Lineare Algebra I & II

Gert-Martin Greuel
Universität Kaiserslautern
Fachbereich Mathematik

Ausarbeitung der je vierstündigen Vorlesung im Wintersemester 1999/2000 und Sommersemester 2000 durch Thomas Keilen

Lineare Algebra I & II

je vierstündige Vorlesung im Wintersemester 1999/2000
und Sommersemester 2000

Gert-Martin Greuel
Universität Kaiserslautern
Fachbereich Mathematik

Inhaltsverzeichnis

Einleitung

Grundlegend für die lineare Algebra sind die Begriffe des Vektorraumes und des linearen Gleichungssystems, die beide eng untereinander zusammenhängen.

Lineare Gleichungssysteme sind in allen Wissenschaften, in denen mathematische Modelle und Methoden verwendet werden unvermeidlich, seien es Natur- und Ingenieurwissenschaften oder Wirtschafts- und Sozialwissenschaften. Allerdings liegen die Probleme nicht in Form von Gleichungssystemen vor, sondern oft ist die Aufstellung der Gleichungssysteme durch geeignete Modellbildung und durch Formulierung in mathematischer Sprache das Hauptproblem. Dies ist aber nicht Gegenstand der linearen Algebra, diese beschäftigt sich mit der Struktur der Lösungsmenge linearer Gleichungssysteme, d. h. mit den Vektorräumen.

Historisch hat der Vektorraumbegriff allerdings nicht diesen operationellen Ursprung, sondern er hat sich aus den geometrischen Objekten wie Gerade, Ebene und dreidimensionaler Anschauungsraum entwickelt. Durch Einführung eines *künstlichen* Ursprungs und gradliniger *kartesischer* Koordinaten durch Pierre de Fermat (1601-1655) und vor allem durch René Descartes (1596-1650) gelang eine Verschmelzung von Geometrie und Algebra und damit wurde es möglich, geometrische Objekte durch Gleichungen zu beschreiben. Aber erst durch den Mathematiker und Gymnasiallehrer Hermann Günther Graßmann (1809-1877) wurde die Theorie des n-dimensionalen Vektorraumes als Theorie der *Ausdehnungsgebiete* erstmals axiomatisch formalisiert und entwickelt. Als fundamentaler Begriff der Mathematik konnte sich der Begriff des Vektorraumes allerdings erst sehr viel später aufgrund des Erfolges des dreidimensionalen Vektorkalküls in der Physik durchsetzen. Der Begriff eines beliebigen reellen Vektorraumes taucht zuerst 1922 bei dem Analytiker Stephan Banach (1892-1945) auf, während sich der Begriff des abstrakten Vektorraumes über einem beliebigen Körper erst im Zuge der *modernen Algebra* in den Arbeiten von Emmy Noether (1882-1935) und Barthel Leendert van der Waerden (1903-1996) als Spezialfall des Moduls über einem Ring entwickelt. Die endgültige Fixierung des Vektorraumbegriffs im heutigen Sinne findet sich erst 1947 in den Werken von Nikolas Bourbaki. Höchstinteressante Aspekte dieser historischen Entwicklung findet man in den Lehrbüchern von Brieskorn [**Bri83**] und Köcher [**Köc83**].

Lineare Gleichungssysteme und sogar ihre Lösung mit Hilfe des, wie man heute sagt, gaußschen Algorithmus waren natürlich schon sehr viel früher bekannt, etwa als Methode des *fang cheng* in einem chinesischen Mathematiklehrbuch des zweiten vorchristlichen Jahrhunderts.

Wir überspringen diese Zeitspanne von über 2000 Jahren und geben mit einigen Beispielen eine Idee von den vielfältigen Anwendungen linearer Gleichungssysteme. Für die exakten Definitionen der verwendeten mathematischen Begriffe verweisen wir auf die späteren Kapitel des Skriptes.

Da praktische Anwendungen der Mathematik sich heute letztlich immer als Computerprogramme in Software niederschlagen, wird die computermäßige Behandlung von Begriffen und Methoden von Beginn an systematisch mitentwickelt.

Beispiele linearer Gleichungssysteme

Wir geben jetzt einige Beispiele von Gleichungssystemen, die zum Teil aus Anwendungen kommen. Sie illustrieren die große praktische Bedeutung linearer Gleichungssysteme.

Beispiel 1.1

Ein Vater hat einen Sohn und eine Tochter. Der Vater ist viermal so alt wie sein Sohn und der Sohn ist fünf Jahre älter als seine Schwester. In fünf Jahren sind Vater und Sohn zusammen sechsmal so alt wie die Tochter.

Wie alt sind Vater, Sohn und Tochter?

Das lineare Gleichungssystem mit v = Alter des Vaters, s = Alter des Sohnes, und t = Alter der Tochter lautet:

$$v = 4s,\ s = t + 5,\ (v + 5) + (s + 5) = 6(t + 5).$$

Das Gleichungssystem schreiben wir systematisch folgendermaßen auf:

$$\begin{aligned} v - 4s + 0 \cdot t &= 0, \\ 0 \cdot v + s - t &= 5, \\ v + s - 6t &= 20. \end{aligned}$$

Dies ist ein lineares Gleichungssystem in den Unbestimmten v, s, t.

Die Lösung mit Hilfe des gaußschen Algorithmus geht wie folgt:

$$\left(\begin{array}{rrr|r} 1 & -4 & 0 & 0 \\ 0 & 1 & -1 & 5 \\ 1 & 1 & -6 & 20 \end{array}\right) \mapsto \left(\begin{array}{rrr|r} 1 & -4 & 0 & 0 \\ 0 & 1 & -1 & 5 \\ 0 & 5 & -6 & 20 \end{array}\right) \mapsto \left(\begin{array}{rrr|r} 1 & -4 & 0 & 0 \\ 0 & 1 & -1 & 5 \\ 0 & 0 & -1 & -5 \end{array}\right)$$

Als Lösung erhalten wir also: $t = 5,\ s = 10,\ v = 40$, d. h. der Vater ist 40 Jahre alt, sein Sohn zehn und seine Tochter fünf.

Beispiel 1.2
Wir definieren eine *Ebene* im $\mathbb{R}^3$ als Lösungsmenge einer linearen Gleichung

$$E : a_1x_1 + a_2x_2 + a_3x_3 = b$$

mit $a_1, a_2, a_3, b \in \mathbb{R}$ und $a_i \neq 0$ für mindestens ein i.

Dies stimmt mit der Anschauung überein (sind alle a_i und b gleich 0, so erhalten wir als Lösungsmenge den ganzen $\mathbb{R}^3$, sind alle $a_i = 0$ und $b \neq 0$, so ist die Lösungsmenge leer).

Um den Schnitt der beiden Ebenen, die durch die Gleichungen $E_1 : x_1 + x_2 + 2x_3 = 2$ und $E_2 : x_1 + x_3 = 4$ gegeben sind, zu bestimmen, müssen wir also das Gleichungssystem aus diesen beiden Gleichungen lösen, wobei wir wie in Abschnitt II.7 beschrieben vorgehen:

$$\left(\begin{array}{ccc|c} 1 & 1 & 2 & 2 \\ 1 & 0 & 1 & 4 \end{array}\right) \mapsto \left(\begin{array}{ccc|c} 1 & 1 & 2 & 2 \\ 0 & -1 & -1 & 2 \end{array}\right).$$

Wir erhalten $-x_2 - x_3 = 2$ also $-x_2 = x_3 + 2$ und $x_1 = -x_2 - 2x_3 + 2 = -x_3 + 4$, wobei $x_3 \in \mathbb{R}$ beliebig sein kann. Wir haben also unendlich viele Lösungen, aber in sehr schöner Form. Schreiben wir t statt x_3 so erhalten wir als Lösungsmenge

$$\begin{aligned} E_1 \cap E_2 &= \left\{ \left(\begin{array}{c} 4-t \\ -2-t \\ t \end{array} \right) \middle| t \in \mathbb{R} \right\} = \left\{ \left(\begin{array}{c} 4 \\ -2 \\ 0 \end{array} \right) + t \left(\begin{array}{c} -1 \\ -1 \\ 1 \end{array} \right) \middle| t \in \mathbb{R} \right\} \\ &= \left(\begin{array}{c} 4 \\ -2 \\ 0 \end{array} \right) + \mathbb{R} \cdot \left(\begin{array}{c} -1 \\ -1 \\ 1 \end{array} \right). \end{aligned}$$

Dies ist offensichtlich die Parameterdarstellung einer Geraden im $\mathbb{R}^3$ durch die Punkte $\left(\begin{smallmatrix} 4 \\ -2 \\ 0 \end{smallmatrix}\right)$ und $\left(\begin{smallmatrix} 3 \\ -3 \\ 1 \end{smallmatrix}\right)$.

Beispiel 1.3
Im allgemeinen werden sich zwei Ebenen, E_1, E_2, im $\mathbb{R}^3$ in einer Geraden schneiden, in Spezialfällen können die Ebenen aber parallel sein ($E_1 \cap E_2 = \emptyset$) oder übereinstimmen ($E_1 = E_2$).

Sei E_1 die Ebene

$$E_1 : x_1 + x_2 + 2x_2 = 3$$

und E_2 eine beliebige Ebene

$$E_2 : a_1x_1 + a_2x_2 + a_3x_3 = b.$$

Wir wollen feststellen für welche a_1, a_2, a_3, b entweder $E_1 \cap E_2$ eine Gerade, leer oder E_1 ist:

$$\left(\begin{array}{ccc|c} 1 & 1 & 2 & 3 \\ a_1 & a_2 & a_3 & b \end{array}\right) \mapsto \left(\begin{array}{ccc|c} 1 & 1 & 2 & 3 \\ 0 & a_2 - a_1 & a_3 - 2a_1 & b - 3a_1 \end{array}\right).$$

Die letzte Gleichung lautet

$$(a_2 - a_1)x_2 + (a_3 - 2a_1)x_3 = b - 3a_1.$$

Ein wenig Überlegung liefert (da die Lösungsmenge der ersten Gleichung E_1 ist, und da die Lösungsmenge der zweiten Gleichung unabhängig von x_1 ist):

$$E_1 \cap E_2 = \emptyset \quad \Leftrightarrow \quad a_2 - a_1 = a_3 - 2a_1 = 0, (b - 3a_1) \neq 0, \tag{1}$$

$$E_1 = E_2 \quad \Leftrightarrow \quad a_2 - a_1 = a_3 - 2a_1 = b - 3a_1 = 0. \tag{2}$$

In allen anderen Fällen ist $E_1 \cap E_2$ eine Gerade.

Im Fall $E_1 = E_2$ haben wir wieder ein Gleichungssystem (2) mit drei Gleichungen in den vier Unbestimmten a_1, a_2, a_3, b zu lösen. Dies ist homogen, wir lassen deshalb die letzte (Null-)Spalte weg:

$$\begin{pmatrix} -1 & 1 & 0 & 0 \\ -2 & 0 & 1 & 0 \\ -3 & 0 & 0 & 1 \end{pmatrix} \mapsto \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -2 & 1 & 0 \\ 0 & -3 & 0 & 1 \end{pmatrix} \mapsto \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -2 & 1 & 0 \\ 0 & 0 & -\frac{3}{2} & 1 \end{pmatrix}.$$

Da wir die letzte Nullspalte weggelassen haben, ergibt sich durch Rücksubstitution, $b = \frac{3}{2}a_3$, $2a_2 = a_3$ und $a_1 = a_2$, oder $(a_1, a_2, a_3, b) = a_1 \cdot (1, 1, 2, 3)$.

Daraus können wir aber alle drei Fälle ablesen:

$E_1 = E_2$ genau dann, wenn die Gleichung von E_2 ein Vielfaches $\neq 0$ der Gleichung von E_1 ist; $E_1 \cap E_2 = \emptyset$ genau dann, wenn der Koeffizientenvektor (a_1, a_2, a_3) ein Vielfaches $\neq 0$ des Koeffizientenvektors von E_1 ist, aber die rechte Seite b von E_2 nicht das gleiche Vielfache der rechten Seite von E_1 ist; und $E_1 \cap E_2$ ist eine Gerade in allen anderen Fällen.

Beispiel 1.4

Im "Computer aided geometric design" (CAGD) werden zum Design von Flächen und Kurven (z. B. im Automobil- oder Flugzeugbau) Flächen- und Kurvenstücke verwendet (meist durch sogenannte kubische Splines realisiert), die dann an den Endpunkten oder Randkurven glatt zusammenpassen müssen. Am bekanntesten sind die Bézier-Kubiken, die von dem französischen Auto-Designer bei Renault, P. Bézier, eingeführt wurden (diese werden heute z. B. auch in der Text-Beschreibungssprache PostScript verwendet).

Ein typisches Problem ist z.B. die Bestimmung einer kubischen Parabel

$$f(x) = ax^3 + bx^2 + cx + d$$

durch zwei Punkte (x_1, y_1), (x_2, y_2) in der Ebene mit vorgegebener Steigung m_1 in (x_1, y_1) und m_2 in (x_2, y_2).

Für $(x_1, y_1) = (0, 2)$, $(x_2, y_2) = (4, 0)$, $m_1 = -3$, $m_2 = -3$ ergibt sich aus

$$f'(x) = 3ax^2 + 2bx + c$$

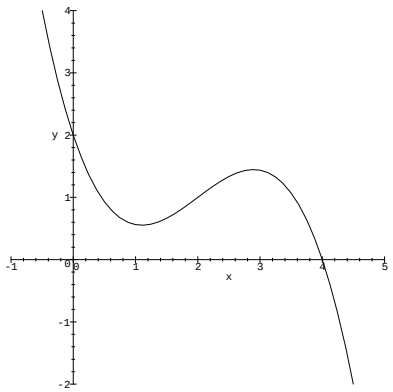
und

$$f(0) = 2, f(4) = 0, f'(0) = -3 \text{ und } f'(4) = -3$$

das lineare Gleichungssystem

$$\begin{aligned} d &= 2, \\ 64a + 16b + 4c + d &= 0, \\ c &= -3, \\ 48a + 8b + c &= -3, \end{aligned}$$

also $d = 2$, $c = -3$, $6a + b = 0$, $32a + 8b = 5$, und damit $a = -\frac{5}{16}$ und $b = \frac{15}{8}$. Die Kurve $y = -\frac{5}{16}x^3 + \frac{15}{8}x^2 - 3x + 2$ hat etwa die folgende Gestalt



Die Aufgabe ist, wie leicht zu sehen ist, stets lösbar und daher können kubische Splines stückweise definiert und glatt aneinander gesetzt werden.

Beispiel 1.5

Die folgende Planungsaufgabe zeigt, daß durchaus Gleichungen mit vielen Veränderlichen in der Praxis auftauchen.

Ein Konzern besitzt n Fabriken $F_1, \ldots, F_n$, in der Fabrik F_i wird das Produkt P_i hergestellt.

Zur Produktion einer Einheit von P_k werden a_{jk} Einheiten von P_j benötigt; wir nehmen an $a_{ii} = 0$.

Am Ende eines Produktionszyklus sind x_k Einheiten von P_k hergestellt, $k = 1, \ldots, n$; wir haben also einen Produktionsvektor $x = (x_1, \ldots, x_n)$.

Zur Herstellung von $x = (x_1, \ldots, x_n)$ werden

$$\sum_{k=1}^{n} a_{jk}x_k = a_{j1}x_1 + \cdots + a_{jn}x_n$$

Einheiten von P_j verbraucht.

Für den Markt verbleiben damit

$$y_j = x_j - \sum_{k=1}^{n} a_{jk}x_k$$

Einheiten von P_j.

Die Planungsaufgabe lautet nun:

Der Mehrbedarf $y = (y_1, \ldots, y_n)$ ist vorgegeben. Gesucht ist ein Produktionsvektor $x = (x_1, \ldots, x_n)$ mit

$$\begin{array}{rcl} x_1 - (a_{11}x_1 + \cdots + a_{1n}x_n) & = & y_1 \\ \vdots & & \vdots \\ x_n - (a_{n1}x_1 + \cdots + a_{nn}x_n) & = & y_n. \end{array}$$

Also ist ein lineares Gleichungssystem zu lösen. Allerdings, und das macht das Problem schwerer, ist zu beachten, daß alle $x_i \geq 0$ sein müssen (natürlich sind auch die y_j und die $a_{j_k} \geq 0$).

(Das Modell heißt Leontieff-Modell und ist nach Vassili Leontieff benannt, der 1973 den Nobelpreis für Wirtschaftswissenschaften erhielt.)

Ein einfaches Beispiel mit zwei Fabriken, Verbrauchsmatrix

$$\begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{3} & 0 \end{pmatrix}$$

und zunächst unbestimmtem Mehrbedarf (y_1, y_2) liefert das Gleichungssystem

$$\left(\begin{array}{cc|c} 1 & -\frac{1}{2} & y_1 \\ -\frac{1}{3} & 1 & y_2 \end{array}\right) \mapsto \left(\begin{array}{cc|c} 1 & -\frac{1}{2} & y_1 \\ 0 & \frac{5}{6} & \frac{1}{3}y_1 + y_2 \end{array}\right) \mapsto \left(\begin{array}{cc|c} 1 & -\frac{1}{2} & y_1 \\ 0 & 1 & \frac{2}{5}y_1 + \frac{6}{5}y_2 \end{array}\right)$$
$$\mapsto \left(\begin{array}{cc|c} 1 & 0 & \frac{6}{5}y_1 + \frac{3}{5}y_2 \\ 0 & 1 & \frac{2}{5}y_1 + \frac{6}{5}y_2 \end{array}\right)$$

Beispiel 1.6
Ein Gleichungssystem besitze die spezielle Lösung $(1, 0, 1)^t$ und das zugehörige homogene System besitze $(1, 1, 1)^t$ als Lösung und habe den Rang zwei. Finde ein Gleichungssystem, das diese Bedingungen erfüllt.

Da die Lösungen Vektoren im $\mathbb{R}^3$ sind, ist es ein System in drei Variablen.

Da der Rang zwei ist, hat die Zeilen-Stufen-Form zwei Zeilen. Da die Lösungsmenge nicht von der Form abhängt, können wir das System in Zeilen-Stufen-Form annehmen:
Problem: Finde Gerade im $\mathbb{R}^3$ die durch $(1, 0, 1)^t$ geht und so, daß die in den Nullpunkt verschobene Gerade durch $(1, 1, 1)^t$ geht.

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1, \\ a_{22}x_2 + a_{23}x_3 &= b_2. \end{aligned}$$

$(1, 0, 1)^t$ ist Lösung:

$$a_{11} + a_{13} = b_1, \tag{1}$$
$$a_{23} = b_2. \tag{2}$$

$(1,1,1)^t$ ist Lösung des homogenen Systems:

$$a_{11} + a_{12} + a_{13} = 0, \tag{3}$$

$$a_{22} + a_{23} = 0. \tag{4}$$

Das zugehörige lineare Gleichungssystem in $a_{11}, a_{12}, , a_{13}, a_{22}, a_{23}, b_1, b_2$ lautet:

$$\begin{array}{c} \\ (1) \\ (3) \\ (4) \\ (2) \end{array} \begin{array}{c} \begin{array}{ccccccc} a_{11} & a_{12} & a_{13} & a_{22} & a_{23} & b_1 & b_2 \end{array} \\ \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & -1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 \end{pmatrix} \end{array} \mapsto \begin{array}{c} \begin{array}{ccccccc} a_{11} & a_{12} & a_{13} & a_{22} & a_{23} & b_1 & b_2 \end{array} \\ \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 \end{pmatrix} \end{array}.$$

Das System hat unendlich viele Lösungen, und da der Rang 2 sein soll, muß $a_{22} \neq 0$ und $a_{23} \neq 0$ sein.

Wir wählen

$$a_{22} = 1 \Rightarrow a_{23} = b_2 = -1,$$
$$a_{12} = 1 \Rightarrow b_1 = -1,$$
$$a_{11} = 1 \Rightarrow a_{13} = -2.$$

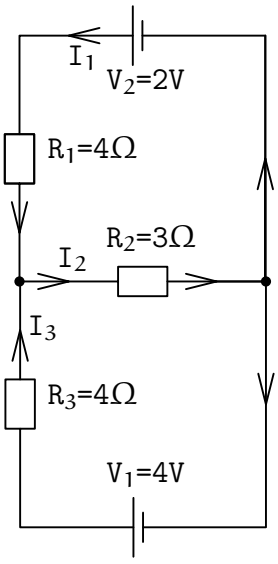
Also ist

$$x_1 + x_2 - 2x_3 = -1,$$
$$x_2 - x_3 = -1$$

ein geeignetes Gleichungssystem.

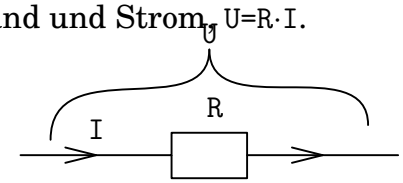
Beispiel 1.7

In einem einfachen elektrischen Netzwerk, wie z. B.

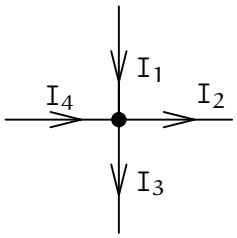


bezeichnet man mit `U` die Spannung, mit `I` den Strom und mit `R` den Widerstand, gemessen in Volt (`V`), Ampere (`A`) und Ohm (Ω) respektive. Dabei gelten folgende Gesetze:

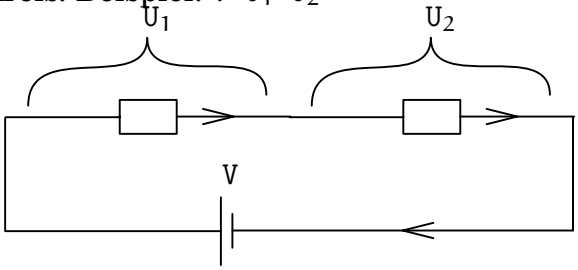
- *Ohmsches Gesetz*: Der Spannungsabfall über einen Widerstand ist das Produkt von Widerstand und Strom, $U=R\cdot I$.



- *1. Kirchhoffsches Gesetz (Knotengleichung):* Die Summe der in einen Knoten hineinfließenden Ströme ist gleich der Summe der hinausfließenden Ströme. Beispiel: $I_1+I_4=I_2+I_3$



- *2. Kirchhoffsches Gesetz (Maschengleichung):* Die Summe der Spannungsverluste in einem geschlossenen Kreis ist gleich der Gesamtspannung in einem Kreis. Beispiel: $V=U_1+U_2$



Im obigen Beispiel stellt man mit Hilfe der drei Gesetze das folgende lineare Gleichungssystem auf:

$$\begin{aligned} I_1 + I_3 &= I_2, && \text{(Knotengleichung)}\\ 4I_1 + 3I_2 &= 2, && \text{(1. Maschengleichung)}\\ 4I_3 + 3I_2 &= 4. && \text{(2. Maschengleichung)} \end{aligned}$$

Wir erhalten das folgende Gleichungssystem:

$$\left(\begin{array}{ccc|c} 1 & -1 & 1 & 0\\ 4 & 3 & 0 & 2\\ 0 & 3 & 4 & 4 \end{array}\right) \mapsto \left(\begin{array}{ccc|c} 1 & -1 & 1 & 0\\ 0 & 7 & -4 & 2\\ 0 & 3 & 4 & 4 \end{array}\right) \mapsto \left(\begin{array}{ccc|c} 1 & -1 & 1 & 0\\ 0 & 7 & -4 & 2\\ 0 & 0 & 40 & 22 \end{array}\right),$$

woraus sich als Lösungen ergeben:

$$I_3 = \tfrac{11}{20}, I_2 = \tfrac{1}{7}\left(\tfrac{44}{20} + \tfrac{40}{20}\right) = \tfrac{3}{5} \text{ und } I_1 = \tfrac{3}{5} - \tfrac{11}{20} = \tfrac{1}{20}.$$

KAPITEL I

Mengen, Zahlen, Gruppen, Körper

Wir beginnen jetzt, grundlegende Begriffe einzuführen und zu besprechen, die für alle mathematischen Disziplinen gleich wichtig sind.

1. Etwas Logik

Die Mathematik verwendet die *axiomatische Methode*, d. h. gewisse Aussagen nennt man *Axiome*. Mit den Regeln der *Logik* werden daraus neue, wahre Aussagen gewonnen. Viele Bemühungen der Mathematik sind darauf gerichtet, in den unterschiedlichen Erscheinungsformen gemeinsame einfache *Strukturen* und Prinzipien zu finden und diese axiomatisch zu fassen. Die Mathematik läßt sich aber nicht auf Logik reduzieren. Mathematik ist wesentlich mehr, als nur aus wahren Aussagen andere wahre Aussagen korrekt zu folgern. Die Mathematik ist eine äußerst kreative Wissenschaft, die ständig neue Strukturen schafft, deren große Bedeutung sich manchmal erst viel später erschließt. Die Mathematik hat ihre gesellschaftliche Relevanz über Jahrtausende bewiesen, und zwar nicht durch korrektes logisches Schließen, sondern durch die Schaffung von *wichtigen* Strukturen. Was wichtig ist, wird nicht durch Logik entschieden, sondern über einen historisch längeren Zeitraum und in einem komplexeren Rückkoppelungsprozeß mit der Realität.

Natürlich ist korrektes logisches Schließen die Grundlage jeder mathematischen Argumentation. Jeder weiß, wie oft in der Umgangssprache etwa die doppelte Verneinung falsch verwendet wird. Das darf in mathematischen Beweisen auf gar keinen Fall passieren. Das korrekte Verneinen sollte deshalb besonders geübt werden.

Einige Begriffe und Notationen, die zum täglichen mathematischen Handwerkszeug gehören, werden jetzt eingeführt.

Definition 1.1

Es seien A und B Aussagen, so lassen sich daraus durch folgende Operationen neue Aussagen gewinnen:

Name	**Symbol**	**Bedeutung**
Konjunktion	$A \wedge B$	"A und B"; sowohl A als auch B
Disjunktion	$A \vee B$	"A oder B" (oder beides); nicht-ausschließendes Oder
Negation	$\neg A$	"nicht A"
Implikation	$A \Rightarrow B$	"aus A folgt B"; "A impliziert B"; in der Bedeutung $(\neg A) \vee B$
Äquivalenz	$A \Leftrightarrow B$	"A ist äquivalent zu B"; " A ist gleichbedeutend zu B"; in der Bedeutung $(A \Rightarrow B) \wedge (B \Rightarrow A)$

Bemerkung 1.2
Man beachte, daß der *Schluß* "aus A folgt B" für jede Aussage B richtig ist, wenn A falsch ist. Das folgt aus der Definition von "⇒". Mit der Wahrheit von B hat die Richtigkeit der *Schlußweise* nichts zu tun!

Beispiel 1.3
Hier nun einige mathematische Aussagen.

A. Jede gerade Zahl ist Summe zweier ungerader Zahlen.
B. Es gibt unendlich viele Primzahlen.
C. Jede gerade Zahl größer zwei ist Summe zweier Primzahlen.
D. Zu jedem Kreis läßt sich, nur mit Zirkel und Lineal, ein Quadrat konstruieren, das den gleichen Flächeninhalt hat.
E. Die Gleichung $x^n + y^n = z^n$ besitzt für $n > 2$ keine Lösung mit positiven ganzen Zahlen x, y, z.
F. Gegeben sei eine Familie nicht-leerer Mengen. Dann läßt sich aus jeder der Mengen ein Element auswählen.

Die Aussage A ist offensichtlich wahr, und auch die Aussage B ist richtig, allerdings ist dies keine triviale Aussage. Sie muß bewiesen werden. Die Aussage C ist die bekannte *Goldbachsche Vermutung* aus dem Jahre 1742. Sie ist bis heute weder bewiesen noch widerlegt.

Die Aussage D ist unter dem Begriff *Quadratur des Kreises* bekannt. Sie ist falsch, was sich daraus ableiten läßt, daß die Kreiszahl π transzendent ist (Lindemann 1882). Umgangssprachlich sollte man also die Quadratur des Kreises nicht als Synonym für etwas extrem Schwieriges verwenden, sondern für etwas Unmögliches.

Die Aussage E hat jahrhundertelang als *Fermatsche Vermutung* die Mathematiker beschäftigt. Sie wurde erst 1995 von dem englischen Mathematiker Wiles als wahr nachgewiesen. Für den Beweis wurden modernste und tiefste mathematische Methoden verwendet.

Die Aussage F, möchte man meinen, ist offensichtlich wahr, eher noch als Aussage A. In gewissem Sinne ist diese Aussage jedoch weder beweisbar noch widerlegbar. Sie ist im Axiomensystem der Mengenlehre von Zermelo und

Fraenkel unabhängig von den anderen Axiomen. In der Tat kann man die Aussage F, die als *Auswahlaxiom* bezeichnet wird, als Axiom der Mengenlehre zulassen (was wir, wie die überwiegende Zahl der Mathematiker, tun wollen) oder auch nicht. Da das Auswahlaxiom, wenn überhaupt, so nur für überabzählbare Mengen strittig ist, sind Zustimmung oder Ablehnung kaum von praktischer Relevanz.

Soweit zu einigen interessanten mathematischen Aussagen. Mit den Mitteln der Logik erhalten wir, daß die nächste Aussage wahr und die übernächste Aussage falsch ist.

Beispiel 1.4 G. Die Aussage A oder die Aussage D ist wahr. $(A \vee D)$

H. Die Aussagen A und D sind wahr. $(A \wedge D)$

Beispiel 1.5

Ein typischer Gebrauch des mathematischen "oder" findet sich bei der Multiplikation von ganzen Zahlen a, b:

$$a \cdot b = 0 \quad \Rightarrow \quad a = 0 \ \vee \ b = 0.$$

Natürlich können beide Zahlen null sein.

Neben Aussagen, die wahr oder falsch sein können, sind *Aussagefunktionen* oder *Prädikate* wichtig, die erst dann wahr oder falsch werden, wenn spezielle Werte eingesetzt werden.

Beispiel 1.6

So ist etwa für ganze Zahlen a und b die Aussage $a > b$ erst dann wahr oder falsch, wenn konkrete Zahlen eingesetzt werden, z. B. $42 > 37$.

Aussagefunktionen werden in der Praxis häufig mit *Quantoren* gebraucht.

Definition 1.7

$\forall$ oder $\vee$: "für alle".

$\exists$ oder $\wedge$: "es gibt".

Ist P eine Aussagefunktion, so bedeutet:

$\forall\, x \,:\, P(x)$: "für alle x gilt $P(x)$",

$\exists\, x \,:\, P(x)$: "es gibt ein x, so daß $P(x)$ gilt".

Beispiel 1.8

$$\forall\, x, \forall\, y, \forall\, z, \forall\, n \,:\, n > 2 \Rightarrow x^n + y^n \neq z^n.$$

Dies ist für positive natürliche Zahlen x, y, z und n die Fermatsche Vermutung.

Bemerkung 1.9

Wichtig ist das richtige Verneinen einer Aussage.

$$\neg\big(\forall\, x \,:\, P(x)\big) \ \Leftrightarrow \ \exists\, x \,:\, \big(\neg P(x)\big).$$

Die Verneinung der Aussage "für alle x gilt die Aussage $P(x)$" ist gleichbedeutend mit "es gibt ein x, für das die Aussage $P(x)$ nicht gilt".

$$\neg\big(\exists\, x \,:\, P(x)\big) \ \Leftrightarrow \ \forall\, x \,:\, \big(\neg P(x)\big).$$

Die Verneinung der Aussage "es gibt ein x, für das die Aussage $P(x)$ gilt" ist gleichbedeutend mit "für alle x gilt die Aussage $P(x)$ nicht" bzw. mit "für kein x gilt die Aussage $P(x)$".

$$(A \Rightarrow B) \;\Leftrightarrow\; (\neg B \Rightarrow \neg A).$$

Die Aussage "aus A folgt B" ist gleichbedeutend mit "aus nicht B folgt nicht A". Letzteres bezeichnet man auch als *Kontraposition* von ersterem.

Notation 1.10
Als Notation haben sich "," sowie "und" anstelle von "$\wedge$" eingebürgert, und "oder" statt "$\vee$" sowie "nicht" statt "$\neg$".

AUFGABEN

Aufgabe 1.11 a. Drücke die folgenden Aussagen in Worten aus und, falls eine Aussage falsch sein sollte, ersetze sie dabei durch ihre Negation.

(i) $\forall\, m \in \mathbb{N}\ \exists\, n \in \mathbb{N}\ : m = n + n$,

(ii) $\forall\, m \in \mathbb{N},\ \forall\, n \in \mathbb{N}\ :\ m \geq n \implies \exists\, l \in \mathbb{N}\ :\ m = n + l$,

(iii) $\exists\, m \in \mathbb{N},\ \exists\, n \in \mathbb{N}\ : (m \neq n) \wedge (m^n = n^m)$.

b. Drücke die folgende Aussage in Symbolen aus: Zwischen je zwei verschiedenen reellen Zahlen gibt es eine weitere reelle Zahl.

Aufgabe 1.12
Negiere die folgenden Aussagen:

a. Jedes Auto, das am Samstag um neun Uhr auf dem Parkplatz parkte, war rot.
b. Mindestens ein Auto, das am Samstag um neun Uhr auf dem Parkplatz parkte, war rot.
c. Am Samstag um neun Uhr parkten rote Autos auf dem Parkplatz.
d. Jede ganze Zahl ist ein Vielfaches von drei.
e. Es gibt keine größte ganze Zahl.
f. Die Summe von je zwei ungeraden Zahlen ist gerade.

Aufgabe 1.13
Beweise oder widerlege die Aussagen d.-f. in Aufgabe 1.12

Aufgabe 1.14
Welche der folgenden Schlußfolgerungen ist korrekt?

a. Falls es anfängt zu regnen, wird die Straße naß. Aber, da die Straße nicht naß werden wird, wird es auch nicht regnen.
b. Falls die Bundesrepublik aus der EU ausscheidet oder falls das Handelsdefizit reduziert wird, wird der Butterpreis fallen. Das Handelsdefizit wird steigen, sofern die Exporte nicht erhöht werden. Also wird der Butterpreis nicht fallen.
c. Einige Politiker sind ehrlich. Einige Frauen sind Politiker. Also sind einige weibliche Politiker ehrlich.
d. Wenn ich nicht hart arbeite, werde ich einschlafen. Ich befürchte, ich werde nicht einschlafen. Also befürchte ich, daß ich hart arbeiten werde.

Aufgabe 1.15
Diskutiere die folgende Aussage und ihren *Beweis*:
"Satz": Für je zwei reelle Zahlen $x, y \in \mathbb{R}$ gilt:

$$\frac{1}{2}(x+y) \geq \sqrt{xy}. \tag{3}$$

Beweis: Man quadriere die Gleichung (3) und multpliziere das Ergebnis mit vier, um

$$x^2 + 2xy + y^2 \geq 4xy$$

zu erhalten. Zieht man auf beiden Seiten $4xy$ ab, so erhält man

$$(x-y)^2 = x^2 - 2xy + y^2 \geq 0.$$

Letzteres ist immer erfüllt, also ist die Aussage bewiesen. □

2. Abbildungen und Mengen

Der folgende "naive" Mengenbegriff des deutschen Mathematikers Cantor (1845-1918) ist praktisch für alle Zwecke der Mathematik ausreichend. Danach ist eine *Menge* eine Zusammenfassung von bestimmten, wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens. Die Objekte heißen *Elemente* der Menge.

Wir führen nun einige wichtige Symbole und Konstruktionen im Zusammenhang mit Mengen ein.
Definition 2.1
Es seien M, N, I, M_i ($i \in I$) Mengen, P eine Aussagefunktion.

$\{x_1, \ldots, x_n\}$	:	Menge aus den (verschiedenen) Elementen $x_1, \ldots, x_n$
	:	z. B. $\{1,1\} = \{1\}$, $\{1,2,3\} = \{3,1,2\}$;
$x \in M$	:	x ist *Element* der Menge M;
$x \notin M$	:	x ist *nicht Element* der Menge M;
$\{x \in M \mid P(x)\}$	:	Menge aller Elemente $x \in M$, für die die Aussage $P(x)$ gilt;
$\emptyset$ oder $\{\,\}$	:	*leere Menge*, die Menge, die keine Elemente enthält;
$M \subset N$ oder $M \subseteq N$	:	M ist *Teilmenge* von N, d. h. jedes Element von M ist auch Element von N, d. h. $x \in M \Rightarrow x \in N$;
$M = N$	:	$M \subseteq N$ und $N \subseteq M$;
$M \neq N$	:	$\neg(M = N)$;
$M \subsetneq N$	:	$M \subseteq N$ und $M \neq N$;
$M \cap N$	:	*Durchschnitt* der Mengen M und N, d. h. $M \cap N = \{x \mid x \in M \wedge x \in N\}$;
$\bigcap_{i \in I} M_i$	:	Durchschnitt aller Mengen M_i mit $i \in I$, wobei I als *Indexmenge* bezeichnet wird, d. h. $\bigcap_{i \in I} M_i = \{x \mid \forall\, i \in I : x \in M_i\} = \{x \mid x \in M_i \;\forall\, i \in I\}$;

$M \cup N$: *Vereinigung* der Mengen M und N, d. h. $M \cup N = \{x \mid x \in M \vee x \in N\}$;

$\bigcup_{i \in I} M_i$: Vereinigung aller Mengen M_i mit $i \in I$, d. h. $\bigcup_{i \in I} M_i = \{x \mid \exists\, i \in I : x \in M_i\}$;

$M \setminus N$: *Differenz* von M und N, d. h. $M \setminus N = \{x \in M \mid x \notin N\}$;

$M \times N$: *kartesisches Produkt* von M und N, Menge aller (geordneten) Paare, d. h. $M \times N = \{(m,n) \mid m \in M \wedge n \in N\}$;

$\prod_{i \in I} M_i$: kartesisches Produkt aller Mengen M_i mit $i \in I$, d. h. $\prod_{i \in I} M_i = \left\{(x_i)_{i \in I} \mid x_i \in M_i \;\forall\, i \in I\right\}$;

$\mathcal{P}(M)$: *Potenzmenge* von M, Menge aller Teilmengen von M, d. h. $\mathcal{P}(M) = \{N \mid N \subseteq M\}$.

Führen wir nun noch einige spezielle Mengen ein:

$\mathbb{N} := \{0,1,2,3,\ldots\}$: die Menge der *natürlichen Zahlen*;

$\mathbb{Z} := \{0, \pm 1, \pm 2, \ldots\}$: die Menge der *ganzen Zahlen*;

$\mathbb{Q} := \left\{\frac{p}{q} \mid p, q \in \mathbb{Z}, q \neq 0\right\}$: die Menge der *rationalen Zahlen*;

$\mathbb{R}$: die Menge der *reellen Zahlen* - diese lassen sich durch endliche oder unendliche Dezimalbrüche darstellen;

$\mathbb{R}_{>0}$ bzw. $\mathbb{R}_{<0}$: die Menge der positiven bzw. negativen reellen Zahlen.

Hier und im Folgenden verwenden wir die folgenden Symbole:

$:=$: "per definitionem gleich", d. h. die linke Seite wird durch die rechte Seite definiert;

$:\Leftrightarrow$: "per definitionem äquivalent", d. h. die linke Seite gilt definitionsgemäß genau dann, wenn die rechte Seite gilt.

Bemerkung 2.2
In Singular ist "=" das, was mathematisch gesehen ":=" ist, nämlich der Zuweisungsoperator. Der Vergleichsoperator "=" ist in Singular hingegen "==".

In Definitionen werden wir häufig statt ":⇔" etwas unexakt "falls" verwenden, siehe etwa Definition 2.7.

Beispiel 2.3 a. $\mathbb{N} \subsetneq \mathbb{Z} \subsetneq \mathbb{Q} \subsetneq \mathbb{R}$.

Die Inklusionen sind klar. Daß die Mengen nicht gleich sind, zeigt man dadurch, daß man ein Element der größeren Menge angibt, das nicht in der kleineren enthalten ist.

$$-1 \in \mathbb{Z},\ -1 \notin \mathbb{N};\ \ \frac{1}{2} \in \mathbb{Q},\ \frac{1}{2} \notin \mathbb{Z};\ \ \sqrt{2} \in \mathbb{R},\ \sqrt{2} \notin \mathbb{Q}.$$

b. Sei für $i \in \mathbb{N}$ die Menge $M_i := [-i, i] := \{x \in \mathbb{R} \mid -i \leq x \leq i\}$. Dann gilt:

$$\bigcap_{i \in \mathbb{N}} M_i = \{0\};\ \ \bigcup_{i \in \mathbb{N}} M_i = \mathbb{R}.$$

c. $\mathbb{R}\times \overset{n}{\cdots} \times\mathbb{R} := \prod_{i=1}^{n} \mathbb{R} := \prod_{i\in\{1,\ldots,n\}} \mathbb{R} = \{(x_1,\ldots,x_n) \mid x_i \in \mathbb{R}\}$.

Definition 2.4
Seien M und N Mengen.

a. Eine *Relation* zwischen M und N ist eine Teilmenge $\Gamma \subseteq M \times N$.
b. Sei $\Gamma \subseteq M \times N$ eine Relation. Das Tripel $f = (M, N, \Gamma)$ heißt *Abbildung* von M in N, falls gilt:
   (i) f ist *linksvollständig*, d. h. $\forall\, x \in M\ \exists\, y \in N\ :\ (x,y) \in \Gamma$, und
   (ii) f ist *rechtseindeutig*, d. h. $\forall\ (x,y) \in \Gamma\ \forall\ (x',y') \in \Gamma$ gilt: $x = x' \ \Rightarrow\ y = y'$.

   Statt $f = (M, N, \Gamma)$ schreibt man gemeinhin auch $f : M \to N$, und statt $(x,y) \in \Gamma$ schreibt man $y = f(x)$ oder $x \mapsto y$.
   Die Menge $\Gamma_f := \Gamma = \{(x,y) \in M \times N \mid y = f(x)\}$ heißt der *Graph* der Abbildung f.
   Wir bezeichnen mit

$$N^M := \{f : M \to N \mid f \text{ ist Abbildung}\}$$

   die Menge der Abbildungen von M nach N.

Bemerkung 2.5
Eine Abbildung $f : M \to N$ besteht also aus drei Daten, dem *Definitionsbereich* M, dem *Wertebereich* N und der *Abbildungsvorschrift*, die *jedem* $x \in M$ *genau ein* $y = f(x) \in N$ zuordnet. Man beachte, daß nicht gefordert wird, daß $f(x)$ in irgendeiner Form aus x (mittels einer universellen Formel) berechenbar sein muß. Mit den Mitteln der Logik läßt sich beweisen, daß es nicht berechenbare Abbildungen gibt.

Für Abbildungen, die auf dem Computer dargestellt werden sollen, kommen natürlich nur berechenbare Abbildungen in Frage. Mehr noch, man braucht einen Algorithmus, der aus gegebenem x den Wert $f(x)$ in endlich vielen Schritten berechnet.

Statt des Begriffs rechtseindeutig verwendet man häufig auch den Begriff *wohldefiniert*.

Man beachte ferner, daß für zwei Abbildungen $f, g : M \to N$ genau dann gilt $f = g$, wenn für alle $x \in M$ gilt $f(x) = g(x)$.

Definition 2.6 a. Es sei M eine Menge. Die Abbildung $\mathrm{id}_M : M \to M : x \mapsto x$ heißt die *Identität* oder *identische Abbildung* auf M.
   Wir schreiben häufig kurz id statt id_M, wenn keine Unklarheiten zu befürchten sind.
b. Ist $N \subseteq M$ eine Teilmenge, so nennen wir $i_{N,M} : N \to M : x \mapsto x$ die (kanonische) *Inklusion* von N in M.
   Wir schreiben manchmal auch i_N oder i statt $i_{N,M}$, sofern keine Mißverständnisse auftreten können.

