{0\}, \cdot),\ (\mathbb{C}\setminus\{0\}, \cdot)$$

sind weitere Beispiele für unendliche abelsche Gruppen in additiver und multiplikativer Schreibweise.

12.1.2.4. Die Menge L_n der quadratischen n-reihigen Matrizen aus rationalen Zahlen mit von 0 verschiedener Determinante bildet bezüglich der Matrizenmultiplikation eine unendliche Gruppe, die im Fall $n > 1$ nicht abelsch ist, während für $n = 1$ $(\mathbb{Q}\setminus\{0\}, \cdot)$ vorliegt.

Das neutrale Element ist die Einheitsmatrix $\boldsymbol{I}$; zur Matrix $\boldsymbol{A}$ invers ist die aus der linearen Algebra bekannte inverse Matrix $\boldsymbol{A}^{-1}$.

12.1.2.5. Es sei M eine Menge und $A \subseteq M$, $B \subseteq M$. Bezüglich der durch

$$A \circ B := (A \cup B)\setminus(A \cap B)$$

definierten Operation bildet die Potenzmenge $\mathfrak{P}(M)$ eine abelsche Gruppe. $A \circ B$ besteht aus denjenigen Elementen von M, die in genau einer der Mengen A, B liegen. $(A \circ B) \circ C$ besteht dann aus denjenigen Elementen von M, die entweder in genau einer der Mengen A, B, C liegen oder in allen drei Mengen enthalten sind. Aus denselben Elementen besteht aber die Menge $A \circ (B \circ C)$.

Neutrales Element ist die leere Menge $\emptyset$.

Da $A \circ A = \emptyset$, ist jedes $A \in \mathfrak{P}(M)$ zu sich selbst invers.

Wegen der Kommutativität von $\cup$ und $\cap$ ist $A \circ B = B \circ A$.

Wenn M eine endliche Menge aus $|M|$ Elementen ist, hat die Gruppe $(\mathfrak{P}(M), \circ)$ die Ordnung $2^{|M|}$, sonst ist sie unendlich.

12.1.2.6. Es bezeichne $\mathfrak{T}(M)$ die Menge aller 1-1-Abbildungen (Permutationen) einer Menge M auf sich. Mit der Nacheinanderausführung als Operation bildet $\mathfrak{T}(M)$ eine Gruppe (vgl. MfL, Bd. 1, 2.4.).

Das neutrale Element e ist die identische Abbildung. Das zu $f \in \mathfrak{T}(M)$ inverse Gruppenelement ist die inverse Abbildung f^{-1}. Ist M eine unendliche Menge, so bildet $\mathfrak{T}(M)$ eine unendliche Gruppe, ist aber M eine endliche Menge aus n Elementen, so bildet $\mathfrak{T}(M)$ eine Gruppe der Ordnung $n!$ (vgl. MfL, Bd. 1, 3.6.), die mit S_n bezeichnet und *symmetrische Gruppe des Grades n* genannt wird.

Bildet die Permutation f der Reihe nach die Elemente $a, b, c, \ldots \in M$ auf $f(a), f(b), f(c), \ldots \in M$ ab, so schreibt man in übersichtlicher Weise

$$f = \begin{pmatrix} a & b & c & \ldots \\ f(a) & f(b) & f(c) & \ldots \end{pmatrix}.$$

Dabei steht in jeder der beiden Zeilen jedes Element aus M genau einmal. Zwei solche Symbole stellen genau dann die gleiche Permutation dar, wenn sie durch Vertauschung der Spalten auseinander hervorgehen.

Ist $M = \{1, 2, 3\}$, so sind

$$e = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad p = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad q = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix},$$

$$r = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad s = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \quad t = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

alle Permutationen aus S_3.

Das Resultat $f \circ g$ der Nacheinanderausführung (f nach g) der Permutationen

$$f = \begin{pmatrix} a & b & c & \ldots \\ f(a) & f(b) & f(c) & \ldots \end{pmatrix}$$

und

$$g = \begin{pmatrix} a & b & c & \ldots \\ g(a) & g(b) & g(c) & \ldots \end{pmatrix}$$

schreiben wir als Produkt

$$gf := f \circ g,$$

wobei der linke Faktor die zuerst auszuführende Permutation g angibt. Da man nach einer Spaltenvertauschung die Permutation f in der Form

$$f = \begin{pmatrix} g(a) & g(b) & g(c) & \ldots \\ f(g(a)) & f(g(b)) & f(g(c)) & \ldots \end{pmatrix}$$

schreiben kann, ist

$$f \circ g = gf = \begin{pmatrix} a & b & c & \ldots \\ g(a) & g(b) & g(c) & \ldots \end{pmatrix} \begin{pmatrix} g(a) & g(b) & g(c) & \ldots \\ f(g(a)) & f(g(b)) & f(g(c)) & \ldots \end{pmatrix}$$

$$= \begin{pmatrix} a & b & c & \ldots \\ f(g(a)) & f(g(b)) & f(g(c)) & \ldots \end{pmatrix}.$$

Beispielsweise ist in der S_3

$$pr = \begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix} = \begin{pmatrix}1&2&3\\3&2&1\end{pmatrix}$$

(gelesen: „1 bei p in 2 und 2 bei r in 3, also 1 bei pr in 3" usw.),

$$rp = \begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix} = \begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}.$$

Man muß also auf die Reihenfolge der Faktoren achten. Die S_3 ist nicht abelsch.

12.1.2.7. Es sei P folgende Menge von Polynomquotienten in x:

$$P = \left\{x, \frac{1}{x}, 1-x, \frac{x-1}{x}, \frac{1}{1-x}, \frac{x}{x-1}\right\}.$$

Für beliebige Elemente $f = f(x)$ und $g = g(x)$ aus P definieren wir

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

Ist etwa $f = 1 - x$ und $g = \dfrac{x}{x-1}$, so ergibt diese *Substitution* von g in f

$$f \circ g = 1 - \frac{x}{x-1} = \frac{1}{1-x}$$

wieder ein Element aus P. Man kann in endlich vielen Schritten nachprüfen, daß durch diese Festlegung jedem geordneten Paar (f, g) von Elementen aus P eindeutig ein Element aus P zugeordnet ist.

Da nur endlich viele Elementtripel existieren, kann man durch Berechnung aller möglichen Fälle die Gültigkeit des Assoziativgesetzes für die Operation $\circ$ nachweisen. Man kann sich den Sachverhalt aber auch durch die folgende Überlegung klarmachen. $f \circ (g \circ h)$ bedeutet, daß zunächst $h(x)$ statt x in $g(x)$ einzusetzen und dann das Ergebnis in $f(x)$ an die Stelle von x zu schreiben ist:

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

Das gleiche Resultat entsteht, wenn zunächst x in $f(x)$ durch $g(x)$ ersetzt und anschließend statt x im Ergebnis $h(x)$ geschrieben wird:

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

Das neutrale Element e ist x. Die Elemente $x, \dfrac{1}{x}, 1-x, \dfrac{x}{x-1}$ sind jeweils zu sich selbst invers, $\dfrac{x-1}{x}$ und $\dfrac{1}{1-x}$ sind zueinander invers. P ist also eine Gruppe der Ordnung 6. Sie ist nicht abelsch, denn

$$(1-x) \circ \frac{1}{1-x} = \frac{x}{x-1} \neq \frac{1}{1-x} \circ (1-x) = \frac{1}{x}.$$

12.1.2.8. Es sei m eine natürliche Zahl > 1 und

$$\varepsilon := \cos\frac{2\pi}{m} + i\sin\frac{2\pi}{m}.$$

Dann ist

$$\varepsilon^k = \cos\frac{2\pi k}{m} + i\sin\frac{2\pi k}{m}$$

(vgl. MfL, Bd. 2, 7.3.) und m der kleinste Exponent aus $\mathbf{N}^*$, für den $\varepsilon^m = 1$ gilt. $\varepsilon^0 = 1, \varepsilon, \varepsilon^2, \ldots, \varepsilon^{m-1}$ bilden bezüglich der Multiplikation der komplexen Zahlen eine zyklische Gruppe der Ordnung m.

12.1.2.9. Es sei m eine natürliche Zahl > 1 und a, b ganze Zahlen. a heißt *kongruent b modulo m* genau dann, wenn m Teiler von $a - b$ ist:

$$a \equiv b \bmod m :\Leftrightarrow m \mid a - b$$

(vgl. MfL, Bd. 1, 3.7.). Die so definierte Kongruenz ist eine Äquivalenzrelation und gibt Anlaß zu einer Einteilung der Menge der ganzen Zahlen in paarweise elementefremde Klassen (vgl. MfL, Bd. 1, 2.5.), die *Restklassen modulo m* genannt werden. Dividiert man a und b durch m, so gelangt man zu der Darstellung

$$a = mq + r \qquad (q \in \mathbf{Z} \wedge r \in \mathbf{N} \wedge 0 \leqq r < m),$$

$$b = mq' + r' \qquad (q' \in \mathbf{Z} \wedge r' \in \mathbf{N} \wedge 0 \leqq r' < m).$$

Daraus ergibt sich

$$m \mid a - b \Leftrightarrow m \mid r - r' \Leftrightarrow r = r'.$$

Also liegen zwei Zahlen a und b genau dann in der gleichen Restklasse mod m, wenn sie bei der Division durch m den gleichen Rest aus der Menge $\{0, 1, 2, \ldots, m - 1\}$ lassen.

Ist $a = sm + a'$ und $b = tm + b'$ $(s, t \in \mathbf{Z})$, so folgt

$$a + b = (s + t)\, m + a' + b'$$

und

$$ab = (stm + a't + b's)\, m + a'b'.$$

Daher gilt:

$$a \equiv a' \bmod m \wedge b \equiv b' \bmod m \Rightarrow a + b \equiv a' + b' \bmod m \wedge ab \equiv a'b' \bmod m \qquad (8)$$

Es bezeichne $[a]$ diejenige Restklasse modulo m, in der a liegt. Jedes Element aus $[a]$ heißt ein *Repräsentant* von $[a]$. Durch

$$[a] + [b] := [a + b] \qquad (9)$$

wird eine Addition in der Menge der Restklassen modulo m erklärt, indem man als Summe der Restklassen von a und b diejenige Restklasse definiert, in der $a + b$ liegt. Dies Ergebnis hängt nicht von der Auswahl der benutzten Restklassenrepräsentanten ab. Sind nämlich a' bzw. b' andere Elemente aus den Restklassen $[a]$ bzw. $[b]$, d. h., ist $a \equiv a' \bmod m$ und $b \equiv b' \bmod m$, so besagt (8), daß $a + b$ und $a' + b'$ in derselben Restklasse liegen. Also gilt:

$$[a] = [a'] \wedge [b] = [b'] \Rightarrow [a] + [b] = [a + b] = [a' + b'] = [a'] + [b'].$$

Durch (9) wird deshalb eine Operation in der Menge der Restklassen modulo m erklärt.

Ganz entsprechend definiert man eine Multiplikation der Restklassen modulo m durch

$$[a]\,[b] := [ab] \tag{10}$$

und zeigt mittels (8), daß auch hier das Ergebnis nicht von der Auswahl der benutzten Restklassenrepräsentanten abhängt.

Da die Addition und Multiplikation der Restklassen mod m mittels der entsprechenden Operationen in der Menge $\mathbf{Z}$ der ganzen Zahlen definiert wurden, weist der Leser leicht nach, daß für sie jeweils das Assoziativgesetz sowie das Kommutativgesetz gelten und auch das Distributivgesetz erfüllt wird. Die Menge Z_m der Restklassen mod m bildet bezüglich der Restklassenaddition (9) als Operation eine abelsche Gruppe der Ordnung m. Weil ihre Elemente $[1], [2], \ldots, [m] = [0]$ wegen

$$[k] = k[1] := [1] + \cdots + [1]$$

sämtlich Vielfache von $[1]$ sind, liegt eine additiv geschriebene zyklische Gruppe vor. Das neutrale Element ist $[0]$, und zu $[a]$ invers ist $[-a]$.

Mit den durch (9) und (10) definierten Operationen ist die Menge der Restklassen mod m ein kommutativer Ring (vgl. 11.2.(18)), der *Restklassenring* mod m genannt wird.

12.1.2.10. Für zwei Elemente a und a' einer Restklasse mod m gilt $a = sm + a'$ ($s \in \mathbf{Z}$). Daher ist jeder gemeinsame Teiler von a und m auch Teiler von a' und jeder gemeinsame Teiler von a' und m Teiler von a. Insbesondere gilt also für die größten gemeinsamen Teiler $a \sqcap m = a' \sqcap m$. Diejenigen Restklassen, deren Elemente zum Modul m teilerfremd sind, heißen *prime Restklassen* mod m. Sind $[a]$ und $[b]$ prime Restklassen mod m, so ist auch $[a]\,[b] = [ab]$ eine prime Restklasse mod m.

Zu zwei teilerfremden ganzen Zahlen a und m gibt es ganze Zahlen u und v, so daß

$$au + mv = 1$$

ist (vgl. MfL, Bd. 1, 3.7.). Multipliziert man mit der ganzen Zahl b und setzt $ub = x$, so ergibt sich daraus

$$ax + m(vb) = b. \tag{11}$$

Mit a und b ist auch x zu m teilerfremd. Daher bedeutet (11), daß es zu den modulo m primen Restklassen $[a]$ und $[b]$ eine prime Restklasse $[x]$ gibt, für die

$$[a]\,[x] = [b]$$

gilt. Weil die Restklassenmultiplikation assoziativ und kommutativ ist, erfüllen die primen Restklassen mod m bezüglich der Restklassenmultiplikation als Operation (1) und (2) und bilden deshalb eine endliche abelsche Gruppe der Ordnung $\varphi(m)$ (*Eulersche Funktion*; vgl. MfL, Bd. 1, 3.7.).

12.1.2.11. E^2 sei die euklidische Ebene, $|P, Q|$ der Abstand ihrer Punkte P, Q und B_2 die Menge aller 1-1-Abbildungen f von E^2 (als Punktmenge aufgefaßt) auf sich, welche die Abstände aller Punktepaare ungeändert lassen:

$$B_2 = \left\{f : f \in \mathfrak{T}(E^2) \wedge \bigwedge_{P,Q \in E^2} |P, Q| = |f(P), f(Q)|\right\}.$$

Mit der Nacheinanderausführung als Operation bildet B_2 eine Gruppe.

Sind $f, g \in B_2$, so gilt

$$|f(g(P)), f(g(Q))| = |g(P), g(Q)| = |P, Q|.$$

Da auch $f \circ g \in \mathfrak{T}(E^2)$, ist $f \circ g \in B_2$.

Für die Nacheinanderausführung von 1-1-Abbildungen gilt das Assoziativgesetz. Neutrales Element ist die identische Abbildung. Zu jedem $f \in B_2$ gibt es in $\mathfrak{T}(E^2)$ die inverse Abbildung f^{-1}, von der noch gezeigt werden muß, daß auch sie die Abstände je zweier Punkte ungeändert läßt. Sind $P, Q \in E^2$, so gilt wegen $f \in B_2$

$$|f^{-1}(P), f^{-1}(Q)| = |f(f^{-1}(P)), f(f^{-1}(Q))| = |P, Q|.$$

Die Elemente von B_2 heißen *Bewegungen* von E^2 und B_2 die *Gruppe der Bewegungen* oder *Bewegungsgruppe* von E^2.

12.1.2.12. Eine Teilmenge F der Punktmenge der euklidischen Ebene E^2 nennen wir eine *Figur*. Bezeichne B_{2F} diejenigen Bewegungen aus B_2, die F auf sich abbilden.

B_{2F} bildet bezüglich der Nacheinanderausführung eine Gruppe.

Offensichtlich ist mit f und g auch $f \circ g$ Element von B_{2F}. Das Assoziativgesetz gilt, und die identische Abbildung ist neutrales Element in B_{2F}. $f \in B_{2F}$ besitzt eine inverse Abbildung f^{-1} in B_2. Ist P ein Punkt von F, so existiert ein $\overline{P} \in F$ mit $f(\overline{P}) = P$. Daher ist $f^{-1}\big(f(\overline{P})\big) = \overline{P} = f^{-1}(P)$, also $f^{-1} \in B_{2F}$.

B_{2F} heißt *Gruppe der Figur* F. Wir betrachten die Gruppe B_{2D} der Eckpunkte eines gleichseitigen Dreiecks in E^2. Ein Element aus B_{2D} ist bereits durch die Angabe der Bilder der drei Eckpunkte festgelegt, denn jeder Punkt von E^2 ist durch die Abstände von drei festen, nicht auf einer Geraden liegenden Punkten bestimmt. Jede Permutation der drei Eckpunkte ist möglich. B_{2D} ist also die symmetrische Gruppe S_3.

Die Gruppe B_{2Q} der Eckpunkte eines Quadrates ist dagegen nicht die S_4. Bezeichnet man die Eckpunkte des Quadrates mit 1, 2, 3, 4 (vgl. Abb. 14), so beschreibt die Permutation

$$\begin{pmatrix}1&2&3&4\\2&1&3&4\end{pmatrix}$$

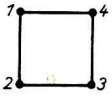


Abb. 14

keine Bewegung aus B_{2Q}, weil der Abstand (1, 4) vom Abstand der Bildpunkte (2, 4) verschieden ist. Die Bewegungen von B_{2Q} werden durch die Permutationen

$$\begin{pmatrix}1&2&3&4\\1&2&3&4\end{pmatrix},\ \begin{pmatrix}1&2&3&4\\2&3&4&1\end{pmatrix},\ \begin{pmatrix}1&2&3&4\\3&4&1&2\end{pmatrix},\ \begin{pmatrix}1&2&3&4\\4&1&2&3\end{pmatrix},$$

$$\begin{pmatrix}1&2&3&4\\1&4&3&2\end{pmatrix},\ \begin{pmatrix}1&2&3&4\\3&2&1&4\end{pmatrix},\ \begin{pmatrix}1&2&3&4\\2&1&4&3\end{pmatrix},\ \begin{pmatrix}1&2&3&4\\4&3&2&1\end{pmatrix}$$

angegeben. Die Gruppe hat also die Ordnung 8 und ist wegen

$$\begin{pmatrix}1&2&3&4\\2&3&4&1\end{pmatrix}\begin{pmatrix}1&2&3&4\\1&4&3&2\end{pmatrix} \neq \begin{pmatrix}1&2&3&4\\1&4&3&2\end{pmatrix}\begin{pmatrix}1&2&3&4\\2&3&4&1\end{pmatrix}$$

nicht abelsch.

12.2. Komplexe und Untergruppen

Da eine Struktur nur wenigen und naheliegenden Axiomen genügen muß, um eine Gruppe zu sein, ist der Gruppenbegriff, wie die Beispiele illustrieren, in vielen Bereichen der Mathematik von Bedeutung. Aus den Grundaussagen werden in der Gruppentheorie viele nichttriviale Resultate gewonnen, die dann in den jeweiligen Modellen Sätze über Abbildungen, Permutationen, Bewegungen, Zahlen, Restklassen u. v. a. ergeben. Um die abzuleitenden Aussagen bequem formulieren zu können, führen wir zunächst einige Begriffe ein und untersuchen einfache Zusammenhänge.

12.2.1. Für die Teilmengen der Menge der Elemente einer Gruppe G kann neben den mengentheoretischen Beziehungen, wie sie durch die Zeichen $=, \subseteq, \supseteq, \subset, \supset, \cap, \cup$ ausgedrückt werden, auch die in G erklärte Operation betrachtet werden. Um diesen Unterschied zur Mengenlehre zu betonen, erfolgt die

Definition 1. In der Gruppe G heißt

$$K \textit{ Komplex von } G :\Leftrightarrow K \subseteq G \wedge K \neq \emptyset .$$

Definition 2. Sind K und L Komplexe der Gruppe G, so heißt

$$KL := \{kl : k \in K \wedge l \in L\}$$

Komplexprodukt von K und L. $K^{-1} := \{k^{-1} : k \in K\}$ heißt *zu K inverser Komplex.*

In der symmetrischen Gruppe S_3 (vgl. 12.1.2.6.) enthält das Komplexprodukt KL der Komplexe $K = \{q, r\}$ und $L = \{p, r\}$ die Permutationen

$$qp = \begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix} = \begin{pmatrix}1&2&3\\1&2&3\end{pmatrix} = e,$$

$$qr = \begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix} = \begin{pmatrix}1&2&3\\2&1&3\end{pmatrix} = t,$$

$$rp = \begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix} = \begin{pmatrix}1&2&3\\2&1&3\end{pmatrix} = t,$$

$$rr = \begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix} = \begin{pmatrix}1&2&3\\1&2&3\end{pmatrix} = e,$$

es ist also $KL = \{e, t\}$. Ebenso berechnet man $LK = \{e, s\}$.

Die Bildung des Komplexproduktes ist eine Operation in der Menge der Komplexe einer Gruppe G. Sie ist assoziativ, da die Gruppenoperation assoziativ ist. Wie unser Beispiel zeigt, ist sie aber im allgemeinen nicht kommutativ.

Ist K ein beliebiger und $A = \{a\}$ ein Komplex von G, der nur aus einem Element a besteht, so schreiben wir statt KA bzw. AK auch Ka bzw. aK.

Definition 3. K und L seien Komplexe von G.

L heißt (unter G) zu K *konjugiert* $:\Leftrightarrow \bigvee_{g \in G} g^{-1}Kg = L$.

K heißt *normal* (oder *invariant*) in $G :\Leftrightarrow \bigwedge_{g \in G} g^{-1}Kg = K$.

Ist $g^{-1}Kg = L$, so sagt man auch, daß K bei der Transformation mit dem Element $g \in G$ in L übergeht, und nennt L *transformierten Komplex*. K ist also genau dann ein invarianter Komplex in G, wenn K bei allen Transformationen mit Elementen aus G in sich übergeht. $g^{-1}Kg = K$ bedeutet nicht, daß die einzelnen Elemente von K bei der Transformation mit g ungeändert bleiben. Davon kann man sich am Beispiel der Gruppe S_3 überzeugen, denn $K = \{e, p, q\}$ ist darin normaler Komplex. Insbesondere prüft man leicht die Beziehungen

$$r^{-1} = r; \quad r^{-1}er = e, \quad r^{-1}pr = q, \quad r^{-1}qr = p$$

nach, aus denen $r^{-1}Kr = K$ folgt.

In jeder Gruppe G besitzen die Gleichungen $ax = b$ und $ya = b$ $(a, b \in G)$ eindeutig bestimmte Lösungen. Daher gilt für beliebige Elemente $g \in G$

$$Gg = gG = G \qquad \text{oder} \qquad g^{-1}Gg = G. \tag{1}$$

Ferner ist

$$GG = G \tag{2}$$

und $(g^{-1})^{-1} = g$ zufolge

$$G^{-1} = G. \tag{3}$$

Neben G ist in jeder Gruppe G der nur aus dem neutralen Element bestehende Komplex invariant.

12.2.2. Von besonderer Bedeutung sind Komplexe, auf die zutrifft die

Definition 4. Ist U Komplex der Gruppe G, so heißt U *Untergruppe* von $G :\Leftrightarrow U$ ist bezüglich der in G definierten Operation eine Gruppe.

Die in G erklärte Multiplikation ist genau dann auch eine Operation in U, wenn für sie gilt:

$$u \in U \wedge v \in U \Rightarrow uv \in U. \tag{4}$$

U ist dann und nur dann Untergruppe, wenn neben (4)

$$u \in U \Rightarrow u^{-1} \in U \tag{5}$$

erfüllt ist. Da es nämlich in U wenigstens ein Element u gibt, folgt aus (5) und (4), daß $uu^{-1} = e \in U$. Das Assoziativgesetz gilt selbstverständlich in U, da es in G gilt. Also ist U eine Gruppe (vgl. 12.1., Satz 1).

Gleichbedeutend mit (4) und (5) ist die Bedingung

$$u \in U \wedge v \in U \Rightarrow uv^{-1} \in U. \tag{6}$$

Denn weil U wenigstens ein Element u enthält, ergibt sich für $v = u$ aus (6), daß $uu^{-1} = e \in U$. Ferner ist mit e und u auch $eu^{-1} = u^{-1}$ in U enthalten. Daher liegt mit u und v auch v^{-1} und nach (6) $u(v^{-1})^{-1} = uv$ in U. Umgekehrt ergibt sich offensichtlich (6) aus (4) und (5). Also gilt der

Satz 1. *Der Komplex U von G ist genau dann Untergruppe von G, wenn*

$$UU^{-1} \subseteq U \tag{6'}$$

gilt.

Die nur aus dem neutralen Element bestehende Gruppe sowie G sind Untergruppen jeder Gruppe G. Sie heißen *triviale*, alle anderen heißen *nichttriviale* oder *eigentliche Untergruppen* von G.

Mit Hilfe dieser Überlegungen ist es leicht, die folgenden Mengen als Untergruppen der gegebenen Gruppen zu erkennen (vgl. 12.1.2.):

1. $\{g\colon \bigvee_{z\in\mathbf{Z}}\, g = 2z\}$ in $(\mathbf{Z}, +)$,
2. $\mathbf{Z}$ in $(\mathbf{Q}, +)$,
3. $\mathbf{Q}$ in $(\mathbf{R}, +)$,
4. $\{\boldsymbol{A}\colon \boldsymbol{A} \in L_n \wedge \det \boldsymbol{A} = 1\}$ in $(L_n, \cdot)$,
5. $\{f\colon f \in \mathfrak{T}(M) \wedge$ für ein festes Element $a \in M$ ist $f(a) = a\}$ in $\mathfrak{T}(M)$, insbesondere $\{e, r\}$ in S_3,
6. $\{e, p, q\}$ in S_3,
7. $B_{\mathfrak{T}^{\mathbf{T}}}$ in B_2,
8. $\left\{\begin{pmatrix}1&2&3&4\\1&2&3&4\end{pmatrix}, \begin{pmatrix}1&2&3&4\\2&3&4&1\end{pmatrix}, \begin{pmatrix}1&2&3&4\\3&4&1&2\end{pmatrix}, \begin{pmatrix}1&2&3&4\\4&1&2&3\end{pmatrix}\right\}$ in B_{2Q}

(Drehungen des Quadrates),

$$\left\{\begin{pmatrix}1&2&3&4\\1&2&3&4\end{pmatrix}, \begin{pmatrix}1&2&3&4\\1&4&3&2\end{pmatrix}\right\} \text{ in } B_{2Q}$$

(Spiegelungen des Quadrates an einer Diagonalen).

Zwei Elemente a, b einer Gruppe heißen *vertauschbar*, wenn $ab = ba$ ist.

Definition 5. In der Gruppe G heißt

$$Z(G) := \{z\colon z \in G \wedge \bigwedge_{g\in G} gz = zg\} \quad \textit{Zentrum} \text{ von } G.$$

$Z(G)$ enthält das neutrale Element $e \in G$. Aus $gz_1 = z_1 g$ und $gz_2 = z_2 g$ ergibt sich $gz_1z_2 = z_1gz_2 = z_1z_2g$, also

$$z_1 \in Z(G) \wedge z_2 \in Z(G) \Rightarrow z_1z_2 \in Z(G).$$

Aus $gz = zg$ folgt $z^{-1}g^{-1} = g^{-1}z^{-1}$. Weil $G^{-1} = G$ ist, gilt dann

$$z \in Z(G) \Rightarrow z^{-1} \in Z(G).$$

$Z(G)$ ist also eine Untergruppe von G. Die abelschen Gruppen stimmen mit ihrem Zentrum überein, $Z(S_3)$ enthält nur das neutrale Element und $Z(B_{2Q})$ besteht aus den Permutationen

$$\begin{pmatrix}1&2&3&4\\1&2&3&4\end{pmatrix} \quad \text{und} \quad \begin{pmatrix}1&2&3&4\\3&4&1&2\end{pmatrix}.$$

Ist U eine Untergruppe von G, so sind auch die konjugierten Komplexe $g^{-1}Ug$, $g \in G$, Untergruppen von G, denn unter Verwendung von 12.1.(7), sowie (2) und (3)

erhält man

$$(g^{-1}Ug)\,(g^{-1}Ug)^{-1} = g^{-1}Ugg^{-1}U^{-1}g = g^{-1}\,UU^{-1}g = g^{-1}Ug,$$

d. h., (6') gilt für den Komplex $g^{-1}Ug$.

Ist U Untergruppe von G, V Untergruppe von U, so ist V auch Untergruppe von G.

Ist $\mathfrak{U}$ eine nichtleere Menge von Untergruppen von G, so ist auch der Durchschnitt $D = \bigcap_{U\in\mathfrak{U}} U$ eine Untergruppe von G, denn liegen u und v in jedem $U \in \mathfrak{U}$, so liegt notwendig nach (6) auch uv^{-1} in jedem $U \in \mathfrak{U}$, und daher ist D eine Untergruppe.

Definition 6. Bezeichnet K einen Komplex der Gruppe G und $\mathfrak{U}$ die Menge aller K umfassenden Untergruppen U von G, so heißt $\langle K\rangle := \bigcap_{U\in\mathfrak{U}} U$ *Erzeugnis* von K.

$\langle K\rangle$ ist die kleinste Untergruppe von G, die K enthält, da $\bigwedge_{U\in\mathfrak{U}} \langle K\rangle \subseteq U$ gilt. Sie besteht als allen endlichen Produkten $k_1^{\varepsilon_1}\cdots k_r^{\varepsilon_r}$, die aus Elementen $k_1, \ldots \in K$ mit Exponenten $\varepsilon_1, \ldots \in \{+1, -1\}$ erzeugt werden können. Nach 12.1.1. ist unmittelbar klar, daß alle diese Produkte in $\langle K\rangle$ liegen müssen, und aus Satz 1 folgt, daß der Komplex dieser Produkte eine K umfassende Untergruppe ist.

Ein Komplex K, für den $\langle K\rangle = G$ ist, heißt ein *Erzeugendensystem* von G. Gibt es einen endlichen Komplex $K = \{k_1, \ldots, k_n\}$ dieser Eigenschaft, so nennt man G *endlich erzeugbar*, und die kleinste dabei auftretende Zahl n wird als *Erzeugendenzahl* von G bezeichnet. Durch die Erzeugendenzahl 1 sind genau die zyklischen Gruppen beschrieben.

Für ein Element a einer Gruppe G besteht $\langle a\rangle$ aus der Menge aller verschiedenen Potenzen a^k ($k \in \mathbf{Z}$) und wird die *durch a erzeugte zyklische Untergruppe* von G genannt. Ist für je zwei verschiedene Zahlen k_1, k_2 aus $\mathbf{Z}$ immer $a^{k_1} \neq a^{k_2}$, so ist $\langle a\rangle$ eine unendliche zyklische Gruppe. Gilt aber eine Gleichung der Form $a^k = a^l$, in der wir $k < l$ annehmen können, so ist $a^{l-k} = e$. Sei m der kleinste Exponent aus $\mathbf{N}^*$ mit

$$a^m = e. \tag{7}$$

Dann sind offenbar $a^0 = e$, $a^1, a^2, \ldots, a^{m-1}$ paarweise verschieden, und jedes a^k ($k \in \mathbf{Z}$) ist gleich einer dieser Potenzen. Denn da es Zahlen $q \in \mathbf{Z}$ und $r \in \{0, 1, \ldots, m-1\}$ gibt, für die $k = mq + r$ ist (vgl. MfL, Bd. 1, 3.7.), gilt

$$a^k = a^{mq+r} = (a^m)^q a^r = ea^r = a^r.$$

$\langle a\rangle$ ist also eine zyklische Gruppe der Ordnung m. In ihr wird durch (7) das Rechnen festgelegt. Diese Gruppe kann daher durch Angabe des erzeugenden Elementes und der *definierenden Relation* (7) vollständig beschrieben werden.

Definition 7. Für ein Element g einer Gruppe G heißt $|g| := |\langle g\rangle|$ *Ordnung* von g, sofern $\langle g\rangle$ eine endliche Gruppe ist. Sonst nennt man g *Element unendlicher Ordnung*.

In einer endlichen Gruppe hat selbstverständlich jedes Element endliche Ordnung, während es in einer unendlichen Gruppe auch Elemente unendlicher Ordnung geben kann.

An einigen Beispielen wollen wir zeigen, daß auch nichtzyklische Gruppen durch Angabe eines Erzeugendensystems und definierender Relationen, welche die Rechnung in der Gruppe vollständig bestimmen, festgelegt werden können.

Für die S_3 (vgl. 12.1.2.6.) sind p und r erzeugende Elemente. Man prüft leicht nach, daß sie den Relationen

$$p^3 = e, \qquad r^2 = e, \qquad rp = p^2r$$

genügen. Daher läßt sich jedes Produkt aus Potenzen von p und r so umrechnen, daß es die Form

$$p^\alpha r^\beta, \qquad \alpha \in \{0, 1, 2\}, \qquad \beta \in \{0, 1\}$$

erhält. Die Elemente der S_3 sind

$$e, \quad p, \quad q = p^2, \quad r, \quad s = pr, \quad t = p^2r.$$

Die Ausnutzung der definierenden Relationen zeigen wir an der Berechnung des Produktes

$$tp = (p^2r)\,p = p^2(rp) = p^3(p^2r) = p^4r = pr.$$

Die Gruppe B_{2Q} (vgl. 12.1.2.12.) kann aus den Elementen

$$r := \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix} \qquad \text{und} \qquad s := \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix}$$

unter Beachtung der definierenden Relationen

$$r^4 = e, \qquad s^2 = e, \qquad sr = r^3s$$

erzeugt werden. Jedes Element der Gruppe B_{2Q} läßt sich auf genau eine Weise in der Form

$$r^\varrho s^\sigma \qquad \text{mit} \qquad \varrho \in \{0, 1, 2, 3\} \qquad \text{und} \qquad \sigma \in \{0, 1\}$$

darstellen.

Die multiplikative Gruppe der primen Restklassen mod 8 kann aus den Elementen $[3]$ und $[5]$ erzeugt werden. Definierende Relationen sind

$$[3]^2 = [1], \qquad [5]^2 = [1], \qquad [3]\,[5] = [5]\,[3].$$

Oft läßt sich die Menge der Untergruppen einer Gruppe G und ihre Einbettung in die Gruppe in übersichtlicher Weise durch *Graphen* (auch *Diagramme* genannt) veranschaulichen. Dabei werden die Untergruppen derart durch Punkte einer Ebene dargestellt, daß das Bild von V unterhalb von U liegt, wenn $V \subset U$ ist. Beide Punkte werden durch eine Strecke verbunden, wenn G keine Untergruppe W mit der Eigen-

schaft $V \subset W$ und $W \subset U$ enthält. Für einige Beispiele aus 12.1.2. geben wir die Graphen an:

Die Abbildungen 15 und 16 zeigen die Diagramme der zyklischen Gruppen der Restklassen nach den Moduln 8 und 6.

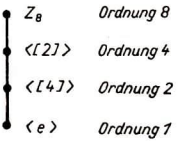


Abb. 15

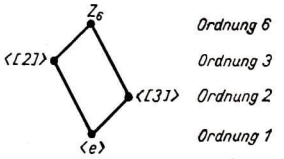


Abb. 16

In den Abbildungen 17, 18 und 19 sind die Untergruppen der Gruppe der primen Restklassen mod 8, der S_3 und der B_{2Q} dargestellt.

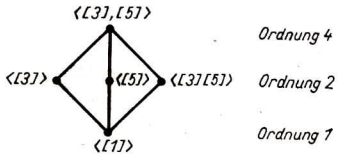


Abb. 17

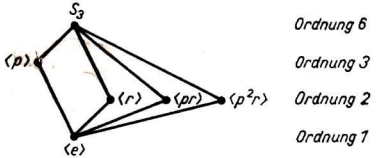


Abb. 18

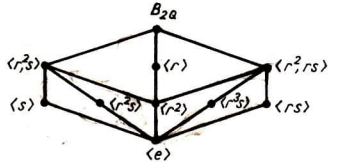


Abb. 19

Mitunter will man nur die Lage einiger Untergruppen in G illustrieren. Um die Übersichtlichkeit zu erhöhen, nimmt man dann nicht alle Untergruppen von G in das Diagramm auf. Ein Beispiel gibt Abb. 20 für zwei echte Untergruppen U, V der Gruppe G.

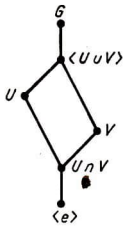


Abb. 20

12.2.3. Definition 8. Bezeichnet U eine Untergruppe, g ein Element der Gruppe G, so heißt der Komplex Ug *Rechtsnebenklasse* nach U und gU *Linksnebenklasse* nach U.

Jedes Element $g \in G$ liegt in einer Rechtsnebenklasse nach U, da $g = eg \in Ug$. Kein Element aus G liegt in zwei verschiedenen Rechtsnebenklassen. Ist nämlich $d \in Ug_1 \cap Ug_2$, so gibt es in U Elemente u_1, u_2, für welche die Gleichungen $d = u_1g_1 = u_2g_2$ gelten. Daher ist $g_2 = u_2^{-1}u_1g_1$, und nach (1) ergibt sich $Ug_2 = Uu_2^{-1}u_1g_1 = Ug_1$. Es erfolgt also eine Einteilung der Menge G in paarweise elementefremde Klassen, die Rechtsnebenklassen nach U.

Dies wird häufig durch

$$G = Ua \cup Ub \cup Uc \cup \cdots \tag{8}$$

ausgedrückt, wobei $\cup$ hier die Vereinigung elementefremder Teilmengen bezeichnet. Zwei Elemente g_1, g_2 von G liegen genau dann in einer Rechtsnebenklasse nach U, wenn $g_1g_2^{-1} \in U$. Aus $g_1g_2^{-1} = u \in U$ und $g_2 = u_2g \in Ug$ folgt nämlich $g_1 = ug_2 = uu_2g \in Ug$. Umgekehrt ergibt sich aus $g_1 = u_1g \in Ug$ und $g_2 = u_2g \in Ug$, daß $g_1g_2^{-1} = u_1gg^{-1}u_2^{-1} = u_1u_2^{-1} \in U$. Eine Nebenklasse ist natürlich U selbst.

Die Elemente $a, b, c, \ldots$ in (8) werden ein *Rechtsrepräsentantensystem* R für G nach U genannt. Da Elemente g_1, g_2 aus G genau dann die gleiche Rechtsnebenklasse nach U repräsentieren, wenn $g_1 = ug_2$ mit $u \in U$ ist, erhält man aus einem gegebenen Rechtsrepräsentantensystem für G nach U alle möglichen Rechtsrepräsentantensysteme, wenn man seine Elemente von links nacheinander mit sämtlichen Elementen aus U multipliziert.

Ganz entsprechend kann man G in Linksnebenklassen nach U zerlegen. Aus (8) erhält man unter Benutzung von (3) durch Übergang zu den inversen Elementen

$$G^{-1} = G = a^{-1}U \cup b^{-1}U \cup c^{-1}U \cup \cdots .$$

Ist also R ein Rechtsrepräsentantensystem, so ist R^{-1} ein *Linksrepräsentantensystem* für G nach U. Die Menge der Linksnebenklassen ist daher genau dann endlich, wenn die Menge der Rechtsnebenklassen endlich ist, und beide Mengen enthalten in diesem Fall die gleiche Anzahl von Nebenklassen.

Diese Mengen sind jedoch im allgemeinen nicht gleich. Beispielsweise sind in der S_3 die Rechtsnebenklassen nach der Untergruppe $U = \langle r \rangle$:

$$U = Ue = \{e, r\}, \qquad Up = \{p, t\}, \qquad Uq = \{q, s\},$$

die Linksnebenklassen aber

$$U = eU = \{e, r\}, \qquad qU = \{q, t\}, \qquad pU = \{p, s\}.$$

Definition 9. Ist in der Gruppe G die Anzahl i der Nebenklassen nach der Untergruppe U endlich, so heißt $[G : U] := i$ *Index von* U *in* G. Sonst heißt U *Untergruppe von unendlichem Index.*

Da jede Nebenklasse nach der Untergruppe $E = \langle e \rangle$ nur aus einem Element besteht, ist in endlichen Gruppen G die Ordnung $|G| = [G : E]$.

Jede Nebenklasse Ug von G nach der endlichen Untergruppe U enthält genau $|U|$ Elemente, denn aus $u_1 g = u_2 g$ $(u_1, u_2 \in U)$ folgt $u_1 = u_2$. Hat U in der endlichen Gruppe G den Index i, so ergibt sich aus (8) durch Vergleich der Elementeanzahlen $|G| = |U|\, i$. Diese Tatsache wurde zuerst von LAGRANGE bewiesen, wir formulieren sie als

Satz 2. *Für jede Untergruppe* U *einer endlichen Gruppe* G *gilt*

$$[G : E] = [G : U]\,[U : E]. \tag{9}$$

Folgerung 1. *In einer endlichen Gruppe sind die Ordnung und der Index jeder Untergruppe Teiler der Gruppenordnung.*

Betrachtet man die von einem Element erzeugte Untergruppe, so ergibt sich

Folgerung 2. *In einer endlichen Gruppe* G *ist die Ordnung jedes Elementes ein Teiler der Gruppenordnung. Daher sind Gruppen von Primzahlordnung zyklisch. Ferner gilt für alle* $g \in G$

$$g^{|G|} = e. \tag{10}$$

Wendet man dieses Ergebnis auf die Gruppe der primen Restklassen mod m an, so erhält man

Folgerung 3. *Für jede zum Modul* m *teilerfremde Zahl* a *ist*

$$a^{\varphi(m)} \equiv 1 \bmod m.$$

Insbesondere ergibt diese zahlentheoretische Aussage

Folgerung 3′. *Ist* $a \in \mathbf{Z}$ *nicht durch die Primzahl* p *teilbar, so gilt*

$$a^{p-1} \equiv 1 \bmod p.$$

Ohne Beweis wurde der letzte Satz schon von dem französischen Mathematiker PIERRE DE FERMAT (1601—1665) angegeben. Daher nennt man (10) den *Fermatschen Satz der Gruppentheorie.*

12.3. Isomorphie von Gruppen

12.3.1. In einer Gruppe G beherrscht man die Rechnung vollständig, wenn man zu jedem geordneten Paar von Elementen g, h auch deren Produkt gh kennt. Für endliche Gruppen liegt es nahe, alle Produkte in einer Tabelle der Form

	$\dots h \dots$
$\vdots$	$\vdots$
g	$\dots gh \dots$
$\vdots$	$\vdots$

anzugeben. Dabei stehen in der Eingangszeile und Eingangsspalte jeweils alle Elemente der Gruppe in irgendeiner Reihenfolge. Im Schnittpunkt der Zeile von g mit der Spalte von h wird das Produkt gh notiert. Eine solche Tabelle heißt *Gruppentafel* von G.

Als Beispiele betrachten wir die additiven Gruppen der Restklassen nach den Moduln 2 und 3. Beide Gruppen Z_2 und Z_3 sind zyklisch, ihre Gruppentafeln lauten

Z_2	[0]	[1]
[0]	[0]	[1]
[1]	[1]	[0]

und

Z_3	[0]	[1]	[2]
[0]	[0]	[1]	[2]
[1]	[1]	[2]	[0]
[2]	[2]	[0]	[1]

.

Die multiplikative Gruppe der primen Restklassen mod 8 hat die Gruppentafel

	[1]	[3]	[5]	[7]
[1]	[1]	[3]	[5]	[7]
[3]	[3]	[1]	[7]	[5]
[5]	[5]	[7]	[1]	[3]
[7]	[7]	[5]	[3]	[1]

Schließlich geben wir noch Gruppentafeln für die Gruppe S_3 in den Bezeichnungen von 12.1.2.6. sowie für die Gruppe B_{2Q} unter Verwendung des in 12.2.2. betrachteten Erzeugendensystems und der zugehörigen definierenden Relationen an.

S_3	e	p	q	r	s	t
e	e	p	q	r	s	t
p	p	q	e	s	t	r
q	q	e	p	t	r	s
r	r	t	s	e	q	p
s	s	r	t	p	e	q
t	t	s	r	q	p	e

B_{2Q}	e	r	r^2	r^3	s	rs	r^2s	r^3s
e	e	r	r^2	r^3	s	rs	r^2s	r^3s
r	r	r^2	r^3	e	rs	r^2s	r^3s	s
r^2	r^2	r^3	e	r	r^2s	r^3s	s	rs
r^3	r^3	e	r	r^2	r^3s	s	rs	r^2s
s	s	r^3s	r^2s	rs	e	r^3	r^2	r
rs	rs	s	r^3s	r^2s	r	e	r^3	r^2
r^2s	r^2s	rs	s	r^3s	r^2	r	e	r^3
r^3s	r^3s	r^2s	rs	s	r^3	r^2	r	e

Da die links-(rechts-)seitige Division in der Gruppe G möglich und eindeutig ist, steht in jeder Zeile (Spalte) der Gruppentafel jedes Element von G genau einmal. Dann und nur dann ist die Gruppentafel symmetrisch bezüglich der Diagonalen von links oben nach rechts unten, wenn G abelsch ist.

12.3.2. Die multiplikative Gruppe mit den Elementen

$$\varepsilon^0 = 1, \qquad \varepsilon = -\frac{1}{2} + \frac{i}{2}\sqrt{3}, \qquad \varepsilon^2 = -\frac{1}{2} - \frac{i}{2}\sqrt{3}$$

(vgl. 12.1.2.8.) besitzt die Gruppentafel

	ε^0	ε^1	ε^2
ε^0	ε^0	ε^1	ε^2
ε^1	ε^1	ε^2	ε^0
ε^2	ε^2	ε^0	ε^1

Die Rechnung verläuft genauso wie in der Gruppe Z_3, denn diese Gruppentafel entsteht aus derjenigen der Gruppe Z_3, wenn darin die Elemente durch ihre Bilder bei der Abbildung $f\colon [k] \mapsto \varepsilon^k$ $(k = 0, 1, 2)$ ersetzt werden. Die Umkehrabbildung überführt entsprechend die vorliegende Gruppentafel in diejenige der Gruppe Z_3.

Der Leser überprüft leicht, daß mit der Abbildung

$$x \mapsto e, \qquad \frac{1}{1-x} \mapsto p, \qquad \frac{x-1}{x} \mapsto q,$$

$$1-x \mapsto r, \qquad \frac{1}{x} \mapsto s, \qquad \frac{x}{x-1} \mapsto t$$

die Gruppentafel der Gruppe P (vgl. 12.1.2.7.) in die Gruppentafel der S_3 übergeht.

In der multiplikativen Bezeichnungsweise bedeutet diese Eigenschaft der angegebenen Abbildungen, daß immer

Bild des Produktes = Produkt der Bilder

gilt.

Zwar unterscheiden sich in beiden Beispielen die jeweiligen zwei Gruppen durch die konkrete Bedeutung und Bezeichnung ihrer Elemente, doch verläuft die Rechnung in ihnen gleichartig. Wir befinden uns damit in derselben Situation wie ein Kind, das die Addition von Zahlen an Hand von realen Dingen (Rechenstäbchen, Fingern, ...) erlernt, dabei erfährt, daß es gar nicht entscheidend ist, *was* addiert wird, und daher zu abstrahieren beginnt.

Uns schafft nun der wichtige Begriff der Isomorphie von Gruppen die Möglichkeit, von der Bedeutung der Gruppenelemente zu abstrahieren.

Definition 1. G und $\overline{G}$ seien Gruppen. Dann heißt
f *Isomorphismus von* G *auf* $\overline{G}$ $:\Leftrightarrow$ f ist 1-1-Abbildung von G auf $\overline{G}$

$$\wedge \bigwedge_{g_1,g_2\in G} f(g_1g_2) = f(g_1)\,f(g_2).$$

Man nennt

G *isomorph* $\overline{G}$ $:\Leftrightarrow$ ein Isomorphismus von G auf $\overline{G}$ existiert

und schreibt in diesem Fall $G \cong \overline{G}$.

Beispiele.

1. Es ist leicht nachzuprüfen, daß die Abbildungen $f_1\colon \alpha^k \mapsto k$ und $f_2\colon \alpha^k \mapsto -k$ ($k \in \mathsf{Z}$) die einzigen Isomorphismen von Z_∞ auf $(\mathsf{Z}, +)$ sind (vgl. 12.1.2.).

2. Aus seiner Schulzeit weiß der Leser, daß die Gruppe $(\mathsf{R}_+{}^*, \cdot)$ der positiven reellen Zahlen mit der Multiplikation als Operation vermittels der Abbildung

$$f = \{(x, y)\colon x \in \mathsf{R}_+{}^* \wedge y = \ln x\}$$

zur Gruppe $(\mathsf{R}, +)$ der reellen Zahlen mit der Addition als Operation isomorph ist, da

$$\bigwedge_{x_1,x_2\in\mathsf{R}_+{}^*} \ln(x_1x_2) = \ln x_1 + \ln x_2$$

gilt.

Ist f ein Isomorphismus der Gruppe G auf die Gruppe $\overline{G}$ und K ein Komplex von G, so sei $f(K) := \{f(k)\colon k \in K\}$.

Satz 1. *Ein Isomorphismus f von G auf $\overline{G}$ hat folgende Eigenschaften:*

e neutrales Element von $G \Rightarrow f(e) = \bar{e}$ neutrales Element von $\overline{G}$, (1)

a^{-1} invers zu a in $G \Rightarrow f(a^{-1}) = [f(a)]^{-1}$ invers zu $f(a)$ in $\overline{G}$, (2)

U Untergruppe von $G \Rightarrow f(U) = \overline{U}$ Untergruppe von $\overline{G}$, (3)

K normaler Komplex von $G \Rightarrow f(K) = \overline{K}$ normaler Komplex von $\overline{G}$, (4)

G abelsch $\Rightarrow f(G) = \overline{G}$ abelsch. (5)

Beweis. Da $f(e)\,f(e) = f(ee) = f(e)$, ist $f(e) = \bar{e}$ neutrales Element von $\bar{G}$. Aus $f(a)\,f(a^{-1}) = f(aa^{-1}) = f(e) = \bar{e}$ ergibt sich (2). Für die Untergruppe U gilt $UU^{-1} \subseteq U$, und unter Verwendung von (2) erhält man daraus

$$f(U)\,[f(U)]^{-1} = f(U)\,f(U^{-1}) = f(UU^{-1}) \subseteq f(U).$$

Nach 12.2., Satz 1, ist daher $f(U)$ Untergruppe von $\bar{G}$. Da jedes Element aus $\bar{G}$ Bild ist, folgt die Behauptung (4) aus

$$\bar{g}^{-1}f(K)\,\bar{g} = [f(g)]^{-1}\,f(K)\,f(g) = f(g^{-1})\,f(K)\,f(g) = f(g^{-1}Kg)$$

für alle $\bar{g} = f(g) \in \bar{G}$. Zu beliebigen Elementen $\bar{g}_1$, $\bar{g}_2$ aus $\bar{G}$ gibt es Urbilder g_1, g_2 bei f, und aus $g_1g_2 = g_2g_1$ erhält man

$$\bar{g}_1\bar{g}_2 = f(g_1)\,f(g_2) = f(g_1g_2) = f(g_2g_1) = f(g_2)\,f(g_1) = \bar{g}_2\bar{g}_1,$$

was (5) beweist.

Für jede Gruppe G gilt

$$G \cong G, \tag{6}$$

denn immer ist die identische Abbildung ein Isomorphismus von G auf G. Bildet der Isomorphismus f von G auf $\bar{G}$ die Elemente g_1, g_2 auf $f(g_1) = \bar{g}_1$, $f(g_2) = \bar{g}_2$ ab, so ergibt $f(g_1g_2) = f(g_1)\,f(g_2)$, daß

$$g_1g_2 = f^{-1}(\bar{g}_1\bar{g}_2)$$

ist und daher

$$f^{-1}(\bar{g}_1)\,f^{-1}(\bar{g}_2) = g_1g_2 = f^{-1}(\bar{g}_1\bar{g}_2)$$

gilt. Mithin ist f^{-1} ein Isomorphismus von $\bar{G}$ auf G, und d. h.

$$G \cong \bar{G} \Rightarrow \bar{G} \cong G. \tag{7}$$

Ist ferner $\bar{f}$ ein Isomorphismus von $\bar{G}$ auf $\bar{\bar{G}}$, so gilt

$$\bar{f} \circ f(g_1g_2) = \bar{f}(\bar{g}_1\bar{g}_2) = \bar{f}(\bar{g}_1)\,\bar{f}(\bar{g}_2) = \big(\bar{f} \circ f(g_1)\big)\,\big(\bar{f} \circ f(g_2)\big).$$

$\bar{f} \circ f$ ist also ein Isomorphismus von $\bar{G}$ auf $\bar{\bar{G}}$, und daher folgt

$$G \cong \bar{G} \wedge \bar{G} \cong \bar{\bar{G}} \Rightarrow G \cong \bar{\bar{G}}. \tag{8}$$

Damit haben wir bewiesen, daß die Isomorphie eine Äquivalenzrelation in jeder Menge von Gruppen ist und daher eine Einteilung dieser Gruppen in disjunkte Klassen vermittelt (vgl. MfL, Bd. 1, 2.5.). Eine solche Klasse isomorpher Gruppen wird *abstrakte Gruppe* genannt. Die einzelnen Gruppen einer Klasse unterscheiden sich zwar durch die Bezeichnung und Bedeutung ihrer Elemente, stellen aber sämtlich Realisierungen desselben abstrakten Rechenschemas dar. Das bedeutet für iso-

morphe endliche Gruppen insbesondere, daß sie (bis auf die Bezeichnung) übereinstimmende Gruppentafeln besitzen.