Definition 2.7
Es sei $f : M \to N$ eine Abbildung, $A \subseteq M$, $B \subseteq N$.

a. $f(A) := \{y \in N \mid \exists\, x \in A \;:\; y = f(x)\}$ heißt das *Bild* von A unter der Abbildung f.
b. $f^{-1}(B) := \{x \in M \mid f(x) \in B\}$ heißt *Urbild* von B unter f.
Ist $B = \{y\}$ für ein $y \in N$, so schreiben wir auch $f^{-1}(y)$ statt $f^{-1}(B)$.
c. Die Abbildung $f_{|A} : A \to N : x \mapsto f(x)$ heißt *Einschränkung* von f auf A.
Es gilt offenbar $\Gamma_{f_{|A}} = \Gamma_f \cap (A \times N)$.
d. f heißt *injektiv*, falls gilt:

$$\forall\, x, x' \in M \;:\; f(x) = f(x') \;\Rightarrow\; x = x',$$

d. h. zwei verschiedene Elemente von M können durch f nicht auf dasselbe Element in N abgebildet werden.
e. f heißt *surjektiv*, falls gilt:

$$\forall\, y \in N\ \exists\, x \in M \;:\; y = f(x),$$

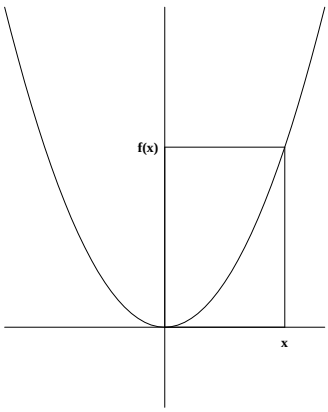
d. h. $f(M) = N$, d. h. jedes Element von N kommt als Bild unter f vor.
f. f heißt *bijektiv*, falls f injektiv und surjektiv ist.

Bemerkung 2.8
Ist $f : M \to N$ eine Abbildung, $A \subseteq M$ und $B \subseteq N$ mit $f(A) \subseteq B$, dann bezeichnen wir hin und wieder auch die Abbildung

$$A \to B : x \mapsto f(x)$$

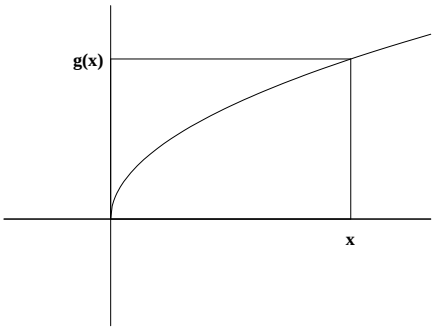
mit $f_{|A}$ und als Einschränkung von f auf A. Das ist zwar etwas unsauber, wird aber in den konkreten Fällen nicht zu Zweideutigkeiten führen.

Beispiel 2.9 a. Sei $f : \mathbb{R} \to \mathbb{R}$ die Abbildung, die durch die Vorschrift $f(x) = x^2$ gegeben ist. Der Graph $\Gamma_f = \left\{(x, y) \in \mathbb{R}^2 \;\middle|\; y = x^2\right\}$ ist die Normalparabel.



f ist weder surjektiv (da etwa $-1 \notin f(\mathbb{R})$) noch injektiv (da z. B. $f(-1) = 1 = f(1)$).

b. $g : \mathbb{R}_{\geq 0} \to \mathbb{R} : x \mapsto \sqrt{x}$ ist eine Abbildung, die injektiv ist (da für $x, x' \in \mathbb{R}_{\geq 0}$ aus $\sqrt{x} = \sqrt{x'}$ folgt, daß $x = x'$), aber nicht surjektiv (da $f\big(\mathbb{R}_{\geq 0}\big) = \mathbb{R}_{\geq 0} \neq \mathbb{R}$).



c. $h : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0} : x \mapsto \sqrt{x}$ unterscheidet sich von g nur durch den Wertebereich. Aber dies reicht, daß h bijektiv ist.

Definition 2.10
Es seien I und M Mengen

a. Eine *Familie* von Elementen in M mit *Indexmenge* I ist eine Abbildung $F : I \to M$.
Für $i \in I$ setze $x_i := F(i) \in M$. Dann schreibt man statt $F : I \to M$ auch $(x_i)_{i\in I}$ (oder kurz (x_i), falls über I kein Zweifel besteht) und nennt dann $(x_i)_{i\in I}$ eine Familie von Elementen in M mit Indexmenge I.
b. Ist $F : I \to M$ eine Abbildung und $J \subseteq I$, so heißt die Einschränkung $F_{|J}$ von F auf J auch eine *Teilfamilie* und wird gemeinhin auch mit $(x_i)_{i\in J}$ bezeichnet.

Bemerkung 2.11
Beachte, daß in der Familie $F = (x_i)_{i\in I}$ für $i, j \in I$ mit $i \neq j$ sehr wohl $x_i = x_j$ gelten kann, während dies in der Menge $\{F\} := F(I) = \{x_i \mid i \in I\}$ nicht der Fall ist.

Wir schreiben meist kurz $x \in F$, wenn wir $x \in F(I)$ meinen.

Beispiel 2.12 a. Für $J = \emptyset$ spricht man von der *leeren Familie*.

b. Die Familien in M mit Indexmenge $I = \{1, \ldots, n\}$ werden mittels der Schreibweise in Definition 2.10 a. mit den Elementen des n-fachen kartesischen Produktes $M \times \overset{n}{\cdots} \times M$ identifiziert, d. h. eine Familie $(x_i)_{i\in I} = (x_1, \ldots, x_n)$ ist das Gleiche wie ein n-Tupel.
c. Eine Familie mit $I = \mathbb{N}$ nennt man eine *Folge*. Somit ist

$$M^{\mathbb{N}} = \{F : \mathbb{N} \to M \mid F \text{ ist Abbildung}\} = \{(x_i)_{i\in\mathbb{N}} \mid x_i \in M\}$$

die Menge aller Folgen in M.
d. Jede Teilmenge $N \subseteq M$ ist eine Familie mittels der kanonischen Inklusion i_N.
e. Ist $M = \{M_i \mid i \in I\}$ und $F : I \to M : i \mapsto M_i$, so heißt $F = (M_i)_{i\in I}$ auch eine *Familie von Mengen*.
In Definition 2.1 haben wir - ohne dies zu erwähnen - bereits Familien von Mengen benutzt und den Schnitt, die Vereinigung sowie das kartesische Produkt von beliebigen Familien von Mengen definiert!

Definition 2.13
Sind $f_1 : M_1 \to M_2$ und $f_2 : M_2 \to M_3$ Abbildungen, so heißt die Abbildung $f_2 \circ f_1 : M_1 \to M_3 : x \mapsto f_2\big(f_1(x)\big)$ die *Komposition* von f_1 und f_2.

Lemma 2.14
Die Komposition ist assoziativ, d. h. sind $f_1 : M_1 \to M_2$, $f_2 : M_2 \to M_3$ *und* $f_3 : M_3 \to M_4$ *Abbildungen, so gilt:*

$$(f_3 \circ f_2) \circ f_1 = f_3 \circ (f_2 \circ f_1).$$

Wir schreiben für die beiden Ausdrücke deshalb auch vereinfacht $f_3 \circ f_2 \circ f_1$.

Beweis: Nach Definition der Komposition gilt für alle $x \in M_1$:

$$\big((f_3 \circ f_2) \circ f_1\big)(x) = (f_3 \circ f_2)\big(f_1(x)\big) = f_3\Big(f_2\big(f_1(x)\big)\Big)$$

$$= f_3\big((f_2 \circ f_1)(x)\big) = \big(f_3 \circ (f_2 \circ f_1)\big)(x).$$

□

Bemerkung 2.15
Man darf die Reihenfolge der Komposition nicht vertauschen! Betrachte etwa:

$$f : \mathbb{R} \to \mathbb{R} : x \mapsto x + 1, \quad g : \mathbb{R} \to \mathbb{R} : x \mapsto x^2.$$

Dann gilt:

$$(f \circ g)(x) = f\big(x^2\big) = x^2 + 1, \quad (g \circ f)(x) = g(x + 1) = (x + 1)^2.$$

Damit ist $f \circ g \neq g \circ f$, da etwa $(f \circ g)(1) = 2 \neq 4 = (g \circ f)(1)$.

Wir führen nun die folgenden Notationen ein.

Definition 2.16
Zwei Mengen M und N heißen *gleichmächtig*, falls es eine bijektive Abbildung $f : M \to N$ gibt. Mit

$$\#M := |M| := \begin{cases} \text{Anzahl der Elemente in } M, \text{ falls } M \text{ endlich ist,} \\ \infty, \text{ falls } M \text{ unendlich viele Elemente enthält,} \end{cases}$$

bezeichnen wir die *Mächtigkeit* der Menge M.[1]

Lemma 2.17
Es seien M *und* N *zwei endliche Mengen.*

a. *Genau dann gilt* $|M| \leq |N|$*, wenn es eine injektive Abbildung* $f : M \to N$ *gibt.*
b. *Genau dann gilt* $|M| \geq |N|$*, wenn es eine surjektive Abbildung* $f : M \to N$ *gibt.*
c. *Genau dann gilt* $|M| = |N|$*, wenn es eine bijektive Abbildung* $f : M \to N$ *gibt.*

[1] Auch für unendliche Mengen gibt es unterschiedliche Mächtigkeiten, sog. *Kardinalzahlen*, auf die wir hier aber nicht eingehen wollen.

Beweis: Es seien $M = \{x_1, \ldots, x_m\}$ und $N = \{y_1, \ldots, y_n\}$ mit paarweise verschiedenen Elementen $x_i \neq x_j$ für $i \neq j$ und $y_i \neq y_j$ für $i \neq j$. Es gilt $|M| = m$ und $|N| = n$.

a. Ist $m \leq n$, so definiere $f : M \to N$ durch $f(x_i) = y_i$ für $i = 1, \ldots, m$. Dann gilt für $i, j \in \{1, \ldots, m\}$ mit $i \neq j$

$$f(x_i) = y_i \neq y_j = f(x_j).$$

Mithin ist f injektiv.
Ist umgekehrt $f : M \to N$ eine injektive Abbildung, so gilt $f(M) = \{f(x_1), \ldots, f(x_m)\} \subseteq N$ eine Teilmenge von paarweise verschiedenen Elementen. Mithin enthält N mindestens m Elemente, und folglich gilt $m \leq n$.

b. Ist $m \geq n$, so definiere $f : M \to N$ durch $f(x_i) = y_i$ für $i = 1, \ldots, n$ und $f(x_i) = y_1$ für $i = n+1, \ldots, m$. Dann gilt offenbar $f(M) = \{y_1, \ldots, y_n\} = N$ und f ist surjektiv.
Ist umgekehrt $f : M \to N$ eine surjektive Abbildung, so gilt $\{y_1, \ldots, y_n\} = N = f(M) = \{f(x_1), \ldots, f(x_m)\}$. Mithin enthält die Menge $\{f(x_1), \ldots, f(x_m)\}$ n verschiedene Elemente, und folglich ist $m \geq n$.

c. Die Aussage folgt unmittelbar aus den ersten beiden Teilen.

□

Bemerkung 2.18
Sind M und N endliche Mengen, so folgt aus $M \subsetneq N$ mittels Lemma 2.17 unmittelbar $|M| < |N|$ und M und N sind nicht gleichmächtig.

Dies gilt für unendliche Mengen nicht mehr, wie das Beispiel $\mathbb{N} \subsetneq \mathbb{Z}$ zeigt. Denn die Abbildung

$$f : \mathbb{Z} \to \mathbb{N} : k \mapsto \begin{cases} 2k, & \text{für } k \geq 0, \\ -2k-1 & \text{für } k < 0, \end{cases}$$

ist bijektiv, wie man sich leicht überzeugt. Also sind $\mathbb{N}$ und $\mathbb{Z}$ gleichmächtig.

Lemma 2.19
Seien M und N zwei nicht-leere Mengen, $f : M \to N$ eine Abbildung.

a. *f ist genau dann injektiv, wenn es eine Abbildung $g : N \to M$ gibt mit $g \circ f = \mathrm{id}_M$.*
b. *f ist genau dann surjektiv, wenn es eine Abbildung $g : N \to M$ gibt mit $f \circ g = \mathrm{id}_N$.*
c. *f ist genau dann bijektiv, wenn es eine Abbildung $g : N \to M$ gibt mit $g \circ f = \mathrm{id}_M$ und $f \circ g = \mathrm{id}_N$.*
d. *Ist f bijektiv, so ist die nach c. existierende Abbildung g eindeutig bestimmt und ebenfalls bijektiv. Sie heißt die* Inverse *oder* Umkehrabbildung *von f und wird mit f^{-1} bezeichnet.*

Beweis: a. "$\Rightarrow$": Es sei f injektiv. Dann gilt für $y \in f(M)$, daß $\left|f^{-1}(y)\right| = 1$, also $f^{-1}(y) = \{x_y\}$ für ein geeignetes $x_y \in M$ und $f(x_y) = y$. Hingegen ist $f^{-1}(y) = \emptyset$ für $y \notin f(M)$. Wähle ein $x_0 \in M \neq \emptyset$ fest und definiere eine Abbildung

$$g : N \to M : y \mapsto \begin{cases} x_y, & \text{falls } y \in f(M), \\ x_0, & \text{falls } y \in N \setminus f(M). \end{cases}$$

Dann gilt für $x \in M$:

$$(g \circ f)(x) = g\big(f(x)\big) = x_{f(x)} = x = \mathrm{id}_M(x).$$

Da $x \in M$ beliebig gewählt war, folgt also $g \circ f = \mathrm{id}_M$.
"$\Leftarrow$": Es sei nun $g : N \to M$ mit $g \circ f = \mathrm{id}_M$ gegeben. Seien ferner $x, x' \in M$ mit $f(x) = f(x')$, dann gilt:

$$x = \mathrm{id}_M(x) = (g \circ f)(x) = g\big(f(x)\big) = g\Big(f\big(x'\big)\Big) = (g \circ f)\big(x'\big) = \mathrm{id}_M\left(x'\right) = x'.$$

Also ist f injektiv.

b. "$\Rightarrow$": Es sei f surjektiv. Dann können wir zu jedem $y \in N = f(M)$ ein $x_y \in M$ wählen mit $f(x_y) = y$. Definiere eine Abbildung

$$g : N \to M : y \mapsto x_y.$$

Dann gilt für $y \in N$:

$$(f \circ g)(y) = f\big(g(y)\big) = f(x_y) = y = \mathrm{id}_N(y).$$

Da $y \in N$ beliebig gewählt war, folgt also $f \circ g = \mathrm{id}_N$.
"$\Leftarrow$": Es sei nun $g : N \to M$ mit $f \circ g = \mathrm{id}_N$ gegeben. Für $y \in N$ definiere $x := g(y) \in M$. Dann gilt:

$$y = \mathrm{id}_N(y) = (f \circ g)(y) = f\big(g(y)\big) = f(x) \in f(M).$$

Also ist f surjektiv.

c. "$\Rightarrow$": Ist f bijektiv, so gilt für jedes $y \in N$, daß $\left|f^{-1}(y)\right| = 1$ und die Definitionen der Abbildungen g in den beiden obigen Teilen stimmen überein, so daß wir eine einzige Abbildung $g : N \to M$ erhalten mit:

$$g \circ f = \mathrm{id}_M \quad \text{und} \quad f \circ g = \mathrm{id}_N\,.$$

"$\Leftarrow$": Dies folgt unmittelbar aus den obigen beiden Teilen.

d. Die Bijektivität von g folgt aus dem in c. bewiesenen Kriterium für Bijektivität. Mithin bleibt die Eindeutigkeit von g zu zeigen, unter der Voraussetzung. Angenommen, $h : N \to M$ sei eine weitere Abbildung mit

$$h \circ f = \mathrm{id}_M \quad \text{und} \quad f \circ h = \mathrm{id}_N\,.$$

Für $y \in N$ beliebig gilt dann:

$$f\big(g(y)\big) = (f \circ g)(y) = \mathrm{id}_N(y) = (f \circ h)(y) = f\big(h(y)\big).$$

Da aber f injektiv ist, folgt damit $g(y) = h(y)$ und schließlich und $g = h$.

□

Bemerkung 2.20
Man beachte, daß die Umkehrabbildung $f^{-1} : N \to M$ nur für eine bijektive Abbildung $f : M \to N$ erklärt ist, daß aber für eine beliebige Abbildung $h : M \to N$ und eine beliebige Teilmenge $B \subseteq N$ das Urbild $h^{-1}(B)$ definiert ist.

Für ein bijektives f stimmen beide Notationen überein, das heißt das Urbild $f^{-1}(B)$ von $B \subseteq N$ unter f ist gleich dem Bild $f^{-1}(B)$ von $B \subseteq N$ unter f^{-1}.

Ist f nicht bijektiv, so ist zwar weiterhin für jedes $y \in N$ das Urbild $f^{-1}(y)$ erklärt, aber die Relation $\{(y, x) \in N \times M \mid x \in f^{-1}(y)\}$ ist keine Abbildung, da sowohl $f^{-1}(y) = \emptyset$ (falls f nicht surjektiv ist) als auch $\left|f^{-1}(y)\right| > 1$ (falls f nicht injektiv ist) möglich ist. In ersterem Fall ist die Linksvollständigkeit verletzt, in letzterem Fall die Rechtseindeutigkeit.

Beispiel 2.21 a. Ist M eine Menge, so ist id_M bijektiv, da offenbar $id_M = id_M \circ id_M$.

b. Die Abbildung $f : \mathbb{Z} \to \mathbb{Z} : x \mapsto 2x$ ist injektiv, da für $x, y \in \mathbb{Z}$ aus $2x = 2y$ unmittelbar $x = y$ folgt. f ist aber nicht surjektiv, da etwa die Zahl 1 kein Urbild besitzt.

c. Im Gegensatz zu b. ist die Abbildung $g : \mathbb{Q} \to \mathbb{Q} : x \mapsto 2x$ sowohl injektiv, als auch surjektiv. Für letzteres beachte man, daß für eine rationale Zahl $y \in \mathbb{Q}$ die rationale Zahl $\frac{y}{2} \in \mathbb{Q}$ ein Urbild von y unter g ist.

Wir kommen noch einmal auf Relationen zurück. Wir hatten schon Abbildungen als Relationen mit besonderen Eigenschaften definiert. Andere wichtige Relationen haben auch einen speziellen Namen.

Definition 2.22
Es sei M ein Menge. Eine *Ordnungsrelation* auf M, auch *Halbordnung* oder *partielle Ordnung* genannt, ist eine Relation $R \subseteq M \times M$, so daß für alle $x, y, z \in M$ gilt:

a. $(x, x) \in R$, ("Reflexivität")
b. $(x, y), (y, x) \in R \Rightarrow x = y$, ("Antisymmetrie")
c. $(x, y), (y, z) \in R \Rightarrow (x, z) \in R$. ("Transitivität")

Notation 2.23
Es sei M eine Menge und R ein Ordnungsrelation auf M. Wir definieren für $x, y \in M$

$$x \leq y \ :\Leftrightarrow \ (x, y) \in R,$$

und sprechen hin und wieder auch von der Ordnungsrelation "$\leq$" statt R, sofern keine Mißverständnisse zu befürchten sind. Ferner sprechen wir von der *partiell* oder *(teil-)geordneten Menge* $(M, \leq)$.

Mit dieser Schreibweise lassen sich die drei Axiome in Definition 2.22 wie folgt formulieren. Für $x, y, z \in M$ soll gelten:

a. $x \leq x$, ("Reflexivität")
b. $x \leq y, y \leq x \Rightarrow x = y$, ("Antisymmetrie")
c. $x \leq y, y \leq z \Rightarrow x \leq z$. ("Transitivität")

Gilt für $x, y \in M$, daß $x \leq y$ und $x \neq y$, so schreiben wir auch $x < y$.

Definition 2.24

Es sei M ein Menge.

a. Eine Ordnungsrelation "$\leq$" heißt *Totalordnung* oder *lineare Ordnung*, falls je zwei Elemente aus M vergleichbar sind, d. h. für je zwei Elemente $x, y \in M$ gilt $x \leq y$ oder $y \leq x$.
b. Ist "$\leq$" eine Ordnungsrelation auf M, $A \subseteq M$ und $x \in A$, so heißt x *minimal* (bzw. *maximal*) in A, falls für alle $y \in A$ mit $y \leq x$ (bzw. $x \leq y$) gilt $x = y$.
c. Eine Totalordnung heißt *Wohlordnung*, falls jede nicht-leere Teilmenge von M ein minimales Element besitzt.

Beispiel 2.25

Die reellen Zahlen $(\mathbb{R}, \leq)$ mit der üblichen Kleiner-Gleich-Relation $\leq$ sind total geordnet, aber nicht wohlgeordnet.

Gleiches trifft auf $(\mathbb{Z}, \leq)$ mit der üblichen Kleiner-Gleich-Relation

$$\ldots -2 < -1 < 0 < 1 < 2 < \ldots$$

zu. Allerdings definiert die "unübliche" Anordnung

$$0 < -1 < 1 < -2 < 2 < -3 < 3 < \ldots$$

in der Tat ein Wohlordnung auf $\mathbb{Z}$.

Die natürlichen Zahlen $(\mathbb{N}, \leq)$ sind bereits mit der üblichen Kleiner-Gleich-Relation wohlgeordnet.

Beispiel 2.26

Ist M eine Menge, so ist die Potenzmenge $\mathcal{P}(M)$ von M durch

$$A \leq B \ :\Leftrightarrow\ A \subseteq B, \text{ für } A, B \in \mathcal{P}(M),$$

partiell geordnet, aber im allgemeinen nicht total geordnet. Z. B. sind im Fall $M = \mathbb{N}$ die Elemente $\{2\}$ und $\{3\}$ in $\mathcal{P}(\mathbb{N})$ nicht vergleichbar.

Allgemeiner gilt, ist N eine Menge, deren Elemente wieder Mengen sind, so wird N mit der analogen Definition von "$\leq$" eine partiell geordnete Menge.

Der Begriff der Ordnungsrelation bereitet in der Regel keine großen begrifflichen Schwierigkeiten. Wohl auch deshalb, da im täglichen Leben alles mögliche verglichen wird - seien es Größen, Entfernungen oder Geschwindigkeiten.

Bei dem folgenden Begriff der Äquivalenzrelation ist das ganz anders. Er bereitet den Studierenden oft extreme Schwierigkeiten. Deshalb zunächst der Versuch einer anschaulichen Erläuterung. Die grundlegende Idee besteht darin, komplexe Probleme dadurch zu vereinfachen, daß man ähnliche (man sagt *äquivalente*) Objekte in einem Paket oder Topf (man sagt in einer *Äquivalenzklasse*) zusammenfaßt und dann mit den Paketen oder Töpfen weiterrechnet anstatt mit deren Inhalt, der aus vielen Elementen bestehen kann. Man

kann auch sagen, daß der Begriff *äquivalent* den Begriff *gleich* abschwächt und äquivalente Objekte so behandelt werden, als wenn sie gleich wären.

Definition 2.27
Es sei M eine Menge. Eine *Äquivalenzrelation* auf M ist eine Relation $R \subseteq M \times M$, so daß für alle $x, y, z \in M$ gilt:

a. $(x, x) \in R$, ("Reflexivität")
b. $(x, y) \in R \Rightarrow (y, x) \in R$, ("Symmetrie")
c. $(x, y), (y, z) \in R \Rightarrow (x, z) \in R$. ("Transitivität")

Notation 2.28
Es sei M eine Menge und R ein Äquivalenzrelation auf M. Wir definieren für $x, y \in M$

$$x \sim y :\Leftrightarrow (x, y) \in R,$$

und sprechen hin und wieder auch von der Äquivalenzrelation "$\sim$" statt R, sofern keine Mißverständnisse zu befürchten sind.

Mit dieser Schreibweise lassen sich die drei Axiome in Definition 2.27 wie folgt formulieren. Für $x, y, z \in M$ soll gelten:

a. $x \sim x$, ("Reflexivität")
b. $x \sim y \Rightarrow y \sim x$, ("Symmetrie")
c. $x \sim y, y \sim z \Rightarrow x \sim z$. ("Transitivität")

Definition 2.29
Es sei M eine Menge und $\sim$ eine Äquivalenzrelation auf M. Für $x \in M$ heißt die Menge

$$[x] := \{y \in M \mid y \sim x\}$$

die *Äquivalenzklasse* von x. Jedes $y \in [x]$ heißt ein *Repräsentant* der Klasse $[x]$. Mit

$$M/\sim := \big\{[x] \mid x \in M\big\}$$

bezeichnen wir die Menge der *Äquivalenzklassen modulo der Äquivalenzrelation* $\sim$.

Beispiel 2.30
Ein gutes Beispiel für Äquivalenzklassen, das allen bekannt ist, sind die rationalen Zahlen. Wir wissen, daß ein Bruch auf verschiedene Weisen als $\frac{p}{q}$ dargestellt werden kann, z. B. ist $\frac{1}{2} = \frac{2}{4}$.

Man kann die rationalen Zahlen wie folgt als Äquivalenzklassen von Paaren ganzer Zahlen definieren. Für $(p, q), (p', q') \in M := \mathbb{Z} \times \big(\mathbb{Z} \setminus \{0\}\big)$ definiere

$$(p, q) \sim (p', q') :\Leftrightarrow pq' = p'q.$$

Man prüft leicht nach, daß dadurch auf M eine Äquivalenzrelation definiert wird.

Wir setzen nun $\mathbb{Q} := M/\sim$ und für $(p, q) \in M$ setzen wir $\frac{p}{q} := \big[(p, q)\big]$, d. h. die rationale Zahl $\frac{p}{q}$ ist die Äquivalenzklasse des Paares (p, q) unter der obigen

Äquivalenzrelation. Dann bedeutet die Definition von $\sim$ soviel wie, daß $\frac{p}{q}$ und $\frac{p'}{q'}$ gleich sind, wenn die kreuzweisen Produkte von Zähler und Nenner, pq' und $p'q$, übereinstimmen.

Auch die Rechenregeln für rationale Zahlen lassen sich mit Hilfe der Äquivalenzklassen definieren. Für $(p,q),(r,s) \in M$ definiere:

$$\big[(p,q)\big] + \big[(r,s)\big] := \big[(ps+qr, qs)\big],$$
$$\big[(p,q)\big] \cdot \big[(r,s)\big] := \big[(pr, qs)\big].$$

Wenn man sich die Äquivalenzklassen als Töpfe vorstellt, in denen man die untereinander äquivalenten Elemente gesammelt hat, so kann man das obige Rechenprinzip als "Topfrechnen" bezeichnen. Will man zwei Töpfe addieren (bzw. multiplizieren), so nimmt man aus jedem der Töpfe ein Element, addiert (bzw. multipliziert) diese Elemente und schaut, in welchen Topf das Resultat gehört. Dieser Topf ist dann die Summe (bzw. das Produkt) der beiden Töpfe.

Was man sich bei diesem Vorgehen allerdings klar machen muß, ist, daß das Ergebnis nicht von der Wahl der Repräsentanten (der Elemente aus den Töpfen) abhängt. Wir führen das für die Addition der rationalen Zahlen vor.

Sind $(p',q') \in \big[(p,q)\big]$ und $(r',s') \in \big[(r,s)\big]$ andere Repräsentanten, dann gilt $p'q = q'p$ und $r's = s'r$. Es ist zu zeigen, daß $(p's' + q'r', q's') \in \big[(ps+qr, qs)\big]$ gilt. Ausmultiplizieren liefert

$$(p's' + q'r')(qs) = p'qs's + q'qr's = q'ps's + q'qs'r = (ps+qr)(q's'),$$

was zu zeigen war.

Abschließend zeigen wir noch, daß die Äquivalenzklassen einer Menge eine Zerlegung dieser Menge in paarweise disjunkte Teilmengen liefert.

Definition 2.31 a. Zwei Mengen M und N heißen *disjunkt*, falls $M \cap N = \emptyset$.

b. Eine Familie $(M_i)_{i\in I}$ von Mengen heißt *paarweise disjunkt*, wenn für alle $i,j \in I$ mit $i \neq j$ gilt M_i und M_j sind disjunkt.

c. Es sei M eine Menge. Eine paarweise disjunkte Familie $(M_i)_{i\in I}$ von Teilmengen von M heißt eine *Zerlegung* von M, falls $M = \bigcup_{i\in I} M_i$.
Wir schreiben in diesem Fall:

$$M = \coprod_{i\in I} M_i.$$

Lemma 2.32
Es sei M eine Menge. Ist $\sim$ eine Äquivalenzrelation auf M, dann bilden die Äquivalenzklassen eine disjunkte Zerlegung von M, d. h. jedes $x \in M$ liegt in genau einer Äquivalenzklasse.

Insbesondere gilt für je zwei Äquivalenzklassen $[x]$ und $[y]$ entweder $[x] = [y]$ oder $[x] \cap [y] = \emptyset$.

Beweis: Sei $x \in M$ beliebig. Aus $x \sim x$ folgt $x \in [x] \subseteq \bigcup_{[y]\in M/\sim}[y]$. Mithin gilt

$$M = \bigcup_{[y]\in M/\sim} [y].$$

Es bleibt also zu zeigen, daß die Äquivalenzklassen paarweise disjunkt sind.

Seien $[x], [y] \in M/\sim$ mit $[x] \cap [y] \neq \emptyset$. Dann gibt es ein $z \in [x] \cap [y]$, und es gilt $z \sim x$ und $z \sim y$. Wegen der Symmetrie gilt aber auch $x \sim z$ und mittels der Transitivität dann $x \sim y$. Sei nun $u \in [x]$ beliebig, dann gilt $u \sim x$ und wieder wegen der Transitivität $u \sim y$. Also $u \in [y]$ und damit $[x] \subseteq [y]$. Vertauschung der Rollen von x und y in der Argumentation liefert schließlich $[x] = [y]$. □

Korollar 2.33
Sei M eine endliche Menge, $\sim$ eine Äquivalenzrelation auf M und $M_1, \ldots, M_s$ seien die paarweise verschiedenen Äquivalenzklassen von $\sim$. Dann gilt:

$$|M| = \sum_{i=1}^{s} |M_i|.$$

Beweis: Mit M sind auch alle M_i endlich und die Behauptung folgt aus Lemma 2.32. □

AUFGABEN

Aufgabe 2.34
Seien $A, B, C \subset M$ Mengen.

a. $A \times (B \cup C) = (A \times B) \cup (A \times C)$,
b. $A \times (B \cap C) = (A \times B) \cap (A \times C)$,
c. $A \times (B \setminus C) = (A \times B) \setminus (A \times C)$,
d. $A \times B = B \times A$ genau dann, wenn einer der folgenden Fälle vorliegt
   (i) $A = B$,
   (ii) $A = \emptyset$, oder
   (iii) $B = \emptyset$.

Aufgabe 2.35 (De Morgansche Gesetze)
Sei X eine Menge und $A, B \subset X$. Zeige:

a. $X \setminus (A \cup B) = (X \setminus A) \cap (X \setminus B)$, und
b. $X \setminus (A \cap B) = (X \setminus A) \cup (X \setminus B)$.

Aufgabe 2.36 (Verallgemeinertes Distributivgesetz)
Seien M eine Menge und $(M_i)_{i\in I}$ eine Familie von Mengen. Beweise:

$$M \cup \left(\bigcap_{i\in I} M_i\right) = \bigcap_{i\in I} (M \cup M_i).$$

Aufgabe 2.37
Es seien M und N Mengen, und $\mathcal{P}(M)$ und $\mathcal{P}(N)$ bezeichnen die Potenzmengen. Beweise die folgenden Aussagen:

a. $\mathcal{P}(M \cap N) = \mathcal{P}(M) \cap \mathcal{P}(N)$,

b. $\mathcal{P}(M \cup N) \supseteq \mathcal{P}(M) \cup \mathcal{P}(N)$,
c. $\mathcal{P}(M \cup N) = \mathcal{P}(M) \cup \mathcal{P}(N)$ genau dann, wenn $M \subseteq N$ oder $N \subseteq M$.

Aufgabe 2.38
Gib die zu den folgenden Mengen gehörenden Potenzmengen an:

$$\emptyset,\ \{\emptyset\},\ \{a, b\} \text{ und } \mathcal{P}\big(\{a, b\}\big).$$

Aufgabe 2.39
Untersuche die folgenden Abbildungen auf Injektivität, Surjektivität und Bijektivität:

a. $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R} \times \mathbb{R} \ : \ (x, y) \mapsto (y, 3)$,
b. $g : \mathbb{R} \times \mathbb{R} \to \mathbb{R} \times \mathbb{R} \ : \ (x, y) \mapsto (x + 3, y - 2)$,
c. $h : \mathbb{R} \times \mathbb{R} \to \mathbb{R} \times \mathbb{R} \ : \ (x, y) \mapsto (xy, x + 1)$, und
d. $k : \mathbb{R} \times \mathbb{R} \to \mathbb{R} \times \mathbb{R} \ : \ (x, y) \mapsto (xy, x + y)$.

Aufgabe 2.40
Es sei M eine Menge, $\mathcal{P}(M)$ die Potenzmenge von M. Wir definieren ein Abbildung:

$$f : \mathcal{P}(M) \to \mathcal{P}(M) \ : \ X \to M \setminus X.$$

Zeige, f ist eine Bijektion.

Aufgabe 2.41
Prüfe, ob die folgenden Abbildungen bijektiv sind:

a. $f : \mathbb{R} \to \mathbb{R} \ : \ x \mapsto 3x + 2$,
b. $g : \mathbb{Z} \to \mathbb{Z} \ : \ x \mapsto 3x + 2$.

Aufgabe 2.42
Finde eine Menge M und zwei Abbildungen $g,\ f : M \to M$, so daß gilt:

a. f ist surjektiv, aber nicht injektiv.
b. g ist injektiv, aber nicht surjektiv.

Aufgabe 2.43
Es sei M eine Menge, $\mathcal{P}(M)$ die Potenzmenge von M. Zeige, daß es keine surjektive Abbildung $f : M \to \mathcal{P}(M)$ gibt.
Hinweis: Man nehme an, daß es eine Surjektion f gibt und betrachte die Menge $\{m \in M \mid m \notin f(m)\}$, d. h. die Menge der Elemente von M, die nicht in ihrem Bild enthalten sind.

Aufgabe 2.44
Seien M, N zwei nicht-leere Mengen, $f : M \to N$ eine Abbildung. Beweise die folgenden beiden Aussagen:

a. f ist genau dann surjektiv, wenn für *alle* nicht-leeren Mengen X *und* für *alle* Abbildungen $g : N \to X$ und $h : N \to X$ aus $g \circ f = h \circ f$ folgt: $g = h$.
b. f ist genau dann injektiv, wenn gilt:

$$\forall\, X \neq \emptyset,\ \forall\, g : X \to M,\ \forall\, h : X \to M \ : (f \circ g) = (f \circ h) \implies g = h.$$

Aufgabe 2.45
Seien L, M, N Mengen und $f : L \to M$, $g : M \to N$ Abbildungen. Dann gelten:

a. Ist $g \circ f$ surjektiv, so ist g surjektiv, aber f nicht notwendigerweise.
b. Ist $g \circ f$ injektiv, so ist f injektiv, aber g nicht notwendigerweise.
c. Ist $g \circ f$ bijektiv, so ist g surjektiv und f injektiv.
d. Ist g surjektiv und f injektiv, so ist $g \circ f$ nicht notwendigerweise bijektiv.

Aufgabe 2.46
Seien M und N zwei *endliche* Mengen mit $|M| = |N|$, $f : M \to N$ eine Abbildung. Zeige, daß die folgenden Aussagen äquivalent sind:

a. f ist bijektiv,
b. f ist injektiv,
c. f ist surjektiv.

Aufgabe 2.47
Wir definieren auf $\mathbb{N} \setminus \{0\}$ eine Relation R durch

$$R := \left\{(n, m) \in (\mathbb{N} \setminus \{0\}) \times (\mathbb{N} \setminus \{0\}) \;\middle|\; n \mid m\right\}.$$

Ist R eine Äquivalenzrelation oder eine Ordnungsrelation (und dann ggf. eine Totalordnung oder gar eine Wohlordnung).

Aufgabe 2.48
Zu gegebenem $n \in \mathbb{N}$ definieren wir auf $\mathbb{Z}$ eine Relation

$$R_n := \left\{(x, y) \in \mathbb{Z} \times \mathbb{Z} \;\middle|\; n \mid (x - y)\right\}.$$

Ist R_n eine Äquivalenzrelation oder eine Ordnungsrelation (und dann ggf. eine Totalordnung oder gar eine Wohlordnung).

Hinweis: Für zwei ganze Zahlen x und y gilt $x \mid y$, in Worten x teilt y, genau dann, wenn es eine ganze Zahl z gibt, so daß $x \cdot z = y$. Vgl. auch Definition 5.3.

Aufgabe 2.49
Sei M eine endliche Menge. Ist die Relation $R = \left\{(A, B) \in \mathcal{P}(M) \times \mathcal{P}(M) \;\middle|\; |A| = |B|\right\}$ auf der Potenzmenge $\mathcal{P}(M)$ eine Ordnungsrelation?

Aufgabe 2.50
Es sei $\mathbb{R}^{\mathbb{R}}$ die Menge aller Abbildungen von $\mathbb{R}$ nach $\mathbb{R}$. Wir definieren eine Relation R auf $\mathbb{R}^{\mathbb{R}}$ durch $R = \{(f, g) \in \mathbb{R}^{\mathbb{R}} \times \mathbb{R}^{\mathbb{R}} \mid f(0) = g(0)\}$. Zeige, daß R eine Äquivalenzrelation ist.

Aufgabe 2.51
Wieviele Äquivalenzrelationen gibt es auf der Menge $\{1, 2, 3\}$.

Aufgabe 2.52
Suche zu jeder Teilmenge T der Menge $M = \{\text{reflexiv, symmetrisch, transitiv}\}$ eine Relation, die zwar die Eigenschaften aus T, nicht aber die aus $M \setminus T$ besitzt.

Aufgabe 2.53
Ist der folgende *Beweis* korrekt, wonach schon aus der Symmetrie und der Transitivität einer Relation R auf einer Menge M die Reflexivität folgt?

Sei $a \in M$.
Aus $(a,b) \in R$ folgt wegen der Symmetrie $(b,a) \in R$.
Aus $(a,b) \in R$ und (wie eben gezeigt) $(b,a) \in R$ folgt dann wegen der Transitivität $(a,a) \in R$, also die Reflexivität.

3. Gruppen und Homomorphismen

Die grundlegendste und wichtigste algebraische Struktur auf einer Menge ist die Gruppenstruktur.

Definition 3.1 a. Eine *Gruppe* ist ein Paar $(G,\cdot)$ bestehend aus einer nichtleeren Menge G und einer zweistelligen Operation "$\cdot$", d. h. einer Abbildung

$$\cdot : G \times G \to G : (x,y) \mapsto x \cdot y,$$

so daß die folgenden *Gruppenaxiome* gelten:

(i) $(x \cdot y) \cdot z = x \cdot (y \cdot z) \quad \forall\, x,y,z \in G,$ ("Assoziativgesetz")

(ii) $\exists\, e \in G \,:\, \forall\, x \in G \,:\, e \cdot x = x,$ ("Existenz eines Neutralen")

(iii) $\forall\, x \in G\ \exists\, x' \in G \,:\, x' \cdot x = e.$ ("Existenz von Inversen")

Ein Element mit der Eigenschaft von e nennt man ein *neutrales Element* der Gruppe G. Ein Element mit der Eigenschaft von x' nennt man ein *Inverses* zu x.

b. Erfüllt das Paar $(G,\cdot)$ nur das Axiom (i) so nennt man $(G,\cdot)$ eine *Halbgruppe*.

c. Wir nennen $(G,\cdot)$ ein *Monoid*, falls nur das Axiom (i) und (ii)' gelten:

(ii)' $\exists\, e \in G \,:\, \forall\, x \in G \,:\, e \cdot x = x \cdot e = x.$ ("Existenz eines Neutralen")

d. Eine (Halb-)Gruppe bzw. ein Monoid $(G,\cdot)$ heißt *abelsch* oder *kommutativ*, wenn $(G,\cdot)$ zudem noch dem folgenden Axiom genügt:

(iv) $x \cdot y = y \cdot x \quad \forall\, x,y \in G$ ("Kommutativitätsgesetz")

e. Eine (Halb-)Gruppe bzw. ein Monoid $(G,\cdot)$ heißt *endlich*, falls $|G| < \infty$, und sonst *unendlich*. $|G|$ heißt die *Ordnung* von G.

Lemma 3.2
Es sei $(G,\cdot)$ eine Gruppe.

a. *Das neutrale Element $e \in G$ ist eindeutig bestimmt und hat zusätzlich die Eigenschaft:*

$$x \cdot e = x \quad \forall\, x \in G.$$

b. *Sei $x \in G$. Das inverse Element x' zu x ist eindeutig bestimmt und hat zusätzlich die Eigenschaft:*

$$x \cdot x' = e.$$

Beweis: Sei $e \in G$ ein neutrales Element, und sei $x \in G$ beliebig. Ferner sei $x' \in G$ ein zu x Inverses.

Da $(G, \cdot)$ eine Gruppe ist, gibt es ein $x'' \in G$ mit $x'' \cdot x' = e$. Also folgt, unter Ausnutzung des Assoziativgesetzes:

$$x \cdot x' = e \cdot (x \cdot x') = (x'' \cdot x') \cdot (x \cdot x') = x'' \cdot ((x' \cdot x) \cdot x')$$

$$= x'' \cdot (e \cdot x') = x'' \cdot x' = e.$$

Damit gilt aber:

$$x \cdot e = x \cdot (x' \cdot x) = (x \cdot x') \cdot x = e \cdot x = x.$$

Folglich sind die zusätzlichen Eigenschaften in a. und b. gezeigt.

Sei nun $\tilde{e} \in G$ ein weiteres neutrales Element, so gilt:

$$\tilde{e} = \tilde{e} \cdot e = e.$$

Sei $\tilde{x}' \in G$ ein weiteres inverses Element zu x, so gilt:

$$\tilde{x}' = \tilde{x}' \cdot e = \tilde{x}' \cdot (x \cdot x') = (\tilde{x}' \cdot x) \cdot x' = e \cdot x' = x'.$$

□

Notation 3.3
Statt $(G, \cdot)$ schreiben wir häufig nur G, sofern keine Unklarheiten über die Operation bestehen. Außerdem schreiben wir, für $x, y \in G$, statt $x \cdot y$ oft verkürzt xy. Das neutrale Element bezeichnen wir auch mit 1 statt mit e, oder mit 1_G bzw. e_G, wenn wir hervorheben wollen, in welcher Gruppe es das Neutrale ist. Und das zu $x \in G$ existierende, eindeutig bestimmte inverse Element wird mit x^{-1} bezeichnet.

Ist die Gruppe abelsch, so bezeichnet man die Operation meist mit $+$ anstatt mit $\cdot$. In diesem Fall verwenden wir die Bezeichnung 0 (bzw. 0_G) für das neutrale Element und $-x$ für das zu $x \in G$ eindeutig bestimmte Inverse.

Lemma 3.4
Sei G *eine Gruppe,* $x, y, a, b \in G$. *Dann gelten:*

a. $\left(x^{-1}\right)^{-1} = x$ *und* $(xy)^{-1} = y^{-1}x^{-1}$.
b. *In* G *gelten die* Kürzungsregeln*:*
   (i) $ax = bx \Rightarrow a = b$, *und*
   (ii) $xa = xb \Rightarrow a = b$.

Beweis: Der Beweis verläuft ähnlich wie der Beweis von Lemma 3.2 und sei dem Leser als Übungsaufgabe überlassen. □

Beispiel 3.5 a. $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$ und $(\mathbb{R}, +)$ mit der üblichen Addition als Gruppenoperation sind abelsche Gruppen. Das neutrale Element ist jeweils die Zahl Null, und das zu einer Zahl x existierende Inverse ist die Zahl $-x$.

b. $(\mathbb{Q}\setminus\{0\},\cdot)$ und $(\mathbb{R}\setminus\{0\},\cdot)$ mit der üblichen Multiplikation als Gruppenoperation sind ebenfalls abelsche Gruppen. Das neutrale Element ist jeweils die Zahl Eins, und das zu einer Zahl x existierende inverse Element ist die Zahl $\frac{1}{x}$.

c. $(\mathbb{Z}\setminus\{0\},\cdot)$ ist hingegen nur ein (abelsches) Monoid mit der Zahl Eins als neutralem Element. Das Axiom (iii) ist nicht erfüllt, da nur die Zahlen $x = 1$ und $x = -1$ in $\mathbb{Z}\setminus\{0\}$ ein Inverses $x^{-1} = \frac{1}{x}$ besitzen.

d. $(\mathbb{N},+)$ ist ebenfalls nur ein (abelsches) Monoid mit der Zahl Null als neutralem Element, da zu $x > 0$ kein Inverses $-x$ in $\mathbb{N}$ existiert.

e. Die einfachste Gruppe ist die *einelementige Gruppe* $G = \{e\}$, deren Gruppenoperation durch $e \cdot e = e$ definiert ist.

f. Sei M ein Menge und $\mathbb{S}(M) := \{f : M \to M \mid f \text{ ist bijektiv}\}$. Die Komposition $\circ$ definiert eine zweistellige Operation auf $\mathbb{S}(M)$.
Behauptung: $(\mathbb{S}(M),\circ)$ ist eine Gruppe. Sie heißt die *symmetrische Gruppe auf der Menge* M.
Zunächst wollen wir uns davon überzeugen, daß die Komposition zweier bijektiver Abbildung wieder bijektiv ist, sprich, daß das Bild der Abbildung $\circ$ auch wirklich wieder in $\mathbb{S}(M)$ liegt.
Sind $f, g : M \to M$ bijektiv, so existieren Abbildungen $f^{-1} : M \to M$ und $g^{-1} : M \to M$ nach Lemma 2.19, und für diese gilt (unter Verwendung der Assoziativität der Komposition, Lemma 2.14):

$$(f\circ g)\circ\left(g^{-1}\circ f^{-1}\right) = f\circ\left(g\circ g^{-1}\right)\circ f^{-1} = f\circ \mathrm{id}_M \circ f^{-1} = f\circ f^{-1} = \mathrm{id}_M,$$

und analog $\left(g^{-1}\circ f^{-1}\right)\circ(f\circ g) = \mathrm{id}_M$. Folglich gilt wieder mit Lemma 2.19, daß $f\circ g$ bijektiv ist.
Die Assoziativität von $\circ$, sprich Axiom (i), ist bereits in Lemma 2.14 gezeigt. Die Identität id_M auf M ist bijektiv (siehe Beispiel 2.21) und hat die Eigenschaft, daß $\mathrm{id}_M \circ f = f$ für alle $f \in \mathbb{S}(M)$. Sie ist mithin das neutrale Element von $(\mathbb{S}(M),\circ)$. Die zu $f \in \mathbb{S}(M)$ nach Lemma 2.19 existierende inverse Abbildung ist, wie dort gezeigt, die Inverse im Sinne von Axiom (iii). Also ist $(\mathbb{S}(M),\circ)$ eine Gruppe.

g. Ist $M = \{1,\dots,n\}$, so schreiben wir $\mathbb{S}_n$ statt $\mathbb{S}(M)$ und wir nennen $\mathbb{S}_n$ die *symmetrische Gruppe auf* n *Ziffern* oder die *Permutationsgruppe vom Grad* n. Die Elemente von $\mathbb{S}_n$ heißen *Permutationen*.

Definition 3.6
Sei $(G,\cdot)$ eine Gruppe, $x \in G$. Wir setzen $x^0 := e$, und für $i \in \mathbb{N}$, $i > 0$, definieren wir rekursiv $x^i := x \cdot x^{i-1}$, und schließlich $x^{-i} := \left(x^{-1}\right)^i$.

Lemma 3.7
Sei $(G,\cdot)$ *eine Gruppe,* $x \in G$*,* $i,j \in \mathbb{Z}$*, so gelten:*

a. $x^i \cdot x^j = x^{i+j}$*, und*
b. $\left(x^i\right)^j = x^{i\cdot j}$.

Beweis: Beachte, aus Definition 3.6 folgt unmittelbar:

$$x^k = \left(x^{-1}\right)^{-k} \quad \forall\, x \in G, \forall\, k \in \mathbb{Z}.$$

a. Seien $i, j \in \mathbb{Z}$. 1. Fall: Sei $x \in G$ beliebig, $i \geq 0$. Beweis durch Induktion nach i.
$i = 0$: Dann gilt $x^i \cdot x^j = x^0 \cdot x^j = e \cdot x^j = x^j = x^{i+j}$.
$i \mapsto i+1$: Nach Definition und Induktionsvoraussetzung gilt:

$$x^{i+1} \cdot x^j = \left(x \cdot x^i\right) \cdot x^j = x \cdot \left(x^i \cdot x^j\right) = x \cdot x^{i+j} = x^{i+1+j}.$$

2. Fall: Sei $x \in G$ beliebig, $i < 0$. Aus dem 1. Fall (angewendet auf x^{-1}) und Definition 3.6 folgt (da $-i > 0$!):

$$x^i \cdot x^j = \left(x^{-1}\right)^{-i} \cdot \left(x^{-1}\right)^{-j} = \left(x^{-1}\right)^{-i-j} = x^{i+j}.$$

b. Seien $i, j \in \mathbb{Z}$, $x \in G$. 1. Fall: $j \geq 0$. Beweis durch Induktion nach j.
$j = 0$: Dann gilt $\left(x^i\right)^j = \left(x^i\right)^0 = e = x^0 = x^{i \cdot j}$.
$j \mapsto j+1$: Nach Definition, Induktionsvoraussetzung und Teil a. gilt:

$$\left(x^i\right)^{j+1} = \left(x^i\right) \cdot \left(x^i\right)^j = x^i \cdot x^{i \cdot j} = x^{i+i \cdot j} = x^{i \cdot (j+1)}.$$

2. Fall: $j < 0$. Nach Teil a. gilt $x^{-i} \cdot x^i = x^{-i+i} = x^0 = e$, und somit $\left(x^i\right)^{-1} = x^{-i}$. Aus dem 1. Fall und Definition 3.6 folgt dann (da $-j > 0$!):

$$\left(x^i\right)^j = \left(\left(x^i\right)^{-1}\right)^{-j} = \left(x^{-i}\right)^{-j} = x^{(-i)\cdot(-j)} = x^{i \cdot j}.$$

□

Bemerkung 3.8
Ist $(H, \cdot)$ eine Halbgruppe (bzw. ein Monoid) und $x \in H$, so definiert man für $0 \neq i \in \mathbb{N}$ (bzw. $i \in \mathbb{N}$) das Element x^i analog und zeigt für $0 \neq i, j \in \mathbb{N}$ (bzw. $i, j \in \mathbb{N}$) die obigen Potenzgesetze mit den gleichen Beweisen.

Definition 3.9
Sei $(G, \cdot)$ eine Gruppe. Eine Teilmenge $U \subseteq G$ heißt *Untergruppe* von G, wenn $(U, \cdot)$ eine Gruppe ist, d. h. die Einschränkung der Operation $\cdot$ auf U macht U zu einer Gruppe.

Proposition 3.10
Sei $(G, \cdot)$ eine Gruppe und $\emptyset \neq U \subseteq G$ eine nicht-leere Teilmenge. Dann sind gleichwertig:

a. *U ist eine Untergruppe von G,*
b. *$\forall\, x, y \in U$ gilt: $xy \in U$ und $x^{-1} \in U$.*

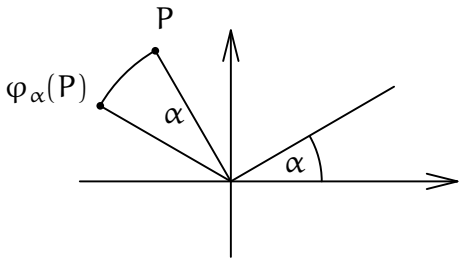
Beweis: "a. $\Rightarrow$ b.": Sei zunächst U eine Untergruppe von G. Nach Definition bedeutet dies, daß das Bild von $U \times U$ unter der Abbildung $\cdot$ in U liegt, d. h. für $x, y \in U$ gilt $xy \in U$. Außerdem gelten in U die Gruppenaxiome. Sei also $e_U \in U$ das Neutrale in U und $e_G \in G$ das Neutrale in G. Ferner bezeichne zu $x \in U$ x^{-1} stets das Inverse von x in G, d. h. $x^{-1}x = xx^{-1} = e_G$. Dann gilt:

$$e_U = e_U e_G = e_U\left(e_U e_U^{-1}\right) = (e_U e_U)e_U^{-1} = e_U e_U^{-1} = e_G.$$

Bezeichne nun zu $x \in U$ $\tilde{x}^{-1}$ das Inverse von x in U, so gilt nach Definition $\tilde{x}^{-1} \in U$. Zudem gilt aber $\tilde{x}^{-1}x = e_U = e_G$, also ist $\tilde{x}^{-1} = x^{-1}$ wegen der Eindeutigkeit des Inversen in G, und damit $x^{-1} \in U$.
"a. $\Leftarrow$ b.": Da $xy \in U$ für alle $x, y \in U$, ist das Bild von $U \times U$ unter der Abbildung $\cdot$ in der Tat in U enthalten. Es bleibt also, die Axiome (i)-(iii) nachzuprüfen. Dabei gilt (iii) nach Voraussetzung und (i) überträgt sich von der größeren Menge G auf die Teilmenge U. Da $U \neq \emptyset$, existiert ein $x \in U$. Nach Voraussetzung gilt dann aber $x^{-1} \in U$ und damit $e_G = x^{-1}x \in U$. Da aber $e_G x = x$ für alle $x \in U$, ist auch (ii) erfüllt. □

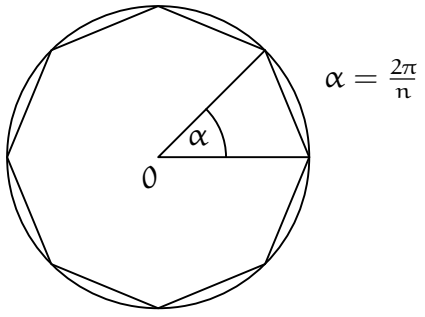
Beispiel 3.11 a. $(\{-1, 1\}, \cdot)$ ist eine Untergruppe von $(\mathbb{Q} \setminus \{0\}, \cdot)$, wie unmittelbar aus Proposition 3.10 folgt.

b. Für $\alpha \in \mathbb{R}$ bezeichne $\varphi_\alpha : \mathbb{R}^2 \to \mathbb{R}^2$ die Drehung der Ebene um den Nullpunkt um den Winkel α im Bogenmaß. Setze $SO(2) := \{\varphi_\alpha : \mathbb{R}^2 \to \mathbb{R}^2 \mid \alpha \in \mathbb{R}\}$.



Offensichtlich gilt $\varphi_\alpha \circ \varphi_\beta = \varphi_{\alpha+\beta}$ für $\alpha, \beta \in \mathbb{R}$, und für $\alpha \in \mathbb{R}$ ist somit $\varphi_{-\alpha} = (\varphi_\alpha)^{-1}$, da $\varphi_0 = \mathrm{id}_{\mathbb{R}^2}$. Insbesondere ist φ_α also bijektiv für jedes $\alpha \in \mathbb{R}$.
Damit folgt aus Proposition 3.10, daß $SO(2)$ eine Untergruppe von $\mathbb{S}(\mathbb{R}^2)$ ist.

c. Sei $E_n \subset \mathbb{R}^2$ das *reguläre* n-*Eck*.



Wir setzen

$$U := \{\varphi_\alpha \in SO(2) \mid \varphi_\alpha(E_n) = E_n\}.$$

Behauptung: $(U, \circ)$ ist eine Untergruppe von $(SO(2), \circ)$.
Für $\varphi_\alpha, \varphi_\beta \in U$ gilt

$$(\varphi_\alpha \circ \varphi_\beta)(E_n) = \varphi_\alpha(\varphi_\beta(E_n)) = \varphi_\alpha(E_n) = E_n$$

und

$$\varphi_\alpha^{-1}(E_n) = \varphi_\alpha^{-1}(\varphi_\alpha(E_n)) = (\varphi_\alpha^{-1} \circ \varphi_\alpha)(E_n) = \mathrm{id}_{\mathbb{R}^2}(E_n) = E_n.$$

Also gilt $\varphi_\alpha \circ \varphi_\beta \in U$ und $\varphi_\alpha^{-1} \in U$, und da $\mathrm{id}_{\mathbb{R}^2} = \varphi_0 \in U$, ist $U \neq \emptyset$ und folglich ist U nach Proposition 3.10 eine Untergruppe von $\mathrm{SO}(2)$.
Offensichtlich besteht U aus allen Drehungen φ_α mit $\alpha = k \cdot \frac{2\pi}{n}$, $k = 0, \ldots, n-1$. Insbesondere gilt also, $|U| = n$.

d. Sei $n \in \mathbb{Z}$ und $n\mathbb{Z} := \{nz \mid z \in \mathbb{Z}\}$ die Menge aller Vielfachen von n.
Behauptung: $(n\mathbb{Z}, +)$ ist eine Untergruppe von $(\mathbb{Z}, +)$.
Seien $nz, nz' \in n\mathbb{Z}$, dann gilt $nz + nz' = n(z + z') \in n\mathbb{Z}$ und $-(nz) = n \cdot (-z) \in n\mathbb{Z}$. Da ferner $\emptyset \neq n\mathbb{Z} \subset \mathbb{Z}$, folgt wieder mit Proposition 3.10 die Behauptung.

e. Die Inklusionen $\mathbb{Z} \subset \mathbb{Q}$, $\mathbb{Z} \subset \mathbb{R}$ und $\mathbb{Q} \subset \mathbb{R}$ machen die Teilmenge bezüglich der Addition als Gruppenstruktur jeweils zu Untergruppen.

Immer wenn man eine Struktur auf einer Menge definiert hat, spielen die *strukturerhaltenden Abbildungen* eine besondere Rolle. Diese werden (Struktur-)*Morphismen* oder (Struktur-)*Homomorphismen* genannt.

Definition 3.12
Es seien $(G, \cdot)$ und $(H, *)$ zwei Gruppen. Eine Abbildung $f : G \to H$ heißt *Gruppenhomomorphismus* (oder kürzer *Homomorphismus* oder nur *Morphismus*), falls für alle $x, y \in G$ gilt:

$$f(x \cdot y) = f(x) * f(y).$$

Beispiel 3.13 a. Ist $(G, \cdot)$ eine Gruppe und $U \subseteq G$ eine Untergruppe, dann ist die kanonische Inklusion $i_U : U \to G$ ein Gruppenhomomorphismus, da für $x, y \in U$ gilt $i_U(x \cdot y) = x \cdot y = i_U(x) \cdot i_U(y)$.

b. Sei $a \in \mathbb{R}$ und $m_a : (\mathbb{R}, +) \to (\mathbb{R}, +) : x \mapsto ax$ die Multiplikation mit a, dann ist m_a ein Gruppenhomomorphismus, da für $x, y \in \mathbb{R}$ gilt

$$m_a(x + y) = a(x + y) = ax + ay = m_a(x) + m_a(y).$$

c. Ist $(G, \cdot)$ eine Gruppe und $g \in G$, so hat man Abbildungen

$$R_g : G \to G : x \mapsto xg \qquad \text{(die “}\textit{Rechtstranslation}\text{”)}$$

und

$$L_g : G \to G : x \mapsto gx \qquad \text{(die “}\textit{Linkstranslation}\text{”)}$$

Für $g \neq e$ gilt jedoch wegen der Kürzungsregel

$$L_g(g \cdot g) = g^3 \neq g^4 = L_g(g) \cdot L_g(g)$$

und entsprechend für R_g. Also sind L_g und R_g für $g \neq e$ *keine* Gruppenhomomorphismen.
Man sieht leicht, daß L_g und R_g bijektiv sind, mit Inverser $L_{g^{-1}}$ bzw. $R_{g^{-1}}$.

d. Ist $(G, \cdot)$ eine Gruppe und $g \in G$, so definiert man

$$i_g : G \to G : x \mapsto g^{-1}xg =: x^g.$$

i_g heißt *innerer Automorphismus* oder *Konjugation* mit g.

Behauptung: Die Konjugation ist ein bijektiver Gruppenhomomorphismus.

Für $x, y \in G$ gilt:

$$\begin{aligned} i_g(xy) &= g^{-1}(xy)g = g^{-1}(xey)g = g^{-1}\Big(x\big(gg^{-1}\big)y\Big)g \\ &= \big(g^{-1}xg\big)\big(g^{-1}yg\big) = i_g(x) \cdot i_g(y), \end{aligned}$$

also ist i_g ein Gruppenhomomorphismus. Außerdem gilt für ein beliebiges $x \in G$:

$$(i_g \circ i_{g^{-1}})(x) = g^{-1}\Big(\big(g^{-1}\big)^{-1}xg^{-1}\Big)g = \big(g^{-1}g\big)x\big(g^{-1}g\big) = exe = x = \mathrm{id}_G(x),$$

also ist $i_g \circ i_{g^{-1}} = \mathrm{id}_G$. Analog sieht man $i_{g^{-1}} \circ i_g = \mathrm{id}_G$, und folglich ist i_g bijektiv nach Lemma 2.19.

Es ist offenbar $i_g = R_g \circ L_{g^{-1}}$. Die Komposition von zwei Nicht-Homomorphismen kann also durchaus ein Homomorphismus sein. Das folgende Lemma sagt, daß umgekehrt die Komposition von zwei Homomorphismen stets wieder ein Homomorphismus ist.

Lemma 3.14

Sind $f_1 : (G_1, \cdot) \to (G_2, *)$ *und* $f_2 : (G_2, *) \to (G_3, \times)$ *Gruppenhomomorphismen, so ist auch* $f_2 \circ f_1 : (G_1, \cdot) \to (G_3, \times)$ *ein Gruppenhomomorphismus.*

Beweis: Seien $x, y \in G_1$, dann gilt:

$$\begin{aligned} (f_2 \circ f_1)(x \cdot y) &= f_2\big(f_1(x \cdot y)\big) = f_2\big(f_1(x) * f_1(y)\big) = f_2\big(f_1(x)\big) \times f_2\big(f_1(y)\big) \\ &= (f_2 \circ f_1)(x) \times (f_2 \circ f_1)(y). \end{aligned}$$

□

Definition 3.15

Sei $f : (G, \cdot) \to (H, *)$ ein Gruppenhomomorphismus.

a. Wir nennen f einen *Monomorphismus*, falls f injektiv ist.
b. Wir nennen f einen *Epimorphismus*, falls f surjektiv ist.
c. Wir nennen f einen *Isomorphismus*, falls f bijektiv ist.
d. Wir nennen f einen *Endomorphismus*, falls $(G, \cdot) = (H, *)$.
e. Wir nennen f einen *Automorphismus*, falls f ein bijektiver Endomorphismus ist.

Beispiel 3.16

In den obigen Beispielen 3.13 ist m_a für $a \neq 0$ ein Automorphismus mit Inverser $m_{\frac{1}{a}}$.

Die Konjugation ist ebenfalls ein Automorphismus.

Proposition 3.17

Es sei $f : (G, \cdot) \to (H, *)$ *ein Gruppenhomomorphismus. Dann gelten:*

a. $f(e_G) = e_H$.
b. $f\big(x^{-1}\big) = \big(f(x)\big)^{-1}$ *für* $x \in G$.
c. *Ist* f *bijektiv, so ist* $f^{-1} : H \to G$ *ein Gruppenhomomorphismus.*

d. *Ist* $U \subseteq G$ *eine Untergruppe von* G*, dann ist* $f(U) \subseteq H$ *eine Untergruppe von* H.
e. *Ist* $V \subseteq H$ *eine Untergruppe von* H*, dann ist* $f^{-1}(V) \subseteq G$ *eine Untergruppe von* G.
f. $\mathrm{Im}(f) := f(G)$*, das* Bild *von* f*, ist eine Untergruppe von* H.
g. $\mathrm{Ker}(f) := f^{-1}(e_H)$*, der* Kern *von* f*, ist eine Untergruppe von* G.

Beweis: a. Es gilt

$$e_H * f(e_G) = f(e_G) = f(e_G \cdot e_G) = f(e_G) * f(e_G).$$

Mit Hilfe der Kürzungsregel 3.4 folgt dann $e_H = f(e_G)$.

b. Für $x \in G$ gilt:

$$f\big(x^{-1}\big) * f(x) = f\big(x^{-1} \cdot x\big) = f(e_G) = e_H.$$

Wegen der Eindeutigkeit der Inversen in H folgt die Behauptung.

c. Ist $f : G \to H$ bijektiv, so existiert die Umkehrabbildung $f^{-1} : H \to G$. Seien $u, v \in H$. Setze $x := f^{-1}(u)$ und $y := f^{-1}(v)$, also $u = f(x)$ und $v = f(y)$. Dann gilt:

$$f^{-1}(u * v) = f^{-1}\big(f(x) * f(y)\big) = f^{-1}\big(f(x \cdot y)\big) = x \cdot y = f^{-1}(u) \cdot f^{-1}(v).$$

Also ist f^{-1} ein Gruppenhomomorphismus.

d. Sind $u, v \in f(U)$, dann existieren $x, y \in U$ mit $f(x) = u$ und $f(y) = v$. Da $x \cdot y \in U$, gilt:

$$u * v = f(x) * f(y) = f(x \cdot y) \in f(U).$$

Außerdem gilt $x^{-1} \in U$ und somit:

$$u^{-1} = \big(f(x)\big)^{-1} = f\big(x^{-1}\big) \in f(U).$$

Da zudem $f(e_G) \in f(U)$, also $f(U) \neq \emptyset$, folgt mit Proposition 3.10, daß $f(U)$ eine Untergruppe von H ist.

e. Seien $x, y \in f^{-1}(V)$, so gilt $f(x \cdot y) = f(x) * f(y) \in V$, da V eine Untergruppe ist. Also gilt $x \cdot y \in f^{-1}(V)$. Außerdem gilt $f\big(x^{-1}\big) = \big(f(x)\big)^{-1} \in V$, wieder da V eine Untergruppe ist. Somit liegt auch x^{-1} in $f^{-1}(V)$. Da das Urbild von V unter f ferner nicht leer ist, alldieweil wegen $f(e_G) = e_H \in V$ gilt, daß $e_G \in f^{-1}(V)$, folgt wieder mit Proposition 3.10, daß $f^{-1}(V)$ eine Untergruppe von G ist.

f. Dies folgt aus d., da G eine Untergruppe von G ist.

g. Dies folgt aus e., da $\{e_H\}$ eine Untergruppe von H ist.

□

Lemma 3.18
Ein Gruppenhomomorphismus $f : (G, \cdot) \to (H, *)$ *ist genau dann injektiv, wenn* $\mathrm{Ker}(f) = \{e_G\}$.

Beweis: Ist f injektiv, so ist $f^{-1}(e_H)$ höchstens einelementig, und wegen $f(e_G) = e_H$ gilt dann $\mathrm{Ker}(f) = f^{-1}(e_H) = \{e_G\}$.

Gilt umgekehrt $\mathrm{Ker}(f) = \{e_G\}$, und sind $x, y \in G$ mit $f(x) = f(y)$, so folgt wegen:

$$e_H = f(x) * \big(f(y)\big)^{-1} = f(x) * f\big(y^{-1}\big) = f\big(x \cdot y^{-1}\big),$$

daß $x \cdot y^{-1} = e_G$, also $x = y$. Somit ist f injektiv. □

AUFGABEN

Aufgabe 3.19

Untersuche, welche der folgenden Verknüpfungen Gruppen definieren:

a. $G := \mathbb{Q} \times \mathbb{Q}$ mit $(a,b) \cdot (a',b') := (aa', bb')$ für $a, a', b, b' \in \mathbb{Q}$,
b. $G := \big(\mathbb{Q}\backslash\{0\}\big) \times \big(\mathbb{Q}\backslash\{0\}\big)$ mit $(a,b) \cdot (a',b') := (ab', ba')$ für $a, a', b, b' \in \mathbb{Q}\backslash\{0\}$,
c. $G := \mathbb{Q} \times \mathbb{Q} \setminus \big\{(0,0)\big\}$ mit $(a,b) \cdot (a',b') := (aa' - bb', ab' + ba')$ für $a, a', b, b' \in \mathbb{Q}$,
d. $G := \mathbb{Q}_{>0} \times \mathbb{Q}_{>0}$ mit $(a,b) \cdot (a',b') := (aa', bb')$ für $a, a', b, b' \in \mathbb{Q}_{>0}$.

Aufgabe 3.20 a. Eine Halbgruppe, in der die beiden Kürzungsregeln gelten, braucht noch keine Gruppe zu sein.

b. Eine endliche Halbgruppe, in der die beiden Kürzungsregeln gelten, ist eine Gruppe.

Aufgabe 3.21 a. Finde eine Halbgruppe $(H, \cdot)$, in der es unendlich viele linksneutrale Elemente gibt, d. h. unendlich viele Elemente $l \in H$ mit $l \cdot h = h$ für alle $h \in H$.

b. Gibt es in einer Halbgruppe ein Linksneutrales und ein Rechtsneutrales, so stimmen sie überein.

Aufgabe 3.22 (Direktes Produkt)

Es seien $(G, \cdot)$ und $(H, \cdot)$ zwei Gruppen. Wir definieren auf der Menge $G \times H = \{(x,y) \mid x \in G, y \in H\}$ eine Operation durch $(x,y) * (x',y') := (x \cdot x', y \cdot y')$ für $(x,y), (x',y') \in G \times H$. Zeige, dann ist $(G \times H, *)$ eine Gruppe.

Anmerkung: Wir bezeichnen $G \times H$ als das *(äußere) direkte Produkt* von G und H

Aufgabe 3.23

Es sei $(G, \cdot)$ eine Gruppe. Genau dann ist G abelsch, wenn für alle $g, h \in G$ gilt: $(g \cdot h)^2 = g^2 \cdot h^2$.

Aufgabe 3.24

Es sei $(G, \cdot)$ ein Gruppe mit neutralem Element e. Gilt $g^2 = e$ für alle $g \in G$, so ist G abelsch.

Aufgabe 3.25

Es sei $M \neq \emptyset$ eine Menge. Zeige, $(\mathbb{S}(M), \circ)$ ist genau dann abelsch, wenn $|M| \leq 2$.

Aufgabe 3.26

Welche der folgenden Mengen sind Untergruppen von $\big(\mathbb{S}(\mathbb{R}), \circ\big)$?

a. $U_1 := \{f \in \mathbb{S}(\mathbb{R}) \mid f(x) \neq x$ für nur endlich viele $x \in \mathbb{R}\}$,
b. $U_2 := \{f \in \mathbb{S}(\mathbb{R}) \mid f(x) < f(y)$ falls $x < y\}$,

c. $U_3 := \{f \in \mathbb{S}(\mathbb{R}) \mid |f(x)| = |x| \text{ für alle } x \in \mathbb{R}\}$.

Aufgabe 3.27 (Untergruppenkriterium)
Es sei $(G, \cdot)$ eine Gruppe und $\emptyset \neq U \subseteq V$ eine nicht-leere Teilmenge. Dann sind gleichwertig:

a. U ist eine Untergruppe von G.
b. Für alle $x, y \in U$ gilt: $x \cdot y^{-1} \in U$.

Aufgabe 3.28
Es sei $G := \{f_{a,b} : \mathbb{R} \to \mathbb{R} : x \mapsto ax + b \mid a, b \in \mathbb{R}, a \neq 0\}$. Zeige, daß G bezüglich der Komposition von Abbildungen eine nicht-abelsche Gruppe ist.

Aufgabe 3.29
Es sei $(G, \cdot)$ eine Gruppe und $\emptyset \neq U \subseteq G$ eine *endliche* Teilmenge. Zeige, genau dann ist U ist eine Untergruppe von G, wenn für alle $u, v \in U$ gilt: $u \cdot v \in U$.

Aufgabe 3.30
[Satz von Lagrange] Es sei $(G, \cdot)$ eine Gruppe und $U \subseteq G$ eine Untergruppe von G. Wir definieren auf G eine Relation $\sim_U$ durch

$$g \sim_U h \;:\Leftrightarrow\; g^{-1} \cdot h \in U$$

für $g, h \in G$. Man beweise die folgenden Aussagen:

a. $\sim_U$ ist eine Äquivalenzrelation mit $[g] = g \cdot U := \{g \cdot u \mid u \in U\}$ für $g \in G$.
b. Für $g \in G$ gilt, U und $g \cdot U$ sind gleichmächtig.
c. Ist $|G| < \infty$, so gilt:
$$|G| = |U| \cdot |G/\sim_U|.$$

Hinweise: In b. gebe man konkret eine Bijektion zwischen den beiden Mengen an, und für c. beachte man Lemma 2.32.

Aufgabe 3.31
Es sei $(G, \cdot)$ ein endliche Gruppe von gerader Ordnung mit neutralem Element e. Zeige die folgenden Aussagen:

a. G enthält mindestens ein Element $g \neq e$ mit $g^2 = e$.
b. Ist G eine Gruppe der Ordnung $|G| = 4$, so ist G abelsch.

Hinweis: Für a. betrachte man die Relation $\sim$ auf G mit $g \sim h \;:\Leftrightarrow\; h = g$ oder $h = g^{-1}$.

Aufgabe 3.32
Es sei $(G, \cdot)$ eine Gruppe mit genau einem Element $g \neq e$, so daß $g^2 = e$. Dann gilt $h \cdot g = g \cdot h$ für alle $h \in G$.

Aufgabe 3.33
Es sei $(G, \cdot)$ eine Gruppe. Prüfe, welche der folgenden Abbildungen ein Gruppenhomomorphismus / Gruppenisomorphismus ist, wobei $g, h \in G$ fest vorgegeben seien:

a. $G \to G : x \mapsto g \cdot x \cdot g$,
b. $G \to G : x \mapsto h \cdot x \cdot g$,
c. $G \to G : x \mapsto h^{-1} \cdot x \cdot g$,

d. $G \to G : x \mapsto x^{-1}$,

e. $G \to G : x \mapsto x^2$.

Aufgabe 3.34

$(G, \cdot)$ sei eine Gruppe, und die Abbildung $\alpha : G \to G$ sei definiert durch $\alpha(g) = g^{-1}$ für $g \in G$. Zeige, genau dann ist α ein Gruppenhomomorphismus, wenn G abelsch ist.

Aufgabe 3.35

Bestimme alle Gruppenhomomorphismen $\alpha : (\mathbb{Q}, +) \to (\mathbb{Z}, +)$.

Aufgabe 3.36

Bestimme alle Automorphismen von $(\mathbb{Q}, +)$.

Aufgabe 3.37

Es sei $(G, \cdot)$ eine Gruppe. Für $g \in G$ definieren wir

$$\alpha_g : G \to G \ : \ x \mapsto g^{-1} \cdot x \cdot g.$$

a. Zeige, für jedes $g \in G$ ist α_g ein Gruppenisomorphismus.

b. Zeige, die Menge $\mathrm{Aut}(G) := \{\alpha \in \mathbb{S}(G) \mid \alpha \text{ ist ein Automorphismus}\}$ ist eine Untergruppe von $\big(\mathbb{S}(G), \circ\big)$.

c. Zeige, die Menge $\mathrm{Inn}(G) := \{\alpha_g \mid g \in G\}$ ist eine Untergruppe von $\mathrm{Aut}(G)$.

d. Die Abbildung $\varphi : G \to \mathrm{Aut}(G) \ : \ g \mapsto \alpha_g$ ist ein Gruppenhomomorphismus mit $\mathrm{Ker}(\varphi) = Z(G) := \{h \in G \mid h \cdot x = x \cdot h \ \ \forall x \in G\}$.

4. Die symmetrische Gruppe

Die symmetrische Gruppe $\mathbb{S}(M)$ der bijektiven Selbstabbildungen einer Menge M ist die "Urmutter" aller Gruppen. Für eine beliebige Menge M ist $\mathbb{S}(M)$ allerdings wenig nützlich, da man außer der Definition kaum etwas über sie aussagen kann.

Für eine endliche Menge M ist das ganz anders. Zunächst einmal ist es egal, ob wir $\mathbb{S}\big(\{x_1, \ldots, x_n\}\big)$, für eine beliebige n-elementige Menge $M = \{x_1, \ldots, x_n\}$, betrachten oder $\mathbb{S}_n = \mathbb{S}\big(\{1, \ldots, n\}\big)$. Die beiden Gruppen sind isomorph, und zwar so offensichtlich, daß wir keinen Unterschied machen - wir identifizieren sie. $\mathbb{S}_n$ ist nicht nur die Urmutter aller endlichen Gruppen,[2] $\mathbb{S}_n$ ist auch für praktische Anwendungen sehr wichtig. In der linearen Algebra werden wir die Gruppe $\mathbb{S}_n$ vor allem im Zusammenhang mit Determinanten benötigen, vgl. Abschnitt II.9.

Definition 4.1 a. Ein Element $\sigma \in \mathbb{S}_n$ heißt eine *Permutation*. Wir beschreiben σ durch das folgende Schema:

$$\begin{pmatrix} 1 & 2 & \ldots & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(n) \end{pmatrix}$$

[2] Der Satz von Cayley sagt, daß jede endliche Gruppe G mit $|G| = n$ isomorph zu einer Untergruppe der $\mathbb{S}_n$ ist, vgl. Aufgabe 4.21.

bzw.

$$\begin{pmatrix} x_1 & x_2 & \dots & x_n \\ \sigma(x_1) & \sigma(x_2) & \dots & \sigma(x_n) \end{pmatrix},$$

falls $x_1, \dots, x_n$ irgendeine Anordnung der Zahlen $1, \dots, n$ ist.

b. Eine Permutation $\sigma \in \mathbb{S}_n$, für die es eine Zerlegung von $\{1, \dots, n\} = \{a_1, \dots, a_k\} \cup \{b_1, \dots, b_{n-k}\}$ gibt, so daß gilt:

$$\sigma = \begin{pmatrix} a_1 & a_2 & \dots & a_{k-1} & a_k & b_1 & \dots & b_{n-k} \\ a_2 & a_3 & \dots & a_k & a_1 & b_1 & \dots & b_{n-k} \end{pmatrix},$$

heißt ein **k-Zyklus**.

Wir schreiben kurz: $\sigma = (a_1 \dots a_k)$.

Beachte: die Zyklen $(a_1 \dots a_k)$, $(a_k a_1 \dots a_{k-1})$, etc. stimmen überein!

c. Ein $2-Zyklus$ wird auch eine **Transposition** genannt.
Eine Transposition $\tau = (i\ j)$ ist mithin eine Permutation, die nur die zwei Zahlen i und j miteinander vertauscht, alle anderen aber fest läßt.

d. Das neutrale Element von $\mathbb{S}_n$, per definitionem $id_{\{1,\dots,n\}}$, wollen wir der Einfachheit halber mit id bezeichnen.

Beispiel 4.2
Die Gruppe $\mathbb{S}_n$ ist für $n \geq 3$ nicht abelsch, denn für die Zyklen $(1\ 2), (1\ 3\ 2) \in \mathbb{S}_n$ gilt:

$$(1\ 2) \circ (1\ 3\ 2) = (1\ 3) \neq (2\ 3) = (1\ 3\ 2) \circ (1\ 2).$$

Bemerkung 4.3
Für kleine Werte n ist $\mathbb{S}_n$ sehr übersichtlich, für große Werte n wird $\mathbb{S}_n$ jedoch riesig.

$\mathbb{S}_1 = \{id\}$ und $\mathbb{S}_2 = \{id, (1\ 2)\}$. $\mathbb{S}_3 = \{id, (1\ 2), (1\ 3), (2\ 3), (1\ 2\ 3), (1\ 3\ 2)\}$ hat schon sechs Elemente, $\mathbb{S}_4$ gar 24 und $\mathbb{S}_{60}$ ungefähr 10^{82}. Letztere Zahl entspricht in etwa der angenommenen Anzahl der Nukleone des Universums.

Bemerkung 4.4 (Prinzip der vollständigen Induktion)
Es gelte eine Aussage für die ganze Zahl n_0 (*Induktionsanfang*), außerdem sei folgendes richtig: gilt die Aussage für ein beliebiges $n \geq n_0$ (*Induktionsvoraussetzung*), so gilt sie auch für $n + 1$ (*Induktionsschluß*). Dann gilt die Aussage für alle ganzen Zahlen $n \geq n_0$.

Proposition 4.5
$|\mathbb{S}_n| = n! = 1 \cdot 2 \cdot 3 \cdots n$.

Beweis: Wir zeigen durch Induktion über n etwas allgemeiner:
Behauptung: Sind $M = \{x_1, \dots, x_n\}$ und $N = \{y_1, \dots, y_n\}$ zwei n-elementige Mengen, so hat die Menge

$$\mathrm{Iso}(M, N) := \{f : M \to N \mid f \text{ ist bijektiv}\}$$

genau $n!$ Elemente.

Induktionsanfang: Sei $n = 1$, dann gilt offensichtlich $|\mathrm{Iso}(M, N)| = 1 = 1!$.

Induktionsschluß: Es sei $n > 1$ beliebig, aber fest, und es gelte $|\operatorname{Iso}(M', N')| = (n-1)!$ für alle $n-1$-elementigen Mengen M' und N'.

Seien nun M und N zwei n-elementige Mengen. Für $i \in \{1, \ldots, n\}$ definieren wir:

$$\operatorname{Iso}_i := \{f \in \operatorname{Iso}(M, N) \mid f(x_1) = y_i\}.$$

Offensichtlich ist die Einschränkung[3]

$$\operatorname{Iso}_i \to \operatorname{Iso}\left(M \setminus \{x_1\}, N \setminus \{y_i\}\right) : f \mapsto f_{|M\setminus\{x_1\}}$$

bijektiv, und daher gilt nach Induktionsvoraussetzung $|\operatorname{Iso}_i| = (n-1)!$. Da nun außerdem

$$\operatorname{Iso}(M, N) = \coprod_{i=1}^{n} \operatorname{Iso}_i,$$

d. h. $(\operatorname{Iso}_i)_{i\in\{1,\ldots,n\}}$ ist eine disjunkte Zerlegung von $\operatorname{Iso}(M, N)$, folgt:

$$|\operatorname{Iso}(M, N)| = \sum_{i=1}^{n} |\operatorname{Iso}_i| = n \cdot (n-1)! = n!.$$

□

Bemerkung 4.6
Wir wollen uns jetzt mit den Transpositionen näher beschäftigen. Zunächst ist klar, daß für eine Transposition $\tau \in \mathbb{S}_n$ gilt $\tau^{-1} = \tau$, also $\tau^2 = \mathrm{id}$.

Allgemeiner gilt für eine Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & \ldots & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(n) \end{pmatrix},$$

daß das Inverse σ^{-1} gegeben ist durch

$$\sigma^{-1} = \begin{pmatrix} \sigma(1) & \sigma(2) & \ldots & \sigma(n) \\ 1 & 2 & \ldots & n \end{pmatrix}.$$

Man kann das Inverse einer Permutation also einfach durch ablesen “von unten nach oben” beschreiben.

Proposition 4.7
Jede Permutation in $\mathbb{S}_n$, $n \geq 2$, läßt sich als Komposition von höchstens n Transpositionen darstellen.

Beweis: Wir führen den Beweis durch Induktion über n.

Induktionsanfang: Sei $n = 2$. Es ist $\mathbb{S}_2 = \{\mathrm{id}, (1\ 2)\}$, und $\mathrm{id} = (1\ 2) \circ (1\ 2)$, also folgt die Behauptung.

Induktionsschluß: Sei nun $n \geq 2$ gegeben, und die Behauptung gelte für n bereits. Ferner sei $\sigma \in \mathbb{S}_{n+1}$ beliebig, aber fest. Es gibt ein $i \in \{1, \ldots, n+1\}$ mit $\sigma(n+1) = i$. Dann gilt mit $\tau = (n+1\ \ i)$

$$(\tau \circ \sigma)(n+1) = n+1,$$

[3] Siehe Bemerkung 2.8.

also können wir die Einschränkung[4] $\sigma' = (\tau \circ \sigma)_{|\{1,\dots,n\}}$ als Element von $\mathbb{S}_n$ auffassen. Mithin gilt nach Induktionsvoraussetzung, es gibt Transpositionen $\tau'_1, \dots, \tau'_k \in \mathbb{S}_n$, $k \leq n$, mit

$$\sigma' = \tau'_1 \circ \cdots \circ \tau'_k.$$

Bezeichnen wir mit τ_j die Fortsetzung von τ'_j, die definiert wird durch

$$\tau_j : \{1, \dots, n+1\} \to \{1, \dots, n+1\} : l \mapsto \begin{cases} \tau'_j(l), & \text{falls } l \leq n \\ n+1, & \text{falls } l = n+1, \end{cases}$$

so folgt unmittelbar

$$\tau \circ \sigma = \tau_1 \circ \cdots \tau_k,$$

und mithin

$$\sigma = (\tau \circ \tau) \circ \sigma = \tau \circ (\tau \circ \sigma) = \tau \circ \tau_1 \circ \cdots \circ \tau_k.$$

D. h. σ ist Komposition von $k+1 \leq n+1$ Transpositionen. □

Korollar 4.8
Jede Permutation läßt sich als Produkt von Transpositionen zweier aufeinanderfolgender Zahlen schreiben.

Beweis: Wegen Proposition 4.7 reicht es, dies für eine Transposition $(i\ j)$ mit $i < j$ zu zeigen. Es gilt aber offenbar

$$\begin{aligned} (i\ j) \;=\; & (i\ i+1) \circ (i+1\ i+2) \circ \cdots \circ (j-2\ j-1) \circ (j-1\ j) \circ \\ & \circ (j-2\ j-1) \circ \cdots \circ (i+1\ i+2) \circ (i\ i+1). \end{aligned}$$

□

Die Darstellung einer Permutation als Komposition von Transpositionen ist also keineswegs eindeutig. Was jedoch unabhängig ist, ist, daß eine Permutation entweder immer durch eine gerade oder immer durch eine ungerade Anzahl von Transpositionen darstellbar ist. Das wollen wir nun beweisen und definieren dazu das Vorzeichen einer Permutation.

Definition 4.9
Es sei $\sigma \in \mathbb{S}_n$ gegeben.

a. Ein Zahlenpaar (i, j) mit $1 \leq i, j \leq n$ heißt ein *Fehlstand* von σ, falls $i < j$, aber $\sigma(i) > \sigma(j)$.
b. Wir definieren das *Signum* oder *Vorzeichen* von σ durch

$$\operatorname{sgn}(\sigma) = \begin{cases} +1, & \text{falls } \sigma \text{ eine gerade Anzahl von Fehlständen besitzt,} \\ -1, & \text{falls } \sigma \text{ eine ungerade Anzahl von Fehlständen besitzt.} \end{cases}$$

Beispiel 4.10
Eine Transposition $\tau = (i\ j) \in \mathbb{S}_n$, mit $i < j$, hat exakt $2 \cdot (j - i - 1) - 1$ Fehlstände, nämlich

$$(i\ i+1), (i\ i+2), \dots, (i\ j), (i+1\ j), (i+2\ j), \dots, (j-1\ j),$$

[4] Siehe Bemerkung 2.8.

und mithin gilt $\operatorname{sgn}(\tau) = -1$.

Die Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}$$

hat die Fehlstände $(1,2)$ und $(3,4)$. Also gilt $\operatorname{sgn}(\sigma) = 1$.

Manchmal ist die folgende geschlossene Formel nützlich, deren Beweis als Übungsaufgabe dem Leser überlassen sei.

Lemma 4.11
Für $\sigma \in \mathbb{S}_n$ gilt:

$$\operatorname{sgn}(\sigma) = \prod_{1 \le i < j \le n} \frac{\sigma(j) - \sigma(i)}{j - i} = \frac{\sigma(2) - \sigma(1)}{2 - 1} \cdot \frac{\sigma(3) - \sigma(1)}{3 - 1} \cdots \frac{\sigma(n) - \sigma(n-1)}{n - (n-1)}.$$

Satz 4.12 a. *Für $\sigma_1, \sigma_2 \in \mathbb{S}_n$ gilt:*

$$\operatorname{sgn}(\sigma_1 \circ \sigma_2) = \operatorname{sgn}(\sigma_1) \cdot \operatorname{sgn}(\sigma_2).$$

b. *Ist $\sigma = \tau_1 \circ \cdots \circ \tau_k \in \mathbb{S}_n$ eine Komposition von k Transpositionen, dann gilt:*

$$\operatorname{sgn}(\sigma) = (-1)^k.$$

Beweis: Es sei $\sigma = \sigma' \circ \tau \in \mathbb{S}_n$ mit $\sigma' \in \mathbb{S}_n$ und $\tau = (i\ i+1)$ für ein $i \in \{1, \ldots, n-1\}$. Ist $(i, i+1)$ ein Fehlstand von σ', so hebt τ diesen auf und σ hat einen Fehlstand weniger als σ'. Ist hingegen $(i, i+1)$ kein Fehlstand von σ', so erzeugt die Komposition mit τ diesen Fehlstand und σ hat einen Fehlstand mehr als σ'. Damit gilt dann aber

$$\operatorname{sgn}(\sigma) = -\operatorname{sgn}(\sigma') = \operatorname{sgn}(\sigma') \cdot \operatorname{sgn}(\tau).$$

Da jede Transposition als Produkt von Transpositionen zweier aufeinanderfolgender Zahlen geschrieben werden kann, läßt sich wegen Proposition 4.7 auch jede Permutation als Produkt solcher Transpositionen schreiben.