In der Gruppentheorie untersucht man hauptsächlich solche abstrakten Gruppen, d. h., man versucht, allein aus dem abstrakten Rechenschema Aussagen zu gewinnen, die dann für sämtliche Gruppen dieser Klasse gelten. Bei einer solchen Betrachtung sieht man zwei isomorphe Gruppen als nicht wesentlich verschieden an. Untersucht man jedoch eine Gruppe G, die zwei isomorphe Untergruppen U_1, U_2 enthält, so wird man U_1 und U_2 als Untergruppen von G sehr wohl zu unterscheiden haben. Beispielsweise ist in der Gruppe B_{2Q} (vgl. 12.2.2.) die Untergruppe $\langle r^2 \rangle$ zur Untergruppe $\langle s \rangle$ isomorph, aber $\langle r^2 \rangle$ bildet das Zentrum von B_{2Q}, während die Elemente s und r nicht vertauschbar sind.

Die wichtigste Aufgabe der Gruppentheorie, das *Strukturproblem*, besteht darin, jede Klasse isomorpher Gruppen so genau zu beschreiben, daß man die Rechnung in allen Gruppen der Klasse vollständig beherrscht. Diese Aufgabe ist bisher nur für wenige spezielle Gruppentypen, z. B. für die endlichen abelschen Gruppen befriedigend gelöst worden. Wir werden hier als ein Beispiel die zyklischen Gruppen betrachten.

12.3.3. Sind f und g Isomorphismen von der Gruppe G auf die Gruppe $\overline{G}$, so ist die durch Nacheinanderausführung gewonnene Abbildung $g^{-1} \circ f$ ein Isomorphismus von G auf G. Ist h ein Isomorphismus von G auf sich und f ein Isomorphismus von G auf $\overline{G}$, so ist $f \circ h$ ein Isomorphismus von G auf $\overline{G}$. Zwei Isomorphismen von G auf $\overline{G}$ unterscheiden sich also nur durch einen Isomorphismus von G auf sich. Daher kann man sämtliche Isomorphismen von G auf $\overline{G}$ aus einem einzigen erhalten, wenn man alle Isomorphismen von G auf sich kennt.

Definition 2. Bezeichnet G eine Gruppe, so heißt

$$f \textit{ Automorphismus} \text{ von } G :\Leftrightarrow f \text{ Isomorphismus von } G \text{ auf } G.$$

Führt man zwei Automorphismen der Gruppe G nacheinander aus, so erhält man wieder einen Automorphismus von G. Für die Nacheinanderausführung gilt das Assoziativgesetz (vgl. MfL, Bd. 1, 2.4.), die identische Abbildung von G ist ein Automorphismus, und zu jedem Automorphismus f ist auch die inverse Abbildung f^{-1} ein Automorphismus. Daher bilden die Automorphismen einer Gruppe G bezüglich der Nacheinanderausführung eine Gruppe, die *Automorphismengruppe* von G.

Die Automorphismengruppe der additiven Gruppe Z_3 der Restklassen mod 3 hat die Ordnung 2, denn $e\colon [x] \mapsto [x]$ und $f\colon [x] \mapsto [3 - x]$ $(x = 0, 1, 2)$ sind die Automorphismen von Z_3. Zu jeder mit Z_3 isomorphen Gruppe G gibt es also zwei Isomorphismen von Z_3 auf G.

Ist g ein festes Element der Gruppe G und ordnet man jedem $a \in G$ das Element $g^{-1}ag$ zu, so erhält man eine Abbildung von G auf sich, die eindeutig umkehrbar ist, denn

$$g^{-1}ag = g^{-1}\bar{a}g \Rightarrow a = \bar{a}\,.$$

Weil

$$g^{-1}(ab)\,g = g^{-1}agg^{-1}bg = (g^{-1}ag)\,(g^{-1}bg)$$

gilt, ist die Abbildung sogar ein Automorphismus von G. Er wird der *durch g erzeugte innere Automorphismus von G* genannt.

In abelschen Gruppen erzeugt jedes Element den identischen Automorphismus. In der Gruppe S_3 erzeugen verschiedene Elemente auch verschiedene innere Automorphismen, und jeder Automorphismus der S_3 ist ein innerer Automorphismus (Übungsaufgabe).

Besitzt die Gruppe G einen inneren Automorphismus, welcher $a \in G$ auf $b \in G$ abbildet, d. h., gibt es ein $g \in G$ mit $g^{-1}ag = b$, so heißt b *zu a konjugiert* (vgl. 12.2., Definition 3). Diese Beziehung ist eine Äquivalenzrelation in G und vermittelt daher eine Zerlegung von G in *Klassen konjugierter Elemente*.

Als *Normalisator* des Komplexes K der Gruppe G bezeichnet man

$$N_G(K) := \{g : g \in G \wedge g^{-1}Kg = K\}.$$

$N_G(K)$ ist eine Untergruppe von G. Da für zwei Elemente p, q von G gilt:

$$p^{-1}Kp = q^{-1}Kq \Leftrightarrow pq^{-1} \in N_G(K),$$

ist in endlichen Gruppen G die Anzahl der verschiedenen zu K konjugierten Komplexe gleich $[G : N_G(K)]$, also ein Teiler der Ordnung von G. Insbesondere ist also die Zahl der zum Element $a \in G$ konjugierten Elemente gleich $[G : N_G(a)]$ und die Zahl der zur Untergruppe $U \subseteq G$ konjugierten Untergruppen gleich $[G : N_G(U)]$.

Eine besondere Rolle spielen solche Untergruppen N von G, die bei allen inneren Automorphismen von G auf sich abgebildet werden. Für sie ist $N_G(N) = G$.

Definition 3. G bezeichne eine Gruppe.

N heißt *Normalteiler* (oder *invariante Untergruppe*) von $G :\Leftrightarrow N$ ist Untergruppe von $G \wedge \bigwedge_{g \in G} g^{-1}Ng = N$. (9)

Jede Gruppe G enthält die *trivialen Normalteiler* G und $\langle e \rangle = E$. Offenbar ist das Zentrum $Z(G)$ Normalteiler von G. Jede Untergruppe N vom Index 2 in G ist Normalteiler von G. Weil nämlich die von N verschiedene Nebenklasse alle nicht in N liegenden Elemente a aus G enthält, ist $aN = Na$ und also $g^{-1}Ng = N$ für alle $g \in G$. Daher ist z. B. die Untergruppe $\langle p \rangle$ Normalteiler der S_3. In einer abelschen Gruppe ist jede Untergruppe Normalteiler. Es gibt jedoch auch nichtabelsche Gruppen, in denen jede Untergruppe Normalteiler ist. Sie heißen *Hamiltonsche Gruppen*.

Ein Beispiel dafür bildet die *Quaternionengruppe* Q. Sie kann aus den Elementen a, b unter Beachtung der definierenden Relationen

$$a^4 = e, \qquad b^2 = a^2, \qquad ba = a^{-1}b$$

erzeugt werden. Eine Realisierung ist die von den Matrizen

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \quad (i^2 = -1)$$

bezüglich der Matrizenmultiplikation erzeugte Gruppe der Ordnung 8 (Übungsaufgabe). Die Quaternionengruppe enthält genau eine Untergruppe der Ordnung 2, nämlich $\langle a^2 \rangle$, sowie drei Untergruppen der Ordnung 4, nämlich $\langle a \rangle$, $\langle b \rangle$, $\langle ab \rangle$. Als Untergruppen vom Index 2 sind sie Normalteiler von Q. $\langle a^2 \rangle$ ist das Zentrum der Quaternionengruppe und daher ebenfalls Normalteiler von Q.

Da aus (9)

$$\bigwedge_{g \in G} gN = Ng \tag{10}$$

folgt, stimmen die Rechtsnebenklassen und Linksnebenklassen nach einem Normalteiler überein.

Bezeichnet $\mathfrak{N}$ eine nichtleere Menge von Normalteilern der Gruppe G, so ergibt sich aus Definition 3, daß

$$D := \bigcap_{N \in \mathfrak{N}} N$$

ebenfalls Normalteiler von G ist (vgl. 12.2.2.).

Das Komplexprodukt zweier Untergruppen einer Gruppe G ist im allgemeinen keine Untergruppe von G, wie man am Beispiel der Gruppe S_3 und ihrer Untergruppen $\langle r \rangle$ und $\langle pr \rangle$ (vgl. 12.2.2.) erkennt. Ist aber U eine Untergruppe und N ein Normalteiler von G, so ist nach (10) $UN = NU$ und UN Untergruppe von G, denn es ist

$$(UN)(UN)^{-1} = (UN)(N^{-1}U^{-1}) = UNU^{-1} = NUU^{-1} = NU = UN$$

(vgl. 12.2.2., Satz 1).

Das Produkt zweier Normalteiler N_1, N_2 von G ist sogar wieder Normalteiler von G, weil für alle $g \in G$

$$g^{-1}(N_1N_2)\, g = (g^{-1} N_1 g)\,(g^{-1} N_2 g) = N_1N_2$$

gilt.

12.4. Zyklische Gruppen

In den Beispielen wurde bereits erklärt

Definition 1. Eine Gruppe G heißt *zyklisch* und wird mit $G = \langle a \rangle$ bezeichnet, wenn ihre Elemente die verschiedenen Potenzen eines festen Elementes $a \in G$ sind.

$G = \langle a \rangle$ ist eine unendliche Gruppe, wenn für verschiedene ganzzahlige Exponenten k_1, k_2 immer $a^{k_1} \neq a^{k_2}$ ist. Die Gruppenelemente sind dann

$$\ldots, a^{-3}, a^{-2}, a^{-1}, a^0 = e, a^1, a^2, a^3, \ldots$$

Da mit ihnen nach der Regel $a^k a^l = a^{k+l}$ $(k, l \in \mathbf{Z})$ gerechnet wird, ist die Abbildung

$$f: a^k \mapsto k \qquad (k \in \mathbf{Z})$$

ein Isomorphismus von G auf die additive Gruppe $(\mathbf{Z}, +)$ der ganzen Zahlen. Sind aber einmal zwei Potenzen von a mit verschiedenen ganzzahligen Exponenten gleich, so gibt es einen minimalen Exponenten $m \in \mathbf{N}^*$, für den $a^m = e$ ist. In diesem Fall ist $G = \langle a \rangle$ eine endliche zyklische Gruppe, die aus den Elementen

$$a^0 = e, a^1, a^2, \ldots, a^{m-1}$$

besteht. Weil mit ihnen unter Beachtung der Relation $a^m = e$ gerechnet wird, ist

$$a^k a^l = \begin{cases} a^{k+l}, & \text{wenn} \quad k + l < m, \\ a^{k+l-m}, & \text{wenn} \quad k + l \geqq m \end{cases}$$

(vgl. 12.2.2.).

Da die Addition der Restklassen $[0], [1], \ldots, [m-1]$ modulo m nach der Regel

$$[k] + [l] = \begin{cases} [k + l], & \text{wenn} \quad k + l < m, \\ [k + l - m], & \text{wenn} \quad k + l \geqq m \end{cases}$$

erfolgt, ist die Abbildung

$$f: a^k \mapsto [k] \qquad (k \in \{0, 1, \ldots, m-1\})$$

ein Isomorphismus von G auf die additive Gruppe $(Z_m, +)$ der Restklassen modulo m. Damit ist bewiesen:

Satz 1. *Für zyklische Gruppen $G = \langle a \rangle$ gilt:*

G hat unendliche Ordnung $\Rightarrow G \cong (\mathbf{Z}, +)$,

G hat endliche Ordnung m $\Rightarrow G \cong (Z_m, +)$.

Daraus ergeben sich sofort die

Folgerungen. *Jede zyklische Gruppe ist abelsch. Zu jeder endlichen Ordnung gibt es bis auf Isomorphie genau eine zyklische Gruppe. Bis auf Isomorphie gibt es genau eine unendliche zyklische Gruppe.*

In einer von $\langle e \rangle$ verschiedenen Untergruppe U der zyklischen Gruppe $G = \langle a \rangle$ sei a^d Potenz von a mit minimalem positiven Exponenten d. Ein solches Element gibt es in U, weil mit jeder Potenz a^k auch a^{-k} in U liegt. Bezeichne a^s ein beliebiges Element aus U. Es gibt ganze Zahlen q und r, so daß $s = qd + r$ mit $0 \leqq r < d$ ist (vgl. MfL, Bd. 1, 3.7.). Aus $a^s = (a^d)^q a^r \in U$ und $a^d \in U$ folgt $a^r \left((a^d)^q\right)^{-1} = a^r \in U$. Wegen der vorausgesetzten Minimalität von d muß dann $r = 0$ und $a^s = (a^d)^q$ sein. Daher ist $U = \langle a^d \rangle$.

Weil jede Zahl $t \in \mathbf{Z}$ eine eindeutige Darstellung der Form

$$t = xd + r \qquad (x \in \mathbf{Z} \text{ und } r \in \{0, 1, 2, \ldots, d-1\})$$

besitzt, kann jedes Element $a^t = (a^d)^x a^r$ aus $G = \langle a \rangle$ auf genau eine Weise als $a^t = ua^r$ mit $u \in U$ geschrieben werden. Daher ist

$$\langle a \rangle = U \cup Ua \cup \cdots \cup Ua^{d-1} \tag{1}$$

die Nebenklassenzerlegung von G nach U (Rechts- und Linksnebenklassen stimmen wegen der Kommutativität von G überein!). Also ist

$$[\langle a \rangle : \langle a^d \rangle] = d.$$

Falls G die endliche Ordnung m besitzt, ist daher der minimale Exponent d von $U = \langle a^d \rangle$ ein Teiler von m (vgl. 12.2.3.).

Umgekehrt gibt es zu jedem vorgegebenen minimalen Exponenten $d \in \mathbf{N}^*$ in der unendlichen zyklischen Gruppe $G = \langle a \rangle$ eine Untergruppe $U = \langle a^d \rangle$. Ist $|\langle a \rangle| = m$ und der minimale Exponent d ein Teiler von m, so bilden die Elemente

$$a^0 = e, a^d, a^{2d}, \ldots, a^{(m-1)d}$$

eine Untergruppe $U = \langle a^d \rangle$ von $\langle a \rangle$.

Durch die minimalen Exponenten d ist die Untergruppe U eindeutig bestimmt. Wir fassen unsere Resultate zusammen.

Satz 2. *Die Untergruppen U einer zyklischen Gruppe $G = \langle a \rangle$ sind zyklisch und von der Form $U = \langle a^d \rangle$ mit $d \in \mathbf{N}$.*

In unendlichen Gruppen $\langle a \rangle$ gibt es zu jedem $d \in \mathbf{N}$ genau eine Untergruppe $\langle a^d \rangle$.

Hat $\langle a \rangle$ die endliche Ordnung m, so ist notwendig $d = 0$ oder d ein positiver Teiler von m, und zu jedem solchen d gibt es genau eine Untergruppe $\langle a^d \rangle$.

Für die von $\langle e \rangle$ verschiedenen Untergruppen ist $[\langle a \rangle : \langle a^d \rangle] = d$.

Da sich in einer endlichen Gruppe der Index und die Ordnung einer Untergruppe gegenseitig bestimmen (vgl. 12.2.3., Satz 2), ergibt sich die

Folgerung. *In einer endlichen zyklischen Gruppe der Ordnung m gibt es zu jedem Teiler t von m genau eine Untergruppe der Ordnung t.*

Abschließend soll untersucht werden, wann zwei Elemente a^{h_1} und a^{h_2} der Gruppe $\langle a \rangle$ dieselbe Untergruppe erzeugen. Es gilt

$$\langle a^{h_1} \rangle = \langle a^{h_2} \rangle \Leftrightarrow \bigvee_{u_1, u_2 \in \mathbf{Z}} (a^{h_1} = a^{u_2 h_2} \wedge a^{h_2} = a^{u_1 h_1}). \tag{2}$$

Ist $\langle a \rangle$ unendliche zyklische Gruppe, so ergibt sich daraus

$$\langle a^{h_1} \rangle = \langle a^{h_2} \rangle \Leftrightarrow \bigvee_{u_1, u_2 \in \mathbf{Z}} (h_1 = u_2 h_2 \wedge h_2 = u_1 h_1).$$

Weil hieraus $u_1 u_2 = 1$ oder $h_1 = h_2 = 0$ folgt, ist

$$\langle a^{h_1} \rangle = \langle a^{h_2} \rangle \Leftrightarrow h_1 = h_2 \vee h_1 = -h_2.$$

Jede von $\langle e\rangle$ verschiedene Untergruppe einer unendlichen zyklischen Gruppe besitzt also genau zwei erzeugende Elemente. Entsprechend gibt es genau zwei Automorphismen von $\langle a\rangle$, die durch $a \mapsto a$ bzw. $a \mapsto a^{-1}$ festgelegt sind.

Hat aber $\langle a\rangle$ die endliche Ordnung m, so kann angenommen werden, daß die Exponenten h_1, h_2 Elemente der Menge $\{0, 1, \ldots, m-1\}$ sind. (2) bedeutet dann

$$\begin{aligned}\langle a^{h_1}\rangle = \langle a^{h_2}\rangle &\Leftrightarrow \bigvee_{u_1,u_2\in Z} (u_2h_2 \equiv h_1 \bmod m \wedge u_1h_1 \equiv h_2 \bmod m)\\ &\Leftrightarrow m \sqcap h_2 \mid h_1 \wedge m \sqcap h_1 \mid h_2\\ &\Leftrightarrow m \sqcap h_2 \mid m \sqcap h_1 \wedge m \sqcap h_1 \mid m \sqcap h_2\\ &\Leftrightarrow m \sqcap h_1 = m \sqcap h_2.\end{aligned}$$

Insbesondere besitzt also jede zyklische Gruppe $\langle a\rangle$ der Ordnung m genau $\varphi(m)$ erzeugende Elemente, nämlich die Potenzen a^h, deren Exponenten mit m den größten gemeinsamen Teiler $m \sqcap h = 1$ besitzen. Zu jedem Teiler t von m liegen genau $\varphi(t)$ Elemente der Ordnung t in $\langle a\rangle$. Läßt man t alle Teiler von m durchlaufen, so erhält man jedes Element der Gruppe genau einmal. Daher ergibt sich über die *Eulersche Funktion* die Aussage

$$\sum_{t\mid m} \varphi(t) = m. \tag{3}$$

Ist $d = m \sqcap h$, so gilt $\langle a^h\rangle = \langle a^d\rangle$, und deshalb ist

$$|\langle a^h\rangle| = \frac{m}{m \sqcap h}. \tag{4}$$

Es gibt genau $\varphi(m)$ Automorphismen der zyklischen Gruppe $\langle a\rangle$ der Ordnung m, die durch

$$a \mapsto a^h \quad \text{mit} \quad m \sqcap h = 1$$

festgelegt sind. Bei der Nacheinanderausführung der beiden durch

$$a \mapsto a^h \quad \text{mit} \quad m \sqcap h = 1 \qquad \text{und} \qquad a \mapsto a^k \quad \text{mit} \quad m \sqcap k = 1$$

bestimmten Automorphismen wird a auf a^{hk} abgebildet. Da es bei der Multiplikation der Exponenten nur auf die Restklassen modulo m ankommt, ist die Automorphismengruppe von $\langle a\rangle$ isomorph zur multiplikativen Gruppe der primen Restklassen modulo m.

12.5. Homomorphie von Gruppen

Verzichtet man auf die bei den Isomorphismen geforderte eindeutige Umkehrbarkeit, so kommt man zu dem sehr wichtigen Begriff der *homomorphen Abbildung*.

Definition 1. Für die Gruppen G und $\overline{G}$ heißt

$$f \textit{ Homomorphismus (oder homomorphe Abbildung) von } G \text{ in } \overline{G} \Leftrightarrow f \text{ Abbildung von } G \text{ in } \overline{G} \wedge \bigwedge_{g_1,g_2 \in G} f(g_1 g_2) = f(g_1)\, f(g_2)$$

ist. Man schreibt

$$G \rightarrow \overline{G} :\Leftrightarrow \text{ein Homomorphismus von } G \text{ in } \overline{G} \text{ existiert.}$$

Sei f ein Homomorphismus von G in $\overline{G}$ und

$$\overline{U} := \left\{ \overline{u} : \overline{u} \in \overline{G} \wedge \bigvee_{g \in G} f(g) = \overline{u} \right\}$$

das Bild von G bei f. Bezeichnen e, $\bar{e}$ die neutralen Elemente von G bzw. $\overline{G}$, so folgt aus $f(e)\, f(e) = f(ee) = f(e)$, daß

$$f(e) = \bar{e} \in \overline{U}.$$

Ist $\overline{u} \in \overline{U}$, so gibt es ein $g \in G$ mit $f(g) = \overline{u}$, und aus $f(g)\, f(g^{-1}) = f(gg^{-1}) = f(e) = \bar{e}$ ergibt sich, daß gilt:

$$\overline{u}^{-1} = [f(g)]^{-1} = f(g^{-1}) \in \overline{U}.$$

Zu $\overline{u}_1, \overline{u}_2 \in \overline{U}$ existieren $g_1, g_2 \in G$, so daß $f(g_1) = \overline{u}_1$ und $f(g_2) = \overline{u}_2$. Dann ist

$$f(g_1 g_2) = f(g_1)\, f(g_2) = \overline{u}_1 \overline{u}_2 \in \overline{U}.$$

Das homomorphe Bild $f(G) = \overline{U}$ der Gruppe G ist also eine Untergruppe von $\overline{G}$. Normalteiler von G werden durch f auf Normalteiler von $\overline{U}$ abgebildet, und ist G abelsch, so auch $\overline{U}$ (vgl. 12.3., Satz 1). Ist $f(G) = \overline{G}$, so heißt f *Homomorphismus von* G *auf* $\overline{G}$.

Beispiele.

1. Die Abbildung $f: z \mapsto |z|$ $(z \in \mathbb{C} \setminus \{0\})$, die jeder von Null verschiedenen komplexen Zahl ihren Betrag zuordnet, ist ein Homomorphismus von $(\mathbb{C} \setminus \{0\}, \cdot)$ in $(\mathbb{R} \setminus \{0\}, \cdot)$ (vgl. MfL, Bd. 2, 7.).

2. Der *Multiplikationssatz für Determinanten* (vgl. 8.2.) besagt, daß die Abbildung $f: A \mapsto \det A$ $(A \in L_n)$ ein Homomorphismus von der Matrizengruppe L_n (vgl. 12.1.2.4.) in die Gruppe der rationalen Zahlen $(\mathbb{Q} \setminus \{0\}, \cdot) = L_1$ ist. Man überzeugt sich leicht davon, daß jedes Element aus L_1 als Bild auftritt. Im Fall $n > 1$ ist diese Abbildung offensichtlich kein Isomorphismus, denn L_n ist dann im Unterschied zu L_1 nicht abelsch.

3. Bildet man von der Gruppe S_3 die erzeugenden Elemente p auf 1, r auf -1 und die übrigen, hieraus als Produkte darstellbaren Elemente auf die entsprechenden

Produkte der Bilder ab, so werden $p^3 = e$ auf 1, r^2 auf 1 sowie rp und p^2r auf -1 abgebildet. Die definierenden Relationen der Gruppe S_3 gehen also in richtige Relationen zwischen den Zahlen 1 und -1 über. Daher ist die Abbildung

$$e \mapsto 1, \qquad p \mapsto 1, \qquad p^2 \mapsto 1,$$
$$r \mapsto -1, \qquad pr \mapsto -1, \qquad p^2r \mapsto -1$$

ein Homomorphismus von der S_3 auf die multiplikative Gruppe der Zahlen 1 und -1.

Sei f ein Homomorphismus von der Gruppe G auf die Gruppe $\bar{G}$ und $\bar{e}$ das neutrale Element von $\bar{G}$. Dann ist

$$N := \{n : n \in G \wedge f(n) = \bar{e}\}$$

ein Normalteiler von G, der *Kern* von f genannt wird. Sind nämlich n_1 und n_2 Elemente von N, so liegt wegen

$$f(n_1 n_2^{-1}) = f(n_1)\, f(n_2^{-1}) = f(n_1)\, [f(n_2)]^{-1} = \bar{e}\bar{e}^{-1} = \bar{e}$$

auch $n_1 n_2^{-1}$ in N. N ist also Untergruppe von G (vgl. 12.2., Satz 1). Weil für beliebige Elemente $g \in G$ und $n \in N$

$$f(g^{-1}ng) = f(g^{-1})\, f(n)\, f(g) = [f(g)]^{-1}\, \bar{e} f(g) = \bar{e}$$

ist, gilt

$$\bigwedge_{g \in G} g^{-1}Ng = N,$$

d. h., N ist Normalteiler von G.

Für Elemente $g_1, g_2 \in G$ ist

$$f(g_1) = f(g_2) \Leftrightarrow f(g_1)\, f(g_2^{-1}) = \bar{e} \Leftrightarrow g_1 g_2^{-1} \in N,$$

d. h., genau solche Elemente aus G, die in derselben Nebenklasse nach dem Kern N von f liegen, haben bei f dasselbe Bild. Insbesondere ist ein Homomorphismus von G auf $\bar{G}$ genau dann ein Isomorphismus, wenn der zugehörige Kern nur aus dem neutralen Element e besteht.

Wir wollen nun zeigen, daß es umgekehrt zu jedem Normalteiler N einer Gruppe G eine Gruppe $\bar{G}$ und einen Homomorphismus f von G auf $\bar{G}$ gibt, dessen Kern N ist.

Die Nebenklassenzerlegung von G nach N sei

$$G = N \cup Na \cup Nb \cup \cdots.$$

Die Menge dieser Nebenklassen bildet bezüglich der Komplexmultiplikation (vgl. 12.2., Definition 2) eine Gruppe. Aus der Normalteilereigenschaft von N folgt nämlich $Na = aN$ für alle $a \in G$. Daher ist das Produkt zweier Nebenklassen

$$(Nb)\,(Nb) = N(aN)\,b = N(Na)\,b = (NN)\,(ab) = Nc$$

gleich derjenigen Nebenklasse Nc, in der das Produkt ab liegt. Für die Komplexmultiplikation dieser Nebenklassen gilt das Assoziativgesetz (vgl. 12.2.). N ist neutrales Element bei der Multiplikation der Nebenklassen, denn für alle $g \in G$ gilt

$$N(Ng) = (NN)\,g = Ng \wedge (Ng)\,N = N(gN) = N(Ng) = Ng.$$

Ist Na eine beliebige Nebenklasse und $a^{-1} \in Nb$, so liegt $aa^{-1} = e$ in $(Na)\,(Nb)$, d. h.

$$(Na)\,(Nb) = N \wedge (Nb)\,(Na) = N.$$

Nb ist also zu Na invers.

Definition 2. Sei N ein Normalteiler der Gruppe G. Die Gruppe der Nebenklassen von G nach N mit der Komplexmultiplikation als Operation wird mit G/N bezeichnet und *Faktorgruppe* von G nach N genannt.

Bilden umgekehrt die Nebenklassen von G nach einer Untergruppe U bezüglich der Komplexmultiplikation eine Gruppe, so ist U Normalteiler von G. Bezeichnet nämlich g ein beliebiges Element von G, so gilt wegen $e \doteqdot eg^{-1}eg \in (Ug^{-1})\,(Ug)$ die Beziehung

$$U(g^{-1}Ug) = (Ug^{-1})\,(Ug) = U,$$

also

$$g^{-1}Ug \subseteq U$$

und daher

$$g^{-1}Ug = U.$$

Ist N ein Normalteiler der Gruppe G und bildet man jedes $g \in G$ auf diejenige Nebenklasse von G nach N ab, in der g liegt, so erhält man eine Abbildung f von G auf $\overline{G} = G/N$. Sie ist sogar ein Homomorphismus, denn aus $f(g_1) = Na$ und $f(g_2) = Nb$ folgt nach der Definition des Produktes zweier Nebenklassen, daß $(Na)\,(Nb) = Nc$ diejenige Nebenklasse ist, in der g_1g_2 liegt. Es ist also

$$f(g_1g_2) = f(g_1)\,f(g_2).$$

Diese Abbildung heißt *natürlicher* (oder *kanonischer*) *Homomorphismus* von G auf G/N.

Bezeichnet f einen Homomorphismus von G auf $\overline{G}$ und N den zugehörigen Kern, so ist

$$G/N \cong \overline{G}.$$

Da nämlich alle Elemente einer Nebenklasse von G nach N dasselbe Bild bei f haben, ist

$$\bar{f}: Na \mapsto f(a) \qquad (Na \in G/N)$$

eine Abbildung von G/N auf $\bar{G}$, die sogar eineindeutig ist, weil Elemente verschiedener Nebenklassen verschiedene Bilder bei f besitzen. Sind Na, Nb beliebige Nebenklassen von G nach N und ist $ab \in Nc$, so gilt

$$\bar{f}(NaNb) = \bar{f}(Nc) = f(ab) = f(a)\,f(b) = \bar{f}(Na)\,\bar{f}(Nb).$$

$\bar{f}$ ist also ein Isomorphismus von G/N auf $\bar{G}$.

Die Ergebnisse fassen wir zusammen zum

Satz 1 (*Homomorphiesatz für Gruppen*). *Durch jede homomorphe Abbildung f einer Gruppe G auf eine Gruppe $\bar{G}$ wird ein Normalteiler N von G bestimmt. Er besteht aus denjenigen Elementen von G, die bei f auf das neutrale Element $\bar{e}$ von $\bar{G}$ abgebildet werden und heißt Kern des Homomorphismus f. Die Faktorgruppe G/N ist isomorph $\bar{G}$. Umgekehrt gibt es zu jedem Normalteiler N von G eine homomorphe Abbildung von G auf G/N, deren Kern N ist.*

In unseren Beispielen besteht der jeweilige Kern aus

1. den komplexen Zahlen z mit dem Betrag 1,
2. den Matrizen $\boldsymbol{A}$ aus L_n, deren Determinante 1 ist,
3. den Elementen e, p, p^2 der zyklischen Untergruppe $\langle p\rangle$.

Besteht der Kern des Homomorphismus f von G auf $\bar{G}$ nur aus dem neutralen Element e von G, so ist f ein Isomorphismus, liegt dagegen die ganze Gruppe G im Kern, so ist f eine Abbildung auf die nur aus dem neutralen Element $\bar{e}$ von $\bar{G}$ bestehende Gruppe $\langle\bar{e}\rangle$.

Jede Faktorgruppe einer zyklischen Gruppe $\langle a\rangle$ ist zyklisch (vgl. 12.4.(1)). Erzeugendes Element der Faktorgruppe $\langle a\rangle/U$ ist Ua.

Bildet man jedes Element g der Gruppe G auf den von g erzeugten inneren Automorphismus ab, so erhält man einen Homomorphismus von G auf die Gruppe der inneren Automorphismen $I(G)$, dessen Kern das Zentrum $Z(G)$ ist (Übungsaufgabe).

Besitzt eine Gruppe G einen nichttrivialen Normalteiler N, so kann man durch Betrachtung der Gruppen N und G/N Aufschlüsse über die Struktur von G erhalten. Diese Gruppen sind meist übersichtlicher gebaut als G selbst; beispielsweise haben sie im endlichen Fall kleinere Ordnungen als G. Von besonderem Interesse sind daher Gruppen G, die keine solche „Zerspaltung" mehr gestatten, die also außer G und $\langle e\rangle$ keine Normalteiler besitzen. Sie heißen *einfache Gruppen*.

Die Gruppen von Primzahlordnung sind einfach, da sie nur triviale Untergruppen enthalten. Umgekehrt ist eine abelsche Gruppe $G \neq \langle e\rangle$ auch nur dann einfach, wenn sie Primzahlordnung hat. Ist nämlich G nicht zyklisch, so erzeugt jedes von e verschiedene Element einen zyklischen Normalteiler, der von G und $\langle e\rangle$ verschieden ist. Wenn aber G zyklisch und nicht von Primzahlordnung ist, so gibt es in G eine nichttriviale invariante Untergruppe.

Beispiele für nichtabelsche einfache Gruppen werden wir noch kennenlernen.

12.6. Kommutatorgruppe, Auflösbarkeit

Für beliebige Elemente a, b einer Gruppe G ist $ab = ba(a^{-1}b^{-1}ab)$. Man nennt

$$[a, b] := a^{-1}b^{-1}ab \text{ den } \textit{Kommutator} \text{ der Elemente } a \text{ und } b.$$

Es gilt

$$[a, b] = e \Leftrightarrow ab = ba.$$

Definition 1. Ist $K := \left\{k\colon k \in G \wedge \bigvee_{a,b \in G} k = a^{-1}b^{-1}ab\right\}$ der Komplex aller Kommutatoren einer Gruppe G, so heißt das Erzeugnis $G' := \langle K\rangle$ *Kommutatorgruppe* von G.

G' beschreibt, grob gesagt, die Abweichung der Gruppe G von der Kommutativität. Es ist

$$G \textit{ abelsche Gruppe} \Leftrightarrow G' = \langle e\rangle.$$

Im anderen Extremfall $G' = G$ wird G *perfekt* genannt.

Da bei einem Automorphismus von G ein Kommutator wieder in einen Kommutator übergeht, wird G' bei jedem Automorphismus von G auf sich abgebildet. Untergruppen mit dieser Eigenschaft heißen *charakteristische Untergruppen*. Da sie insbesondere bei jedem inneren Automorphismus von G auf sich abgebildet werden, sind sie Normalteiler. Nichtabelsche einfache Gruppen sind demnach perfekt.

Satz 1. *Bezeichnet G' die Kommutatorgruppe der Gruppe G, so ist G/G' abelsch, und für jeden Normalteiler N von G mit abelscher Faktorgruppe G/N gilt $G' \subseteq N$.*

Beweis. Sind $G'a$ und $G'b$ beliebige Elemente von G/G', so ist

$$(G'a)\,(G'b) = G'ab = G'(aba^{-1}b^{-1})\,ba = G'ba = (G'b)\,(G'a).$$

Weil G/N abelsch ist, gilt für einen beliebigen Kommutator $a^{-1}b^{-1}ab$ aus G

$$Na^{-1}b^{-1}ab = (Na^{-1})\,(Nb^{-1})\,(Na)\,(Nb) = (Na^{-1})\,(Na)\,(Nb^{-1})\,(Nb) = N.$$

Daher ist $a^{-1}b^{-1}ab \in N$ und also $G' \subseteq N$.

Aus dem Satz ergibt sich sofort, daß die Gruppen $\langle p\rangle$ bzw. $\langle r^2\rangle$ (vgl. 12.2.2.) die Kommutatorgruppen von S_3 bzw. B_{2Q} sind.

G'' bezeichnet die Kommutatorgruppe von G'. Allgemein heißt

$$G^{(i)} := (G^{(i-1)})' \qquad (i = 1, 2, \ldots)$$

die i-te Kommutatorgruppe der Gruppe $G^{(0)} := G$.

Definition 2. Bezeichnet G eine Gruppe, so heißt G *auflösbar* $:\Leftrightarrow \bigvee_{n \in \mathbb{N}} G^{(n)} = \langle e\rangle$.

In diesem Fall wird durch die Kommutatorgruppen

$$G \supset G' \supset \cdots \supset G^{(n)} = \langle e\rangle$$

eine *Kette* abelscher Faktorgruppen

$$G/G',\ G'/G'',\ \ldots,\ G^{(n-1)}/G^{(n)} \cong G^{(n-1)}$$

bestimmt. Insbesondere sind abelsche Gruppen G auflösbar, da bei ihnen schon $G' = \langle e \rangle$ ist. S_3 und B_{2Q} sind nichtabelsche auflösbare Gruppen.

Satz 2. *Untergruppen und Faktorgruppen auflösbarer Gruppen sind auflösbar.*

Beweis. Ist U Untergruppe von G und $G^{(n)} = \langle e \rangle$, so ist sicher auch $U^{(n)} = \langle e \rangle$. Beim natürlichen Homomorphismus von G auf die Faktorgruppe G/N wird ein Kommutator $a^{-1}b^{-1}ab$ auf einen Kommutator $Na^{-1}Nb^{-1}NaNb$ abgebildet. Deshalb ist $(G/N)^{(i)}$ das Bild von $G^{(i)}$ ($i = 0, 1, 2, \ldots$), und aus $G^{(n)} = \langle e \rangle$ ergibt sich, daß $(G/N)^{(n)} = N$ das neutrale Element der Faktorgruppe G/N ist.

Satz 3. *Ist N Normalteiler der Gruppe G und sind N und G/N auflösbar, so ist auch G auflösbar.*

Beweis. Es gibt eine natürliche Zahl m, für die $(G/N)^{(m)} = N$ gilt. Daher ist $G^{(m)} \subseteq N$. Ferner existiert ein $n \in \mathbb{N}$, so daß $N^{(n)} = \langle e \rangle$. Daher ist $G^{(m+n)} = \langle e \rangle$.

Definition 3. Eine Gruppe G mit der Primzahlpotenzordnung p^n ($n \in \mathbb{N}^*$) heißt *p-Gruppe*.

Satz 4. *Jede p-Gruppe G hat ein von $\langle e \rangle$ verschiedenes Zentrum $Z(G)$.*

Beweis. Seien $K_1, \ldots, K_r$ die Klassen konjugierter Elemente von G. Die Anzahl k_i der Elemente in der Klasse K_i ($i = 1, 2, \ldots, r$) ist ein Teiler von $|G|$ (vgl. 12.3.3.), also 1 oder eine Potenz von p. Da jedes Element von G in genau einer Klasse konjugierter Elemente liegt, ist

$$|G| = p^n = k_1 + \cdots + k_r.$$

Eine Klasse enthält genau dann nur ein Element z wenn $g^{-1}zg = z$ für alle $g \in G$ gilt, d. h., wenn $z \in Z(G)$ ist. Da sicher $e \in Z(G)$ ist, hat wenigstens ein k_i den Wert 1. Weil aber $p \mid (k_1 + k_2 + \cdots + k_r)$, gibt es mindestens p Klassen, die nur aus einem Element bestehen, d. h. $|Z(G)| \geqq p$.

Satz 5. *G ist p-Gruppe $\Rightarrow$ G ist auflösbar.*

Beweis durch vollständige Induktion nach dem Exponenten n der Gruppenordnung p^n. Jede Gruppe der Ordnung p ist als zyklische Gruppe abelsch und daher auflösbar. Wir nehmen nun an, daß alle Gruppen der Ordnung p^k ($k \in \{1, 2, \ldots, n-1\}$) auflösbar sind. Sei G eine Gruppe der Ordnung p^n. Dann hat $G/Z(G)$ die Ordnung p^k mit $k \in \{0, 1, \ldots, n-1\}$ und ist also auflösbar. $Z(G)$ ist als abelsche Gruppe auflösbar. Nach Satz 3 ist daher auch G auflösbar.

12.7. Direkte Produkte

Definition 1. Eine Gruppe G heißt genau dann *direktes Produkt* ihrer Untergruppen A und B (Bezeichnung: $G = A \times B$), wenn jedes Element $g \in G$ auf genau eine Weise in der Form

$$g = ab \qquad (a \in A,\ b \in B)$$

dargestellt werden kann und die Elemente aus A mit den Elementen aus B vertauschbar sind, d. h., wenn

$$\bigwedge_{a\in A}\bigwedge_{b\in B} ab = ba$$

gilt.

a heißt *A-Komponente*, b heißt *B-Komponente* von g. Aus $g = ab$ und $\bar{g} = \bar{a}\bar{b}$ $(a, \bar{a} \in A;\ b, \bar{b} \in B)$ folgt

$$g\bar{g} = (ab)(\bar{a}\bar{b}) = a(b\bar{a})\bar{b} = a(\bar{a}b)\bar{b} = (a\bar{a})(b\bar{b}),$$

d. h., man beherrscht die Rechnung in $A \times B$ bereits, wenn man sie in A und B kennt, da komponentenweise gerechnet wird.

Offensichtlich ist im Fall endlicher Gruppen $|A \times B| = |A|\,|B|$.

Das direkte Produkt läßt sich auch durch andere Eigenschaften der direkten Faktoren A und B beschreiben.

Satz 1. *$G = A \times B \Leftrightarrow A, B$ sind Normalteiler von $G \wedge AB = G \wedge A \cap B = \langle e\rangle$.*

Beweis. Sei $G = A \times B$. Wenn $g \in A$, so ist $g = ge$ und wenn $g \in B$, so ist $g = eg$ die einzige Darstellung von g als Produkt eines Elementes aus A mit einem Element aus B. Wenn also $g \in A \cap B$, so ist $g = ee = e$. Für ein beliebiges $g = ab \in G$ gilt

$$g^{-1}Ag = b^{-1}a^{-1}Aab = b^{-1}Ab = b^{-1}bA = A.$$

Analog ergibt sich die Normalteilereigenschaft von B.

Sei nun G das Produkt seiner Normalteiler A, B und $A \cap B = \langle e\rangle$. Hat ein Element $g \in G$ die Darstellungen $g = ab = \bar{a}\bar{b}$ $(a, \bar{a} \in A;\ b, \bar{b} \in B)$, so ist $\bar{a}^{-1}a = \bar{b}b^{-1} \in A \cap B = \langle e\rangle$, also $\bar{a} = a$ und $\bar{b} = b$. Für beliebige Elemente $a \in A$ und $b \in B$ folgt aus der Normalteilereigenschaft von A und B

$$a^{-1}b^{-1}ab \in A \cap B = \langle e\rangle,$$

also $ab = ba$.

In der Gruppe B_{2Q} (vgl. 12.2.2.) ist die Untergruppe $\langle r^2, s\rangle = \langle r^2\rangle \times \langle s\rangle$ direktes Produkt. Die Gruppe $S_3 = \langle p\rangle\langle r\rangle$ ist zwar Produkt ihrer Untergruppen $\langle p\rangle$ und $\langle r\rangle$, aber nicht direktes Produkt, da $\langle r\rangle$ kein Normalteiler von S_3 ist.

Definition 2. Eine Gruppe G heißt genau dann *direktes Produkt* ihrer Untergruppen $A_1, \ldots, A_n$ (Bezeichnung: $G = A_1 \times \cdots \times A_n$) $(n \in \mathbb{N}^*)$, wenn jedes $g \in G$ auf genau eine Weise in der Form

$$g = a_1 \cdots a_n \quad (a_i \in A_i;\ i = 1, \ldots, n)$$

dargestellt werden kann und jedes $a_i \in A_i$ mit jedem $a_j \in A_j$ $(i \neq j)$ vertauschbar ist.

Wir setzen

$$A_i' := A_1 \cdots A_{i-1}A_{i+1} \cdots A_n \qquad (i = 1, \ldots, n).$$

Satz 2. $G = A_1 \times \cdots \times A_n \Leftrightarrow \bigwedge_{i\in\{1,\ldots,n\}} A_i$ *ist Normalteiler von* G
$\wedge\, A_1 \cdots A_n = G \wedge \bigwedge_{i\in\{1,\ldots,n\}} A_i \cap A_i' = \langle e\rangle.$

Der Beweis kann wie im Fall von zwei Faktoren geführt werden.

Wie man unmittelbar aus den homomorphen Abbildungen

$$a_1 \cdots a_{i-1} a_i a_{i+1} \cdots a_n \mapsto a_1 \cdots a_{i-1} a_{i+1} \cdots a_n$$

bzw.

$$a_1 \cdots a_{i-1} a_i a_{i+1} \cdots a_n \mapsto a_i$$

abliest, ist

$$G/A_i \cong A_i' \qquad \text{und} \qquad G/A_i' \cong A_i \qquad (i = 1, \ldots, n).$$

Sind A und B gegebene Gruppen, so bildet die Menge

$$G = \{(a, b) : a \in A \wedge b \in B\}$$

mit der durch $(a, b)\,(\bar{a}, \bar{b}) := (a\bar{a}, b\bar{b})$ für die Elemente (a, b) und $(\bar{a}, \bar{b})$ aus G beschriebenen Operation eine Gruppe. Bezeichnen e_A, e_B die neutralen Elemente von A bzw. B, so sind $\bar{A} := \{(a, e_B) : a \in A\}$ und $\bar{B} := \{(e_A, b) : b \in B\}$ zu A bzw. B isomorphe Untergruppen von G, und es gilt

$$G = \bar{A} \times \bar{B}.$$

Allgemeiner kann man nach der Konstruktion solcher Gruppen G fragen, die zu gegebenen Gruppen U und N isomorphe Untergruppen $\overline{U}$ und $\overline{N}$ derart enthalten, daß $\overline{N}$ Normalteiler von G und $G/\overline{N} \cong \overline{U}$ ist. Eine solche Gruppe G nennt man eine *Erweiterung von N mit U*. Das direkte Produkt ist ein einfaches Beispiel einer derartigen Erweiterung. Andere Lösungen dieser Aufgabe werden in der *Erweiterungstheorie* gegeben.

Mit Hilfe des direkten Produktes können wir die Struktur der endlichen abelschen Gruppen übersichtlich beschreiben. Zur Vereinfachung der Sprechweise wollen wir nachstehend die Gruppe G auch dann als direktes Produkt zyklischer Gruppen bezeichnen, wenn G selbst zyklisch ist, das Produkt also nur einen Faktor besitzt. Mit dieser Festlegung gilt der

Hilfssatz 1. *Jede abelsche p-Gruppe P ist direktes Produkt zyklischer Gruppen.*

Beweis. Wir zeigen, daß es zu jeder Untergruppe U von P, welche direktes Produkt zyklischer Gruppen oder gleich $\langle e\rangle$ ist, im Fall $U \neq P$ ein $a \in P \setminus U$ gibt, für das $\langle U, a\rangle$ direktes Produkt zyklischer Gruppen ist. Weil $P \setminus U \neq \emptyset$ ist, existiert ein $\bar{a} \in P \setminus U$. Unter dessen Potenzen $\bar{a}, \bar{a}^p, \bar{a}^{p^2}, \bar{a}^{p^3}, \ldots$ tritt e auf, und es gibt daher ein minimales $i \in \mathbf{N}$, für das

$$\bar{a}^{p^i} \notin U \wedge \bar{a}^{p^{i+1}} \in U$$

gilt. Für $a := \bar{a}^{p^i} \in P \setminus U$ ist also

$$a \notin U \wedge a^p \in U \tag{1}$$

und

$$|\langle U, a\rangle| = p\,|U|. \tag{2}$$

Nach der Voraussetzung ist

$$U = \langle e\rangle \vee U = \langle b_1\rangle \times \cdots \times \langle b_n\rangle. \tag{3}$$

Ist $a^p = e$, so gilt

$$\langle U, a\rangle = U \times \langle a\rangle.$$

Aus (3) folgt dann, daß $\langle U, a\rangle$ direktes Produkt zyklischer Gruppen ist. Im Fall $a^p \neq e$ ist bei passender Reihenfolge der Faktoren in (3)

$$a^p = b_1^{i_1} \cdots b_t^{i_t} \quad (i_k \in \{1, \ldots, |b_k| - 1\};\ k = 1, 2, \ldots, t;\ 1 \leqq t \leqq s). \tag{4}$$

Man kann sich a so gewählt denken, daß keiner der Exponenten $i_1, \ldots, i_t$ durch p teilbar ist. Wäre z. B. $i_1 = pi_1'$ ($i_1' \in \mathbb{N}^*$), so folgte aus (4)

$$\left(ab_1^{-i_1'}\right)^p = b_2^{i_2} \cdots b_t^{i_t}.$$

Weil

$$ab_1^{-i_1'} \notin U \wedge \left(ab_1^{-i_1'}\right)^p \in U$$

ist, könnte man $ab_1^{-i_1'}$ statt a benutzen.

Da o.B.d.A. $|b_1| \geqq |b_2| \geqq \cdots \geqq |b_t|$ angenommen werden kann, ist nach (4) $|a| = p\,|b_1|$ und also wegen (2) und (3)

$$|\langle U, a\rangle| = |a|\,|b_2| \cdots |b_s|.$$

Weil nach (4) $b_1 \in \langle a, b_2, \ldots, b_t\rangle$, also $\langle U, a\rangle = \langle a, b_2, \ldots, b_s\rangle$ ist, folgt

$$\langle U, a\rangle = \langle a\rangle \times \langle b_2\rangle \times \cdots \times \langle b_s\rangle.$$

Hilfssatz 2. *Hat das Element x der Gruppe G die Ordnung $|x| = mn$, wo m, n teilerfremde natürliche Zahlen bezeichnen, so gibt es eindeutig bestimmte Elemente y und z der Ordnungen $|y| = m$ und $|z| = n$ in G, für die*

$$x = yz = zy$$

gilt. y und z sind Potenzen von x.

Beweis. Aus $m \sqcap n = 1$ folgt die Existenz ganzer Zahlen u, v, für die $um + vn = 1$ ist (vgl. MfL, Bd. 1, 3.7.). Setzt man $y := x^{vn}$ und $z := x^{um}$, so gilt demnach

$$yz = zy = x^{umvn} = x \wedge |y| = m \wedge |z| = n.$$

Sind $\overline{y}, \overline{z} \in G$ und ist

$$\overline{yz} = \overline{zy} = x \wedge |\overline{y}| = m \wedge |\overline{z}| = n,$$

so folgt aus

$$x\overline{y} = \overline{yz}\overline{y} = \overline{y}x \quad \text{und} \quad x\overline{z} = \overline{zy}\overline{z} = \overline{z}x,$$

daß $\overline{y}$ und $\overline{z}$ mit x und daher auch mit $y = x^{vn}$ und $z = x^{um}$ vertauschbar sind. Die Gleichung

$$yz = \overline{y}\overline{z}$$

ergibt, daß

$$w = \overline{y}^{-1}y = \overline{z}z^{-1}$$

ist. Aus der Vertauschbarkeit von y mit $\overline{y}$ folgt

$$w^m = (\overline{y}^{-1}y)^m = (\overline{y}^m)^{-1}y^m = ee = e$$

und aus der Vertauschbarkeit von z mit $\overline{z}$ ebenso $w^n = e$. Dann ist aber auch $w = w^{mu}w^{nv} = e$, d. h.

$$y = \overline{y}, \quad z = \overline{z}.$$

Folgerung. *Hat das Element x der Gruppe G die Ordnung $|x| = p_1^{e_1} \cdots p_r^{e_r}$, wo $p_1, \ldots, p_r$ paarweise verschiedene Primzahlen bezeichnen und $\{e_1, \ldots, e_r\} \subset \mathbb{N}^*$ ist, so gibt es eindeutig bestimmte Elemente $x_1, \ldots, x_r$ der Ordnungen $|x_1| = p_1^{e_1}, \ldots, |x_r| = p_r^{e_r}$ in G, für die*

$$x = x_1 x_2 \cdots x_r$$

gilt. $x_1, \ldots, x_r$ sind Potenzen von x.

Bezeichnet G eine endliche abelsche Gruppe und p einen Primteiler von $|G|$, so ist

$$P := \left\{g : g \in G \wedge \bigvee_{i \in \mathbb{N}} g^{p^i} = e\right\}$$

eine Untergruppe von G, denn sind g_1 und g_2 Elemente von P mit den Ordnungen $|g_1| = p^{i_1}$ und $|g_2| = p^{i_2}$, so hat $g_1 g_2^{-1}$ die Ordnung $p^{\max(i_1, i_2)}$ und liegt also in P. Nach der Beweisführung von Hilfssatz 1 ist P direktes Produkt zyklischer Gruppen von p-Potenzordnung und also $|P| = p^e$ ($e \in \mathbb{N}$). Man nennt P die *p-Sylowgruppe* der abelschen Gruppe G. Sind $p_1, \ldots, p_r$ die sämtlichen verschiedenen Primteiler von $|G|$ und $P_1, \ldots, P_r$ die zugehörigen Sylowgruppen, so ist $P_i \cap P_k = \langle e \rangle$, wenn $i \neq k$, und

$$P_1 \times \cdots \times P_r \subseteq G.$$

Nach der Folgerung aus Hilfssatz 2 ist

$$P_1 \cdots P_r = G.$$

Daher gilt der

Satz 3. *Bezeichnet G eine endliche abelsche Gruppe mit den Sylowgruppen $P_1, \ldots, P_r$, so ist*

$$G = P_1 \times P_2 \times \cdots \times P_r.$$

Aus Satz 3 und Hilfssatz 1 folgt unmittelbar der

Satz 4 (Hauptsatz für endliche abelsche Gruppen). *Jede endliche abelsche Gruppe ist das direkte Produkt zyklischer Gruppen von Primzahlpotenzordnung.*

Damit ist das Strukturproblem für endliche abelsche Gruppen in befriedigender Weise gelöst. Die Gruppen $P_1, \ldots, P_r$ sind durch G eindeutig bestimmt. Weiter gilt der

Satz 5. *Ist P eine endliche abelsche Gruppe von Primzahlpotenzordnung und sind $A_1, \ldots, A_r$, $B_1, \ldots, B_s$ von $\langle e \rangle$ verschiedene zyklische Gruppen, für die*

$$P = A_1 \times \cdots \times A_r = B_1 \times \cdots \times B_s$$

gilt, so ist $r = s$ und bei geeigneter Numerierung $|A_i| = |B_i|$ $(i = 1, \ldots, r)$.

Beweis durch Induktion nach dem Exponenten n der Ordnung p^n von P. Offenbar ist der Satz für Gruppen der Ordnung p richtig. Induktionsannahme: Er gilt für alle Gruppen, deren Ordnung ein echter Teiler von $|P|$ ist.

$$P_p := \{x : x \in P \wedge x^p = e\} \quad \text{und} \quad P^p := \left\{y : \bigvee_{x \in P} x^p = y\right\}$$

sind Untergruppen von P.

Sei $A_i = \langle a_i \rangle$, $|A_i| = p^{e_i}$ $(i = 1, \ldots, r)$ und die Numerierung so gewählt, daß $e_1 \geqq \cdots \geqq e_r$. Der Leser prüft leicht nach, daß folgendes gilt:

$$P_p = \langle a_1^{p^{e_1 - 1}} \rangle \times \cdots \times \langle a_r^{p^{e_r - 1}} \rangle$$

und $|P_p| = p^r$. Ist $e_1 = 1$, so $P^p = \langle e \rangle$. Anderenfalls sei e_m das letzte von 1 verschiedene e_i, d. h. $e_1 \geqq \cdots \geqq e_m > e_{m+1} = \cdots = e_r = 1$. Dann ist

$$P^p = \langle a_1{}^p \rangle \times \cdots \times \langle a_m{}^p \rangle .$$

Sei $B_i = \langle b_i \rangle$, $|B_i| = p^{f_i}$ $(i = 1, \ldots, s)$ und die Numerierung so gewählt, daß $f_1 \geqq \cdots \geqq f_s$. Wie eben erkennt man, daß $|P_p| = p^s$ ist. Daher folgt $r = s$. Im Fall $e_1 = f_1 = 1$ ist der Satz damit bewiesen.

Anderenfalls sei $f_1 \geqq \cdots \geqq f_n > f_{n+1} = \cdots = f_s = 1$. P^p ist dann einerseits direktes Produkt von m zyklischen Gruppen der Ordnungen $p^{e_1-1}, \ldots, p^{e_m-1}$ und andererseits direktes Produkt zyklischer Gruppen der Ordnungen $p^{f_1-1}, \ldots, p^{f_n-1}$. Aus der Induktionsannahme folgt $m = n$ und $e_1 - 1 = f_1 - 1, \ldots, e_m - 1 = f_m - 1$. Hieraus ergibt sich wegen $r = s$: $e_1 = f_1, \ldots, e_r = f_r$.

12.8. Permutationsgruppen

Jede Untergruppe G einer symmetrischen Gruppe S_n (vgl. 12.1.2.6.) wird *Permutationsgruppe des Grades n* genannt. Ihre Elemente, die ja 1-1-Abbildungen einer n-elementigen Menge M auf sich sind, heißen *Permutationen* von M. Besteht G aus Permutationen der Menge $M = \{a_1, \ldots, a_n\}$, $\overline{G}$ aus solchen der Menge $\overline{M} = \{\bar{a}_1, \ldots, \bar{a}_n\}$ und gibt es eine 1-1-Abbildung

$$f: a_i \mapsto \bar{a}_i := f(a_i) = \bar{a}_i \qquad (i = 1, 2, \ldots, n)$$

von M auf $\overline{M}$, für die

$$p := \begin{pmatrix} a_1 & \ldots & a_n \\ p(a_1) & \ldots & p(a_n) \end{pmatrix} \in G \Leftrightarrow \overline{p} := \begin{pmatrix} \bar{a}_1 & \ldots & \bar{a}_n \\ \overline{p}(\bar{a}_1) & \ldots & \overline{p}(\bar{a}_n) \end{pmatrix} = \begin{pmatrix} \overline{a_1} & \ldots & \overline{a_n} \\ \overline{p(a_1)} & \ldots & \overline{p(a_n)} \end{pmatrix} \in \overline{G}$$

gilt, gehen also die Permutationen von $\overline{G}$ aus denjenigen von G durch Umbezeichnung der permutierten Elemente hervor, so heißen G und $\overline{G}$ *ähnlich*. Man zeigt leicht, daß diese Ähnlichkeit eine Äquivalenzrelation ist. Insbesondere gehen konjugierte Permutationen durch Umbenennung der permutierten Elemente auseinander hervor, denn ist

$$t = \begin{pmatrix} a_1 \ldots a_n \\ \overline{a_1} \ldots \overline{a_n} \end{pmatrix}$$

eine feste Permutation von M und $p \in G$, so ist

$$\begin{aligned} t^{-1}pt &= \begin{pmatrix} \overline{a_1} \ldots \overline{a_n} \\ a_1 \ldots a_n \end{pmatrix} \begin{pmatrix} a_1 & \ldots & a_n \\ p(a_1) & \ldots & p(a_n) \end{pmatrix} \begin{pmatrix} a_1 \ldots a_n \\ \overline{a_1} \ldots \overline{a_n} \end{pmatrix} \\ &= \begin{pmatrix} \overline{a_1} \ldots \overline{a_n} \\ a_1 \ldots a_n \end{pmatrix} \begin{pmatrix} a_1 & \ldots & a_n \\ p(a_1) & \ldots & p(a_n) \end{pmatrix} \begin{pmatrix} p(a_1) \ldots p(a_n) \\ \overline{p(a_1)} \ldots \overline{p(a_n)} \end{pmatrix} \\ &= \begin{pmatrix} \overline{a_1} & \ldots & \overline{a_n} \\ \overline{p(a_1)} & \ldots & \overline{p(a_n)} \end{pmatrix} \end{aligned} \tag{1}$$

und daher $t^{-1}Gt$ eine zu G ähnliche Permutationsgruppe. Die vermöge t zu p konjugierte Permutation entsteht also aus p, indem auf die Elemente der beiden Zeilen von p die Permutation t angewendet wird, d. h. die permutierten Elemente gemäß t umbezeichnet werden. Wir werden ähnliche Permutationsgruppen als nicht wesentlich verschieden ansehen und häufig $M = \{1, 2, \ldots, n\}$ wählen.

Es bezeichne G_{a_i} die Untergruppe aller Permutationen von G, die ein festes Element $a_i \in M$ auf sich abbilden. Das Element $a_i \in M$ heißt durch die Permutationsgruppe G mit $a_k \in M$ *verbunden*, wenn es in G eine Permutation p gibt, die a_i auf a_k abbildet. Jedes $a_i \in M$ ist vermöge der identischen Permutation mit sich verbunden. Ist a_i durch die Permutation p mit a_k verbunden, so a_k durch p^{-1} mit a_i. Wird a_k bei der Permutation q auf a_l abgebildet, so a_i bei der Permutation pq in a_l. Das Verbundensein ist also eine Äquivalenzrelation in M. Die zugehörigen Äquivalenzklassen werden *Transitivitätssysteme* genannt. Besteht M nur aus einem Transitivitätssystem, so heißt die Permutationsgruppe G *transitiv*, anderenfalls *intransitiv*. Beispielsweise sind B_{2Q} (vgl. 12.1.2.12.) und deren zyklische Untergruppe $\langle r \rangle$ (vgl. 12.2.2.) transitiv. Die Untergruppe $\langle r^2, s \rangle$ von B_{2Q} ist intransitiv.

In einer transitiven Gruppe G von Permutationen der Menge $M = \{1, 2, \ldots, n\}$ gibt es zu jedem $i \in M$ eine Permutation p_i, bei der 1 auf i abgebildet wird.

$$G = G_1 p_1 \cup G_1 p_2 \cup \cdots \cup G_1 p_n \tag{2}$$

ist die Zerlegung von G in Rechtsnebenklassen nach der Untergruppe G_1 aller Permutationen, welche 1 auf sich abbilden, denn bildet p die 1 auf i ab, so ist $pp_i^{-1} \in G_1$ und $p \in G_1 p_i$. Damit ist bewiesen:

Satz 1. *Ist G eine transitive Permutationsgruppe des Grades n, so gilt*

$$n \mid |G|.$$

Eine transitive Permutationsgruppe G des Grades n heißt *regulär*, wenn $n = |G|$. Zum Beispiel ist die durch

$$\begin{pmatrix} 1 & 2 & \ldots & n-1 & n \\ 2 & 3 & \ldots & n & 1 \end{pmatrix}$$

erzeugte zyklische Permutationsgruppe regulär.

Eine Permutationsgruppe G ist genau dann regulär, wenn es zu jedem Paar permutierter Elemente a_i, a_k eine und nur eine Permutation in G gibt, die a_i auf a_k abbildet. In den Bezeichnungen von (2) besteht $p_i^{-1}G_1p_i$ $(i = 1, 2, \ldots, n)$ aus allen Permutationen von G, die i auf i abbilden. Falls G abelsch ist, gilt $p_i^{-1}G_1p_i = G_1$ $(i = 1, 2, \ldots, n)$, d. h., G_1 besteht nur aus der identischen Permutation. Daher ist eine transitive abelsche Permutationsgruppe regulär.

Unter einer *Darstellung* der Gruppe G durch die Gruppe $\bar{G}$ versteht man eine homomorphe Abbildung von G auf $\bar{G}$. Insbesondere spricht man von Darstellungen, wenn die Bildgruppen aus Matrizen oder Permutationen bestehen.

Sei U eine Untergruppe von G und $[G:U]=n$ endlich. Ist $M=\{Ur_1,\ldots,Ur_n\}$ die Menge der Rechtsnebenklassen von G nach U und bezeichnet a ein Element aus G, so ist

$$p_a=\begin{pmatrix}Ur_1 & \ldots & Ur_n\\ Ur_1a & \ldots & Ur_na\end{pmatrix}$$

eine Permutation von M. Die Abbildung

$$a\mapsto p_a \tag{3}$$

ist ein Homomorphismus von G in die Gruppe aller Permutationen der Menge M. Ist nämlich b ein weiteres Element aus G und

$$p_b=\begin{pmatrix}Ur_1 & \ldots & Ur_n\\ Ur_1b & \ldots & Ur_nb\end{pmatrix},$$

so ist

$$\begin{aligned}p_ap_b&=\begin{pmatrix}Ur_1 & \ldots & Ur_n\\ Ur_1a & \ldots & Ur_na\end{pmatrix}\begin{pmatrix}Ur_1 & \ldots & Ur_n\\ Ur_1b & \ldots & Ur_nb\end{pmatrix}\\&=\begin{pmatrix}Ur_1 & \cdots & Ur_n\\ Ur_1a & \cdots & Ur_na\end{pmatrix}\begin{pmatrix}Ur_1a & \ldots & Ur_na\\ Ur_1ab & \ldots & Ur_nab\end{pmatrix}\\&=\begin{pmatrix}Ur_1 & \ldots & Ur_n\\ Ur_1ab & \ldots & Ur_nab\end{pmatrix}=p_{ab}.\end{aligned}$$

Die Bildgruppe ist transitiv, denn o. B. d. A. ist $Ur_1=U$, und deshalb wird Ur_1 bei der Permutation p_{r_i} auf Ur_i $(i=1,2,\ldots,n)$ abgebildet.

Der Kern N des Homomorphismus (3) ist ein Normalteiler von G, der aus allen Elementen $k\in G$ besteht, für die

$$\bigwedge_{i\in\{1,\ldots,n\}} Ur_ik=Ur_i,$$

d. h.

$$\bigwedge_{i\in\{1,\ldots,n\}} k\in r_i^{-1}Ur_i$$

gilt. Daher ist

$$N=\bigcap_{i\in\{1,\ldots,n\}} r_i^{-1}Ur_i$$

der größte in U enthaltene Normalteiler von G.

Wählt man in der endlichen Gruppe G speziell $U=\langle e\rangle$, so erhält man die sogenannte *reguläre Darstellung* von G durch eine zu G isomorphe transitive Permutationsgruppe. Es gilt also der folgende, schon von Arthur Caley (1821—1895) bewiesene Satz.

Satz 2. *Jede endliche Gruppe läßt sich isomorph durch eine transitive Permutationsgruppe darstellen.*

Genau dann, wenn die Untergruppen U und V in G konjugiert sind, liefern sie bei dem angegebenen Verfahren ähnliche Darstellungen von G.

Sind die von U und V erzeugten Permutationsdarstellungen von G ähnlich, so ist notwendig $[G:U] = [G:V]$. Seien $\{Ur_1, \ldots, Ur_n\}$, $\{Vs_1, \ldots, Vs_n\}$ die Mengen der Rechtsnebenklassen von G nach U bzw. V.

$$\begin{pmatrix} Ur_1 & \ldots & Ur_n \\ Ur_1a & \ldots & Ur_na \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} Vs_1 & \ldots & Vs_n \\ Vs_1a & \ldots & Vs_na \end{pmatrix}$$

seien für alle $a \in G$ bis auf die Bezeichnung der permutierten Elemente dieselbe Permutation. Dann gilt insbesondere für die gleichen $a \in G$

$$Ur_1a = Ur_1, \qquad Vs_1a = Vs_1,$$

d. h.

$$r_1^{-1}Ur_1a = r_1^{-1}Ur_1, \qquad s_1^{-1}Vs_1a = s_1^{-1}Vs_1.$$

Es ist also

$$a \in r_1^{-1}Ur_1 \Leftrightarrow a \in s_1^{-1}Vs_1$$

und daher

$$V = (r_1s_1^{-1})^{-1}\, U(r_1s_1^{-1}).$$

Sind aber U und V konjugiert, ist also $g^{-1}Ug = V$ mit einem geeigneten $g \in G$, so bilden $s_1 := g^{-1}r_1g, \ldots, s_n := g^{-1}r_ng$ ein Repräsentantensystem für die Rechtsnebenklassen von V in G. Für ein festes Element $a \in G$ ist Ur_ia ($i \in \{1, \ldots, n\}$) eine wohlbestimmte Nebenklasse aus der Menge $\{Ur_1, \ldots, Ur_n\}$, die wir mit $Ur_{i'} := Ur_ia$ bezeichnen.

$$\begin{pmatrix} 1 & \ldots & n \\ 1' & \ldots & n' \end{pmatrix}$$

ist dann eine durch a festgelegte Permutation der Menge $\{1, 2, \ldots, n\}$. Weil

$$Vs_ig^{-1}ag = g^{-1}Ugg^{-1}r_igg^{-1}ag = g^{-1}Ur_iag$$
$$= g^{-1}Ur_i{}'g = g^{-1}Ugg^{-1}r_i{}'g = Vs_{i'} \qquad (i = 1, \ldots, n)$$

ist, unterscheiden sich die Permutationen

$$\begin{pmatrix} Ur_1 & \ldots & Ur_n \\ Ur_1a & \ldots & Ur_na \end{pmatrix} = \begin{pmatrix} Ur_1 & \ldots & Ur_n \\ Ur_1{}' & \ldots & Ur_n{}' \end{pmatrix}$$

und

$$\begin{pmatrix} Vs_1 & \ldots & Vs_n \\ Vs_1g^{-1}ag & \ldots & Vs_ng^{-1}ag \end{pmatrix} = \begin{pmatrix} Vs_1 & \ldots & Vs_n \\ Vs_1{}' & \ldots & Vs_n{}' \end{pmatrix}$$

nur durch die Bezeichnung der permutierten Elemente. Die letzte Permutation ist zu

$$\begin{pmatrix} Vs_1 & \ldots & Vs_n \\ Vs_1a & \ldots & Vs_na \end{pmatrix}$$

in der Gruppe aller Permutationen von $\{Vs_1, \ldots, Vs_n\}$ konjugiert, denn sie kann in der Form

$$\begin{pmatrix} Vs_1 & \ldots & Vs_n \\ Vs_1 g & \ldots & Vs_n g \end{pmatrix}^{-1} \begin{pmatrix} Vs_1 & \ldots & Vs_n \\ Vs_1 a & \ldots & Vs_n a \end{pmatrix} \begin{pmatrix} Vs_1 & \ldots & Vs_n \\ Vs_1 g & \ldots & Vs_n g \end{pmatrix}$$

geschrieben werden. Daher werden durch die Untergruppen U und V ähnliche Permutationsdarstellungen von G erzeugt.

Bis auf Ähnlichkeit erhält man durch das angegebene Verfahren alle transitiven Permutationsdarstellungen von G. Sei nämlich Q eine transitive Gruppe von Permutationen der Menge $\{1, \ldots, n\}$,

$$a \mapsto q_a \qquad (a \in G,\ q_a \in Q)$$

eine homomorphe Abbildung von G auf Q und U die Untergruppe aller Elemente u von G, deren q_u die 1 auf sich abbildet. Wegen der Transitivität von Q gibt es Permutationen $q_{r_1}, q_{r_2}, \ldots, q_{r_n}$, die die 1 der Reihe nach auf $1, 2, \ldots, n$ abbilden. $r_1, r_2, \ldots, r_n$ ist ein Repräsentantensystem für die Rechtsnebenklassen von G nach U. Wenn $a \in G$ die Permutation

$$q_a = \begin{pmatrix} 1 & \ldots & n \\ 1' & \ldots & n' \end{pmatrix}$$

zugeordnet ist, bilden die Permutationen, welche den Elementen aus $Ur_i a$ zugeordnet sind, die 1 auf i' ab, und es ist also $Ur_i a = Ur_{i'}$ $(i = 1, 2, \ldots, n)$. Daher stimmt

$$p_a = \begin{pmatrix} Ur_1 & \ldots & Ur_n \\ Ur_1 a & \ldots & Ur_n a \end{pmatrix} = \begin{pmatrix} Ur_1 & \ldots & Ur_n \\ Ur_{1'} & \ldots & Ur_{n'} \end{pmatrix}$$

bis auf die Bezeichnung der permutierten Elemente mit der Permutation q_a überein, und die Behauptung ist bewiesen.

Die bisherigen Ergebnisse geben Veranlassung, sich noch etwas näher mit dem Aufbau der symmetrischen Gruppe S_n zu beschäftigen, die aus allen Permutationen der Menge $M = \{1, 2, \ldots, n\}$ besteht. Da S_1 nur aus einem Element besteht, wollen wir $n > 1$ annehmen. S_n ist transitiv und hat die Ordnung $|S_n| = n!$ (vgl. MfL, Bd. 1, 3.6.(16)).

Unter einem *Zyklus* $(a_1, a_2, \ldots, a_l)$ *der Länge* l $(l \geqq 2)$ versteht man diejenige Permutation, welche a_1 auf a_2, a_2 auf $a_3, \ldots, a_{l-1}$ auf a_l und a_l auf a_1 abbildet. Zum Beispiel ist

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 2 & 6 & 3 & 5 & 1 \end{pmatrix} = (1, 4, 3, 6)$$

ein Zyklus der Länge 4 aus S_6. Jede von der identischen Permutation verschiedene Permutation $p \in S_n$ läßt sich als (eventuell in einen Faktor ausgeartetes) Produkt paarweise elementefremder Zyklen darstellen (vgl. MfL, Bd. 1, 3.6.(19)). Beispielsweise ist

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 10 & 5 & 7 & 8 & 4 & 1 & 3 & 2 & 9 & 6 \end{pmatrix} = (1, 10, 6)\,(2, 5, 4, 8)\,(3, 7). \tag{4}$$

Das Produkt elementefremder Zyklen ist kommutativ. Ferner gilt

$$(a_1, a_2, \ldots, a_l) = (a_2, a_3, \ldots, a_l, a_1) = \cdots = (a_l, a_1, \ldots, a_{l-1}).$$

Die Darstellung einer Permutation als Produkt elementefremder Zyklen ist bis auf die Reihenfolge der Faktoren eindeutig.

Die Ordnung eines Zyklus $(a_1, a_2, \ldots, a_l)$ ist gleich seiner Länge l. Ist

$$p = (a_{11}, a_{12}, \ldots, a_{1l_1})\,(a_{21}, a_{22}, \ldots, a_{2l_2}) \cdots (a_{r1}, a_{r2}, \ldots, a_{rl_r}) \tag{5}$$

die Darstellung der Permutation p durch elementefremde Zyklen, so nennt man $\{l_1, l_2, \ldots, l_r\}$ den *Typus der Permutation* p. Wegen der Vertauschbarkeit elementefremder Zyklen gilt

$$p^m = (a_{11}, a_{12}, \ldots, a_{1l_1})^m\,(a_{21}, a_{22}, \ldots, a_{2l_2})^m \cdots (a_{r1}, a_{r2}, \ldots, a_{rl_r})^m.$$

p^m ist genau dann die identische Permutation, wenn

$$l_1 \mid m \wedge l_2 \mid m \wedge \cdots \wedge l_r \mid m.$$

Daher ist die Ordnung von p gleich dem kleinsten gemeinsamen Vielfachen der Zyklenlängen $l_1, l_2, \ldots, l_r$. Die Permutation im Beispiel (4) hat also die Ordnung $3 \cdot 4 = 12$.

Nach (1) erhält man die zur Permutation $p \in S_n$ vermöge $t \in S_n$ konjugierte Permutation $t^{-1}pt$, indem man in beiden Zeilen von p die Permutation t ausführt. Ist p in der Form (5) als Produkt elementefremder Zyklen dargestellt und

$$t = \begin{pmatrix} 1 & 2 & \ldots & n \\ \bar{1} & \bar{2} & \ldots & \bar{n} \end{pmatrix},$$

so ist

$$(\bar{a}_{11}, \bar{a}_{12}, \ldots, \bar{a}_{1l_1})\,(\bar{a}_{21}, \bar{a}_{22}, \ldots, \bar{a}_{2l_2}) \cdots (\bar{a}_{r1}, \bar{a}_{r2}, \ldots, \bar{a}_{rl_r}) \tag{6}$$

die Darstellung von $t^{-1}pt$ als Produkt elementefremder Zyklen. Umgekehrt sind offenbar zwei Permutationen gleichen Typus, die durch (5) und (6) gegeben sind, vermöge t konjugiert. Zwei Permutationen aus der S_n sind also genau dann konjugiert, wenn sie vom gleichen Typus sind. Beispielsweise ist $p = (1, 3)$ zu $q = (1, 2)$ in der S_3 konjugiert, denn mit

$$t = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$$

ist $t^{-1}pt = q$.

Zyklen der Länge 2 heißen *Transpositionen*. Sind a_1 und a_2 verschiedene permutierte Elemente, so läßt sich die identische Permutation e von S_n darstellen als

$$e = (a_1, a_2)\,(a_1, a_2).$$

Da ferner im Fall $l > 2$

$$(a_1, a_2, \ldots, a_l) = (a_1, a_l)\,(a_2, a_l)\,(a_3, a_l) \cdots (a_{l-1}, a_l)$$

ist, kann jede Permutation aus S_n als ein Produkt von Transpositionen geschrieben werden. Aus

$$(2,3)\,(1,3)\,(2,3)\,(1,3) = (1,2)\,(1,3) \neq (1,3)\,(1,2)$$

entnimmt man, daß es dabei auf die Reihenfolge der Transpositionen ankommt, da diese im allgemeinen nicht elementefremd sind und daß die Darstellung nicht eindeutig ist.

S_n $(n > 2)$ kann bereits von zwei Permutationen, beispielsweise $t = (a_1, a_2)$ und $z = (a_1, a_2, \ldots, a_n)$ erzeugt werden, denn es ist $zt = (a_2, a_3, \ldots, a_n)$ und daher

$$(zt)^{-(k-2)}\, t (zt)^{k-2} = (a_1, a_k) \qquad (k = 2, \ldots, n)$$

sowie

$$(a_1, a_i)\,(a_1, a_k)\,(a_1, a_i) = (a_i, a_k) \qquad (2 \leqq i < k \leqq n).$$

Wir wollen nun zeigen, daß jede Darstellung einer Permutation $p \in S_n$ als Produkt von Transpositionen stets eine gerade oder stets eine ungerade Anzahl von Faktoren enthält. Dazu betrachten wir das Differenzenprodukt

$$\Delta := \prod_{\substack{i,k \in \{1,\ldots,n\} \\ i<k}} (i - k).$$

Ist

$$p = \begin{pmatrix} 1 & 2 & \ldots & n \\ 1' & 2' & \ldots & n' \end{pmatrix} \in S_n,$$

so entsteht

$$\Delta^p := \prod_{\substack{i,k \in \{1,\ldots,n\} \\ i<k}} (i' - k')$$

aus Δ, indem die Zahlen i, k aus $\{1, \ldots, n\}$ durch ihre Bilder i', k' bei p ersetzt werden. Δ und Δ^p stimmen bis auf das Vorzeichen überein: $\Delta^p = \chi(p)\,\Delta$, $\chi(p) \in \{1, -1\}$. Die Abbildung $\chi : p \mapsto \chi(p)$ ist ein Homomorphismus von S_n auf die zyklische Gruppe $\langle -1 \rangle$ der Ordnung 2. Mit einer weiteren Permutation $q \in S_n$ ist nämlich

$$\chi(pq)\,\Delta = \Delta^{pq} = (\Delta^p)^q = \chi(p)\,\Delta^q = \chi(p)\,\chi(q)\,\Delta$$

und also

$$\chi(pq) = \chi(p)\,\chi(q).$$

Der Ausdruck $\chi(p)$ heißt *Charakter der Permutation* p. Der Kern des Homomorphismus ist die *alternierende Gruppe* A_n, deren Elemente die Menge $\{a : a \in S_n \wedge \chi(a) = 1\}$ bilden.[1])

---

[1]) Für jede Permutation $p \in S_n$ gilt offenbar $\chi(p) = \operatorname{sgn} p$ (vgl. 8. 1., S. 114). Daher gehört $a \in S_n$ genau dann zu A_n, wenn a eine gerade Anzahl von Inversionen enthält.

Offensichtlich ist $\chi\big((1,2)\big) = -1$. Bezeichnet (i,k) eine beliebige Transposition aus S_n, t eine Permutation der Form

$$t = \begin{pmatrix} 1 & 2 & \dots \\ i & k & \dots \end{pmatrix} \in S_n,$$

so ist $t^{-1}(1,2)\,t = (i,k)$ und also

$$\begin{aligned} \chi\big((i,k)\big) &= \chi(t^{-1})\,\chi\big((1,2)\big)\,\chi(t) \\ &= \chi(t)^{-1}\,\chi\big((1,2)\big)\,\chi(t) \\ &= \chi\big((1,2)\big) = -1. \end{aligned}$$

Ein Produkt einer geraden Anzahl von Transpositionen liegt daher immer in A_n, ein Produkt einer ungeraden Anzahl von Transpositionen in der einzigen von A_n verschiedenen Nebenklasse $A_n(1,2)$. Weil demnach die Permutationen aus A_n nur durch Produkte einer stets geraden Anzahl, die Permutationen aus $A_n(1,2)$ nur durch Produkte einer stets ungeraden Anzahl von Transpositionen darstellbar sind, heißen die ersten *gerade*, die zweiten *ungerade Permutationen*. S_n $(n > 1)$ enthält ebensoviele gerade wie ungerade Permutationen, nämlich

$$|A_n| = \frac{n!}{2}.$$

Weil S_2 eine zyklische Gruppe der Ordnung 2 ist, gilt $A_2 = S_2' = \langle e \rangle$. S_3 ist nicht abelsch. $S_3' = A_3 = \langle (1,2,3) \rangle$ ergibt $S_3'' = \langle e \rangle$. S_4 enthält den Normalteiler A_4 vom Index 2. Daher ist $S_4' \subseteq A_4$. Aus

$$(a_1, a_2)^{-1}(a_2, a_3)^{-1}(a_1, a_2)\,(a_2, a_3) = (a_1, a_2)\,(a_1, a_3)$$

bzw.

$$(a_1, a_2)^{-1}[(a_1, a_3)\,(a_2, a_4)]^{-1}\,(a_1, a_2)\,[(a_1, a_3)\,(a_2, a_4)] = (a_1, a_2)\,(a_3, a_4)$$

folgt, daß jedes Produkt von zwei Transpositionen aus S_4 in S_4' liegt. Also ist $S_4' = A_4$. A_4 ist einzige Untergruppe vom Index 2 in S_4, denn jede Untergruppe vom Index 2 ist Normalteiler mit abelscher Faktorgruppe und umfaßt deshalb S_4'. Für die Permutationen

$$e, \quad a = (1,2)\,(3,4), \quad b = (1,3)\,(2,4), \quad ab = (1,4)\,(2,3)$$

gelten die Relationen $a^2 = b^2 = e$, $ab = ba$. Sie bilden daher eine abelsche Untergruppe V von A_4, die *Kleinsche Vierergruppe*. Weil V aus e und sämtlichen Permutationen vom Typus $\{2,2\}$ von S_4 besteht, ist V Normalteiler von S_4 und erst recht von A_4. A_4/V ist zyklisch von der Ordnung 3 und daher $S_4'' = A_4' \subseteq V$. Aus $V' = \langle e \rangle$ ergibt sich $S_4''' = A_4'' = \langle e \rangle$. Also gilt:

$$S_n, A_n \textit{ sind auflösbare Gruppen, wenn } n \leqq 4 \textit{ ist.} \tag{7}$$

Seien a_1, a_2, a_3, a_4 paarweise verschiedene Elemente von $\{1, 2, \dots, n\}$. Es ist

$$(a_1, a_2)\,(a_1, a_3) = (a_1, a_2, a_3)$$

und

$$(a_1, a_2)\,(a_3, a_4) = (a_1, a_2)\,(a_1, a_3)\,(a_3, a_1)\,(a_3, a_4) = (a_1, a_2, a_3)\,(a_3, a_1, a_4).$$

Weil jede Permutation aus A_n als Produkt einer geraden Anzahl von Transpositionen darstellbar ist, kann sie also auch als Produkt von Zyklen der Länge 3 (= Dreierzyklen) geschrieben werden. Andererseits ist jeder Dreierzyklus eine gerade Permutation. Daher ist A_n ($n \geqq 3$) das Erzeugnis sämtlicher Dreierzyklen von S_n.

Satz 3. *Die alternierenden Gruppen A_n sind außer für $n = 4$ einfach.*

Beweis. $A_2 = \langle e \rangle$ und A_3 als Gruppe der Ordnung 3 sind einfach. A_4 ist nicht einfach.

Sei nun $n \geqq 5$ und N ein von $\langle e \rangle$ verschiedener Normalteiler von A_n. Dann gibt es in N eine Permutation p von Primzahlordnung. Sie ist als Produkt von lauter elementefremden Zyklen der Länge p darstellbar. Wir zeigen, daß es in N eine Permutation vom Typus $\{3\}$, d. h. einen einzelnen Dreierzyklus gibt.

1. $|p| = 2$. Dann ist

$$p = (a_1, a_2)\,(a_3, a_4) \cdots,$$

wobei die punktierten Stellen auch leer sein können. Weil $t_1 := (a_1, a_2, a_3) \in A_n$, ist auch

$$q := t_1^{-1} p t_1 = (a_2, a_3)\,(a_1, a_4) \cdots \in N \wedge r := pq^{-1} = (a_1, a_3)\,(a_2, a_4) \in N.$$

Weil $n \geqq 5$ ist, enthält A_n die Permutation $t_2 := (a_1, a_5)\,(a_2, a_4)$, und es folgt

$$s := t_2^{-1} r t_2 = (a_5, a_3)\,(a_4, a_2) \in N \quad \text{sowie} \quad rs = (a_1, a_5, a_3) \in N.$$

2. $|p| = 3$. Ist p nicht schon selbst ein Dreierzyklus, so ergibt sich aus

$$p = (a_1, a_2, a_3)\,(a_4, a_5, a_6) \cdots$$

mit $t_1 := (a_1, a_4)\,(a_2, a_3) \in A_n$ und $t_2 := (a_2, a_5)\,(a_3, a_6) \in A_n$, daß auch

$$q := t_1^{-1} p t_1 = (a_4, a_3, a_2)\,(a_1, a_5, a_6) \cdots \in N \wedge r := t_2^{-1} p t_2 = (a_1, a_5, a_6)\,(a_4, a_2, a_3) \cdots \in N$$

und daher $qr^{-1} = (a_2, a_3, a_4) \in N$.

3. $|p| > 3$. Weil $t_1 := (a_2, a_4, a_3) \in A_n$, ist mit

$$p = (a_1, a_2, a_3, a_4, a_5, \ldots) \cdots \in N$$

auch

$$q := t_1^{-1} p t_1 = (a_1, a_4, a_2, a_3, a_5, \ldots) \cdots \in N$$

und daher

$$pq^{-1} = (a_1, a_4, a_3) \in N.$$

Sei $p = (a_1, a_2, a_3)$ ein in N enthaltener Dreierzyklus, $(a_{1'}, a_{2'}, a_{3'})$ ein beliebiger Dreierzyklus aus A_n. Weil sich die Permutationen

$$s = \begin{pmatrix} a_1 & a_2 & a_3 & \ldots \\ a_{1'} & a_{2'} & a_{3'} & \ldots \end{pmatrix} \qquad \text{und} \qquad t = \begin{pmatrix} a_1 & a_2 & a_3 & \ldots \\ a_{2'} & a_{1'} & a_{3'} & \ldots \end{pmatrix}$$

aus S_n genau um eine Transposition unterscheiden, liegt eine in A_n, also ist

$$s^{-1} p s = (a_{1'}, a_{2'}, a_{3'}) \in N \vee t^{-1} p t = (a_{2'}, a_{1'}, a_{3'}) \in N.$$

Weil aber mit $(a_{2'}, a_{1'}, a_{3'}) \in N$ auch

$$(a_{2'}, a_{1'}, a_{3'})\,(a_{2'}, a_{1'}, a_{3'}) = (a_{1'}, a_{2'}, a_{3'}) \in N$$

gilt, ist also jedenfalls $(a_{1'}, a_{2'}, a_{3'}) \in N$. Daher ist $N = A_n$ und folglich A_n für $n \geqq 5$ einfach.

Aus (1, 2, 3) (2, 3, 4) $\neq$ (2, 3, 4) (1, 2, 3) ergibt sich, daß die alternierenden Gruppen A_n ($n \geqq 4$) nicht abelsch sind. Für $n \geqq 5$ folgt dann aus der Einfachheit von A_n: $S_n' = A_n = A_n'$, d. h.,

$$S_n,\ A_n \textit{ sind nicht auflösbar, wenn } n \geqq 5 \textit{ ist,} \tag{8}$$

und A_n ist in diesen Fällen sogar perfekt (vgl. 12.6.).

Zu einer beliebigen Untergruppe U von endlichem Index in G liefert (3) eine Darstellung von G durch eine transitive Permutationsgruppe, deren Ordnung ein Teiler von $[G: U]!$ sein muß. Der Kern N des Homomorphismus ist der größte in U enthaltene Normalteiler von G. Nach dem Homomorphiesatz ist deshalb $[G : N] \mid [G : U]!$. In einer einfachen Gruppe G gibt es demnach keine nichttriviale Untergruppe U, für die $[G : U]! < [G : \langle e \rangle]$ ist. Deshalb enthält die A_5 keine Untergruppen der Indizes 2, 3, 4, d. h. der Ordnungen 30, 20, 15.

Es gibt also nicht notwendig in jeder endlichen Gruppe G zu jedem Teiler t von $|G|$ eine Untergruppe U mit der Ordnung t. Ist aber $t = p^n$ ($n \in \mathbb{N}^*$) eine Potenz der Primzahl p und p^n teilerfremd zu $|G| : p^n$, so enthält G eine Untergruppe P der Ordnung p^n, die *p-Sylowgruppe* von G genannt wird. Zu jedem Exponenten $r \in \{1, \ldots, n\}$ gibt es wenigstens eine Untergruppe der Ordnung p^r in G. Nach dem *Satz von* SYLOW (1872) ist außerdem jede Untergruppe von p-Potenzordnung in einer p-Sylowgruppe von G enthalten und sind je zwei p-Sylowgruppen in G konjugiert.

Für auflösbare endliche Gruppen G bewies P. HALL (1928): Ist der Teiler t von $|G|$ zu $|G| : t$ teilerfremd, so besitzt G eine Untergruppe H der Ordnung t, je zwei Untergruppen der Ordnung t sind in G konjugiert, und jede Untergruppe U von G mit $|U| \mid t$ ist in einer solchen Untergruppe der Ordnung t enthalten. Untergruppen H, deren Ordnungen zu ihrem Index teilerfremd sind, werden *Hallgruppen* von G genannt. Auf die Angabe von Beweisen verzichten wir hier. Der Leser findet sie in Lehrbüchern der Gruppentheorie.

12.9. Endliche Drehgruppen

In diesem Abschnitt verwenden wir einige Begriffe und Sachverhalte aus der Geometrie, die dort ausführlich dargestellt werden (vgl. MfL, Bde. 6 und 7), zunächst unter unmittelbarer Berufung auf die Anschauung.

Sei K die Menge aller Punkte der Oberfläche einer Kugel mit dem Mittelpunkt 0. Wir betrachten alle Drehungen des dreidimensionalen euklidischen Raumes um Achsen durch 0. Ordnet man jedem Punkt des Raumes denjenigen Punkt zu, in welchen er bei einer solchen Drehung übergeführt wird, so erhält man eine eineindeutige Abbildung der Menge aller Punkte des Raumes auf sich, die wir in diesem Abschnitt auch als *Drehung* bezeichnen werden. Insbesondere vermittelt jede solche Drehung (als Abbildung) eine 1—1-Abbildung von K auf sich und ist umgekehrt durch diese bereits vollständig bestimmt. Das Ergebnis zweier nacheinander ausgeführter Drehungen läßt sich auch durch eine einzige Drehung erzielen. Die *identische Drehung* e bildet jeden Punkt von K auf sich ab. Zu jeder Drehung d gibt es eine Drehung d^{-1}, für die das Ergebnis der Nacheinanderausführung von d und d^{-1} sowie von d^{-1} und d die identische Drehung e ist. Bezüglich der Nacheinanderausführung bilden die sämtlichen Drehungen um Achsen durch 0 daher eine

Gruppe, deren Untergruppen *Drehgruppen* genannt werden. Wir wollen alle endlichen Drehgruppen angeben.

Jede von e verschiedene Drehung d läßt genau zwei Punkte von K, nämlich die Schnittpunkte der Drehachse mit der Kugeloberfläche, fest. Sie werden *Pole der Drehung* d genannt. Das Bild $d(\alpha)$ des Punktes $\alpha \in K$ bei der Drehung d werde in diesem Abschnitt mit αd bezeichnet. Bei der Nacheinanderausführung der Drehungen d_1 und d_2 wird α auf $\alpha d_1 d_2 := (\alpha d_1)\, d_2$ abgebildet. Liegen $\alpha, \beta \in K$ einander diametral gegenüber, so auch αd und βd.

Wir betrachten nun eine endliche Drehgruppe G von der Ordnung $n > 1$. Bei den Drehungen aus G treten dann auch nur endlich viele Pole auf. Ist t eine beliebige Drehung aus G und $\alpha \in K$ ein Pol bei G, d. h., gibt es in G eine Drehung $d \neq e$ mit $\alpha d = \alpha$, so ist $\alpha t = \alpha\, dt = \alpha t\, (t^{-1} dt)$ ebenfalls ein Pol bei G. G vermittelt also eine Gruppe von Permutationen der Pole. Entsprechend zerfällt die Menge der Pole in Transitivitätssysteme $P_1, \ldots, P_k$. Die Anzahl der Pole in P_i sei m_i $(i = 1, \ldots, k)$. Alle Drehungen aus G, die einen festen Punkt $\alpha \in P_i$ als Pol haben, bilden zusammen mit e eine Untergruppe Z_i. In G gibt es solche Elemente $t_1, \ldots, t_{m_i}$ $(t_1 \in Z_i)$, daß $\alpha = \alpha t_1, \ldots, \alpha_{m_i}$ alle Pole aus dem Transitivitätssystem P_i sind. Ein beliebiges $t \in G$ bildet jedenfalls α auf einen dieser m_i Pole, etwa αt_j ab. Dann ist $\alpha t t_j^{-1} = \alpha$, also $t t_j^{-1} \in Z_i$ und $t \in Z_i t_j$. Daher ist

$$G = Z_i t_1 \cup \cdots \cup Z_i t_{m_i}$$

die Zerlegung von G in Rechtsnebenklassen nach Z_i. Mit $|Z_i| = n_i$ ergibt sich daraus

$$n = m_i n_i \qquad (i = 1, \ldots, k). \tag{1}$$

Für ein Element $t \in G$ gilt: $\alpha t_j t = \alpha t_j \Leftrightarrow t \in t_j^{-1} Z_i t_j$. Daher tritt jedes αt_j aus P_i bei der gleichen Anzahl von Drehungen aus G als Pol auf, nämlich bei $n_i - 1$ nichtidentischen Drehungen. Weil jede nichtidentische Drehung zwei Pole besitzt und insgesamt $n - 1$ solche Drehungen zu G gehören, ist

$$2(n-1) = \sum_{i=1}^{k} m_i (n_i - 1).$$

Wegen (1) ergibt sich daraus

$$2\left(1 - \frac{1}{n}\right) = \sum_{i=1}^{k} \left(1 - \frac{1}{n_i}\right)$$

und

$$k - 2 + \frac{2}{n} = \sum_{i=1}^{k} \frac{1}{n_i}. \tag{2}$$

Aus dieser Gleichung liest man unmittelbar ab, daß $k \geqq 2$ sein muß. Weil $2 \leqq n_i \leqq n$ $(i = 1, \ldots, k)$ ist, folgt weiter $k - 2 + \frac{2}{n} \leqq \frac{k}{2}$, d. h. $k \leqq 3$. Wir brauchen also nur die Fälle $k = 2$ und $k = 3$ zu betrachten.

Fall 1. $k = 2$. Dann heißt die Gleichung (2)

$$\frac{2}{n} = \frac{1}{n_1} + \frac{1}{n_2}. \tag{3}$$

Weil $n_i \leqq n$ und daher $\frac{1}{n_i} \geqq \frac{1}{n}$ $(i = 1, 2)$ ist, ergibt (3) $n_1 = n_2 = n$. Jedes der beiden Transitivitätssysteme P_1, P_2 enthält genau einen Pol. G besteht aus lauter Drehungen um die Achse durch diese beiden Pole. Unter ihnen gibt es eine Drehung a mit minimalem Drehwinkel φ $(0 < \varphi$

$< 2\pi$). Der Drehwinkel ψ ($0 \leqq \psi < 2\pi$) einer beliebigen Drehung $b \in G$ läßt sich in der Form $\psi = z\varphi + \varrho$ ($z \in \mathbf{N}$, $0 \leqq \varrho < \varphi$) schreiben. Dann ist $ba^{-z} \in G$ eine Drehung um den Winkel ϱ. Wegen der Minimaleigenschaft von a muß $\varrho = 0$, also $b = a^z$ sein. Daher liegt eine zyklische Gruppe $G = \langle a \rangle$ vor, die von der Drehung a um den Winkel $\varphi = \dfrac{2\pi}{n}$ erzeugt wird.

Fall 2. $k = 3$. Gleichung (2) besagt dann

$$1 + \frac{2}{n} = \frac{1}{n_1} + \frac{1}{n_2} + \frac{1}{n_3}. \tag{4}$$

O.B.d.A. kann $2 \leqq n_1 \leqq n_2 \leqq n_3 \leqq n$ angenommen werden. Es muß $n_1 = 2$ und $n_2 \leqq 3$ sein, weil sonst die rechte Seite von (4) höchstens gleich 1 wäre. Daher kann (4) nur die folgenden Lösungen in natürlichen Zahlen besitzen:

a) $n_1 = 2,\ n_2 = 2,\ n_3 = \dfrac{n}{2}$.

Mit $n_1 = 2$, $n_3 = 3$ folgt aus (4)

$$12n_3 = n(6 - n_3),$$

also $3 \leqq n_3 < 6$, was noch die Lösungen

b) $n_1 = 2,\ n_2 = 3,\ n_3 = 3,\ n = 12$;

c) $n_1 = 2,\ n_2 = 3,\ n_3 = 4,\ n = 24$;

d) $n_1 = 2,\ n_2 = 3,\ n_3 = 5,\ n = 60$

ergibt.

Zu a). Das Transitivitätssystem P_3 besteht aus zwei Polen α_1, α_2, die bei jeder Drehung aus G einzeln festbleiben oder miteinander vertauscht werden. α_1 und α_2 liegen einander also diametral gegenüber. Alle Drehungen aus G, die α_1 unverändert lassen, bilden eine Untergruppe Z von G. Sie besteht aus Drehungen um die Achse durch α_1 und α_2. Wie im Fall 1 überlegt man sich, daß Z zyklisch von der Ordnung $m = \dfrac{n}{2}$ ist. $u \in G$ vertausche α_1 und α_2 miteinander. Dann ist

$$G = Z \cup Zu.$$

$Z = \langle a \rangle$ wird von einer Drehung a um die Achse durch α_1 und α_2 mit dem Drehwinkel $\varphi = \dfrac{4\pi}{n}$ erzeugt. Weil die Drehungen zu ($z \in Z$) die Pole α_1 und α_2 miteinander vertauschen, haben sie von diesen verschiedene Pole. Die Drehung $(zu)^2$ hat dieselben Pole wie zu und läßt außerdem α_1 und α_2 fest. Deshalb muß $(zu)^2 = e$ sein. Insbesondere folgt $u^2 = e$ und $u^{-1}au = a^{-1}$.

Wir betrachten ein *reguläres m-Eck* in der zur Achse durch α_1 und α_2 senkrechten Ebene durch 0, dessen Ecken auf der Kugel K liegen. G besteht dann aus allen Drehungen, die dieses „Dieder" in sich überführen und wird deshalb *Diedergruppe* genannt. Die m Drehungen um die Achse durch α_1 und α_2 bilden Z, während Zu alle „Umklappungen" enthält. Ist m gerade, so sind das Drehungen mit dem Winkel π um Achsen durch gegenüberliegende Ecken oder gegenüberliegende Seitenmitten des m-Ecks. Ist m ungerade, so gehen die Achsen dieser Drehungen durch je eine Ecke und die Mitte der gegenüberliegenden Seite des m-Ecks. Das Transitivitätssystem P_1 besteht aus den Ecken des m-Ecks, P_2 aus den Schnittpunkten der Strahlen von 0 durch die Seitenmitten des m-Ecks mit der Kugeloberfläche.

Als abstrakte Gruppen der Ordnung $n = 2m$ können die *Diedergruppen* G beschrieben werden durch zwei erzeugende Elemente a, u und die definierenden Relationen

$$a^m = e,\quad u^2 = e,\quad u^{-1}au = a^{-1}.$$

In den übrigen Fällen ist die Gruppenordnung n eindeutig bestimmt. Jedes Transitivitätssystem P_i besteht aus $\frac{n}{n_i} > 2$ Polen ($i = 1, 2, 3$). Nur die identische Drehung läßt also alle Pole eines Transitivitätssystems fest. Daher ist G isomorph zu den transitiven Permutationsgruppen, die von G in den drei Transitivitätssystemen induziert werden.

Zu b). Zu P_3 gehören vier Pole. Als Permutationsgruppe dieser vier Pole hat G in der Gruppe S_4 aller Permutationen der Pole aus P_3 den Index 2 und ist daher Normalteiler in S_4. Weil aber A_4 der einzige Normalteiler dieser Art in S_4 ist, folgt $G \cong A_4$.

G besteht aus allen Drehungen, die ein K einbeschriebenes reguläres Tetraeder mit sich zur Deckung bringen. Deshalb heißt G auch *Tetraedergruppe*. Die Pole des Transitivitätssystems P_3 sind die vier Ecken des Tetraeders, die diametral gegenüberliegenden Pole bilden das Transitivitätssystem P_2. Die zugehörigen Drehachsen gehen also jeweils durch eine Ecke und den Mittelpunkt der gegenüberliegenden Fläche des Tetraeders. Die von e verschiedenen Drehungen um diese Achsen haben die Ordnung 3. Verbindet man die Mittelpunkte gegenüberliegender Kanten des Tetraeders, so erhält man drei weitere Drehachsen für Drehungen der Ordnung 2. Sie bilden zusammen mit e die *Kleinsche Vierergruppe*. Die zugehörigen Pole liegen im Transitivitätssystem P_1.

Zu c). Im Transitivitätssystem P_2 liegen acht Pole. Da $n_2 = 3$ ist, haben die von e verschiedenen Drehungen, welche einen Pol aus P_2 besitzen, die Ordnung 3. Die Pole in den beiden anderen Transitivitätssystemen gehören zu Drehungen der Ordnung 2 oder 4. Folglich liegen jeweils beide Pole einer Drehung der Ordnung 3 in P_2. Wir fassen die diametral gegenüberliegenden Pole, die zu einer Drehachse gehören, zu einem Polpaar zusammen. Da bei jeder Drehung aus G ein solches Polpaar wieder in ein Polpaar übergeführt wird, induziert G eine Gruppe $\overline{G}$ von Permutationen dieser vier Polpaare, die homomorphes Bild von G ist.

Läßt eine Drehung $u \in G$ alle vier Polpaare fest, so bleiben jeweils die beiden Pole eines Paares einzeln fest oder werden miteinander vertauscht. Dann bleiben sicher bei u^2 alle Pole aus P_2 einzeln fest, und folglich ist $u^2 = e$. Ist nicht schon $u = e$, so hat u die Ordnung 2, und die Pole der Drehung u liegen in P_1 oder P_3. Dann werden bei u also die beiden Pole jedes Paares aus P_2 miteinander vertauscht. Sind $\alpha_1, \alpha_2; \beta_1, \beta_2; \gamma_1, \gamma_2; \delta_1, \delta_2$ die vier Polpaare, so ist in Zyklenschreibweise

$$u = (\alpha_1, \alpha_2)\,(\beta_1, \beta_2)\,(\gamma_1, \gamma_2)\,(\delta_1, \delta_2).$$

Weil bei jeder Drehung aus G ein Polpaar in ein Polpaar übergeführt wird, gilt für jede Drehung $d \in G$

$$d^{-1}ud = u.$$

Insbesondere wäre u mit einer Drehung a von der Ordnung 3 vertauschbar, und deshalb hätte ua die Ordnung 6, während es in G nur Drehungen der Ordnungen 2, 3, 4 gibt. Es muß also $u = e$ sein. Weil bei keiner von e verschiedenen Drehung aus G alle vier Polpaare fest bleiben, ist $G \cong \overline{G}$. Die Gruppe S_4 aller Permutationen der vier Polpaare hat ebenso wie G die Ordnung 24. Daher ist $G \cong S_4$.

Die Gruppe G besteht aus allen Drehungen, die ein K einbeschriebenes reguläres Oktaeder mit sich zur Deckung bringen. Daher nennt man G *Oktaedergruppe*. Verbindet man die Mitten gegenüberliegender Flächen, so erhält man vier Drehachsen, deren Pole das Transitivitätssystem P_2 bilden. Jede Drehung aus G permutiert diese Drehachsen untereinander. Die Verbindungslinien der Mittelpunkte gegenüberliegender Kanten liefern sechs Drehachsen für Drehungen der Ordnung 2. Die zugehörigen Pole bilden das Transitivitätssystem P_1. Schließlich ergeben die Verbindungslinien gegenüberliegender Ecken drei Drehachsen, deren Pole in P_3 liegen. Zu jeder dieser drei Achsen gehört eine zyklische Untergruppe, die von einer Drehung um den Winkel $\frac{\pi}{2}$ erzeugt wird.

Tatsächlich erfahren die vier Verbindungslinien gegenüberliegender Flächenmitten bei den Drehungen aus G die volle Permutationsgruppe S_4, denn bei Drehungen um $\frac{\pi}{2}$ um die Achsen

durch gegenüberliegende Ecken des Oktaeders werden die Verbindungslinien zyklisch vertauscht, während bei Drehungen um π um die Achsen durch gegenüberliegende Kantenmitten zwei solche Verbindungslinien jeweils auf sich abgebildet und die beiden anderen vertauscht werden. S_4 kann aber von einem Zyklus der Länge 4 und einer dieser Transpositionen erzeugt werden (vgl. 12.8.).

Zu d). Die 30 Pole im Transitivitätssystem P_1 gehören sämtlich zu Drehungen der Ordnung 2. Die beiden anderen Transitivitätssysteme bestehen aus Polen von Drehungen der Ordnungen 3 und 5. Daher enthält P_1 mit jedem Pol auch den diametral gegenüberliegenden. Wir fassen beide zu einem Polpaar zusammen. Bei jeder Drehung aus G geht ein Polpaar wieder in ein Polpaar über. G induziert also eine zu G homomorphe Gruppe von Permutationen der 15 Polpaare aus P_1.