Seien nun $\sigma_1 = \tilde{\tau}_1 \circ \cdots \circ \tilde{\tau}_r$ und $\sigma_2 = \tilde{\tau}_{r+1} \circ \cdots \circ \tilde{\tau}_{r+s}$ als Produkte solcher Transpositionen aufeinanderfolgender Zahlen gegeben. Dann folgt mit Induktion

$$\operatorname{sgn}(\sigma_1 \circ \sigma_2) = (-1)^{r+s} = (-1)^r \cdot (-1)^s = \operatorname{sgn}(\sigma_1) \cdot \operatorname{sgn}(\sigma_2).$$

Damit ist a. gezeigt und b. folgt mittels Induktion nach k.

□

Bemerkung 4.13
Die Aussage in Satz 4.12 a. läßt sich eleganter auch wie folgt formulieren. Betrachte sgn als eine Abbildung

$$\operatorname{sgn} : \mathbb{S}_n \to \{-1, 1\},$$

und betrachte $(\{-1, 1\}, \cdot)$ als Gruppe (siehe Beispiel 3.11). Dann besagt 4.12 a., daß sgn ein Gruppenhomomorphismus ist.

Der Kern dieses Homomorphismus besteht aus allen Permutationen mit positivem Vorzeichen, man nennt diese auch *gerade* Permutationen, und ist nach Proposition 3.17 eine Untergruppe der $\mathbb{S}_n$.

Die Menge $\{\sigma \in \mathbb{S}_n \mid \operatorname{sgn}(\sigma) = -1\}$ ist keine Untergruppe der $\mathbb{S}_n$, da sie etwa das neutrale Element id nicht enthält.

Definition 4.14

$\mathbb{A}_n := \operatorname{Ker}(\operatorname{sgn}) = \{\sigma \in \mathbb{S}_n \mid \operatorname{sgn}(\sigma) = 1\}$ heißt die *alternierende Gruppe* vom Grad n.

Lemma 4.15

$|\mathbb{A}_n| = \frac{n!}{2}$, *für* $n \geq 2$.

Beweis: Es sei $\tau \in \mathbb{S}_n$. Wegen Satz 4.12 gilt:

$$\mathbb{S}_n \setminus \mathbb{A}_n = \{\sigma \in \mathbb{S}_n \mid \operatorname{sgn}(\sigma) = -1\} = \{\tau \circ \sigma' \mid \sigma' \in \mathbb{A}_n\} =: \tau\mathbb{A}_n.$$

Ferner ist die Abbildung

$$\mathbb{A}_n \to \tau\mathbb{A}_n : \sigma \mapsto \tau \circ \sigma$$

offensichtlich eine Bijektion. Also gilt

$$|\mathbb{S}_n| = |\mathbb{S}_n \setminus \mathbb{A}_n| + |\mathbb{A}_n| = |\tau\mathbb{A}_n| + |\mathbb{A}_n| = 2 \cdot |\mathbb{A}_n|.$$

□

AUFGABEN

Aufgabe 4.16 (Zyklenzerlegung einer Permutation)

Sei $0 \neq n \in \mathbb{N}$ und $(\mathbb{S}_n, \circ)$ die Symmetrische Gruppe vom Grad n.

Sei $\sigma \in \mathbb{S}_n$ eine beliebige Permutation. Beweise die folgenden Aussagen a.–c. und löse Teil d.:

a. Für $a, b \in \{1, \ldots, n\}$ definieren wir

$$a \sim b \;:\Leftrightarrow\; \exists\, \nu \in \mathbb{Z} \;:\; a = \sigma^{\nu}(b).$$

$\sim$ ist eine Äquivalenzrelation auf $\{1, \ldots, n\}$.

b. Für $a \in \{1, \ldots, n\}$ bezeichne $[a]$ die Äquivalenzklasse von a bez. $\sim$. Dann gilt:

$$\exists\, k \in \mathbb{N} \;:\; [a] = \left\{a, \sigma(a), \ldots, \sigma^{k-1}(a)\right\}.$$

c. Es gibt eine disjunkte Zerlegung $\{1, \ldots, n\} = \bigcup_{i=1}^{t} \{a_{i1}, \ldots, a_{ik_i}\}$, so daß

$$\sigma = (a_{11} \cdots a_{1k_1}) \circ \ldots \circ (a_{t1} \cdots a_{tk_t}).$$

(Wir nennen diese Darstellung die *Zyklenzerlegung* von σ.)

d. Es seien $\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 4 & 3 & 1 \end{pmatrix}, \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 2 & 4 & 1 \end{pmatrix} \in \mathbb{S}_5$ gegeben. Berechne die folgenden Permutationen und gib ihre Zyklenzerlegung (gemäß c.) sowie eine Zerlegung in Transpositionen (= 2-Zykel) (gemäß Proposition 4.7) an:

$$\pi \circ \sigma, \qquad \sigma \circ \pi, \qquad \pi^{-1}, \qquad \sigma^{-1}.$$

Hinweis: Die Teile a. bis c. bauen jeweils aufeinander auf. Wer Teil b. sauber beweisen möchte, denke über die folgenden Aussagen nach: (i) $\exists\, k_0 \geq 0$ minimal mit $\sigma^{k_0}(a) = a$; (ii) für alle $\lambda \in \mathbb{Z}$ gilt $\sigma^{\lambda k_0}(a) = a$; und (iii) für die ganzen Zahlen ist Division mit Rest möglich.

Aufgabe 4.17

Sei $0 \neq n \in \mathbb{N}$ eine natürliche Zahl und $\mathrm{sgn} : \mathbb{S}_n \to \{-1, 1\}$ der Signum-Homomorphismus. Zeige, falls $\pi \in \mathbb{S}_n$ die Zyklenzerlegung $\pi = (a_{11} \cdots a_{1k_1}) \circ \ldots \circ (a_{t1} \cdots a_{tk_t})$ besitzt, so gilt

$$\mathrm{sgn}(\pi) = (-1)^{k_1 - 1} \cdots (-1)^{k_t - 1}.$$

Insbesondere, ein k-Zyklus hat genau dann Signum 1, wenn k ungerade ist.

Aufgabe 4.18 (Ordnung einer Permutation)

Sei $(G, \cdot)$ eine Gruppe, $g \in G$. Wir definieren

$$o(g) := \begin{cases} \infty, & \text{falls } g^k \neq e \ \ \forall\, 0 \neq k \in \mathbb{N}, \\ \min\{k \in \mathbb{N} \setminus \{0\} \mid g^k = e\}, & \text{sonst.} \end{cases}$$

Sei nun $(G, \cdot) = (\mathbb{S}_n, \circ)$ und $\pi \in \mathbb{S}_n$ habe die Zyklenzerlegung $\pi = (a_{11} \cdots a_{1k_1}) \circ \ldots \circ (a_{t1} \cdots a_{tk_t})$, dann gilt

$$o(g) = \mathrm{kgV}(k_1, \ldots, k_t).$$

Insbesondere, ein k-Zyklus hat Ordnung k.

Aufgabe 4.19 (Typ einer Permutation)

Es sei $\pi \in \mathbb{S}_n$ eine Permutation mit Zyklenzerlegung $\pi = (a_{11} \cdots a_{1k_1}) \circ \ldots \circ (a_{t1} \cdots a_{tk_t})$, so daß gilt $k_1 \leq k_2 \leq \cdots \leq k_t$. Dann heißt das Tupel $(k_1, \ldots, k_t)$ der Typ der Permutation π. Für zwei Permutationen $\pi, \sigma \in \mathbb{S}_n$ mit Typ $(k_1, \ldots, k_t)$ bzw. $(l_1, \ldots, l_s)$ gilt:

$$\exists \zeta \in \mathbb{S}_n \ : \ \pi = \zeta^{-1} \circ \sigma \circ \zeta \quad \Longleftrightarrow \quad (k_1, \ldots, k_t) = (l_1, \ldots, l_s).$$

(In Worten, zwei Permutationen sind genau dann konjugiert, wenn sie den gleichen Typ haben.)

Aufgabe 4.20 (Kleinsche Vierergruppe)

Die Teilmenge $\mathbb{K}_4 := \{(1), (1\ 2) \circ (3\ 4), (1\ 3) \circ (2\ 4), (1\ 4) \circ (2\ 3)\} \subset \mathbb{S}_4$ ist eine Untergruppe der Symmetrischen Gruppe vom Grad 4, die sogenannte Kleinsche Vierergruppe.

Aufgabe 4.21 (Satz von Cayley)

Ist $(G, \cdot)$ eine endliche Gruppe mit $|G| = n$, so ist G isomorph zu einer Untergruppe von $\mathbb{S}_n$.

5. Faktorgruppen und Kongruenzen

Wir kommen in diesem Abschnitt noch einmal auf den Begriff der Äquivalenzrelation, d. h. einer reflexiven, symmetrischen und transitiven Relation auf einer Menge, zurück. Allerdings wollen wir jetzt Relationen auf einer Gruppe betrachten, die mit der Gruppenstruktur verträglich sind.

Definition 5.1
Es sei G eine Gruppe und $\sim$ eine Äquivalenzrelation auf G. $\sim$ heißt *mit der Gruppenstruktur verträgliche Äquivalenzrelation*, falls für alle $x, y, z \in G$ gilt:

$$x \sim y \;\Rightarrow\; zx \sim zy.$$

Lemma 5.2
Es sei G eine Gruppe.

a. *Ist $\sim$ eine mit der Gruppenstruktur verträgliche Äquivalenzrelation, dann gelten:*
   (i) *Die Äquivalenzklasse $U := [e]$ des neutralen Elementes $e \in G$ ist eine Untergruppe von G.*
   (ii) *Für $x, y \in G$ gilt:*

$$x \sim y \;\Leftrightarrow\; x^{-1}y \in U.$$

b. *Ist umgekehrt $U \subseteq G$ eine Untergruppe von G, so wird durch*

$$x \sim y \;:\Leftrightarrow\; x^{-1}y \in U \quad \textit{für } x, y \in G$$

eine mit der Gruppenstruktur verträgliche Äquivalenzrelation definiert.

In beiden Fällen gilt für $x \in G$

$$[x] = x \cdot U := \{xu \mid u \in U\}.$$

Die Menge der Äquivalenzklassen wird mit G/U bezeichnet, $|G : U| := \#G/U$ heißt der Index *von U in G, und die zu $x \in G$ gehörende Äquivalenzklasse $x \cdot U$ wird* Restklasse modulo U *genannt.*

Beweis: a. (i) Sind $x, y \in U$, dann gilt $x \sim e$ und $y \sim e$. Damit gilt aber

$$xy \sim xe = x \sim e,$$

also $xy \in U$. Außerdem gilt

$$x^{-1} = x^{-1}e \sim x^{-1}x = e,$$

also auch $x^{-1} \in U$. Da ferner U nicht leer ist, wegen $e \sim e$, ist mit Proposition 3.10 U eine Untergruppe von G.

(ii) Für $x, y \in G$ mit $x \sim y$ gilt

$$e = x^{-1}x \sim x^{-1}y,$$

also $x^{-1}y \in U$. Umgekehrt folgt aus $x^{-1}y \in U$

$$x = xe \sim x\big(x^{-1}y\big) = y.$$

b. Seien $x, y, z \in G$. Da $x^{-1}x = e \in U$, ist $\sim$ reflexiv. Da mit $x^{-1}y \in U$ auch $y^{-1}x = \big(x^{-1}y\big)^{-1} \in U$ gilt, ist $\sim$ auch symmetrisch. Ferner folgt aus $x^{-1}y, y^{-1}z \in U$, daß $x^{-1}z = \big(x^{-1}y\big)\big(y^{-1}z\big) \in U$, und damit die Transitivität von $\sim$. Mithin ist $\sim$ eine Äquivalenzrelation.

Ferner gilt für $x, y, z \in G$ mit $x \sim y$

$$(zx)^{-1}(zy) = \left(x^{-1}z^{-1}\right)(zy) = x^{-1}\left(z^{-1}z\right)y = x^{-1}ey = x^{-1}y \in U.$$

Also ist $\sim$ mit der Gruppenstruktur verträglich.

□

Ein ganz wichtiges Beispiel hierzu ist die Kongruenz ganzer Zahlen.

Definition 5.3

Sei $n \in \mathbb{Z}$ fest gewählt. $x, y \in \mathbb{Z}$ heißen *kongruent modulo* n, falls $x - y \in n\mathbb{Z} = \{np \mid p \in \mathbb{Z}\}$.

Wir verwenden die Notation

$$x \equiv y \ (n) \quad \text{oder} \quad x \equiv y \ (\mathrm{mod}\ n).$$

Bemerkung 5.4

Da $n\mathbb{Z} \subseteq \mathbb{Z}$ eine Untergruppe von $(\mathbb{Z}, +)$ ist, ist nach Lemma 5.2 Kongruenz modulo n eine mit der Gruppenstruktur verträgliche Äquivalenzrelation auf $\mathbb{Z}$. Weiter folgt für $x \in \mathbb{Z}$

$$[x] = x + n\mathbb{Z}$$

ist die zu x gehörende Äquivalenzklasse und wird auch *Kongruenzklasse* modulo n genannt. $x + n\mathbb{Z}$ ist die Menge aller $y \in \mathbb{Z}$, so daß $x - y$ durch n teilbar ist.

Für die Menge der Kongruenzklassen modulo n schreiben wir

$$\mathbb{Z}_n := \mathbb{Z}/n := \mathbb{Z}/(n) := \mathbb{Z}/n\mathbb{Z}.$$

Aus der *Division mit Rest*, siehe Lemma 5.5, folgt für jedes $x \in \mathbb{Z}$ existiert ein $q \in \mathbb{Z}$ und ein $r \in \mathbb{Z}$ mit $0 \leq r < n - 1$, so daß gilt

$$x = qn + r.$$

Es folgt

$$\mathbb{Z}_n = \{[0], [1], \ldots, [n-1]\},$$

d. h. $\mathbb{Z}_n$ hat genau n Elemente, d. h. n Kongruenzklassen modulo n in $\mathbb{Z}$, die eindeutig durch die Zahlen $0, 1, \ldots, n-1$ repräsentiert werden. So repräsentiert z. B. $[0]$ die durch n teilbaren Zahlen in $\mathbb{Z}$.

Lemma 5.5 (Division mit Rest)

Zu $x, y \in \mathbb{Z}$ mit $y \neq 0$ existieren eindeutig bestimmte Zahlen $q, r \in \mathbb{Z}$ mit

$$x = qy + r, \quad 0 \leq r < |y|. \tag{4}$$

Beweis: Die Existenz von q und r ist klar. Es bleibt also die Eindeutigkeit zu zeigen. Ist $\left(q', r'\right)$ ein weiteres Zahlenpaar, für das (4) gilt, so folgt:

$$x = qy + r = q'y + r'. \tag{5}$$

O. E. gilt $r' \geq r$, also $0 \leq r \leq r' < |y|$. Dann folgt aber aus (5)

$$0 \leq \left(q - q'\right)y = r' - r < |y|.$$

Da $q - q' \in \mathbb{Z}$, muß folglich $q - q' = 0$ gelten, also $q = q'$ und dann auch $r = r'$. □

Definition 5.6
Eine Untergruppe $U \subseteq G$ von G heißt *normal* oder *Normalteiler*, falls für alle $x \in G$ und $u \in U$ gilt:

$$xu \in Ux := U \cdot x := \{u'x \mid u' \in U\}.$$

Wir schreiben in diesem Falle $U \trianglelefteq G$.

Bemerkung 5.7 a. Aus der Definition folgt für eine Untergruppe U unmittelbar, daß $U \trianglelefteq G$ genau dann, wenn für alle $x \in G$ und $u \in U$ gilt:

$$xux^{-1} \in U. \tag{6}$$

b. Man beachte, daß eine Teilmenge $U \subseteq G$ zunächst einmal eine Untergruppe sein muß, um ein Normalteiler sein zu können!
c. Ist G abelsch, so ist jede Untergruppe ein Normalteiler.
d. Die Untergruppe $U := \{\mathrm{id}, (1\ 2)\} \subset \mathbb{S}_3$ ist kein Normalteiler der $\mathbb{S}_3$, denn für $\sigma = (2\ 3) \in \mathbb{S}_3$ gilt

$$\sigma \circ (1\ 2) \circ \sigma^{-1} = (2\ 3) \circ (1\ 2) \circ (2\ 3) = (1\ 3) \notin U.$$

Satz 5.8
Es sei $(G, \cdot)$ *eine Gruppe.*

a. *Eine Untergruppe* $U \subseteq G$ *von* G *ist genau dann ein Normalteiler, wenn auf der Menge* G/U *die Operation*

$$[x] \cdot [y] := [x \cdot y], \quad \textit{für } [x], [y] \in G/U,$$

wohldefiniert ist, d. h. die Vorschrift $G/U \times G/U \to G/U : \big([x], [y]\big) \mapsto [x \cdot y]$ *definiert eine Abbildung.*

b. *Ist* $U \subseteq G$ *ein Normalteiler von* G*, dann ist* $(G/U, \cdot)$ *mit der in a. definierten Gruppenoperation eine Gruppe mit neutralem Element* $[e] = U$*. Für* $[x] \in G/U$ *gilt ferner* $[x]^{-1} = \big[x^{-1}\big]$.
Man nennt G/U *die* Faktorgruppe *von* G *nach* U.
Die Restklassenabbildung

$$\pi : G \to G/U : x \mapsto [x]$$

ist ein Gruppenhomomorphismus mit $\mathrm{Ker}(\pi) = U$.

Beweis: a. Sei zunächst $U \trianglelefteq G$ und seien $[x] = xU, [y] = yU \in G/U$. Für beliebige $u, v \in U$ ist zu zeigen:

$$(xu)(yv) \in (xy)U.$$

Wegen $v \in U$ ist das äquivalent zu

$$xuy \in (xy)U.$$

Dies ist nun wiederum gleichwertig zu

$$uy = x^{-1}(xuy) \in x^{-1}(xyU) = yU,$$

was erfüllt ist, da $U \trianglelefteq G$.
Gilt umgekehrt, daß die Operation wohldefiniert ist, und sind $x \in G$ und $u \in U$ gegeben, so ist zu zeigen:

$$xux^{-1} \in U.$$

Da $[xu] = [x]$, gilt

$$U = [e] = \left[x \cdot x^{-1}\right] = [x] \cdot \left[x^{-1}\right] = [xu] \cdot \left[x^{-1}\right] = \left[xux^{-1}\right],$$

was gleichbedeutend zu $xux^{-1} \in U$ ist. Also ist die Untergruppe U ein Normalteiler.

b. Nach Teil a. ist die Gruppenoperation wohldefiniert, und für $[x], [y], [z] \in G/U$ gilt:

$$\big([x] \cdot [y]\big) \cdot [z] = [xy] \cdot [z] = \big[(xy)z\big] = \big[x(yz)\big] = [x] \cdot [yz] = [x] \cdot \big([y] \cdot [z]\big).$$

Außerdem ist $[e] \cdot [x] = [ex] = [x]$, so daß $[e]$ das Neutrale von G/U ist, und es gilt

$$\left[x^{-1}\right] \cdot [x] = \left[x^{-1} \cdot x\right] = [e],$$

und somit besitzt $[x]$ ein Inverses, nämlich $[x]^{-1} = \left[x^{-1}\right]$.
Schließlich folgt aus den Definitionen

$$\pi(xy) = [xy] = [x] \cdot [y] = \pi(x) \cdot \pi(y)$$

und $\mathrm{Ker}(\pi) = \pi^{-1}\big([e]\big) = [e] = U$, so daß π ein Gruppenhomomorphismus mit $\mathrm{Ker}(\pi) = U$ ist.

□

Betrachten wir nun wieder den Spezialfall $G/U = \mathbb{Z}_n$

Definition 5.9
Wir definieren auf $\mathbb{Z}_n$ eine Addition und eine Multiplikation repräsentantenweise, d. h. für $[x], [y] \in \mathbb{Z}_n$ definieren wir

$$[x] + [y] := [x + y] \quad \text{und} \quad [x] \cdot [y] = [x \cdot y].$$

Satz 5.10 a. *Die Addition und die Multiplikation sind wohldefiniert, d. h. unabhängig von der Wahl der Repräsentanten.*

b. $(\mathbb{Z}_n, +)$ *ist eine abelsche Gruppe mit* $[0]$ *als neutralem Element. Zu* $[x] \in \mathbb{Z}_n$ *ist* $[-x] \in \mathbb{Z}_n$ *das Inverse.*

c. $\big(\mathbb{Z}_n \setminus \{[0]\}, \cdot\big)$ *ist eine Halbgruppe mit neutralem Element* $[1]$.

Beweis: a. Die Wohldefiniertheit von $+$ folgt aus Satz 5.8, da $n\mathbb{Z}$ ein Normalteiler von $\mathbb{Z}$ ist.

Bleibt also die Wohldefiniertheit von $\cdot$ zu zeigen. Seien $x' = x + na$ und $y' = y + nb$ mit $a, b, x, y \in \mathbb{Z}$ gegeben, d. h. $[x'] = [x]$ und $[y'] = [y]$. Dann gilt

$$x'y' = xy + n(xb + ay + nab),$$

also $[x'y'] = [xy]$.

b./c. Die nachzuprüfenden Axiome lassen sich mittels der Definition sofort daraus ableiten, daß $(\mathbb{Z}, +)$ eine abelsche Gruppe ist und $(\mathbb{Z} \setminus \{0\}, \cdot)$ eine Halbgruppe.

□

Beispiel 5.11

Für Gruppen und Halbgruppen kleiner Ordnung, d. h. mit wenig Elementen, ist es sinnvoll sog. Verknüpfungstafeln aufzustellen, aus denen zu je zwei gegebenen Elementen die Summe bzw. das Produkt der beiden Elemente abgelesen werden kann. Im Falle von $\mathbb{Z}_n$ erhalten wir für $n = 2, 3, 4$ die folgenden Verknüpfungstafeln.

$n = 2$**:**

$+$	$[0]$	$[1]$
$[0]$	$[0]$	$[1]$
$[1]$	$[1]$	$[0]$

$\cdot$	$[0]$	$[1]$
$[0]$	$[0]$	$[0]$
$[1]$	$[0]$	$[1]$

$n = 3$**:**

$+$	$[0]$	$[1]$	$[2]$
$[0]$	$[0]$	$[1]$	$[2]$
$[1]$	$[1]$	$[2]$	$[0]$
$[2]$	$[2]$	$[0]$	$[1]$

$\cdot$	$[0]$	$[1]$	$[2]$
$[0]$	$[0]$	$[0]$	$[0]$
$[1]$	$[0]$	$[1]$	$[2]$
$[2]$	$[0]$	$[2]$	$[1]$

$n = 4$**:**

$+$	$[0]$	$[1]$	$[2]$	$[3]$
$[0]$	$[0]$	$[1]$	$[2]$	$[3]$
$[1]$	$[1]$	$[2]$	$[3]$	$[0]$
$[2]$	$[2]$	$[3]$	$[0]$	$[1]$
$[3]$	$[3]$	$[0]$	$[1]$	$[2]$

$\cdot$	$[0]$	$[1]$	$[2]$	$[3]$
$[0]$	$[0]$	$[0]$	$[0]$	$[0]$
$[1]$	$[0]$	$[1]$	$[2]$	$[3]$
$[2]$	$[0]$	$[2]$	$[0]$	$[2]$
$[3]$	$[0]$	$[3]$	$[2]$	$[1]$

Der folgende Isomorphiesatz ist einer von mehreren wichtigen Isomorphiesätzen der Gruppentheorie, er ist der grundlegendste.

Satz 5.12 (Homomorphiesatz)

Ist $f : G \to H$ *ein Gruppenhomomorphismus, dann gilt:*

a. $\mathrm{Ker}(f)$ *ist ein Normalteiler von* G.

b. *Die durch* f *induzierte Abbildung*

$$\tilde{f} : G/\mathrm{Ker}(f) \to \mathrm{Im}(f) : [x] \mapsto f(x)$$

ist wohldefiniert und ein Isomorphismus.

Beweis: a. Sind $u \in \mathrm{Ker}(f)$ und $x \in G$, so gilt:

$$f\big(xux^{-1}\big) = f(x)f(u)f\big(x^{-1}\big) = f(x)e_H f\big(x^{-1}\big) = f(x)f\big(x^{-1}\big)$$

$$= f\big(xx^{-1}\big) = f(e_G) = e_H.$$

Mithin gilt $xux^{-1} \in \operatorname{Ker}(f)$, und $\operatorname{Ker}(f)$ ist ein Normalteiler von G, da $\operatorname{Ker}(f) \subseteq G$ nach Proposition 3.17 bereits eine Untergruppe von G ist.

b. Wir zeigen zunächst, daß $\tilde{f}$ wohldefiniert ist. Sei dazu $[x] = \big[x'\big] \in G/\operatorname{Ker}(f)$ gegeben. Dann gilt also $x^{-1}x' \in \operatorname{Ker}(f)$ und damit

$$e_H = f\big(x^{-1}x'\big) = f\big(x^{-1}\big)f\big(x'\big) = \big(f(x)\big)^{-1}f\big(x'\big).$$

Mithin gilt $f(x) = f\big(x'\big)$, und $\tilde{f}$ ist somit wohldefiniert.
Für $[x], [y] \in G/\operatorname{Ker}(f)$ gilt ferner

$$\tilde{f}\big([x] \cdot [y]\big) = \tilde{f}\big([xy]\big) = f(xy) = f(x)f(y) = \tilde{f}\big([x]\big) \cdot \tilde{f}\big([y]\big).$$

Also ist $\tilde{f}$ ein Gruppenhomomorphismus.
$\tilde{f}$ ist offensichtlich surjektiv. Bleibt also noch zu zeigen, daß $\tilde{f}$ injektiv ist. Seien dazu $[x], [y] \in G/\operatorname{Ker}(f)$ mit $f(x) = \tilde{f}\big([x]\big) = \tilde{f}\big([y]\big) = f(y)$, so gilt

$$e_H = \big(f(x)\big)^{-1}f(y) = f\big(x^{-1}\big)f(y) = f\big(x^{-1}y\big).$$

D. h. $x^{-1}y \in \operatorname{Ker}(f)$, also $[x] = [y]$. Mithin ist $\tilde{f}$ injektiv.

□

Bemerkung 5.13
Betrachten wir den surjektiven Gruppenhomomorphismus $\operatorname{sgn} : \mathbb{S}_n \to \{-1, 1\}$, vgl. Bemerkung 4.13, dann gilt $\operatorname{Ker}(\operatorname{sgn}) = \mathbb{A}_n$. Mithin folgt aus dem Homomorphiesatz 5.12 insbesondere $|\mathbb{S}_n/\mathbb{A}_n| = |\{-1, 1\}| = 2$. Da nach dem Satz von Lagrange, siehe Aufgabe 3.30, zudem gilt $|\mathbb{S}_n/\mathbb{A}_n| = \frac{|\mathbb{S}_n|}{|\mathbb{A}_n|}$ erhalten wir so einen zweiten Beweis der Aussage von Lemma 4.15.

AUFGABEN

Aufgabe 5.14 (Normalteiler)
Es sei $(G, \cdot)$ eine Gruppe, $N \subseteq G$ eine Untergruppe von G. Zeige die Äquivalenz der folgenden Aussagen:

a. $N \trianglelefteq G$.
b. $x^{-1}nx \in N$ für alle $x \in G$ und $n \in N$.
c. $x^{-1}Nx \subseteq N$ für alle $x \in G$.
d. $xN = Nx$ für alle $x \in G$.
e. $(xN) \cdot (yN) = (xy)N$ für alle $x, y \in G$.

Aufgabe 5.15
Zeige durch ein Gegenbeispiel, daß für eine Teilmenge N einer Gruppe $(G, \cdot)$ aus $x^{-1}nx \in N$ für alle $x \in G$ noch nicht folgt, daß N ein Normalteiler ist.

Aufgabe 5.16
Zeige, die Kleinsche Vierergruppe $\mathbb{K}_4$ aus Aufgabe 4.20 ist ein Normalteiler von $\mathbb{S}_4$.

Aufgabe 5.17
Es sei $(G, \cdot)$ eine Gruppe und $N \subseteq G$ eine Untergruppe von G mit der Eigenschaft $|G/N| = 2$, dann ist N ein Normalteiler von G.

Aufgabe 5.18
Es sei $(G, \cdot)$ eine Gruppe und $n > 0$ eine (fest gegebene) natürliche Zahl. Zeige, ist $N \subseteq G$ die einzige Untergruppe von G mit $|N| = n$, dann ist N ein Normalteiler von G.

Aufgabe 5.19
Ist $(G, \cdot)$ eine Gruppe, $U \subseteq G$ eine Untergruppe und $N \trianglelefteq G$. Zeige:

a. $U \cdot N := \{un \mid u \in U, n \in N\}$ ist eine Untergruppe von G.
b. $N \trianglelefteq UN$.
c. $U \cap N \trianglelefteq U$.

Aufgabe 5.20 (1. Isomorphiesatz)
Ist $(G, \cdot)$ eine Gruppe, $U \subseteq G$ eine Untergruppe und $N \trianglelefteq G$. Dann ist

$$U/U \cap N \cong UN/N$$

Aufgabe 5.21 (2. Isomorphiesatz)
Es seien $(G, \cdot)$ eine Gruppe, $M \subseteq N \subseteq G$ zwei Normalteiler von G. Zeige, dann ist auch N/M ein Normalteiler von G/M und es gilt:

$$(G/M)/(N/M) \cong G/N.$$

Aufgabe 5.22
Es seien G und H endliche Gruppen von teilerfremder (siehe Definition 6.1) Ordnung. Zeige, daß es genau einen Homomorphismus von G nach H gibt.
Hinweis: Verwende den Homomorphiesatz 5.12 und den Satz von Lagrange 3.30.

6. Der Euklidische Algorithmus

Neben dem gaußschen Algorithmus zur Lösung linearer Gleichungssysteme, siehe Abschnitt 9, ist der Euklidische Algorithmus zur Bestimmung des größten gemeinsamen Teilers der wichtigste elementare mathematische Algorithmus.

Definition 6.1
Für $a, b, p \in \mathbb{Z}$ definiert man:

a. a *teilt* b (oder a ist *Teiler* von b), falls es ein $c \in \mathbb{Z}$ gibt mit $b = ac$. Wir schreiben kurz: $a \mid b$.
b. p heißt *Primzahl*, falls $p > 1$ und falls p außer p und 1 keine weiteren positiven Teiler besitzt.
c. a und b heißen *teilerfremd* (oder *coprim*), falls 1 und -1 die einzigen gemeinsamen Teiler von a und b sind.
d. p heißt *größter gemeinsamer Teiler* von a und b, in Zeichen $p = \mathrm{ggT}(a, b)$, falls gilt:
   (i) $p \mid a$ und $p \mid b$, und
   (ii) $\forall\, p' \in \mathbb{Z}$ mit $p' \mid a$ und $p' \mid b$ gilt: $p' \mid p$.
   D. h. falls p ein gemeinsamer Teiler von a und b ist, so daß jeder andere Teiler von a und b auch ein Teiler von p ist.

Bemerkung 6.2
Für $a, b \in \mathbb{Z} \setminus \{0\}$ ist der größte gemeinsame Teiler $\mathrm{ggT}(a, b)$ bis aufs Vorzeichen eindeutig bestimmt.

Denn, sind $p \in \mathbb{Z}$ und $p' \in \mathbb{Z}$ beides größte gemeinsame Teiler von a und b, so gilt $p \mid p'$ und $p' \mid p$, also gibt es $c, c' \in \mathbb{Z}$ mit $p = c'p'$ und $p' = cp$. Mithin gilt $p = c'cp$, also $c'c = 1$, woraus $c = c' \in \{-1, 1\}$ folgt.

Algorithmus 6.3 (Euklidischer Divisionsalgorithmus)
INPUT: $a, b \in \mathbb{Z} \setminus \{0\}$.

OUTPUT: $\mathrm{ggT}(a, b)$.

1. Schritt: Falls $|a| < |b|$, so vertausche a und b.
2. Schritt: $r = a \pmod b$ ("Rest der Division von a durch b")
3. Schritt: [5]
```
while (r > 0)
{
  a = b;
  b = r;
  r = a (mod b);
}
```
4. Schritt: `return (b);`

Beweis: O. E. $b \geq 0$. Im Algorithmus erhalten wir durch Division mit Rest, siehe Lemma 5.5:

$$a = bq + r, \quad 0 \leq r < b.$$

Setze $r_0 := b$ und $r_1 := r$, dann erhalten wir im 3. Schritt bei Division mit Rest:

$$\begin{aligned} r_0 &= r_1 q_1 + r_2, & 0 \leq r_2 < r_1, \\ r_1 &= r_2 q_2 + r_3, & 0 \leq r_3 < r_2, \\ &\vdots \\ r_{n-2} &= r_{n-1} q_{n-1} + r_n, & 0 \leq r_n < r_{n-1}, \\ r_{n-1} &= r_n q_n + r_{n+1}, & r_{n+1} = 0. \end{aligned}$$

Wegen $r_1 > r_2 > r_3 > \ldots$ stoppt der Algorithmus mit $r_{n+1} = 0$ und gibt r_n zurück.

Behauptung: $r_n = \mathrm{ggT}(a, b)$.

Beweis durch Induktion über n, d. h. die Anzahl der Iterationen in der While-Schleife.

Induktionsanfang: $n = 0$. Dann ist $0 = r_{n+1} = r_1$, also $b \mid a$ und $r_0 = b = \mathrm{ggT}(a, b)$.

Induktionsschluß: Sei nun $n \geq 1$ und die Behauptung gelte für alle Paare von Zahlen, für die der Algorithmus nach $n - 1$ Iterationen stoppt. Die Betrachtung der letzten $n - 1$ Schritte unseres Algorithmus liefert mithin durch

[5] Die Beschreibung des Algorithmus orientiert sich an der Sprache des Computeralgebra-Systems SINGULAR.

Anwendung der Induktionsvoraussetzung auf r_0 und r_1:

$$r_n = ggT(r_0, r_1) = ggT(b, r).$$

Wegen $a = bq + r$ ist r_n mithin auch ein Teiler von a. Ist nun $r' \in \mathbb{Z}$ ein weiterer Teiler von a und b, dann gilt:

$$r' \mid (a - bq) = r,$$

also folgt:

$$r' \mid r_n.$$

Mithin ist $r_n = ggT(a, b)$. □

Korollar 6.4
Zu $a, b \in \mathbb{Z} \setminus \{0\}$ *gibt es ganze Zahlen* $x, y \in \mathbb{Z}$ *mit*

$$ggT(a, b) = xa + yb.$$

Beweis: Wir führen den Beweis wieder durch Induktion über die Anzahl n der Iterationen im Euklidischen Algorithmus und verwenden die Notation von 6.3

Induktionsanfang: $n = 0$. Dann gilt $a = bq$ und $b = ggT(a, b) = 0 \cdot a + 1 \cdot b$.

Induktionsschluß: Sei die Aussage wieder richtig für $n - 1$. Dann gilt nach Induktionsvoraussetzung, es gibt $x', y' \in \mathbb{Z}$ mit

$$r_n = ggT(r_0, r_1) = x'r_0 + y'r_1 = x'b + y'(a - bq) = y'a + \big(x' - y'q\big)b.$$

D. h. mit $x = y'$ und $y = (x' - y'q)$ ist die Behauptung bewiesen. □

Aus dem Beweis des Korollars läßt sich unmittelbar der erweiterte Euklidische Algorithmus als rekursiver Algorithmus zur Bestimmung des $ggT(a, b)$ sowie der Zahlen x und y ableiten.

Algorithmus 6.5 (Erweiterter Euklidischer Algorithmus)
INPUT: $a, b \in \mathbb{Z} \setminus \{0\}$, $|a| \geq |b|$.

OUTPUT: $\big[ggT(a, b), x, y\big]$ mit $ggT(a, b) = xa + yb$.

1. Schritt: $r = a \pmod b$ (“Rest der Division von a durch b”)

2. Schritt:

```
if (r == 0)
{
  return ([b,0,1]);
}
else
{
  list l = ggT(b,r);

  return ([l[1],l[3],l[2] − l[3] ∗ (a − r)/b]);
}
```

Beispiel 6.6
Es seien $a = 5$ und $b = 3$. Der Euklidische Algorithmus liefert:

$$\begin{aligned} 5 &= 1 \cdot 3 + 2 \\ 3 &= 1 \cdot 2 + 1 \\ 2 &= 2 \cdot 1 + 0 \quad \Rightarrow \ 1 = \mathrm{ggT}(3,5). \end{aligned}$$

Der erweiterte Euklidische Algorithmus liefert

$$1 = 3 - 2 = 3 - (5 - 3) = 2 \cdot 3 + (-1) \cdot 5.$$

Korollar 6.7
Für $a, b \in \mathbb{Z} \setminus \{0\}$ *sind äquivalent:*

a. a *und* b *sind teilerfremd.*
b. $\mathrm{ggT}(a,b) = \pm 1$.
c. $\exists\, x, y \in \mathbb{Z} \ : \ 1 = xa + yb$. *(*"Bézout Identität"*)*

Beweis: Die Äquivalenz von a. und b. ist klar, und die Implikation "b. $\Rightarrow$ c." folgt aus Korollar 6.4.

Es bleibt also "c. $\Rightarrow$ b." zu zeigen. Seien dazu $c = \mathrm{ggT}(a,b)$ und $x, y \in \mathbb{Z}$ mit $1 = xa + yb$. Da c sowohl a als auch b teilt, folgt

$$c \mid xa + yb = 1,$$

und mithin ist $c \in \{-1, 1\}$. □

Satz 6.8
Sei $n \in \mathbb{Z}$, $n > 1$.

a. *Es seien* $[0] \neq [a] \in \mathbb{Z}_n$. *Genau dann gibt es ein* $[b] \in \mathbb{Z}_n$ *mit* $[a] \cdot [b] = [1]$, *wenn* a *und* n *teilerfremd sind.*
b. *Genau dann ist* $(\mathbb{Z}_n \setminus \{[0]\}, \cdot)$ *eine Gruppe, wenn* n *eine Primzahl ist.*

Beweis: a. Für $[0] \neq [a] \in \mathbb{Z}_n$ gilt wegen Korollar 6.7:

$$\exists\, [b] \in \mathbb{Z}_n \ : \ [a] \cdot [b] = [1] \Leftrightarrow \ \exists\, b \in \mathbb{Z} \ : \ ab - 1 \in n\mathbb{Z}$$

$$\Leftrightarrow \ \exists\, b, c \in \mathbb{Z} \ : \ ab - nc = 1 \Leftrightarrow \ \mathrm{ggT}(a,n) = \pm 1.$$

b. $\big(\mathbb{Z}_n \setminus \{[0]\}, \cdot\big)$ ist genau dann eine Gruppe, wenn jedes $[0] \neq [a] \in \mathbb{Z}_n$ ein multiplikatives Inverses besitzt. Wegen Teil a. und Bemerkung 5.4 ist das gleichwertig dazu, daß n teilerfremd zu den Zahlen $1, \ldots, n-1$ ist, mit anderen Worten, daß n eine Primzahl ist.

□

Bemerkung 6.9
Der erweiterte Euklidische Algorithmus angewandt auf a und n berechnet das Inverse zu $[a]$ in $\mathbb{Z}_n$, falls es existiert.

Lemma 6.10
Seien $a, b_1, \ldots, b_r \in \mathbb{Z} \setminus \{0\}$ *mit* $\mathrm{ggT}(a, b_i) = \pm 1$ *für* $i = 1, \ldots, r$. *Dann gilt*

$$\mathrm{ggT}(a, b_1 \cdots b_r) = \pm 1.$$

Beweis: Nach Korollar 6.4 gibt es ganze Zahlen $x_i, y_i \in \mathbb{Z}$ mit

$$1 = x_i b_i + y_i a, \quad \text{für } i = 1, \ldots, r.$$

Folglich gilt:

$$(x_1 \cdots x_r)(b_1 \cdots b_r) = (1 - y_1 a) \cdots (1 - y_r a) = 1 - ya,$$

für ein geeignetes $y \in \mathbb{Z}$. Die Behauptung folgt dann aus Korollar 6.7. □

Proposition 6.11 a. *Sind* $a_1, \ldots, a_r \in \mathbb{Z} \setminus \{0\}$ *paarweise teilerfremd und* $b \in \mathbb{Z} \setminus \{0\}$ *mit* $a_i \mid b$ *für* $i = 1, \ldots, r$*, so gilt:*

$$a_1 \cdots a_r \mid b.$$

b. *Ist* p *eine Primzahl und sind* $a_1, \ldots, a_r \in \mathbb{Z} \setminus \{0\}$ *mit* $p \mid a_1 \cdots a_r$*, so gibt es ein* $i \in \{1, \ldots, r\}$ *mit*

$$p \mid a_i.$$

Beweis: a. Wir führen den Beweis durch Induktion über r.
Induktionsanfang: Für $r = 1$ ist die Aussage trivial. Sei nun $r = 2$. Dann gibt es $x, y \in \mathbb{Z}$ mit

$$xa_1 + ya_2 = 1.$$

Ferner existieren ganze Zahlen $c_1, c_2 \in \mathbb{Z}$ mit $b = c_i a_i$, $i = 1, 2$. Daraus folgt:

$$b = a_1 c_1 (xa_1 + ya_2) = a_1 (a_1 c_1) x + (a_1 a_2) c_1 y$$

$$= a_1 (a_2 c_2) x + (a_1 a_2) c_1 y = a_1 a_2 (c_2 x + c_1 y).$$

Induktionsschluß: Sei nun $r \geq 2$ und die Behauptung sei für r bereits bewiesen. Wir setzen $a' := a_1 \cdots a_r$. Mit Lemma 6.10 folgt

$$\mathrm{ggT}\left(a_{r+1}, a'\right) = 1.$$

Nach Induktionsvoraussetzung gilt aber $a' \mid b$, und aus dem Fall $r = 2$ angewandt auf a' und a_{r+1} folgt somit:

$$a_1 \cdots a_{r+1} = a' \cdot a_{r+1} \mid b.$$

b. Da p eine Primzahl ist, gilt für $i \in \{1, \ldots, r\}$ entweder $\mathrm{ggT}(a_i, p) = \pm 1$ oder $\mathrm{ggT}(a_i, p) = \pm p$. Würde für alle $i = 1, \ldots, r$ gelten $\mathrm{ggT}(a_i, p) = \pm 1$, so würde mit Lemma 6.10 folgen $\mathrm{ggT}(a_1 \cdots a_r, p) = \pm 1$, im Widerspruch zur Voraussetzung.

□

Satz 6.12
In $\mathbb{Z}$ gilt die eindeutige Primfaktorzerlegung, d. h. jede Zahl $a \in \mathbb{Z}$ mit $a > 1$ läßt sich als Produkt $a = p_1 \cdots p_r$ von Primzahlen $p_1, \ldots, p_r$ darstellen und diese Darstellung ist (bis auf die Reihenfolge) eindeutig.

Beweis: Wir zeigen zunächst die Existenz der Zerlegung. Nehmen wir an, nicht jede Zahl $a > 1$ sei Produkt von endlich vielen Primzahlen. Dann gibt es eine kleinste Zahl $a > 1$, die nicht Produkt von endlich vielen Primzahlen ist, und insbesondere ist a keine Primzahl. Aber dann gibt es Zahlen $1 < a_1, a_2 < a$ mit $a = a_1 \cdot a_2$. Da a eine minimale "schlechte" Zahl war, sind a_1 und a_2 mithin das Produkt von endlich vielen Primzahlen, und somit trifft selbiges auch auf a zu, im Widerspruch zur Annahme.

Zum Beweis der Eindeutigkeit seien Primzahlen $p_1, \ldots, p_r$ und $q_1, \ldots, q_s$ gegeben mit

$$a = p_1 \cdots p_r = q_1 \cdots q_s.$$

Dann gilt $p_r \mid q_1 \cdots q_s$, und aus Proposition 6.11 b. folgt, da p_r eine Primzahl ist, daß es ein $i \in \{1, \ldots, s\}$ gibt mit $p_r | q_i$. Da q_i eine Primzahl ist und $p_r > 1$, gilt also $p_r = q_i$. Durch Umnumerieren können wir $i = s$ erreichen. Induktion über r liefert dann die Behauptung. □

Bemerkung 6.13
Der Satz gibt keinerlei Hinweis darauf, wie man die Primfaktorzerlegung von a findet. Der Beweis ist nicht konstruktiv.