Die Drehungen aus G, welche die 15 Polpaare einzeln festlassen, bilden einen Normalteiler H von G. Jedes $h \in H$ vertauscht höchstens die beiden Pole einzelner Polpaare untereinander. Daher ist $h^2 = e$. Wäre $h \neq e$, so hätte diese Drehung die Ordnung 2 und ihre Pole lägen in P_1. Da sie keine weiteren Fixpunkte besitzen kann, vertauscht sie in den 14 übrigen Paaren jeweils beide Pole. Weil die Polpaare aus einem Transitivitätssystem stammen und H Normalteiler von G ist, gäbe es dann in H zu jedem der 15 Polpaare eine Drehung der Ordnung 2, deren Achse durch die Pole dieses Paares bestimmt ist und die die beiden Pole jedes anderen Paares vertauscht. Sind h und h_1 zwei solche Drehungen mit verschiedenen Achsen, dann ist $hh_1 \neq e$ eine Drehung, die 26 von den 30 Polen aus P_1 festläßt. Weil es eine solche Drehung nicht geben kann, ist $H = \langle e \rangle$ und daher die durch G induzierte Gruppe von Permutationen der 15 Polpaare aus P_1 sogar zu G isomorph.

Es sei $\alpha = (\alpha_1, \alpha_2)$ ein Polpaar aus P_1. Da die zu den Polpaaren aus P_1 gehörigen Drehungen die Ordnung 2 haben, gibt es genau eine Drehung $a \neq e$ in G, deren Pole α_1 und α_2 sind. Weil beide Pole im Transitivitätssystem P_1 liegen, gibt es in G eine Drehung b, die α_1 in α_2 überführt. Bei b geht dann notwendig α_2 in α_1 über, b^2 läßt neben α_1 und α_2 auch noch die davon verschiedenen Pole β_1 und β_2 von b fest. Daher ist $b^2 = e$ und $\beta = (\beta_1, \beta_2)$ ein Polpaar aus P_1; $b^{-1}ab$ läßt α_1 und α_2 einzeln fest, d. h. $b^{-1}ab = a$. Daraus folgt

$$\beta_1 a = \beta_1 ba = (\beta_1 a)b,$$

$\beta_1 a$ ist also ein Pol von b. Weil β_1 kein Pol von a ist, muß $\beta_1 a = \beta_2$ sein. Ferner ist $\beta_2 a = \beta_1$. Es vertauscht also b die Pole von a und a die Pole von b.

Es sei $c = ab$. Dann ist $c^2 = e$. Daher liegen die Pole γ_1, γ_2 von c in P_1 und bilden das Polpaar γ. c vertauscht α_1 mit α_2 und β_1 mit β_2. Daher sind γ_1, γ_2 von $\alpha_1, \alpha_2, \beta_1, \beta_2$ verschieden. Aus

$$\gamma_1 a = \gamma_1 ca = (\gamma_1 a)c$$

folgt, daß $\gamma_1 a$ Pol von c ist. Weil aber γ_1 kein Pol von a ist, muß $\gamma_1 a = \gamma_2$ sein. Damit ist $\gamma_2 a = \gamma_1$. Ebenso zeigt man, daß γ_1 und γ_2 bei b vertauscht werden.

Die Drehungen a, b, c bilden zusammen mit e eine zur *Kleinschen Vierergruppe* isomorphe Gruppe $V_{\alpha\beta\gamma}$.

Eine Drehung $d \neq e$ von G, die eines der Polpaare α, β, γ festläßt, liegt in $V_{\alpha\beta\gamma}$ und überführt daher jedes der drei Paare in sich. Bleiben nämlich bei d beide Pole eines Paares einzeln fest, so muß d eine der Drehungen a, b, c sein. Wenn aber d etwa α_1 und α_2 vertauscht, bleiben α_1 und α_2 bei db einzeln fest. Daher ist $db = e$ oder $db = a$, d. h. $d = b$ oder $d = c$.

Da P_1 ein Transitivitätssystem ist, gibt es eine Drehung $s \in G$, die das Polpaar α in β überführt und daher kein Element von $V_{\alpha\beta\gamma}$ ist. Dann gilt $s^{-1}as = b$. $s^{-1}bs$ und $s^{-1}cs$ vertauschen β_1 und β_2 miteinander und liegen daher ebenfalls in $V_{\alpha\beta\gamma}$. Demnach gilt

$$s^{-1}V_{\alpha\beta\gamma}s = V_{\alpha\beta\gamma}.$$

Ist $\beta_1 s = \xi_1$ und $\beta_2 s = \xi_2$, so bleiben ξ_1 und ξ_2 bei $s^{-1}bs$ einzeln fest und sind also die Pole dieser nichtidentischen Drehung aus $V_{\alpha\beta\gamma}$. Daher stimmt (ξ_1, ξ_2) mit einem der Polpaare α oder γ überein. Ebenso erkennt man, daß γ bei s in eines der Paare α oder γ übergeht. Da $s \notin V_{\alpha\beta\gamma}$, läßt s keines

der Polpaare α, β, γ fest und vertauscht sie deshalb zyklisch. Dann ist $s^3 \in V_{\alpha\beta\gamma}$, und weil es in G keine Drehungen der Ordnung 6 gibt, folgt daraus $s^3 = e$.

$T_{\alpha\beta\gamma} = \langle V_{\alpha\beta\gamma}, s \rangle$ hat die Ordnung 12. Man überzeugt sich leicht, daß $T_{\alpha\beta\gamma}$ zur *Tetraedergruppe* isomorph ist, indem man im Fall 2b) die Permutationen der sechs Pole des Transitivitätssystems P_1 betrachtet. Diese sechs Pole bilden drei Polpaare, welche zu Drehungen der Ordnung 2 gehören.

Jede Drehung $g \neq e$ aus G, die ein Polpaar der Menge $\{\alpha, \beta, \gamma\}$ in ein Polpaar derselben Menge überführt, liegt in $T_{\alpha\beta\gamma}$ und permutiert daher die Paare α, β, γ nur untereinander. Geht nämlich $\xi \in \{\alpha, \beta, \gamma\}$ bei g in sich über, so liegt g sogar in $V_{\alpha\beta\gamma}$, geht ξ aber in ein anderes Paar η über, so gibt es ein solches $i \in \{1, 2\}$, daß auch s^i das Paar ξ in η überführt. Dann läßt aber gs^{-i} das Paar ξ fest, liegt also in $V_{\alpha\beta\gamma}$ und es ist $g \in V_{\alpha\beta\gamma} s^i \subset T_{\alpha\beta\gamma}$.

Geometrisch bedeuten die bisherigen Ergebnisse, daß jede der drei Drehachsen, die von den Polpaaren α, β, γ bestimmt werden, auf den beiden anderen senkrecht steht. Die Drehungen aus $T_{\alpha\beta\gamma}$ überführen dieses orthogonale System als Ganzes in sich. Ist g eine nicht in $T_{\alpha\beta\gamma}$ gelegene Drehung von G, so geht das orthogonale System in ein anderes orthogonales System über, das durch drei von α, β, γ verschiedene Polpaare α', β', γ' gegeben wird. Dieses System wird genau durch die Drehungen aus $g^{-1}T_{\alpha\beta\gamma}g = T_{\alpha'\beta'\gamma'}$ als Ganzes in sich übergeführt.

Die 15 Polpaare aus P_1 zerfallen also in fünf Tripel von Polpaaren. Jedes Tripel bestimmt ein orthogonales System von Drehachsen. Jede Drehung aus G läßt ein solches System als Ganzes fest oder überführt es in ein anderes System. Daher induziert G eine zu G homomorphe Gruppe von Permutationen dieser fünf orthogonalen Systeme.

Der Kern des Homomorphismus ist der Durchschnitt aller unter G zu $T_{\alpha\beta\gamma}$ konjugierten Gruppen. Bezeichnen α, β, γ und α', β', γ' zwei verschiedene Tripel von Polpaaren, so ist $V_{\alpha\beta\gamma} \cap V_{\alpha'\beta'\gamma'} = \langle e \rangle$. Daher kann der Durchschnitt aller zu $T_{\alpha\beta\gamma}$ konjugierten Gruppen nur noch die Ordnung 3 oder 1 haben. Hätte er die Ordnung 3, so wäre $T_{\alpha\beta\gamma}$ direktes Produkt abelscher Gruppen der Ordnungen 3 und 4 und daher abelsch. Das stimmt aber nicht, da bereits gezeigt wurde, daß $s^{-1}as = b \neq a$ ist.

G ist also einer Untergruppe der symmetrischen Gruppe S_5 aller Permutationen der fünf orthogonalen Systeme isomorph. Weil $|G| = 60$ ist, hat diese Untergruppe den Index 2 in S_5 und muß daher deren einziger nichttrivialer Normalteiler, nämlich die alternierende Gruppe A_5 sein. Folglich ist $G \cong A_5$.

Die Gruppe G besteht aus allen Drehungen, die ein K einbeschriebenes reguläres Ikosaeder mit sich zur Deckung bringen. Daher nennt man G die *Ikosaedergruppe*. Die Verbindungslinien der Mittelpunkte gegenüberliegender Kanten bestimmen 15 Drehachsen, deren zugehörige Pole das Transitivitätssystem P_1 bilden. Jede Drehung aus G permutiert die aus je drei dieser Achsen gebildeten fünf orthogonalen Systeme untereinander. Die zehn Drehachsen zu Drehungen der Ordnung 3, deren Pole in P_2 liegen, führen durch die Mitten gegenüberliegender Flächen. Die Verbindungslinien gegenüberliegender Ecken legen sechs Drehachsen für Drehungen der Ordnung 5 fest, deren Pole P_3 ausmachen.

Die Mittelpunkte der Ikosaederflächen können als die 20 Ecken eines regulären Dodekaeders aufgefaßt werden, das bei den Drehungen der Ikosaedergruppe ebenfalls in sich übergeführt wird. Entsprechend sind die acht Flächenmittelpunkte des regulären Oktaeders die Ecken eines Würfels, der genau bei den Drehungen der Oktaedergruppe in sich übergeht. Konstruiert man in entsprechender Weise zu einem regulären Tetraeder den „dualen" regelmäßigen Körper, so erhält man erneut ein reguläres Tetraeder.

12.10. Übungsaufgaben

1. Man prüfe nach, daß die Menge der Restklassen modulo 10 bezüglich der Restklassenaddition und Restklassenmultiplikation einen kommutativen Ring bildet.
2. Man bestimme alle abstrakten Gruppen der Ordnungen 4 und 6.
3. Man zeige, daß

$$K := \{(a, b) : a \in G \wedge b \in G \wedge \bigvee_{g \in G} g^{-1}ag = b\}$$

eine Äquivalenzrelation in der Menge der Elemente einer Gruppe G ist.

4. $\mathfrak{U}$ sei die Menge aller Untergruppen einer gegebenen Gruppe G. Durch

$$U \wedge V := U \cap V \quad \text{und} \quad U \vee V := \langle U \cup V \rangle$$

für beliebige U, V aus $\mathfrak{U}$ werden in $\mathfrak{U}$ binäre Operationen definiert. Man zeige, daß $(\mathfrak{U}, \wedge, \vee)$ ein Verband ist.
Für die zyklische Gruppe der Ordnung 4 und die Kleinsche Vierergruppe beschreibe man diese Untergruppenverbände durch Angabe

a) der Operationen in Tabellenform,

b) der Untergruppendiagramme.

5. Gilt für jedes Element g der Gruppe G: $g^2 = e$, so ist G abelsch.
6. Man gebe Beispiele für unendliche Gruppen an, in denen

a) jedes Element endliche Ordnung hat,

b) jedes Element $\neq e$ unendliche Ordnung hat,

c) Elemente $\neq e$ von endlicher Ordnung und Elemente unendlicher Ordnung enthalten sind.

7. Man bestimme die Automorphismengruppe der Gruppe $V = \langle a, b \rangle$ mit den definierenden Relationen $a^2 = b^2 = e$ und $ab = ba$.
8. Welche Ordnung hat die Gruppe B_{2S} aller Bewegungen einer euklidischen Ebene, die ein regelmäßiges Sechseck dieser Ebene auf sich abbilden?
Welche natürlichen Zahlen treten als Ordnungen von Elementen dieser Gruppe auf?
Man beschreibe die Elemente der Gruppe durch Angabe der Permutationen, denen die Eckpunkte bei den einzelnen Abbildungen unterworfen werden.
Man gebe mindestens drei verschiedene nichttriviale Untergruppen dieser Gruppe an.
9. Man gebe sämtliche Normalteiler der Gruppe B_{2S} (vgl. Aufgabe 8) an und bestimme (bis auf Isomorphie) alle homomorphen Bilder dieser Gruppe.
10. Man zeige, daß die inneren Automorphismen einer Gruppe G eine Untergruppe $I(G)$ der Automorphismengruppe von G bilden und daß $I(G) \cong G/Z(G)$ ist (dabei bezeichnet $Z(G)$ das Zentrum von G).
11. Man bestimme die Gruppe der inneren Automorphismen der Quaternionengruppe.
12. Man zeige, daß die durch ein festes $n \in \mathbb{N}^*$ bestimmte Abbildung $q \mapsto nq$ $(q \in \mathbb{Q})$ ein Automorphismus der additiven Gruppe $(\mathbb{Q}, +)$ der rationalen Zahlen ist.
13. Man beweise, daß die additive Gruppe $(\mathbb{Q}, +)$ der rationalen Zahlen keine echte Untergruppe von endlichem Index enthält.
14. Eine endliche Folge

$$G = G_0 \supseteq G_1 \supseteq G_2 \supseteq \cdots \supseteq G_l = \langle e \rangle$$

ineinanderliegender Untergruppen der Gruppe G, die mit $\langle e \rangle$ endet, heißt *Normalreihe* von G, wenn jede Untergruppe G_i Normalteiler von G_{i-1} ($i = 1, 2, \ldots, l$) ist. Die G_{i-1}/G_i heißen *Faktorgruppen* der Normalreihe.

Man beweise: G ist auflösbar $\Leftrightarrow$ G besitzt eine Normalreihe mit abelschen Faktorgruppen.

15. Man bestimme die reguläre Darstellung
    a) der Kleinschen Vierergruppe,
    b) der symmetrischen Gruppe S_3.
16. Man gebe eine homomorphe Abbildung der Gruppe B_{2Q} (vgl. 12.1.2.12.) auf die multiplikative Gruppe der Zahlen 1 und -1 an.

13. Ringe, Integritätsbereiche, Körper

13.1. Ringe

Eine nichtleere Menge R mit zwei binären Operationen, die üblicherweise als Addition und Multiplikation bezeichnet werden, heißt *Ring*, wenn $(R, +)$ eine abelsche Gruppe ist, $(R, \cdot)$ eine Halbgruppe und wenn in $(R, +, \cdot)$ die Distributivgesetze gelten (vgl. 11.3.). Wir geben noch einmal die

Definition 1. Eine nichtleere Menge R mit zwei binären Operationen $+, \cdot$ heißt *Ring*, wenn folgende Axiome erfüllt sind:

$$\bigwedge_{a,b,c\in R} (a+b)+c = a+(b+c) \quad (\textit{Assoziativgesetz der Addition}), \tag{1}$$

$$\bigwedge_{a,b\in R} a+b = b+a \quad (\textit{Kommutativgesetz der Addition}), \tag{2}$$

$$\bigwedge_{a,b\in R} \bigvee_{x\in R} a+x = b \quad (\textit{Ausführbarkeit der Subtraktion}), \tag{3}$$

$$\bigwedge_{a,b,c\in R} (ab)\,c = a(bc) \quad (\textit{Assoziativgesetz der Multiplikation}), \tag{4}$$

$$\bigwedge_{a,b,c\in R} a(b+c) = ab+ac, \quad \bigwedge_{a,b,c\in R} (b+c)\,a = ba+ca \ (\textit{Distributivgesetze}). \tag{5}$$

Gilt überdies noch

$$\bigwedge_{a,b\in R} ab = ba \quad (\textit{Kommutativgesetz der Multiplikation}), \tag{6}$$

so heißt R *kommutativer Ring*.

$(R, +)$ nennt man die *additive Gruppe des Ringes* R. Aus der Gruppeneigenschaft von $(R, +)$ folgt sofort, daß das Element x in (3) durch a und b eindeutig bestimmt ist. Es gibt genau ein neutrales Element o in $(R, +)$, für das

$$\bigwedge_{a\in R} a+o = a \tag{7}$$

gilt. o wird *Nullelement des Ringes* genannt und oft auch mit 0 bezeichnet. Die nur aus dem Nullelement bestehende Menge erfüllt mit den Operationen $o + o = o$ und $oo = o$ alle Ringaxiome. Sie bildet den *Nullring*. Wir werden hier im allgemeinen voraussetzen, daß R vom Nullring verschieden ist. Für das Nullelement o jedes Ringes R gilt

$$\bigwedge_{a \in R} ao = oa = o \tag{8}$$

(vgl. 11.3.(15)).

Da $(R, +)$ eine Gruppe ist, gibt es zu jedem $a \in R$ genau ein Element $-a$ in R, so daß $a + (-a) = o$ ist. Für beliebige Elemente a, b, c aus R gelten dann die Regeln

$$a(b - c) = ab - ac, \qquad (b - c)\,a = ba - ca, \tag{9}$$

$$(-a)\,b = a(-b) = -ab, \qquad (-a)\,(-b) = ab \tag{10}$$

(vgl. 11.3.(14)), wobei $x - y := x + (-y)$ bedeutet.

Allein aus den entsprechenden Assoziativgesetzen folgt, daß in R Summen und Produkte von je endlich vielen Elementen aus R unabhängig von der Beklammerung eindeutig definiert sind (vgl. 12.1.1.). Daher kann für eine Summe aus n Summanden $a \in R$ ($n \in \mathbb{N}^*$) kurz

$$na := a + \cdots + a$$

geschrieben werden. Erklärt man noch

$$0a := o, \qquad (-n)\,a := n(-a),$$

so ergibt sich aus 12.1.1., daß für beliebige $a, b \in R$ und $m, n \in \mathbb{Z}$ die Regeln

$$(m + n)\,a = ma + na, \qquad (mn)\,a = m(na)$$

sowie

$$n(a + b) = na + nb$$

gelten.

Bezeichnet

$$a^n := a \cdots a \qquad (a \in R \wedge n \in \mathbb{N}^*)$$

ein Produkt aus n Faktoren a, so gelten für beliebige $a, b \in R$ und $m, n \in \mathbb{N}^*$ die *Potenzgesetze*

$$a^m a^n = a^{m+n} = a^n a^m, \qquad a^{mn} = (a^m)^n$$

sowie in kommutativen Ringen

$$(ab)^n = a^n b^n.$$

Durch Induktion können die distributiven Gesetze auf mehr als zwei Summanden und auf Klammerprodukte ausgedehnt werden. Es ist

$$a(b + c + \cdots + d) = ab + ac + \cdots + ad$$

und

$$(a + b)(c + d) = a(c + d) + b(c + d) = ac + ad + bc + bd$$

$$(a, b, c, \ldots, d \in R).$$

Daraus ergeben sich die bekannten Regeln für das Ausmultiplizieren von Klammern, wobei aber zu beachten ist, daß das Produkt nicht kommutativ zu sein braucht.

Beispiele. Die folgenden Mengen sind bezüglich der in ihnen jeweils erklärten Addition und Multiplikation Ringe.

1. $\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{C}$.

2. Die *geraden ganzen Zahlen* $G = \{x : x \in \mathbf{Z} \wedge 2 \mid x\}$ und allgemeiner alle durch $t \in \mathbf{N}^*$ teilbaren ganzen Zahlen $T = \{x : x \in \mathbf{Z} \wedge t \mid x\}$.

3. Die *Gaußschen ganzen komplexen Zahlen* $\{a + bi : a \in \mathbf{Z} \wedge b \in \mathbf{Z}\}$.

4. Die *Restklassen* modulo m ($m \in \mathbf{N}^*$) (vgl. 12.1.2.9.).

5. Die *quadratischen Matrizen* $\boldsymbol{A} = (a_{\mu\nu})$ einer festen Zeilenzahl $n \in \mathbf{N}^*$ aus rationalen Zahlen.

6. Die *Polynome*

$$p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \qquad (n \in \mathbf{N})$$

mit Koeffizienten $a_0, \ldots, a_n$ aus $\mathbf{Z}$. Dieser Ring wird mit $\mathbf{Z}[x]$ bezeichnet. (Im Abschnitt 14 werden solche *Polynomringe* näher untersucht.)

Sei K ein Körper (vgl. 11.3.(20)). Die Elemente des Vektorraumes K^n (vgl. 9.) sind dann alle Vektoren $\boldsymbol{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ mit Koordinaten $a_1, \ldots, a_n$ aus K. Sie bilden bezüglich der durch

$$\boldsymbol{a} + \boldsymbol{b} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} := \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix}$$

erklären Addition einen Modul (vgl. 11.3.(12)). Außerdem ist durch

$$k\boldsymbol{a} = k \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} := \begin{pmatrix} ka_1 \\ \vdots \\ ka_n \end{pmatrix}$$

eine Verknüpfung der Elemente von K mit denjenigen von K^n definiert, die für beliebige Elemente $k, l \in K$ und $\boldsymbol{a}, \boldsymbol{b} \in K^n$ den Regeln

$$k(\boldsymbol{a} + \boldsymbol{b}) = k\boldsymbol{a} + k\boldsymbol{b},$$

$$(k + l)\boldsymbol{a} = k\boldsymbol{a} + l\boldsymbol{a},$$

$$(kl)\,\boldsymbol{a} = k(l\boldsymbol{a}),$$

$$1\boldsymbol{a} = \boldsymbol{a} \text{ für das Einselement 1 aus } K$$

genügt.

Sei $\boldsymbol{e}_i$ $(i = 1, 2, \ldots, n)$ derjenige Vektor, dessen i-te Koordinate 1 und dessen übrige Koordinaten 0 sind. Jeder Vektor $\boldsymbol{a} \in K^n$ besitzt dann genau eine Darstellung der Form

$$\boldsymbol{a} = a_1\boldsymbol{e}_1 + a_2\boldsymbol{e}_2 + \cdots + a_n\boldsymbol{e}_n,$$

weshalb $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ eine *Basis* von K^n genannt werden.

Man definiert durch

$$\boldsymbol{e}_i\boldsymbol{e}_j := \sum_{k=1}^{n} c_{ijk}\boldsymbol{e}_k$$

mit beliebigen Elementen $c_{ijk} \in K$ $(i, j, k \in \{1, \ldots, n\})$ Produkte der Basisvektoren miteinander. Durch

$$\boldsymbol{a}\boldsymbol{b} = \left(\sum_{i=1}^{n} a_i\boldsymbol{e}_i\right)\left(\sum_{j=1}^{n} b_j\boldsymbol{e}_j\right) := \sum_{i=1}^{n}\sum_{j=1}^{n} a_ib_j\boldsymbol{e}_i\boldsymbol{e}_j = \sum_{k=1}^{n}\left(\sum_{i=1}^{n}\sum_{j=1}^{n} a_ib_jc_{ijk}\right)\boldsymbol{e}_k$$

wird dann ein Produkt in K^n erklärt. Mit der Vektoraddition und dieser Multiplikation erfüllt K^n alle Ringaxiome bis auf (4). Das Assoziativgesetz der Multiplikation gilt genau dann, wenn für beliebige $i, j, k \in \{1, \ldots, n\}$

$$\boldsymbol{e}_i(\boldsymbol{e}_j\boldsymbol{e}_k) = (\boldsymbol{e}_i\boldsymbol{e}_j)\boldsymbol{e}_k$$

ist. Aus

$$\boldsymbol{e}_i(\boldsymbol{e}_j\boldsymbol{e}_k) = \boldsymbol{e}_i\sum_{l=1}^{n} c_{jkl}\boldsymbol{e}_l = \sum_{l=1}^{n} c_{jkl}\boldsymbol{e}_i\boldsymbol{e}_l$$

$$= \sum_{l=1}^{n} c_{jkl}\sum_{m=1}^{n} c_{ilm}\boldsymbol{e}_m = \sum_{m=1}^{n}\left(\sum_{l=1}^{n} c_{jkl}\,c_{ilm}\right)\boldsymbol{e}_m$$

und

$$(\boldsymbol{e}_i\boldsymbol{e}_j)\boldsymbol{e}_k = \sum_{l=1}^{n} c_{ijl}(\boldsymbol{e}_l\boldsymbol{e}_k) = \sum_{l=1}^{n} c_{ijl}\sum_{m=1}^{n} c_{lkm}\boldsymbol{e}_m$$

$$= \sum_{m=1}^{n}\left(\sum_{l=1}^{n} c_{ijl}c_{lkm}\right)\boldsymbol{e}_m$$

folgt, daß das Assoziativgesetz der Multiplikation genau dann gilt, wenn

$$\sum_{l=1}^{n} c_{jkl}c_{ilm} = \sum_{l=1}^{n} c_{ijl}c_{lkm} \quad \text{für alle } i, j, k, m \in \{1, 2, \ldots, n\} \tag{11}$$

ist. Sind diese n^4 Bedingungen erfüllt, so nennt man den erhaltenen Ring $A = (K^n, +, \cdot)$ eine *Algebra* des Ranges n über dem Körper K.

Die Algebren bilden eine umfangreiche Klasse von Ringen. So können die komplexen Zahlen $\mathbf{C}$ als eine Algebra des Ranges 2 über $\mathbf{R}$ mit den Basiselementen $e_1 = 1$, $e_2 = i$ und der Multiplikationsvorschrift

$$e_1^2 = e_1, \quad e_1 e_2 = e_2 e_1 = e_2, \quad e_2^2 = -e_1$$

aufgefaßt werden. Die im Beispiel 5 genannten Matrizenringe sind Algebren des Ranges n^2 über $\mathbf{Q}$. Basiselemente sind die Matrizen $E_{\mu\nu}$ $(\mu, \nu = 1, 2, \ldots, n)$, die im Schnittpunkt der μ-ten Zeile mit der ν-ten Spalte eine 1 und sonst lauter Nullen enthalten. Es ist

$$A = (a_{\mu\nu}) = \sum_{\mu=1}^{n} \sum_{\nu=1}^{n} a_{\mu\nu} E_{\mu\nu}.$$

Die Multiplikation der Basiselemente erfolgt nach den Regeln der Matrizenmultiplikation. Es ist also

$$E_{\varkappa\lambda} E_{\mu\nu} = \begin{cases} E_{\varkappa\nu}, & \text{wenn } \mu = \lambda \\ O, & \text{wenn } \mu \neq \lambda, \end{cases}$$

wobei O die Nullmatrix bezeichnet.

Ein Element e des Ringes R heißt *Einselement* von R, wenn

$$\bigwedge_{a \in R} ae = ea = a \tag{12}$$

ist. Jeder Ring besitzt höchstens ein Einselement (vgl. 11.3.(4)), das wir oft auch mit 1 bezeichnen werden. Beispiel 2 zeigt, daß es Ringe ohne Einselement gibt. Im Nullring ist das Nullelement 0 auch Einselement. In jedem Ring, der nicht nur aus 0 besteht und ein Einselement 1 enthält, ist nach (8) und (12) $0 \neq 1$. Wenn im folgenden Ringe betrachtet werden, die ein Einselement enthalten, sei immer der Nullring ausgeschlossen.

Nicht jeder Ring ist kommutativ, wie man aus Beispiel 5 im Fall $n \geqq 2$ ersehen kann.

Es kann in gewissen Ringen vorkommen, daß ein Produkt gleich dem Nullelement ist, obwohl seine Faktoren vom Nullelement verschieden sind. Im Ring der Restklassen modulo 10 ist beispielsweise $[2]\,[5] = [10] = [0]$.

Definition 2. Bezeichnet 0 das Nullelement des Ringes R und sind $a, b \in R$, so heißt

$$a \textit{ linker } \text{und } b \textit{ rechter Nullteiler} :\Leftrightarrow ab = 0 \wedge a \neq 0 \wedge b \neq 0.$$

Alle Restklassenringe modulo m enthalten Nullteiler, wenn $m \in \mathbf{N}^*$ nicht 1 oder Primzahl ist. Ein weiteres Beispiel liefern die zweireihigen quadratischen Matrizen aus rationalen Zahlen. Nullelement dieses Ringes ist die Nullmatrix $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$. Es ist

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

Eine Teilmenge U eines Ringes R, die bezüglich der Operationen von R bereits selbst einen Ring bildet, für die also insbesondere Summe und Produkt zweier Elemente aus U wieder in U liegen, heißt *Unterring* (oder *Teilring*) von R. Die Menge aller durch t ($t \in \mathbf{N}^*$) teilbaren Zahlen bildet einen Unterring des Ringes $\mathbf{Z}$ aller ganzen Zahlen.

Es kann vorkommen, daß ein Unterring U des Ringes R ein anderes Einselement besitzt als R. So enthält der Ring aller zweireihigen quadratischen Matrizen $\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$ ($a, b \in \mathbf{Z}$) mit dem Einselement $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ den Unterring U aller Matrizen $\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}$ ($a \in \mathbf{Z}$), dessen Einselement $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ ist. R enthält Nullteiler, während U nullteilerfrei ist.

In dem Ring R bildet die Menge $Z(R) := \{z\colon z \in R \wedge \bigwedge_{r \in R} rz = zr\}$ einen Unterring. Sind nämlich $z_1, z_2 \in Z(R)$, also

$$\bigwedge_{r \in R} (rz_1 = z_1 r \wedge rz_2 = z_2 r),$$

so folgt

$$\bigwedge_{r \in R} r(z_1 + z_2) = (z_1 + z_2)\, r$$

und

$$\bigwedge_{r \in R} r(z_1 z_2) = (z_1 z_2)\, r.$$

$Z(R)$ heißt *Zentrum des Ringes* R. Es enthält mindestens das Nullelement von R und stimmt mit R überein, wenn R kommutativ ist.

13.2. Integritätsbereiche, Körper

Wir beweisen folgende Aussage:

In einem Ring R hat jede Gleichung $ax = b$ und $ya = b$ ($a, b \in R \wedge a \neq 0$) genau dann höchstens eine Lösung, wenn R keine Nullteiler enthält.

Besitzt nämlich R keine Nullteiler und sind x_1, x_2 Lösungen von $ax = b$, gilt also $ax_1 = b$ und $ax_2 = b$, so folgt $a(x_1 - x_2) = 0$ und wegen der Nullteilerfreiheit dann $x_1 = x_2$. Enthält R Nullteiler, d. h. Elemente $a \neq 0$ und $c \neq 0$, für die $ac = 0$ ist, so gilt $ax = a(x + c)$ für jedes $x \in R$. Es gibt dann also jedenfalls Gleichungen $ax = b$ ($a, b \in R \wedge a \neq 0$), die in R zwei verschiedene Lösungen besitzen. Für Gleichungen der anderen Form schließt man analog. Im Ring der Restklassen modulo 10 ergibt sich beispielsweise aus $[2]\,[3] = [6]$ und $[2]\,[5] = [0]$, daß $[2]\,x = [6]$ die Lösungen $x = [3]$ und $x = [3] + [5] = [8]$ hat.

Unser Ergebnis besagt, daß in nullteilerfreien Ringen die „*Kürzungsregel*"

$$ax_1 = ax_2 \wedge a \neq 0 \Rightarrow x_1 = x_2, \qquad y_1 a = y_2 a \wedge a \neq 0 \Rightarrow y_1 = y_2 \tag{1}$$

gilt.

Weil die Struktur eines Ringes durch die Existenz von Nullteilern wesentlich beeinflußt wird, kennzeichnet man nullteilerfreie kommutative Ringe mit einem eigenen Namen.

Definition 1. I heißt *Integritätsbereich* $:\Leftrightarrow$ I ist kommutativer Ring $\wedge$ I enthält keine Nullteiler.

Beispiele für Integritätsbereiche sind die Ringe $\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{C}$ und deren Unterringe sowie die Polynomringe $\mathbf{Z}[x]$, $\mathbf{Q}[x]$, $\mathbf{C}[x]$. Der Restklassenring mod m ($m \in \mathbf{N} \setminus \{0, 1\}$) ist offensichtlich genau dann ein Integritätsbereich, wenn m eine Primzahl ist.

Weil in einem Integritätsbereich I mit dem Einselement 1 die Gleichung $ex = 1$ ($e \in I$) höchstens eine Lösung besitzt, kann ein Element $e \in I$ in I höchstens ein (bezüglich der Ringmultiplikation) inverses Element e^{-1} besitzen.

Definition 2. Ist I ein Integritätsbereich mit dem Einselement 1 und $e \in I$, so heißt

$$e \textit{ Einheit} \text{ von } I :\Leftrightarrow \bigvee_{e^{-1} \in I} ee^{-1} = 1.$$

Sind e_1 und e_2 Einheiten von I, so liegt mit den Elementen e_1^{-1} und e_2^{-1} auch $(e_1e_2)^{-1} = e_2^{-1}e_1^{-1}$ in I, es ist

$$(e_1e_2)\,(e_1e_2)^{-1} = e_1(e_2e_2^{-1})\,e_1^{-1} = e_1e_1^{-1} = 1$$

und daher e_1e_2 Einheit. Aus $1 \cdot 1 = 1$ folgt, daß 1 Einheit ist. Aus $ee^{-1} = e^{-1}e$ ergibt sich, daß mit e zusammen auch e^{-1} Einheit ist. Ferner gilt für die Multiplikation das Assoziativgesetz. Also bilden die Einheiten von I eine abelsche Gruppe.

Die Einheiten von $\mathbf{Z}$ sind 1 und -1. In $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{C}$ sind alle von 0 verschiedenen Zahlen Einheiten. Unter den Gaußschen ganzen komplexen Zahlen sind $1, i, -1, -i$ Einheiten, und im Restklassenring modulo einer Primzahl ist jede Restklasse $[a] \neq [0]$ eine Einheit (vgl. 12.2.3., Folgerung 3').

Bezeichnet a eine Einheit und b ein Element des Integritätsbereiches I, so hat die Gleichung $ax = b$ in I die (eindeutig bestimmte) Lösung $x = a^{-1}b$.

Sei R ein kommutativer Ring, der nicht nur aus dem Nullelement besteht. Hat in R jede Gleichung

$$ax = b \qquad (a, b \in R \wedge a \neq 0) \tag{2}$$

genau eine Lösung, so ist R nach der Eingangsbemerkung nullteilerfrei, also Integritätsbereich. Daher ist das Produkt zweier von 0 verschiedener Elemente nicht 0. $(R \setminus \{0\}, \cdot)$ ist also eine abelsche Gruppe. Ist umgekehrt $(R \setminus \{0\}, \cdot)$ eine abelsche Gruppe, so hat jede Gleichung (2) in R genau eine Lösung, die im Fall $b = 0$ selbst 0 ist und sonst in $R \setminus \{0\}$ liegt. Daher ist die in 11.3.(20) gegebene Erklärung gleichwertig mit

Definition 3. K heißt *Körper* $:\Leftrightarrow K$ ist vom Nullring verschiedener kommutativer Ring $\wedge$ jede Gleichung $ax = b$ $(a, b \in K \wedge a \neq 0)$ besitzt in K genau eine Lösung.

Weil für das neutrale Element 1 der Gruppe $(K \setminus \{0\}, \cdot)$

$$\bigwedge_{a \in K \setminus \{0\}} 1a = a1 = a$$

und nach 13.1.(8)

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

gilt, ist es Einselement von K. Zu jedem a aus $K \setminus \{0\}$ gibt es in K genau ein inverses Element a^{-1} mit

$$aa^{-1} = a^{-1}a = 1.$$

Ein Körper ist also ein Integritätsbereich mit Einselement, in dem zu jedem von 0 verschiedenen Element ein inverses Element liegt, also jedes Element $\neq 0$ Einheit ist.

Beispiele für Körper sind $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$.

Satz 1. *Jeder vom Nullring verschiedene Integritätsbereich I mit endlich vielen Elementen ist ein Körper.*

Beweis. Multipliziert man sämtliche Elemente $x_1, \ldots, x_n$ $(n \in \mathbb{N}^*)$ von I mit einem Element $a \neq 0$ von I, so sind nach der Kürzungsregel (1) $ax_1, \ldots, ax_n$ paarweise voneinander verschieden und liefern also wieder sämtliche Elemente von I, darunter das Element b. Die Gleichung $ax = b$ hat also in I eine (und wegen der Nullteilerfreiheit auch nur eine) Lösung.

Folgerung. *Der Restklassenring modulo einer Primzahl p ist ein Körper.*

Verzichtet man auf die Kommutativität des vom Nullring verschiedenen Ringes R, so zeigt sich, daß $(R \setminus \{0\}, \cdot)$ genau dann eine Gruppe ist, wenn jede Gleichung $ax = b$ und jede Gleichung $ya = b$ $(a, b \in R \wedge a \neq 0)$ genau eine Lösung hat. Ringe mit dieser Eigenschaft heißen *Schiefkörper* (vgl. 11.3.(19)). Wie bei den Körpern ergibt sich, daß sie genau ein Einselement und zu jedem Element $a \neq 0$ genau ein inverses Element a^{-1} enthalten.

Man kann zeigen, daß jeder Schiefkörper, der nur endlich viele Elemente enthält, ein Körper ist (Satz von Wedderburn). Ein Beispiel für einen Schiefkörper bilden die *Quaternionen*. Sie sind eine Algebra Q vom Rang 4 über $\mathbb{R}$. Die Multiplikationsvorschrift für die Basiselemente e_i $(i = 1, 2, 3, 4)$ lautet

$$e_1e_i = e_ie_1 = e_i \qquad (i = 1, 2, 3, 4),$$

$$e_1e_1 = e_1, \quad e_ie_i = -e_1 \qquad (i = 2, 3, 4),$$

$$e_2e_3 = e_4 = -e_3e_2, \quad e_3e_4 = e_2 = -e_4e_3, \quad e_4e_2 = e_3 = -e_2e_4.$$

Statt e_1 kann einfach die Zahl 1 geschrieben werden. Ersetzt man die Basiselemente e_2, e_3, e_4 durch die Symbole i, j, k, so erhält man eine übliche Schreibweise für Quaternionen. Es gelten die

Multiplikationsregeln

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

$$ij = k,\ jk = i,\ ki = j,\ ji = -k,\ kj = -i,\ ik = -j.$$

Die Quaternionen sind dann die sämtlichen Ausdrücke der Form

$$\alpha = a_1 + a_2 i + a_3 j + a_4 k \qquad (a_1, a_2, a_3, a_4 \in \mathbf{R}).$$

Ist

$$\beta = b_1 + b_2 i + b_3 j + b_4 k \qquad (b_1, b_2, b_3, b_4 \in \mathbf{R}),$$

so gilt nach den Festlegungen über Algebren

$$\alpha = \beta \Leftrightarrow a_1 = b_1 \wedge a_2 = b_2 \wedge a_3 = b_3 \wedge a_4 = b_4,$$

$$\alpha + \beta = (a_1 + b_1) + (a_2 + b_2)i + (a_3 + b_3)j + (a_4 + b_4)k,$$

$$\begin{aligned}\alpha \cdot \beta = {} & (a_1b_1 - a_2b_2 - a_3b_3 - a_4b_4) + (a_2b_1 + a_1b_2 - a_4b_3 + a_3b_4)i \\ & + (a_3b_1 + a_4b_2 + a_1b_3 - a_2b_4)j + (a_4b_1 - a_3b_2 + a_2b_3 + a_1b_4)k.\end{aligned}$$

Mit den die Quaternionen darstellenden formalen Summen wird also nach den Assoziativ- und Distributivgesetzen gerechnet, wobei die reellen Zahlen mit i, j, k vertauschbar sind und die angegebenen Multiplikationsregeln für i, j, k ausgenutzt werden.

Das Nullelement der Quaternionenalgebra Q ist offenbar $0 = 0 + 0i + 0j + 0k$. Sind $\alpha, \beta \in Q$ und ist $\alpha \neq 0$, so gibt es immer genau ein $\xi = x_1 + x_2 i + x_3 j + x_4 k$ in Q, so daß $\alpha\xi = \beta$ ist. Die Bestimmung eines solchen ξ bedeutet nämlich die Lösung des Gleichungssystems

$$\begin{aligned} a_1x_1 - a_2x_2 - a_3x_3 - a_4x_4 &= b_1, \\ a_2x_1 + a_1x_2 - a_4x_3 + a_3x_4 &= b_2, \\ a_3x_1 + a_4x_2 + a_1x_3 - a_2x_4 &= b_3, \\ a_4x_1 - a_3x_2 + a_2x_3 + a_1x_4 &= b_4. \end{aligned}$$

Da dessen Determinante

$$\begin{vmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{vmatrix} = (a_1^2 + a_2^2 + a_3^2 + a_4^2)^2 \neq 0$$

ist, gibt es genau ein $\xi \in Q$ mit der geforderten Eigenschaft. Analog ergibt sich, daß auch jede Gleichung $\eta\alpha = \beta$ $(\alpha, \beta \in Q \wedge \alpha \neq 0)$ in Q genau eine Lösung hat.

Die Quaterionen bilden also einen Schiefkörper, der sicher kein Körper ist. Sie wurden von Sir William Rowan Hamilton (1805–1865) entdeckt und spielten eine Rolle als Vorläufer unserer heutigen Vektorrechnung. Georg Frobenius (1849–1917) bewies, daß unter allen Algebren über $\mathbf{R}$ bis auf Isomorphie (vgl. 13.3.) $\mathbf{R}$, $\mathbf{C}$ und die Quaternionen die einzigen Schiefkörper sind.

13.3. Isomorphie von Ringen und Körpern

Für Ringe ist der Begriff der Isomorphie von gleicher Bedeutung wie für Gruppen (vgl. 12.3.).

Definition 1. R und $\bar{R}$ seien Ringe. Dann heißt

$$f \textit{ Isomorphismus von } R \textit{ auf } \bar{R} :\Leftrightarrow f \text{ ist 1-1-Abbildung von } R \text{ auf } \bar{R}$$

$$\wedge \bigwedge_{r_1,r_2 \in R} f(r_1 + r_2) = f(r_1) + f(r_2) \quad \wedge \bigwedge_{r_1,r_2 \in R} f(r_1 r_2) = f(r_1)\, f(r_2).$$

Man nennt

$$R \textit{ isomorph } \bar{R} :\Leftrightarrow \text{ein Isomorphismus von } R \text{ auf } \bar{R} \text{ existiert}$$

und schreibt in diesem Fall $R \cong \bar{R}$.

Insbesondere vermittelt f einen Isomorphismus der additiven Gruppe $(R, +)$ des Ringes R auf die additive Gruppe von $\bar{R}$. Daher ist $f(0) = \bar{0}$ das Nullelement von $\bar{R}$, und für alle $r \in R$ gilt $f(-r) = -f(r)$ (vgl. 12.3., Satz 1). Der Leser weist leicht die Übertragung weiterer Eigenschaften von R auf $\bar{R}$ nach. Ist beispielsweise 1 Einselement von R, so ist $f(1) = \bar{1}$ Einselement von $\bar{R}$; ist R Integritätsbereich, so auch $\bar{R}$, ist R Körper, so auch $\bar{R}$.

Beispiele.

1. Die Abbildung

$$f: a \mapsto \boldsymbol{A} = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \qquad (a \in \mathbf{Z})$$

ist offensichtlich ein Isomorphismus vom Ring der ganzen Zahlen auf den Ring der Matrizen $\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$ mit ganzzahligen Koeffizienten a.

2. Durch

$$f: a + bi \mapsto \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \qquad (a \in \mathbf{Z} \wedge b \in \mathbf{Z})$$

wird der Ring der Gaußschen ganzen komplexen Zahlen ebenfalls auf einen Matrizenring abgebildet.

Es ist sofort klar, daß das Bild einer Summe von komplexen Zahlen gleich der Summe der Bilder dieser Zahlen im Matrizenring ist. Aus

$$(a + bi)\,(c + di) = (ac - bd) + (ad + bc)\,i$$

und

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \begin{pmatrix} c & d \\ -d & c \end{pmatrix} = \begin{pmatrix} ac - bd & ad + bc \\ -(ad + bc) & ac - bd \end{pmatrix}$$

folgt, daß auch das Bild eines Produktes gleich dem Produkt der Bilder ist.

3. Wir überlassen es dem Leser als Aufgabe, zu zeigen, daß die Abbildung

$$f\colon a_1 + a_2 i + a_3 j + a_4 k \mapsto \begin{pmatrix} a_1 + a_2 i & a_3 + a_4 i \\ -(a_3 - a_4 i) & a_1 - a_2 i \end{pmatrix} \quad (a_1, a_2, a_3, a_4 \in \mathbf{R})$$

ein Isomorphismus des Schiefkörpers der Quaternionen auf einen Schiefkörper von Matrizen ist.

Die Isomorphie von Ringen ist wie im Fall der Gruppen (vgl. 12.3.1.) reflexiv, symmetrisch und transitiv, vermittelt als Äquivalenzrelation in jeder Menge von Ringen also eine Einteilung dieser Ringe in disjunkte Klassen (vgl. MfL, Bd. 1, 2.5.). Eine solche Klasse isomorpher Ringe heißt *abstrakter Ring*. Die einzelnen Vertreter einer Klasse stellen verschiedene Realisierungen desselben abstrakten Rechenschemas dar. Deshalb werden isomorphe Ringe oft als nicht wesentlich verschieden angesehen und in der Ringtheorie abstrakte Ringe betrachtet.

Definition 2. Bezeichnet R einen Ring, so heißt

f *Automorphismus* von R: $\Leftrightarrow$ f ist Isomorphismus von R auf sich.

Beispiele.

1. Die Menge der Zahlen $a + b\sqrt{2}$ $(a, b \in \mathbf{Z})$ bildet einen Unterring U von $\mathbf{R}$.

$$f\colon a + b\sqrt{2} \mapsto a - b\sqrt{2}$$

vermittelt einen Automorphismus von U.

2. Die Abbildung

$$f\colon a + bi \mapsto a - bi \qquad (a, b \in \mathbf{R}),$$

die jeder komplexen Zahl ihre konjugiert komplexe Zahl zuordnet, ist ein Automorphismus von $\mathbf{C}$.

3. Der identische Automorphismus, der jedes Element eines Ringes auf sich abbildet, ist der einzige Automorphismus von $\mathbf{Z}$, $\mathbf{Q}$ und allen Restklassenringen mod m $(m \in \mathbf{N}^*)$ (Übungsaufgabe).

Sind f und g Automorphismen eines Ringes R, so ist $f \circ g$ eine 1-1-Abbildung von R auf sich, und es gilt

$$\bigwedge_{r_1, r_2 \in R} f \circ g(r_1 + r_2) = f\big(g(r_1) + g(r_2)\big) = f \circ g(r_1) + f \circ g(r_2)$$

sowie

$$\bigwedge_{r_1, r_2 \in R} f \circ g(r_1 r_2) = f\big(g(r_1)\, g(r_2)\big) = f \circ g(r_1)\, f \circ g(r_2),$$

d. h., $f \circ g$ ist Automorphismus von R. Die Nacheinanderausführung von Automorphismen ist assoziativ (vgl. MfL, Bd. 1, 2.4.), und die identische Abbildung ist offenbar ein Automorphismus. Jedes Element aus R besitzt genau eine Darstellung

der Form $f(r)$. Sind $f(r_1)$, $f(r_2)$ beliebige Elemente aus R, so gilt für die zu f inverse Abbildung f^{-1}:

$$f^{-1}\big(f(r_1)+f(r_2)\big)=f^{-1}\big(f(r_1+r_2)\big)=r_1+r_2=f^{-1}\big(f(r_1)\big)+f^{-1}\big(f(r_2)\big)$$

und

$$f^{-1}\big(f(r_1)\,f(r_2)\big)=f^{-1}\big(f(r_1r_2)\big)=r_1r_2=f^{-1}\big(f(r_1)\big)\,f^{-1}\big(f(r_2)\big).$$

Daher ist auch f^{-1} ein Automorphismus. Die Automorphismen eines Ringes R bilden also bezüglich der Nacheinanderausführung eine Gruppe, die *Automorphismengruppe* von R.

13.4. Homomorphie von Ringen

13.4.1. Wie bei den Gruppen kommt man auf den wichtigen Begriff der homomorphen Abbildung, wenn man auf die bei den Isomorphismen geforderte eindeutige Umkehrbarkeit verzichtet.

Definition 1. Für die Ringe R und $\overline{R}$ heißt

f *Homomorphismus* (oder *homomorphe Abbildung*) von R in $\overline{R}$

$:\Leftrightarrow f$ Abbildung von R in $\overline{R} \wedge \bigwedge_{r_1,r_2\in R} f(r_1+r_2)=f(r_1)+f(r_2) \wedge \bigwedge_{r_1,r_2\in R} f(r_1r_2)=f(r_1)\,f(r_2)$

ist. Man schreibt

$R \simeq \overline{R} :\Leftrightarrow$ ein Homomorphismus von R in $\overline{R}$ existiert.

Sei f ein Homomorphismus von R in $\overline{R}$. Dann ist das Bild

$$\overline{U} := \{\overline{u} : \overline{u}\in\overline{R} \wedge \bigvee_{r\in R} f(r)=\overline{u}\}$$

von R bei f ein Unterring von $\overline{R}$. Wie bei den Gruppen (vgl. 12.5.) ergibt sich nämlich, daß mit zwei Elementen $\overline{u}_1$ und $\overline{u}_2$ aus $\overline{U}$ auch deren Summe und Produkt in $\overline{U}$ liegen. Die Gültigkeit der Ringaxiome wird durch die Abbildung f von R nach $\overline{U}$ übertragen. Da f insbesondere einen Homomorphismus der additiven Gruppe von R in diejenige von $\overline{R}$ vermittelt, gilt für die Nullelemente 0 und $\bar{0}$ von R bzw. $\overline{R}$

$$f(0)=\bar{0}$$

(vgl. 12.5.). Besitzt R ein Einselement 1, so ist

$$r\cdot 1=1\cdot r=r \qquad \text{für alle } r\in R.$$

Daraus folgt

$$\overline{u}f(1)=f(1)\,\overline{u}=\overline{u} \qquad \text{für alle } \overline{u}\in f(R)=\overline{U}.$$

Ist also $\overline{U}$ nicht der Nullring, so enthält $\overline{U}$ das Einselement $\bar{1}=f(1)$.

Beispiele.

1. Es sei $m \in \mathbf{N}^*$, und für $a \in \mathbf{Z}$ bezeichne $[a]$ die Restklasse von a mod m. Die Abbildung

$$f: a \mapsto [a]$$

ist ein Homomorphismus von $\mathbf{Z}$ auf den Restklassenring mod m, denn für die Elemente $a, b \in \mathbf{Z}$ gilt

$$[a + b] = [a] + [b], \qquad [ab] = [a]\,[b]$$

(vgl. 12.1.2.9.).

2. Die Abbildung

$$f: p(x) = a_0 + a_1 x + \cdots + a_n x^n \mapsto a_0 \qquad (n \in \mathbf{N} \wedge a_0, \ldots, a_n \in \mathbf{Z}),$$

die jedem Polynom aus $\mathbf{Z}[x]$ sein „Absolutglied" a_0 zuordnet, ist ein Homomorphismus von $\mathbf{Z}[x]$ auf $\mathbf{Z}$.

Sei f ein Homomorphismus des Ringes R auf den Ring $\bar{R}$. Er bestimmt insbesondere einen Homomorphismus der additiven Gruppe von R auf diejenige von $\bar{R}$. Bezeichnet $\bar{0}$ das Nullelement von $\bar{R}$, so ist

$$\mathfrak{n} := \{n: n \in R \wedge f(n) = \bar{0}\}$$

als Kern des Gruppenhomomorphismus eine Untergruppe der additiven Gruppe von R. Nach dem Homomorphiesatz für Gruppen haben zwei Elemente $r_1, r_2 \in R$ genau dann dasselbe Bild bei f, wenn sie in derselben Nebenklasse $\mathfrak{n} + a$ $(a \in R)$ der additiven Gruppe von R nach dem Kern $\mathfrak{n}$ liegen.

Wir beziehen nun auch die Multiplikation in unsere Betrachtungen mit ein. Weil f ein Ringhomomorphismus ist, gilt für beliebige $r \in R$ und $n \in \mathfrak{n}$

$$f(rn) = f(r)\,f(n) = f(r)\,\bar{0} = \bar{0}$$

sowie

$$f(nr) = f(n)\,f(r) = \bar{0} f(r) = \bar{0},$$

d. h. $rn \in \mathfrak{n}$ und $nr \in \mathfrak{n}$. Für ein festes $r \in R$ sei

$$r\mathfrak{n} := \{rn: n \in \mathfrak{n}\}, \qquad \mathfrak{n}r := \{nr: n \in \mathfrak{n}\}.$$

Dann hat $\mathfrak{n}$ für beliebige $r \in R$ die Eigenschaft $r\mathfrak{n} \subseteq \mathfrak{n}$ und $\mathfrak{n}r \subseteq \mathfrak{n}$. Wählt man insbesondere $r \in \mathfrak{n}$, so ergibt sich, daß das Produkt zweier Elemente aus $\mathfrak{n}$ wieder in $\mathfrak{n}$ liegt und demzufolge $\mathfrak{n}$ ein Unterring von R ist. $\mathfrak{n}$ wird *Kern des Ringhomomorphismus* f genannt.

Definition 2. In dem Ring R heißt

$$\mathfrak{n} \textit{ Ideal von } R :\Leftrightarrow \mathfrak{n} \text{ ist Unterring von } R \wedge \bigwedge_{r \in R} r\mathfrak{n} \subseteq \mathfrak{n} \wedge \mathfrak{n}r \subseteq \mathfrak{n}.$$

Sei R ein kommutativer Ring mit Einselement 1 und $a \in R$. Die Menge

$$\mathfrak{n} = \left\{n: \bigvee_{k \in R} ka = n\right\}$$

aller *Vielfachen* ka ($k \in R$) von a bildet ein Ideal von R. Es heißt das von a erzeugte ***Hauptideal*** von R und wird mit (a) bezeichnet. Das nur aus dem Nullelement 0 bestehende ***Nullideal*** (0) ist ein spezielles Hauptideal. Das von einer Einheit $e \in R$ erzeugte Hauptideal ist der ganze Ring R, denn in (e) liegt $ee^{-1} = 1$ und daher jedes Element $1r = r$ aus R.

Im Polynomring $\mathbf{Z}[x]$ bilden die Polynome der Form

$$a(x) = a_1 x + \cdots + a_n x^n \qquad (n \in \mathbf{N}^* \wedge a_1, \ldots, a_n \in \mathbf{Z})$$

das Ideal (x). (x^2) besteht aus allen Polynomen

$$b(x) = b_2 x^2 + \cdots + b_n x^n \qquad (n \in \mathbf{N}^* \wedge n \geqq 2 \wedge b_2, \ldots, b_n \in \mathbf{Z}),$$

und die Elemente von (k) sind die Polynome

$$c(x) = kc_0 + kc_1 x + \cdots + kc_n x^n \qquad (n \in \mathbf{N} \wedge c_0, \ldots, c_n \in \mathbf{Z}),$$

deren sämtliche Koeffizienten durch eine feste Zahl $k \in \mathbf{Z}$ teilbar sind.

Die Nebenklassen $\mathfrak{n} + a$ ($a \in R$) der additiven Gruppe von R nach $\mathfrak{n}$ nennt man (als Teilmengen des Ringes R) *Restklassen* modulo $\mathfrak{n}$.

Diejenigen Elemente des Ringes R, die bei dem Homomorphismus f von R auf $\bar{R}$ auf das Nullelement $\bar{0}$ abgebildet werden, bilden also ein Ideal von R, und die Restklassen modulo $\mathfrak{n}$ entsprechen eineindeutig den Elementen von $\bar{R}$.

Im Beispiel 1 besteht das zur Abbildung $a \mapsto f(a) = [a]$ gehörige Ideal aus allen ganzzahligen Vielfachen von m und ist also das Hauptideal (m). Die Restklassen mod (m) sind dann genau die früher eingeführten Restklassen mod m.

Im Beispiel 2 ist (x) das zugehörige Ideal, denn im Kern liegen alle Polynome mit dem Absolutglied 0. Sie haben die Form $xp(x)$ $\big(p(x) \in \mathbf{Z}[x]\big)$. Zwei Polynome aus $\mathbf{Z}[x]$ liegen genau dann in derselben Restklasse mod (x), wenn ihre Absolutglieder übereinstimmen.

13.4.2. Wir wollen nun zeigen, daß es umgekehrt zu jedem Ideal $\mathfrak{n}$ eines Ringes R einen Ring $\bar{R}$ und einen Homomorphismus f von R auf $\bar{R}$ gibt, dessen Kern $\mathfrak{n}$ ist.

Dazu betrachten wir die Menge $\{\mathfrak{n}, \mathfrak{n} + a, \mathfrak{n} + b, \ldots\}$ aller Restklassen mod $\mathfrak{n}$. Aus der Kommutativität der additiven Gruppe von R folgt, daß $\mathfrak{n}$ Normalteiler in $(R, +)$ ist. Durch Übertragung der Ergebnisse von 12.5. in die additive Schreibweise ergibt sich sofort, daß die Menge der Restklassen mod $\mathfrak{n}$ bezüglich der „Komplexaddition" eine Gruppe bildet. Sind $(\mathfrak{n} + a)$, $(\mathfrak{n} + b)$ zwei beliebige Restklassen

mod $\mathfrak{n}$, so bedeutet dabei

$$\begin{aligned}(\mathfrak{n}+a)+(\mathfrak{n}+b) &:= \{(n_1+a)+(n_2+b): n_1 \in \mathfrak{n} \wedge n_2 \in \mathfrak{n}\}\\ &= \{n+(a+b): n \in \mathfrak{n}\}\\ &= \mathfrak{n}+(a+b)\\ &= \mathfrak{n}+c\end{aligned}$$

diejenige Restklasse mod $\mathfrak{n}$, in der $a+b$ liegt.

Sind $n_1, n_2 \in \mathfrak{n}$ und $a, b \in R$, so folgt aus den Idealeigenschaften von $\mathfrak{n}$, daß $n := n_1 n_2 + n_1 b + a n_2$ ein Element von $\mathfrak{n}$ ist. Daher liegen alle Produkte

$$(n_1+a)(n_2+b) = n_1 n_2 + n_1 b + a n_2 + ab = n + ab \qquad (n_1, n_2 \in \mathfrak{n})$$

in derselben Restklasse mod $\mathfrak{n}$ wie ab. Diese sei $\mathfrak{n}+d$. Durch

$$(\mathfrak{n}+a)(\mathfrak{n}+b) := (\mathfrak{n}+ab) = (\mathfrak{n}+d)$$

wird also eine zweistellige Operation in der Menge der Restklassen mod $\mathfrak{n}$ erklärt, die wir *Restklassenmultiplikation* nennen werden.

Aus der Assoziativität der Multiplikation in R ergibt sich sofort die Assoziativität der Restklassenmultiplikation. Ebenso folgt aus der Gültigkeit der Distributivgesetze in R, daß auch die Addition und Multiplikation der Restklassen mod $\mathfrak{n}$ diesen Gesetzen genügen. Die Restklassen mod $\mathfrak{n}$ bilden also einen Ring.

Definition 3. Sei $\mathfrak{n}$ ein Ideal des Ringes R. Der Ring der Restklassen von R modulo $\mathfrak{n}$ mit der Restklassenaddition und Restklassenmultiplikation als Operationen wird mit $R/\mathfrak{n}$ bezeichnet und *Restklassenring* von R modulo $\mathfrak{n}$ genannt.

Kurz gesagt, man rechnet im Restklassenring, indem man aus jeder Klasse einen Vertreter wählt und als Summe (Produkt) von zwei Restklassen diejenige Klasse nimmt, in der die Summe (das Produkt) der Vertreter liegt. Wir haben gesehen, daß das Ergebnis der Rechnung von der Auswahl der Vertreter unabhängig ist.

Bildet man nun jedes Element $r \in R$ auf diejenige Restklasse von R modulo $\mathfrak{n}$ ab, in der r liegt, so erhält man eine Abbildung von R auf $\overline{R} = R/\mathfrak{n}$. Sie ist sogar ein Homomorphismus, denn aus $f(r_1) = \mathfrak{n}+a$ und $f(r_2) = \mathfrak{n}+b$ folgt nach der Definition der Summe und des Produktes zweier Restklassen, daß $r_1 + r_2$ in

$$(\mathfrak{n}+a)+(\mathfrak{n}+b) = (\mathfrak{n}+c)$$

und $r_1 r_2$ in

$$(\mathfrak{n}+a)(\mathfrak{n}+b) = (\mathfrak{n}+d)$$

liegt. Es ist also

$$f(r_1+r_2) = f(r_1)+f(r_2) \qquad \text{und} \qquad f(r_1 r_2) = f(r_1)\,f(r_2).$$

Diese Abbildung heißt *natürlicher* (oder *kanonischer*) *Homomorphismus von R auf $R/\mathfrak{n}$*.

Bezeichnet f einen Homomorphismus des Ringes R auf den Ring $\bar{R}$ und $\mathfrak{n}$ den zugehörigen Kern, so ist

$$R/\mathfrak{n} \cong \bar{R}.$$

Da nämlich zwei Elemente von R genau dann dasselbe Bild bei f haben, wenn sie in derselben Restklasse modulo $\mathfrak{n}$ liegen, ist

$$\bar{f}: (\mathfrak{n} + a) \mapsto f(a) \qquad \big((\mathfrak{n} + a) \in R/\mathfrak{n}\big)$$

eine 1-1-Abbildung von $R/\mathfrak{n}$ auf $\bar{R}$. Sind $(\mathfrak{n} + a)$, $(\mathfrak{n} + b)$ beliebige Restklassen modulo $\mathfrak{n}$ und ist $a + b \in (\mathfrak{n} + c)$, $ab \in (\mathfrak{n} + d)$, so gilt

$$\begin{aligned}\bar{f}\big((\mathfrak{n} + a) + (\mathfrak{n} + b)\big) &= \bar{f}(\mathfrak{n} + c) = f(a + b)\\ &= f(a) + f(b) = \bar{f}(\mathfrak{n} + a) + \bar{f}(\mathfrak{n} + b)\end{aligned}$$

und

$$\bar{f}\big((\mathfrak{n} + a)(\mathfrak{n} + b)\big) = \bar{f}(\mathfrak{n} + d) = f(ab) = f(a) f(b) = \bar{f}(\mathfrak{n} + a)\,\bar{f}(\mathfrak{n} + b).$$

$\bar{f}$ ist also ein Isomorphismus von $R/\mathfrak{n}$ auf $\bar{R}$.

Die Ergebnisse fassen wir zusammen zum

Satz 1 (Homomorphiesatz für Ringe). *Durch jede homomorphe Abbildung f eines Ringes R auf einen Ring $\bar{R}$ wird ein Ideal $\mathfrak{n}$ von R bestimmt. Es besteht aus denjenigen Elementen von R, die bei f auf das Nullelement $\bar{0}$ von $\bar{R}$ abgebildet werden und heißt Kern des Homomorphismus f. Der Restklassenring $R/\mathfrak{n}$ ist isomorph $\bar{R}$.*

Umgekehrt gibt es zu jedem Ideal $\mathfrak{n}$ von R eine homomorphe Abbildung von R auf $R/\mathfrak{n}$, deren Kern $\mathfrak{n}$ ist.

Jeder Ring R ist Ideal von sich und enthält das nur aus dem Nullelement bestehende Nullideal (0). Es ist $R/R \cong o$ und $R/(0) \cong R$, wobei o den Nullring bezeichnet. Schiefkörper und Körper enthalten auch nur diese beiden *trivialen Ideale*. Ist nämlich R ein Schiefkörper und enthält das Ideal $\mathfrak{n}$ von R ein Element $a \neq 0$, so gibt es in R zu a ein inverses Element a^{-1}, und wegen der Idealeigenschaft liegt das Einselement $1 = aa^{-1}$ von R in $\mathfrak{n}$. Dann gehören aber auch alle Elemente $1r = r$ $(r \in R)$ zu $\mathfrak{n}$, und es ist $\mathfrak{n} = R$. Damit haben wir bewiesen:

Folgerung. *Jedes homomorphe Bild eines Körpers K ist zu K oder dem Nullring o isomorph.*

Definition 4. Es sei $\mathfrak{p}$ ein Ideal des kommutativen Ringes R.
$\mathfrak{p}$ heißt *Primideal* von $R :\Leftrightarrow R/\mathfrak{p}$ Integritätsbereich ist.

R ist immer Primideal von R, weil $R/R \cong o$ nullteilerfrei ist. Das Nullideal $\mathfrak{o}$ ist genau dann Primideal, wenn R Integritätsbereich ist. Der Restklassenring $\mathbf{Z}/(m)$ der ganzen Zahlen modulo m $(m \in \mathbf{N} \wedge m > 1)$ ist dann und nur dann nullteilerfrei, wenn m eine Primzahl ist (vgl. 13.1., 13.2.). Für natürliche Zahlen $m \in \mathbf{N}$ gilt also

$$(m) \textit{ ist Primideal von } \mathbf{Z} \Leftrightarrow m \textit{ Primzahl} \vee m = 1 \vee m = 0. \tag{1}$$

13.4.3. Um ein weiteres Beispiel für einen Restklassenring zu erhalten, betrachten wir im Polynomring $\mathbf{Z}[x]$ das Hauptideal $(1 + x^2)$. Es besteht aus allen Polynomen der Form $(1 + x^2)\, q(x)$ mit $q(x) \in \mathbf{Z}[x]$ $f(x), g(x) \in \mathbf{Z}[x]$ liegen genau dann in derselben Restklasse mod $(1 + x^2)$, wenn $f(x) - g(x) \in (1 + x^2)$ (vgl. 12.2.3.), wenn es also in $\mathbf{Z}[x]$ ein $q(x)$ gibt, so daß

$$f(x) = (1 + x^2)\, q(x) + g(x)$$

ist. Weil es zu jedem $p(x) \in \mathbf{Z}[x]$ Polynome $q(x)$ und $r(x) = a_0 + a_1 x$ in $\mathbf{Z}[x]$ gibt, für die

$$p(x) = (1 + x^2)\, q(x) + r(x)$$

gilt (Polynomdivision mit Rest!), kann jede Restklasse $[p(x)] \in \mathbf{Z}[x]/(1 + x^2)$ in der Form $[a_0 + a_1 x]$ $(a_0, a_1 \in \mathbf{Z})$ geschrieben werden. Dabei ist

$$[a_0 + a_1 x] = [b_0 + b_1 x] \Leftrightarrow a_0 = b_0 \wedge a_1 = b_1 \qquad (a_0, a_1, b_0, b_1 \in \mathbf{Z}).$$

Die Operationen im Restklassenring $\mathbf{Z}[x]/(1 + x^2)$ sind durch

$$[a_0 + a_1 x] + [b_0 + b_1 x] = [(a_0 + b_0) + (a_1 + b_1)\, x]$$

und

$$\begin{aligned}[a_0 + a_1 x]\,[b_0 + b_1 x] &= [a_0 b_0 + (a_1 b_0 + a_0 b_1)\, x + a_1 b_1 x^2] \\ &= [(a_0 b_0 - a_1 b_1) + (a_1 b_0 + a_0 b_1)\, x] \quad (a_0, a_1, b_0, b_1 \in \mathbf{Z})\end{aligned}$$

festgelegt, denn $a_1 b_1 x^2 = a_1 b_1 (x^2 + 1) - a_1 b_1$.

Insbesondere ist dann

$$[x]\,[x] = [-1].$$

Daher ist die Abbildung

$$f\colon [a_0 + a_1 x] \mapsto a_0 + a_1 i \qquad (a_0, a_1 \in \mathbf{Z})$$

offenbar ein Isomorphismus von $\mathbf{Z}[x]/(1 + x^2)$ auf den Ring der Gaußschen ganzen komplexen Zahlen. $(1 + x^2)$ ist daher ein Primideal von $\mathbf{Z}[x]$.

Analog ergibt sich die Isomorphie

$$\mathbf{R}[x]/(1 + x^2) \cong \mathbf{C}$$

des Restklassenringes aller Polynome mit reellen Koeffizienten nach dem von $x^2 + 1$ erzeugten Hauptideal mit dem Körper der komplexen Zahlen.

13.4.4. Nach dem Homomorphiesatz kann man sich (bis auf Isomorphie) eine Übersicht über alle homomorphen Bilder eines Ringes R verschaffen, wenn man sämtliche Ideale von R kennt. Als Beispiel betrachten wir den Ring $\mathbf{Z}$ der ganzen Zahlen.

Sei $\mathfrak{m} \neq (0)$ ein Ideal von $\mathbf{Z}$. Es gibt also ein Element $a \neq 0$ in $\mathfrak{m}$. Da aus $a \in \mathfrak{m}$ folgt $-a \in \mathfrak{m}$, kann man sogar $a \in \mathbf{N}^*$ annehmen. Unter den in $\mathfrak{m}$ enthaltenen

Zahlen aus $\mathbf{N}^*$ sei m die kleinste. Zu jeder Zahl $x \in \mathfrak{m}$ gibt es ganze Zahlen a, r, so daß

$$x = am + r \qquad (0 \leqq r < m)$$

ist. Da x und m in $\mathfrak{m}$ liegen, ist auch $r = x - am \in \mathfrak{m}$. Wegen der Minimaleigenschaft von m muß dann $r = 0$ und demnach

$$\mathfrak{m} = (m)$$

sein. Die Ideale von $\mathbf{Z}$ sind also die von den Zahlen $m \in \mathbf{N}$ erzeugten Hauptideale (m). Jedes homomorphe Bild des Ringes $\mathbf{Z}$ ist daher einem Restklassenring $\mathbf{Z}/(m)$ isomorph. Im Fall $m = 1$ ist das der Nullring und für $m = 0$ ist $\mathbf{Z}/(0) \cong \mathbf{Z}$.

Die Elemente von $\mathbf{Z}/(m)$ bezeichnen wir mit $[0], [1], \ldots, [m-1]$. Ist $a \sqcap m = d > 1$, so kann die Gleichung

$$[a]\,[x] = [1] \tag{1}$$

sicher keine Lösung $[x] \in \mathbf{Z}/(m)$ besitzen, da sonst d ein Teiler von 1 sein müßte. Ist aber $a \sqcap m = 1$, so hat (1) eine (nach 11.3.(6) eindeutig bestimmte) Lösung $[\bar{a}]$ (vgl. 12.1.2.10.), d. h., genau zu den primen Restklassen gibt es inverse Elemente in $\mathbf{Z}/(m)$. Weil aus $[a]\,[b] = [0]$ und $[\bar{a}]\,[a] = [1]$ folgt, daß

$$[b] = [\bar{a}]\,([a]\,[b]) = [\bar{a}]\,[0] = [0]$$

ist, kann eine prime Restklasse nicht Nullteiler von $\mathbf{Z}/(m)$ sein.

Besitzt die Gleichung

$$[a]\,[x] = [b] \tag{2}$$

eine Lösung $[x] \in \mathbf{Z}/(m)$, gibt es also Zahlen $x, k \in \mathbf{Z}$, für die $ax = b + km$ ist, so muß notwendig $d = a \sqcap m$ ein Teiler von b sein. Ist umgekehrt $d = a \sqcap m$ Teiler von b, so gibt es Zahlen $u, v \in \mathbf{Z}$ und $b' \in \mathbf{N}$ mit der Eigenschaft $au + mv = d$ und $db' = b$ (vgl. MfL, Bd. 1, 3.7.). Daraus folgt

$$a(ub') + m(vb') = db' = b, \tag{3}$$

also ist

$$[x] = [ub']$$

eine Lösung von (2). Mit $a = a'd$ und $m = m'd$ ergibt sich weiter aus (3)

$$a(ub' + km') + m(vb' - ka') = b,$$

d. h., die Restklassen $[ub' + km']$ $(k = 0, 1, \ldots, d-1)$ sind Lösungen von (2). Da $[a']$ prime Restklasse ist, folgt andererseits für zwei Lösungen $[x]$ und $[\bar{x}]$ von (2) aus

$$[a]\,[x] - [a]\,[\bar{x}] = [a]\,([x] - [\bar{x}]) = [a']\,[d(x - \bar{x})] = [0],$$

daß $[d(x - \bar{x})] = [0]$ ist und sich x und $\bar{x}$ daher nur um ein ganzzahliges Vielfaches von m' unterscheiden.

Damit haben wir bewiesen

Satz 2. *Im Restklassenring* $\mathbf{Z}/(m)$ *ist* $[a]\,[x] = [b]$ *lösbar* $\Leftrightarrow a \sqcap m \mid b$.

Bezeichnet $[x]$ *eine Lösung, so sind* $\left[x + k\,\frac{m}{a \sqcap m}\right]$ $(k = 0, 1, \ldots, a \sqcap m - 1)$ *sämtliche Lösungen.*

Insbesondere sieht man noch einmal (vgl. 13.2.), daß

$$\mathbf{Z}/(m) \textit{ nullteilerfrei} \Leftrightarrow m = 0 \vee m = 1 \vee m \textit{ Primzahl}$$

und

$$\mathbf{Z}/(m) \textit{ Körper} \Leftrightarrow m \textit{ Primzahl}.$$

Sei p eine Primzahl. Nach dem Satz von Fermat (vgl. 12.2.3., Folgerung 3') gilt für alle von [0] verschiedenen Elemente $[a]$ des *Restklassenkörpers* $\mathbf{Z}/(p)$

$$[a]^{p-1} - [1] = [0] \tag{4}$$

und für alle $[a] \in \mathbf{Z}/(p)$ dann $[a]^p = [a]$.

Sei $[a] \in \mathbf{Z}/(p) \wedge [a] \neq [0]$. Wie bei der Polynomdivision in $\mathbf{Z}[x]$ ergibt sich, daß es ein Polynom $q_2(x) = x^{p-2} + \cdots$ mit Koeffizienten aus $\mathbf{Z}/(p)$ und ein $[r] \in \mathbf{Z}/(p)$ gibt, so daß

$$q_1(x) = x^{p-1} - [1] = q_2(x)\,(x - [a]) + [r]$$

ist, wobei $x^n := [1]\,x^n$ $(n \in \mathbf{N}^*)$ bedeutet. Wegen (4) folgt daraus für $x = [a]$, daß $[r] = [0]$ ist. Die Nullteilerfreiheit von $\mathbf{Z}/(p)$ ergibt weiter, daß für alle von [0] und $[a]$ verschiedenen Elemente $[b]$

$$q_2([b]) = [0]$$

gilt. Daher kann die auf $q_1(x)$ angewendete Überlegung für $q_2(x)$ wiederholt werden usw. Nach $p - 2$ Schritten erhält man

$$x^{p-1} - [1] = (x - [1])\,(x - [2]) \cdots (x - [p-1]).$$

Hieraus ergibt sich durch Vergleich der Koeffizienten vor den Potenzen von x

$$[1] + [2] + \cdots + [p-1] = [0],$$

$$[1]\,[2] + [1]\,[3] + \cdots + [p-2]\,[p-1] = [0],$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$[1]\,[2] \cdots [p-1] = [-1]^p.$$

Die letzte Gleichung liefert den folgenden Satz.

Satz 3 (Satz von Wilson). *Für jede Primzahl* p *gilt* $(p-1)! \equiv (-1)^p \bmod p$.

Es sei p eine Primzahl der Form $p = 4n + 1$ $(n \in \mathbf{N}^*)$. Dann ist

$$\begin{aligned}(p-1)! &= (1 \cdot 2 \cdots (2n)) \cdot ((p-1)\,(p-2) \cdots (p-2n)) \\ &\equiv ((2n)!) \cdot ((-1)^{2n}(2n)!) \bmod p \\ &\equiv ((2n)!)^2 \bmod p\end{aligned}$$

und nach dem Satz von WILSON also

$$((2n)!)^2 \equiv -1 \mod p.$$

Folgerung. *Für Primzahlen p der Form $p = 4n+1$ ($n \in \mathbb{N}^*$) wird in $\mathbb{Z}/(p)$ die Gleichung $x^2 + [1] = [0]$ durch $x = [(2n)!]$ gelöst.*

13.5. Teilbarkeitstheorie in Integritätsbereichen

13.5.1. Die Teilbarkeit ist eine wichtige Relation im Ring der ganzen Zahlen. Sie kann auf andere Integritätsbereiche I übertragen werden.

Definition 1. Seien a, b Elemente des Integritätsbereiches I. Dann heißt

$$a \textit{ Teiler von } b :\Leftrightarrow \bigvee_{x \in I} ax = b.$$

Man schreibt in diesem Fall $a \mid b$ und nennt b *teilbar durch* a.[1])

Wir beschränken uns hier auf Integritätsbereiche, weil in diesen die Gleichung $ax = b$ ($a \neq 0$) höchstens eine Lösung besitzt, was gegenüber allgemeinen Ringen eine Vereinfachung bedeutet. In Körpern ist die Teilbarkeitsbeziehung trivial, da die Gleichung $ax = b$ für $a \neq 0$ immer genau eine Lösung hat.

Definition 2. Sei I ein Integritätsbereich mit Einselement 1 und $a, a' \in I$. a heißt *assoziiert* zu $a' :\Leftrightarrow \bigvee_{e \in I} e$ ist Einheit von $I \wedge ae = a'$.

Weil aus $ae_1 = a' \wedge a'e_2 = a$ die Gleichung $ae_1e_2 = a$ und nach 13.2. (1) dann $a = 0 \vee e_1e_2 = 1$ folgt, ist

$$a \text{ assoziiert zu } a' \Leftrightarrow a \mid a' \wedge a' \mid a. \tag{1}$$

Da die Einheiten von I eine abelsche Gruppe mit dem neutralen Element 1 bilden, weist man leicht nach, daß die Assoziiertheit eine Äquivalenzrelation in I ist. Die zu $a \in I$ assoziierten Elemente und sämtliche Einheiten e von I sind wegen $ae^{-1}e = a$ immer Teiler von a. Sie werden *triviale Teiler* von a genannt. t heißt *echter Teiler* von a, wenn $t \mid a$ und t nicht assoziiert zu a ist.

Definition 3. Ist I ein Integritätsbereich mit Einselement und $a \neq 0$ aus I, so heißt

a *unzerlegbar* (oder *irreduzibel*) in $I :\Leftrightarrow a$ besitzt nur triviale Teiler.

Anderenfalls wird a *zerlegbar* (oder *reduzibel*) genannt.

Im Ring $\mathbb{Z}$ der ganzen Zahlen sind z. B. alle Primzahlen p und die Zahlen $-p$ unzerlegbar. Die Menge $\{a + b\sqrt{-3} : a \in \mathbb{Z} \wedge b \in \mathbb{Z}\}$ bildet mit der üblichen Addition

[1]) $a \nmid b$ bedeutet, daß b nicht durch a teilbar ist.

und Multiplikation der komplexen Zahlen einen Integritätsbereich $\mathbf{Z}\left[\sqrt{-3}\right]$ mit Einselement $1 = 1 + 0\sqrt{-3}$. Einheiten sind nur 1 und -1 (Übungsaufgabe). In diesem Integritätsbereich ist

$$3 = \left(0 + 1\sqrt{-3}\right)\left(0 - 1\sqrt{-3}\right)$$

zerlegbar, 2 aber unzerlegbar, da

$$2 = \left(a + b\sqrt{-3}\right)\left(c + d\sqrt{-3}\right) = (ac - 3bd) + (ad + bc)\sqrt{-3}$$

nur die ganzzahligen Lösungen $b = d = 0$, $a = \pm 2$, $c = \pm 1$ und $b = d = 0$, $a = \pm 1, c = \pm 2$ hat.

Definition 4. Ist I ein Integritätsbereich mit Einselement und $p \in I$ weder Nullelement noch Einheit von I, so heißt

$$p \textit{ Primelement} \text{ von } I :\Leftrightarrow \bigwedge_{a,b \in I} (p \mid ab \Rightarrow p \mid a \vee p \mid b).$$

Satz 1. *Sei I Integritätsbereich mit Einselement.*

p Primelement von $I \Rightarrow p$ ist unzerlegbar in I.

Beweis. Ist $p = ab$ $(a, b \in I)$, so gilt $p \mid a$ oder $p \mid b$. Sei etwa $a = pc$ $(c \in I)$. Dann ist $p = pcb$ und nach 13.2.(1) $cb = 1$. b ist also Einheit von I, und a, b sind triviale Teiler von p.

Im Integritätsbereich $\mathbf{Z}\left[\sqrt{-3}\right]$ ist 2 unzerlegbar, aber kein Primelement, weil das Produkt $\left(1 + \sqrt{-3}\right)\left(1 - \sqrt{-3}\right) = 4$ durch 2 teilbar ist, ohne daß einer der Faktoren $\left(1 + \sqrt{-3}\right)$ oder $\left(1 - \sqrt{-3}\right)$ durch 2 teilbar ist (Übungsaufgabe).

In 13.4.4. wurde gezeigt, daß sämtliche Ideale von $\mathbf{Z}$ Hauptideale sind.

Definition 5. Ein Integritätsbereich I mit Einselement, dessen sämtliche Ideale Hauptideale sind, wird *Hauptidealring* genannt.

Grundlegend für die Teilbarkeitslehre in $\mathbf{Z}$ sind die Sätze vom größten gemeinsamen Teiler und von der eindeutigen Primzahlzerlegung (vgl. MfL, Bd. 1, 3.7.). Sie gelten in allen Hauptidealringen.

Definition 6. Sei I ein Integritätsbereich. $d \in I$ heißt *größter gemeinsamer Teiler* der Elemente $a_1, a_2, \ldots, a_n \in I$ (Bezeichnung: $a_1 \sqcap a_2 \sqcap \cdots \sqcap a_n$)

$$:\Leftrightarrow (d \mid a_1 \wedge d \mid a_2 \wedge \cdots \wedge d \mid a_n) \tag{2}$$

$$\wedge \bigwedge_{t \in I} (t \mid a_1 \wedge t \mid a_2 \wedge \cdots \wedge t \mid a_n) \Rightarrow t \mid d. \tag{3}$$

Ein Element $d \in I$ mit der Eigenschaft (2) heißt *gemeinsamer Teiler* von $a_1, a_2, \ldots, a_n$.

Satz 2 (Satz vom größten gemeinsamen Teiler). *Sind $a_1, a_2, \ldots, a_n$ Elemente eines Hauptidealringes I, so gilt:*

$$\text{Es gibt größte gemeinsame Teiler von } a_1, a_2, \ldots, a_n \text{ in } I. \tag{4}$$

$$\text{Je zwei größte gemeinsame Teiler von } a_1, a_2, \ldots, a_n \text{ sind assoziiert.} \tag{5}$$

$$d \in I \text{ ist größter gemeinsamer Teiler von } a_1, a_2, \ldots, a_n$$

$$\Leftrightarrow d \in I \text{ ist gemeinsamer Teiler von } a_1, a_2, \ldots, a_n$$

$$\wedge \bigvee_{x_1,\ldots,x_n \in I} x_1 a_1 + x_2 a_2 + \cdots + x_n a_n = d. \tag{6}$$

Beweis. Sei $\mathfrak{d}$ der Durchschnitt aller Ideale von I, die $(a_1), (a_2), \ldots, (a_n)$ umfassen. $\mathfrak{d}$ ist Ideal von I und wird daher von einem Element $d \in I$ erzeugt: $\mathfrak{d} = (d)$. Aus $(d) \supseteq (a_i)$ folgt $a_i \in (d)$, d. h. $\bigvee_{k \in I} dk_i = a_i$ $(i = 1, 2, \ldots, n)$. Demnach ist

$$d \mid a_1 \wedge d \mid a_2 \wedge \cdots \wedge d \mid a_n. \tag{7}$$

Ferner liegen alle Vielfachen $x_i a_i$ $(x_i \in I)$ von a_i $(i = 1, 2, \ldots, n)$ in (d). Da (d) ein Ideal ist, gehören dann auch alle *Vielfachsummen*

$$x_1 a_1 + x_2 a_2 + \cdots + x_n a_n \tag{8}$$

mit beliebigen Faktoren $x_1, x_2, \ldots, x_n \in I$ zu (d), d. h.

$$\mathfrak{v} := \left\{ v : \bigvee_{x_1,\ldots,x_n \in I} v = x_1 a_1 + x_2 a_2 + \cdots + x_n a_n \right\} \subseteq (d). \tag{9}$$

Andererseits ist $\mathfrak{v}$ ein Ideal von I, das die Ideale $(a_1), (a_2), \ldots, (a_n)$ umfaßt. Daher gilt

$$\mathfrak{v} \supseteq (d). \tag{10}$$

Weil nach (9) und (10) $\mathfrak{v} = (d)$ ist, besitzt d eine Vielfachsummendarstellung (8), d. h.

$$\bigvee_{x_1,\ldots,x_n \in I} x_1 a_1 + x_2 a_2 + \cdots + x_n a_n = d.$$

Ist $t \in I$ gemeinsamer Teiler von $a_1, a_2, \ldots, a_n$, gibt es also Elemente $f_i \in I$ $(i = 1, 2, \ldots, n)$, für die $a_1 = f_1 t, \ldots, a_n = f_n t$ gilt, so folgt aus

$$d = x_1 a_1 + x_2 a_2 + \cdots + x_n a_n = (x_1 f_1 + x_2 f_2 + \cdots + x_n f_n)\, t,$$

daß $t \mid d$ und daher d größter gemeinsamer Teiler von $a_1, a_2, \ldots, a_n$ ist. Damit sind (4), die Darstellbarkeit eines größten gemeinsamen Teilers von $a_1, a_2, \ldots, a_n$ als Vielfachsumme und die Implikation „$\Leftarrow$" aus (6) bewiesen.

Bezeichnen d, d' größte gemeinsame Teiler von $a_1, a_2, \ldots, a_n$, so ist

$$d \mid d' \wedge d' \mid d,$$

also nach (1) d assoziiert zu d'. Damit ist (5) bewiesen.

Ist d' ein größter gemeinsamer Teiler von $a_1, a_2, \ldots, a_n$, so ist nach (5) $d' = de$ mit einer Einheit $e \in I$, und man erhält sofort eine Vielfachsummendarstellung von d' aus einer solchen von d. Daher ist auch die Implikation „$\Rightarrow$" aus (6) bewiesen.

Die Elemente $a_1, a_2, \ldots, a_n$ aus I heißen *teilerfremd*, wenn ihre größten gemeinsamen Teiler die Einheiten von I sind.

Satz 3. *Sei I Hauptidealring und $p \in I$ weder Nullelement noch Einheit von I. Dann gilt:*

$$p \text{ unzerlegbar in } I \Rightarrow p \text{ Primelement von } I.$$

Beweis. Sei $p \mid ab$, etwa $ab = kp\,(a, b, k \in I)$. Da p unzerlegbar ist, ist p nur durch Einheiten von I und zu p assoziierte Elemente teilbar. Ist $p \sqcap a$ eine Einheit, so gibt es nach Satz 2 Elemente $x, y \in I$ derart, daß $xp + ya = 1$ und also

$$xbp + yab = (xb + yk)\,p = b,$$

d. h. $p \mid b$.

Folgerung. *Für jedes unzerlegbare Element p des Hauptidealringes I folgt aus $p \mid ab\ (a, b \in I)$, daß $p \mid a$ oder $p \mid b$.*

In einem Hauptidealring I gilt (vgl. 13.4.2.(1)):

$$(p) \text{ ist Primideal von } I$$
$$\Leftrightarrow p \text{ ist Primelement von } I \vee p = 0 \vee p \text{ ist Einheit von } I. \tag{11}$$

Das Ideal (0) ist ein Primideal. Bezeichnet p eine Einheit von I, so ist $(p) = I$ und daher ebenfalls Primideal. Ist aber p ein Primelement von I und $a \in I$, aber $a \notin (p)$, so sind p und a teilerfremd, und nach Satz 1 gibt es eine Vielfachsummendarstellung

$$xa + yp = 1 \tag{12}$$

mit Elementen x, y aus I. Bezeichne $[a]$ diejenige Restklasse von $I/(p)$, in der a liegt. Weil die Nullklasse [0] aus den Elementen von (p) besteht, folgt aus (12)

$$[x]\,[a] = [1],$$

d. h., jedes $[a] \neq [0]$ aus $I/(p)$ besitzt ein inverses Element. Daher ist $I/(p)$ ein Körper, also (p) Primideal.

Ist umgekehrt $p \neq 0$ als Produkt $p = rs\ (r, s \in I)$ darstellbar und ist keiner der Faktoren Einheit von I, so kann auch keiner der Faktoren in (p) liegen, weil er nach (1) sonst zu p assoziiert und der andere Faktor Einheit wäre. Für die Restklassen modulo (p) bedeutet das

$$[r] \neq [0] \wedge [s] \neq [0], \quad \text{aber} \quad [r]\,[s] = [0].$$

Daher ist (p) kein Primideal.

13.5.2. In Hauptidealringen gilt der folgende Satz.

Satz 4 (Teilerkettensatz). *Sei I ein Hauptidealring. Ist in*

$$a_1, a_2, \dots \qquad (a_i \in I; \quad i = 1, 2, \dots) \tag{13}$$

jedes Element a_{i+1} echter Teiler von a_i $(i = 1, 2, \dots)$, so kann die Folge (13) *nur endlich viele Glieder enthalten.*

Beweis. Wegen der Voraussetzung über die a_{i+1} ist

$$(a_i) \subset (a_{i+1}) \qquad (i = 1, 2, \dots). \tag{14}$$

Die Vereinigungsmenge

$$\mathfrak{a} := (a_1) \cup (a_2) \cup \cdots$$

ist ein Ideal von I. Ist nämlich $a \in \mathfrak{a}$, so liegt a in einem (a_i), und daher sind alle Vielfachen xa_i $(x \in I)$ in (a_i), also auch in $\mathfrak{a}$. Bezeichnen a, b Elemente von $\mathfrak{a}$, so liegt a in einem (a_i), b in einem (a_k). O. B. d. A. kann $k \geq i$ angenommen werden. Aus (14) folgt dann, daß a und b in (a_k) liegen, also auch $a - b \in (a_k) \subseteq \mathfrak{a}$.

$\mathfrak{a}$ ist nach der Voraussetzung über I ein Hauptideal: $\mathfrak{a} = (a)$. a muß als Element von $\mathfrak{a}$ in einem (a_i) enthalten sein. Sei n der kleinste Index mit der Eigenschaft $a \in (a_n)$. Dann ist $(a) \subseteq (a_n)$. Gäbe es noch einen echten Teiler a_{n+1} von a_n, so wäre $(a_n) \subset (a_{n+1})$, woraus der Widerspruch $(a) \subseteq (a_n) \subset (a_{n+1}) \subseteq \mathfrak{a} = (a)$ folgen würde. a_n muß also das letzte Glied der *Teilerkette* (13) sein.

Aus diesem Satz ergibt sich nun leicht der

Satz 5 (Satz von der eindeutigen Primelementzerlegung). *Ist I ein Hauptidealring und $a \in I$ weder Nullelement noch Einheit von I, so läßt sich a als Produkt von Primelementen $p_i \in I$ darstellen:*

$$a = p_1 \cdots p_m. \tag{15}$$

Bezeichnet

$$a = q_1 \cdots q_n$$

eine weitere derartige Darstellung, so ist $m = n$ und (bei geeigneter Numerierung) q_i assoziiert zu p_i $(i = 1, \dots, m)$.

Beweis. Die Existenz einer Primelementzerlegung (15) beweisen wir indirekt. Hätte a keine solche Zerlegung, so könnte a jedenfalls kein Primelement sein, da sonst bereits eine Primelementzerlegung (mit einem Faktor) vorläge. Nach Satz 3 muß dann a in der Form $a = a'a''$ zerlegbar sein, wobei a', a'' echte Teiler von a bezeichnen. Mindestens einer dieser Faktoren kann keine Primelementzerlegung besitzen, denn sonst wäre auch a als Produkt von Primelementen darstellbar. Auf diesen Faktor a_1 kann die Überlegung erneut angewendet werden usw. Es entsteht eine unendliche Folge $a, a_1, a_2, \dots$ von Elementen aus I, in der jedes Glied echter

Teiler des vorhergehenden ist. Die Existenz einer solchen Folge widerspricht Satz 4. Daher muß a eine Primelementzerlegung besitzen.

Daß die Faktoren in (15) bis auf Assoziiertheit eindeutig bestimmt sind, kann durch Induktion nach m bewiesen werden. Im Fall $m = 1$ ist a Primelement und besitzt also keine echte Zerlegung $a = q_1 \cdots q_n$. Daher ist $n = 1$ und $q_1 = p_1$.

Induktionsannahme: Für Produkte aus weniger als m Primelementen ist die Eindeutigkeit der Darstellung bereits bewiesen. Aus

$$a = p_1 p_2 \cdots p_m = q_1 q_2 \cdots q_n \tag{16}$$

folgt $p_1 \mid q_1 q_2 \cdots q_n$, und da p_1 Primelement ist, existiert ein q_i ($i \in \{1, \ldots, n\}$), welches durch p_1 teilbar ist. O. B. d. A. kann $p_1 \mid q_1$ angenommen werden. Da andererseits q_1 als Primelement unzerlegbar ist, besitzt es nur triviale Teiler. Es gibt also eine solche Einheit $e_1 \in I$, daß $q_1 = p_1 e_1$ ist. Aus $a = p_1 p_2 \cdots p_n = p_1 e_1 q_2 \cdots q_n$ ergibt sich dann nach 13.2.(1)

$$p_2 \cdots p_m = (e_1 q_2) \cdots q_n .$$

Nach der Induktionsannahme folgt daraus, daß $m - 1 = n - 1$, also $m = n$ und q_i assoziiert zu p_i ($i = 2, \ldots, m$) ist.

Mitunter wählt man aus jeder Äquivalenzklasse assoziierter Primelemente von I eines aus und schreibt die Primelementzerlegung (15) mit Hilfe dieser „normierten" Primelemente (und einer Einheit als Faktor) auf. Beispielsweise werden im Ring der ganzen Zahlen häufig die Primzahlen aus $\mathbf{N}^*$ als normierte Primelemente verwendet.

Integritätsbereiche mit Einselement, in denen der Satz von der eindeutigen Primelementzerlegung gilt, werden *ZPE-Ringe* genannt. Wir merken hier ohne Beweise an, daß es ZPE-Ringe gibt, die nicht Hauptidealringe sind.

13.5.3. In 13.4.4. wurde gezeigt, daß der Ring $\mathbf{Z}$ der ganzen Zahlen ein Hauptidealring ist. Der Beweis beruht auf der *Division mit Rest* und nutzt die durch die $<$-Relation gegebene Ordnung der Elemente von $\mathbf{Z}$ aus. Man kann ihn auf solche Ringe übertragen, die beschrieben werden durch

Definition 7. Sei I ein Integritätsbereich und 0 sein Nullelement.

I heißt *euklidischer Ring*

$:\Leftrightarrow$ eine Abbildung h von $I \setminus \{0\}$ in $\mathbf{N}$ existiert, so daß

$$\bigwedge_{a,b \in I} \Big(a \neq 0 \Rightarrow \bigvee_{q,r \in I} b = aq + r \wedge \big(r = 0 \vee h(r) < h(a) \big) \Big) .$$

Beispiele für euklidische Ringe sind:

1. Der Ring der ganzen Zahlen $\mathbf{Z}$ vermöge der Abbildung $h\colon a \mapsto |a|$.

2. Der Ring $\mathbf{Q}[x]$ aller Polynome in x mit Koeffizienten aus dem Körper der rationalen Zahlen $\mathbf{Q}$. $h(a)$ sei dabei der Grad des Polynoms $a \in \mathbf{Q}[x]$.

3. Jeder Körper K ist trivialerweise euklidischer Ring. Man setze nämlich $q = a^{-1}b$ und $r = 0$.

Satz 6. *Jeder euklidische Ring I ist ein Hauptidealring.*

Beweis. Sei $\mathfrak{a} \neq (0)$ ein Ideal in I. Dann existiert in $\mathfrak{a}$ ein Element $a \neq 0$, für das $h(a)$ minimal ist. Bezeichne b ein beliebiges Element aus $\mathfrak{a}$. In dem euklidischen Ring I gibt es Elemente q und r, so daß

$$b = aq + r \wedge \big(r = 0 \vee h(r) < h(a)\big)$$

ist. Weil mit b und a auch $b - aq = r$ in $\mathfrak{a}$ liegt, muß $r = 0$ sein. $\mathfrak{a}$ besteht also aus allen Vielfachen von a.

Außerdem muß gezeigt werden:

Jeder euklidische Ring I besitzt ein Einselement. (17)

I ist selbst ein Ideal. Daher muß es nach dem Vorstehenden ein $a \in I$ geben, von dem alle Ringelemente Vielfache qa sind. Insbesondere existiert ein solches $e \in I$, daß $a = ae$ ist. Bezeichnet $b = qa$ ein beliebiges Element aus I, so folgt $be = qae = qa = b$, d. h., e ist Einselement von I.

Während Satz 2 nur die Existenz des größten gemeinsamen Teilers in Hauptidealringen sichert, liefert das schon von Euklid angegebene *Verfahren der sukzessiven Divisionen* (*euklidischer Algorithmus*, vgl. MfL, Bd. 1, 3.7.) für euklidische Ringe sogar eine Methode zur Berechnung des größten gemeinsamen Teilers und seiner Vielfachsummendarstellung.

Sind a_1, a_2 $(a_1 \neq 0 \wedge a_2 \neq 0)$ Elemente des euklidischen Ringes I und ist $h(a_1) \geqq h(a_2)$, so bestimme man Elemente $q_1, q_2, \ldots$ und $r_1, r_2, \ldots$ aus I derart, daß

$$\begin{aligned}
a_1 &= a_2 q_1 + r_1 \quad \text{mit} \quad h(r_1) < h(a_2), \\
a_2 &= r_1 q_2 + r_2 \quad \text{mit} \quad h(r_2) < h(r_1), \\
r_1 &= r_2 q_3 + r_3 \quad \text{mit} \quad h(r_3) < h(r_2), \\
&\ldots\ldots\ldots\ldots
\end{aligned}$$

und fahre damit solange fort, bis einmal die Division mit dem Rest Null aufgeht:

$$\begin{aligned}
r_{k-1} &= r_k q_{k+1} + r_{k+1} \quad \text{mit} \quad h(r_{k+1}) < h(r_k), \\
r_k &= r_{k+1} q_{k+2},
\end{aligned}$$

was nach endlich vielen Divisionsschritten eintreten muß, weil die $h(r_i)$ $(i = 1, 2, \ldots)$ eine monoton abnehmende Folge natürlicher Zahlen sind.

Aus den Gleichungen (von unten nach oben ausgenutzt) ergibt sich sukzessive

$$r_{k+1} \mid r_k, \qquad r_{k+1} \mid r_{k-1}, \ldots, \qquad r_{k+1} \mid a_2, \qquad r_{k+1} \mid a_1.$$

Weiter entnimmt man den Gleichungen (von oben nach unten), daß aus $t \mid a_1 \wedge t \mid a_2$ auch folgt

$$t \mid r_1, \qquad t \mid r_2, \ldots, \qquad t \mid r_{k+1}.$$

Daher ist $r_{k+1} = a_1 \sqcap a_2$.

Die vorletzte Gleichung liefert eine Vielfachsummendarstellung von r_{k+1} als Vielfachsumme von r_{k-1} und r_k. Die drittletzte Gleichung ergibt eine Darstellung von r_{k+1} als Vielfachsumme von r_{k-1} und r_{k-2}. Benutzt man die Gleichungen weiter der Reihe nach (von unten nach oben), so erhält man schließlich r_{k+1} als Vielfachsumme von a_1 und a_2 ausgedrückt.

Wir wollen die letzten Ergebnisse am Ring G der Gaußschen ganzen komplexen Zahl illustrieren. Als *Norm* der komplexen Zahl $\alpha = a_0 + a_1 i$ $(a_0, a_1 \in \mathbf{R})$ bezeichnet man

$$N(\alpha) := a_0^2 + a_1^2 = (a_0 + a_1 i)(a_0 - a_1 i) = \alpha\bar{\alpha}.$$

Es ist

$$N(\alpha) = 0 \Leftrightarrow a_0 = a_1 = 0 \Leftrightarrow \alpha = 0.$$

Ist $\beta = b_0 + b_1 i$ eine weitere komplexe Zahl, so folgt aus den Regeln für das Rechnen mit komplexen Zahlen (vgl. MfL, Band 2)

$$N(\alpha\beta) = \alpha\beta\,\overline{\alpha\beta} = \alpha\bar{\alpha}\beta\bar{\beta} = N(\alpha)\,N(\beta). \tag{18}$$

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

$$N(\alpha^{-1}) = N(\alpha)^{-1} \tag{19}$$

Wir betrachten nun die Gaußschen ganzen komplexen Zahlen $a_0 + a_1 i$ $(a_0, a_1 \in \mathbf{Z})$. Die Normen dieser Zahlen liegen in $\mathbf{N}$. Nach (19) müssen daher notwendig die Einheiten die Norm 1 besitzen. Gilt umgekehrt für

$$\varepsilon = e_0 + e_1 i \qquad (e_0, e_1 \in \mathbf{Z})$$

die Gleichung

$$N(\varepsilon) = e_0^2 + e_1^2 = 1,$$

so ist $\bar{\varepsilon} = e_0 - e_1 i$ zu ε invers und daher ε Einheit. Genau diejenigen Zahlen $\varepsilon = e_0 + e_1 i$ aus G sind also Einheiten von G, für die $e_0^2 + e_1^2 = 1$ ist. Demnach sind die Einheiten von G

$$1, -1, i, -i.$$

Als nächstes zeigen wir, daß G ein euklidischer Ring ist. Sind $\beta = b_0 + b_1 i$ und $\alpha = a_0 + a_1 i$ aus G und ist $\alpha \neq 0$, so soll nachgewiesen werden, daß es eine Zahl $\varkappa = k_0 + k_1 i$ $(k_0, k_1 \in \mathbf{Z})$ gibt, für die

$$N(\beta - \alpha\varkappa) < N(\alpha)$$

gilt. Bekanntlich ist

$$\frac{\beta}{\alpha} = q_0 + q_1 i$$

mit

$$q_0 = \frac{b_0 a_0 + b_1 a_1}{a_0^2 + a_1^2} \in \mathbf{Q}$$

und

$$q_1 = \frac{b_1 a_0 - b_0 a_1}{a_0^2 + a_1^2} \in \mathbf{Q}.$$

Offensichtlich liegen in $\mathbf{Z}$ solche Zahlen k_0, k_1, daß folgendes gilt:

$$|q_0 - k_0| \leqq \frac{1}{2}, \qquad |q_1 - k_1| \leqq \frac{1}{2}.$$

Setzen wir $\varkappa = k_0 + k_1 i$, so ist

$$N\left(\frac{\beta}{\alpha} - \varkappa\right) = N\left((q_0 - k_0) + (q_1 - k_1)i\right) \leqq \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

und nach (18)

$$N(\beta - \alpha\varkappa) = N\left(\frac{\beta}{\alpha} - \varkappa\right) N(\alpha) < N(\alpha).$$

Damit ist gezeigt, daß G ein euklidischer Ring ist. Man bilde nämlich jedes $\alpha \in G$ ab auf $h(\alpha) := N(\alpha)$. In G ist also jedes Ideal ein Hauptideal, und es gelten die Sätze vom größten gemeinsamen Teiler und von der eindeutigen Primelementzerlegung.

Es sollen nun die Primelemente von G bestimmt werden. Da G ein Hauptidealring ist, sind die Primelemente genau die von den Einheiten verschiedenen unzerlegbaren Elemente. Aus (18) folgt sofort:

$$\alpha \in G \wedge N(\alpha) \text{ } \textit{Primzahl} \Rightarrow \alpha \text{ } \textit{Primelement von } G. \tag{20}$$

Für $\pi \in G$ ergibt sich aus $N(\pi) = \pi\bar{\pi}$, daß $\pi \mid N(\pi)$, wobei $N(\pi)$ als Element von G aufgefaßt wird. Ist π insbesondere Primelement von G, so muß (vgl. Definition 4) π Teiler einer derjenigen Primzahlen sein, die ihrerseits Teiler von $N(\pi)$ sind. π kann nicht Teiler zweier verschiedener Primzahlen p und q sein, denn weil die 1 als Vielfachsumme

$$1 = xp + yq \qquad (x, y \in \mathbf{Z})$$

darstellbar ist, müßte sonst $\pi \mid 1$ gelten. Das ist unmöglich, da π als Primelement keine Einheit ist. Also gilt:

$$\pi \text{ } \textit{Primelement von } G \Rightarrow \textit{es gibt genau eine Primzahl } p, \textit{ die durch } \pi \textit{ teilbar ist.} \tag{21}$$

Daher hat man in G die (nichtassoziierten) Primteiler der Primzahlen p zu bestimmen.

1. Sei $p = 2$. Es ist $2 = (1 + i)(1 - i)$. Da $N(1 + i) = 2$, ist $(1 + i)$ Primelement von G. $(1 - i) = (-i)(1 + i)$ ist zu $(1 + i)$ assoziiert. Weil in G der Satz von der eindeutigen Primelementzerlegung gilt, sind daher alle Primelemente, die 2 teilen, zu $(1 + i)$ assoziiert.

2. Sei $p \equiv 3 \mod 4$. Da $N(p) = p^2$ ist und $N(\pi) \mid N(p)$ aus $\pi \mid p$ folgt, muß $N(\pi) = p$ oder $N(\pi) = p^2$ sein. Für $n \in \mathbf{N}$ gilt:

$$2 \mid n \Rightarrow n^2 \equiv 0 \mod 4$$

und

$$2 \nmid n \Rightarrow n^2 \equiv 1 \mod 4.$$

Ist $\pi = x + yi$, so ergibt sich daraus

$$N(\pi) = x^2 + y^2 \not\equiv 3 \mod 4,$$

d. h., es muß $N(\pi) = p^2$ sein. Aus $p = \pi\alpha$ erhält man dann

$$N(\pi) = p^2 = N(p) = N(\pi)N(\alpha),$$

also $N(\alpha) = 1$. α ist daher Einheit und p Primelement von G.