Zwar gibt es Algorithmen, Zahlen in Primfaktoren zu zerlegen, aber alle bekannten Algorithmen sind für große Zahlen sehr langsam. In der Tat beruhen heute viele Verschlüsselungsmethoden auf der Tatsache, daß es einfach ist, große Zahlen zu multiplizieren, aber sehr schwierig, sie wieder in ihre Faktoren zu zerlegen.

Satz 6.14 (Euklid, $\approx$ 300 v. Chr.)
Es gibt unendlich viele Primzahlen.

Beweis: Annahme, es gäbe nur endlich viele Primzahlen, $p_1, \ldots, p_r$. Nach Satz 6.12 ist die Zahl $c = p_1 \cdots p_r + 1$ Produkt von Primzahlen. Also existiert ein $i \in \{1, \ldots, r\}$ mit

$$p_i \mid p_1 \cdots p_r + 1,$$

aber dann gilt auch $p_i \mid 1$ im Widerspruch zu $p_i > 1$. □

Lemma 6.15
Es seien $a, n \in \mathbb{Z} \setminus \{0\}$ mit $\mathrm{ggT}(a, n) = \pm 1$. Dann hat die Kongruenzgleichung

$$ax \equiv b \pmod{n}$$

für jedes $b \in \mathbb{Z}$ eine Lösung $x \in \mathbb{Z}$, die modulo n eindeutig ist.

Beweis: Die Behauptung ist, daß es genau eine Klasse $[x] \in \mathbb{Z}_n$ mit $[a] \cdot [x] = [b]$ gibt.

Betrachte die Abbildung

$$\varphi_{[a]} : \mathbb{Z}_n \to \mathbb{Z}_n : [x] \mapsto [a] \cdot [x] = [ax].$$

Behauptung: $\varphi_{[a]}$ ist ein Isomorphismus.

Man sieht leicht, daß $\varphi_{[a]}$ ein Gruppenhomomorphismus von $(\mathbb{Z}_n, +)$ ist. Es bleibt also zu zeigen, daß $\varphi_{[a]}$ bijektiv ist. Da $\mathbb{Z}_n$ endlich ist, reicht es, zu zeigen, daß $\varphi_{[a]}$ injektiv ist,[6] d. h. $\operatorname{Ker}\big(\varphi_{[a]}\big) = \big\{[0]\big\}$, siehe Lemma 3.18. Nach Satz 6.8 gibt es ein $[c] \in \mathbb{Z}_n$ mit $[c][a] = [1]$, da $\operatorname{ggT}(a, n) = \pm 1$. Sei nun $[x] \in \operatorname{Ker}\big(\varphi_{[a]}\big)$, dann gilt $[a][x] = [0]$, und damit

$$[x] = [1][x] = \big([c][a]\big)[x] = [c]\big([a][x]\big) = [c][0] = [0].$$

Also ist $\operatorname{Ker}\big(\varphi_{[a]}\big) = \big\{[0]\big\}$.

Da nun $\varphi_{[a]}$ bijektiv ist, gibt es zu $b \in \mathbb{Z}$ beliebig jeweils genau ein $[x] \in \mathbb{Z}_n$ mit

$$[b] = \varphi_{[a]}\big([x]\big) = [a][x].$$

□

Satz 6.16 (Chinesischer Restsatz)
Seien $n_1, \ldots, n_r \in \mathbb{Z} \setminus \{0\}$ paarweise teilerfremd. Dann existiert zu beliebig vorgegebenen ganzen Zahlen $a_1, \ldots, a_r \in \mathbb{Z}$ eine Lösung $x \in \mathbb{Z}$ des Kongruenzgleichungssystems

$$\begin{aligned} x &\equiv a_1 \pmod{n_1}, \\ x &\equiv a_2 \pmod{n_2}, \\ &\vdots \\ x &\equiv a_r \pmod{n_r}. \end{aligned} \tag{7}$$

Diese Lösung ist modulo $N := n_1 \cdots n_r$ eindeutig bestimmt.

Außerdem gilt, wenn $N_i := \frac{N}{n_i}$ und wenn x_i die Kongruenzgleichung

$$N_i x_i \equiv a_i \pmod{n_i}, \quad \textit{d. h.} \ [x_i] = [a_i][N_i]^{-1} \in \mathbb{Z}_{n_i},$$

löst, $i = 1, \ldots, r$, dann ist

$$x = \sum_{i=1}^{r} N_i x_i$$

eine Lösung!

Beweis: Wir zeigen zunächst die Eindeutigkeit der Lösung modulo N. Seien $x, x' \in \mathbb{Z}$ zwei Lösungen von (7). Dann gilt für $i = 1, \ldots, r$

$$x - a_i, x' - a_i \in n_i\mathbb{Z}.$$

[6] Siehe eine Aufgabe 2.46.

Damit gilt aber $x - x' \in n_i\mathbb{Z}$, d. h. $n_i \mid (x - x')$, für alle $i = 1, \ldots, r$. Aus Proposition 6.11 folgt dann $N \mid (x - x')$, d. h. $x - x' \in N\mathbb{Z}$, und damit

$$x \equiv x' \pmod{N}.$$

Es bleibt also die Existenz einer Lösung x zu zeigen. Da N_i und n_i teilerfremd sind, existiert nach Lemma 6.15 eine Lösung $x_i \in \mathbb{Z}$ für die Gleichung

$$N_i x_i \equiv a_i \pmod{n_i}.$$

Setze nun

$$x := \sum_{i=1}^{r} N_i x_i.$$

Wegen $n_i \mid N_j$ für $j \neq i$ gilt:

$$x \equiv N_i x_i \pmod{n_i} \equiv a_i \pmod{n_i}.$$

□

In der Praxis wählt man die n_i als paarweise verschiedene Primzahlen. Falls man vorher abschätzen kann, daß für die auszurechnende Lösung x gilt $x \leq N$, dann ist x sogar eindeutig und nicht nur eindeutig modulo N.

Beispiel 6.17

Gegeben sei das folgende Kongruenzgleichungssystem:

$$\begin{aligned} x &\equiv a_1 = 1 \pmod{2}, \\ x &\equiv a_2 = 2 \pmod{3}, \\ x &\equiv a_3 = 4 \pmod{7}. \end{aligned}$$

Es sind $n_1 = 2, n_2 = 3, n_3 = 7$ paarweise teilerfremd, und $N = 2 \cdot 3 \cdot 7 = 42$, $N_1 = 21$, $N_2 = 14$ und $N_3 = 6$.

Die Berechnung der Inversen von $[N_i]$ in $\mathbb{Z}_{n_i}$ geschieht mit Hilfe des erweiterten Euklidischen Algorithmus. Da n_i und N_i teilerfremd sind, gilt nach Korollar 6.7

$$x_i N_i + y_i n_i = 1$$

für geeignete $x_i \in \mathbb{Z}$ (und $y_i \in \mathbb{Z}$, die hier nicht interessieren).

$$\begin{aligned} [21]^{-1} = [1]^{-1} = [1] \in \mathbb{Z}_2 &\Rightarrow x_1 = a_1 \cdot 1 = 1, \\ [14]^{-1} = [2]^{-1} = [2] \in \mathbb{Z}_3 &\Rightarrow x_2 = a_2 \cdot 2 = 4, \\ [6]^{-1} = [6] \in \mathbb{Z}_7 \qquad &\Rightarrow x_3 = a_3 \cdot 6 = 24. \end{aligned}$$

Es folgt:

$$x = 21 \cdot 1 + 14 \cdot 4 + 6 \cdot 24 = 221 \equiv 11 \pmod{42}.$$

Also ist $x = 11$ die modulo 42 eindeutig bestimmte Lösung.

Bemerkung 6.18

Angewandt wird der Chinesische Restsatz wie folgt. Man führt eine Rechnung statt mit ganzen Zahlen r-mal modulo verschiedener Primzahlen $p_1, \ldots, p_r$ durch. Für jedes p_i sei das Ergebnis a_i. Man bestimmt mit dem erweiterten Euklidischen Algorithmus x_i und mittels Chinesischem Restsatz x. Dann

stimmt x mit dem wirklichen Ergebnis modulo $p_1 \cdots p_r$ überein. Weiß man, daß das wirkliche Ergebnis, d. h. das Ergebnis bei Durchführung der Rechnung mit ganzen Zahlen, kleiner als $p_1 \cdots p_r$ ist, dann ist x das gesuchte Ergebnis.

Diese Art ist oft viel schneller, als direkt mit ganzen Zahlen zu rechnen (z. B. bei der Berechnung großer ganzzahliger Determinanten).

In SINGULAR liefert `prime(n);` die größte Primzahl p mit $2 \leq p \leq \min\{n, 32003\}$. Mit `LIB ''general.lib'';` und `primes(n,m);` erhält man alle Primzahlen im Intervall $[\texttt{prime(n)}, m]$, sofern $2 \leq n < m \leq 32003$.

AUFGABEN

Aufgabe 6.19

Bestimme den größten gemeinsamen Teiler der Zahlen 10219176 und 2377188 mit Hilfe des euklidischen Algorithmus.

Aufgabe 6.20 (ggT)

Es seien $a_1, \ldots, a_n \in \mathbb{Z}$, $n \geq 2$. Eine Zahl $g \in \mathbb{Z}$ heißt größter gemeinsamer Teiler von $a_1, \ldots, a_n$ (kurz: $\mathrm{ggT}(a_1, \ldots, a_n)$), falls gilt:

(i) $g \mid a_i$ für alle $i = 1, \ldots, n$, und
(ii) für alle $g' \in \mathbb{Z}$ mit $g' \mid a_i$ für alle $i = 1, \ldots, n$ gilt $g' \mid g$.

a. Zeige, es gilt $\mathrm{ggT}(a_1, \ldots, a_n) = \mathrm{ggT}\big(\mathrm{ggT}(a_1, \ldots, a_{n-1}), a_n\big)$, falls $n > 2$.[7]
b. Zeige, $g = \mathrm{ggT}(a_1, \ldots, a_n)$ genau dann, wenn $g\mathbb{Z} = a_1\mathbb{Z} + \cdots + a_n\mathbb{Z} := \{a_1x_1 + \cdots a_nx_n \mid x_1, \ldots, x_n \in \mathbb{Z}\}$.
c. Genau dann besitzt die *diophantische Gleichung* $a_1x_1 + \cdots + a_nx_n = b$ (mit $b \in \mathbb{Z}$) eine Lösung $(x_1, \ldots, x_n) \in \mathbb{Z}^n$, wenn $\mathrm{ggT}(a_1, \ldots, a_n) \mid b$.

Aufgabe 6.21 (kgV)

Es seien $a, b \in \mathbb{Z}$. Eine Zahl $k \in \mathbb{Z}$ heißt *kleinstes gemeinsames Vielfaches* von a und b (kurz $\mathrm{kgV}(a, b)$), wenn gilt:

(i) $a \mid k$ und $b \mid k$, und
(ii) für alle $k' \in \mathbb{Z}$ mit $a \mid k'$ und $b \mid k'$ gilt $k \mid k'$.

Zeige, $k \in \mathbb{Z}$ ist genau dann $\mathrm{kgV}(a, b)$, wenn $k\mathbb{Z} = a\mathbb{Z} \cap b\mathbb{Z}$.

Aufgabe 6.22

Für $a, b \in \mathbb{N} \setminus \{0\}$ seien $a = p_1^{\nu_1} \cdots p_r^{\nu_r}$ und $b = p_1^{\mu_1} \cdots p_r^{\mu_r}$ mit $\nu_i, \mu_i \in \mathbb{N}$, $p_i \neq p_j$ für $i \neq j$, Primfaktorzerlegungen von a und b. Dann ist $k = p_1^{\lambda_1} \cdots p_r^{\lambda_r}$ mit $\lambda_i = \max\{\nu_i, \mu_i\}$ ein kgV von a und b.

Aufgabe 6.23

Schreibe eine Prozedur `ggT`, die zwei ganze Zahlen a und b einliest und eine Liste zurückgibt, die als ersten Eintrag den größten gemeinsamen Teiler d

[7] Beachte, ein ggT ist nur bis auf's Vorzeichen eindeutig bestimmt, mithin soll die Aussage bedeuten, wenn $h \in \mathbb{Z}$ ein ggT von $a_1, \ldots, a_{n-1}$ ist, dann gilt: genau dann ist $g \in \mathbb{Z}$ ein ggT von h und a_n, wenn g auch ein ggT von $a_1, \ldots, a_n$ ist.

von a und b enthält sowie zwei Zahlen p und q, so daß gilt

$$d = pa + qb.$$

Hinweis: Der Operator `div` in Singular arbeitet zur Zeit noch nicht in der gewünschten Weise, wenn die betragsgrößere Zahl negativ ist! Will man an einer Stelle den Ausdruck `a div b` verwenden, so schreibe man statt dessen `(a-(a mod b)) div b`. Beachte die Sonderfälle $a = 0$ oder $b = 0$ oder $|a| < |b|$!

Aufgabe 6.24

Finde eine ganze Zahl z, so daß $z \equiv a_i \pmod{n_i}$ für $i = 1, 2, 3$, wo $(a_1, a_2, a_3) = (7, -7, 9)$ und $(n_1, n_2, n_3) = (2, 15, 7)$.

Aufgabe 6.25

Schreibe eine Prozedur `restsatz`, die zwei Vektoren `a` und `n` ganzer Zahlen einliest, und, falls die Vektoren die gleiche Anzahl an Komponenten haben und die Einträge in `n` paarweise teilerfremd und nicht Null sind, eine ganze Zahl x zurück gibt, so daß gilt:

$$x \equiv a_i \pmod{n_i}, \quad \forall\, i = 1, \ldots, r,$$

wobei `a`$= [a_1, \ldots, a_r]$ und `n`$= [n_1, \ldots, n_r]$. Sind die Vektoren nicht gleich lang, ist einer der Einträge in `n` Null oder sind die Einträge in `n` nicht paarweise teilerfremd, so soll eine entsprechende Fehlermeldung ausgegeben werden.

Man verwende zur Berechnung von x den folgenden Algorithmus:

INPUT: $n_1, \ldots, n_r \in \mathbb{Z}$ paarweise teilerfremd, $a_1, \ldots, a_r \in \mathbb{Z}$

OUTPUT: $x \in \mathbb{Z}$ mit $x \equiv a_i \pmod{n_i}$ für $i = 1, \ldots, r$

- INITIALISIERUNG: $y := n_1$, $x := a_1$
- HAUPTSCHLEIFE: Für $i = 2, \ldots, r$ tue
  - Berechne $p, q \in \mathbb{Z}$ mittels erweitertem Euklidischen Algorithmus, so daß gilt: $py + qn_i = 1$.
  - $x := pya_i + qn_ix$
  - $y := yn_i$
  - $x := x \pmod{y}$
- Gib x zurück.

7. Ringe und Körper

Der Begriff des Ringes verallgemeinert die ganzen Zahlen, während der Begriff des Körpers die rationalen und reellen Zahlen verallgemeinert. In einem Ring gibt es zwei Operationen, eine Addition und eine Multiplikation, die gewissen, aus den ganzen Zahlen vertrauten, Regeln genügen. Ein Körper ist ein Ring, in dem zusätzlich jedes Element ungleich Null ein Inverses bezüglich der Multiplikation besitzt. Dies ist die intuitive Vorstellung von einem Ring bzw. einem Körper, die exakten Definitionen folgen jetzt.

Definition 7.1 a. Eine Menge R mit zwei zweistelligen Operationen

$$+ : R \times R \to R : (x, y) \mapsto x + y, \qquad (\text{“}\textit{Addition}\text{”})$$

und

$$\cdot : R \times R \to R : (x, y) \mapsto x \cdot y, \qquad (\text{“}\textit{Multiplikation}\text{”})$$

heißt ein *Ring*, falls die folgenden Axiome erfüllt sind

(i) $(R, +)$ ist eine abelsche Gruppe (deren neutrales Element mit 0 bzw. 0_R bezeichnet wird),

(ii) $x \cdot (y \cdot z) = (x \cdot y) \cdot z$ für alle $x, y, z \in R$, (“*Assoziativität der Multiplikation*”)

(iii) $x \cdot (y + z) = x \cdot y + x \cdot z$ und $(y + z) \cdot x = y \cdot x + z \cdot x$ für alle $x, y, z \in R$. (“*Distributivität*”)

b. $(R, +, \cdot)$ heißt ein *Ring mit Eins*, falls es ein Element $1 = 1_R \in R$ gibt mit $1 \cdot x = x \cdot 1 = x$ für alle $x \in R$.

c. $(R, +, \cdot)$ heißt *kommutativ*, falls $x \cdot y = y \cdot x$ für alle $x, y \in R$ gilt.

Beispiel 7.2 a. $(\mathbb{Z}, +, \cdot)$, $(\mathbb{Q}, +, \cdot)$ und $(\mathbb{R}, +, \cdot)$ mit der üblichen Addition und Multiplikation sind kommutative Ringe mit Eins. Für $n > 1$ ist $(n\mathbb{Z}, +, \cdot)$ ein kommutativer Ring ohne Eins.

b. $\mathbb{Z}_n$ ist mit den in Definition 5.9 definierten Operationen ein kommutativer Ring mit Eins [1], wie sich aus Satz 5.10 ergibt.

c. Ist M eine beliebige Menge und $(R, +, \cdot)$ ein Ring, so ist

$$R^M := \{f \mid f : M \to R \text{ ist eine Abbildung}\}$$

mit den punktweise definierten Operationen

$$+ : R^M \times R^M \to R^M : (f, g) \mapsto \big(f + g : M \to R : x \mapsto f(x) + g(x)\big),$$

und

$$\cdot : R^M \times R^M \to R^M : (f, g) \mapsto \big(f \cdot g : M \to R : x \mapsto f(x) \cdot g(x)\big),$$

ein Ring mit der Nullfunktion $0 : M \to R : x \mapsto 0_R$ als neutralem Element der Addition, wie man mit etwas Fleiß nachprüft.

Bemerkung 7.3

Wir werden in Ringen für die Addition stets das Zeichen $+$ und für die Multiplikation das Zeichen $\cdot$ verwenden, auch wenn wir gleichzeitig verschiedene Ringe betrachten. Wir verzichten im Folgenden deshalb darauf, die Ringoperationen jeweils in Form von “ein Ring $(R, +, \cdot)$” anzugeben.

Ist R ein Ring und sind $x, y \in R$, so schreiben wir statt $x + (-y)$ auch kurz $x - y$.

Mit einer neuen Struktur definieren wir auch gleich die strukturerhaltenden Abbildungen.

Definition 7.4
Es seien R und S Ringe. Eine Abbildung $f : R \to S$ heißt *Ringhomomorphismus* oder *Ringmorphismus* oder *Ringabbildung*, falls für alle $x, y \in R$ gilt:

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

Bemerkung 7.5
Ist $f : R \to S$ ein bijektiver Ringhomomorphismus, dann ist auch $f^{-1} : S \to R$ ein Ringhomomorphismus.

Daß f^{-1} mit der Addition verträglich ist, folgt aus Proposition 3.17 c., da f ein Homomorphismus von der abelschen Gruppe $(R, +)$ nach $(S, +)$ ist. Für die Verträglichkeit mit der Multiplikation kopiere man den dortigen Beweis.

Beispiel 7.6 a. Die kanonischen Inklusionen $i_{\mathbb{Z}} : \mathbb{Z} \to \mathbb{Q}$, $i_{\mathbb{Q}} : \mathbb{Q} \to \mathbb{R}$ und $i_{n\mathbb{Z}} : n\mathbb{Z} \to \mathbb{Z}$ sind Ringhomomorphismen.

b. Die Abbildung $\mathbb{Z} \to n\mathbb{Z} : x \mapsto nx$ ist für $n > 1$ kein Ringhomomorphismus, da etwa $n \cdot (1 \cdot 2) \neq (n \cdot 1) \cdot (n \cdot 2)$.

Die Inklusionen führen uns gleich zum nächsten Begriff.

Definition 7.7
Sei R ein Ring und $S \subseteq R$. S heißt ein *Unterring* von R, falls für alle $x, y \in S$ gilt

$$x + y \in S \quad \text{sowie} \quad x \cdot y \in S$$

und falls zudem S mit diesen Operationen selbst ein Ring ist.

Bemerkung 7.8
Ist $S \subseteq R$ ein Unterring des Ringes R, so ist die kanonische Inklusion $i_S : S \to R$ ein Ringhomomorphismus.

Proposition 7.9
Es sei R ein Ring und $\emptyset \neq S \subseteq R$. Genau dann ist S ein Unterring von R, wenn für alle $x, y \in S$ gilt

a. $x + y \in S$,
b. $-x \in S$, *und*
c. $x \cdot y \in S$.

Ist R zudem kommutativ, so ist auch S kommutativ.

Beweis: Dies folgt unmittelbar aus der Definition des Begriffes Ring und Proposition 3.10. □

Bevor wir einige allgemeine Aussagen über Ringe beweisen, definieren wir gleich den Begriff des Körpers. Grob gesprochen sind Körper Ringe, in denen man zusätzlich noch dividieren darf (außer durch Null).

Definition 7.10 a. Eine Menge K mit zwei zweistelligen Operationen $+$ und $\cdot$ heißt ein *Körper*, falls gilt

(i) $(K, +, \cdot)$ ist ein kommutativer Ring mit Eins, und
(ii) $(K \setminus \{0\}, \cdot)$ ist eine abelsche Gruppe.

b. Eine Abbildung zwischen Körpern heißt ein *Körperhomomorphismus*, falls sie ein Ringhomomorphismus ist.
c. Erfüllt $(K, +, \cdot)$ nur die Bedingungen
   (i') $(K, +, \cdot)$ ist ein Ring mit Eins, und
   (ii') $(K \setminus \{0\}, \cdot)$ ist eine Gruppe,
   d. h. die Multiplikation ist nicht notwendig kommutativ, so nennen wir $(K, +, \cdot)$ einen *Schiefkörper*.
d. Eine Teilmenge $F \subseteq K$ eines Körpers K, die ein Unterring ist und für die gilt $(F \setminus \{0\}, \cdot)$ ist eine abelsche Gruppe, heißt ein *Unterkörper* oder *Teilkörper* von K.

Notation 7.11
Ist K ein Körper oder ein Schiefkörper und ist $x \in K \setminus \{0\}$, so bezeichnen wir das multiplikative Inverse zu x mit x^{-1} oder $\frac{1}{x}$. (Vgl. Notation 3.3.)

Bemerkung 7.12
Durch Auflisten und vergleichen der Axiome sieht man unmittelbar, daß eine Menge K mit zwei zweistelligen Operationen $+$ und $\cdot$ genau dann ein Körper ist, wenn gilt:

a. $(K, +)$ ist eine abelsche Gruppe.
b. $(K \setminus \{0\}, \cdot)$ ist eine abelsche Gruppe.
c. Für alle $x, y, z \in K$ gilt:

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

Entsprechend ist K genau dann ein Schiefkörper, wenn gilt:

a. $(K, +)$ ist eine abelsche Gruppe.
b. $(K \setminus \{0\}, \cdot)$ ist eine Gruppe.
c. Für alle $x, y, z \in K$ gilt:

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

Man beachte, daß bei einem Körper aufgrund der Kommutativität der Multiplikation, das zweite Distributivgesetz aus dem ersten folgt, während das bei Schiefkörpern nicht der Fall ist.

Schließlich ist K genau dann ein Ring mit Eins, wenn gilt:

a. $(K, +)$ ist eine abelsche Gruppe.
b. $(K \setminus \{0\}, \cdot)$ ist eine Halbgruppe.
c. Für alle $x, y, z \in K$ gilt:

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

Beispiel 7.13
$(\mathbb{Q}, +, \cdot)$ und $(\mathbb{R}, +, \cdot)$ sind Körper, während $(\mathbb{Z}, +, \cdot)$ kein Körper ist, da etwa die Zahl Zwei in $\mathbb{Z}$ kein Multiplikatives Inverses besitzt.

Bevor wir weitere Beispiele behandeln, notieren wir erst einmal die wichtigsten Rechenregeln für Ringe und Körper.

Lemma 7.14 (Rechenregeln)
Es sei R *ein Ring. Für* $x, y, z, u, v \in R$ *gelten:*

a. $-(-x) = x$,
b. $x + y = z \Leftrightarrow x = z - y$,
c. $-(x + y) = -x - y$,
d. $0 \cdot x = x \cdot 0 = 0$,
e. $(-x) \cdot y = x \cdot (-y) = -(x \cdot y)$,
f. $(-x) \cdot (-y) = x \cdot y$, *und*
g. $x \cdot (y - z) = x \cdot y - x \cdot z$.

Ist R *gar ein Körper, so gelten zusätzlich:*

h. $1 \neq 0$, *insbesondere hat ein Körper mindestens zwei Elemente,*
i. $\left(x^{-1}\right)^{-1} = x$, *für* $x \neq 0$,
j. $x \cdot y = 0 \Leftrightarrow x = 0$ *oder* $y = 0$,
k. $z \cdot x = z \cdot y,\ z \neq 0 \Rightarrow x = y$,
l. $\frac{x}{u} + \frac{y}{v} = \frac{x \cdot v + y \cdot u}{u \cdot v}$, *für* $u, v \neq 0$,
m. $x^2 = 1 \Leftrightarrow x \in \{-1, 1\}$.

Beweis: Die Aussagen a., b., c. und k. folgen unmittelbar aus Lemma 3.4.

d. Für $x \in R$ gilt $0 \cdot x = (0 + 0) \cdot x = 0 \cdot x + 0 \cdot x$, also folgt $0 \cdot x = 0$ mittels der Kürzungsregeln in $(R, +)$. Analog sieht man $x \cdot 0 = 0$.
e. Für $x, y \in R$ gilt wegen d.:

$$x \cdot y + (-x) \cdot y = (x - x) \cdot y = 0 \cdot y = 0,$$

also $-(x \cdot y) = (-x) \cdot y$. Die Gleichheit des Ausdrucks zu $x \cdot (-y)$ folgt analog.
f. Für $x, y \in R$ folgt unter Zuhilfenahme von a. und e.:

$$(-x) \cdot (-y) = -\big(x \cdot (-y)\big) = -\big(-(x \cdot y)\big) = x \cdot y.$$

g. Für $x, y, z \in R$ impliziert e.:

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

h. Sei $e \in R \setminus \{0\}$ das neutrale Element der Multiplikation in der Gruppe $(R \setminus \{0\}, \cdot)$. Dann gilt mit Lemma 3.2:

$$1 = 1 \cdot e = e \in R \setminus \{0\}.$$

i. Es seien $x, y, z \in R$ mit $z \neq 0$ und $z \cdot x = z \cdot y$. Dann gilt:

$$x = 1 \cdot x = \left(z^{-1} \cdot z\right) \cdot x = z^{-1} \cdot (z \cdot x) = z^{-1} \cdot (z \cdot y) = \left(z^{-1} \cdot z\right) \cdot y = 1 \cdot y = y.$$

j. Ist $x = 0$ oder $y = 0$, so gilt mit d. auch $x \cdot y = 0$. Seien also $x, y \neq 0$. Wir müssen zeigen, daß dann auch $x \cdot y \neq 0$. Angenommen, $x \cdot y = 0$. Dann gilt:

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

im Widerspruch zu h.

l./m. Der Beweis der Aussagen sei dem Leser als Übungsaufgabe empfohlen.

□

Lemma 7.15
Der Ring $\mathbb{Z}_n$ ist genau dann ein Körper, wenn n eine Primzahl ist.

Beweis: Nach Beispiel 7.2 ist $\mathbb{Z}_n$ ein kommutativer Ring mit Eins für jedes $n \in \mathbb{N}$, und nach Beispiel 7.13 ist $\mathbb{Z}_0 = \mathbb{Z}$ kein Körper. Ferner ist $\mathbb{Z}_1 = \{[0]\}$ kein Körper nach Teil h. von Lemma 7.14, und für $n > 1$ gilt nach Satz 6.8, $(\mathbb{Z}_n \setminus \{[0]\}, \cdot)$ genau dann eine abelsche Gruppe, wenn n eine Primzahl ist. □

Bemerkung 7.16
Ist R ein Ring und gilt $1_R = 0_R$, dann ist $R = \{0\}$, d. h. der *Nullring* ist der einzige Ring mit Eins, in dem das Nullelement gleich dem Einselement ist.

Ist $R \neq \{0\}$ ein Ring mit Eins, so besitzt 0_R wegen Lemma 7.14 d. und h. niemals ein Inverses Element bezüglich der Multiplikation. Diejenigen Elemente, die ein Inverses besitzen, haben einen speziellen Namen.

Definition 7.17
Sei R ein Ring mit Eins. Ein Element $x \in R$ heißt *invertierbar* oder *Einheit*, falls es ein $y \in R$ gibt mit

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

Die Menge der Einheiten von R wird mit R^* bezeichnet und heißt die *Einheitengruppe* von R. Der Name wird durch das folgende Lemma gerechtfertigt.

Lemma 7.18
Ist R ein Ring mit Eins, so ist $(R^*, \cdot)$ eine Gruppe.

Beweis: Der Beweis stellt eine leicht Übung im Nachweis der Gruppenaxiome unter Verwendung der Rechenregeln in Lemma 7.14 dar. □

Bemerkung 7.19
Ein kommutativer Ring K mit Eins ist also genau dann ein Körper, wenn $K^* = K \setminus \{0\}$.

Entsprechend ist ein Ring K mit Eins genau dann ein Schiefkörper, wenn $K^* = K \setminus \{0\}$.

Korollar 7.20
Für $n \in \mathbb{N}$, $n > 0$, gilt:

$$\mathbb{Z}_n^* = \{[k] \in \mathbb{Z}_n \mid \mathrm{ggT}(n, k) = 1\}.$$

Beweis: Dies folgt unmittelbar aus der Definition von $\mathbb{Z}_n^*$ und Satz 6.8. □

Lemma 7.21
$\mathbb{Z}_2$ ist der kleinste Körper, d. h. jeder Körper mit zwei Elementen ist isomorph zu $\mathbb{Z}_2$.

Beweis: Nach Lemma 7.14 h. besitzt jeder Körper mindestens zwei Elemente, 0 und 1, also reicht es in der Tat zu zeigen, daß jeder Körper mit zwei Elementen isomorph zu $\mathbb{Z}_2$ ist.

Sei nun $K = \{0, 1\}$ ein Körper mit zwei Elementen. Wir behaupten, die Abbildung

$$f : K \to \mathbb{Z}_2 : 0 \mapsto [0],\ 1 \mapsto [1]$$

ist ein Isomorphismus. Da sie offenbar bijektiv ist, reicht es zu zeigen, daß f ein Ringhomomorphismus ist. D. h. man muß für $a, b \in \{0, 1\}$ zeigen:

$$f(a + b) = f(a) + f(b) \quad \text{und} \quad f(a \cdot b) = f(a) \cdot f(b),$$

was durch Ausprobieren der jeweils vier Möglichkeiten leicht zu sehen ist. □

Lemma 7.22
Es sei $f : F \to K$ *ein Körperhomomorphismus.*

a. *Ist* $f \neq 0$*, so ist* f *injektiv. Man nennt* f *dann auch eine* Körpererweiterung.
b. *Ist* f *bijektiv, so ist* $f^{-1} : K \to F$ *ebenfalls ein Körperhomomorphismus.*

Beweis: a. Wegen $1 \cdot f(1) = f(1) = f(1 \cdot 1) = f(1) \cdot f(1)$, gilt mit Lemma 7.14 k.

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

Angenommen, $f(1) = 0$, so ist für $x \in F$ $f(x) = f(x \cdot 1) = f(x) \cdot f(1) = f(x) \cdot 0 = 0$. Mithin ist in diesem Fall $f = 0$ die Nullabbildung, im Widerspruch zur Voraussetzung $f \neq 0$.
Also gilt $f(1) = 1$. Da f ein Gruppenhomomorphismus ist, gilt f ist genau dann injektiv, wenn $\{0\} = \mathrm{Ker}(f) = \{x \in F \mid f(x) = 0\}$. Es reicht also $\mathrm{Ker}(f) = \{0\}$ zu zeigen. Für $x \in F \setminus \{0\}$ gilt aber:

$$f(x) \cdot f\left(x^{-1}\right) = f\left(x \cdot x^{-1}\right) = f(1) = 1 \neq 0.$$

Also gilt mit Lemma 7.14 j., daß $f(x) \neq 0$, also $x \notin \mathrm{Ker}(f)$.

b. Dies folgt aus Proposition 3.17 c., der entsprechenden Aussage für Gruppenhomomorphismen.

□

Die Körper $\mathbb{Q}$ und $\mathbb{R}$ auf der einen Seite und $\mathbb{Z}_p$, p Primzahl, auf der anderen Seite besitzen einen wichtigen Unterschied. In $\mathbb{Q}$ ist $1 + 1 + 1 + \ldots$ niemals 0, in $\mathbb{Z}_p$ gilt hingegen $p \cdot 1 = 1 + \overset{p}{\ldots} + 1 = 0$. Das führt uns auf die folgende Definition.

Definition 7.23
Es sei R ein Ring mit Eins. Die kleinste natürliche Zahl $n > 0$ mit $n \cdot 1 := \sum_{i=1}^{n} 1 = 1 + \overset{n}{\ldots} + 1 = 0$ heißt, so sie existiert, die *Charakteristik* des Ringes R. Existiert keine solche Zahl n, so sagen wir, der Ring hat die Charakteristik null.

Wir schreiben $\mathrm{char}(R)$ für die Charakteristik von R.

Bemerkung 7.24
Wir haben für jeden Ring R eine kanonische Abbildung

$$\varphi : \mathbb{Z} \to R : k \mapsto k \cdot 1.$$

Man sieht sofort, daß φ ein Ringhomomorphismus ist. Zudem gilt für $n \geq 0$

$$\text{char}(R) = n \;\Leftrightarrow\; \text{Ker}(\varphi) = n\mathbb{Z}.$$

Dies ist nur eine andere Formulierung der Definition.

Beispiel 7.25
Es gilt: $\text{char}(\mathbb{Q}) = \text{char}(\mathbb{R}) = 0$, $\text{char}(\mathbb{Z}_n) = n$ für $n > 1$.

Lemma 7.26
Ist K *ein Körper mit* $\text{char}(K) \neq 0$*, so ist* $\text{char}(K)$ *eine Primzahl.*

Beweis: Wir nehmen an, $n := \text{char}(K)$ sei keine Primzahl. Wegen $1 \neq 0$ gilt $n > 1$. Dann gilt aber $n = m \cdot k$ für geeignete $m, k \in \mathbb{N}$ mit $1 < m, k < n$.

Da n die kleinste positive Zahl mit $n \cdot 1 = 0$ ist, so sind $k' := k \cdot 1 \in K$ und $m' := m \cdot 1 \in K$ beide nicht null. Andererseits gilt

$$k' \cdot m' = (k \cdot m) \cdot 1 = n \cdot 1 = 0,$$

im Widerspruch zu Teil j. von Lemma 7.14. □

Wir kommen jetzt zum Körper $\mathbb{C}$ der komplexen Zahlen, dem neben $\mathbb{R}$ wichtigsten Körper. Warum reichen eigentlich die reellen Zahlen nicht aus, wozu braucht man die komplexen Zahlen? Ja, man kann sogar fragen, warum wir überhaupt die reellen Zahlen benötigen, wenn wir doch ohnehin nur mit endlichen Dezimalbrüchen, also rationalen Zahlen, rechnen können? Die Antwort auf die zweite Frage ist schnell gegeben. Wir wissen alle, daß etwa ganz natürlich auftretende Größen wie die Länge der Diagonalen eines Quadrates mit Seitenlänge eins, sprich die Zahl $\sqrt{2}$, oder das Verhältnis von Umfang zum Durchmesser eines Kreises, sprich die Kreiszahl π, keine rationalen Zahlen sind. Sie sind aber reelle Zahlen und die reellen Zahlen sind in gewissen Sinne (wie der Leser in der Analysis lernen kann), eine Vervollständigung der rationalen Zahlen. Wir brauchen also die reellen Zahlen, da die rationalen Zahlen Lücken aufweisen. Die komplexen Zahlen werden nun deshalb eingeführt, um einen Mangel, den die reellen Zahlen immer noch haben, zu beheben. Hierbei geht es um das Lösen von Gleichungen, aber nicht mehr linearen, sondern quadratischen. Es ist bekannt, daß das Quadrat einer reellen Zahl stets nicht-negativ ist. Also kann es keine reelle Zahl x geben, die die Gleichung $x^2 = -1$ löst.

Als Lösung genau dieser Gleichung wird nun eine neue Größe eingeführt, die *imaginäre Einheit* i. Definitionsgemäß ist sie diejenige Zahl, für die $i^2 = -1$ gilt. Wenn man nun eine solche Größe i einführt, dann ist damit alleine gar nichts gewonnen. Man will ja mit i auch rechnen können, und zwar will man möglichst alle Rechenregeln von $\mathbb{R}$ übertragen. Man will nicht nur $i^2 = i \cdot i$,

sondern auch $i + i$ oder Ausdrücke wie $37 + 42i$ bilden können. Dabei sollen die so zu konstruierenden *komplexen Zahlen* die reellen Zahlen als Teilmenge enthalten.

Daß es wirklich ein solches Zahlsystem komplexer Zahlen, in unserer Sprache den Körper der komplexen Zahlen, gibt, ist überhaupt nicht klar und wurde historisch erst spät realisiert und auch akzeptiert.[8] Gauß hat die Zahlen geometrisch, als Punkte in der Ebene, eingeführt, weshalb die komplexen Zahlen heute noch *gaußsche Zahlenebene* heißen. Wir führen die komplexen Zahlen ebenfalls als reelle Zahlenpaare ein, definieren die Addition und die Multiplikation aber algebraisch und werden die Definitionen erst im Anschluß daran geometrisch interpretieren.

Definition 7.27
Die Menge $\mathbb{C} := \{(x, y) \mid x, y \in \mathbb{R}\}$ zusammen mit der durch

$$(x, y) + (u, v) := (x + u, y + v), \quad \text{für } (x, y), (u, v) \in \mathbb{C},$$

und

$$(x, y) \cdot (u, v) := (xu - yv, xv + yu), \quad \text{für } (x, y), (u, v) \in \mathbb{C},$$

definierten Addition und Multiplikation heißt der *Körper der komplexen Zahlen*. Für $z = (x, y) \in \mathbb{C}$ heißt $\Re(z) := x$ der Realteil von z und $\Im(z) := y$ der Imaginärteil.

Lemma 7.28
$(\mathbb{C}, +, \cdot)$ *ist ein Körper.*

Beweis: Man sieht sofort, daß $(\mathbb{C}, +)$ eine abelsche Gruppe ist mit $(0, 0)$ als neutralem Element und $(-x, -y)$ als Inversem zu $(x, y) \in \mathbb{C}$.

Etwas mehr ist zu zeigen, um zu sehen, daß $(\mathbb{C} \setminus \{(0, 0)\}, \cdot)$ eine abelsche Gruppe ist mit $(1, 0)$ als neutralem Element und $\left(\frac{x}{x^2+y^2}, -\frac{y}{x^2+y^2}\right)$ als Inversem zu $(x, y) \in \mathbb{C} \setminus \{(0, 0)\}$. Wir überlassen den Nachweis dem Leser als Übungsaufgabe. □

Bemerkung 7.29
Wir wollen nun sehen, daß $\mathbb{C}$ ein Erweiterungskörper von $\mathbb{R}$ ist. Dazu betrachten wir die Abbildung

$$\varphi : \mathbb{R} \to \mathbb{C} : x \mapsto (x, 0).$$

Man prüft leicht nach, daß φ ein injektiver Körperhomomorphismus ist.

Wir identifizieren $\mathbb{R}$ mit dem Bild $\varphi(\mathbb{R}) = \mathbb{R} \times \{0\} \subset \mathbb{R} \times \mathbb{R} = \mathbb{C}$. Damit ist $\mathbb{R}$ ein *Unterkörper* von $\mathbb{C}$.

Notation 7.30
Praktischer als das Rechnen mit Paaren von Zahlen ist die folgende Notation

[8] Erstmals taucht $\sqrt{-1}$ wohl um 1540 bei Cardano auf. Wirklich als Zahlsystem wurden die komplexen Zahlen aber erst durch Gauß, 1777-1855, etabliert. Hierzu und zu vielen weiteren interessanten Tatsachen um die komplexen Zahlen vgl. [**Ebb92**] § 3.

für komplexe Zahlen. Wir setzen $x := (x, 0)$ für $x \in \mathbb{R}$ und $i := (0, 1)$. Dann gilt für $z = (x, y) \in \mathbb{C}$

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

Bemerkung 7.31
Mit dieser Schreibweise gilt zunächst:

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

Ferner ergibt sich die etwas willkürlich anmutende Definition der Multiplikation ganz "natürlich" aus

$$(x + iy)(u + iv) = \left(xu + i^2yv\right) + i(xv + yu) = (xu - yv) + i(xv + yu).$$

Bemerkung 7.32
Auf $\mathbb{R}$ und $\mathbb{C}$ hat man noch andere wichtige Strukturen, die man auf beliebigen Körpern nicht hat.

Auf $\mathbb{R}$ hat man die *Ordnungsrelation* $\leq$, die eine totale Ordnung auf $\mathbb{R}$ ist, und die mit den Operationen auf $\mathbb{R}$ verträglich ist, d. h. für $x, y, z \in \mathbb{R}$ gilt:

a. $x \leq y \;\Rightarrow\; x + z \leq y + z$, und
b. $0 < x, 0 < y \;\Rightarrow\; 0 < xy$.

Außerdem hat man auf $\mathbb{R}$ die *Betragsfunktion*

$$|\cdot| : \mathbb{R} \to \mathbb{R}_{\geq 0} : x \mapsto \begin{cases} x, & x \geq 0 \\ -x, & x < 0. \end{cases}$$

Die Ordnungsrelation $\leq$ auf $\mathbb{R}$ läßt sich nicht so auf $\mathbb{C}$ fortsetzen, daß die obigen Gesetze a. und b. erhalten bleiben.[9] Im Gegensatz dazu besitzt $\mathbb{C}$ aber eine Betragsfunktion,

$$|\cdot| : \mathbb{C} \to \mathbb{R}_{\geq 0} : x + iy \mapsto \sqrt{x^2 + y^2},$$

die die Betragsfunktion auf $\mathbb{R}$ fortsetzt.

Außerdem gibt es auf $\mathbb{C}$ eine weitere wichtige Abbildung, die *komplexe Konjugation*

$$\bar{\cdot} : \mathbb{C} \to \mathbb{C} : z = x + iy \mapsto \overline{z} := x - iy.$$

Für $z \in \mathbb{C}$ heißt $\overline{z}$ die zu z *konjugiert komplexe Zahl*.

Die folgenden Eigenschaften der komplexen Zahlen sind einfach nachzuweisen, und ihr Nachweis sei dem Leser überlassen.

Lemma 7.33
Für $z, w \in \mathbb{C}$ gelten:

a. $\overline{z} + \overline{w} = \overline{z + w}$,
b. $\overline{z} \cdot \overline{w} = \overline{z \cdot w}$,
c. $\overline{\overline{z}} = z$,

[9] Sonst würde entweder $0 < i$ oder $0 < -i$ gelten, und somit $0 < i^2 = -1$ oder $0 < (-i)^2 = -1$, was im Widerspruch zur Definition von $\leq$ auf $\mathbb{R}$ steht.

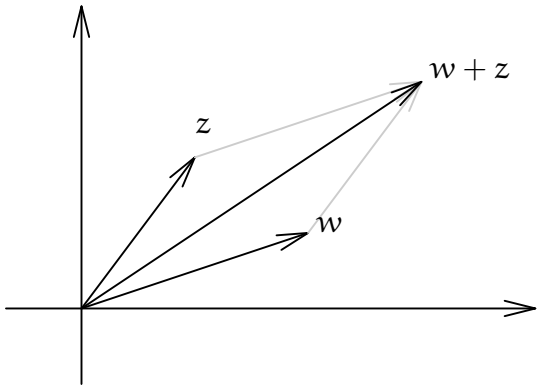
d. $\overline{z} = z \Leftrightarrow z \in \mathbb{R}$,
e. $z \cdot \overline{z} = |z|^2$,
f. $|z| \cdot |w| = |zw|$,
g. $|z + w| \leq |z| + |w|$, *und*
h. $z = 0 \Leftrightarrow |z| = 0$.

Teil g. nennt man die *Dreiecksungleichung*. Sie wird vor allem in der Analysis von großer Bedeutung sein. Elementargeometrisch wird ihre Bedeutung im Folgenden augenscheinlich.