3. Sei $p \equiv 1 \bmod 4$. Dann gibt es eine Zahl $z \in \mathbf{N}$, für die

$$z^2 \equiv -1 \bmod p$$

gilt (vgl. 13.4.4., Satz 3, Folgerung). Weil $z^2 + 1 = (z + i)(z - i)$ ist, folgt für ein Primelement π aus $\pi \mid p$ und $p \mid z^2 + 1$

$$\pi \mid z + i \qquad \text{oder} \qquad \pi \mid z - i.$$

Wäre π zu p assoziiert, so müßte auch

$$p \mid z + i \qquad \text{oder} \qquad p \mid z - i$$

sein, was wegen

$$\frac{z}{p} + \frac{i}{p} \notin G \wedge \frac{z}{p} - \frac{i}{p} \notin G$$

unmöglich ist. Daher kann p kein Primelement von G sein und muß also eine Zerlegung der Form

$$p = \alpha\pi \qquad \text{mit} \qquad \alpha \in G \wedge \pi \in G \wedge N(\alpha) > 1 \wedge N(\pi) > 1$$

gestatten. Wegen

$$N(p) = p^2 = N(\pi)N(\alpha)$$

ist dann

$$N(\pi) = N(\alpha) = p,$$

d. h., π und α sind Primelemente von G. Aus $\pi = x + yi$ ergibt sich

$$N(\pi) = p = x^2 + y^2 = (x + yi)\alpha$$

und also

$$\alpha = \frac{x^2 + y^2}{x + yi} = x - yi = \overline{\pi}.$$

Daher ist $p = \pi\overline{\pi}$. Man prüft sofort nach, daß π und $\overline{\pi}$ nicht assoziiert sind.

Sämtliche Primelemente von G sind also: $1 + i$; die Primzahlen p mit $p \equiv 3 \bmod 4$; die Zahlen $x + yi$ und $x - yi$, wo x, y ganzzahlige Lösungen der Gleichung $x^2 + y^2 = p$ (p Primzahl $\wedge\ p \equiv 1 \bmod 4$) bedeuten, sowie alle dazu assoziierten Elemente aus G.

Nach Satz 3 ist in einem Hauptidealring jedes unzerlegbare Element, das nicht Einheit ist, Primelement. Daher ist der Ring $\mathbf{Z}[\sqrt{-3}]$ aller Zahlen der Form $a + b\sqrt{-3}$ ($a, b \in \mathbf{Z}$) kein Hauptidealring (und erst recht kein euklidischer Ring), denn er enthält das Element 2, welches unzerlegbar, aber weder Einheit noch Primelement ist (vgl. 13.5.1.).

13.6. Quotientenkörper

Bekanntlich sind die Integritätsbereiche der ganzen Zahlen, der geraden ganzen Zahlen und der Gaußschen ganzen komplexen Zahlen Teilbereiche von Körpern. Sei der Integritätsbereich I Teilbereich des Körpers K. In diesem Fall heißt I *eingebettet* in K. I bestehe nicht nur aus dem Nullelement 0. Dann besitzt in K insbeson-

dere jede Gleichung

$$ax = b \qquad (a \in I \wedge b \in I \wedge a \neq 0)$$

genau eine Lösung

$$x = ba^{-1} = a^{-1}b,$$

die *Quotient* $\frac{b}{a} := ba^{-1}$ genannt wird.

Für die Quotienten gelten in K die Rechenregeln:

$$\frac{b}{a} = \frac{d}{c} \Leftrightarrow bc = ad, \tag{1}$$

$$\frac{b}{a} \pm \frac{d}{c} = \frac{bc \pm ad}{ac}, \tag{2}$$

$$\frac{b}{a} \cdot \frac{d}{c} = \frac{bd}{ac}. \tag{3}$$

Sie ergeben sich wegen der Nullteilerfreiheit von K aus den Gleichungen

$$ac\left(\frac{b}{a} - \frac{d}{c}\right) = bc - ad,$$

$$ac\left(\frac{b}{a} \pm \frac{d}{c} - \frac{bc \pm ad}{ac}\right) = 0,$$

$$ac\left(\frac{b}{a} \cdot \frac{d}{c} - \frac{bd}{ac}\right) = 0.$$

Aus den Regeln erkennt man, daß die Menge aller Quotienten $\frac{b}{a}$ $(a \neq 0)$ einen Teilkörper Q von K bildet. Das Einselement sind die Quotienten $\frac{a}{a}$ $(a \neq 0)$, zu $\frac{b}{a}$ $(a \neq 0 \wedge b \neq 0)$ invers ist $\frac{a}{b}$. I ist in Q eingebettet, denn die Elemente k aus I erscheinen als Quotienten $\frac{kx}{x}$ $(x \neq 0)$ in Q.

Die Regeln (1) bis (3) bedeuten, daß die Rechnung in Q bereits völlig durch die Rechnung in I festgelegt ist. Jeder Körper, der I umfaßt, enthält auch Q. In diesem Sinne ist Q der „kleinste" Körper, in welchen I eingebettet ist. Er wird durch I bis auf Isomorphie eindeutig bestimmt und heißt *Quotientenkörper* von I. Zu zwei verschiedenen Integritätsbereichen können gleiche (bzw. isomorphe) Quotientenkörper gehören; beispielsweise ist der Körper $\mathbf{Q}$ Quotientenkörper des Integritätsbereiches $\mathbf{Z}$ und des Integritätsbereiches aller geraden Zahlen aus $\mathbf{Z}$.

Bisher hatten wir vorausgesetzt, daß I in einen Körper eingebettet ist. Es gilt aber sogar der

Satz 1. *Zu jedem Integritätsbereich I gibt es einen bis auf Isomorphie eindeutig bestimmten Quotientenkörper.*

Beweis. Wir konstruieren einen Körper, der einen zu I isomorphen Integritätsbereich umfaßt. Die Existenz und Eindeutigkeit (bis auf Isomorphie) des Quotientenkörpers ergibt sich dann aus den bisherigen Feststellungen.

Es wird hier formal ebenso vorgegangen wie bei der Konstruktion des Körpers $\mathbf{Q}$ der rationalen Zahlen aus dem Integritätsbereich $\mathbf{Z}$ der ganzen Zahlen.

Es sei I nicht der Nullring. Dann bilden wir die Menge

$$P := \{(b, a) : b \in I \wedge a \in I \setminus \{0\}\}$$

aller geordneten Paare von Elementen aus I, deren zweite Komponente ungleich Null ist. Man rechnet leicht nach, daß durch

$$(b, a)\ R\ (d, c) :\Leftrightarrow bc = ad \qquad \big((b, a) \in P \wedge (d, c) \in P\big)$$

eine Äquivalenzrelation R in P gegeben ist. Sie vermittelt eine Einteilung der Elemente von P in Äquivalenzklassen (vgl. MfL, Bd. 1, 2.5.). Sei $[b, a]$ diejenige Äquivalenzklasse, in der $(b, a) \in P$ liegt. Die Menge dieser Äquivalenzklassen bezeichnen wir mit Q. Dann ist in Q

$$[b, a] = [d, c] \Leftrightarrow bc = ad. \tag{4}$$

Ferner gelten für die Elemente aus Q die Implikationen

$$[b, a] = [b', a'] \wedge [d, c] = [d', c'] \Rightarrow [bc + ad, ac] = [b'c' + a'd', a'c'] \tag{5}$$

und

$$[b, a] = [b', a'] \wedge [d, c] = [d', c'] \Rightarrow [bd, ac] = [b'd', a'c'], \tag{6}$$

denn es gilt

$$ba' = ab' \wedge dc' = cd' \Rightarrow bca'c' + ada'c' = acb'c' + aca'd'$$

und

$$ba' = ab' \wedge dc' = cd' \Rightarrow bda'c' = acb'd'.$$

Unter Verwendung der in I erklärten Addition kann man durch

$$[b, a] + [d, c] := [bc + ad, ac] \qquad ([b, a] \in Q, [d, c] \in Q) \tag{7}$$

eine „Addition" genannte Operation in Q definieren. Um die Summe zweier Elemente aus Q zu bestimmen, wählt man also aus der Klasse $[b, a]$ einen Vertreter (mit (b, a) bezeichnet) sowie aus der Klasse $[d, c]$ einen Vertreter (mit (d, c) bezeichnet) und bildet mit Hilfe der in I erklärten Addition das Paar $(bc + ad, ac)$, welches wegen der Nullteilerfreiheit von I in P liegt. Schließlich sucht man die Äquivalenzklasse $[bc + ad, ac]$ dieses Paares auf. Nach (5) ergibt sich dieselbe Klasse auch,

wenn man aus den Klassen $[b, a]$ und $[d, c]$ andere Vertreter auswählt. Die Definition (7) ist also „vertreterunabhängig". Sie ordnet je zwei Elementen aus Q eindeutig ein Element aus Q zu und ist demnach eine Operation in Q.

Ganz entsprechend wird durch

$$[b, a]\,[d, c] := [bd, ac] \qquad ([b, a] \in Q, [d, c] \in Q) \tag{8}$$

eine Multiplikation in Q erklärt. Die Vertreterunabhängigkeit dieser Definition folgt aus (6).

Für diese Operationen in Q gelten die Assoziativgesetze, Kommutativgesetze und das Distributivgesetz, weil sie für die Operationen in I gelten (Übungsaufgabe). Die Gleichung

$$[b, a] + [y, x] = [d, c] \qquad ([b, a] \in Q, [d, c] \in Q)$$

hat in Q die Lösung

$$[y, x] = [da - bc, ac].$$

Daher ist Q bezüglich der eben erklärten Operationen ein kommutativer Ring. Nullelement ist die Klasse $[0, n]$ aller Paare, deren erste Komponente 0 ist. Da

$$[b, a]\,[d, c] = [bd, ac] = [0, n] \Leftrightarrow bdn = 0$$

ist, folgt aus $n \neq 0$ und der Nullteilerfreiheit von I

$$[b, a]\,[d, c] = [0, n] \Leftrightarrow [b, a] = [0, n] \vee [d, c] = [0, n],$$

d. h., Q ist nullteilerfrei. Einselement von Q ist die Klasse $[n, n]$ aller Paare, deren beide Komponenten gleich sind. Ist $[b, a] \neq [0, n]$, so liegt auch $[a, b]$ in Q und ist wegen $[b, a]\,[a, b] = [ab, ab]$ zu $[b, a]$ invers. Q ist also ein Körper.

Bezeichnet a ein festes Element aus I und durchläuft x die Menge $I \setminus \{0\}$, so bilden die Paare (ax, x) eine Äquivalenzklasse $[ax, x]$. Die 1-1-Abbildung

$$f: a \mapsto [ax, x] \qquad (a \in I)$$

ist ein Isomorphismus von I auf einen Teilintegritätsbereich I' von Q, denn nach (7) und (8) gilt für alle $a, b \in I$

$$[ax, x] + [bx, x] = [(a + b)\,x^2, x^2] = [(a + b)\,x, x],$$

$$[ax, x]\,[bx, x] = [abx^2, x^2] = [abx, x].$$

Damit ist die Existenz eines Körpers bewiesen, der den zu I isomorphen Teilbereich I' umfaßt.

Ist $a \neq 0$, so hat die Gleichung $[ax, x]\,[v, u] = [by, y]$ die Lösung $[v, u] = [b, a]$. K besteht also aus allen Quotienten der Elemente von I' und ist demnach Quotientenkörper von I'.

13.7. Primkörper

Eine Teilmenge T von Elementen des Körpers K, die bezüglich der Operationen von K bereits selbst einen Körper bildet, heißt *Teilkörper* von K. Insbesondere spricht man im Fall $T \neq K$ von einem *echten Teilkörper*.

Definition 1. K heißt *Primkörper* $:\Leftrightarrow K$ ist Körper $\wedge$ K enthält keine echten Teilkörper.

Beispiele für Primkörper sind die Restklassenkörper $\mathbf{Z}/(p)$ nach einer Primzahl p und der Körper der rationalen Zahlen $\mathbf{Q}$. Enthält ein Teilkörper von $\mathbf{R}$ die Zahl 1, so alle Elemente der von ihr erzeugten additiven Gruppe, d. h. alle ganzen Zahlen und den durch sie bestimmten Quotientenkörper $\mathbf{Q}$. Jeder Restklassenkörper $\mathbf{Z}/(p)$ ist ebenfalls ein Primkörper, denn ein Teilkörper enthält zusammen mit der Restklasse [1] alle durch Addition daraus zu gewinnenden Restklassen und stimmt deshalb mit $\mathbf{Z}/(p)$ überein.

Satz 1. *Jeder Körper enthält genau einen Primkörper.*

Beweis. Bezeichnet K einen Körper und $\mathfrak{T}$ die Menge aller Teilkörper von K, so ist

$$D := \bigcap_{T \in \mathfrak{T}} T$$

ein Teilkörper von K. D ist Primkörper, denn ein echter Teilkörper von D wäre auch Teilkörper von K. Es ist D durch K eindeutig bestimmt.

Sei K ein Körper. Der Primkörper von K enthält sicher alle durch Addition und Subtraktion aus dem Einselement e von K entstehenden Elemente

$$ke := \begin{cases} e + e + \cdots + e & (k \text{ Summanden}), \text{ wenn } k > 0, \\ 0, & \text{wenn } k = 0, \\ -(e + e + \cdots + e) & (|k| \text{ Summanden}), \text{ wenn } k < 0. \end{cases}$$

Die Abbildung

$$f: k \mapsto ke \qquad (k \in \mathbf{Z})$$

ist ein Homomorphismus von $\mathbf{Z}$ in K, denn für $k, l \in \mathbf{Z}$ gilt

$$f(k + l) = (k + l)\, e = ke + le = f(k) + f(l)$$

und

$$f(kl) = (kl)\, e = (ke)\,(le) = f(k)\, f(l).$$

Die Elemente ke $(k \in \mathbf{Z})$ bilden also einen Teilintegritätsbereich I von K, und es gibt ein Ideal $\mathfrak{n}$ von $\mathbf{Z}$, so daß $\mathbf{Z}/\mathfrak{n} \cong I$ ist.

Fall 1: $\mathfrak{n} = (0)$. Dann ist $I \cong \mathbf{Z}$. Der Quotientenkörper von I ist der Primkörper von K. Er ist isomorph dem Quotientenkörper $\mathbf{Q}$ von $\mathbf{Z}$. Man nennt K in diesem Fall einen *Körper der Charakteristik* 0.

Fall 2: $\mathfrak{n} \neq (0)$. Da I nullteilerfrei ist und nicht nur aus dem Nullelement besteht, muß $\mathfrak{n} = (p)$ sein, wo p eine Primzahl bezeichnet (vgl. 13.4.2.). Aus der Isomorphie von I zum Restklassenkörper modulo p folgt, daß I bereits der Primkörper von K ist. In diesem Fall wird K als *Körper der Charakteristik* p bezeichnet. $\mathbf{Q}$ und $\mathbf{Z}/(p)$ (p Primzahl) sind also (bis auf Isomorphie) die einzigen Primkörper.

In einem Körper K der Charakteristik p gilt

$$(a + b)^p = a^p + b^p \qquad (a \in K \wedge b \in K),$$

da die Binomialkoeffizienten

$$\binom{p}{i} = \frac{p(p-1)\cdots(p-i+1)}{1 \cdot 2 \cdots i} \qquad (i = 1, 2, \ldots, p-1)$$

durch p teilbar sind. Außerdem ist $(ab)^p = a^p b^p$, und aus $a^p = b^p$ folgt wegen $0 = a^p - b^p = (a-b)^p$ und der Nullteilerfreiheit in K, daß $a = b$ sein muß. Daher ist die Abbildung

$$f\colon a \mapsto a^p \qquad (a \in K)$$

ein Isomorphismus von K auf den Teilkörper K^p, der aus allen p-ten Potenzen der Elemente von K besteht. Weil f 1-1-Abbildung ist, muß für einen endlichen Körper $K = K^p$ sein, d. h., in einem endlichen Körper K der Charakteristik p existiert zu jedem $a \in K$ genau ein $b \in K$, für das

$$b^p = a$$

gilt.

13.8. Übungsaufgaben

1. Sei L ein Integritätsbereich mit dem Einselement $\bar{1}$ und I ein vom Nullring verschiedener Teilintegritätsbereich von L mit dem Einselement 1. Man zeige, daß dann $\bar{1} = 1$ ist.
2. Man bestimme sämtliche Einheiten des Ringes $\mathbf{Z}\left[\sqrt{-3}\right]$ (vgl. 13.5.1.).
   In diesem Ring ist 2 ein Teiler des Produktes $\left(1 + \sqrt{-3}\right)\left(1 - \sqrt{-3}\right) = 4$. Man zeige, daß aber weder $1 + \sqrt{-3}$ noch $1 - \sqrt{-3}$ in $\mathbf{Z}\left[\sqrt{-3}\right]$ durch 2 teilbar ist.
3. Es sei

   $$P := \{(a,b) : a \in \mathbf{R} \wedge b \in \mathbf{R}\}$$

   die Menge aller geordneten Paare reeller Zahlen. Für beliebige Elemente (a_1, b_1) und (a_2, b_2) dieser Menge gilt

   $$(a_1, b_1) = (a_2, b_2) \Leftrightarrow a_1 = a_2 \wedge b_1 = b_2.$$

   Durch

   $$(a_1, b_1) + (a_2, b_2) := (a_1 + a_2, b_1 + b_2)$$

und

$$(a_1, b_1)\,(a_2, b_2) := (a_1a_2,\ a_1b_2 + a_2b_1)$$

werden in P ***Addition*** und ***Multiplikation*** genannte zweistellige Operationen erklärt. Man beweise, daß P bezüglich dieser Operationen ein kommutativer Ring ist. Ferner zeige man, daß er einen zum Körper der reellen Zahlen isomorphen Teilkörper enthält, und bestimme die Einheiten sowie die Nullteiler dieses Ringes.

4. In einem Integritätsbereich I kann jede Gleichung

$$ax^2 + bx + c = 0 \qquad (a, b, c \in I)$$

höchstens zwei Lösungen haben.

5. Man bestimme sämtliche Lösungen der Gleichung $x^2 = -1$ im Schiefkörper der Quaternionen (vgl. 13.2.).

6. Im Ring $\mathbf{Z}$ der ganzen rationalen Zahlen bestimme man das mengenmäßig kleinste Ideal $\mathfrak{n}$, welches die Zahlen 546, 498 und 210 enthält und gebe die Elemente des Restklassenringes $\mathbf{Z}/\mathfrak{n}$ an.

7. Man zeige, daß die Menge

$$\left\{x + y\sqrt{-2} : x \in \mathbf{Z} \wedge y \in \mathbf{Z}\right\}$$

mit der Addition und Multiplikation der komplexen Zahlen einen euklidischen Ring $\mathbf{Z}\left[\sqrt{-2}\right]$ bildet. Man bestimme die Einheiten dieses Ringes.
Ist auch $\mathbf{Z}\left[\sqrt{-3}\right]$ ein euklidischer Ring?

8. Im Ring G der Gaußschen ganzen komplexen Zahlen bestimme man die größten gemeinsamen Teiler der Elemente $2 + 4i$ und $5 + 5i$ und stelle sie als Vielfachsummen dar.

9. Man gebe die Primelementzerlegungen von $2 + 4i$ und $5 + 5i$ im Ring der Gaußschen ganzen komplexen Zahlen an.

10. Man bestimme den Quotientenkörper des Ringes der Gaußschen ganzen komplexen Zahlen.

11. Wie viele Körper aus genau zwei, drei und vier Elementen gibt es? Man stelle in jedem Fall die Additions- und Multiplikationstabellen auf.

14. Polynome

14.1. Polynome in einer Unbestimmten

Es sei I ein Integritätsbereich mit Einselement 1 und

$$P := \{a : a = (a_0, a_1, a_2, \ldots) \wedge \bigwedge_{n \in \mathbb{N}} a_n \in I \wedge \bigvee_{m \in \mathbb{N}} \bigwedge_{n \in \mathbb{N}} n > m \Rightarrow a_n = 0\}$$

die Menge aller unendlichen Folgen von Elementen aus I, in denen nur endlich viele Glieder vom Nullelement $0 \in I$ verschieden sind. Bezeichnen

$$a = (a_0, a_1, a_2, \ldots), \quad b = (b_0, b_1, b_2, \ldots)$$

zwei beliebige Elemente aus P, so ist (vgl. MfL, Bd. 1, 2.4.)

$$a = b \Leftrightarrow \bigwedge_{n \in \mathbb{N}} a_n = b_n. \tag{1}$$

Durch

$$a + b := (a_0 + b_0, a_1 + b_1, a_2 + b_2, \ldots) \tag{2}$$

wird eine Addition und durch

$$ab := \left(a_0 b_0, a_0 b_1 + a_1 b_0, a_0 b_2 + a_1 b_1 + a_2 b_0, \ldots, \sum_{i=0}^{n} a_i b_{n-i}, \ldots\right) \tag{3}$$

eine Multiplikation in P erklärt. Es ist leicht nachzuprüfen, daß P bezüglich dieser beiden Operationen einen kommutativen Ring bildet. Er hat das Nullelement $(0, 0, 0, \ldots)$ und das Einselement $(1, 0, 0, \ldots)$.

Die Teilmenge

$$I' = \{a : a = (a_0, 0, 0, \ldots) \wedge a_0 \in I\}$$

aller Folgen, in denen höchstens das vorderste Glied $a_0 \neq 0$ ist, bildet vermöge der Abbildung

$$f : a_0 \mapsto (a_0, 0, 0, \ldots)$$

einen zu I isomorphen Unterring von P, denn

$$f : a_0 + b_0 \mapsto (a_0 + b_0, 0, 0, \ldots) = (a_0, 0, 0, \ldots) + (b_0, 0, 0, \ldots)$$

und

$$f\colon a_0 b_0 \mapsto (a_0 b_0, 0, 0, \ldots) = (a_0, 0, 0, \ldots)\,(b_0, 0, 0, \ldots).$$

Identifizieren wir diese speziellen Folgen mit den Elementen aus I und schreiben also für alle Elemente a_0 aus I

$$a_0 = (a_0, 0, 0, \ldots), \tag{4}$$

so ist I ein echter Unterring von P. Daher nennt man P auch einen *Erweiterungsring* von I.

Setzen wir die Bezeichnung

$$x := (0, 1, 0, 0, \ldots)$$

fest, so ist

$$x^2 = (0, 0, 1, 0, \ldots),$$
$$x^3 = (0, 0, 0, 1, \ldots)$$

und für beliebige $n \in \mathbb{N}$ bedeutet

$$x^n = (a_0, a_1, a_2, a_3, \ldots) \qquad \text{mit} \qquad a_n = 1 \wedge \bigwedge_{i \in \mathbb{N} \setminus \{n\}} a_i = 0.$$

Dann ist

$$(a_0, a_1, a_2, \ldots) = (a_0, 0, 0, \ldots) + (a_1, 0, 0, \ldots)\,x + (a_2, 0, 0, \ldots)\,x^2 + \cdots$$

oder wegen (4)

$$(a_0, a_1, a_2, \ldots) = a_0 + a_1 x + a_2 x^2 + \cdots. \tag{5}$$

Die Summe auf der rechten Seite enthält höchstens endlich viele Summanden, die vom Nullelement verschieden sind.

Nach (4) wird das Nullelement von P mit 0 bezeichnet. Wegen (1) gilt

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m = 0 \Leftrightarrow a_0 = a_1 = \cdots = a_m = 0. \tag{6}$$

Man sagt für diesen Sachverhalt auch, daß je endlich viele der Elemente $1, x, x^2, \ldots$ über I *linear unabhängig* sind, und nennt x ein bezüglich I *transzendentes Element* oder eine *Unbestimmte*. Die Elemente von P, für die wir in Zukunft die durch (5) gegebene Schreibweise

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m \tag{7}$$

benutzen werden, heißen *Polynome in der Unbestimmten* x. Die $a_0, a_1, \ldots, a_m$ werden als *Koeffizienten* des Polynoms bezeichnet. Insbesondere nennt man a_0 das *absolute Glied* von $f(x)$. Ist $a_m \neq 0$, so heißt m *Grad des Polynoms* und a_m *höchster Koeffizient* von $f(x)$. Polynome des Grades 0 sind also die von 0 verschiedenen Elemente aus I. Dem Nullpolynom 0 wird kein Grad zugeschrieben. Die Operationen (2) und (3) bedeuten in der neuen Schreibweise einfach die bekannte Addition und Multiplikation von Polynomen.

Die eben durchgeführte Konstruktion, die zu dem Integritätsbereich I den kommutativen Ring P aller Polynome in der Unbestimmten x mit Koeffizienten aus I liefert, wird *Adjunktion* von x zu I (da das Resultat ein Ring ist, genauer: *Ringadjunktion*) genannt. Wir bezeichnen diesen Polynomring mit $P = I[x]$.

Satz 1. *I Integritätsbereich mit Einselement $1 \Rightarrow I[x]$ Integritätsbereich mit Einselement 1.*

Bezeichnen $f(x)$, $g(x)$ Polynome $\neq 0$ aus $I[x]$, so gilt:

$$\text{Grad}\,\big(f(x)\,g(x)\big) = \text{Grad}\,f(x) + \text{Grad}\,g(x).$$

Beweis. Nach den vorstehenden Bemerkungen braucht nur noch die Nullteilerfreiheit von $I[x]$ sowie die Aussage über die Grade gezeigt zu werden. Das Produkt zweier Polynome der Grade m und n ist

$$\begin{aligned} f(x)\,g(x) &= (a_0 + a_1x + \cdots + a_mx^m)\,(b_0 + b_1x + \cdots + b_nx^n) \\ &= a_0b_0 + (a_0b_1 + a_1b_0)\,x + \cdots + a_mb_nx^{m+n}. \end{aligned}$$

Weil I Integritätsbereich ist, folgt aus $a_m \neq 0$ und $b_n \neq 0$ auch $a_mb_n \neq 0$. Daher hat das Produkt $f(x)\,g(x)$ den Grad $m + n$ und ist nicht das Nullpolynom.

Nach der Definition 1 aus 13.5. ist $f(x) \in I[x]$ genau dann Teiler von $g(x) \in I[x]$ (Bezeichnung: $f(x) \mid g(x)$), wenn es in $I[x]$ ein Polynom $q(x)$ gibt, für das

$$f(x)\,q(x) = g(x)$$

gilt. Da das Einselement 1 von $I[x]$ ein Polynom vom Grade 0 ist, muß jede Einheit $e(x)$ von $I[x]$ nach der Bemerkung über den Grad eines Produktes vom Grade 0, also aus I sein. Die Einheiten von $I[x]$ sind also genau die Einheiten von I.

Sei L ein I (nicht notwendig echt) umfassender Integritätsbereich und

$$f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_mx^m \in I[x].$$

Durch

$$f\colon \alpha \mapsto f(\alpha) = a_0 + a_1\alpha + a_2\alpha^2 + \cdots + a_m\alpha^m \qquad (\alpha \in L)$$

ist eine Abbildung von L in L gegeben, die als *ganze rationale Funktion* mit Koeffizienten aus I über dem Definitionsbereich L bezeichnet wird. Ist

$$g\colon \alpha \mapsto g(\alpha) = b_0 + b_1\alpha + b_2\alpha^2 + \cdots + b_n\alpha^n \qquad (\alpha \in L)$$

eine weitere solche Funktion, so sind auch die Summe

$$f + g\colon \alpha \mapsto f(\alpha) + g(\alpha) = (a_0 + b_0) + (a_1 + b_1)\,\alpha + \cdots \qquad (\alpha \in L)$$

und das Produkt

$$fg\colon \alpha \mapsto f(\alpha)\,g(\alpha) = a_0b_0 + (a_0b_1 + a_1b_0)\,\alpha + \cdots \qquad (\alpha \in L)$$

ganze rationale Funktionen mit Koeffizienten aus I über dem Definitionsbereich L. Man prüft leicht nach, daß die Menge dieser Funktionen bezüglich der beiden ge-

nannten Operationen einen kommutativen Ring bildet. Sein Nullelement ist die „Nullfunktion", die jedes $\alpha \in L$ auf $0 \in L$ abbildet.

$$F: f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m \mapsto f \qquad \big(f(x) \in I[x]\big)$$

ist ein Homomorphismus von $I[x]$ auf den Ring der ganzen rationalen Funktionen mit Koeffizienten aus I über dem Definitionsbereich L. Genau dann liegt das Polynom

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m \in I[x]$$

im Kern dieses Homomorphismus, wenn

$$\bigwedge_{\alpha \in L} f(\alpha) = a_0 + a_1 \alpha + a_2 \alpha^2 + \cdots + a_m \alpha^m = 0.$$

Ist I (und daher auch L) ein unendlicher Integritätsbereich (z. B. einer der Zahlenbereiche $\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{C}$), so gilt

$$\bigwedge_{\alpha \in L} f(\alpha) = a_0 + a_1 \alpha + a_2 \alpha^2 + \cdots + a_m \alpha^m = 0 \Leftrightarrow a_0 = a_1 = \cdots = a_m = 0.$$

Betrachtet man nämlich die Gleichung

$$a_0 + a_1 \alpha + a_2 \alpha^2 + \cdots + a_m \alpha^m = 0$$

für $m + 1$ verschiedene $\alpha \in L$, so erhält man ein homogenes lineares Gleichungssystem für die Koeffizienten $a_0, a_1, \ldots, a_m$, das nur die Lösung $a_0 = a_1 = \cdots = a_m = 0$ besitzt. In diesem Fall ist daher der Polynomring $I[x]$ zum Ring der ganzen rationalen Funktionen mit Koeffizienten aus I über dem Definitionsbereich L isomorph.

In der Analysis interessiert man sich für die Funktionseigenschaft der ganzen rationalen Funktionen. Hier kommt es uns jedoch darauf an, daß es einen Erweiterungsring $I[x]$ von I gibt, der durch Adjunktion eines über I transzendenten Elementes x zu I entsteht. Seine Elemente nennen wir Polynome. Ist I ein unendlicher Bereich (z. B. $\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{C}$), so kann man sie sich als ganze rationale Funktionen mit Koeffizienten aus I vorstellen. Wie unsere Konstruktion zeigt, muß man das aber nicht tun, sondern kann sich die Elemente des Polynomrings $I[x]$ auch als spezielle Folgen von Elementen aus I denken, d. h., die Funktionseigenschaft ist für uns unwesentlich. Bezeichnen x und y Unbestimmte, so sind offensichtlich $I[x]$ und $I[y]$ isomorph.

14.2. Polynome über einem Körper und einem Integritätsbereich. Zerlegung in irreduzible Faktoren

14.2.1. Für den Integritätsbereich aller Polynome in x mit Koeffizienten aus einem Körper gilt der

Satz 1. *Bezeichnet K einen Körper, so ist $K[x]$ ein euklidischer Ring.*

Beweis. h bilde jedes vom Nullpolynom verschiedene $f(x) \in K[x]$ auf $h\big(f(x)\big)$:= Grad von $f(x)$ ab. Wir haben zu zeigen, daß zu beliebigen Polynomen $a(x) \neq 0$ und $b(x)$ aus $K[x]$ immer Polynome $q(x)$ und $r(x)$ in $K[x]$ existieren, für die

$$b(x) - a(x)\,q(x) = r(x)$$

mit $r(x) = 0$ oder $h\big(r(x)\big) < h\big(a(x)\big)$ gilt. Ist $b(x) = 0$, so wird diese Bedingung von $q(x) = r(x) = 0$ erfüllt. Daher sei

$$a(x) = a_0 + a_1 x + \cdots + a_m x^m \qquad (a_m \neq 0)$$

und

$$b(x) = b_0 + b_1 x + \cdots + b_n x^n \qquad (b_n \neq 0).$$

Ist $n < m$, so wird die Bedingung von $q(x) = 0$ und $r(x) = b(x)$ erfüllt, da ja $h\big(b(x)\big) = n < m = h\big(a(x)\big)$ ist. Ist schließlich $n \geqq m$, so hat

$$b_1(x) := b(x) - a_m^{-1} b_n x^{n-m} a(x)$$

einen Grad $n_1 < n$. Ist $n_1 < m$, so setze man $q(x) = a_m^{-1} b_n x^{n-m}$ und $r(x) = b_1(x)$. Ist aber $n_1 \geqq m$, so kann man von $b_1(x)$ wieder ein geeignetes Vielfaches von $a(x)$ subtrahieren, so daß ein Polynom $b_2(x)$ mit einem Grad $< n_1$ entsteht. Nach endlich vielen Wiederholungen findet man ein solches $q(x) = a_m^{-1} b_n x^{n-m} + \cdots$, so daß für $r(x) = b(x) - a(x)\,q(x)$ gilt: $r(x) = 0 \vee h\big(r(x)\big) < m$. Aus früheren Überlegungen (vgl. 13.5., Sätze 2, 5, 6) ergibt sich die

Folgerung 1. *$K[x]$ ist Hauptidealring. Daher gelten in $K[x]$ die Sätze vom größten gemeinsamen Teiler und von der eindeutigen Primelementzerlegung.*

Einheiten von $K[x]$ sind alle von 0 verschiedenen Elemente aus K, und $p(x) \in K[x]$ ist genau dann Primelement von $K[x]$, wenn Grad $p(x) > 0$ ist und $p(x)$ nur triviale Teiler besitzt (vgl. 13.5., Sätze 1, 3). Solche Polynome $p(x)$ werden *in $K[x]$ unzerlegbar* oder *irreduzibel* (vgl. 13.5., Definition 3) oder auch *Primpolynome von $K[x]$* genannt. Die Zerlegbarkeit eines Polynoms hängt von dem Körper ab, der für die Koeffizienten der Faktoren zugelassen wird. So ist $p(x) = x^2 - 2$ in $\mathbb{Q}[x]$ unzerlegbar, wegen $x^2 - 2 = \big(x - \sqrt{2}\big)\big(x + \sqrt{2}\big)$ aber in $\mathbb{R}[x]$ zerlegbar.

Sei L ein K umfassender Körper und bezeichne $d(x)$ einen größten gemeinsamen Teiler der Polynome $f(x)$, $g(x)$ aus $K[x]$ in $L[x]$. Dann gibt es einen zu $d(x) \in L[x]$ assoziierten größten gemeinsamen Teiler $d_1(x)$, der bereits in $K[x]$ liegt, denn bei der Bestimmung eines größten gemeinsamen Teilers von $f(x)$ und $g(x)$ mit dem eukli-

dischen Algorithmus treten nur Polynome aus $K[x]$ auf. Insbesondere folgt: Sind $f(x)$ und $g(x)$ in $K[x]$ teilerfremd (d. h., die von 0 verschiedenen Elemente aus K sind die einzigen gemeinsamen Teiler), so auch in $L[x]$.

14.2.2. Wir betrachten jetzt einen Integritätsbereich I mit Einselement 1 und eindeutiger Primelementzerlegung. Aus dieser Voraussetzung folgt sofort, daß je endlich viele Elemente aus I einen bis auf Einheitsfaktoren eindeutig bestimmten größten gemeinsamen Teiler besitzen (Übungsaufgabe). Jedes unzerlegbare Element $p \neq 0$ aus I, das nicht Einheit ist, muß notwendig Primelement sein. Es sei $p(x) \in I[x]$ und $p(x) \neq 0$. Dann heißt

$$p(x) = a_0 + a_1 x + \cdots + a_m x^m \textit{ primitives Polynom}: \Leftrightarrow a_0 \sqcap a_1 \sqcap \cdots \sqcap a_m = 1.$$

Als *Hilfssatz von* Gauss bezeichnet man den

Satz 2. *In $I[x]$ gilt: $f(x)$, $g(x)$ primitive Polynome $\Rightarrow$ $f(x)\,g(x)$ primitives Polynom.*

Beweis. Die Koeffizienten des Polynoms $f(x)\,g(x) = h(x)$ haben genau dann keinen von einer Einheit verschiedenen größten gemeinsamen Teiler, wenn es kein Primelement von I gibt, welches gemeinsamer Teiler aller dieser Koeffizienten ist. Sei

$$f(x) = a_0 + a_1 x + \cdots + a_i x^i + \cdots + a_m x^m \qquad (a_m \neq 0),$$
$$g(x) = b_0 + b_1 x + \cdots + b_k x^k + \cdots + b_n x^n \qquad (b_n \neq 0)$$

und p ein beliebiges Primelement von I. Unter den nicht durch p teilbaren Koeffizienten von $f(x)$ bzw. $g(x)$ habe a_i bzw. b_k maximalen Index. Da $f(x)$ und $g(x)$ primitiv sind, gibt es solche Koeffizienten. Setzt man

$$a_{m+1} = a_{m+2} = \cdots = b_{n+1} = b_{n+2} = \cdots = 0,$$

so kann der Koeffizient von x^{i+k} im Produkt

$$f(x)\,g(x) = \sum_{j=0}^{m+n} c_j x^j$$

in der Form

$$c_{i+k} = a_0 b_{i+k} + a_1 b_{i+k-1} + \cdots + a_i b_k + \cdots + a_{i+k-1} b_1 + a_{i+k} b_0$$

geschrieben werden. Aus $p \nmid a_i \wedge p \nmid b_k$ folgt $p \nmid a_i b_k$. Alle übrigen Summanden der rechten Seite sind durch p teilbar. Daher ist $p \nmid c_{i+k}$. $f(x)\,g(x)$ muß also primitiv sein.

Aus dem Beweis folgt sofort:

$$p \mid f(x)\,g(x) \Rightarrow p \mid f(x) \vee p \mid g(x).$$

Daher sind die Primelemente von I auch Primelemente von $I[x]$.

Die Einheiten von I und $I[x]$ stimmen überein.

Sei K der Quotientenkörper von I. Jedes von 0 verschiedene Polynom $\varphi(x) \in K[x]$ läßt sich in der Form

$$\varphi(x) = \varrho f^*(x) \qquad (\varrho \in K \wedge f^*(x) \textit{ primitives Polynom aus } I[x]) \tag{1}$$

schreiben. Dabei sind ϱ und $f^*(x)$ bis auf Einheitsfaktoren aus $I[x]$ eindeutig bestimmt.

Zum Beweis nehme man ein von 0 verschiedenes $a \in I$, für das $a\varphi(x)$ in $I[x]$ liegt. (Ein solches a ist z. B. das Produkt aller in den Koeffizienten von $\varphi(x)$ auftretenden Nenner.) Ist dann b größter gemeinsamer Teiler der Koeffizienten von $a\varphi(x)$, so ist $f^*(x)$ ein primitives Polynom aus $I[x]$, für das $a\varphi(x) = bf^*(x)$ gilt. (1) wird also mit $\varrho = a^{-1}b$ erfüllt. Gilt neben (1) auch noch

$$\varphi(x) = \sigma g^*(x) \qquad (\sigma \in K \wedge g^*(x) \text{ primitives Polynom aus } I[x]),$$

so ist $\varrho f^*(x) = \sigma g^*(x)$. Es sei $c \neq 0$ ein Element aus I, für das $c\varrho$ und $c\sigma$ in I liegen. Wegen

$$c\varphi(x) = c\varrho f^*(x) = c\sigma g^*(x)$$

sind $c\varrho$ und $c\sigma$ größte gemeinsame Teiler der Koeffizienten desselben Polynoms aus $I[x]$ und daher in I assoziiert. Hieraus folgt, daß auch $f^*(x)$ und $g^*(x)$ in $I[x]$ assoziiert sind.

Ist $f^*(x)$ ein primitives Polynom aus $I[x]$ und sind b und $a \neq 0$ aus I, so gilt

$$a^{-1}bf^*(x) \in I[x] \Leftrightarrow a \mid b. \tag{2}$$

Es sei $a^{-1}bf^*(x) \in I[x]$. Ist $c \in I$ größter gemeinsamer Teiler der Koeffizienten dieses Polynoms, so gibt es eine Darstellung

$$a^{-1}bf^*(x) = cg^*(x)$$

mit einem primitiven Polynom $g^*(x)$ aus $I[x]$. Nach der eben bewiesenen Eindeutigkeitsaussage ist $a^{-1}b = ce$ mit einer Einheit $e \in I$, d. h. $a \mid b$. Die andere Richtung der Äquivalenz (2) ist trivial.

Satz 3 (Satz von Gauss). *Besitzt ein Polynom $f(x) \in I[x]$ in $K[x]$ eine Zerlegung in Faktoren positiven Grades, so gibt es bereits in $I[x]$ eine solche Zerlegung.*

Beweis. Es sei

$$f(x) = \varphi_1(x)\,\varphi_2(x) \qquad (\varphi_1(x), \varphi_2(x) \in K[x]).$$

Wendet man (1) auf $\varphi_1(x)$ und $\varphi_2(x)$ an, so erhält man die Existenz von Elementen $a_1 \neq 0$, $a_2 \neq 0$, b_1, b_2 in I und primitiver Polynome positiven Grades $f_1^*(x)$, $f_2^*(x)$ in $I[x]$, für die

$$f(x) = a_1^{-1}b_1f_1^*(x)\,a_2^{-1}b_2f_2^*(x) = (a_1a_2)^{-1}\,(b_1b_2)\,f_1^*(x)\,f_2^*(x)$$

gilt. Weil $f_1^*(x)\,f_2^*(x)$ nach Satz 2 primitiv ist, folgt aus (2), daß $a_1a_2 \mid b_1b_2$ und daher $f(x)$ sogar in $I[x]$ zerlegbar ist.

Ist also $f(x) \in I[x]$ ein primitives, in $I[x]$ unzerlegbares Polynom positiven Grades, so ist $f(x)$ unzerlegbar in $K[x]$. Nach Folgerung 1 ist $K[x]$ Hauptidealring und daher $f(x)$ Primpolynom von $K[x]$. Aus

$$f(x) \mid g(x)\, h(x) \qquad (g(x), h(x) \in I[x])$$

folgt daher bei Betrachtung in $K[x]$

$$f(x) \mid g(x) \vee f(x) \mid h(x).$$

O. B. d. A. sei $f(x) \mid g(x)$, d. h., es gibt ein Polynom $\varphi_1(x) \in K[x]$, für das $f(x)\, \varphi_1(x) = g(x)$ gilt. Anwendung von (1) auf $\varphi_1(x)$ ergibt die Existenz von Elementen $a \neq 0$ und b aus I sowie eines primitiven Polynoms $f_1^*(x) \in I[x]$ mit der Eigenschaft

$$a^{-1} b f(x)\, f_1^*(x) = g(x).$$

Nach Satz 2 ist $f(x)\, f_1^*(x)$ ein primitives Polynom aus $I[x]$, und aus (2) folgt dann $a \mid b$. Daher gilt schon in $I[x]$

$$f(x) \mid g(x).$$

$f(x)$ ist also ein Primpolynom von $I[x]$.

Damit ist gezeigt, daß die Primelemente von I und die primitiven, irreduziblen Polynome positiven Grades in $I[x]$ Primelemente von $I[x]$ sind. Da jedes Primelement notwendig unzerlegbar ist (vgl. 13.5., Satz 1), gibt es keine weiteren Primelemente in $I[x]$.

Nach diesen Vorbereitungen ergibt sich der

Satz 4. *Ist I ein Integritätsbereich mit Einselement und eindeutiger Primelementzerlegung, so besitzt auch $I[x]$ eine eindeutige Primelementzerlegung.*

Beweis. K bezeichne wieder den Quotientenkörper von I. Nach Folgerung 1 gilt in $K[x]$ der Satz von der eindeutigen Primelementzerlegung. Ist $f(x) \neq 0$ aus $I[x]$, so besitzt es eine Darstellung der Form

$$f(x) = a^{-1} b \varrho_1(x) \cdots \varrho_n(x),$$

wobei $a \neq 0$ und b in I liegen und $\varrho_1(x), \ldots, \varrho_n(x)$ bis auf Einheitsfaktoren aus K eindeutig bestimmte Primpolynome aus $K[x]$ sind. Nach (1) kann jedes $\varrho_i(x)$ mit Elementen $b_i, a_i \neq 0$ aus I als

$$\varrho_i(x) = a_i^{-1} b_i r_i^*(x) \qquad (i = 1, \ldots, n)$$

geschrieben werden. Dabei bezeichnen die $r_i^*(x)$ primitive, in $I[x]$ unzerlegbare Polynome aus $I[x]$, die bis auf Einheitsfaktoren aus $I[x]$ eindeutig bestimmt sind. Dann ist

$$f(x) = (a a_1 \cdots a_n)^{-1} (b b_1 \cdots b_n)\, r_1^*(x) \cdots r_n^*(x).$$

Weil $r_1^*(x) \cdots r_n^*(x)$ nach Satz 2 ein primitives Polynom aus $I[x]$ ist, folgt aus (2)

$$c := (aa_1 \cdots a_n)^{-1} (bb_1 \cdots b_n) \in I.$$

Da die $r_i^*(x)$ bis auf Einheiten aus $I[x]$ eindeutig festliegen, ist auch c durch $f(x)$ bis auf einen Einheitsfaktor eindeutig bestimmt (vgl. 13.2.(1)). c besitzt in I eine bis auf Einheitsfaktoren eindeutige Primelementzerlegung

$$c = p_1 \cdots p_m.$$

Also gibt es eine Darstellung

$$f(x) = p_1 \cdots p_m r_1^*(x) \cdots r_n^*(x)$$

von $f(x)$ als Produkt von Primelementen aus $I[x]$, welche bis auf Einheitsfaktoren aus $I[x]$ durch $f(x)$ eindeutig bestimmt sind.

14.3. Nullstellen von Polynomen

14.3.1. Sei I ein Integritätsbereich mit Einselement 1 und L ein I umfassender Integritätsbereich, wobei auch $L = I$ zugelassen ist. Bezeichnet

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m \qquad (a_m \neq 0)$$

ein Polynom aus $I[x]$ und α ein Element aus L, so heißt der durch die Ersetzung von x durch α entstehende Ausdruck

$$f(\alpha) = a_0 + a_1 \alpha + a_2 \alpha^2 + \cdots + a_m \alpha^m$$

ein *Polynom in* α. $f(\alpha)$ ist ein wohlbestimmtes Element aus L. Man rechnet sofort nach (Übungsaufgabe), daß die durch

$$f(x) \mapsto f(\alpha)$$

gegebene Abbildung ein Homomorphismus von $I[x]$ in L ist. $f(\alpha)$ wird der *Wert des Polynoms* $f(x)$ *an der Stelle* $x = \alpha$ genannt.

Definition 1. Seien $I \subseteq L$ Integritätsbereiche mit Einselement. Bezeichnet 0 das Nullelement und sind $\alpha \in L$, $f(x) \in I[x]$, so heißt

$$\alpha \textit{ Nullstelle von } f(x) \text{ (oder } \textit{Wurzel der Gleichung } f(x) = 0) :\Leftrightarrow f(\alpha) = 0.$$

Die Division von $f(x)$ durch $x - \alpha$ ist in $L[x]$ ausführbar, da $x - \alpha$ den höchsten Koeffizienten 1 hat. Es gibt also ein Polynom $f_1(x) \in L[x]$ und ein $\varrho \in L$ mit der Eigenschaft

$$f(x) = (x - \alpha) f_1(x) + \varrho.$$

Für eine Nullstelle α von $f(x)$ ergibt sich

$$f(\alpha) = 0 f_1(\alpha) + \varrho = \varrho = 0,$$

und daher gilt:

$$\alpha \in L \textit{ ist Nullstelle von } f(x) \in I[x] \Rightarrow f(x) \textit{ ist in } L[x] \textit{ teilbar durch } (x - \alpha). \tag{1}$$

Ist α auch noch Nullstelle von $f_1(x)$, so ist $f_1(x)$ ebenfalls durch $(x - \alpha)$ teilbar:

$$f(x) = (x - \alpha)^2 f_2(x) \qquad \big(f_2(x) \in L[x]\big).$$

Falls $f_2(\alpha) = 0$ ist, läßt sich nochmals ein Faktor $(x - \alpha)$ abspalten. Weil dieses Verfahren nach spätestens $m = \operatorname{Grad} f(x)$ Schritten abbrechen muß, gibt es ein $k \in \mathbb{N}^*$ und ein $f_k(x) \in L[x]$, so daß

$$f(x) = (x - \alpha)^k f_k(x) \qquad \text{und} \qquad f_k(\alpha) \neq 0$$

ist. α wird dann eine *k-fache Nullstelle* von $f(x)$ genannt.

Ist $\beta \in L$ eine von α verschiedene l-fache Nullstelle von $f(x)$, so muß $f(x)$ durch $(x - \beta)^l$ teilbar sein. Da $(x - \alpha)^k$ wegen $(\beta - \alpha)^k \neq 0$ nicht durch $(x - \beta)$ teilbar sein kann, ist

$$f_k(x) = (x - \beta)^l f_{k+l}(x)$$

und also

$$f(x) = (x - \alpha)^k (x - \beta)^l f_{k+l}(x) \qquad \big(f_{k+l}(x) \in L[x]\big).$$

Liegen noch weitere Nullstellen von $f(x)$ in L, so kann man wiederum *Linearfaktoren*, das sind Faktoren ersten Grades, von $f_{k+l}(x)$ abspalten. $f(x)$ kann als Polynom m-ten Grades aber niemals als Produkt mit mehr als m Linearfaktoren darstellbar sein. Damit ist folgender Satz bewiesen.

Satz 1. *Seien $L \supseteq I$ Integritätsbereiche mit Einselement und $f(x) \in I[x]$. Dann gilt:*

$$\operatorname{Grad} f(x) = m \Rightarrow f(x) \textit{ besitzt in } L \textit{ höchstens } m \textit{ Nullstellen.}$$

14.3.2. Definition 2. Sei I ein Integritätsbereich mit Einselement 1 und

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m \in I[x].$$

Dann heißt

$$f'(x) \in I[x] \textit{ Ableitung von } f(x) :\Leftrightarrow f'(x) = a_1 + 2a_2 x + \cdots + m a_m x^{m-1}.$$

Diese Definition bedeutet, daß für Polynome $f(x)$ vom Grade 0 und für das Nullpolynom die Ableitung $f'(x) = 0$ ist. Für $f(x), g(x) \in I[x]$ ergeben sich allein aus der Definition die Regeln

$$\big(f(x) + g(x)\big)' = f'(x) + g'(x) \tag{2}$$

und

$$\big(f(x)\, g(x)\big)' = f'(x)\, g(x) + f(x)\, g'(x). \tag{3}$$

Zum Beweis schreiben wir die Polynome in der Form

$$f(x) = \sum_{i=0}^{\infty} a_i x^i, \qquad g(x) = \sum_{i=0}^{\infty} b_i x^i,$$

wobei $x^0 := 1$ und jede der Summen nur endlich viele vom Nullelement 0 verschiedene Koeffizienten a_i bzw. b_i enthält. Dann ist

$$f'(x) = \sum_{i=1}^{\infty} i a_i x^{i-1}, \qquad g'(x) = \sum_{i=1}^{\infty} i b_i x^{i-1},$$

und es ergibt sich

$$\begin{aligned}(f(x) + g(x))' &= \left(\sum_{i=0}^{\infty} (a_i + b_i) x^i\right)' \\ &= \sum_{i=1}^{\infty} i(a_i + b_i) x^{i-1} \\ &= \sum_{i=1}^{\infty} i a_i x^{i-1} + \sum_{i=1}^{\infty} i b_i x^{i-1} \\ &= f'(x) + g'(x).\end{aligned}$$

(3) kann ebenfalls durch einfache Rechnung bewiesen werden:

$$\begin{aligned}(f(x)\,g(x))' &= \left(\left(\sum_{i=0}^{\infty} a_i x^i\right)\left(\sum_{k=0}^{\infty} b_k x^k\right)\right)' = \left(\sum_{l=0}^{\infty}\left(\sum_{i+k=l} a_i b_k\right) x^l\right)' \\ &= \sum_{l=1}^{\infty} l\left(\sum_{i+k=l} a_i b_k\right) x^{l-1} = \sum_{l=1}^{\infty}\left(\sum_{i+k=l} (i+k)\, a_i b_k\right) x^{l-1} \\ &= \sum_{l=1}^{\infty}\left(\sum_{i+k=l} i a_i b_k\right) x^{l-1} + \sum_{l=1}^{\infty}\left(\sum_{i+k=l} k a_i b_k\right) x^{l-1} \\ &= \sum_{m=0}^{\infty}\left(\sum_{i+k=m} (i+1)\, a_{i+1} b_k\right) x^m + \sum_{m=0}^{\infty}\left(\sum_{i+k=m} a_i (k+1)\, b_{k+1}\right) x^m \\ &= \left(\sum_{i=0}^{\infty} (i+1)\, a_{i+1} x^i\right)\left(\sum_{k=0}^{\infty} b_k x^k\right) + \left(\sum_{i=0}^{\infty} a_i x^i\right)\left(\sum_{k=0}^{\infty} (k+1)\, b_{k+1} x^k\right) \\ &= f'(x)\, g(x) + f(x)\, g'(x).\end{aligned}$$

Die Ableitung $f'(x)$ von $f(x) = a_0 + a_1 x + \cdots + a_m x^m \in I[x]$ ist genau dann das Nullpolynom 0, wenn

$$i a_i = 0 \qquad (i = 1, \ldots, m) \tag{4}$$

(vgl. 14.1.(6)). Als Charakteristik von I bezeichnen wir die bereits durch I festgelegte Charakteristik des Quotientenkörpers von I. Ist sie 0, so bedeutet (4) $a_1 = a_2 = \cdots = a_m = 0$. Ist sie aber eine Primzahl p, so folgt aus (4) nur $a_i = 0$,

wenn $p \nmid i$. Damit ist gezeigt:

I hat die Charakteristik $0 \Rightarrow \big(f'(x) = 0 \Leftrightarrow f(x) = a_0 \in I\big)$, (5)

I hat die Charakteristik p

$$\Rightarrow \big(f'(x) = 0 \Leftrightarrow f(x) = a_0 + a_p x^p + a_{2p} x^{2p} + \cdots + a_{lp} x^{lp}\big). \tag{6}$$

I und L seien wie in 14.3.1. erklärt. Bezeichnet $\alpha \in L$ eine k-fache Nullstelle von $f(x) \in I[x]$, ist also

$$f(x) = (x - \alpha)^k f_k(x) \qquad \text{und} \qquad f_k(\alpha) \neq 0 \qquad \big(f_k(x) \in L[x]\big),$$

so folgt aus (3)

$$\begin{aligned} f'(x) &= k(x - \alpha)^{k-1} f_k(x) + (x - \alpha)^k f_k'(x) \\ &= (x - \alpha)^{k-1} [k f_k(x) + (x - \alpha) f_k'(x)], \end{aligned}$$

d. h., $f'(x)$ ist in $L[x]$ mindestens durch $(x - \alpha)^{k-1}$ teilbar. Hat I die Charakteristik 0, so folgt aus $k f_k(\alpha) \neq 0$, daß $k f_k(x) + (x - \alpha) f_k'(x)$ nicht durch $(x - \alpha)$ teilbar ist. In diesem Fall ist α also genau $(k - 1)$-fache Nullstelle von $f'(x)$. Die Aussage stimmt nicht mehr, wenn I die Charakteristik p besitzt. In diesem Fall hat beispielsweise das Polynom $f(x) = (x - \alpha)^p$ die Ableitung $f'(x) = 0$.