Geometrische Deutung der komplexen Zahlen

Wir betrachten $z = (x, y)$ als Richtungsvektor in der Zahlenebene $\mathbb{R}^2$.

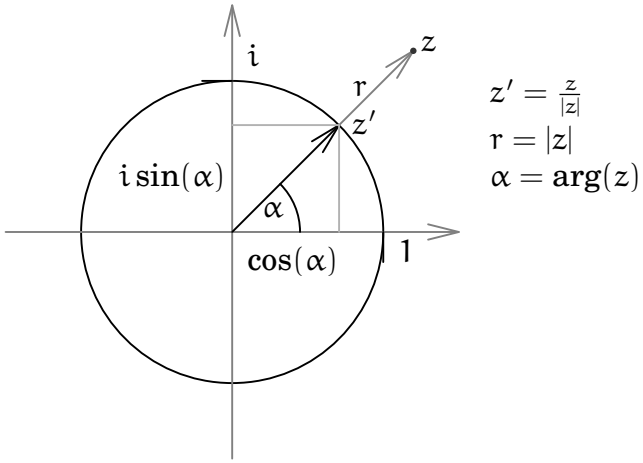
Die Addition ist einfach die komponentenweise Addition, also die Addition der Vektoren.



Zur geometrischen Interpretation der Multiplikation brauchen wir die Betragsfunktion. Der Betrag $r := |z|$ einer komplexen Zahl z ist die Länge des Vektors z (Pythagoras). Für $z \neq 0$ hat $z' := \frac{z}{|z|}$ die Länge eins, und es gilt

$$z = |z| \cdot z' = r \cdot z'.$$

D. h. z ist das Produkt eines Vektors von Länge eins mit einer nicht-negativen reellen Zahl. Dabei ist z' vollständig durch den Winkel α bestimmt, den z' mit der x-Achse einschließt, nämlich $z' = \big(\cos(\alpha), \sin(\alpha)\big)$. Also ist jede komplexe Zahl $z \neq 0$ eindeutig durch ihren Betrag und den Winkel $\alpha =: \arg(z)$, das *Argument* von z, bestimmt. Das Paar $(r, \alpha) = \big(|z|, \arg(z)\big)$ nennt man die *Polarkoordinaten* von z.



Die komplexen Zahlen vom Betrag eins sind genau die Punkte auf dem Einheitskreis. Für $z' \in \mathbb{C}$ mit $|z'| = 1$ gibt es also genau ein $0 \leq \alpha < 2\pi$ mit

$$z' = \cos(\alpha) + i\sin(\alpha) = e^{i\alpha},$$

wie man in der Analysis lernt. Damit gilt für ein beliebiges $c \in \mathbb{C}$

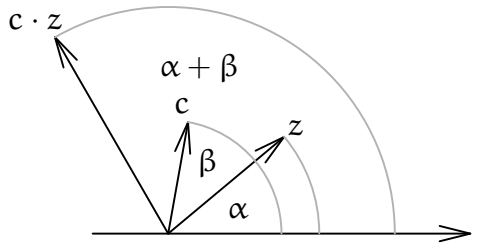
$$c = |c| \cdot \big(\cos(\beta) + i\sin(\beta)\big) = |c| \cdot e^{i\beta},$$

für $\beta = \arg(c)$.

Daraus ergibt sich für die Multiplikationsabbildung mit einer festen komplexen Zahl $z = |z| \cdot e^{i\alpha}$,

$$m_z : \mathbb{C} \to \mathbb{C} : c \mapsto |z| \cdot |c| \cdot e^{i(\alpha+\beta)}.$$

Die Multiplikation mit z ist also eine *Drehstreckung*, daß heißt der Vektor c wird um den Winkel $\alpha = \arg(z)$ gedreht und um den Faktor $|z|$ gestreckt.



AUFGABEN

Aufgabe 7.34 (Direkte Summe)
Es seien $(R, +, \cdot)$ und $(S, +, \cdot)$ zwei Ringe. Wir definieren auf der Menge $R \times S$ zwei Operationen $+$ und $\cdot$ durch:

$$(r, s) + (r', s') := (r + r', s + s')$$

und

$$(r, s) \cdot (r', s') := (r \cdot r', s \cdot s')$$

für $(r, s), (r', s') \in R \times S$. Zeige, $(R \times S, +, \cdot)$ wird dadurch zu einem Ring.

Anmerkung: Wir nennen $R \times S$ die *direkte Summe* von R und S. Als Schreibweise ist $R \oplus S$ gebräuchlich.

Aufgabe 7.35 (Gaußsche Zahlen)
a. Es sei $\mathbb{Q}[i] := \{a + ib \mid a, b \in \mathbb{Q}\}$, wobei $i \in \mathbb{C}$ die imaginäre Einheit ist. Zeige, $\mathbb{Q}[i]$ ist ein Teilkörper von $(\mathbb{C}, +, \cdot)$.
b. Es sei $\mathbb{Z}[i] := \{a + ib \mid a, b \in \mathbb{Z}\}$. Zeige, $\mathbb{Z}[i]$ ist ein Teilring von $(\mathbb{Q}[i], +, \cdot)$.
c. Definiere $\varphi : \mathbb{Z}[i] \to \mathbb{N} \;:\; a + ib \mapsto |a + ib|^2 = a^2 + b^2$. Zeige, für $A, B \in \mathbb{Z}[i] \setminus \{0\}$ gibt es $Q, R \in \mathbb{Z}[i]$ mit $A = Q \cdot B + R$ und $0 \leq \varphi(R) < \varphi(B)$.

Anmerkung: Ringe zusammen mit einer Funktion wie in c. nennt man *euklidische Ringe*!

Aufgabe 7.36
Es seien M eine Menge und K ein Körper. Ist dann K^M ein Körper?

Aufgabe 7.37
Es sei p eine Primzahl und $K = \mathbb{Z}_p$.

a. Zeige, $a^p - a = 0$ für alle $a \in K$.

b. Zeige, $(a+b)^p = a^p + b^p$ für alle $a, b \in K$.

c. Zeige, ist $n \in \mathbb{N}$, so gilt $p \mid n^p - n$.

Aufgabe 7.38

Ein Ring $(R, +, \cdot)$ heißt *nullteilerfrei*, falls gilt: $\forall\, x, y \in R$ mit $x \cdot y = 0$ gilt $x = 0$ oder $y = 0$.

Zeige, ein endlicher, nullteilerfreier Ring mit 1 ist ein Schiefkörper.

Aufgabe 7.39

Ist K ein endlicher Körper ungerader Charakteristik. Zeige, in K gibt es genau $\frac{|K|+1}{2}$ Quadrate und $\frac{|K|-1}{2}$ Nicht-Quadrate.

Hinweis: Betrachte den Gruppenhomomorphismus $K^* \to K^* : a \mapsto a^2$.

Aufgabe 7.40

Es seien $z = 1 - i, z' = 1 + 2 \cdot i \in \mathbb{C}$. Bestimme $z + z'$, $z \cdot z'$, $|z|$, $\overline{z} \cdot z$, z^{-1}.

Aufgabe 7.41

Bestimme alle Körperautomorphismen $\alpha : \mathbb{C} \to \mathbb{C}$ für die gilt $\alpha(r) = r \;\; \forall\, r \in \mathbb{R}$.

8. Matrixgruppen

Die für die lineare Algebra wichtigsten Gruppen sind die Matrixgruppen. Wir werden hier nur die elementarsten Grundtatsachen bereitstellen. Die Theorie der Matrixgruppen geht weit darüber hinaus und ist heute noch ein hochaktueller Gegenstand der mathematischen Forschung.

Generalvoraussetzung: Im Verlauf dieses Abschnitts bezeichne R stets einen kommutativen Ring mit Eins.

Definition 8.1 a. Eine $m \times n$-*Matrix* über R ist ein rechteckiges Schema A mit Einträgen aus R der Form

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

Wir schreiben verkürzt auch

$$A = (a_{ij})_{i=1,\dots,m;j=1,\dots,n} = (a_{ij}),$$

wenn keine Unklarheiten zu befürchten sind.

b. Ist $A = (a_{ij})$ eine $m \times n$-Matrix, dann bezeichnen wir

$$a_i := (a_{i1}, \dots, a_{in})$$

als den i-ten *Zeilenvektor* von A und

$$a^j := \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

als den j-ten *Spaltenvektor* von A.

c. Die Menge aller $m \times n$-Matrizen über R wird mit

$$\mathrm{Mat}(m \times n, R)$$

bezeichnet, und falls $m = n$, dann auch kurz mit $\mathrm{Mat}_n(R) = \mathrm{Mat}(n, R)$ und man spricht von *quadratischen Matrizen*.

d. Ist $A = (a_{ij}) \in \mathrm{Mat}(m \times n, R)$, so heißt die $n \times m$-Matrix

$$A^t := \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix},$$

d. h. für $A^t = (a'_{ij})$ gilt $a'_{ij} = a_{ji}$, die *Transponierte* oder *transponierte Matrix* von A.

e. Schließlich definieren wir

$$R^n := \mathrm{Mat}(n \times 1, R) = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \;\middle|\; x_1, \dots, x_n \in R \right\}.$$

Die Elemente von R^n heißen *Vektoren* oder *Punkte* im R^n. x_i heißt die i-te *Komponente* des Vektors x.

Wir definieren und untersuchen zunächst allgemeiner, als wir es für die Matrixgruppen selbst brauchen, die Matrixaddition, die Skalarmultiplikation und die Matrixmultiplikation.

Definition 8.2 a. Es seien $A = (a_{ij}), B = (b_{ij}) \in \mathrm{Mat}(m \times n, R)$ und $\lambda \in R$. Dann definiert man

$$A + B := (a_{ij} + b_{ij}) = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2n} + b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \dots & a_{mn} + b_{mn} \end{pmatrix},$$

sowie

$$\lambda \cdot A := (\lambda a_{ij}) = \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \dots & \lambda a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \dots & \lambda a_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda a_{m1} & \lambda a_{m2} & \dots & \lambda a_{mn} \end{pmatrix}.$$

b. Sind $A = (a_{ij}) \in \mathrm{Mat}(m \times n, R)$ und $B = (b_{jk}) \in \mathrm{Mat}(n \times p, R)$ zwei Matrizen, wobei A genauso viele Spalten wie B Zeilen hat. Dann definieren wir das *Matrixprodukt* durch

$$A \circ B := C, \quad \text{mit } C = (c_{ik}) \in \mathrm{Mat}(m \times p, R) \text{ und } c_{ik} := \sum_{j=1}^{n} a_{ij} b_{jk}.$$

Bemerkung 8.3 a. Die in Definition 8.2 a. definierte Addition zweier Matrizen definiert auf $\mathrm{Mat}(m \times n, R)$ offensichtlich eine zweistellige Operation,

bezüglich derer $\mathrm{Mat}(m \times n, R)$ eine abelsche Gruppe $(\mathrm{Mat}(m \times n, R), +)$ wird, wie man leicht nachprüft.

b. Die in Definition 8.1 a. definierte *Skalarmultiplikation* kann als zweistellige Operation

$$\cdot : R \times \mathrm{Mat}(m \times n, R) \to \mathrm{Mat}(m \times n, R) : (\lambda, A) \mapsto \lambda \cdot A$$

aufgefaßt werden. Diese wird später näher betrachtet werden.
Wir werden meist kurz λA bzw. λx schreiben, statt $\lambda \cdot A$ bzw. $\lambda \cdot x$, wenn $\lambda \in R$, $A \in \mathrm{Mat}(m \times n, R)$ und $x \in R^n$.

c. Die in Definition 8.1 b. eingeführte Matrixmultiplikation kann ebenfalls als zweistellige Operation

$$\circ : \mathrm{Mat}(m \times n, R) \times \mathrm{Mat}(n \times p, R) \to \mathrm{Mat}(m \times p, R) : (A, B) \mapsto A \circ B$$

aufgefaßt werden. Wir schreiben statt $A \circ B$ häufig kurz AB, insbesondere auch Ax statt $A \circ x$ für $x \in R^n$.

d. Man beachte, daß die Vektoren im R^n Spaltenvektoren sind, jedoch schreiben wir sie aus Platzgründen in aller Regel als "transponierte Zeilenvektoren". Ferner schreiben wir statt $(x_1 \ \dots \ x_n)^t$ meist $(x_1, \dots, x_n)^t$, um die Trennung zwischen den einzelnen Vektoreinträgen deutlicher zu machen.

e. Sind $x = (x_1, \dots, x_n)^t, y = (y_1, \dots, y_n)^t \in R^n$, dann definieren wir das *Skalarprodukt*

$$\langle x, y \rangle := \sum_{i=1}^{n} x_i y_i$$

von x und y. Dieses definiert eine zweistellige Operation auf R^n

$$\langle \cdot, \cdot \rangle : R^n \times R^n \to R : (x, y) \mapsto \langle x, y \rangle.$$

f. Sind $A = (a_{ij}) \in \mathrm{Mat}(m \times n, R)$ und $B = (b_{jk}) \in \mathrm{Mat}(n \times p, R)$, und gilt $A \circ B = (c_{ik})$, dann ist das Element $c_{ik} = \langle a_i^t, b^j \rangle$ das Skalarprodukt aus dem i-ten Zeilenvektor von A und dem k-ten Spaltenvektor von B.

g. Man beachte, daß das Produkt nur dann definiert ist, wenn A soviele Spalten wie B Zeilen hat. Das Produkt $A \circ B$ hat dann soviele Zeilen wie A und soviele Spalten wie B.

Beispiel 8.4

Folgende Matrizen $A, B \in \mathrm{Mat}(2 \times 3, \mathbb{R})$ und $C \in \mathrm{Mat}(3 \times 2, \mathbb{R})$ seien gegeben:

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 3 & 1 & 2 \end{pmatrix}, B = \begin{pmatrix} 3 & 2 & 1 \\ 0 & 4 & 5 \end{pmatrix}, C = \begin{pmatrix} 1 & 0 \\ 2 & 1 \\ 3 & 1 \end{pmatrix}.$$

Dann gilt:

$$A + B = \begin{pmatrix} 4 & 2 & 3 \\ 3 & 5 & 7 \end{pmatrix}, \ 3 \cdot A = \begin{pmatrix} 3 & 0 & 6 \\ 9 & 3 & 6 \end{pmatrix} \ \text{ und } \ A \circ C = \begin{pmatrix} 7 & 2 \\ 11 & 3 \end{pmatrix}.$$

Jede Matrix definiert wie folgt eine Abbildung.

Definition 8.5
Ist $A \in \mathrm{Mat}(m \times n, R)$, so definieren wir

$$f_A : R^n \to R^m : x \mapsto Ax = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n a_{1j}x_j \\ \vdots \\ \sum_{j=1}^n a_{mj}x_j \end{pmatrix}.$$

f_A heißt die *zu* A *assoziierte* oder *zu* A *gehörige Abbildung*.

Bemerkung 8.6
Um den Zusammenhang zwischen A und f_A besser zu verstehen, betrachten wir für $i = 1, \dots, n$ den i-ten *Einheitsvektoren* $e^i = (\delta_{1i}, \cdots, \delta_{ni})^t \in R^n$, wo δ_{ij} das *Kronecker Symbol*[10] ist, d. h.

$$e^i := \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix},$$

wo die Eins in der i-ten Komponente steht.

Es ist dann

$$f_A(e^i) = \begin{pmatrix} \sum_{i=1}^n a_{1j}\delta_{ji} \\ \vdots \\ \sum_{i=1}^n a_{mj}\delta_{ji} \end{pmatrix} = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix} = a^i,$$

d. h. die i-te Spalte von A ist das Bild des i-ten Einheitsvektors unter f_A.

Hieraus folgt insbesondere, daß A durch f_A eindeutig bestimmt ist.

Lemma 8.7
Für $x, y \in R^n$, $A, B \in \mathrm{Mat}(m \times n, R)$, $C \in \mathrm{Mat}(n \times p, R)$ *und* $\lambda \in R$ *gelten:*

a. $A(x + y) = Ax + Ay$ *und* $A(\lambda x) = \lambda Ax$,
b. $\lambda \cdot (A \circ C) = (\lambda \cdot A) \circ C = A \circ (\lambda \cdot C)$,
c. $f_{A+B} = f_A + f_B$, *und*
d. $f_{\lambda A} = \lambda f_A$.

Beweis: Es seien $A = (a_{ij})$, $B = (b_{ij})$, $x = (x_1, \dots, x_n)^t$ und $y = (y_1, \dots, y_n)^t$.

a./b. Dies sei dem Leser als Übungsaufgabe überlassen.

c. Es gilt:

$$f_{A+B}(x) = (A + B)x = \begin{pmatrix} a_{11} + b_{11} & \dots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \dots & a_{mn} + b_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

[10] Das Kronecker Symbol δ_{ij} für $i, j \in \mathbb{N}$ ist definiert durch

$$\delta_{ij} := \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

$$= \begin{pmatrix} \sum_{j=1}^{n}(a_{1j}+b_{1j})x_j \\ \vdots \\ \sum_{j=1}^{n}(a_{mj}+b_{mj})x_j \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{n} a_{1j}x_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj}x_j \end{pmatrix} + \begin{pmatrix} \sum_{j=1}^{n} b_{1j}x_j \\ \vdots \\ \sum_{j=1}^{n} b_{mj}x_j \end{pmatrix} = Ax + Bx.$$

d. Es gilt:

$$f_{\lambda A}(x) = (\lambda A)x = \begin{pmatrix} \lambda a_{11} & \dots & \lambda a_{1n} \\ \vdots & & \vdots \\ \lambda a_{m1} & \dots & \lambda a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{j=1}^{n}(\lambda a_{1j})x_j \\ \vdots \\ \sum_{j=1}^{n}(\lambda a_{mj})x_j \end{pmatrix} = \lambda \begin{pmatrix} \sum_{j=1}^{n} a_{1j}x_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj}x_j \end{pmatrix} = \lambda(Ax).$$

□

Wir wollen jetzt sehen, wie sich die Multiplikation von Matrizen mit den zugehörigen Abbildungen verträgt.

Satz 8.8

Für $A \in \mathrm{Mat}(m \times n, R)$ *und* $B \in \mathrm{Mat}(n \times p, R)$ *gilt:*

$$f_{A\circ B} = f_A \circ f_B.$$

Beweis: Da Definitionsbereich und Wertebereich von beiden Abbildungen übereinstimmen, reicht es zu zeigen:

$$(f_{A\circ B})(x) = (f_A \circ f_B)(x), \quad \text{für alle } x \in R^p.$$

Seien $A = (a_{ij})$ und $B = (b_{jk})$, und sei $x = (x_1, \dots, x_p)^t \in R^p$ gegeben.

$$(f_{A\circ B})(x) = (A \circ B)x = \begin{pmatrix} \sum_{j=1}^{n} a_{1j}b_{j1} & \dots & \sum_{j=1}^{n} a_{1j}b_{jp} \\ \vdots & & \vdots \\ \sum_{j=1}^{n} a_{mj}b_{j1} & \dots & \sum_{j=1}^{n} a_{mj}b_{jp} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{k=1}^{p}\sum_{j=1}^{n} a_{1j}b_{jk}x_k \\ \vdots \\ \sum_{k=1}^{p}\sum_{j=1}^{n} a_{mj}b_{jk}x_k \end{pmatrix}.$$

Ferner gilt:

$$(f_A \circ f_B)(x) = f_A(Bx) = A(Bx) = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} \sum_{k=1}^{p} b_{1k}x_k \\ \vdots \\ \sum_{k=1}^{p} b_{nk}x_k \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{j=1}^{n} a_{1j}\sum_{k=1}^{p} b_{jk}x_k \\ \vdots \\ \sum_{j=1}^{n} a_{mj}\sum_{k=1}^{p} b_{jk}x_k \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{n}\sum_{k=1}^{p} a_{1j}b_{jk}x_k \\ \vdots \\ \sum_{j=1}^{n}\sum_{k=1}^{p} a_{mj}b_{jk}x_k \end{pmatrix}.$$

Beide Ausdrücke stimmen (bis auf die Reihenfolge der Summation) überein, was zu zeigen war. □

Korollar 8.9
Die Matrixmultiplikation ist assoziativ, d. h. für $A \in \mathrm{Mat}(m \times n, R)$, $B \in \mathrm{Mat}(n \times p, R)$ *und* $C \in \mathrm{Mat}(p \times q, R)$ *gilt*

$$A \circ (B \circ C) = (A \circ B) \circ C.$$

Beweis: Dies folgt aus Satz 8.8, da die Komposition von Abbildungen assoziativ ist und da eine Matrix A durch die Abbildung f_A eindeutig bestimmt ist. □

Man kann die Aussage des Korollars natürlich auch direkt nachweisen, was auf die gleiche Rechnung wie in 8.8 führt - statt des einen Vektors x hat man die q Spaltenvektoren von C zu multiplizieren, was etwas mehr Schreibarbeit bedeutet.

Lemma 8.10
Sind $A, B \in \mathrm{Mat}(m \times n, R)$ *und* $C, D \in \mathrm{Mat}(n \times p, R)$*, so gelten die Distributivgesetze:*

$$A \circ (C + D) = A \circ C + A \circ D,$$

sowie

$$(A + B) \circ C = A \circ C + B \circ C.$$

Beweis: Die Aussage kann wie Korollar 8.9 aus Lemma 8.7 und Satz 8.8 abgeleitet werden und sei dem Leser als Übung anempfohlen. □

Von jetzt an wollen wir uns auf quadratische Matrizen beschränken. Sind alle Matrizen aus $\mathrm{Mat}(n, R)$, d. h. haben alle die gleiche Größe, dann können wir diese beliebig untereinander multiplizieren.

Lemma 8.11
$\big(\mathrm{Mat}(n, R), \circ\big)$ *ist eine Halbgruppe mit neutralem Element*

$$\mathbb{1}_n := E_n := \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix},$$

d. h. $\mathbb{1}_n = (\delta_{ij})$*. Wir nennen* $\mathbb{1}_n$ *die* Einheitsmatrix.

Beweis: Wir haben bereits in Bemerkung 8.3 gesehen, daß $\circ$ eine zweistellige Operation auf $\mathrm{Mat}(n, R)$ definiert, die zudem nach 8.9 assoziativ ist. Es bleibt also zu zeigen, daß $\mathbb{1}_n$ ein neutrales Element bezüglich der Multiplikation ist.

Sei dazu $A = (a_{jk}) \in \mathrm{Mat}(n, R)$, dann gilt $\mathbb{1}_n \circ A = (c_{ik})$ mit

$$c_{ik} = \sum_{j=1}^{n} \delta_{ij} a_{jk} = a_{ik}.$$

Also ist $\mathbb{1}_n \circ A = (c_{ik}) = A$. □

Korollar 8.12
$\big(\mathrm{Mat}(n,R),+,\circ\big)$ *ist ein (i. a. nicht-kommutativer) Ring mit Eins* $\mathbb{1}_n$.

Beweis: Aus Bemerkung 8.3 folgt, daß $\big(\mathrm{Mat}(n,R),+\big)$ eine abelsche Gruppe ist. Da zudem $\big(\mathrm{Mat}(n,R),\circ\big)$ nach Lemma 8.11 eine Halbgruppe ist und nach Lemma 8.10 die Distributivgesetze gelten, folgt die Behauptung - vgl. Bemerkung 7.12. □

Definition 8.13
Eine Matrix $A \in \mathrm{Mat}(n,R)$ heißt *invertierbare Matrix* oder *reguläre Matrix*, falls es eine Matrix $B \in \mathrm{Mat}(n,R)$ gibt mit $B \circ A = \mathbb{1}_n$ und $A \circ B = \mathbb{1}_n$.

Die Menge

$$\mathrm{Gl}_n(R) := \mathrm{Gl}(n,R) := \{A \in \mathrm{Mat}(n,R) \mid A \text{ ist invertierbar}\}$$

der invertierbaren Matrizen heißt *allgemeine lineare Gruppe*.

Wir werden später sehen (vgl. Korollar II.10.7), daß im Falle $R = K$, ein Körper, A genau dann invertierbar ist, wenn es ein $B \in \mathrm{Mat}(n,K)$ gibt mit $B \circ A = \mathbb{1}_n$.

Der Name allgemeine lineare *Gruppe* ist durch den folgenden Satz begründet. Im Besonderen folgt damit, daß die Matrix B aus der obigen Definition eindeutig bestimmt ist und gerade das Inverse A^{-1} von A bezüglich der Gruppenoperation $\circ$ ist.

Satz 8.14
$\big(\mathrm{Gl}_n(R),\circ\big)$ *ist eine Gruppe mit neutralem Element* $\mathbb{1}_n$.

Insbesondere gelten für $A, B \in \mathrm{Gl}_n(R)$

$$(A \circ B)^{-1} = B^{-1} \circ A^{-1}, \quad \textit{und} \quad \big(A^{-1}\big)^{-1} = A.$$

Beweis: Wegen $\mathbb{1}_n \circ \mathbb{1}_n = \mathbb{1}_n$ gilt $\mathbb{1}_n \in \mathrm{Gl}_n(R)$ und somit ist $\mathrm{Gl}_n(R) \neq \emptyset$.

Sind $A, A' \in \mathrm{Gl}_n(R)$, so gibt es Matrizen $B, B' \in \mathrm{Mat}(n,R)$ mit $A \circ B = B \circ A = A' \circ B' = B' \circ A' = \mathbb{1}_n$. Damit gilt dann für $B' \circ B \in \mathrm{Mat}(n,R)$

$$\big(A \circ A'\big) \circ \big(B' \circ B\big) = A \circ \big(A' \circ B'\big) \circ B = A \circ \mathbb{1}_n \circ B = A \circ B = \mathbb{1}_n,$$

und analog sieht man $\big(B' \circ B\big) \circ \big(A \circ A'\big) = \mathbb{1}_n$. Also ist $A \circ A' \in \mathrm{Gl}_n(R)$, und $\mathrm{Gl}_n(R)$ ist somit abgeschlossen gegenüber der Operation $\circ$.

Wegen Korollar 8.9 gilt zudem, daß die Operation $\circ$ assoziativ auf $\mathrm{Mat}(n,R)$ ist, also erst recht auf der Teilmenge $\mathrm{Gl}_n(R)$.

Wir haben bereits in Lemma 8.11 gesehen, daß $\mathbb{1}_n \circ A = A$ für alle $A \in \mathrm{Gl}_n(R) \subset \mathrm{Mat}(n,R)$. Da zudem $\mathbb{1}_n \in \mathrm{Gl}_n(R)$ gilt, ist $\mathbb{1}_n$ das neutrale Element von $\mathrm{Gl}_n(R)$.

Sei $A \in \mathrm{Gl}_n(R)$, so gibt es per definitionem ein $B \in \mathrm{Mat}(n,R)$ mit

$$A \circ B = B \circ A = \mathbb{1}_n. \tag{8}$$

Wegen (8) gilt aber zum einen $B \in \mathrm{Gl}_n(R)$ und zum anderen erfüllt B die Bedingung des zu A gehörenden inversen Elementes.

Also ist $\big(\mathrm{Gl}_n(R), \circ\big)$ eine Gruppe, und die angegebenen Eigenschaften folgen aus Lemma 3.4. □

Bemerkung 8.15
Für $n = 1$ ist $(\mathrm{Gl}_n(R), \circ)$ nichts anderes als $(R^*, \cdot)$, d. h. die Abbildung

$$\mathrm{Gl}_n(R) \to R^* : (\lambda) \mapsto \lambda,$$

die eine 1×1-Matrix (λ) auf das Ringelement λ abbildet, ist ein Gruppenisomorphismus.

Für $n > 1$ ist die Matrixmultiplikation und damit $\mathrm{Gl}_n(R)$ i. a. nicht kommutativ. Für 2×2-Matrizen gilt etwa:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \circ \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} aa' + bc' & ab' + bd' \\ ca' + dc' & cb' + dd' \end{pmatrix}$$

und

$$\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \circ \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a'a + b'c & a'b + b'd \\ c'a + d'c & c'b + d'd \end{pmatrix},$$

was im allgemeinen ungleich ist. Sei etwa $R = \mathbb{R}$ dann gilt

$$\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \circ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix} \neq \begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}.$$

Man beachte, daß in obigem Beispiel die Multiplikation von rechts mit $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ die Spalten vertauscht, während Multiplikation mit der gleichen Matrix von links die Vertauschung der Zeilen bewirkt. Es handelt sich dabei um ein allgemeineres Phänomen, das wir nun näher untersuchen wollen.

Definition 8.16 a. Es sei $\lambda \in R^*$. Dann heißt die Abbildung

$$\mathrm{Mat}(m \times n, R) \to \mathrm{Mat}(m \times n, R) : \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_m \end{pmatrix} \mapsto \begin{pmatrix} a_1 \\ \vdots \\ \lambda a_i \\ \vdots \\ a_m \end{pmatrix},$$

die die i-te Zeile einer Matrix mit λ multipliziert, eine *elementare Zeilenoperation vom Typ I*.

b. Es sei $i \neq j$. Die Abbildung

$$\mathrm{Mat}(m \times n, R) \to \mathrm{Mat}(m \times n, R) : \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_m \end{pmatrix} \mapsto \begin{pmatrix} a_1 \\ \vdots \\ a_i + a_j \\ \vdots \\ a_m \end{pmatrix},$$

die die j-te Zeile zur i-ten Zeile addiert, nennt man eine *elementare Zeilenoperation vom Typ II*.

c. Es sei $\lambda \in R$ und $i \neq j$. Dann heißt die Abbildung

$$\mathrm{Mat}(m \times n, R) \to \mathrm{Mat}(m \times n, R) : \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_m \end{pmatrix} \mapsto \begin{pmatrix} a_1 \\ \vdots \\ a_i + \lambda a_j \\ \vdots \\ a_m \end{pmatrix},$$

die das λ-fache der j-ten Zeile zur i-ten addiert, eine *elementare Zeilenoperation vom Typ III*.

d. Es sei $i < j$. Dann heißt die Abbildung

$$\mathrm{Mat}(m \times n, R) \to \mathrm{Mat}(m \times n, R) : \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \\ a_m \end{pmatrix} \mapsto \begin{pmatrix} a_1 \\ \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \\ a_m \end{pmatrix},$$

die die i-te und j-te Zeile vertauscht, eine *elementare Zeilenoperation vom Typ IV*.

Entsprechend definiert man *elementare Spaltenoperationen*, indem man in den obigen Definitionen überall Zeile durch Spalte ersetzt.

Definition 8.17

Es seien $\lambda \in R^*$, $n > 0$ und $1 \leq i, j \leq n$ mit $i \neq j$. Wir definieren die folgenden quadratischen Matrizen in $\mathrm{Mat}(n, R)$:

a. $E_i^j := (e_{lk})_{l,k=1,\ldots,n}$ mit $e_{lk} = \delta_{il} \cdot \delta_{jk}$,
b. $S_i(\lambda) := \mathbb{1}_n + (\lambda - 1) \cdot E_i^i$,
c. $Q_i^j := \mathbb{1}_n + E_i^j$,
d. $Q_i^j(\lambda) := \mathbb{1}_n + \lambda \cdot E_i^j$, und
e. $P_i^j := \mathbb{1}_n - E_i^i - E_j^j + E_i^j + E_j^i$.

Die Matrizen $S_i(\lambda)$, Q_i^j und $Q_i^j(\lambda)$ heißen *Elementarmatrizen*, und die P_i^j heißen *Permutationsmatrizen*.

Die Namen sind durch folgendes Lemma gerechtfertigt. Der Beweis dieses und des folgenden Lemmas erfolgt durch einfaches Ausmultiplizieren und sei dem Leser als Übungsaufgabe überlassen.

Lemma 8.18

Es seien $\lambda \in R^*$, $1 \leq i, j \leq n$ *mit* $i \neq j$ *und* $A \in \mathrm{Mat}(n \times m, R)$. A_{I}, A_{II}, A_{III} *und* A_{IV} *mögen die durch folgende* elementare Zeilenoperationen *entstandenen Matrizen bezeichnen:*

A_I *entstehe durch Multiplikation der* i*-ten Zeile von* A *mit* λ*,*
A_{II} *entstehe durch Addition der* j*-ten Zeile von* A *zur* i*-ten Zeile von* A*,*
A_{III} *entstehe durch Addition des* λ*-fachen der* j*-ten Zeile von* A *zur* i*-ten Zeile von* A*, und*
A_{IV} *entstehe durch Vertauschen der* i*-ten Zeile von* A *mit der* j*-ten Zeile.*

Dann gilt:

$$A_I = S_i(\lambda) \circ A, \qquad A_{II} = Q_i^j \circ A,$$
$$A_{III} = Q_i^j(\lambda) \circ A, \qquad A_{IV} = P_i^j \circ A.$$

Bemerkung 8.19
Man kann analog *elementare Spaltenoperation* von A betrachten und definiert A^I, A^{II}, A^{III} und A^{IV}, indem man in der Definition von A_I, A_{II}, A_{III} und A_{IV} das Wort "Zeile" durch "Spalte" ersetzt. Dann gilt analog für $A \in \mathrm{Mat}(m \times n, R)$:

$$A^I = A \circ S_i(\lambda), \qquad A^{II} = A \circ Q_i^j,$$
$$A^{III} = A \circ Q_i^j(\lambda), \qquad A^{IV} = A \circ P_i^j.$$

Die Wirkung der Matrix $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ in Bemerkung 8.15 ist also ein Spezialfall der Permutationsmatrizen.

Lemma 8.20
Es seien $\lambda \in R^*$*,* $1 \leq i, j \leq n$ *mit* $i \neq j$ *und* $A \in \mathrm{Mat}(n \times m, R)$*. Dann gelten:*

$$Q_i^j(\lambda) = S_j\left(\lambda^{-1}\right) \circ Q_i^j \circ S_j(\lambda),$$

und

$$P_i^j = Q_j^i \circ Q_i^j(-1) \circ Q_j^i \circ S_j(-1).$$

Die Aussage dieses Lemmas läßt sich so interpretieren, daß die elementaren Zeilen-/Spaltenoperationen vom Typ III für $\lambda \in R^*$ und IV auch durch mehrfache Ausführung von Operationen vom Typ I und II erzielt werden können.

Die folgende Aussage benötigen wir später, um zu zeigen, daß die Lösungsmenge eines linearen Gleichungssystems invariant gegenüber elementaren Umformungen ist (siehe Satz II.7.13).

Proposition 8.21
Es seien $\lambda \in R^*$*,* $1 \leq i, j \leq n$ *mit* $i \neq j$*. Dann gelten:*

a. $S_i\left(\lambda^{-1}\right) \circ S_i(\lambda) = S_i(\lambda) \circ S_i\left(\lambda^{-1}\right) = \mathbb{1}_n$,
b. $Q_i^j(-1) \circ Q_i^j = Q_i^j \circ Q_i^j(-1) = \mathbb{1}_n$,
c. $Q_i^j(-\lambda) \circ Q_i^j(\lambda) = Q_i^j(\lambda) \circ Q_i^j(-\lambda) = \mathbb{1}_n$*, und*
d. $P_i^j \circ P_i^j = \mathbb{1}_n$.

D. h. die Elementar- und die Permutationsmatrizen sind invertierbar und die Inversen sind wiederum Elementar- / Permutationsmatrizen vom gleichen Typ.

Beweis: Wir führen den Beweis für c. vor. Die übrigen Teile lassen sich dann analog zeigen.

Für $\lambda \in R^*$ gilt, vermittels der Distributivität der Matrixmultiplikation:

$$Q_i^j(-\lambda) \circ Q_i^j(\lambda) = \left(\mathbb{1}_n - \lambda \cdot E_i^j\right) \circ \left(\mathbb{1}_n + \lambda \cdot E_i^j\right) = \mathbb{1}_n - \lambda^2 \cdot E_i^j \circ E_i^j = \mathbb{1}_n,$$

da $E_i^j \circ E_i^j = 0$ wegen $i \neq j$. Beachte dazu, daß für $E_i^j \circ E_i^j = (c_{lk})$ gilt:

$$c_{lk} = \sum_{p=1}^{n} \delta_{il}\delta_{jp}\delta_{ip}\delta_{jk},$$

und daß für $i \neq j$ und p beliebig gilt $\delta_{jp}\delta_{ip} = 0$.

Vertauschen von λ und $-\lambda$ zeigt, daß auch $Q_i^j(\lambda) \circ Q_i^j(-\lambda) = \mathbb{1}_n$ gilt. □

Bemerkung 8.22
Die Proposition impliziert also, daß die Elementarmatrizen und die Permutationsmatrizen Elemente der allgemeinen linearen Gruppe $Gl_n(R)$ sind. Im Falle, daß $R = K$ ein Körper ist, gilt aber noch mehr. Wir zeigen später, daß die Elementarmatrizen die $Gl_n(K)$ erzeugen - vgl. dazu Bemerkung II.5.30.

Abschließend beweisen wir einen Hilfssatz, der beim Rechnen mit Matrizen Anwendung finden wird.

Lemma 8.23 a. *Für* $A \in \mathrm{Mat}(m \times n, R)$ *und* $B \in \mathrm{Mat}(n \times p, R)$ *gilt:*

$$(A \circ B)^t = B^t \circ A^t.$$

b. *Für* $A \in Gl_n(R)$ *gilt:*

$$\left(A^t\right)^{-1} = \left(A^{-1}\right)^t.$$

Beweis: a. Seien $A = (a_{ij}) \in \mathrm{Mat}(m \times n, R)$ und $B = (b_{jk}) \in \mathrm{Mat}(n \times p, R)$, dann ist $A \circ B = (c_{ik}) \in \mathrm{Mat}(m \times p, R)$ mit

$$c_{ik} = \sum_{j=1}^{n} a_{ij}b_{jk}.$$

Ferner ist $B^t \circ A^t = (d_{ki}) \in \mathrm{Mat}(p \times m, R)$ mit

$$d_{ki} = \sum_{j=1}^{n} b_{jk}a_{ij}.$$

Also gilt $c_{ik} = d_{ki}$ und damit

$$(A \circ B)^t = (c_{ik})^t = (d_{ki}) = B^t \circ A^t.$$

b. Aus $A \circ A^{-1} = A^{-1} \circ A = \mathbb{1}_n$ folgt mit Teil a.

$$\left(A^{-1}\right)^t \circ A^t = A^t \circ \left(A^{-1}\right)^t = \mathbb{1}_n,$$

also ist A^t invertierbar mit Inverser $\left(A^{-1}\right)^t$.

□

AUFGABEN

Aufgabe 8.24 (Diëdergruppe vom Grad 8)
Es sei $G = Gl_2(\mathbb{R})$. Wir setzen $A := \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right) \in G$, $B := \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right) \in G$ und $E := \mathbb{1}_2$. Prüfe die folgenden Eigenschaften nach:

a. $A^4 = B^2 = E$ und $B \cdot A \cdot B^{-1} = A^{-1}$.
b. $\mathbb{D}_8 := \{A^i \cdot B^j \mid i = 0,1,2,3;\ \ j = 1,2\}$ ist eine Untergruppe von G.
c. $|\mathbb{D}_8| = 8$.

Hinweis: Wenn man geschickt vorgeht, kommt man in Teil b. ohne eine einzige Matrixmultiplikation aus!

Aufgabe 8.25 (Diëdergruppe vom Grad 8)
Die Diëdergruppe $\mathbb{D}_8$ besitzt genau drei Untergruppen der Ordnung 4 und 5 Untergruppen der Ordnung 2.

Aufgabe 8.26 (Diëdergruppe vom Grad 6)
Es sei $G = Gl_2(\mathbb{R})$. Wir setzen $A := \begin{pmatrix} \frac{1}{2}\sqrt{3} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2}\sqrt{3} \end{pmatrix} \in G$, $B := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \in G$ und $E := \mathbb{1}_2$. Prüfe die folgenden Eigenschaften nach:

a. $A^3 = B^2 = E$ und $B^{-1} \cdot A \cdot B = A^{-1}$.
b. $\mathbb{D}_6 := \{A^j,\ A^j \cdot B \mid j = 0,1,2\}$ ist eine Untergruppe von G mit $|\mathbb{D}_6| = 6$.
c. $\mathbb{D}_6$ besitzt genau drei Untergruppen der Ordnung zwei und eine Untergruppe der Ordnung drei.
d. $\mathbb{D}_6$ ist isomorph zur Symmetrischen Gruppe $\mathbb{S}_3$.

Anmerkung: Die Diëdergruppe vom Grad 6 ist die Symmetriegruppe des gleichseitigen Dreiecks.

Aufgabe 8.27 (Quaternionengruppe)
Es sei $G = Gl_2(\mathbb{C})$. Wir setzen $A := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in G$, $B := \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix} \in G$ und $E := \mathbb{1}_2$. Prüfe die folgenden Eigenschaften nach:

a. $A^4 = B^4 = E$ und $B^{-1} \cdot A \cdot B = A^{-1}$.
b. $\mathbb{Q}_8 := \{A^j,\ A^j \cdot B \mid j = 0,1,2,3\}$ ist eine Untergruppe von G mit $|\mathbb{Q}_8| = 8$.
c. $\mathbb{Q}_8$ besitzt genau drei Untergruppen der Ordnung 4 und nur eine Untergruppe der Ordnung 2.
d. Alle Untergruppen von $\mathbb{Q}_8$ sind Normalteiler.

Anmerkung: Die Quaternionengruppe ist neben der Diëdergruppe der Ordnung 8 bis auf Isomorphie die einzige nicht-abelsche Gruppe der Ordnung 8.

Aufgabe 8.28
Es sei R ein beliebiger Ring, $A = \left(a_{ij}\right)_{i,j=1,\ldots,n} \in \mathrm{Mat}(n, R)$ mit $a_{ij} = 0$ für $i \geq j$. Zeige, $A^n = 0$.

9. Der Gauß-Algorithmus

In diesem Abschnitt wollen wir zeigen, daß man jede Matrix durch elementare Zeilenoperationen in Zeilen-Stufen-Form transformieren kann. Wir zeigen aber nicht nur, daß das im Prinzip möglich ist, sondern wir geben ein Verfahren an, wie man die elementaren Zeilenoperationen durchzuführen hat, so daß man nach endlich vielen Schritten die Zeilen-Stufen-Form erhält. Ein solches konkretes Verfahren nennt man einen *Algorithmus* - in diesem Fall den *Gauß-Algorithmus*.

Generalvoraussetzung: In diesem Abschnitt sei K ein Körper.

Wir definieren jetzt exakt, was es heißt, daß eine Matrix Zeilen-Stufen-Form hat.

Definition 9.1

Es sei $A = (a_{ij}) \in \mathrm{Mat}(m \times n, K)$.

a. A besitzt *Zeilen-Stufen-Form*, kurz ZSF, falls es ein r, mit $0 \leq r \leq m$ und Indizes $j_1, \ldots, j_r$ mit $1 \leq j_1 < j_2 < \ldots < j_r \leq m$ gibt, so daß folgendes gilt:
   (i) $a_{ij} = 0$ für $1 \leq i \leq r$ und $1 \leq j < j_i$,
   (ii) $a_{ij} = 0$ für $r < i \leq m$ und $j = 1, \ldots, n$, und
   (iii) $a_{ij_i} \neq 0$ für $i = 1, \ldots, r$.
   Die Zahl r nennen wir den *Zeilenrang*[11] von A, die Körperelemente a_{ij_i} heißen die *Pivots* oder *Angelpunkte* der Zeilen-Stufen-Form.

b. Eine Zeilen-Stufen-Form von A heißt *reduziert* oder *speziell*, falls zusätzlich gilt:
   (iv) $a_{ij_i} = 1$ für $i = 1, \ldots, r$, und
   (v) $a_{kj_i} = 0$ für $k < i$ und $i = 1, \ldots, r$.

Bemerkung 9.2

Eine Matrix A in Zeilen-Stufen-Form ist also von der folgenden Gestalt:

$$A = \begin{pmatrix}
0 & \ldots & 0 & a_{1j_1} & * & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & * \\
0 & \ldots & \ldots & 0 & \ldots & 0 & a_{2j_2} & * & \ldots & \ldots & \ldots & \ldots & \ldots & * \\
0 & \ldots & \ldots & \ldots & \ldots & \ldots & 0 & \ldots & 0 & a_{3j_3} & * & \ldots & \ldots & * \\
\vdots & & & & & & & & & & \ddots & & & \vdots \\
0 & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & 0 & a_{rj_r} & * & \ldots & * \\
0 & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & 0 \\
\vdots & & & & & & & & & & & & & & \vdots \\
0 & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & 0
\end{pmatrix}$$

Hat A reduzierte Zeilen-Stufen-Form, so sind die Pivots alle Eins und die Einträge in der Spalte oberhalb der Pivots sind alle Null.

Der Zeilenrang von A darf Null sein. Das ist genau dann der Falle, wenn A die Nullmatrix ist.

Beispiel 9.3

Betrachte die Matrizen $A, B, C \in \mathrm{Mat}(4 \times 5, K)$ mit

$$A = \begin{pmatrix}
0 & 1 & 0 & 3 & 0 \\
0 & 0 & 1 & 2 & 0 \\
0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0
\end{pmatrix}, B = \begin{pmatrix}
1 & 2 & 3 & 4 & 5 \\
0 & 2 & 3 & 4 & 5 \\
0 & 0 & 3 & 0 & 0 \\
0 & 0 & 0 & 4 & 1
\end{pmatrix}, \text{ und } C = \begin{pmatrix}
0 & 1 & 0 & 0 & 0 \\
1 & 0 & 3 & 4 & 0 \\
0 & 0 & 2 & 2 & 0 \\
0 & 0 & 0 & 0 & 0
\end{pmatrix}.$$

Die Matrix A ist in reduzierter ZSF mit $\mathrm{Zrang}(A) = r = 3$, $j_1 = 2$, $j_2 = 3$ und $j_3 = 5$.

[11] Vgl. Bemerkung II.5.24.

Die Matrix B ist in ZSF mit $\operatorname{Zrang}(B) = r = 4$ und $j_1 = 1$, $j_2 = 2$, $j_3 = 3$ und $j_4 = 4$. Die ZSF ist aber nicht reduziert.

Die Matrix C ist nicht in ZSF. Aber durch Vertauschen der beiden ersten Zeilen entsteht eine Matrix, die ZSF hat.

Wir wollen nun zeigen, daß man Matrizen über einem Körper K mittels elementarer Zeilenoperationen auf Zeilen-Stufen-Form bringen kann.

Satz 9.4

Jede Matrix $A \in \operatorname{Mat}(m \times n, K)$ läßt sich mittels endlich vieler elementarer Zeilenoperationen in reduzierte Zeilen-Stufen-Form überführen.

Diese ist eindeutig bestimmt und wird mit $\operatorname{RZSF}(A)$ *bezeichnet.*

Bemerkung 9.5

Der Beweis des Satzes ist konstruktiv, daß heißt, aus dem Beweis läßt sich ein Algorithmus zur Berechnung einer ZSF von A herleiten, der sogenannte *Gauß-Algorithmus*.

Man beachte, daß Zeilen-Stufen-Formen i. a. nur über einem *Körper* existieren, da z. B. für die Auslöschung der Einträgen unterhalb der Pivots die Division durch beliebige Elemente ungleich Null notwendig ist.

Beweis von Satz 9.4: Sei $A \in \operatorname{Mat}(m \times n, K)$. Ist $A = 0$, so hat A bereits ZSF mit $r = 0$ und wir sind fertig.

Ist $A \neq 0$, so führe folgende Schritte durch:

1. **Schritt:** Durchlaufe die Spalten von oben nach unten, mit der ersten Spalte beginnend, bis der erste Eintrag $a_{i_1 j_1} \neq 0$ gefunden ist:

$$\begin{pmatrix} 0 & \dots & 0 & 0 & * & \dots & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \vdots & & \vdots & 0 & * & \dots & * \\ \vdots & & \vdots & a_{i_1 j_1} & * & \dots & * \\ \vdots & & \vdots & * & * & \dots & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & * & * & \dots & * \end{pmatrix}$$

2. **Schritt:** Steht $a_{i_1 j_1}$ nicht in der ersten Zeile, d. h. $i_1 \neq 1$, dann vertausche die Zeilen a_1 und a_{i_1} - Zeilenoperation vom Typ IV. Die so entstandene Matrix heiße $\widetilde{A}_1 = (\tilde{a}_{ij})$. Dann ist $\tilde{a}_{1 j_1}$ unser erstes Pivot.
3. **Schritt:** Erzeuge in der Spalte $\tilde{a}^{j_1}$ von $\widetilde{A}_1$ unterhalb von $\tilde{a}_{1 j_1}$ Nullen durch elementare Operationen vom Typ III, d. h. addiere für $k = 2, \dots, m$ zur k-ten Zeile das $-\frac{\tilde{a}_{k j_1}}{\tilde{a}_{1 j_1}}$-fache der ersten Zeile. Die Spalten mit Index

kleiner als j_1 werden dadurch nicht geändert. Das Ergebnis ist dann eine Matrix von der Form:

$$A^{(1)} := \left(\begin{array}{cccc|ccc} 0 & \dots & 0 & a^{(1)}_{1j_1} & * & \dots & * \\ \hline 0 & \dots & 0 & 0 & & & \\ \vdots & & \vdots & \vdots & & A_2 & \\ 0 & \dots & 0 & 0 & & & \end{array}\right),$$

wobei A_2 eine $(m-1)\times(n-j_1)$-Matrix ist, sofern $j_1 \neq n$.

Ist $n - j_1 = 0$ oder $m - 1 = 0$ oder $A^{(2)} = 0$, so sind wir fertig.

Andernfalls ist $A_2 \neq 0$, und wir führen Schritt 1-3 mit A_2 durch. Dabei kann man alle Zeilenoperationen auf die Matrix $A^{(1)}$ ausdehnen, ohne daß sich in den ersten j_1 Spalten etwas ändert, da dort nur Nullen stehen. Ist A_2 umgeformt, so erhält man eine Matrix $A^{(2)}$ der Form:

$$A^{(2)} = \left(\begin{array}{cccccccc|ccc} 0 & \dots & 0 & a^{(2)}_{1j_1} & * & \dots & \dots & & \dots & \dots & * \\ 0 & \dots & 0 & 0 & \dots & 0 & a^{(2)}_{2j_2} & & * & \dots & * \\ \hline 0 & \dots & 0 & 0 & \dots & 0 & 0 & & & & \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & & A_3 & \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & & & & \end{array}\right)$$

mit einem Pivot $a^{(2)}_{2j_2}$ und, sofern nicht $m - 2 = 0$ oder $n - j_2 = 0$, einer Matrix A_3, die eine Zeile und mindestens eine Spalte weniger als A_2 hat.

Ist $A_3 = 0$, so sind wir fertig. ansonsten fahren wir fort wie bisher und erhalten Matrizen $A^{(3)}$, A_4, $A^{(4)}$, etc.. Das Verfahren stoppt, falls nach r-maligem Durchlaufen der Schritte 1-3 entweder $r = m$ oder $r = n$ oder $A_{r+1} = 0$. In jedem der drei Fälle ist die Matrix $A^{(r)}$ in ZSF.

Um die Matrix $A^{(r)} = \left(a^{(r)}_{ij}\right)$ in reduzierte ZSF zu bringen, multiplizieren wir zunächst die Zeilen $a^{(r)}_i$, für $i = 1, \dots, r$, mit $\frac{1}{a^{(r)}_{ij_i}}$, was einer elementaren Zeilenoperation vom Typ I entspricht. Die so entstehende Matrix heiße $A' = (a'_{ij})$. Sodann addiert man für $i = 1, \dots, r$ und $k = 1, \dots, i-1$ zur k-ten Zeile das $-a'_{kj_i}$-fache der i-ten Zeile – elementare Operationen vom Typ III – und nennt in jedem Schritt die neue Matrix wieder A'. Man sieht unmittelbar, daß die entstehende Matrix $A'' = \left(a''_{ij}\right)$ reduzierte ZSF hat, da in Spalte j_i die Elemente a'_{kj_i} in $a''_{kj_i} = 0$, für $k < i$, übergegangen sind.

Damit ist die Existenz einer reduzierten ZSF gezeigt. Es bleibt also, deren Eindeutigkeit zu zeigen. Den Beweis dafür können wir allerdings erst in Kapitel II.5 führen, siehe Seite 143 □

Aus Satz 9.4 folgt unter Berücksichtigung von Lemma 8.18 sowie der Tatsache, daß jede Permutationsmatrix das Produkt von Elementarmatrizen ist, unmittelbar Korollar 9.6.

Korollar 9.6
Es sei $A \in \mathrm{Mat}(m \times n, K)$*, dann gibt es Elementarmatrizen* $T_1, \ldots, T_k$*, so daß*

$$\mathrm{RZSF}(A) = T_1 \circ \ldots \circ T_k \circ A.$$

Insbesondere hat $T_1 \circ \ldots \circ T_k \circ A$ *also reduzierte ZSF.*

Beispiel 9.7
Wir überführen nun die folgende Matrix in reduzierte ZSF.

$$\begin{pmatrix} 0 & 0 & -1 & 2 & 3 \\ \underline{-1} & 1 & -3 & 0 & 2 \\ 1 & -1 & 1 & 4 & 3 \end{pmatrix} \xrightarrow{\mathrm{I} \leftrightarrow \mathrm{II}} \begin{pmatrix} \underline{-1} & 1 & -3 & 0 & 2 \\ 0 & 0 & -1 & 2 & 3 \\ 1 & -1 & 1 & 4 & 3 \end{pmatrix} \xrightarrow{\mathrm{III} \mapsto \mathrm{III}+\mathrm{I}}$$

$$\left(\begin{array}{cc|ccc} -1 & 1 & -3 & 0 & 2 \\ \hline 0 & 0 & \underline{-1} & 2 & 3 \\ 0 & 0 & -2 & 4 & 5 \end{array}\right) \xrightarrow{\mathrm{III} \mapsto \mathrm{III}-2\cdot\mathrm{II}} \begin{pmatrix} \underline{-1} & 1 & -3 & 0 & 2 \\ 0 & 0 & \underline{-1} & 2 & 3 \\ 0 & 0 & 0 & 0 & \underline{-1} \end{pmatrix} \xrightarrow[\substack{\mathrm{II} \mapsto -\mathrm{II} \\ \mathrm{III} \mapsto -\mathrm{III}}]{\mathrm{I} \mapsto -\mathrm{I}}$$

$$\begin{pmatrix} 1 & -1 & 3 & 0 & -2 \\ 0 & 0 & 1 & -2 & -3 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow[\mathrm{II} \mapsto \mathrm{II}+3\cdot\mathrm{III}]{\mathrm{I} \mapsto \mathrm{I}+2\cdot\mathrm{III}} \begin{pmatrix} 1 & -1 & 3 & 0 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{\mathrm{I} \mapsto \mathrm{I}-3\cdot\mathrm{II}}$$

$$\begin{pmatrix} 1 & -1 & 0 & 6 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Die vierte Matrix besitzt bereits ZSF mit unterstrichenen Pivots, die letzte ist in reduzierter ZSF.

Wir bemerken, daß wir auch auf anderem Weg zum Ziel gekommen wären, und zwar durch andere Wahl der Pivots.

$$\begin{pmatrix} 0 & 0 & -1 & 2 & 3 \\ -1 & 1 & -3 & 0 & 2 \\ \underline{1} & -1 & 1 & 4 & 3 \end{pmatrix} \xrightarrow{\mathrm{I} \leftrightarrow \mathrm{III}} \begin{pmatrix} 1 & -1 & 1 & 4 & 3 \\ -1 & 1 & -3 & 0 & 2 \\ 0 & 0 & -1 & 2 & 3 \end{pmatrix} \xrightarrow{\mathrm{II} \mapsto \mathrm{II}+\mathrm{I}}$$

$$\left(\begin{array}{cc|ccc} 1 & -1 & 1 & 4 & 3 \\ \hline 0 & 0 & \underline{-2} & 4 & 5 \\ 0 & 0 & -1 & 2 & 3 \end{array}\right) \xrightarrow{\mathrm{III} \mapsto \mathrm{III}-\frac{1}{2}\cdot\mathrm{II}} \begin{pmatrix} \underline{1} & -1 & 1 & 4 & 3 \\ 0 & 0 & \underline{-2} & 4 & 5 \\ 0 & 0 & 0 & 0 & \underline{\frac{1}{2}} \end{pmatrix} \xrightarrow[\mathrm{III} \mapsto 2\cdot\mathrm{III}]{\mathrm{II} \mapsto -\frac{1}{2}\cdot\mathrm{II}}$$

$$\begin{pmatrix} 1 & -1 & 1 & 4 & 3 \\ 0 & 0 & 1 & -2 & +\frac{5}{2} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow[\mathrm{II} \mapsto \mathrm{II}-\frac{5}{2}\cdot\mathrm{III}]{\mathrm{I} \mapsto \mathrm{I}-3\cdot\mathrm{III}} \begin{pmatrix} 1 & -1 & 1 & 4 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{\mathrm{I} \mapsto \mathrm{I}-\mathrm{II}}$$

$$\begin{pmatrix} 1 & -1 & 0 & 6 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

In der Praxis sind 1000×1000-Matrizen keine Seltenheit. Dort wird mit einer festen Stellenzahl gerechnet und deshalb treten bei großen Matrizen unter Umständen erhebliche Rundungsfehler auf. Es kommt der Wahl der richtigen Pivots eine große Bedeutung zu. Ist das gewählte Pivot zu klein, so kann bei Division durch dieses Pivot im dritten Schritt der Rundungsfehler riesig werden - für den Computer bedeutet dies in etwa, als ob man durch Null zu dividieren versuche. Deshalb wählt man in der Praxis das betragsmäßig größte Element als Pivot.

Rechnet man allerdings in Computeralgebrasystemen mit exakter Arithmetik, so spielt die Auslöschung durch Rundungsfehler keine Rolle. Dort muß man eher dafür sorgen, daß die Zahlen, d. h. die Zähler und Nenner, nicht zu groß werden, da dies zu erheblichen Geschwindigkeitsverlusten führen würde.

Wir wollen abschließend den Gauß-Algorithmus in leicht abgewandelter Form als rekursiven Algorithmus zur Bestimmung der reduzierten ZSF einer Matrix formulieren.

Algorithmus 9.8 (Gauß-Algorithmus)
INPUT: $A \in \mathrm{Mat}(m \times n, K)$.

OUTPUT: $\mathrm{RZSF}(A)$, die reduzierte Zeilen-Stufen-Form von A.

1. **Schritt:** Falls $A = 0$ oder $m = 1$, gehe zu Schritt 8.
2. **Schritt:** Durchlaufe die erste Spalte von oben nach unten, bis ein Element ungleich Null a_{i1} gefunden wurde oder das Ende der Spalte erreicht ist.
3. **Schritt:** Wurde kein $a_{i1} \neq 0$ gefunden, bilde eine Untermatrix B von A durch Streichen der ersten Spalte von A und gehe zu Schritt 6. Andernfalls, vertausche die Zeilen a_1 und a_i.
4. **Schritt:** Für $k = 2, \ldots, m$ addiere zur k-ten Zeile das $-\frac{a_{k1}}{a_{11}}$-fache der ersten Zeile.
5. **Schritt:** Falls $n = 1$, gehe zu Schritt 7. Andernfalls bilde eine Untermatrix $B := A^1_1$ von A, durch Streichen der ersten Zeile und der ersten Spalte von A.
6. **Schritt:** Wende den Algorithmus auf die Untermatrix B an.[12]

[12] Dies ist der Rekursionsschritt, indem der Algorithmus mit einer kleineren Untermatrix aufgerufen wird. Das Ergebnis, das man dabei zurück erhält, wird wieder in die Matrix A eingefügt. – Für eine Definition von A^1_1 siehe auch Definition II.10.2.

7. Schritt: Die Matrix A liegt nun in ZSF vor. Für $i = m$ bis $i = 1$, d. h. rückwärts zählend, durchlaufe die Zeile a_i, beginnend mit der ersten Spalte, bis ein Element $a_{ij} \neq 0$ gefunden wurde oder das Ende der Zeile erreicht ist.
In letzterem Fall tue nichts, in ersterem multipliziere die Zeile a_i mit $\frac{1}{a_{ij}}$ und addiere für $k = 1, \ldots, i-1$ zur k-ten Zeile das $-a_{kj}$-fache der i-ten Zeile.

8. Schritt: Gib die (veränderte) Matrix A zurück.

AUFGABEN

Aufgabe 9.9

Bestimme eine ZSF sowie die RZSF von den folgenden Matrizen:

$$\begin{pmatrix} 1 & 3 & 0 \\ 3 & 0 & 1 \\ 2 & 1 & 0 \\ 0 & 3 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 & 4 & 8 \\ 0 & 3 & 2 & 8 & 5 \\ 2 & 3 & 4 & 1 & 0 \end{pmatrix}.$$

Aufgabe 9.10

Schreibe eine Singular-Prozedur `gauss_reduction`, die eine Matrix A einliest und die mittels Gauß-Elimination ermittelte Zeilen-Stufen-Form der Matrix ausgibt. Die Einträge der Matrizen sollen vom Typ `poly` sein. Es dürfen die in `matrix.lib` definierten Prozeduren `multrow`, `addrow` und `permrow` verwendet werden. Teste Deine Ergebnisse mit der Prozedur `gnf` aus der Bibliothek `linalg.lib`.

Aufgabe 9.11

Ändere die Prozedur `gauss_reduction` aus Aufgabe 9.10 so ab, daß sie eine Matrix und eine ganze Zahl einliest. Ist die Zahl `0`, so soll die Prozedur (wie bisher) die nicht reduzierte Zeilen-Stufen-Form der Matrix ausgeben. Ansonsten soll nun die vollständig reduzierte Zeilen-Stufenform zurückgegeben werden.

Aufgabe 9.12

Bestimme die vollständig reduzierte Zeilen-Stufen-Form folgender 10×15-Matrizen A und B:

a. Definiere A durch:
```
ring r = 0,x,dp;
ideal I = 1..150;
matrix A[10][15] = I;
```
b. B entstehe aus A dadurch, daß in A die Diagonalelemente zu 1 gesetzt werden.

c. Man beweise das Ergebnis für A ohne Zuhilfenahme von Singular.

Hinweis: Durch den Befehl `pagewidth=500;` wird erreicht, daß die Matrizen vollständig angezeigt werden. B läßt sich aus A mit Hilfe einer Schleife erzeugen.

KAPITEL II

Vektorräume

In diesem Kapitel werden Vektorräume und lineare Abbildungen über beliebigen Körpern eingeführt. Wesentlich ist der abstrakte Zugang ohne Basen, der die strukturellen und weniger die rechnerischen Aspekte in den Vordergrund stellt.

Den Begriff der Basis eines Vektorraumes werden wir definieren und zeigen, daß jeder Vektorraum eine Basis besitzt. Nach Auswahl einer Basis verhalten sich dann (endlich-dimensionale) Vektorräume wie der $\mathbb{R}^n$ und lineare Abbildungen wie Matrizen. Allerdings besitzt ein Vektorraum viele Basen und je nach Wahl der Basis ändert sich das *Erscheinungsbild* der Vektoren und der Matrizen; wie sich das ändert, werden wir auch untersuchen.

Für uns wichtig sind in diesem Kapitel diejenigen Eigenschaften von Vektorräumen und linearen Abbildungen, die unabhängig vom Erscheinungsbild, also basisunabhängig, sind. Man nennt dies die *invariante Theorie* der Vektorräume. Nicht nur aus strukturellen Gründen, auch für die Anwendungen ist es wichtig, zu wissen, welche Eigenschaften invariant sind, sich also bei Basiswechsel nicht ändern.

Ein wichtiges Beispiel eines invarianten Begriffs ist der Rang einer Matrix. Wir werden mit Hilfe der abstrakten Theorie der Vektorräume den Rang für lineare Abbildungen definieren, so daß er per definitionem invariant ist, und wir werden dann zeigen, daß man ihn mit Hilfe des gaußschen Algorithmus aus der Zeilen-Stufen-Form einer jeden darstellenden Matrix ablesen kann.

Dies allein zeigt schon den Wert der abstrakten Begriffsbildungen. Die Begriffe werden viel klarer, letztlich einfacher und unbelastet von speziellen Wahlen, schwerfälligen Notationen, etc.. Der wichtigste Vorteil der abstrakten Theorie ist jedoch, daß in konkreten Anwendungen Vektorräume oft nicht mit einer Basis gegeben sind, sondern man weiß nur, daß die untersuchten Objekte die Struktur eines Vektorraumes besitzen, ja manchmal ist eine Basis gar nicht bekannt.

Will man jedoch konkret etwas ausrechnen, z. B. den Rang einer gegebenen linearen Abbildung, dann braucht man wieder Basen und Matrizen. Wir werden deshalb am Schluß dieses Kapitels zeigen, wie man z. B. Basen bestimmen kann, und dann, mit Hilfe des gaußschen Algorithmus, die vorher abstrakt definierten Begriffe konkret berechnen kann.

1. Vektorräume und Moduln

Ein Vektorraum ist nicht nur eine Menge mit einer Struktur, wie etwa eine Gruppe, ein Ring oder ein Körper, sondern zu einem Vektorraum gehört immer ein Körper, der auf dem Vektorraum operiert. Da viele der Definitionen ohne Mehraufwand für Ringe statt für Körper gelten, behandeln wir diesen allgemeineren Fall gleich mit. Statt von einem Vektorraum spricht man dann von einem Modul.

Generalvoraussetzung: Sofern keine weiteren Angaben gemacht sind, sei K in diesem Abschnitt ein kommutativer Ring mit Eins.

Zwar machen die Definitionen auch für beliebige Ringe Sinn, und die meisten Aussagen gelten ebenfalls, aber die Beweise sind teilweise länglicher.

Definition 1.1

Ein K-*Modul* (oder Modul über K) besteht aus einer nicht-leeren Menge V sowie einer Verknüpfung

$$+ : V \times V \to V : (x, y) \mapsto x + y,$$

die *Vektoraddition* genannt wird, und einer Operation von K auf V, d. h. einer Abbildung

$$\cdot : K \times V \to V : (\lambda, x) \mapsto \lambda \cdot x = \lambda x,$$

die *Skalarmultiplikation* genannt wird, so daß die folgenden Gesetze gelten:

a. $(V, +)$ ist eine abelsche Gruppe,

b. für $\lambda, \mu \in K$ und $x, y \in V$ gelten:

(i) $(\lambda + \mu)x = \lambda x + \mu x$, ("*verallgemeinertes Distributivgesetz*")

(ii) $\lambda(x + y) = \lambda x + \lambda y$, und ("*verallgemeinertes Distributivgesetz*")

(iii) $(\lambda\mu)x = \lambda(\mu x)$. ("*verallgemeinertes Assoziativgesetz*")

Besitzt K eine Eins - wie z. B. jeder Körper - so fordern wir zudem:

b. (iv) $1 \cdot x = x$ für alle $x \in V$.

Ist K ein Körper, so nennt man einen K-Modul auch einen K-*Vektorraum* oder einen Vektorraum über K.

Die Elemente aus V nennt man *Vektoren* und die aus K *Skalare*. Der *Nullvektor*, d. h. das neutrale Element aus V bezüglich der Addition, wird mit 0 bzw. mit 0_V bezeichnet und das neutrale Element von $(K, +)$ ebenfalls mit 0 bzw. mit 0_K.

Beispiel 1.2

a. Der *Nullmodul* $V = \{0\}$ mit $\lambda \cdot 0 = 0$ für alle $\lambda \in K$ ist für jeden Ring K ein K-Modul. Man bezeichnet den Nullmodul auch mit K^0.

b. Der Ring (bzw. Körper) K selbst mit der Ringaddition als Vektoraddition und der Ringmultiplikation als Skalarmultiplikation ist ein K-Modul (bzw. K-Vektorraum).

c. Für $n \geq 1$ ist die Menge K^n (siehe Definition I.8.1) mit der in Definition I.8.2 definierten komponentenweisen Addition und Skalarmultiplikation ein K-Modul, bzw. ein K-Vektorraum, falls K ein Körper ist.
Speziell sind $\mathbb{R}^n$, $\mathbb{C}^n$ und $\mathbb{Z}_p^n$ (p eine Primzahl) Vektorräume über $\mathbb{R}$, $\mathbb{C}$ bzw. $\mathbb{Z}_p$. Ferner ist $\mathbb{Z}^n$ ein $\mathbb{Z}$-Modul.

d. Allgemeiner ist für $n, m \geq 1$ die Menge $\mathrm{Mat}(m \times n, K)$ der $m \times n$-Matrizen über K mittels der in Definition I.8.2 a. definierten Addition und Skalarmultiplikation ein K-Modul mit der *Nullmatrix*

$$0 := 0_n := \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix}$$

als neutralem Element bezüglich der Addition.

e. Ist K ein Unterkörper des Körpers L, so ist L mit der Körperaddition als Vektoraddition und mit der Körpermultiplikation (eingeschränkt auf $K \times L$) als Skalarmultiplikation ein K-Vektorraum. Entsprechendes gilt natürlich für Ringe.
Z. B. können wir auf diese Weise $\mathbb{R}$ als $\mathbb{Q}$-Vektorraum oder $\mathbb{C}$ als $\mathbb{R}$-Vektorraum auffassen.

f. Ist M eine Menge und K ein Ring (bzw. Körper), so wird der Ring K^M (siehe Beispiel I.7.2) mit der Ringaddition als Vektoraddition und der folgenden Operation

$$\cdot : K \times K^M \to K^M : (\lambda, f) \mapsto \big(\lambda f : M \to K : x \mapsto \lambda \cdot f(x)\big)$$

als Skalarmultiplikation ein K-Modul (bzw. K-Vektorraum), wie man sofort sieht.

g. Allgemeiner zeigt man, daß für eine Menge M und einen K-Modul V durch die Operationen

$$+ : V^M \times V^M \to V^M : (f, g) \mapsto \big(f + g : M \to V : x \mapsto f(x) + g(x)\big)$$

und

$$\cdot : K \times V^M \to V^M : (\lambda, f) \mapsto \big(\lambda f : M \to V : x \mapsto \lambda \cdot f(x)\big)$$

die Menge V^M aller Abbildungen von M nach V zu einem K-Modul wird.

h. Da man für $M = \{1, \dots, n\}$ eine Abbildung $f : M \to K$ in eindeutiger Weise durch die Familie von Bildelementen $\big(f(1), \dots, f(n)\big)$ beschreiben kann, sieht man leicht, daß die Zuordnung

$$K^M \to K^n : f \mapsto \big(f(1), \dots, f(n)\big)^t$$

in diesem Falle eine Bijektion ist. Man prüft überdies leicht nach, daß diese Abbildung ein Vektorraumhomomorphismus im Sinne von Definition 1.5 ist. K^M und K^n sind dann also isomorph.

Lemma 1.3

In einem K-Vektorraum gelten folgende Rechenregeln:

a. $0_K \cdot x = 0_V$ *und* $\lambda \cdot 0_V = 0_V$ *für alle* $x \in V$, $\lambda \in K$.
b. *Für* $\lambda \in K$ *und* $x \in V$ *gilt:*

$$\lambda \cdot x = 0_V \;\Rightarrow\; \lambda = 0 \quad \textit{oder} \quad x = 0.$$

c. $(-1) \cdot x = -x$ *für alle* $x \in V$.

Beweis: Es seien $x \in V$ und $\lambda \in K$ gegeben.

a. Es gilt:

$$0_V + 0_K \cdot x = 0_K \cdot x = (0_K + 0_K) \cdot x = 0_K \cdot x + 0_K \cdot x,$$

also $0_V = 0_K \cdot x$, wie aus den Kürzungsregeln für $(V, +)$ folgt. Analog gilt:

$$0_V + \lambda \cdot 0_V = \lambda \cdot 0_V = \lambda \cdot (0_V + 0_V) = \lambda \cdot 0_V + \lambda \cdot 0_V,$$

und damit $0_V = \lambda \cdot 0_V$.

b. Ist $\lambda \in K$ mit $\lambda \neq 0$, dann gibt es ein Inverses $\lambda^{-1} \in K$. Aus $\lambda \cdot x = 0$ folgt dann aber wegen a. und den Vektorraumaxiomen

$$0_V = \lambda^{-1} \cdot 0_V = \lambda^{-1} \cdot (\lambda \cdot x) = \left(\lambda^{-1} \cdot \lambda\right) \cdot x = 1 \cdot x = x.$$

c. Für $x \in K$ gilt:

$$x + (-1) \cdot x = 1 \cdot x + (-1) \cdot x = \left(1 + (-1)\right) \cdot x = 0_K \cdot x = 0_V.$$

Also ist $(-1) \cdot x$ das (eindeutig bestimmte) additive Inverse zu x.

□

Bemerkung 1.4
Ist K ein Ring mit Eins, so folgt mit den gleichen Beweisen, daß die Aussagen von Lemma 1.3 a. und c. ebenfalls gelten.

Die Eigenschaft b. gilt für Moduln aber im Allgemeinen nicht, wie die folgenden Beispiele zeigen.

a. Wir betrachten $K = \mathbb{Z}_n$ mit $n = p \cdot q$, $p, q > 1$. Dann gilt für den $\mathbb{Z}_n$-Modul $\mathbb{Z}_n$ mit $\lambda = [p]$ und $x = [q]$

$$\lambda \cdot x = [p \cdot q] = [n] = [0], \quad \text{aber} \quad \lambda \neq [0], x \neq [0].$$

In diesem Fall hat der Ring $\mathbb{Z}_n$ sogenannte (nicht-triviale) *Nullteiler*, d. h. es gibt Elemente $\lambda, \mu \in \mathbb{Z}_n \setminus \{0\}$ mit $\lambda \cdot \mu = 0$.

b. Auch wenn der Ring K keine Nullteiler (außer 0) besitzt, gilt die Eigenschaft b. in obigem Lemma nicht unbedingt. Sei etwa $K = \mathbb{Z}$ und $V = \mathbb{Z}_n$. Mittels

$$\lambda \cdot [k] := [\lambda \cdot k], \quad \text{für } \lambda \in \mathbb{Z}, [k] \in \mathbb{Z}_n,$$

wird V zu einem K-Modul, und es gilt:

$$n \cdot [k] = [n \cdot k] = [0], \quad \text{für alle } [k] \in \mathbb{Z}_n,$$

obwohl $n \neq 0$.

Wie stets, betrachten wir auch bei Vektorräumen (bzw. Moduln) sogleich die strukturerhaltenden Abbildungen.

Definition 1.5
V und W seien K-Moduln. Eine Abbildung $f : V \to W$ heißt K-*lineare Abbildung* oder *Modulhomomorphismus*, wenn für alle $\lambda \in K$ und $x, y \in V$ gilt

a. $f(x+y) = f(x) + f(y)$ und (“*Additivität*”)
b. $f(\lambda x) = \lambda f(x)$.

Eine injektive (bzw. surjektive bzw. bijektive) K-lineare Abbildung heißt auch *Monomorphismus* (bzw. *Epimorphismus* bzw. *Isomorphismus*). Gilt $V = W$, so nennen wir eine K-lineare Abbildung auch einen *Endomorphismus*, und ist sie zudem bijektiv, so sprechen wir von einem *Automorphismus*.

Die Menge aller K-linearen Abbildungen von V nach W bezeichnen wir mit $\mathrm{Hom}_K(V,W)$ und die Menge aller Endomorphismen von V mit $\mathrm{End}_K(V)$.

Ist K ein Körper und sind V und W entsprechend Vektorräume, dann spricht man natürlich von *Vektorraumhomomorphismen* statt von Modulhomomorphismen.

Bemerkung 1.6
Ist K ein Ring *mit Eins*, so lassen sich die beiden Bedingungen a. und b. in obiger Definition zusammenfassen zu der Bedingung:

$$f(\lambda x + \mu y) = \lambda f(x) + \mu f(y), \quad \text{für alle } \lambda, \mu \in K, x, y \in V.$$

Lemma 1.7
Es seien U, V *und* W K-*Moduln, und* $f, f' \in \mathrm{Hom}_K(U,V)$ *und* $g, g' \in \mathrm{Hom}_K(V,W)$. *Ferner seien* $x, x_1, \ldots, x_n \in U$ *und* $\lambda, \lambda', \lambda_1, \ldots, \lambda_n \in K$. *Dann gelten:*

a. $f(0_U) = 0_V$ *und* $f(-x) = -f(x)$.
b. $f(\lambda_1 x_1 + \ldots + \lambda_n x_n) = \lambda_1 f(x_1) + \ldots + \lambda_n f(x_n)$.
c. *Ist* f *bijektiv, so gilt* $f^{-1} \in \mathrm{Hom}_K(V,U)$.
d. $g \circ f \in \mathrm{Hom}_K(U,W)$.
e. $f + f', \lambda \cdot f \in \mathrm{Hom}_K(U,V)$.
f. $g \circ (\lambda f + \lambda' f') = \lambda(g \circ f) + \lambda'(g \circ f')$ *und* $(\lambda g + \lambda' g') \circ f = \lambda(g \circ f) + \lambda'(g' \circ f)$.
g. $\lambda(g \circ f) = (\lambda g) \circ f = g \circ (\lambda f)$.

Beweis: Wegen Bedingung a. in der Definition von K-linearen Abbildungen gilt, daß f und g Gruppenhomomorphismen der den Moduln unterliegenden additiven Gruppen sind. Wir können also die Ergebnisse zu Gruppenhomomorphismen aus Proposition I.3.17 anwenden.

a. folgt dann aus den Teilen a. und b. von Proposition I.3.17.
b. folgt mittels Induktion aus den beiden Bedingungen für Linearität.

c. Wieder aus Proposition I.3.17, diesmal Teil c., folgt, daß f^{-1} der Bedingung a. der Definition von Linearität genügt. Es bleibt zu zeigen:

$$f^{-1}(\lambda y) = \lambda f^{-1}(y), \quad \text{für alle } \lambda \in K, y \in V.$$

Sei also $\lambda \in K$ und $y \in V$ gegeben. Setze $x := f^{-1}(y)$. Wegen der Linearität von f gilt

$$f^{-1}(\lambda y) = f^{-1}\big(\lambda f(x)\big) = f^{-1}\big(f(\lambda x)\big) = \lambda x = \lambda f^{-1}(y).$$

d. Die Additivität von $g \circ f$ folgt aus Lemma I.3.14. Weiter gilt wegen der Linearität von f und g für $\lambda \in K$ und $x \in U$:

$$(g \circ f)(\lambda x) = g\big(f(\lambda x)\big) = g(\lambda f(x)\big) = \lambda g\big(f(x)\big) = \lambda(g \circ f)(x).$$

e. Es gilt:

$$\begin{aligned} \big(f + f'\big)(x_1 + x_2) &= f(x_1 + x_2) + f'(x_1 + x_2) \\ &= \big(f(x_1) + f(x_2)\big) + \big(f'(x_1) + f'(x_2)\big) \\ &= \big(f(x_1) + f'(x_1)\big) + \big(f(x_2) + f'(x_2)\big) \\ &= \big(f + f'\big)(x_1) + \big(f + f'\big)(x_2) \end{aligned}$$

und

$$\begin{aligned} \big(f + f'\big)(\lambda x) &= f(\lambda x) + f'(\lambda x) = \lambda f(x) + \lambda f'(x) \\ &= \lambda\big(f(x) + f'(x)\big) = \lambda\big(f + f'\big)(x). \end{aligned}$$

Also ist $f + f'$ linear.
Außerdem gilt:

$$(\lambda \cdot f)(x_1 + x_2) = \lambda\big(f(x_1 + x_2)\big) = \lambda\big(f(x_1) + f(x_2)\big)$$

$$\lambda f(x_1) + \lambda f(x_2) = (\lambda \cdot f)(x_1) + (\lambda \cdot f)(x_2),$$

sowie

$$(\lambda \cdot f)(\lambda_1 x_1) = \lambda\big(f(\lambda_1 x_1)\big) = \lambda\big(\lambda_1 f(x_1)\big)$$

$$(\lambda\lambda_1)f(x_1) = (\lambda_1\lambda)f(x_1) = \lambda_1\big(\lambda f(x_1)\big) = \lambda_1(\lambda \cdot f)(x_1),$$

und damit ist auch $\lambda \cdot f$ linear.

f./g. Die Nachweise sind elementar und seien dem Leser überlassen.

□

Die wichtigsten Beispiele linearer Abbildungen sind die durch eine Matrix A definierten Abbildungen f_A.

Proposition 1.8
Für $A \in \mathrm{Mat}(m \times n, K)$ gilt:

$$f_A \in \mathrm{Hom}_K\left(K^n, K^m\right).$$

Beweis: Die Aussage folgt aus Teil a. Lemma I.8.7. □

Beispiel 1.9
Im Fall $n = 1$ und $A = (a)$ ist die K-lineare Abbildung $f_A : K \to K : x \mapsto a \cdot x$ gerade die Multiplikation mit a.

Behauptung: Ist K ein kommutativer Ring mit Eins, so ist jede lineare Abbildung $f : K \to K$ von der Form $f = f_A$ für $A = (a) \in \mathrm{Mat}(1, K)$ geeignet.

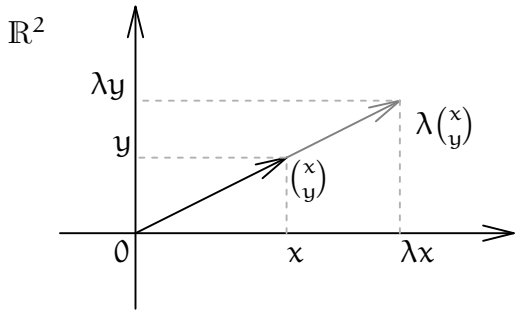
Dazu setzt man $a := f(1)$. Dann gilt für $x \in K$ beliebig:

$$f(x) = x \cdot f(1) = Ax = f_A(x),$$

also $f = f_A$.

Wir werden später sehen, daß die Behauptung allgemein für lineare Abbildungen $f : K^n \to K^m$ zutrifft, siehe Korollar 2.12.

Beispiel 1.10 a. Die lineare Abbildung $f_{\lambda \mathbb{1}_2} : \mathbb{R}^2 \mapsto \mathbb{R}^2 : (x, y)^t \mapsto (\lambda x, \lambda y)$ ist eine *Streckung* um den Faktor λ.



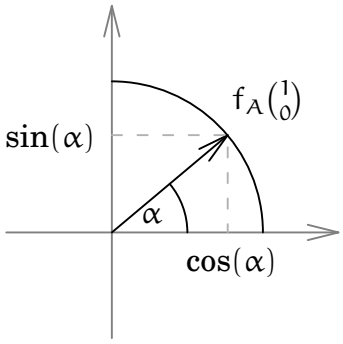
b. Für $\alpha \in \mathbb{R}$ setzen wir

$$A_\alpha := \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}.$$

Dann ist die lineare Abbildung $f_{A_\alpha} : \mathbb{R}^2 \to \mathbb{R}^2$ eine Drehung um den Winkel α. Beachte dazu, daß

$$Ae^1 = \big(\cos(\alpha), \sin(\alpha)\big)^t \quad \text{und} \quad Ae^2 = \big(-\sin(\alpha), \cos(\alpha)\big)^t,$$

woraus die Aussage für die *Einheitsvektoren* e^1 und e^2 unmittelbar folgt.



Daraus leitet sich die Aussage für einen beliebigen Vektor $(x, y)^t$ mittels der Linearität von f_A ab:

$$f_A\big((x, y)^t\big) = x f_A\big(e^1\big) + y f_A\big(e^2\big).$$

c. Ist $n \geq m$, so ist die Abbildung

$$\mathrm{pr} : K^n \to K^m : (x_1, \ldots, x_n)^t \mapsto (x_1, \ldots, x_m)^t$$

eine K-lineare Abbildung, genannt die kanonische *Projektion*.

Ist $m \geq n$, dann ist die kanonische *Inklusion*

$$i_{K^n} : K^n \to K^m : (x_1, \ldots, x_n)^t \mapsto (x_1, \ldots, x_n, 0, \ldots, 0)^t$$

ebenfalls K-linear. Beides prüft man leicht nach.

Wie bei Gruppen und Ringen definiert man auch bei Moduln und Vektorräumen Unterstrukturen.

Definition 1.11
Es sei V ein Modul über K. Eine Teilmenge $U \subseteq V$ von V heißt *Untermodul*, wenn U mit der Einschränkung der Addition und Skalarmultiplikation auf U selbst ein K-Modul ist.

Ist $V = K$ (vgl. Beispiel 1.2), so nennen wir einen Untermodul auch ein *Ideal*.

Ist K ein Körper, so nennt man einen Untermodul von V einen *Untervektorraum* oder kurz *Unterraum* von V.

Proposition 1.12
Es sei V ein K-Modul und $\emptyset \neq U \subseteq V$ eine nicht-leere Teilmenge von V.

a. *Dann sind gleichwertig:*
   (i) U *ist ein Untermodul von* V.
   (ii) *Für alle* $\lambda \in K$ *und* $x, y \in U$ *gilt:*

$$x - y \in U \quad \textit{und} \quad \lambda x \in U.$$

b. *Ist* K *ein Ring* mit Eins*, so sind gleichwertig:*
   (i) U *ist ein Untermodul von* V.
   (ii) *Für alle* $\lambda, \mu \in K$ *und* $x, y \in U$ *gilt:*

$$\lambda x + \mu y \in U.$$

Beweis: Die Richtung $(i) \Rightarrow (ii)$ ist in beiden Fällen klar aus der Definition. Bleibt also jeweils $(ii) \Rightarrow (i)$ zu zeigen. In Fall a. folgt aus Aufgabe I.3.27 mittels (ii), daß $(U, +)$ eine Untergruppe von $(V, +)$ ist, im Fall b. folgt dies aus Proposition I.3.10 mit $(\lambda, \mu) = (1, 1)$ sowie $(\lambda, \mu) = (-1, 0)$ in (ii). U ist ferner abgeschlossen bezüglich der Skalarmultiplikation (in b. wählt man in (ii) dazu $\mu = 0$), und somit ist U ein Untermodul, da die übrigen Modulaxiome sich von V auf U übertragen. □

Beispiel 1.13 a. Ist V ein K-Modul, so ist $\{0_V\}$ stets ein Untermodul von V. Ferner ist V selbst ein Untermodul. Man nennt diese beiden auch die *trivialen Untermoduln*

b. Ist V ein K-Modul und $x \in V$, dann ist

$$Kx := K \cdot x := \{\lambda x \mid \lambda \in K\}$$

ein Untermodul von V. Beachte dazu, daß für $x_i = \lambda_i x \in Kx$ $(i = 1, 2)$ und $\lambda \in K$ beliebig gilt:

$$x_1 + x_2 = (\lambda_1 + \lambda_2)x \in Kx \text{ und } \lambda x_1 = (\lambda\lambda_1)x \in Kx,$$

und daß $Kx \neq \emptyset$ wegen $0_V = 0_K \cdot x \in Kx$.

c. Sind V und W zwei K-Moduln, dann ist $\mathrm{Hom}_K(V,W)$ ein Untermodul des K-Moduls W^V (siehe Beispiel 1.2), wie aus Proposition 1.12 mit Hilfe von Teil e. von Lemma 1.7 folgt. Man beachte dabei, daß $\mathrm{Hom}_K(V,W)$ nicht leer ist, da die Nullabbildung linear ist.

d. In der Analysis zeigt man, daß Summe sowie skalare Vielfache von stetigen bzw. k-fach stetig differenzierbaren Funktionen wieder stetig bzw. k-fach stetig differenzierbar sind, $0 \neq k \in \mathbb{N} \cup \{\infty\}$. Da ferner die Funktion $\mathbb{R} \to \mathbb{R} : x \mapsto 0$ stetig und k-fach differenzierbar ist für jedes $0 \neq k \in \mathbb{N} \cup \{\infty\}$, gilt

$$\mathcal{C}^0(\mathbb{R}) := \{f : \mathbb{R} \to \mathbb{R} \mid f \text{ ist stetig}\}$$

und

$$\mathcal{C}^k(\mathbb{R}) := \{f : \mathbb{R} \to \mathbb{R} \mid f \text{ ist } k\text{-fach stetig differenzierbar}\}, 0 \neq k \in \mathbb{N} \cup \{\infty\},$$

folgt aus Proposition 1.12, daß $\mathcal{C}^0(\mathbb{R})$ und $\mathcal{C}^k(\mathbb{R})$ Unterräume des $\mathbb{R}$-Vektorraumes $\mathbb{R}^{\mathbb{R}}$ sind.
Solche Funktionenräume spielen in der Analysis eine große Rolle. Sie sind für kein n isomorph zu $\mathbb{R}^n$, und sie sind ein gutes Beispiel für den Wert der abstrakten Theorie der Vektorräume.

Wichtige Untermoduln und Unterräume erhält man durch lineare Abbildungen.