Erklärt man nach der Festsetzung $f^{(0)}(x) := f(x)$ durch

$$f^{(i)}(x) := \big(f^{(i-1)}(x)\big)' \qquad (i = 1, 2, \ldots)$$

sukzessive *höhere Ableitungen* von $f(x)$, so ergibt sich im Fall der Charakteristik 0, daß eine k-fache Nullstelle von $f(x) \in I[x]$ $(k - i)$-fache Nullstelle von $f^{(i)}(x)$ $(i = 1, 2, \ldots, k - 1)$ ist.

Wir fassen unsere Ergebnisse zusammen im

Satz 2. *I sei ein Integritätsbereich mit Einselement und α eine k-fache Nullstelle von $f(x) \in I[x]$. Hat I die Charakteristik 0, so ist α $(k - 1)$-fache Nullstelle von $f'(x)$, und es gilt*

$$f(\alpha) = f(\alpha) = \cdots = f^{(k-1)}(\alpha) = 0, \qquad f^{(k)}(\alpha) \neq 0.$$

Hat I die Charakteristik p, so ist $f'(x)$ mindestens durch $(x - \alpha)^{k-1}$ teilbar.

Definition 3. Bezeichnen K und L Körper, so heißt
L Erweiterungskörper (oder *Körpererweiterung*) von $K :\Leftrightarrow K$ ist Teilkörper von L.

Satz 3. *Sei K ein Körper, $f(x) \in K[x]$ und $p(x)$ ein Primpolynom von $K[x]$. Besitzen $f(x)$, $p(x)$ in einem Erweiterungskörper von K eine gemeinsame Nullstelle, so gilt bereits in $K[x]$: $p(x) \mid f(x)$.*

Beweis. Sei α aus dem Erweiterungskörper L von K gemeinsame Nullstelle von $f(x)$ und $p(x)$. Dann sind $f(x)$ und $p(x)$ in $L[x]$ nicht teilerfremd, weil sie den gemeinsamen Teiler $(x - \alpha)$ besitzen. Nach der Bemerkung am Schluß von 14.2.1. haben $f(x)$ und $p(x)$ daher schon in $K[x]$ einen größten gemeinsamen Teiler $d(x)$ von posi-

tivem Grad. Da $p(x)$ Primpolynom von $K[x]$ ist, muß $d(x)$ zu $p(x)$ assoziiert sein. Deshalb ist auch $p(x)$ ein Teiler von $f(x)$.

Dieser Satz gestattet eine interessante

Folgerung. *Für einen Körper K der Charakteristik 0 gilt: Ist $p(x)$ Primpolynom aus $K[x]$, so besitzt $p(x)$ in jedem Erweiterungskörper von K nur einfache Nullstellen.*

Beweis. Wegen der Voraussetzung über die Charakteristik ist $p'(x)$ nicht das Nullpolynom. Hätte $p(x)$ in einem Erweiterungskörper von K eine k-fache Nullstelle α $(k \geqq 2)$, so wäre α $(k-1)$-fache Nullstelle von $p'(x)$, und nach Satz 3 würde gelten $p(x) \mid p'(x)$ in $K[x]$. Das ist unmöglich, weil Grad $p(x) >$ Grad $p'(x)$ ist.

Es bezeichne K weiterhin einen Körper der Charakteristik 0. $f(x) \in K[x]$ sei ein Polynom des Grades $m > 0$. Wir wollen uns für die Nullstellen von $f(x)$ interessieren. Daher können wir o. B. d. A. annehmen, daß $f(x)$ den höchsten Koeffizienten 1 hat. (Sonst dividiere man durch diesen Koeffizienten. Dabei ändern sich die Nullstellen nicht.) Weiter wollen wir annehmen, daß es einen Erweiterungskörper L von K gibt, in dem $f(x)$ genau m Nullstellen besitzt, wobei jede so oft gezählt wird, wie ihre Vielfachheit angibt. Später werden wir zeigen, daß es solche Erweiterungskörper immer gibt. Sind $\alpha_1, \ldots, \alpha_r$ die verschiedenen Nullstellen von $f(x)$ in L, so zerfällt $f(x)$ in $L[x]$ in ein Produkt

$$f(x) = (x-\alpha_1)^{k_1}(x-\alpha_2)^{k_2}\cdots(x-\alpha_r)^{k_r} \tag{7}$$

von Linearfaktoren, wobei k_i die Vielfachheit der Nullstelle α_i $(i = 1, \ldots, r)$ angibt.

Bezeichnet $f_k(x)$ das Produkt derjenigen Linearfaktoren, die in (7) mit dem Exponenten k auftreten, so kann $f(x)$ in der Form

$$f(x) = f_1(x)\, f_2^2(x)\cdots f_t^t(x)$$

geschrieben werden, wobei $f_k(x) := 1$ bedeuten soll, wenn in (7) kein Linearfaktor mit dem Exponenten k vorkommt. Weil k-fache Nullstellen von $f(x)$ genau $(k-1)$-fache Nullstellen von $f'(x)$ sind, ist

$$d_1(x) := f(x) \sqcap f'(x) = f_2(x)\, f_3^2(x)\cdots f_t^{t-1}(x).$$

Seine Berechnung kann mit dem euklidischen Algorithmus bereits in $K[x]$ erfolgen.

$$g(x) := \frac{f(x)}{d_1(x)} = f_1(x)\, f_2(x)\cdots f_t(x) \tag{8}$$

ist dann ein Polynom aus $K[x]$, das die gleichen Nullstellen wie $f(x)$ besitzt, aber jede nur mit der Vielfachheit 1.

Auch die Faktoren $f_k(x)$ können durch Rechnungen in $K[x]$ bestimmt werden. Sei nämlich

$$\begin{aligned} &d_0(x) := f(x), \\ &d_i(x) := f(x) \sqcap f'(x) \sqcap \cdots \sqcap f^{(i)}(x) \qquad (i = 1, \ldots, t-1), \\ &d_t(x) := 1. \end{aligned}$$

Dann ist

$$d_i(x) = f_{i+1}(x)\, f_{i+2}^2(x) \cdots f_t^{t-i}(x) \in K[x] \qquad (i = 0, \ldots, t-1),$$

$$q_i(x) := \frac{d_{i-1}(x)}{d_i(x)} = f_i(x)\, f_{i+1}(x) \cdots f_t(x) \in K[x] \qquad (i = 1, \ldots, t)$$

und also

$$f_i(x) = \frac{q_i(x)}{q_{i+1}(x)} \in K[x] \qquad (i = 1, \ldots, t-1),$$

$$f_t(x) = q_t(x) \in K[x].$$

Besitzt $f(x) \in K[x]$ mehrfache Nullstellen, so kann man also deren Bestimmung durch Rechnungen in $K[x]$ auf die Bestimmung der Nullstellen von Polynomen kleineren Grades zurückführen.

14.3.3. Sei K ein Körper und $f(x) \in K[x]$ ein Polynom des Grades n. Nach Satz 1 besitzt $f(x)$ in jedem Erweiterungskörper von K höchstens n Nullstellen. Ein Polynom n-ten Grades, dessen Nullstellen die (nicht notwendig paarweise verschiedenen) Elemente $\alpha_1, \ldots, \alpha_n$ sind, ist

$$(x - \alpha_1)\,(x - \alpha_2) \cdots (x - \alpha_n).$$

Haben die Polynome

$$f(x) = a_0 + a_1 x + \cdots + a_n x^n \qquad (a_n \neq 0),$$

$$g(x) = b_0 + b_1 x + \cdots + b_n x^n \qquad (b_n \neq 0)$$

aus $K[x]$ dieselben n Nullstellen, so sind diese auch Nullstellen von

$$d(x) = b_n f(x) - a_n g(x).$$

Da aber der Grad von $d(x)$ höchstens $n - 1$ beträgt, muß $d(x)$ das Nullpolynom und daher

$$g(x) = \frac{b_n}{a_n} f(x)$$

sein, d. h., $f(x)$ und $g(x)$ sind in $K[x]$ assoziiert.

Wir verallgemeinern jetzt die Problemstellung und fragen nach der Existenz von Polynomen aus $K[x]$, die an n verschiedenen vorgegebenen Stellen $\alpha_1, \ldots, \alpha_n \in K$ vorgeschriebene Werte $\beta_1, \ldots, \beta_n \in K$ besitzen. Hat ein Polynom $f(x) \in K[x]$ diese Eigenschaft, so auch $f(x) + g(x) \in K[x]$, wobei $g(x)$ ein beliebiges Polynom aus $K[x]$ bezeichnet, für das $g(\alpha_i) = 0$ $(i = 1, \ldots, n)$ ist. Es gibt aber höchstens ein Polynom $f(x) \in K[x]$ mit einem Grad $\leqq n - 1$, das die Bedingungen

$$f(\alpha_i) = \beta_i \qquad (i = 1, \ldots, n)$$

erfüllt. Gilt nämlich auch für das Polynom $\bar{f}(x) \in K[x]$ mit dem Grad $\leqq n-1$

$$\bar{f}(\alpha_i) = \beta_i \qquad (i = 1, \ldots, n),$$

so ist $f(x) - \bar{f}(x) \in K[x]$ als Polynom eines Grades $\leqq n-1$ mit wenigstens n Nullstellen das Nullpolynom und also $f(x) = \bar{f}(x)$.

Da der Grad von

$$h_i := \frac{(x-\alpha_1)\cdots(x-\alpha_{i-1})(x-\alpha_{i+1})\cdots(x-\alpha_n)}{(\alpha_i-\alpha_1)\cdots(\alpha_i-\alpha_{i-1})(\alpha_i-\alpha_{i+1})\cdots(\alpha_i-\alpha_n)} \in K[x] \qquad (i = 1, \ldots, n). \tag{9}$$

gleich $n-1$ ist und

$$h_i(\alpha_i) = 1, \qquad h_i(\alpha_k) = 0 \qquad (i \neq k)$$

gilt, ist der Grad von

$$l(x) := \beta_1 h_1(x) + \beta_2 h_2(x) + \cdots + \beta_n h_n(x) \in K[x] \tag{10}$$

höchstens $n-1$, und es ist

$$l(\alpha_i) = \beta_i \qquad (i = 1, \ldots, n).$$

Damit ist der folgende Satz bewiesen:

Satz 4. *Bezeichnet K einen Körper und $\alpha_1, \ldots, \alpha_n$ paarweise verschiedene, $\beta_1, \ldots, \beta_n$ beliebige Elemente aus K, so gibt es in $K[x]$ genau ein Polynom $l(x)$, dessen Grad $n-1$ nicht übersteigt und für das $l(\alpha_i) = \beta_i$ $(i = 1, \ldots, n)$ gilt.*

Die Aufgabe, ein Polynom zu finden, das an vorgegebenen Stellen bestimmte Werte annimmt, tritt bei der *Interpolation* von Funktionstafeln auf. (10) heißt in Verbindung mit (9) *Lagrangesche Interpolationsformel*. Sie beschreibt im Fall $n = 2$ die *lineare Interpolation*, welche häufig bei der Benutzung von Tafelwerken angewendet wird.

14.4. Irreduzibilitätskriterien

Es bezeichne I einen Integritätsbereich mit Einselement und $f(x)$ ein Polynom positiven Grades aus $I[x]$. Gesucht werden Methoden, mit deren Hilfe in endlich vielen Schritten entschieden werden kann, ob $f(x)$ in $I[x]$ irreduzibel ist.

Wir beschränken uns auf den Fall, daß I der Integritätsbereich $\mathbf{Z}$ der ganzen Zahlen ist und die Koeffizienten von $f(x) \in \mathbf{Z}[x]$ teilerfremd sind. Hat $f(x)$ den Grad $m > 0$ und ist $f(x)$ als Produkt zweier Polynome positiven Grades aus $\mathbf{Z}[x]$ darstellbar, so hat einer der Faktoren, er werde mit $g(x)$ bezeichnet, einen Grad $\leqq \frac{m}{2}$. Sei

r diejenige ganze Zahl, für welche $r \leqq \frac{m}{2} < r+1$ gilt. Dann betrachte man $r+1$ paarweise verschiedene Zahlen $\alpha_0, \alpha_1, \ldots, \alpha_r$ aus $\mathbf{Z}$. Weil $g(x) \mid f(x)$ angenommen wurde, ist

$$g(\alpha_i) \mid f(\alpha_i) \qquad (i = 0, 1, \ldots, r).$$

$f(\alpha_i)$ besitzt nur endlich viele Teiler $\varphi_{i1}, \varphi_{i2}, \ldots, \varphi_{it_i}$ in $\mathbf{Z}$. Daher gibt es zu jedem $i \in \{0, 1, \ldots, r\}$ für $g(\alpha_i)$ nur $t_i \in \mathbf{N}^*$ Möglichkeiten. Insgesamt gibt es also in $\mathbf{Q}[x]$ nur $t_0 t_1 \cdots t_r$ Polynome, die als $g(x)$ in Betracht kommen (vgl. 14.3., Satz 4). Sie können mit Hilfe der Lagrangeschen Interpolationsformel gewonnen werden. In endlich vielen Schritten kann überprüft werden, ob sie Teiler von $f(x)$ sind. Ist keines dieser Polynome ein Teiler von $f(x)$, so ist $f(x)$ irreduzibel in $I[x]$. Wenn aber $f(x)$ durch eines der Polynome teilbar ist, gibt es bereits in $\mathbf{Z}[x]$ eine Zerlegung von $f(x)$ (vgl. 14.2., Satz 3).

Dieses *Verfahren von* KRONECKER (1823—1891) verwendet nur folgende Tatsachen: $\mathbf{Z}$ enthält unendlich viele Elemente (nämlich zu beliebigem $m \in \mathbf{N}^*$ mehr als $\frac{m}{2}$ Elemente), jedes von 0 verschiedene Element von $\mathbf{Z}$ besitzt nur endlich viele Teiler, die in endlich vielen Rechenschritten bestimmbar sind. Daher lassen sich die Überlegungen auf jeden Integritätsbereich I mit diesen Eigenschaften übertragen.

Es ist eine größere Anzahl von *Irreduzibilitätskriterien* (d. h. hinreichenden Bedingungen für die Irreduzibilität von Polynomen entwickelt worden. Wir beweisen als Beispiel den

Satz 1 (Kriterium von EISENSTEIN). *Gibt es eine solche Primzahl $p \in \mathbf{N}^*$, daß für die Koeffizienten des Polynoms*

$$f(x) = a_0 + a_1 x + \cdots + a_m x^m \in \mathbf{Z}[x]$$

die Bedingungen

$$p \nmid a_m \wedge \bigwedge_{i \in \{0,1,\ldots,m-1\}} p \mid a_i \wedge p^2 \nmid a_0 \tag{1}$$

gelten, so ist $f(x)$ in $\mathbf{Q}[x]$ irreduzibel.

Beweis. Wäre $f(x)$ reduzibel, so gäbe es Polynome

$$g(x) = b_0 + b_1 x + \cdots + b_r x^r \in \mathbf{Q}[x] \qquad (b_r \neq 0, \quad r > 0)$$

und

$$h(x) = c_0 + c_1 x + \cdots + c_s x^s \in \mathbf{Q}[x] \qquad (c_s \neq 0, \quad s > 0),$$

für die

$$f(x) = g(x)h(x)$$

gilt. Es kann sogar $g(x) \in \mathbf{Z}[x] \wedge h(x) \in \mathbf{Z}[x]$ angenommen werden (vgl. 14.2., Satz 3). Aus

$$a_0 = b_0 c_0 \qquad \text{und} \qquad p \mid a_0 \wedge p^2 \nmid a_0$$

folgt, daß genau eine der Zahlen b_0, c_0 durch p teilbar ist. Es sei o.B.d.A. $p \mid b_0$. Da $p \nmid a_m$, können nicht alle b_i $(i = 0, 1, \ldots, r)$ durch p teilbar sein. Es gibt also einen Index $k \leqq r < m$ mit der Eigenschaft

$$p \mid b_0 \wedge p \mid b_1 \wedge \cdots \wedge p \mid b_{k-1} \wedge p \nmid b_k. \tag{2}$$

Es ist

$$a_k = b_0c_k + b_1c_{k-1} + \cdots + b_{k-1}c_1 + b_kc_0. \tag{3}$$

Da wir bereits wissen, daß $p \nmid c_0$, folgt aus (2) und (3) $p \nmid a_k$. Weil $k < m$ ist, widerspricht das der Voraussetzung (1).

Beispiele.

1. Ist $a \in \mathbf{Z}$ und gibt es eine Primzahl p, für die $p \mid a \wedge p^2 \nmid a$ gilt, so ist $x^m + a$ ($m \in \mathbf{N}^*$) in $\mathbf{Q}[x]$ irreduzibel.

2. Als *p-tes Kreisteilungspolynom* bezeichnet man

$$F_p(x) := \frac{x^p - 1}{x - 1} = x^{p-1} + x^{p-2} + \cdots + x + 1 \qquad (p \text{ Primzahl}).$$

Es ist in $\mathbf{Q}[x]$ irreduzibel, denn wäre es zerlegbar, so müßte auch

$$F_p(x+1) = \frac{(x+1)^p - 1}{(x+1) - 1} = x^{p-1} + \binom{p}{1}x^{p-2} + \cdots + \binom{p}{p-2}x + \binom{p}{p-1}$$

zerlegbar sein. Aus

$$\bigwedge_{i \in \{1,\ldots,p-1\}} p \,\Big|\, \binom{p}{i} \;\wedge\; p^2 \nmid \binom{p}{p-1}$$

folgt nach Satz 1 aber die Irreduzibilität von $F_p(x+1)$.

14.5. Körpererweiterungen

Wir wollen hier von einem beliebigen Körper K ausgehen und dazu Erweiterungskörper mit gewissen vorgegebenen Eigenschaften konstruieren.

14.5.1. Durch Adjunktion einer Unbestimmten x erhält man aus K den Integritätsbereich $K[x]$ aller Polynome mit Koeffizienten aus K (vgl. 14.1.) Es sei

$$K(x) := \text{Quotientenkörper von } K[x].$$

$K(x)$ besteht aus allen Quotienten

$$\frac{f(x)}{g(x)} = \frac{a_0 + a_1x + a_2x^2 + \cdots + a_mx^m}{b_0 + b_1x + b_2x^2 + \cdots + b_nx^n} \qquad (g(x) \neq 0)$$

von Polynomen aus $K[x]$ mit vom Nullpolynom verschiedenen Nenner und wird *Körper der rationalen Funktionen* in der Unbestimmten x mit Koeffizienten aus K genannt.[1]) Der Körper $K(x)$ ist bis auf Isomorphie durch $K[x]$ und daher sogar schon durch K festgelegt. Nach 13.6. enthält er einen zu K isomorphen Teilkörper. Identifizieren wir diesen mit K, so erscheint $K(x)$ als Erweiterungskörper von K.

[1]) Im Unterschied zu $K(x)$ bilden aber die rationalen Funktionen $r: \alpha \mapsto r(\alpha) := \frac{f(\alpha)}{g(\alpha)}$ ($\alpha \in K$) mit Koeffizienten aus K über dem Definitionsbereich K bezüglich der üblichen Addition und Multiplikation (vgl. 14.1.) keinen Körper. Beispielsweise gibt es zur identischen Funktion $i: \alpha \mapsto \alpha$ ($\alpha \in K$) keine bezüglich der Multiplikation inverse Funktion k: $\alpha \mapsto k(\alpha)$ ($\alpha \in K$), denn für alle $\alpha \in K$ müßte $\alpha k(\alpha) = 1$ sein.

Die hier skizzierte Konstruktion, die zu dem Körper K den Körper $K(x)$ der rationalen Funktionen in x mit Koeffizienten aus K liefert, heißt *Adjunktion* von x zu K (und da das Resultat ein Körper ist, genauer: *Körperadjunktion*). Weil nur ein Element adjungiert wurde, spricht man von einer *einfachen Körpererweiterung*, und da dieses Element bezüglich K transzendent ist (vgl. 14.1.), wird $K(x)$ eine *einfache transzendente Erweiterung* von K genannt.

Definition 1. Seien L und L' Erweiterungskörper des Körpers K. Dann heißt L *äquivalent* L' *bezüglich* $K :\Leftrightarrow$ ein Isomorphismus φ von L auf L' mit $\bigwedge_{a \in K} \varphi(a) = a$ existiert.

Bezeichnen $K(x)$ und $K(\vartheta)$ einfache transzendente Erweiterungen von K, so ist die Abbildung

$$\varphi: f(x) = a_0 + a_1 x + \cdots + a_m x^m \mapsto f(\vartheta) = a_0 + a_1 \vartheta + \cdots + a_m \vartheta^m$$

($f(x) \in K[x]$) ein Isomorphismus von $K[x]$ auf $K[\vartheta]$. Daher ist auch

$$K(x) \cong K(\vartheta),$$

und φ läßt sich zu einem Isomorphismus der Quotientenkörper fortsetzen. Es ergibt sich also:

Je zwei einfache transzendente Erweiterungen von K sind äquivalent. (1)

14.5.2. Es sei jetzt $p(x) \in K[x]$ ein in $K[x]$ irreduzibles Polynom mit einem Grad $n > 1$. $K[x]$ ist Hauptidealring (vgl. 14.2., Satz 1, 13.5., Satz 6) und daher $p(x)$ Primelement von $K[x]$ (vgl. 13.5., Satz 3). Dann ist $(p(x))$ ein Primideal von $K[x]$ und $K[x]/(p(x))$ sogar ein Körper (vgl. 13.5.1.). Da alle Polynome, die sich von $p(x)$ nur um einen konstanten Faktor $\neq 0$ aus K unterscheiden, dasselbe Ideal erzeugen, können wir o. B. d. A. annehmen, daß $p(x)$ den höchsten Koeffizienten 1 hat. Sei

$$p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0. \tag{2}$$

Zu jedem $f(x) \in K[x]$ gibt es eindeutig bestimmte Polynome $q(x)$ und $r(x)$ in $K[x]$, so daß

$$f(x) = q(x)\, p(x) + r(x), \qquad r(x) = 0 \vee \operatorname{Grad} r(x) < n$$

gilt. Daher liegen $f(x)$ und $r(x)$ in derselben Restklasse modulo $(p(x))$. Weil in keiner Restklasse modulo $(p(x))$ zwei verschiedene Polynome, deren Grad $< n$ ist, enthalten sein können, läßt sich also jede Restklasse modulo $(p(x))$ durch genau ein Polynom

$$b_0 + b_1 x + \cdots + b_{n-1}x^{n-1} \in K[x]$$

repräsentieren, dessen Grad höchstens $(n-1)$ beträgt. Beispielsweise liegen

$$x^n = (x^n + a_{n-1}x^{n-1} + \cdots + a_0) + (-a_{n-1}x^{n-1} - \cdots - a_0) \tag{3}$$

und $-a_{n-1}x^{n-1} - \cdots - a_1 x - a_0$ in derselben Restklasse modulo $(p(x))$.

$\overline{f(x)}$ bezeichne die Restklasse von $f(x) \in K[x]$ modulo $(p(x))$. Jedes Element aus $K[x]/(p(x))$ kann auf genau eine Weise in der Form

$$\overline{\bar{b}_0 + \bar{b}_1 x + \cdots + \bar{b}_{n-1}x^{n-1}} = \bar{b}_0 + \bar{b}_1\bar{x} + \cdots + \bar{b}_{n-1}\bar{x}^{n-1}$$

geschrieben werden. Für zwei Elemente des Restklassenkörpers gilt also

$$\bar{b}_0 + \bar{b}_1\bar{x} + \cdots + \bar{b}_{n-1}\bar{x}^{n-1} = \bar{c}_0 + \bar{c}_1\bar{x} + \cdots + \bar{c}_{n-1}\bar{x}^{n-1} \Leftrightarrow \bigwedge_{i\in\{0,\ldots,n-1}} \bar{b}_i = \bar{c}_i. \quad (4)$$

Ferner ist

$$\begin{aligned}&(\bar{b}_0 + \bar{b}_1\bar{x} + \cdots + \bar{b}_{n-1}\bar{x}^{n-1}) + (\bar{c}_0 + \bar{c}_1\bar{x} + \cdots + \bar{c}_{n-1}\bar{x}^{n-1})\\ &= (\bar{b}_0 + \bar{c}_0) + (\bar{b}_1 + \bar{c}_1)\,\bar{x} + \cdots + (\bar{b}_{n-1} + \bar{c}_{n-1})\bar{x}^{n-1}. \end{aligned} \quad (5)$$

Aus (3) folgen die Beziehungen

$$\bar{x}^{n+k} = -\bar{a}_{n-1}\bar{x}^{n-1+k} - \cdots - \bar{a}_1\bar{x}^{1+k} - \bar{a}_0\bar{x}^k \quad (k = 0, 1, \ldots). \quad (6)$$

Daher können in

$$\begin{aligned}\overline{m(x)} :&= (\bar{b}_0 + \bar{b}_1\bar{x} + \cdots + \bar{b}_{n-1}\bar{x}^{n-1})\,(\bar{c}_0 + \bar{c}_1\bar{x} + \cdots + \bar{c}_{n-1}\bar{x}^{n-1})\\ &= \bar{b}_0\bar{c}_0 + (\bar{b}_1\bar{c}_0 + \bar{b}_0\bar{c}_1)\,\bar{x} + \cdots + \bar{b}_{n-1}\bar{c}_{n-1}\bar{x}^{2n-2}\end{aligned} \quad (7)$$

die Potenzen von $\bar{x}$, deren Exponenten $n-1$ übersteigen, schrittweise durch Potenzen von $\bar{x}$ mit kleineren Exponenten ausgedrückt werden, bis die Darstellung des Produktes in der Form

$$\overline{m(x)} = \bar{d}_0 + \bar{d}_1\bar{x} + \cdots + \bar{d}_{n-1}\bar{x}^{n-1} \quad (7')$$

erreicht ist. Zu $\overline{f(x)} \neq \overline{0}$ aus $K[x]/(p(x))$ kann das inverse Element folgendermaßen bestimmt werden:

Da $p(x) \nmid f(x)$, sind $p(x)$ und $f(x)$ teilerfremd. Durch den euklidischen Algorithmus lassen sich in $K[x]$ Polynome $g(x)$ und $h(x)$ berechnen, für die

$$f(x)\,g(x) + p(x)\,h(x) = 1 \quad (8)$$

gilt. Daher ist $\overline{g(x)}$ in $K[x]/(p(x))$ zu $\overline{f(x)}$ invers.

Da $\overline{p(x)} = \overline{0}$ ist, bedeutet (2)

$$\bar{x}^n + \bar{a}_{n-1}\bar{x}^{n-1} + \cdots + \bar{a}_1\bar{x} + \bar{a}_0 = \overline{0}. \quad (9)$$

Der eben konstruierte Körper $K[x]/(p(x))$ enthält einen zu K isomorphen Teilkörper $\bar{K}$, bestehend aus allen Restklassen $\bar{b}$ $(b \in K)$. Wir bilden die Menge der Elemente

$$\left(K[x]/(p(x))\right) \setminus \bar{K}) \cup K,$$

d. h., wir ersetzen in $K[x]/(p(x))$ alle Restklassen $\bar{b}$ durch die Vertreter b $(b \in K)$. Ersetzen wir dann auch in allen Gleichungen, die die Operationen in $K[x]/(p(x))$ beschreiben, die auftretenden Elemente $\bar{b}$ aus $\bar{K}$ durch ihre Vertreter b aus K,

so werden durch diese neuen Gleichungen Operationen in $\big(K[x]/\big(p(x)\big) \setminus \overline{K}\big) \cup K$ erklärt, bezüglich derer diese Menge offensichtlich ein zu $K[x]/\big(p(x)\big)$ isomorpher Erweiterungskörper von K ist. Schließlich wollen wir noch das Element $\overline{x}$ dieses Körpers mit ϑ und den Körper selbst mit $K(\vartheta)$ bezeichnen. Jedes Element aus $K(\vartheta)$ kann dann auf genau eine Weise in der Form

$$b_0 + b_1\vartheta + b_2\vartheta^2 + \cdots + b_{n-1}\vartheta^{n-1} \qquad (b_i \in K;\ \ i = 0, 1, \ldots, n-1)$$

angegeben werden. Die Rechnung in $K(\vartheta)$ wird durch (4), (5), (6), (7), (7') beschrieben, wenn darin ϑ an die Stelle von $\overline{x}$ gesetzt und die weiteren Querstriche weggelassen werden. Die Gleichung (9) bedeutet in $K(\vartheta)$

$$\vartheta^n + a_{n-1}\vartheta^{n-1} + \cdots + a_1\vartheta + a_0 = 0. \tag{10}$$

ϑ ist also eine Nullstelle von $p(x)$.

Definition 2. Sei L ein Erweiterungskörper des Körpers K und $\lambda \in L$. Dann heißt

λ algebraisch bezüglich K $:\Leftrightarrow$ λ ist Nullstelle eines Polynoms $f(x) \neq 0$ aus $K[x]$.

Ist ein Element $\lambda \in L$ nicht algebraisch bezüglich K, so folgt aus jeder Gleichung

$$a_m\lambda^m + a_{m-1}\lambda^{m-1} + \cdots + a_1\lambda + a_0 = 0 \left(\bigwedge_{i\in\{0,\ldots,m\}} a_i \in K\right)$$

die Beziehung $a_m = a_{m-1} = \cdots = a_0 = 0$, d. h., λ ist transzendent bezüglich K (vgl. 14.1.).

Beispiel. $\sqrt{2} \in \mathbb{R}$ ist algebraisch bezüglich $\mathbb{Q}$, da $\sqrt{2}$ Nullstelle von $x^2 - 2 \in \mathbb{Q}[x]$ ist. e und π sind transzendent bezüglich $\mathbb{Q}$. Die Transzendenz von e wurde 1873 durch Charles Hermite (1822—1901), diejenige von π 1882 durch Ferdinand von Lindemann (1852—1939) bewiesen. Wir gehen hier nicht auf die Beweise ein.

Die Gleichung (10) besagt also, daß ϑ algebraisch bezüglich K ist. Daher heißt *$K(\vartheta)$ algebraische Erweiterung* von K.

Die skizzierte Konstruktion, die zu dem Körper K und dem irreduziblen Polynom $p(x) \in K[x]$ den Körper $K(\vartheta)$ liefert, heißt *Adjunktion einer Nullstelle von $p(x)$ zu K*. Da nur ein Element adjungiert wurde, nennt man $K(\vartheta)$ eine *einfache algebraische Erweiterung* von K mit dem *definierenden Polynom* $p(x)$.

Zwei einfache algebraische Erweiterungen von K mit dem definierenden Polynom $p(x) \in K[x]$ sind äquivalent. (11)

Sind nämlich ϑ und ϑ' Nullstellen von $p(x)$, so ist nach unserer Konstruktion

$$K(\vartheta) \cong K[x]/\big(p(x)\big) \qquad \text{und} \qquad K(\vartheta') \cong K[x]/\big(p(x)\big).$$

Bei diesen Isomorphismen entsprechen sich die Elemente von K und die durch sie repräsentierten Restklassen modulo $\big(p(x)\big)$. Ferner entsprechen ϑ bzw. ϑ' der Restklasse $\overline{x}$ von x. Daraus ergibt sich die Behauptung. Insgesamt liefern unsere Resultate den folgenden Satz.

Satz 1. *Bezeichne K einen Körper. Zu K gibt es bis auf Äquivalenz genau eine einfache transzendente Erweiterung. Ist $p(x)$ ein in $K[x]$ irreduzibles Polynom, so gibt es bis auf Äquivalenz genau eine einfache algebraische Erweiterung $K(\vartheta)$ mit dem definierenden Polynom $p(x)$, d. h., in $K(\vartheta)$ gilt die Gleichung $p(\vartheta) = 0$.*

Beispiel. $x^3 - 2$ ist in $\mathbf{Q}[x]$ irreduzibel (vgl. 14.4., Satz 1). Die Elemente der Erweiterung $\mathbf{Q}(\vartheta)$ von $\mathbf{Q}$ mit der definierenden Gleichung $\vartheta^3 - 2 = 0$ können in der Form

$$b_0 + b_1\vartheta + b_2\vartheta^2 \qquad (b_0, b_1, b_2 \in \mathbf{Q})$$

geschrieben werden. Mit ihnen wird wie mit Polynomen gerechnet, wobei aber die Gleichung $\vartheta^3 = 2$ zu beachten ist. Der Teilkörper $\mathbf{Q}\left(\sqrt[3]{2}\right)$ von $\mathbf{R}$, der aus sämtlichen reellen Zahlen der Form

$$b_0 + b_1\sqrt[3]{2} + b_2\left(\sqrt[3]{2}\right)^2 \qquad (b_0, b_1, b_2 \in \mathbf{Q})$$

besteht, sowie der Teilkörper $\mathbf{Q}\left(\sqrt{2}\left(-\frac{1}{2} + \frac{i}{2}\sqrt{3}\right)\right)$ von $\mathbf{C}$, der aus allen komplexen Zahlen

$$b_0 + b_1\left(\sqrt[3]{2}\,\xi\right) + b_2\left(\sqrt[3]{2}\,\xi\right)^2 \qquad \left(b_0, b_1, b_2 \in \mathbf{Q};\ \xi = -\frac{1}{2} + \frac{i}{2}\sqrt{3}\right)$$

gebildet wird, sind bezüglich $\mathbf{Q}$ äquivalent. Beide Körper stellen Realisierungen der durch $p(x) = x^3 - 2$ definierten einfachen algebraischen Erweiterung von $\mathbf{Q}$ dar.

14.5.3. Bisher sind wir von einem festen Körper K ausgegangen und haben dazu einfache transzendente und einfache algebraische Erweiterungen konstruiert. Die Elemente der entstandenen Erweiterungskörper sind Quotienten von Polynomen in einer Unbestimmten bzw. die Restklassen des Polynomringes $K[x]$ nach einem Primideal. Nach Satz 1 sind dies bis auf Isomorphie auch die einzigen Erweiterungen der betrachteten Art. Das Beispiel zeigt jedoch, daß es zu einer so gewonnenen Erweiterung eines gegebenen Körpers in einem bereits bekannten, diesen umfassenden Körper äquivalente Erweiterungen geben kann, deren Elemente nicht Polynomquotienten oder Restklassen zu sein brauchen.

Daher liegt es nahe, einen Körper L sowie einen darin enthaltenen Teilkörper K zu betrachten und nach Erweiterungskörpern von K zu fragen, die in L liegen. Jeden Teilkörper Z von L, der K umfaßt, nennen wir einen *Zwischenkörper* von K und L. Bezeichnet M eine beliebige Menge von Elementen aus L und

$$\mathfrak{Z} := \{Z : Z \supseteq M \wedge Z \text{ ist Zwischenkörper von } K \text{ und } L\},$$

so ist

$$K(M) := \bigcap_{Z \in \mathfrak{Z}} Z$$

ein M umfassender Zwischenkörper von K und L. Offenbar liegen in jedem Z und daher auch in $K(M)$ alle diejenigen Elemente von L, die sich durch endlich viele Rechenschritte (d. h. Anwendungen der Operationen von L sowie Bildung der inversen Elemente bezüglich dieser Operationen) aus den Elementen von $K \cup M$ darstellen lassen. Diese Elemente bilden aber bereits einen M umfassenden Zwischenkörper von K und L, der dann $K(M)$ sein muß. Wir wollen sagen, daß $K(M)$ aus K durch *Adjunktion* (genauer: *Körperadjunktion*) von M entsteht.

Da ein Element aus $K(M)$ zu seiner Darstellung im eben genannten Sinn nur endlich viele Elemente aus M benötigt, gibt es zu jedem Rechenschritt, der in $K(M)$ ausgeführt wird, eine

solche endliche Teilmenge M' von M, daß dieser Rechenschritt bereits in dem Teilkörper $K(M')$ von $K(M)$ abläuft. Beherrscht man also die Rechnung in allen Körpern $K(M')$, die aus K durch Adjunktion endlicher Teilmengen M' von M aus K entstehen, so beherrscht man sie auch in $K(M)$.

Bezeichnen M', M'' Teilmengen von M, so ist $K(M') \subseteq K(M' \cup M'')$ und $M'' \subseteq K(M' \cup M'')$, also

$$K(M')(M'') \subseteq K(M' \cup M'').$$

Andererseits ist $K \subseteq K(M')$ (M'') und $M' \cup M'' \subseteq K(M')$ (M''), also

$$K(M' \cup M'') \subseteq K(M')\,(M'').$$

Aus beiden Beziehungen folgt

$$K(M' \cup M'') = K(M')\,(M''),$$

d. h., die Adjunktion endlicher Mengen kann auf die schrittweise Adjunktion einelementiger Mengen zurückgeführt werden. Ist $M = \{\alpha_1, \ldots, \alpha_n\}$ eine endliche Menge, so sei $K(\alpha_1, \ldots, \alpha_n) := K(\{\alpha_1, \ldots, \alpha_n\})$. Für ein beliebiges Element $\vartheta \in L$ betrachten wir die Erweiterung $K(\vartheta)$. Die Menge aller Elemente der Form

$$f(\vartheta) = a_0 + a_1\vartheta + \cdots + a_k\vartheta^k \quad (\{a_0, a_1, \ldots, a_k\} \subseteq K;\ k \in \mathbb{N})$$

bildet einen Teilintegritätsbereich $K[\vartheta]$ von $K(\vartheta)$.

Wir versuchen nun, einen Zusammenhang zu den vorn konstruierten Erweiterungen herzustellen, und betrachten daher parallel den Integritätsbereich $K[x]$ aller Polynome in der Unbestimmten x mit Koeffizienten aus K. Durch

$$f(x) = a_0 + a_1x + \cdots + a_kx^k \mapsto f(\vartheta) = a_0 + a_1\vartheta + \cdots + a_k\vartheta^k$$

wird ein Homomorphismus von $K[x]$ auf $K[\vartheta]$ beschrieben, dessen Kern $\mathfrak{p}$ aus allen Polynomen $f(x)$ besteht, deren Bild $f(\vartheta) = 0$ ist. Aus dem Homomorphiesatz für Ringe ergibt sich

$$K[x]/\mathfrak{p} \cong K[\vartheta].$$

$\mathfrak{p}$ ist Primideal von $K[x]$, denn $K[\vartheta]$ ist Integritätsbereich. Weil $K[x]$ Hauptidealring ist, muß $\mathfrak{p} = K[x]$ oder $\mathfrak{p} = (0)$ oder $\mathfrak{p} = (p(x))$ sein, wobei $p(x)$ ein in $K[x]$ irreduzibles Polynom positiven Grades bezeichnet. Da $K[\vartheta]$ als K umfassender Integritätsbereich nicht der Nullring sein kann, ist $\mathfrak{p} \neq K[x]$. Es bleiben also die Möglichkeiten $\mathfrak{p} = (0)$ oder $\mathfrak{p} = (p(x))$.

Fall 1. $\mathfrak{p} = (0)$. In diesem Fall ist die gegebene Abbildung ein Isomorphismus von $K[x]$ auf $K[\vartheta]$. $K(x)$ bzw. $K(\vartheta)$ sind gerade die Quotientenkörper von $K[x]$ bzw. $K[\vartheta]$. Nach 13.6. ist dann

$$K(\vartheta) \cong K(x).$$

Es liegt also eine *einfache transzendente Erweiterung* von K vor. Mit ϑ wird ebenso gerechnet wie mit der Unbestimmten x. ϑ ist ein bezüglich K transzendentes Element von L.

Fall 2. $\mathfrak{p} = (p(x))$. Schon vorn wurde gezeigt, daß in diesem Fall $K[x]/\mathfrak{p}$ ein Körper ist und daher

$$K(\vartheta) = K[\vartheta] \cong K[x]/(p(x)).$$

Aus $p(x) \mapsto p(\vartheta) = 0$ folgt, daß ϑ Nullstelle des Polynoms $p(x)$ ist. ϑ ist also ein bezüglich K algebraisches Element von L und $K(\vartheta)$ eine *einfache algebraische Erweiterung* von K mit dem definierenden Polynom $p(x)$. Da $p(x)$ durch $\mathfrak{p}$ nur bis auf Faktoren $\neq 0$ aus K bestimmt ist, kann o.B.d.A. angenommen werden, daß $p(x)$ den höchsten Koeffizienten 1 hat. Hat $p(x)$ den Grad 1, so ist bei dieser Normierung $p(x) = x - \vartheta$ und wegen $p(x) \in K[x]$ also $\vartheta \in K$. Dann ist aber $K(\vartheta) = K$, d. h., es liegt keine echte Erweiterung vor.

Durch die angegebenen Konstruktionen wurden also schon alle Möglichkeiten für *einfache Erweiterungen*, das sind Körpererweiterungen, die durch Adjunktion einer einelementigen Menge entstehen, erfaßt.

14.5.4. Sei K ein Körper und $f(x) \in K[x]$ ein Polynom des Grades $n > 0$.

$$f(x) = f_1(x)\, f_2(x) \cdots f_r(x)$$

bezeichne die Zerlegung von $f(x)$ in irreduzible Faktoren aus $K[x]$ (vgl. 14.2.1.). Sind diese Faktoren sämtlich linear, so liegen alle Nullstellen von $f(x)$ in K. Anderenfalls sei Grad $f_1(x) > 1$. Dann gibt es einen Erweiterungskörper $K(\alpha_1)$ von K, der wenigstens eine Nullstelle α_1 von $f_1(x)$ enthält. Die Zerlegung von $f(x)$ in irreduzible Faktoren aus $K(\alpha_1)\,[x]$ sei

$$f(x) = (x - \alpha_1)\, g_1(x) \cdots g_s(x)$$

(vgl. 14.3.1). Falls noch nicht alle Faktoren den Grad 1 haben, adjungiere man zu $K(\alpha_1)$ eine Nullstelle α_2 eines nichtlinearen Faktors. In $K(\alpha_1)\,(\alpha_2)\,[x]$ zerfällt dann $f(x)$ in das Produkt irreduzibler Faktoren

$$f(x) = (x - \alpha_1)\,(x - \alpha_2)\, h_1(x) \cdots h_t(x).$$

Hat hierin noch ein Faktor einen Grad > 1, so kann man das geschilderte Verfahren fortsetzen und erhält spätestens nach $n - 1$ Schritten einen Erweiterungskörper Z von K mit der Eigenschaft, daß $f(x)$ in $Z[x]$ in ein Produkt von Linearfaktoren zerfällt:

$$f(x) = a(x - \alpha_1)\,(x - \alpha_2) \cdots (x - \alpha_n) \qquad (a \in K).$$

Definition 3. Seien $Z \supseteq K$ Körper und $f(x) \in K[x]$ ein Polynom positiven Grades. Z heißt genau dann *Zerfällungskörper* (im weiteren Sinn) von $f(x)$, wenn $f(x)$ in $Z[x]$ in Linearfaktoren zerfällt. Genau dann heißt Z *kleinster Zerfällungskörper* von $f(x)$, wenn kein echter Teilkörper von Z Zerfällungskörper von $f(x)$ ist.

Satz 2. *Ist K ein Körper und $f(x) \in K[x]$ ein Polynom des Grades $n > 0$, so gibt es einen kleinsten Zerfällungskörper Z von $f(x)$.*

Beweis. Nach den obigen Überlegungen gibt es jedenfalls einen Zerfällungskörper Z von $f(x)$. Der Durchschnitt aller in Z enthaltenen Zerfällungskörper ist dann ein kleinster Zerfällungskörper von $f(x)$.

Je zwei kleinste Zerfällungskörper von $f(x) \in K[x]$ sind bezüglich K äquivalent. Dieser Eindeutigkeitssatz, dessen Beweis (vgl. etwa Kochendörffer: Einführung in die Algebra, 4. Aufl., 1974, S. 166) wir hier übergehen wollen, besagt, daß in jedem Zerfällungskörper von $f(x)$ bei der Zerlegung von $f(x)$ in ein Produkt von Linearfaktoren die gleiche Anzahl verschiedener Linearfaktoren und die gleichen Vielfachheiten dieser Faktoren auftreten, die Zerlegung von $f(x)$ in Linearfaktoren also im wesentlichen in allen Zerfällungskörpern dieselbe ist.

14.6. Polynome in mehreren Unbestimmten

14.6.1. Es sei I ein Integritätsbereich mit Einselement. Durch Adjunktion der Unbestimmten x_1 entsteht daraus der Polynomring $I[x_1]$. Hierzu kann man erneut das in 14.1. geschilderte Verfahren anwenden (vgl. 14.1., Satz 1) und erhält nach Adjunktion der Unbestimmten x_2 den Polynomring $I[x_1][x_2]$, den wir kürzer mit $I[x_1, x_2]$ bezeichnen wollen. Adjungiert man weiter der Reihe nach die Unbestimmten x_3, $x_4, \ldots, x_n$, so erhält man den Polynomring

$$I[x_1, \ldots, x_n] := I[x_1, \ldots, x_{n-1}][x_n] \qquad (n - 1 \in \mathbb{N}^*).$$

Seine Elemente heißen *Polynome in den Unbestimmten* $x_1, x_2, \ldots, x_n$ und haben die Form

$$f(x_1, x_2, \ldots, x_n) = \sum_{i_1, i_2, \ldots, i_n} a_{i_1 i_2 \ldots i_n} x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}, \tag{1}$$

wobei die Koeffizienten $a_{i_1 i_2 \ldots i_n} \in I$ sind und über endlich viele n-Tupel von Exponenten $i_1, i_2, \ldots, i_n$ aus $\mathbb{N}$ summiert wird.

Es sei L ein I umfassender Integritätsbereich. Setzt man in der Gleichung (1) $\alpha_i \in L$ an die Stelle von x_i ($i = 1, 2, \ldots, n$), so wird durch $(\alpha_1, \alpha_2, \ldots, \alpha_n) \mapsto f(\alpha_1, \alpha_2, \ldots, \alpha_n)$ eine Abbildung von $L \times L \times \cdots \times L$ in L beschrieben, die man *ganze rationale Funktion von n Argumenten* über dem Definitionsbereich $L \times L \times \cdots \times L$ nennt. Wie in 14.1. kann gezeigt werden, daß diese Funktionen einen kommutativen Ring bilden, der im Fall eines unendlichen Integritätsbereiches I (z. B. Zahlenbereich $\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{C}$) zum Polynomring $I[x_1, x_2, \ldots, x_n]$ isomorph ist.

Aus 14.1. ergibt sich, daß $I[x_1, \ldots, x_n]$ ein Integritätsbereich ist. Sein Quotientenkörper, dessen Elemente also die Gestalt

$$\frac{f(x_1, \ldots, x_n)}{g(x_1, \ldots, x_n)}$$

haben, wobei $f(x_1, \ldots, x_n)$, $g(x_1, \ldots, x_n) \in I[x_1, \ldots, x_n]$ und $g(x_1, \ldots, x_n) \neq 0$ ist, wird durch $I(x_1, \ldots, x_n)$ bezeichnet und wegen des Zusammenhanges der Polynome mit den ganzen rationalen Funktionen meistens *Körper der rationalen Funktionen* (von n Unbestimmten) über dem Koeffizientenbereich I genannt.

Die Exponentensumme $i_1 + \cdots + i_n$ heißt *Grad des Gliedes* $a_{i_1 i_2 \ldots i_n} x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}$. Unter dem *Grad des Polynoms* $f(x_1, x_2, \ldots, x_n)$ versteht man den größten Grad, welcher bei den Gliedern von $f(x_1, x_2, \ldots, x_n)$ auftritt. Wie in 14.1. wird dem Nullpolynom kein Grad zugeschrieben. Eine *Form* oder ein *homogenes Polynom* ist ein Polynom, dessen sämtliche Glieder den gleichen Grad haben.

14.6.2. Definition 1. $s(x_1, \ldots, x_n) \in I[x_1, \ldots, x_n]$ heißt *symmetrisches Polynom* $:\Leftrightarrow s(x_1, \ldots, x_n)$ bleibt bei jeder Permutation der Unbestimmten $x_1, \ldots, x_n$ ungeändert.

Beispiele sind:
die *Potenzsummen*

$$\pi_k(x_1, \ldots, x_n) := x_1^k + x_2^k + \cdots + x_n^k \qquad (k = 0, 1, 2, \ldots),$$

die *Diskriminante*

$$\delta(x_1, \ldots, x_n) := \prod_{i<k} (x_i - x_k)^2 \qquad (i, k = 1, \ldots, n)$$

sowie die sogenannten *symmetrischen Grundfunktionen*

$$\sigma_1(x_1, \ldots, x_n) := x_1 + x_2 + \cdots + x_n,$$
$$\sigma_2(x_1, \ldots, x_n) := x_1x_2 + x_1x_3 + \cdots + x_{n-1}x_n = \sum_{0<i_1<i_2\leq n} x_{i_1}x_{i_2},$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\sigma_n(x_1, \ldots, x_n) := x_1x_2 \cdots x_n,$$

die allgemein durch

$$\sigma_k(x_1, \ldots, x_n) := \sum_{0<i_1<i_2<\cdots<i_k\leq n} x_{i_1}x_{i_2} \cdots x_{i_k} \qquad (k = 1, 2, \ldots, n) \tag{2}$$

gegeben sind, wobei die Summe über alle $\binom{n}{k}$ möglichen Indexsysteme zu erstrecken ist.

Sei I ein Integritätsbereich mit Einselement, Q der Quotientenkörper von I und $K \supseteq Q$ Zerfällungskörper des Polynoms

$$f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 \in I[x]. \tag{3}$$

Dann gibt es also Elemente $\alpha_1, \ldots, \alpha_n$ in K, so daß

$$f(x) = (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n) \tag{4}$$

gilt. Multipliziert man dieses Produkt aus, ordnet nach Potenzen von x und vergleicht die Koeffizienten gleicher Potenzen von x in (3) und (4), so ergeben sich folgende Beziehungen zwischen den Koeffizienten und den Nullstellen von $f(x)$, die bequem durch die symmetrischen Grundfunktionen ausgedrückt werden können:

$$\begin{aligned}
&a_{n-1} = (-1)\,\sigma_1(\alpha_1, \ldots, \alpha_n) = (-1)(\alpha_1 + \alpha_2 + \cdots + \alpha_n),\\
&a_{n-2} = (-1)^2\,\sigma_2(\alpha_1, \ldots, \alpha_n) = (-1)^2(\alpha_1\alpha_2 + \alpha_1\alpha_3 + \cdots + \alpha_{n-1}\alpha_n),\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&a_{n-k} = (-1)^k\,\sigma_k(\alpha_1, \ldots, \alpha_n) \qquad (k = 1, \ldots, n),\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&a_0 = (-1)^n\,\sigma_n(\alpha_1, \ldots, \alpha_n) = (-1)^n\,\alpha_1\alpha_2 \cdots \alpha_n.
\end{aligned} \tag{5}$$

Die Polynome $\sigma_k(x_1, \ldots, x_n)$ sind grundlegend für alle symmetrischen Polynome, denn es gilt

Satz 1 (Hauptsatz über symmetrische Polynome). *Zu jedem symmetrischen Polynom*

$$s(x_1, \ldots, x_n) \in I[x_1, \ldots, x_n]$$

gibt es genau ein Polynom

$$f(x_1, \ldots, x_n) \in I[x_1, \ldots, x_n],$$

so daß

$$s(x_1, \ldots, x_n) = f\big(\sigma_1(x_1, \ldots, x_n), \ldots, \sigma_n(x_1, \ldots, x_n)\big) \tag{6}$$

ist.

Beweis. Sei $g \in \mathbf{N} \wedge g > 1$. Wir betrachten symmetrische Polynome

$$s(x_1, \ldots, x_n) = \sum_{i_1, \ldots, i_n} a_{i_1 \ldots i_n} x_1^{i_1} \cdots x_n^{i_n},$$

deren Grad $< g$ ist. Daher sind nur solche Koeffizienten $a_{i_1 \ldots i_n} \neq 0$, für deren Indizes $0 \leqq i_k < g$ $(k = 1, \ldots, n)$ gilt. Die Menge dieser n-Tupel von Indizes wird durch

$$(i_1, \ldots, i_n) \mapsto \sum_{k=1}^{n} i_k g^{n-k}$$

offenbar eineindeutig in die Menge aller nichtnegativen ganzen Zahlen $< g^n$ abgebildet (Darstellung ganzer Zahlen in einem Zahlensystem mit der Grundzahl g). Das Bild eines n-Tupels bezeichnen wir als seine Nummer. Die größte Nummer eines n-Tupels mit $a_{i_1 \ldots i_n} \neq 0$ aus $s(x_1, \ldots, x_n)$ sei h.

Durch vollständige Induktion nach h wird nun die Existenz von Polynomen $f(x_1, \ldots, x_n)$ mit der Eigenschaft (6) gezeigt.

Ist 0 die größte auftretende Nummer, so folgt $i_1 = \cdots = i_n = 0$ und daher $s(x_1, \ldots, x_n) = a_{0 \ldots 0}$. In diesem Fall ist also

$$f(x_1, \ldots, x_n) := a_{0 \ldots 0}$$

ein Polynom aus $I[x_1, \ldots, x_n]$, für das (6) gilt.

Wir machen die Induktionsannahme, daß sämtliche symmetrischen Polynome, in denen nur Koeffizienten mit Indexnummern $< h$ von 0 verschieden sind, gemäß (6) durch die symmetrischen Grundfunktionen ausgedrückt werden können. Es sei dann $s(x_1, \ldots, x_n)$ ein symmetrisches Polynom, das nur von 0 verschiedene Koeffizienten mit Indexnummern $\leqq h$ enthält. Ist $(i_1, \ldots, i_n)$ das n-Tupel mit der Nummer h, so muß $i_k \geqq i_{k+1}$ $(k = 1, \ldots, n-1)$ sein, denn wäre $i_k < i_{k+1}$, so hätte das durch Vertauschung von i_k mit i_{k+1} aus $(i_1, \ldots, i_n)$ entstehende n-Tupel eine Nummer $> h$, und der zugehörige Koeffizient müßte 0 sein, während er wegen der Symmetrie von $s(x_1, \ldots, x_n)$ gleich $a_{i_1 \ldots i_n} \neq 0$ ist.

Offenbar ist

$$s^*(x_1, \ldots, x_n) := s(x_1, \ldots, x_n) - a_{i_1 \ldots i_n} \sigma_1^{i_1 - i_2} \sigma_2^{i_2 - i_3} \cdots \sigma_{n-1}^{i_{n-1} - i_n} \sigma_n^{i_n} \tag{7}$$

ein symmetrisches Polynom, wobei σ_i $(i = 1, \ldots, n)$ jeweils für $\sigma_i(x_1, \ldots, x_n)$ steht. Der Grad von σ_i ist i. Weil

$$1(i_1 - i_2) + 2(i_2 - i_3) + \cdots + (n-1)(i_{n-1} - i_n) + n i_n = \sum_{k=1}^{n} i_k < g$$

und Grad $s(x_1, \ldots, x_n) < g$ ist, hat auch $s^*(x_1, \ldots, x_n)$ einen Grad $< g$. In einem Produkt von σ_k-Faktoren entsteht das Glied $\neq 0$, welches die größte Indexnummer besitzt, durch Multiplikation derjenigen Glieder aus den Summendarstellungen (2) von σ_k $(k = 1, \ldots, n)$ mit den niedrigsten Indizes $i_1, \ldots, i_k$, also der Glieder $x_1 \cdots x_k$ $(k = 1, \ldots, n)$ miteinander. Für das in (7) subtrahierte Produkt ist dies das Glied

$$a_{i_1 \ldots i_n} x_1^{i_1 - i_2} (x_1 x_2)^{i_2 - i_3} \cdots (x_1 \cdots x_{n-1})^{i_{n-1} - i_n} (x_1 \cdots x_n)^{i_n} = a_{i_1 \ldots i_n} x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}.$$

Daher hat jeder Koeffizient $\neq 0$ in $s^*(x_1, \ldots, x_n)$ eine Indexnummer $< h$. Nach der Induktionsannahme gibt es ein Polynom $f^*(x_1, \ldots, x_n) \in I[x_1, \ldots, x_n]$ für das

$$s^*(x_1, \ldots, x_n) = f^*(\sigma_1, \ldots, \sigma_n)$$

gilt. Nach (7) ist dann

$$s(x_1, \ldots, x_n) = f^*(\sigma_1, \ldots, \sigma_n) + a_{i_1 \ldots i_n} \sigma_1^{i_1 - i_2} \sigma_2^{i_2 - i_3} \cdots \sigma_{n-1}^{i_{n-1} - i_n} \sigma_n^{i_n}$$

eine Darstellung von $s(x_1, \ldots, x_n)$ durch ein Polynom in den symmetrischen Grundfunktionen mit Koeffizienten aus I.

Wir beweisen nun noch die Eindeutigkeit der Darstellung. Wären

$$s(x_1, \ldots, x_n) = f(\sigma_1, \ldots, \sigma_n) = g(\sigma_1, \ldots, \sigma_n)$$

zwei Darstellungen von $s(x_1, \ldots, x_n)$ durch Polynome in den symmetrischen Grundfunktionen, so wäre

$$f(\sigma_1, \ldots, \sigma_n) - g(\sigma_1, \ldots, \sigma_n) = d(\sigma_1, \ldots, \sigma_n) = D(x_1, \ldots, x_n)$$

das Nullpolynom in den $x_1, \ldots, x_n$. In dem Potenzprodukt

$$\sigma_1^{i_1 - i_2} \sigma_2^{i_2 - i_3} \cdots \sigma_{n-1}^{i_{n-1} - i_n} \sigma_n^{i_n} \qquad (i_1 \geqq i_2 \geqq \cdots \geqq i_n \geqq 0)$$

ist, wie wir oben gesehen haben, $x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}$ das von 0 verschiedene Glied mit der größten Indexnummer. Verschiedene Potenzprodukte haben also auch verschiedene Glieder mit größter Indexnummer. Daraus folgt, daß sämtliche Koeffizienten von $d(\sigma_1, \ldots, \sigma_n)$ gleich 0 sein müssen und also

$$f(\sigma_1, \ldots, \sigma_n) = g(\sigma_1, \ldots, \sigma_n)$$

gilt. Anderenfalls bestände nämlich $d(\sigma_1, \ldots, \sigma_n)$ aus einer Summe von Potenzprodukten der $\sigma_1, \ldots, \sigma_n$ mit Koeffizienten $\neq 0$. Von diesen Potenzprodukten lieferte eines einen Summanden $c_{i_1 \ldots i_n} x_1^{i_1} \cdots x_n^{i_n}$ mit einem Koeffizienten $c_{i_1 \ldots i_n} \neq 0$

und maximaler Indexnummer. Er könnte nach der vorstehenden Bemerkung nicht gegen einen anderen Summanden fortfallen und $D(x_1, \ldots, x_n)$ also nicht das Nullpolynom sein.

Der Beweis liefert auch sofort ein Verfahren, um die Darstellung eines symmetrischen Polynoms als Polynom in den symmetrischen Grundfunktionen mit Koeffizienten aus I tatsächlich zu finden. Wir zeigen das am Beispiel

$$s(x_1, \ldots, x_n) = x_1^3 + x_2^3 + x_3^3.$$

Der Grad dieses Polynoms ist 3. Daher genügt es, $g = 10$ zu wählen. Die Indexnummern der Summanden sind dann

$$\begin{aligned} x_1^3&: \quad 3 \cdot 10^2 = 300, \\ x_2^3&: \quad 3 \cdot 10^1 = 30, \\ x_3^3&: \quad 3 \cdot 10^0 = 3. \end{aligned}$$

Nach (7) ist

$$\begin{aligned} s^* &= s - 1 \cdot \sigma_1^{3-0}\sigma_2^{0-0}\sigma_3^0 = x_1^3 + x_2^3 + x_3^3 - (x_1 + x_2 + x_3)^3 \\ &= -3x_1^2x_2 - 3x_1^2x_3 - 3x_2^2x_3 - 3x_1x_2^2 - 3x_1x_3^2 - 3x_2x_3^2 - 6x_1x_2x_3. \end{aligned}$$

Unter diesen Summanden hat $-3x_1^2x_2$ die größte Indexnummer. Sie beträgt

$$2 \cdot 10^{3-1} + 1 \cdot 10^{3-2} = 210.$$

Wenden wir das Verfahren auf s^* erneut an, so erhalten wir nach (7)

$$\begin{aligned} s^{**} &= s^* - (-3)\,\sigma_1^{2-1}\sigma_2^{1-0}\sigma_3^0 = s^* + 3\sigma_1\sigma_2 \\ &= s^* + 3(x_1 + x_2 + x_3)\,(x_1x_2 + x_1x_3 + x_2x_3) \\ &= 3x_1x_2x_3 \\ &= 3\sigma_3. \end{aligned}$$

Daher ist

$$\begin{aligned} s &= s^* + \sigma_1^3 = s^{**} - 3\sigma_1\sigma_2 + \sigma_1^3 \\ &= \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3. \end{aligned}$$

14.7. Fundamentalsatz der Algebra

Nach 14.1. ist $\mathbf{C}[x]$ isomorph zum Ring der ganzen rationalen Funktionen mit Koeffizienten aus $\mathbf{C}$. Zur Untersuchung der Elemente dieses Ringes stehen Hilfsmittel der Analysis zur Verfügung, durch die spezielle Eigenschaften des Körpers $\mathbf{C}$ ausgenutzt werden können.

Wir werden zuerst einige Aussagen über solche ganzen rationalen Funktionen und über Nullstellen von Polynomen aus $\mathbf{C}[x]$ beweisen.

Hilfssatz 1. *Sei*

$$g(x) = a_1 x + \cdots + a_m x^m \in \mathbf{C}[x], \qquad \mu := \max\{|a_1|, \ldots, |a_m|\}$$

und ε eine beliebige positive reelle Zahl. Dann gilt für $x \in \mathbf{C}$:

$$|x| < \frac{\varepsilon}{\varepsilon + \mu} \Rightarrow |g(x)| < \varepsilon.$$

Beweis. Es ist

$$\begin{aligned} |g(x)| &\leqq |a_1|\,|x| + \cdots + |a_m|\,|x|^m \\ &\leqq \mu\,|x|\,(1 + |x| + \cdots + |x|^{m-1}) = \mu|x|\,\frac{1 - |x|^m}{1 - |x|}. \end{aligned}$$

Im Fall $|x| < \dfrac{\varepsilon}{\varepsilon + \mu} \leqq 1$ folgt daraus

$$|g(x)| \leqq \mu|x|\,\frac{1}{1 - |x|} < \mu\,\frac{\varepsilon}{\mu} = \varepsilon.$$

Die zu $a \in \mathbf{C}$ konjugiert komplexe Zahl werde mit $\bar{a}$ bezeichnet. Dann gilt $\overline{a + b} = \bar{a} + \bar{b}$ und $\overline{ab} = \bar{a}\bar{b}$. Ist $f(x) \in \mathbf{C}[x]$, so sei $\bar{f}(x)$ dasjenige Polynom, welches entsteht, wenn in $f(x)$ jeder Koeffizient durch die konjugiert komplexe Zahl ersetzt wird.

Hilfssatz 2. *Sei $f(x) \in \mathbf{C}[x]$. Dann ist $F(x) := f(x)\,\bar{f}(x) \in \mathbf{R}[x]$, und es gilt:*

$f(x)$ besitzt eine Nullstelle aus $\mathbf{C} \Leftrightarrow F(x)$ besitzt eine Nullstelle aus $\mathbf{C}$.

Beweis. Offenbar hat $f(x)\,\bar{f}(x)$ reelle Koeffizienten, wenn $f(x)$ das Nullpolynom ist oder den Grad 0 besitzt.

Induktionsannahme: Ist der Grad von $f(x) \leqq m - 1$, so hat $f(x)\,\bar{f}(x)$ reelle Koeffizienten. Es sei

$$f(x) = a_m x^m + g(x) \qquad (a_m \neq 0 \wedge \operatorname{Grad} g(x) \leqq m - 1)$$

ein Polynom vom Grade m. Dann ist

$$\begin{aligned} F(x) = f(x)\,\bar{f}(x) &= \left(a_m x^m + g(x)\right)\left(\bar{a}_m x^m + \bar{g}(x)\right) \\ &= a_m \bar{a}_m x^{2m} + \left(\bar{a}_m g(x) + a_m \bar{g}(x)\right) x^m + g(x)\,\bar{g}(x). \end{aligned}$$

Die Koeffizienten gleicher Potenzen von x in $\bar{a}_m g(x)$ und $a_m \bar{g}(x)$ sind zueinander konjugiert komplexe Zahlen. Daher hat $\bar{a}_m g(x) + a_m \bar{g}(x)$ reelle Koeffizienten. Weil $a_m \bar{a}_m \in \mathbf{R}$ und nach der Induktionsannahme $g(x)\,\bar{g}(x) \in \mathbf{R}[x]$ ist, besitzt $f(x)\,\bar{f}(x)$ reelle Koeffizienten.

Ist $\alpha \in \mathbf{C}$ und $f(\alpha) = 0$, so auch $F(\alpha) = f(\alpha)\,\bar{f}(\alpha) = 0$. Ist umgekehrt $\alpha \in \mathbf{C}$ Nullstelle von $F(x)$, so ist α oder $\bar{\alpha}$ Nullstelle von $f(x)$. Aus $F(\alpha) = f(\alpha)\,\bar{f}(\alpha) = 0$ folgt nämlich $f(\alpha) = 0$ oder $\bar{f}(\alpha) = 0$. $\bar{f}(\alpha) = 0$ ergibt wegen $\overline{0} = 0$, daß

$$\overline{\bar{f}(\alpha)} = \bar{\bar{f}}(\bar{\alpha}) = f(\bar{\alpha}) = 0$$

ist.

Hilfssatz 3. $f(x) \in \mathsf{R}[x] \wedge 2 \nmid \operatorname{Grad} f(x) \Rightarrow \bigvee\limits_{\alpha \in \mathsf{R}} f(\alpha) = 0.$

Beweis. O.B.d.A. kann $f(x)$ mit dem höchsten Koeffizienten 1, also in der Form

$$f(x) = x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0 \qquad (2 \nmid m)$$

angenommen werden. Sei

$$g(t) := a_{m-1}t + \cdots + a_1 t^{m-1} + a_0 t^m$$

und

$$\mu := \max\{|a_{m-1}|, \ldots, |a_0|\}.$$

Dann ist

$$f(x) = x^m\left(1 + a_{m-1}\frac{1}{x} + \cdots + a_1\frac{1}{x^{m-1}} + a_0\frac{1}{x^m}\right) = x^m\left(1 + g\left(\frac{1}{x}\right)\right).$$

Nach Hilfssatz 1 gilt:

$$\left|\frac{1}{x}\right| < \frac{1}{1+\mu} \Rightarrow \left|g\left(\frac{1}{x}\right)\right| < 1,$$

also

$$|x| > |1 + \mu| \Rightarrow 1 + g\left(\frac{1}{x}\right) > 0.$$

Da m ungerade ist, folgt:

$$x < -(1+\mu) \Rightarrow f(x) < 0; \qquad x > 1 + \mu \Rightarrow f(x) > 0.$$

Als stetige Funktion (vgl. MfL, Bd. 4) besitzt $f(x)$ daher wenigstens eine reelle Nullstelle α aus dem Intervall $[\![-1-\mu, 1+\mu]\!]$.

Nach diesen Vorbereitungen beweisen wir den folgenden Satz.

Satz 1 (Fundamentalsatz der Algebra). *Jedes Polynom positiven Grades aus* $\mathbf{C}[x]$ *besitzt in* $\mathbf{C}$ *mindestens eine Nullstelle.*

Beweis. Da die Multiplikation mit einem konstanten Faktor $\neq 0$ die Nullstellen nicht beeinflußt, kann o. B. d. A. angenommen werden, daß der höchste Koeffizient des Polynoms 1 ist. Nach Hilfssatz 2 genügt es, den Satz für Polynome mit reellen Koeffizienten zu beweisen. Es sei

$$f(x) = x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0$$

ein beliebiges Polynom positiven Grades mit Koeffizienten aus dem Körper der reellen Zahlen R. Nach 14.5.3. gibt es einen Erweiterungskörper L von R, der sämtliche Nullstellen von $f(x)$ enthält. In $L[x]$ zerfällt $f(x)$ daher in Linearfaktoren

$$f(x) = (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_m).$$

Wir werden zeigen, daß mindestens ein α_i im Körper der komplexen Zahlen C liegt.

Der Grad m von $f(x)$ kann in der Form

$$m = 2^l u \qquad (l \in \mathbb{N} \wedge 2 \nmid u)$$

geschrieben werden. Wir führen unseren Beweis nun durch vollständige Induktion nach l. Im Fall $l = 0$ ergibt sich aus Hilfssatz 3, daß C eine (sogar reelle) Nullstelle von $f(x)$ enthält.

Induktionsannahme: Jedes Polynom aus $\mathsf{R}[x]$, dessen Grad nicht durch 2^l teilbar ist, besitzt in C wenigstens eine Nullstelle. Wir betrachten nun das Polynom $f(x) \in \mathsf{R}[x]$ vom Grade $m = 2^l u$ ($l \in \mathbb{N}^*$) und bilden mit Hilfe einer beliebigen reellen Zahl c im Zerfällungskörper L die Elemente

$$\beta_{ij} := \alpha_i \alpha_j + c(\alpha_i + \alpha_j) \qquad (i, j \in \{1, \ldots, m\} \wedge i < j).$$

Diese

$$\frac{m(m-1)}{2} = 2^{l-1} u(2^l u - 1) = 2^{l-1} v \qquad (2 \nmid v)$$

Elemente sind Nullstellen des Polynoms

$$g(x) = \prod (x - \beta_{ij}) \qquad (i, j \in \{1, \ldots, m\} \wedge i < j)$$

vom Grade $2^{l-1}v$, dessen Koeffizienten durch die symmetrischen Grundfunktionen der β_{ij} ausgedrückt werden können (vgl. 14.6.2.(5)). Weil bei einer Permutation der α_i die β_{ij} nur untereinander vertauscht werden, sind die Koeffizienten von $g(x)$ auch symmetrische Polynome der α_i mit Koeffizienten aus R und können daher als Polynome der symmetrischen Grundfunktionen der α_i mit Koeffizienten aus R dargestellt werden (vgl. 14.6.2., Satz 1). Die symmetrischen Grundfunktionen der α_i sind aber bis auf das Vorzeichen die reellen Koeffizienten von $f(x)$. Daher sind auch die Koeffizienten von $g(x)$ reell. Nach der Induktionsannahme besitzt $g(x)$ dann wenigstens eine komplexe Nullstelle, d. h., ein β_{ij} liegt in C. Es könnte eventuell von der Wahl der Zahl c abhängen, für welches Indexpaar i, j das zutrifft. Es treten $\frac{m(m-1)}{2}$ verschiedene Indexpaare auf. Führt man die Überlegung für $\frac{m(m-1)}{2} + 1$ verschiedene reelle Zahlen c durch und notiert jedesmal diejenigen Indexpaare k, l, für welche die zugehörigen Elemente β_{kl} in C liegen, so muß wenigstens ein Paar zweimal vorkommen. Folglich gibt es Zahlen $i < j$ aus der Menge $\{1, \ldots, m\}$ sowie zwei verschiedene reelle Zahlen c, c', so daß

$$\beta_{ij} := \alpha_i \alpha_j + c(\alpha_i + \alpha_j),$$
$$\beta_{ij}' := \alpha_i \alpha_j + c'(\alpha_i + \alpha_j)$$

Elemente aus C sind. Dann liegen aber auch

$$-a := \alpha_i + \alpha_j = \frac{\beta_{ij} - \beta_{ij'}}{c - c'}$$

und

$$b := \alpha_i \alpha_j = \frac{c\beta_{ij'} - c'\beta_{ij}}{c - c'}$$

in C, d. h., die Koeffizienten des Polynoms

$$z^2 + az + b = (z - \alpha_i)(z - \alpha_j)$$

sind komplexe Zahlen. Seine Nullstellen α_i und α_j sind bekanntlich (vgl. MfL, Bd. 2, 7.5.) die komplexen Zahlen

$$-\frac{1}{2}\left(a + \sqrt{a^2 - 4b}\right) \quad \text{und} \quad -\frac{1}{2}\left(a - \sqrt{a^2 - 4b}\right).$$

$f(x)$ besitzt also mindestens eine Nullstelle α_i in C. Damit ist der Fundamentalsatz bewiesen.

Folgerung 1. *Sei* $m \in \mathbb{N}^*$. *Dann gilt:*

$f(x) \in \mathbb{C}[x]$ *hat den Grad* $m \Rightarrow f(x)$ *besitzt in* C *genau* m *Nullstellen.*

(*Dabei wird jede Nullstelle in ihrer Vielfachheit gezählt.*)

Beweis. Die Aussage stimmt für Polynome ersten Grades.

Induktionsannahme: Sie stimmt auch für alle Polynome, deren Grad $\leqq m - 1$ ist. Das Polynom $f(x)$ vom Grad m hat nach Satz 1 wenigstens eine Nullstelle α_1 in C und ist daher durch $(x - \alpha_1)$ teilbar (vgl. 14.3.(1)):

$$f(x) = (x - \alpha_1)\, g(x) \qquad \big(g(x) \in \mathbb{C}[x]\big).$$

Da $g(x)$ den Grad $m - 1$ hat, besitzt es nach der Induktionsannahme genau $m - 1$ Nullstellen $\alpha_2, \ldots, \alpha_m$ in C. Daher hat $f(x)$ wenigstens die m Nullstellen $\alpha_1, \alpha_2, \ldots, \alpha_m$ in C. Weitere Nullstellen kann es aber in C nicht geben (vgl. 14.3.1., Satz 1).

Hat $f(x) \in \mathbb{C}[x]$ den höchsten Koeffizienten a_m, so ist es in der Form

$$f(x) = a_m(x - \alpha_1) \cdots (x - \alpha_m)$$

als Produkt von Linearfaktoren aus $\mathbb{C}[x]$ darstellbar. Daher gilt:

Folgerung 2. C *ist Zerfällungskörper (im weiteren Sinn) für alle Polynome aus* $\mathbb{C}[x]$.

Die irreduziblen Polynome aus $\mathbb{C}[x]$ haben also alle den Grad 1. Deshalb gibt es keine echte algebraische Erweiterung des Körpers C. Körper mit dieser Eigenschaft heißen *algebraisch abgeschlossen.*

Ist umgekehrt K ein algebraisch abgeschlossener Körper, so hat jedes irreduzible Polynom aus $K[x]$ den Grad 1, da jedes irreduzible Polynom größeren Grades die Möglichkeit einer echten algebraischen Erweiterung von K bieten würde.

Der Satz 1 ist also gleichwertig mit der Aussage:

Der Körper der komplexen Zahlen ist algebraisch abgeschlossen. (1)

Das ist ein algebraischer Fundamentalsatz der komplexen Zahlen. Der Name „Fundamentalsatz der Algebra" ist entstanden, als sich die Algebra noch auf die Untersuchung des Körpers der komplexen Zahlen, seiner Teilkörper, Teilringe und zugehörigen Polynomringe beschränkte. Da der Körper $\mathbf{C}$ diese zentrale Stellung während der neueren Entwicklung der Algebra verloren hat, trifft eigentlich auch jener Name nicht mehr zu.

Den ersten Beweis des „Fundamentalsatzes der Algebra" gab GAUSS 1799 in seiner Dissertation an. Er hat später noch weitere Beweise für diesen Satz gefunden. Die Hilfsmittel der Funktionentheorie (Theorie der Funktionen komplexer Veränderlicher) gestatten heute sehr kurze Beweise des Fundamentalsatzes.

STEINITZ bewies in seiner Arbeit „Algebraische Theorie der Körper" (J. reine angew. Math. 137 (1910), 167—309), daß es zu jedem Körper K einen Erweiterungskörper L gibt, der algebraisch abgeschlossen ist und dessen sämtliche Elemente sogar algebraisch über K sind (vgl. 14.5.2., Definition 2).

14.8. Das Problem der Auflösung algebraischer Gleichungen durch Radikale

K bezeichne einen Körper. Unter einer *reinen Gleichung* in $K[x]$ versteht man eine Gleichung der Form

$$x^n - a = 0 \quad (a \in K \wedge n \in \mathbf{N}^*).$$

Die Lösungen reiner Gleichungen heißen *Radikale*. Wir wollen hier nur Teilkörper des Körpers der komplexen Zahlen $\mathbf{C}$ betrachten. Die Lösungen von

$$x^n - 1 = 0$$

sind die *n-ten Einheitswurzeln*

$$\cos\frac{2k\pi}{n} + i\sin\frac{2k\pi}{n} \qquad (k = 0, 1, 2, \ldots, n-1).$$

Sie bilden bezüglich der Multiplikation eine zyklische Gruppe, deren erzeugende Elemente *primitive n-te Einheitswurzeln* genannt werden.

Ist

$$a = r(\cos\varphi + i\sin\varphi) \qquad (r \in \mathbf{R} \wedge r \geqq 0 \wedge 0 \leqq \varphi < 2\pi)$$

eine komplexe Zahl, so hat die Gleichung

$$x^n - a = 0 \qquad (n \in \mathbf{N}^*)$$

die Radikale

$$\sqrt[n]{a} := \sqrt[n]{r}\left(\cos\frac{\varphi + 2k\pi}{n} + i\sin\frac{\varphi + 2k\pi}{n}\right) \qquad (k = 0, 1, \ldots, n-1).$$

Man beachte, daß hierbei das Symbol $\sqrt[n]{a}$ im Fall $a \neq 0$ n verschiedene komplexe Zahlen bezeichnet, die *n-te Wurzeln aus a* genannt werden. Im Fall $a = 0$ ist $\sqrt[n]{a} := 0$. Dagegen bedeutet $\sqrt[n]{r}$ die reelle n-te Wurzel aus r, d. h. diejenige (eindeutig bestimmte) nichtnegative reelle Zahl w, für die $w^n = r$ ist. Offensichtlich erhält man die sämtlichen Wurzeln der Gleichung $x^n - a = 0$ aus einer einzigen, indem man diese mit den Potenzen ε^k $(k = 0, 1, \ldots, n-1)$ einer primitiven n-ten Einheitswurzel ε multipliziert. (vgl. MfL. Bd. 2, 7.4.).

Der Fundamentalsatz der Algebra garantiert die Existenz von Lösungen der algebraischen Gleichung

$$f(x) := a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0 \qquad \big(f(x) \in \mathsf{C}[x], n \in \mathsf{N}^*, a_n \neq 0\big) \quad (1)$$

im Körper der komplexen Zahlen. Der angegebene Beweis liefert aber keine Methode zur Berechnung solcher Lösungen aus den Koeffizienten des Polynoms $f(x)$. Daher ergibt sich das *Problem der Auflösung algebraischer Gleichungen durch Radikale:*

Kann man die Lösungen der Gleichung (1) erhalten, indem man von den Koeffizienten ausgehend im Körper C nur endlich oft addiert, subtrahiert, multipliziert, dividiert und Lösungen reiner Gleichungen bestimmt (radiziert)?

Es sei K der Durchschnitt aller Teilkörper von C, welche die Koeffizienten a_n, $a_{n-1}, \ldots, a_1, a_0$ der gegebenen Gleichung (1) enthalten. K ist ein Q umfassender Teilkörper von C und $f(x) \in K[x]$. Sodann adjungieren wir zu K die Nullstellen $\alpha_1, \ldots, \alpha_n$ von $f(x)$ und erhalten den Teilkörper

$$\Omega = K(\alpha_1, \ldots, \alpha_n)$$

von C. Ω ist kleinster Zerfällungskörper von $f(x)$. Er heißt *Normalkörper* oder *Galoisscher Körper der Gleichung* (1).

Definition 1. Die Gleichung

$$f(x) := a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0 \quad \big(f(x) \in K[x], n \in \mathsf{N}^*, a_n \neq 0\big)$$

heißt (*durch Radikale*) *auflösbar* $:\Leftrightarrow$ es gibt eine endliche Kette $K = K_0 \subset K_1 \subset \cdots \subset K_t$ durch reine Gleichungen definierter Körpererweiterungen, deren letztes Glied den Normalkörper Ω von $f(x)$ umfaßt: $\Omega \subseteq K_t$.

Bemerkung. Der Körper K_i entsteht also aus K_{i-1} durch Adjunktion eines gewissen Radikals ϱ_i, d. h., es ist

$$K_i = K_{i-1}(\varrho_i) \qquad \text{mit} \qquad \varrho_i^{n_i} = c_i \in K_{i-1} \qquad (i = 1, \ldots, t).$$

Die Relation $\Omega \subseteq K_t$ bedeutet, daß die Lösungen der Gleichung $f(x) = 0$ in endlich vielen Schritten aus den Koeffizienten von $f(x)$ und rationalen Zahlen durch Addition, Subtraktion, Multiplikation, Division und Auflösung reiner Gleichungen berechnet werden können. Sind umgekehrt die Lösungen von $f(x) = 0$ auf diese Weise berechenbar, so gibt es eine endliche Kette von Körpererweiterungen, die bei K beginnt und deren letztes Glied $K_t \supseteq \Omega$ ist.

Beispiele.

1. Die Lösungen x_1, x_2 der Gleichung

$$x^2 + px + q = 0 \qquad (p, q \in \mathbf{C})$$

werden bekanntlich durch die Formel

$$x_{1,2} = -\frac{p}{2} \pm \frac{1}{2}\sqrt{p^2 - 4q}$$

gegeben. Hierbei ist $K = \mathbf{Q}(p, q)$, $c_1 = p^2 - 4q \in K$ und ϱ_1 einer der zwei Werte der komplexen Wurzel $\sqrt{p^2 - 4q}$. Dann ist $\Omega \subseteq K_1 = K(\varrho_1)$. Falls ϱ_1 schon in K liegt, ist $K_1 = K$. Jede quadratische Gleichung aus $\mathbf{C}[x]$ ist also durch Radikale auflösbar.

2. Für die Gleichung

$$x^3 + px + q = 0 \qquad (p, q \in \mathbf{C}) \tag{2}$$

ist $K = \mathbf{Q}(p, q)$. Weil $x^3 = 0$ sicher durch Radikale auflösbar ist, betrachten wir nur Fälle mit $p \neq 0$ oder $q \neq 0$. Der Wert

$$c_1 = \left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3$$

liegt in K, ϱ_1 sei einer der zwei Werte der komplexen Wurzel $\sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}$. Wenn $\varrho_1 \in K$ ist, sei $K_1 = K$. Anderenfalls ist $K_1 = K(\varrho_1)$ eine echte Erweiterung von K. Der Wert

$$c_2 = -\frac{q}{2} + \varrho_1$$

liegt in K_1. Das Radikal ϱ_2 sei einer der drei Werte der komplexen Wurzel $\sqrt[3]{-\frac{q}{2} + \varrho_1}$. Wenn $\varrho_2 \in K_1$ ist, sei $K_2 = K_1$. Anderenfalls ist $K_2 = K_1(\varrho_2)$ eine echte Erweiterung von K_1. ϱ_3 und ϱ_4 seien die beiden anderen Werte der komplexen Wurzel $\sqrt[3]{-\frac{q}{2} + \varrho_1}$. Liegen sie auch in K_2, so sei $K_3 = K_2$. Sonst adjungieren wir zu K_2 eine Lösung ϱ_5

der Gleichung $x^2 + 3 = 0$. In $K_3 = K_2(\varrho_5)$ sind dann die beiden primitiven dritten Einheitswurzeln $-\frac{1}{2} \pm \frac{1}{2}\sqrt{-3}$ und daher neben ϱ_2 auch ϱ_3 und ϱ_4 enthalten.

Bekanntlich werden die Lösungen x_1, x_2, x_3 der Gleichung (2) durch die Formeln

$$x_1 = \varrho_2 - \frac{p}{3\varrho_2}, \qquad x_2 = \varrho_3 - \frac{p}{3\varrho_3}, \qquad x_3 = \varrho_4 - \frac{p}{3\varrho_4}$$

gegeben, wobei $\varrho_2, \varrho_3, \varrho_4$ die drei Werte der komplexen Wurzel

$$\sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}}$$

sind (vgl. MfL, Bd. 2, 7.5.). Weil x_1, x_2, x_3 in K_3 liegen, ist der Normalkörper Ω der Gleichung (2) in K_3 enthalten und daher die Gleichung durch Radikale auflösbar.

Auch die algebraischen Gleichungen vierten Grades mit Koeffizienten aus C sind durch Radikale auflösbar, und es gibt sogar eine allgemeine Auflösungsformel, welche die Wurzeln dieser Gleichungen durch Radikale darstellt (vgl. MfL, Bd. 2, 7.6.).

Allgemeine Auflösungsformeln für Gleichungen dritten und vierten Grades wurden bereits im 16. Jahrhundert von CARDANO und FERRARI angegeben. Daher vermutete man, daß es auch für die Gleichungen höheren Grades möglich sein müßte, Formeln zu finden. welche die Lösungen solcher Gleichungen durch Radikale darstellen. Die Vermutung wurde 1826 durch ABEL widerlegt, nachdem bereits PAOLO RUFFINI (1765—1822) Untersuchungen in dieser Richtung veröffentlicht hatte.

Ohne Beweis geben wir hier den folgenden Satz von ABEL an:

Satz 1. *Für die algebraischen Gleichungen* (1), *deren Grad* $n \geqq 5$ *ist, gibt es keine allgemeine Auflösungsformel, durch die jeweils eine Wurzel jeder dieser Gleichungen durch Radikale dargestellt wird.*

Einen Beweis, der mit den hier erarbeiteten Hilfsmitteln zugänglich ist, findet der Leser in der „Enzyklopädie der Elementarmathematik", Band II, 5. Aufl., Berlin 1972, S. 255—263. Der Satz 1 läßt noch die Möglichkeit offen, daß jede algebraische Gleichung (1) eine von der gegebenen Gleichung abhängige Auflösung durch Radikale besitzt.

Über Satz 1 hinaus kann man aber zeigen, daß das *allgemeine Polynom* (das ist ein Polynom mit Unbestimmten als Koeffizienten) vom Grade n im Fall $n \geqq 5$ nicht durch Radikale auflösbar ist. Der Beweis beruht auf Hilfsmitteln der *Galois-Theorie*, die im Rahmen dieses Buches nicht dargestellt werden kann.

Es lassen sich davon unabhängig auch konkrete Gleichungen angeben, die nicht durch Radikale auflösbar sind, beispielsweise alle Gleichungen der Form

$$x^5 - p^2x - p = 0,$$

wobei p eine beliebige Primzahl bezeichnet (s. o., S. 282).

Für die Auflösung von Gleichungen fünften und höheren Grades ist man daher auf Näherungsverfahren angewiesen, wie sie in der numerischen Mathematik entwickelt wurden und auch schon zur Lösung von Gleichungen dritten und vierten Grades angewendet werden.

14.9. Partialbruchzerlegung

$K(x)$ sei der Körper der rationalen Funktionen in x mit Koeffizienten aus dem Körper K (vgl. 14.5.1.). Seine Elemente haben die Gestalt

$$\varrho(x) = \frac{h(x)}{g(x)} \qquad \big(h(x) \in K[x] \wedge g(x) \in K[x] \setminus \{0\}\big).$$

Falls Grad $h(x) \geqq$ Grad $g(x)$ ist, ergibt die Division von $h(x)$ durch $g(x)$ mit Rest

$$h(x) = q(x)\, g(x) + f(x) \qquad \big(q(x) \in K[x] \wedge f(x) \in K[x] \wedge \text{Grad } f(x) < \text{Grad } g(x)\big).$$

Daher kann $\varrho(x)$ in der Form

$$\varrho(x) = \frac{h(x)}{g(x)} = q(x) + \frac{f(x)}{g(x)}$$

dargestellt werden.

Weil

$$f(x)\,(g(x)u(x)) = f(x)\,(u(x)g(x)) = (f(x)u(x))\,g(x)$$

ist, gilt für $u(x) \in K[x] \setminus \{0\}$

$$\frac{f(x)}{g(x)} = \frac{f(x)\,u(x)}{g(x)\,u(x)}$$

(vgl. 13.6.(1)). Deshalb kann $f(x) \sqcap g(x) = 1$ angenommen werden. Überdies ist o.B.d.A. 1 der höchste Koeffizient von $g(x)$.

Grundlage für die weitere Zerlegung von $\dfrac{f(x)}{g(x)}$ in eine Summe von Partialbrüchen ist der

Satz 1. *Bezeichne K einen Körper und $k(x)$, $l(x)$ teilerfremde Polynome aus $K[x]$ mit Grad $k(x) = a$ und Grad $l(x) = b$. Ist $f(x)$ ein beliebiges Polynom aus $K[x]$, dessen Grad kleiner als $a + b$ ist, so gibt es in $K[x]$ Polynome $s(x)$, $t(x)$ mit Grad $s(x) < b$ und Grad $t(x) < a$, für die*

$$f(x) = s(x)\, k(x) + t(x)\, l(x) \tag{1}$$

gilt.

Beweis. Da $k(x) \sqcap l(x) = 1$ ist, gibt es in $K[x]$ Polynome $c(x)$, $d(x)$, für die

$$1 = c(x)\,k(x) + d(x)\,l(x)$$

gilt. Dann ist

$$f(x) = f(x)\,c(x)\,k(x) + f(x)\,d(x)\,l(x). \tag{2}$$

Division von $f(x)\,c(x)$ durch $l(x)$ mit Rest ergibt

$$f(x)\,c(x) = v(x)\,l(x) + s(x) \qquad \big(v(x), s(x) \in K[x]\big) \tag{3}$$

wobei der Grad $s(x) <$ Grad $l(x) = b$ ist. Setzt man (3) in (2) ein, so folgt mit

$$t(x) := f(x)\,d(x) + k(x)\,v(x)$$

die Gleichung

$$\begin{aligned} f(x) &= s(x)\,k(x) + \big(f(x)\,d(x) + k(x)\,v(x)\big)\,l(x) \\ &= s(x)\,k(x) + t(x)\,l(x). \end{aligned}$$

Da die Grade von $f(x)$ und $s(x)\,k(x)$ kleiner als $a + b$ sind, ist Grad $t(x) < a$.

Dividiert man auf beiden Seiten der Gleichung (1) durch $k(x)\,l(x)$, so erhält man

$$\frac{f(x)}{k(x)\,l(x)} = \frac{s(x)}{l(x)} + \frac{t(x)}{k(x)}. \tag{4}$$

Auf der linken Seite der Gleichung (4) ist voraussetzungsgemäß der Grad des Zählers kleiner als der Grad des Nenners. Dasselbe gilt für die beiden *Partialbrüche* $\frac{s(x)}{l(x)}$ und $\frac{t(x)}{k(x)}$.

Sind $f(x)$ und $k(x)\,l(x)$ teilerfremd, so auch $s(x)$ und $l(x)$ sowie $t(x)$ und $k(x)$. Hat $k(x)\,l(x)$ den höchsten Koeffizienten 1, so kann man o.B.d.A. annehmen, daß die höchsten Koeffizienten von $k(x)$ und $l(x)$ ebenfalls 1 sind.

Läßt sich nun in einem der Teilbrüche der Nenner wiederum als Produkt von zwei teilerfremden Faktoren schreiben, so kann er erneut als Summe von zwei Partialbrüchen dargestellt werden. Nach endlich vielen Wiederholungen erhält man eine Summe von Partialbrüchen, deren Nenner sämtlich Potenzen von Primpolynomen aus $K[x]$ sind. Damit ist der folgende Satz bewiesen.

Satz 2. *Sei K ein Körper und $\frac{f(x)}{g(x)} \in K(x)$. Ist Grad $f(x) <$ Grad $g(x)$, so kann $\frac{f(x)}{g(x)}$ als Summe von Partialbrüchen geschrieben werden, deren Nenner diejenigen Potenzen von Primpolynomen aus $K[x]$ sind, welche in der Primelementzerlegung von $g(x)$ auftreten. In jedem Partialbruch ist der Grad des Zählers kleiner als der Grad des Nenners.*

Die erhaltenen Partialbrüche $\frac{z(x)}{p(x)^t}$ ($t \in \mathbb{N}^*$), deren Nenner Potenzen von Primpolynomen sind, lassen sich noch weiter aufspalten. Bezeichnet $l \in \mathbb{N}^*$ den Grad von $p(x)$, so ist Grad $z(x) < lt$. Dividiert man $z(x)$ durch $p(x)^{t-1}$, so erhält man einen Rest $z_1(x) \in K[x]$, dessen Grad kleiner als $l(t-1)$ ist. Dividiert man diesen durch $p(x)^{t-2}$, so entsteht ein Rest von einem Grad $< l(t-2)$ usw. Es ergeben sich die Gleichungen

$$\begin{aligned}
z(x) &= q_1(x)\, p(x)^{t-1} + z_1(x),\\
z_1(x) &= q_2(x)\, p(x)^{t-2} + z_2(x),\\
&\cdots\cdots\cdots\cdots\cdots\cdots\\
z_{t-2}(x) &= q_{t-1}(x)\, p(x) + z_{t-1}(x),\\
z_{t-1}(x) &= q_t(x),
\end{aligned}$$

aus denen

$$z(x) = q_1(x)\, p(x)^{t-1} + q_2(x)\, p(x)^{t-2} + \cdots + q_{t-1}(x)\, p(x) + q_t(x)$$

und

$$\frac{z(x)}{p(x)^t} = \frac{q_1(x)}{p(x)} + \frac{q_2(x)}{p(x)^2} + \cdots + \frac{q_{t-1}(x)}{p(x)^{t-1}} + \frac{q_t(x)}{p(x)^t} \tag{5}$$

folgt. Die $q_i(x)$ ($i = 1, \ldots, t$) haben sämtlich einen Grad $< l$. In allen Partialbrüchen auf der rechten Seite der Gleichung (5) ist daher der Grad des Zählers kleiner als der Grad des Nenners. Damit ist bewiesen:

Satz 3 (Satz von der Partialbruchzerlegung). *Es sei K ein Körper, $\frac{f(x)}{g(x)} \in K(x)$, Grad $f(x) <$ Grad $g(x)$ und*

$$g(x) = p_1(x)^{t_1}\, p_2(x)^{t_2} \cdots p_n(x)^{t_n}$$

die Primfaktorzerlegung von $g(x)$ in $K[x]$. Dann ist $\frac{f(x)}{g(x)}$ als Summe von Partialbrüchen darstellbar:

$$\frac{f(x)}{g(x)} = \sum_{\nu=1}^{n} \left(\frac{q_{\nu 1}(x)}{p_\nu(x)} + \frac{q_{\nu 2}(x)}{p_\nu(x)^2} + \cdots + \frac{q_{\nu t_\nu}(x)}{p_\nu(x)^{t_\nu}} \right), \tag{6}$$

wobei für jedes $\nu \in \{1, \ldots, n\}$ die Zähler $q_{\nu\tau_\nu}(x)$ ($\tau_\nu = 1, 2, \ldots, t_\nu$) entweder Null sind oder einen kleineren Grad als $p_\nu(x)$ besitzen.

Wenn die Primpolynome $p_\nu(x)$ sämtlich den Grad 1 haben, sind alle Zähler der Partialbrüche Elemente aus K. In diesem Fall kann die Partialbruchzerlegung sehr einfach gewonnen werden. Tritt $(x - \alpha)$ mit dem Exponenten $t \in \mathbb{N}^*$ in der Primfaktorzerlegung von $g(x)$ auf, ist also $g(x) = (x-\alpha)^t h(x)$ mit einem nicht mehr durch $(x-\alpha)$ teilbaren Polynom $h(x)$ aus $K[x]$, so kann in

$$\frac{f(x)}{g(x)} = \frac{f(x)}{(x-\alpha)^t h(x)} = \frac{\beta}{(x-\alpha)^t} + \frac{f(x) - \beta h(x)}{(x-\alpha)^t\, h(x)}$$

die Konstante $\beta \in K$ so bestimmt werden, daß

$$f(\alpha) - \beta h(\alpha) = 0$$

ist. Bei dieser Wahl von β hat $f(x) - \beta h(x)$ die Nullstelle α und ist daher durch $(x - \alpha)$ teilbar:

$$f(x) - \beta h(x) = (x - \alpha)\, f_1(x) \qquad \big(f_1(x) \in K[x]\big).$$

Mit

$$\frac{f(x) - \beta h(x)}{(x - \alpha)^t h(x)} = \frac{f_1(x)}{(x - \alpha)^{t-1} h(x)}$$

wird nun ebenso verfahren, bis die vollständige Partialbruchzerlegung (6) von $\dfrac{f(x)}{g(x)}$ erreicht ist.

Da nach dem Fundamentalsatz der Algebra alle Primpolynome von $\mathbf{C}[x]$ den Grad 1 haben, gilt die

Folgerung 1. *Hat $g(x) \in \mathbf{C}[x]$ die (paarweise verschiedenen) Nullstellen $\alpha_1, \ldots, \alpha_n$ mit den Vielfachheiten $t_1, \ldots, t_n$ und ist der Grad von*

$$g(x) = (x - \alpha_1)^{t_1} \cdots (x - \alpha_n)^{t_n}$$

größer als der Grad von $f(x) \in \mathbf{C}[x]$, so kann $\dfrac{f(x)}{g(x)} \in \mathbf{C}(x)$ als Summe von Partialbrüchen dargestellt werden, deren Zähler $\beta_{\nu\tau_\nu}$ $(\tau_\nu = 1, \ldots, t_\nu;\ \nu = 1, \ldots, n)$ Elemente aus $\mathbf{C}$ sind:

$$\frac{f(x)}{g(x)} = \sum_{\nu=1}^{n} \left(\frac{\beta_{\nu 1}}{x - \alpha_\nu} + \frac{\beta_{\nu 2}}{(x - \alpha_\nu)^2} + \cdots + \frac{\beta_{\nu t_\nu}}{(x - \alpha_\nu)^{t_\nu}} \right).$$

14.10. Übungsaufgaben

1. Man bestimme den größten gemeinsamen Teiler der Polynome

$$x^5 - 2x^4 - x + 2$$

und

$$x^4 - 5x^3 + 7x^2 - 5x + 6$$

und stelle ihn in $\mathbf{Q}[x]$ als Vielfachsumme dar.

2. I sei ein Integritätsbereich mit Einselement und habe die Charakteristik 0. L bezeichne einen Integritätsbereich, der Erweiterungsring von I ist. $\alpha \in L$ sei eine Nullstelle des vom Nullpolynom verschiedenen Polynoms $f(x)$ aus $I[x]$. Man beweise:
Ist α $(k-1)$-fache Nullstelle von $f'(x)$, so ist α k-fache Nullstelle von $f(x)$ $(k \in \mathbf{N}^*)$.

3. Sei p eine Primzahl. Man gebe in $\mathbf{Z}/(p)\,[x]$ vom Nullpolynom verschiedene Polynome an, die an sämtlichen Stellen $\alpha \in \mathbf{Z}/(p)$ den Wert $[0]$ besitzen.

4. Man bestimme in $\mathbf{Q}[x]$ dasjenige Polynom $f(x)$ mit Grad $f(x) \leqq 4$, für welches

$$f(-2) = 16,\ f(-1) = 2,\ f(0) = 0,\ f(1) = 4,\ f(2) = 32$$

gilt.

5. Mit Hilfe des Verfahrens von KRONECKER beweise man, daß die Polynome

$$x^3 - 2x + 2 \qquad \text{und} \qquad x^5 - x^2 + 1$$

in $\mathbf{Q}[x]$ irreduzibel sind.

6. Man stelle das Polynom

$$6x^5 - 12x^4 - 6x + 12$$

als ein Produkt von Primelementen aus $\mathbf{Z}[x]$ dar.

7. Hat das Polynom

$$f(x) = x^6 - 3x^4 + 4$$

mehrfache Nullstellen?
Gegebenenfalls bestimme man ein Polynom, das genau dieselben Nullstellen wie $f(x)$ besitzt, aber jede mit Vielfachheit 1.

8. Man zerlege die Polynome

$$f(x) = x^6 - 3x^4 + 4$$

und

$$g(x) = x^5 - 4x^4 + 2x^3 - 8x^2 + x - 4$$

in Produkte von Primpolynomen aus $\mathbf{Q}[x]$ und gebe sämtliche Nullstellen beider Polynome an.

9. Ein Polynom mit Koeffizienten aus einem Ring mit Nullteilern kann mehr Nullstellen besitzen, als sein Grad beträgt.
Als Beispiel gebe man etwa ein quadratisches Polynom mit Koeffizienten aus $\mathbf{Z}/(6)$ an, das drei Nullstellen in diesem Ring besitzt.

10. Die Elemente des Restklassenkörpers $K = \mathbf{Z}/(2)$ seien mit $[0]$ und $[1]$ bezeichnet. Das Polynom

$$p(x) = [1]\,x^2 + [1]\,x + [1]$$

ist in $K[x]$ irreduzibel.
Man konstruiere $K[x]/(p(x))$ und stelle für diese algebraische Erweiterung von K die Additions und Multiplikationstabellen auf. (Vgl. 13.8., Aufgabe 11.)

Literatur

[1] Boseck, H., Einführung in die Theorie der linearen Vektorräume, 3. Aufl., VEB Deutscher Verlag der Wissenschaften, Berlin 1973.

[2] Enzyklopädie der Elementarmathematik, Bd. II: Algebra, 6. Aufl., VEB Deutscher Verlag der Wissenschaften, Berlin 1977 (Übersetzung aus dem Russischen).

[3] Grundzüge der Mathematik, Bd. 1: Grundlagen der Mathematik, Arithmetik und Algebra, 2. Aufl., Vandenhoeck & Ruprecht, Göttingen 1962.

[4] Hasse, H., Höhere Algebra, Bd. 1, 2, W. de Gruyter, Berlin 1926/1927.

[5] Hasse, H., Aufgabensammlung zur höheren Algebra, W. de Gruyter, Berlin 1934.

[6] Keller, O.-H., Analytische Geometrie und lineare Algebra, 3. Aufl., VEB Deutscher Verlag der Wissenschaften, Berlin 1968.

[7] Kochendörffer, R., Einführung in die Algebra, 4. Aufl., VEB Deutscher Verlag der Wissenschaften, Berlin 1974.

[8] Kochendörffer, R., Lehrbuch der Gruppentheorie, Akademische Verlagsgesellschaft Geest & Portig, Leipzig 1966.

[9] Kuroš, A. G., Vorlesungen über allgemeine Algebra, B. G. Teubner Verlagsgesellschaft, Leipzig 1964 (Übersetzung aus dem Russischen).

[10] Lichnerowicz, A., Lineare Algebra und lineare Analysis, VEB Deutscher Verlag der Wissenschaften, Berlin 1956 (Übersetzung aus dem Französischen).

[11] Lugowski, H., und H. J. Weinert, Grundzüge der Algebra, Teil 1: Allgemeine Gruppentheorie, 4. Aufl., 1968. Teil 2: Allgemeine Ring- und Körpertheorie, 4. Aufl., 1971. Teil 3: Auflösungstheorie algebraischer Gleichungen, 2. Aufl., 1967. B. G. Teubner Verlagsgesellschaft, Leipzig.

[12] Rédei, L., Algebra, Teil 1, Akademische Verlagsgesellschaft Geest & Portig, Leipzig 1959.

[13] Vieregge, H., Einführung in die klassische Algebra, VEB Deutscher Verlag der Wissenschaften, Berlin 1972.

[14] van der Waerden, B. L., Algebra, I, II, 8. bzw. 5. Aufl. der Modernen Algebra), Springer-Verlag Berlin—Heidelberg—New York 1971 bzw. 1967.

[15] Weber, H., Lehrbuch der Algebra, Bd. 1,2, Vieweg, Braunschweig 1895/1896.

Namen- und Sachverzeichnis