Lemma 1.14
Sind V und W K-Moduln, $f \in \mathrm{Hom}_K(V,W)$, dann sind Ker$(f) = f^{-1}(0_W)$*, der* Kern *von f, und* Im$(f) = f(V)$*, das* Bild *von f, Untermoduln von V bzw. W.*

Beweis: Wir wollen wieder Proposition 1.12 anwenden. Dazu beachten wir zunächst, daß wegen $f(0_V) = 0_W$ Ker(f) und Im(f) beide nicht leer sind.

Sind $\lambda \in K$ und $x, y \in \mathrm{Ker}(f)$, so gilt:

$$f(x+y) = f(x) + f(y) = 0_W + 0_W = 0_W \quad \text{und} \quad f(\lambda x) = \lambda f(x) = \lambda \cdot 0_W = 0_W.$$

Also ist $x + y, \lambda x \in \mathrm{Ker}(f)$.

Sind nun $x, y \in \mathrm{Im}(f)$, so gibt es Vektoren $u, v \in V$ mit $f(u) = x$ und $f(v) = y$. Dann gilt aber

$$x + y = f(u) + f(v) = f(u+v) \in \mathrm{Im}(f) \quad \text{und} \quad \lambda x = \lambda f(u) = f(\lambda u) \in \mathrm{Im}(f).$$

Es folgt aus Proposition 1.12, daß Ker(f) und Im(f) Untermoduln sind. □

Lemma 1.15
Ist V ein K-Modul und $(U_i)_{i\in I}$ eine Familie von Untermoduln in V, so ist der Durchschnitt $\bigcap_{i\in I} U_i$ ein Untermodul von V.

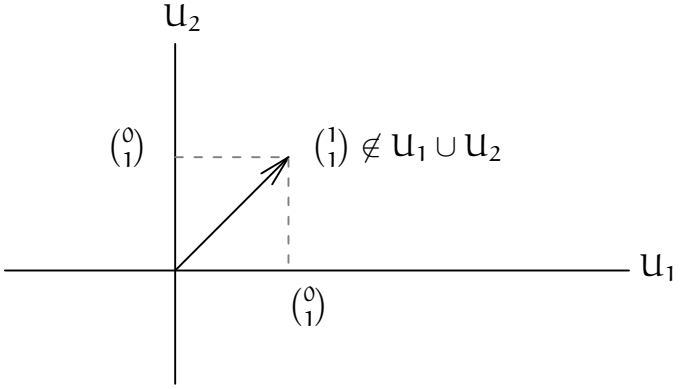
Beweis: Da $0_V \in U_i$ für alle $i \in I$, ist $U := \bigcap_{i\in I} U_i$ nicht die leere Menge. Es bleibt also zu zeigen, daß für $x, y \in U$ und $\lambda \in K$ gilt:

$$x + y \in U \quad \text{und} \quad \lambda x \in U.$$

Für ein beliebiges $i \in I$ gilt, da U_i ein Untermodul von V ist und da $x, y \in U \subseteq U_i$, daß $x + y \in U_i$ und $\lambda x \in U_i$. Also liegen die Vektoren im Durchschnitt U. □

Bemerkung 1.16
Die Vereinigung von zwei Untermoduln ist i. a. kein Untermodul mehr!

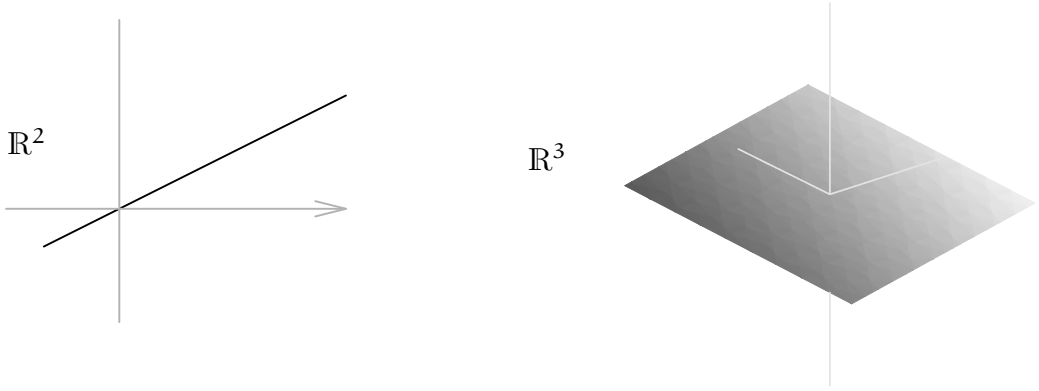
Sei etwa $U_1 := \mathbb{R}e^1 \subset \mathbb{R}^2$ die x-Achse und $U_2 := \mathbb{R}e^2 \subset \mathbb{R}^2$ die y-Achse. Dann liegt $(1,1)^t = e^1 + e^2$ nicht in $U_1 \cup U_2$, und mithin kann $U_1 \cup U_2$ kein Unterraum von $\mathbb{R}^2$ sein.



Es ist durchaus nützlich, sich Eigenschaften von Vektorräumen über einem beliebigen Körper am Beispiel des $\mathbb{R}^2$ bzw. des $\mathbb{R}^3$ klar zu machen. Dies ist zwar in keinem Fall beweiskräftig, aber kann durchaus wichtige Einsichten oder Ideen über zu beweisende oder vermutete Sachverhalte geben.

Die Geraden durch den Ursprung 0 in der Ebene $\mathbb{R}^2$ sind die nicht-trivialen Unterräume von $\mathbb{R}^2$. Entsprechend sind die Geraden und Ebenen durch den Ursprung im Raum $\mathbb{R}^3$ die nicht-trivialen Unterräume des $\mathbb{R}^3$.

Geraden und Ebenen, die nicht durch den Ursprung gehen, sind keine Unterräume.



Wir wollen uns von nun an auf die Theorie der Vektorräume konzentrieren.

AUFGABEN

Aufgabe 1.17

Welche der folgende Teilmengen von $\mathbb{R}^4$ sind Unterräume des $\mathbb{R}^4$? Begründe Deine Aussagen.

a. $\{(1,0,0,0)^t, (0,1,0,0)^t, (1,1,0,0)^t, (0,0,0,0)^t\}$,
b. $\{(x,0,0,0)^t, (0,y,0,0)^t \mid x,y \in \mathbb{R}\}$,
c. $\{(x,x,y,y)^t \mid x,y \in \mathbb{R}\}$,
d. $\{(x_1,x_2,x_3,x_4)^t \mid x_i \in \mathbb{R}, \sum_{i=1}^4 x_i = 1\}$,
e. $\{(x_1,x_2,x_3,x_4)^t \mid x_i \in \mathbb{R}, \sum_{i=1}^4 x_i = 0\}$.
f. Ist $\{(1,0,0,0)^t, (0,1,0,0)^t, (1,1,0,0)^t, (0,0,0,0)^t\}$ aufgefaßt als Teilmenge von $(\mathbb{Z}_2)^4$ ein Unterraum des $\mathbb{Z}_2$-Vektorraums $(\mathbb{Z}_2)^4$?

Aufgabe 1.18

Es sei $V = \mathbb{R}^{\mathbb{R}}$ der $\mathbb{R}$-Vektorraum der Abbildungen von $\mathbb{R}$ nach $\mathbb{R}$. Welche der folgenden Teilmengen von V sind Unterräume von V? Beweise Deine Aussagen.

a. $U_1 := \{f \in V \mid f(0) = 2\}$,
b. $U_2 := \{f \in V \mid f(1) = 0\}$,
c. $U_3 := \{f \in V \mid f \text{ ist beschränkt}\}$,
d. $U_4 := \{f \in V \mid |f(x)| \leq 4 \ \forall\, x \in \mathbb{R}\}$.

Aufgabe 1.19 (Gerade und Ungerade Funktionen)

Es sei $V = \mathbb{R}^{\mathbb{R}}$ der $\mathbb{R}$-Vektorraum aller Abbildungen von $\mathbb{R}$ nach $\mathbb{R}$, $U := \{f \in V \mid f(-x) = f(x) \ \forall\, x \in \mathbb{R}\}$ und $U' := \{f \in V \mid f(-x) = -f(x) \ \forall\, x \in \mathbb{R}\}$.[1] Zeige, U und U' sind Unterräume von V.

Aufgabe 1.20

Es sei $(K,+,\cdot)$ ein Körper, $(V,+,\cdot)$ ein K-Vektorraum und $U, U' \subseteq V$ Unterräume von V. Zeige, genau dann ist $U \cup U'$ ein Unterraum von V, wenn $U \subseteq U'$ oder $U' \subseteq U$.

Aufgabe 1.21

Gib alle Elemente und alle Unterräume des $\mathbb{Z}_2$-Vektorraums $(\mathbb{Z}_2)^2$ an.

Aufgabe 1.22

Es seien $f, g \in \mathrm{End}_K(V)$, V ein K-Vektorraum, $0 \neq \lambda \in K$. Zeige:

a. $\mathrm{Im}(f+g) \subseteq \mathrm{Im}(f) + \mathrm{Im}(g)$,
b. $\mathrm{Ker}(f+g) \supseteq \mathrm{Ker}(f) \cap \mathrm{Ker}(g)$,
c. $\mathrm{Im}(\lambda f) = \mathrm{Im}(f)$,
d. $\mathrm{Ker}(\lambda f) = \mathrm{Ker}(f)$,
e. $\mathrm{Im}(f \circ g) \subseteq \mathrm{Im}(f)$, und
f. $\mathrm{Ker}(f \circ g) \supseteq \mathrm{Ker}(g)$.

Aufgabe 1.23 (Projektionen)

Es sei V ein K-Vektorraum. $f \in \mathrm{End}_K(V)$ heißt *Projektion*, falls $f^2 = f$ gilt. Zeige, die folgenden Aussagen sind äquivalent:

[1] Die Funktionen in U nennt man *gerade Funktionen* , die in U' *ungerade*.

a. f ist eine Projektion,
b. $\mathrm{id}_V - f$ ist eine Projektion,
c. $\mathrm{Im}\left(\mathrm{id}_V - f\right) = \mathrm{Ker}(f)$,
d. $\mathrm{Ker}\left(\mathrm{id}_V - f\right) = \mathrm{Im}(f)$.

Aufgabe 1.24
Es sei V ein K-Vektorraum und $f \in \mathrm{End}_K(V)$. Zeige, für $m, n \in \mathbb{N}$ mit $m < n$ gilt:

$$\mathrm{Ker}(f^m) \subseteq \mathrm{Ker}(f^n) \text{ und } \mathrm{Im}(f^m) \supseteq \mathrm{Im}(f^n).$$

Finde Beispiele, so daß die Inklusionen stets strikt sind.

Aufgabe 1.25
Es seien U, V, W drei K-Vektorräume, $f \in \mathrm{Hom}_K(U, V)$ und $g \in \mathrm{Hom}_K(U, W)$. Zeige, genau dann gibt es ein $h \in \mathrm{Hom}_K(W, V)$ mit $h \circ g = f$, wenn $\mathrm{Ker}(g) \subseteq \mathrm{Ker}(f)$.

Aufgabe 1.26
Es sei $(R, +, \cdot)$ ein kommutativer Ring mit Eins.

a. Zeige, die Abbildung $\varphi : (R, +, \cdot) \to (\mathrm{Mat}(n \times n, R), +, \circ) : \lambda \mapsto \lambda \cdot \mathbb{1}_n$ ist ein Ring*mono*morphismus.
b. Betrachten wir nun R und $\mathrm{Mat}(n \times n, R)$ als R-Moduln. Ist dann φ auch ein R-Modul-Homomorphismus?

Aufgabe 1.27 (Funktionenraum)
Es sei $V = C^\infty(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} \mid f \text{ ist unendlich oft differenzierbar}\}$ der $\mathbb{R}$-Vektorraum der unendlich oft differenzierbaren Funktionen auf $\mathbb{R}$. Zeige, die Abbildung $D : V \to V : f \mapsto f'$, wobei f' die Ableitung von f bezeichne, ist eine $\mathbb{R}$-lineare Abbildung.
Es gilt $\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} \in V$. Folgt aus der Linearität von D, daß $D\big(\exp(x)\big) = \sum_{n=0}^{\infty} D\big(\frac{x^n}{n!}\big) = \exp(x)$?

Aufgabe 1.28
Es sei $(R, +, \cdot)$ ein kommutativer Ring mit $1 \neq 0$. Zeige, genau dann ist R ein Körper, wenn R nur die beiden trivialen Ideale $\{0\}$ und R besitzt.

2. Lineare Abhängigkeit und Basen

Sind $x_1, \ldots, x_n$ Vektoren eines K-Vektorraumes V, so ist auch jede *Linearkombination* $x = \lambda_1 x_1 + \ldots + \lambda_r x_r$ ein Vektor in V, wie wir in Lemma 1.7 gesehen haben. In diesem Abschnitt geht es nun um Familien von Vektoren $B = (x_i)_{i \in I}$, I eine eventuell unendliche Indexmenge, so daß jeder Vektor $x \in V$ als endliche Linearkombination $x = \lambda_1 x_{i_1} + \ldots + \lambda_r x_{i_r}$ für geeignete $\lambda_i \in K$ und $x_{i_j} \in B$ darstellbar ist. Ist diese Darstellung zudem für alle x eindeutig, so heißt B eine Basis. Wir werden sehen, daß jeder Vektorraum eine Basis besitzt und daß die Anzahl der Elemente einer Basis nur vom Vektorraum V, nicht aber von der Basis abhängt. Diese Anzahl nennt man die Dimension von V.

Der Nachweis der Existenz und der Unabhängigkeit der Anzahl der Elemente ist nicht ganz einfach. Da diese Aussagen auch nur für Vektorräume, nicht aber für allgemeine Moduln gelten, werden wir für diesem Abschnitt folgende Generalvoraussetzung treffen.[2]

Generalvoraussetzung: In diesem Abschnitt sei R stets ein kommutativer Ring mit Eins und K ein Körper. V bezeichne einen R-Modul respektive einen K-Vektorraum, je nach aktuellem Kontext.

Definition 2.1 a. Wir nennen $x \in V$ eine *Linearkombination* von $x_1, \ldots, x_r \in V$, falls es $\lambda_1, \ldots, \lambda_r \in R$ gibt mit

$$x = \lambda_1 x_1 + \ldots + \lambda_r x_r.$$

Ist eines der λ_i ungleich Null, so nennen wir die Linearkombination *nicht-trivial*

b. Ist $F = (x_i)_{i \in I}$ eine Familie von Vektoren aus V, so heißt $x = \sum_{i \in I} \lambda_i x_i$ eine *endliche Linearkombination* von Elementen aus F, wenn $\lambda_i \in R$ für alle $i \in I$ und wenn nur endlich viele der λ_i ungleich Null sind.
Wir verwenden die Notation $\sum'_{i \in I} \lambda_i x_i$, $\lambda_i \in R$, oder $\sum_{j=1}^{r} \lambda_j x_{i_j}$, $\lambda_j \in R$, um eine *endliche* Linearkombination von Elementen in F anzusprechen.

c. Vektoren $x_1, \ldots, x_r \in V$ heißen *linear abhängig*, falls es $\lambda_1, \ldots, \lambda_r \in R$ gibt mit $\lambda_i \neq 0$ für mindestens ein $i \in \{1, \ldots, r\}$, so daß

$$\lambda_1 x_1 + \ldots + \lambda_r x_r = 0.$$

Andernfalls heißen sie *linear unabhängig*.

d. Eine Familie $F = (x_i)_{i \in I}$ heißt *linear unabhängig*, wenn für jede endliche Teilfamilie $(x_{i_1}, \ldots, x_{i_r})$ von F gilt, daß die Vektoren $x_{i_1}, \ldots, x_{i_r}$ linear unabhängig sind. Andernfalls heißt sie linear abhängig.

Bemerkung 2.2 a. Es sei $F = (x_i)_{i \in I}$ eine Familie von Vektoren. Dann sind gleichwertig:

(i) F ist linear unabhängig.

(ii) Für jede endliche Teilfamilie $(x_{i_1}, \ldots, x_{i_r})$ von F gilt:

$$\lambda_1, \ldots, \lambda_r \in R \text{ mit } \sum_{j=1}^{r} \lambda_j x_{i_j} = 0 \;\Rightarrow\; \lambda_1 = \ldots = \lambda_r = 0.$$

Entsprechend sind ebenfalls gleichwertig:

(i) F ist linear abhängig.

[2] Der Grund dafür, daß etwa Lemma 2.6 sowie alle Aussagen zur Existenz von Basen über beliebigen Ringen falsch sind, ist der Umstand, daß es nicht ausreicht, daß ein Ringelement ungleich Null ist, um bereits dadurch dividieren zu dürfen. Dies schlägt sich z. B. darin nieder, daß für $\lambda \in R$ und $x \in V$ durchaus $\lambda x = 0$ gelten kann, ohne daß λ oder x Null sein müßten. Man sollte bei Aussagen über Moduln über beliebigen Ringen also streng darauf achten, daß man diese scheinbar *selbstverständliche* Tatsache nicht benutzt!

(ii) Es *gibt* eine endliche Teilfamilie $(x_{i_1}, \ldots, x_{i_r})$ von F und $\lambda_1, \ldots, \lambda_r \in R \setminus \{0\}$, so daß

$$\sum_{j=1}^{r} \lambda_j x_{i_j} = 0.$$

Dies ist nur eine formalere Schreibung der Definition.

b. Nach Definition ist die leere Familie linear unabhängig.

c. Aus der Definition von linearer Unabhängigkeit folgt unmittelbar, daß jede Teilfamilie einer linear unabhängigen Familie wieder linear unabhängig ist.

d. Aus dem letzten Punkt folgt wiederum, daß für eine endliche Familie $(x_1, \ldots, x_r)$ gleichwertig sind:
   (i) Die Familie $(x_1, \ldots, x_r)$ ist linear (un)abhängig.
   (ii) Die Vektoren $x_1, \ldots, x_r$ sind linear (un)abhängig.

Beispiel 2.3

Die *Einheitsvektoren* $e^1, \ldots, e^n \in R^n$ sind linear unabhängig, denn für $\lambda_1, \ldots, \lambda_n \in R$ mit $\sum_{i=1}^{n} \lambda_i e^i = 0$ gilt

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = \lambda_1 e^1 + \ldots + \lambda_n e^n = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix},$$

also $\lambda_1 = \ldots = \lambda_n = 0$.

Obwohl die Einheitsvektoren e^i Spaltenvektoren sind, ist es in der mathematischen Literatur gebräuchlich, den Index als Subskript zu schreiben. Wir werden uns dieser Notation anschließen und schreiben von nun an für die Einheitsvektoren

$$e_1, \ldots, e_n \quad \text{statt} \quad e^1, \ldots, e^n,$$

sofern wir nicht besonderes hervorheben wollen, daß es sich um Spaltenvektoren handelt.

Beispiel 2.4

Die Vektoren $x_1 = (1, 0, 1)^t$, $x_2 = (1, 1, 1)^t$ und $x_3 = (3, 1, 3)^t$ in R^3 sind hingegen linear abhängig, da offensichtlich $x_3 = 2x_1 + x_2$, d. h. wir haben eine nicht-triviale Linearkombination

$$2x_1 + x_2 - x_3 = 0,$$

die Null ergibt.

Lemma 2.5

Sei $F = (x_i)_{i \in I}$ *eine Familie von Vektoren des* R*-Moduls* V.

a. *Ist* $x_i = 0$ *für ein* $i \in I$*, so ist* F *linear abhängig.*

b. *Gilt* $x_i = x_j$ *für zwei* $i, j \in I$ *mit* $i \neq j$*, so ist* F *linear abhängig.*

Beweis: a. Ist $x_i = 0$, so ist $1 \cdot x_i = 0$ eine nicht-triviale Linearkombination, die Null ergibt, also ist F linear abhängig.

b. Ist $x_i = x_j$, so ist $x_i - x_j = 0$ eine nicht-triviale Linearkombination, die Null ergibt, und F ist linear abhängig.

□

Die folgende Aussage ist nur in K-Vektorräumen wahr, da zu ihrer Herleitung die Division durch Elemente ungleich Null notwendig ist.

Lemma 2.6
Für eine Familie $F = (x_i)_{i\in I}$ von Vektoren in einem K-Vektorraum V sind gleichwertig:

a. *F ist linear abhängig.*
b. *Es gibt ein x_i, $i \in I$, das eine endliche Linearkombination der Elemente in $\{x_j \mid j \in I, j \neq i\}$ ist.*

Beweis: Ist F linear abhängig, dann gibt es $i_1, \ldots, i_r \in I$ und $\lambda_1, \ldots, \lambda_r \in K^*$ mit

$$\sum_{j=1}^{r} \lambda_j x_{i_j} = 0.$$

Aber dann gilt:

$$x_{i_1} = \sum_{j=2}^{r} -\frac{\lambda_j}{\lambda_1} x_{i_j},$$

d. h. x_{i_1} ist eine endliche Linearkombination der übrigen x_j.

Gilt hingegen für ein $i \in I$

$$\exists\, \lambda_j \in K \;:\; x_i = \sum_{j\in I\setminus\{i\}}{}' \lambda_j x_j,$$

dann ist mit $\lambda_i := -1$

$$\sum_{j\in I}{}' \lambda_j x_j = 0$$

eine endliche nicht-triviale Linearkombination, die Null ergibt, also ist F linear abhängig. □

Definition 2.7 a. Sei $M \subseteq V$ eine Teilmenge von V und sei $\mathcal{M} = \{U \subseteq V \mid M \subseteq U, U \text{ Untermodul von } V\}$ die Menge aller Untermoduln, die M enthalten, dann setzen wir

$$\langle M\rangle := \langle M\rangle_R := \bigcap_{U\in\mathcal{M}} U$$

und nennen $\langle M\rangle$ das *Erzeugnis* von M (oder den von M *erzeugten Untermodul* oder die *lineare Hülle* von M).

b. Ist $F = (x_i)_{i\in I}$ eine Familie von Vektoren, dann setzen wir

$$\langle F\rangle := \langle F(I)\rangle.$$

c. Eine Familie F von Vektoren in V heißt *Erzeugendensystem* von V, falls $\langle F\rangle = V$.

d. Ein linear unabhängiges Erzeugendensystem von V heißt eine *Basis* von V.

Proposition 2.8
Sei $F = (x_i)_{i\in I}$ *eine Familie von Vektoren in* V.

a. $\langle F\rangle$ *ist ein Untermodul von* V.
b. $\langle F\rangle = \left\{ \sum'_{i\in I} \lambda_i x_i \mid \lambda_i \in R\right\}$, *d. h.* $\langle F\rangle$ *besteht aus den endlichen Linearkombinationen von Elementen in* F.[3]
c. *Genau dann ist* F *linear unabhängig, wenn sich jedes Element in* $\langle F\rangle$ *in* eindeutiger *Weise als Linearkombination der Elemente von* F *schreiben läßt.*
d. *Genau dann ist* F *eine Basis von* V*, wenn sich jedes Element von* V *in* eindeutiger *Weise als Linearkombination der Elemente von* F *schreiben läßt.*

Beweis: a. Die Behauptung folgt aus Lemma 1.15.

b. Wir setzen

$$W := \left\{ \sum_{i\in I}{}' \lambda_i x_i \;\middle|\; \lambda_i \in R \right\}.$$

Aus Proposition 1.12 folgt unmittelbar, daß W ein Untermodul von V ist, da eine Linearkombination zweier endlicher Linearkombinationen wieder eine endliche Linearkombination ist und da $0 \in W \neq \emptyset$. Außerdem gilt $F \subset W$, also gilt nach Definition des Erzeugnisses

$$\langle F\rangle \subseteq W.$$

Andererseits gilt für jeden Untermodul U von V, der F enthält, daß er auch jede endliche Linearkombination von Elementen aus F enthält, sprich $W \subset U$. Aber dann ist W in jedem Untermodul enthalten, der F enthält und mithin gilt

$$W \subseteq \langle F\rangle.$$

c. Sei zunächst F linear unabhängig. Da es für jedes $x \in \langle F\rangle$ nach a. $\lambda_i \in R$ mit $x = \sum'_{i\in I} \lambda_i x_i$ gibt, bleibt nur die Eindeutigkeit der Darstellung zu zeigen. Seien also auch $\mu_i \in R$ mit $\sum'_{i\in I} \mu_i x_i = x$. Dann gilt:

$$0_V = x - x = \sum_{i\in I}{}' \lambda_i x_i - \sum_{i\in I}{}' \mu_i x_i = \sum_{i\in I}{}' (\lambda_i - \mu_i) x_i.$$

Da F linear unabhängig ist, muß mithin gelten:

$$\lambda_i - \mu_i = 0 \quad \text{für alle } i \in I.$$

Also ist die Darstellung eindeutig.

[3] In diesem Zusammenhang definieren wir die *leere Summe* $\sum_{i\in\emptyset} \lambda_i x_i := 0$ als den Nullvektor, so daß die Aussage auch für die leere Familie F gilt.

Sei umgekehrt die Eindeutigkeit der Darstellung eines jeden Vektors in $\langle F\rangle$ gegeben, und seien ferner $\lambda_i \in R$ mit $\sum'_{i\in I} \lambda_i x_i = 0$. Wir müssen zeigen,

$$\lambda_i = 0 \quad \text{für alle } i \in I. \tag{9}$$

Da aber auch $\sum_{i\in I} 0 \cdot x_i = 0$ eine endliche Linearkombination der Null mit Elementen aus F ist, folgt (9) aus der Eindeutigkeit der Darstellung.

d. Folgt aus c..

□

Beispiel 2.9 a. Im R^n bilden die Einheitsvektoren eine Basis $E = (e_1, \ldots, e_n)$, die sogenannte *kanonische Basis* oder *Standardbasis*.
Beachte dazu, daß E nach Beispiel 2.3 linear unabhängig ist. Außerdem gilt für $x = (x_1, \ldots, x_n)^t \in R^n$ beliebig

$$x = x_1 e_1 + \ldots x_n e_n \in \langle E\rangle.$$

Also ist E auch ein Erzeugendensystem von V.

b. Analog sieht man, daß für $n, m \geq 1$ die Familie

$$\big(E_i^j \mid i = 1, \ldots, m; j = 1, \ldots, n\big),$$

wobei $E_i^j := (e_{lk})_{l=1,\ldots,m;k=1,\ldots,n}$ mit $e_{lk} = \delta_{il} \cdot \delta_{jk}$ (siehe auch Definition I.8.17), eine Basis des R-Moduls $\mathrm{Mat}(m \times n, R)$ ist.

Bemerkung 2.10
Da in einer linear unabhängigen Familie $F = (x_i)_{i\in I}$ von Vektoren in V die Elemente paarweise verschieden sind, ist F durch die Teilmenge $F(I) = \{x_i \mid i \in I\} \subseteq V$ bis auf die *Reihenfolge* festgelegt, und wir identifizieren gelegentlich F mit $F(I)$. Wenn es uns auch nicht auf die Indexmenge ankommt, sprechen wir einfach von einer *linear unabhängigen Teilmenge* von V. Diese Sprechweise verwenden wir insbesondere bei Basen gelegentlich.

Lemma 2.11
Seien V und W zwei R-Moduln, $f, g \in \mathrm{Hom}_R(V, W)$ und $F = (x_i)_{i\in I}$ eine Familie von Vektoren in V. Dann gelten:

a. $f\big(\langle F\rangle\big) = \langle f(x_i) \mid i \in I\rangle$.
b. *Gilt $V = \langle F\rangle$ und gilt $f(x_i) = g(x_i)$ für alle $i \in I$, dann gilt $g = f$, d. h. f ist eindeutig bestimmt durch $f_{|F(I)}$.*
c. *Ist F eine Basis von V und ist f injektiv, dann ist die Familie $f(F) := \big(f(x_i) \mid i \in I\big)$ eine Basis von $\mathrm{Im}(f)$.*
d. *Ist F eine Basis von V, dann gelten:*
   (i) *f ist genau dann injektiv, wenn $f(F)$ linear unabhängig ist.*
   (ii) *f ist genau dann surjektiv, wenn $f(F)$ ein Erzeugendensystem ist.*
   (iii) *f ist genau dann bijektiv, wenn $f(F)$ eine Basis ist.*

Beweis: Der Beweis ist dem Leser als Übungsaufgabe überlassen. □

Satz 2.12 (Existenz- und Eindeutigkeitssatz für lineare Abbildungen)
Seien V *und* W *Moduln über* R, $B = (x_i)_{i\in I}$ *eine Basis von* V *und* $F = (y_i)_{i\in I}$ *eine Familie von Vektoren in* W.

Dann existiert genau eine lineare Abbildung $f \in \mathrm{Hom}_R(V,W)$ *mit*

$$f(x_i) = y_i \quad \textit{für alle } i \in I.$$

Beweis: Die Eindeutigkeit der Abbildung f, so sie existiert, folgt aus Teil b. in Lemma 2.11.

Es bleibt also die Existenz zu zeigen. Dazu konstruieren wir f durch *lineare Fortsetzung*. Sei $x \in V$, dann läßt sich x in eindeutiger Weise schreiben als

$$x = \sum_{i\in I}{}' \lambda_i x_i$$

mit $\lambda_i \in R$. Folglich wird mittels der Festsetzung

$$f(x) := \sum_{i\in I}{}' \lambda_i y_i \in W$$

eine Abbildung $f : V \to W$ definiert.
Behauptung: f ist R-linear.

Seien dazu $\lambda, \mu \in R$ und $x = \sum'_{i\in I} \lambda_i x_i$, $x' = \sum'_{i\in I} \lambda'_i x_i \in V$. Dann folgt aus der Definition von f:

$$f\big(\lambda x + \mu x'\big) = f\left(\sum_{i\in I}{}' \big(\lambda\lambda_i + \mu\lambda'_i\big)x_i\right) = \sum_{i\in I}{}' \big(\lambda\lambda_i + \mu\lambda'_i\big)y_i$$

$$= \lambda \sum_{i\in I}{}' \lambda_i y_i + \mu \sum_{i\in I}{}' \lambda'_i y_i = \lambda f(x) + \mu f\big(x'\big).$$

□

Beispiel 2.13
Sei $E = (e_1, \ldots, e_n)$ die kanonische Basis von R^n, und seien $a^i = (a_{i1,}, \ldots, a_{im})^t \in R^m$, $i = 1, \ldots, n$, beliebige Vektoren in R^m. Satz 2.12 besagt dann, daß es genau eine lineare Abbildung $f : R^n \to R^m$ mit $e_i \mapsto a^i$ gibt.

Diese Abbildung kennen wir bereits. Ist nämlich $A \in \mathrm{Mat}(m \times n, R)$ die Matrix, deren Spaltenvektoren die a^i, $i = 1, \ldots, m$, sind, dann gilt wegen der Eindeutigkeitsaussage in 2.12 $f = f_A$, denn $f_A(e_i) = Ae_i = a^i$ (siehe Bemerkung I.8.6).
Korollar 2.14
Jede lineare Abbildung $f : R^n \to R^m$ *ist von der Form* $f = f_A$ *für eine eindeutig bestimmte Matrix* $A \in \mathrm{Mat}(m \times n, R)$.

Es gilt dabei, daß $f(e_i) = a^i$, $i = 1, \ldots, n$, *der* i*-te Spaltenvektor von* A *ist.*

Beweis: Ist $f \in \mathrm{Hom}_R\left(R^n, R^m\right)$ gegeben, so setzen wir $a^i := f(e_i)$, $i = 1, \ldots, n$, und die Aussage folgt aus Beispiel 2.13 unter Berücksichtigung der Eindeutigkeitsaussage aus Satz 2.12. □

Proposition 2.15
Für $A \in \mathrm{Mat}(n, R)$ *gilt,* f_A *ist genau dann bijektiv, wenn* A *invertierbar ist.*

In diesem Fall gilt $(f_A)^{-1} = f_{A^{-1}}$.

Beweis: Ist A invertierbar, dann folgt aus Satz I.8.8:

$$f_A \circ f_{A^{-1}} = f_{A \circ A^{-1}} = f_{\mathbb{1}_n} = \mathrm{id}_{R^n},$$

und analog $f_{A^{-1}} \circ f_A = \mathrm{id}_{R^n}$. Also ist f_A invertierbar und die eindeutig bestimmte Inverse ist $f_{A^{-1}}$.

Sei nun umgekehrt f_A bijektiv. Dann gilt wegen Lemma 1.7, daß $(f_A)^{-1} : R^n \to R^n$ linear ist. Also folgt mittels Korollar 2.14, daß eine Matrix $B \in \mathrm{Mat}(m \times n, R)$ existiert mit $(f_A)^{-1} = f_B$. Aber dann gilt:

$$f_{\mathbb{1}_n} = \mathrm{id}_{R^n} = f_A \circ (f_A)^{-1} = f_A \circ f_B = f_{A \circ B}.$$

Die Eindeutigkeitsaussage in Korollar 2.14 impliziert dann $A \circ B = \mathbb{1}_n$, und analog $B \circ A = \mathbb{1}_n$. Also ist A invertierbar und $B = A^{-1}$. □

Bemerkung 2.16
Die elementaren Zeilenoperationen lassen sich durch invertierbare Matrizen repräsentieren, vgl. Lemma I.8.18, sind mithin also Automorphismen von $\mathrm{Mat}(n \times m, R)$.

Von jetzt ab betrachten wir in diesem Abschnitt nur noch Vektorräume über einem Körper K. Unser Ziel ist es, zu zeigen, daß jeder Vektorraum eine Basis besitzt, was auf Moduln über beliebigen Nicht-Körpern i. a. nicht zutrifft.

Lemma 2.17
Für eine Familie $F = (x_i)_{i \in I}$ *von Vektoren in* V *sind die folgenden Aussagen gleichwertig:*

a. F *ist eine Basis von* V.
b. F *ist ein minimales Erzeugendensystem von* V*, d. h.* $V = \langle F \rangle$ *und für* $J \subsetneqq I$ *gilt* $V \neq \langle x_i \mid i \in J \rangle$.
c. F *ist eine maximale linear unabhängige Familie in* V*, d. h.* F *ist linear unabhängig und jede echte Oberfamilie* $(x_j \mid j \in J)$ *mit* $I \subsetneqq J$ *ist linear abhängig.*

Beweis: "a. ⇒ b.": Ist F eine Basis, so erzeugt F den Vektorraum V per definitionem. Ist $(x_j \mid j \in J)$ eine echte Teilfamilie von F und ist $i \in I \setminus J$, so gibt es wegen der linearen Unabhängigkeit von F keine Darstellung

$$x_i - \sum_{j \in J}{}' \lambda_j x_j = 0, \quad \lambda_j \in K,$$

also ist $x_i \notin \langle x_j \mid j \in J \rangle$

"b. ⇒ c.": Wir zeigen zunächst, daß F linear unabhängig ist. Angenommen, dies sei nicht der Fall, dann gibt es nach Lemma 2.6[4] ein $i \in I$ und $\lambda_j \in K$,

[4]Hier geht ein, daß K ein Körper ist!

$j \in I \setminus \{i\}$, mit

$$x_i = \sum_{j \in I \setminus \{i\}}{}' \lambda_j x_j.$$

Dann gilt aber

$$\langle x_j \mid j \in I \setminus \{i\} \rangle = \langle x_j \mid j \in I \rangle = V,$$

im Widerspruch zur Minimalität von F.

Sei nun $(x_j \mid j \in J)$ mit $I \subsetneqq J$ eine echte Oberfamilie von F und $j \in J \setminus I$, so gibt es $\lambda_i \in K$, $i \in I$, mit

$$x_j = \sum_{i \in I}{}' \lambda_i x_i,$$

da F ein Erzeugendensystem ist. Folglich ist

$$-x_j + \sum_{i \in I}{}' \lambda_i x_i = 0,$$

und damit ist $(x_j \mid j \in J)$ linear abhängig.

"c. $\Rightarrow$ a.": Da F linear unabhängig ist, bleibt zu zeigen, daß $\langle F \rangle = V$. Dies folgt aus dem folgenden Lemma 2.18 □

Lemma 2.18
Ist $F = (x_i)_{i \in I}$ *eine linear unabhängige Familie in* V *mit* $\langle F \rangle \subsetneqq V$*, so ist für jedes* $x \in V \setminus \langle F \rangle$ *die Familie* $(x, x_i \mid i \in I)$ *linear unabhängig.*

Beweis: Seien dazu $\lambda, \lambda_i \in K$, $i \in I$, mit

$$\lambda x + \sum_{i \in I}{}' \lambda_i x_i = 0.$$

Dann folgt[5] $\lambda = 0$, wegen $x \notin \langle F \rangle$. Da aber F linear unabhängig ist, folgt damit auch $\lambda_i = 0$ für $i \in I$. Also ist $(x, x_i \mid i \in I)$ linear unabhängig. □

Wir wollen jetzt zeigen, daß jeder Vektorraum eine Basis besitzt. Da wir dies nicht nur für Vektorräume mit einem endlichen Erzeugendensystem beweisen wollen, benötigen wir das folgende *Zornsche Lemma*, eine Aussage, die zu den logischen Grundlagen der Mathematik gehört. Grob gesprochen gehört es zu den (im Rahmen einer formalen Mengenlehre) nicht aus anderen Axiomen herleitbaren Axiomen. Man kann aber zeigen, daß das Zornsche Lemma äquivalent zum *Wohlordnungssatz* und zum *Auswahlaxiom* ist, vgl. [**Moo82, Sze50**]. Ohne diese Axiome läßt sich der Existenzsatz über Basen nicht für beliebige Vektorräume beweisen.

Für endlich erzeugbare Vektorräume, die für die Praxis natürlich ausreichen, werden wir im nächsten Abschnitt einen konstruktiven Beweis geben. Nicht nur in der Theorie der Vektorräume ist das Zornsche Lemma ein wichtiges Beweismittel für unendliche Objekte.

Definition 2.19
Es sei $(M, \leq)$ eine nicht-leere, teilgeordnete Menge.

[5] Hier geht ein, daß K ein Körper ist!

a. Eine Teilmenge K von M heißt eine *Kette*, wenn sie bezüglich $\leq$ total geordnet ist.
b. Ist K eine Kette in M, so heißt ein Element $x \in M$ eine *obere Schranke* von K, falls für alle $y \in K$ gilt $y \leq x$.
c. Ein Element $x \in K \subseteq M$ einer Teilmenge K von M heißt *maximal*, falls für alle Elemente $y \in K$ gilt:

$$x \leq y \;\Rightarrow\; x = y.$$

Bemerkung 2.20
Man beachte, daß ein maximales Element x in K nicht unbedingt größer als alle anderen Elemente in K ist. Es kann durchaus sein, daß es mit einigen Elementen nicht vergleichbar ist.

Axiom 2.21 (Lemma von Zorn)
Es sei $(M, \leq)$ *eine nicht-leere, teilgeordnete Menge. Gilt, daß jede Kette in* M *eine obere Schranke besitzt, so besitzt* M *ein maximales Element.*

Die Hauptarbeit für den Existenzbeweis für Basen wird in der folgenden Proposition geleistet.

Proposition 2.22
Sei F *eine linear unabhängige Familie in* V *und* E *ein Erzeugendensystem von* V*, das* F *enthält. Dann existiert eine Basis* B *von* V *mit* $F \subseteq B \subseteq E$.

Beweis: Wir betrachten das Mengensystem $M = \{G \subset V \mid G \text{ ist linear unabhängig}, F \subseteq G \subseteq E\}$. Da $F \in M$, ist M nicht leer. Außerdem ist M bezüglich der Inklusion "$\subseteq$" teilgeordnet (siehe Beispiel I.2.26).

Behauptung: Jede Kette $\mathcal{K}$ in M besitzt eine obere Schranke.

Sei also $\mathcal{K}$ eine Kette in M. Wir setzen $X := \bigcup_{G \in \mathcal{K}} G$. Dann gilt sicher $G \subseteq X$ für alle $G \in \mathcal{K}$. Sobald wir gezeigt haben, daß $X \in M$, ist also X eine obere Schranke von $\mathcal{K}$. Wegen $F \subseteq G \subseteq E$ für alle $G \in \mathcal{K}$, gilt auch

$$F \subseteq X \subseteq E.$$

Es bleibt also noch zu zeigen, daß X linear unabhängig ist. Seien dazu $x_1, \ldots, x_r \in X$ und $\lambda_1, \ldots, \lambda_r \in K$ gegeben mit

$$\sum_{i=1}^{r} \lambda_i x_i = 0. \tag{10}$$

Dann gibt es $G_1, \ldots, G_r \in \mathcal{K}$ mit $x_i \in G_i$. Da $\mathcal{K}$ eine Kette - also total geordnet - ist, gibt es ein $G \in \mathcal{K}$ mit $x_i \in G_i \subseteq G$. Da aber G linear unabhängig ist und (10) eine Gleichung in $\langle G \rangle$ ist, gilt:

$$\lambda_1 = \ldots = \lambda_r = 0.$$

Also ist X linear unabhängig, d. h. in M, und damit eine obere Schranke von $\mathcal{K}$.

Wir können also das Zornsche Lemma auf M anwenden und erhalten, daß M ein maximales Element B besitzt.
Behauptung: B ist eine Basis von V.

Da $B \in M$, ist also B linear unabhängig, und es bleibt zu zeigen, daß B ein Erzeugendensystem von V ist. Nehmen wir an, dies sei nicht der Fall. Dann gilt insbesondere $B \subsetneqq E$ und es gibt ein Element $x \in E \setminus \langle B \rangle$. Nach Lemma 2.18 gilt dann aber, daß $B \cup \{x\}$ linear unabhängig ist und somit, wegen $F \subset B \cup \{x\} \subseteq E$,

$$B \cup \{x\} \in M,$$

im Widerspruch zur Maximalität von B. Also war die Annahme falsch und B ist eine Basis. □

Korollar 2.23 (Basisergänzungssatz) a. *Jede linear unabhängige Teilmenge F' von V kann zu einer Basis ergänzt werden.*
b. *Jedes Erzeugendensystem E' von V enthält eine Basis.*

Beweis: Für a. wähle in Proposition 2.22 $F = F'$ und $E = V$, und für b. wähle dort $F = \emptyset$ und $E = E'$. □

Damit haben wir insbesondere den folgenden Satz bewiesen.
Satz 2.24
Jeder Vektorraum besitzt eine Basis.

Wir notieren noch den folgenden Austauschsatz, daß man in einer Basis geeignete Elemente durch andere linear unabhängige austauschen kann.
Satz 2.25 (Schwacher Austauschsatz)
Ist B eine Basis von V und F' eine linear unabhängige Familie, so existiert eine Teilfamilie B' von B, so daß $B' \cup F'$ eine Basis von V ist.

Beweis: Wähle in Proposition 2.22 $F = F'$ und $E = B \cup F'$. □

Bemerkung 2.26
Auch wenn jeder Vektorraum eine Basis besitzt, kann nicht notwendigerweise für jeden Vektorraum eine Basis angegeben werden. $\mathbb{R}$ als $\mathbb{Q}$-Vektorraum ist ein gutes Beispiel für einen Vektorraum, bei dem man keine Basis angeben kann.
Behauptung: Eine Basis von $\mathbb{R}$ als $\mathbb{Q}$-Vektorraum ist überabzählbar.[6]

Hierzu argumentieren wir wie folgt:

a. $\mathbb{R}$ *ist überabzählbar*. Dies zeigt man mit Hilfe des *Cantorschen Diagonalverfahrens*. $\mathbb{R}$ ist sicherlich nicht endlich. Wäre $\mathbb{R}$ abzählbar, so gäbe es

[6]Eine Menge M heißt *abzählbar*, falls es eine bijektive Abbildung $\mathbb{N} \to M$ gibt. Sie heißt *höchstens abzählbar*, wenn sie endlich oder abzählbar ist, und sie heißt *überabzählbar*, falls sie nicht höchstens abzählbar ist.

eine bijektive Abbildung von $\varphi : \mathbb{N} \to \mathbb{R}$, und wir schreiben dann $\varphi(i)$, $i \in \mathbb{N}$, in Dezimaldarstellung:

$$\begin{array}{rcccccccc}
\varphi(0) & = & a_{0,-p_0} & a_{0,-p_0+1} & \dots & \underline{a_{0,0}}, & a_{01} & a_{02} & a_{03} & \cdots \\
\varphi(1) & = & a_{1,-p_1} & a_{1,-p_1+1} & \dots & a_{1,0}, & \underline{a_{11}} & a_{12} & a_{13} & \cdots \\
\varphi(2) & = & a_{2,-p_2} & a_{2,-p_2+1} & \dots & a_{2,0}, & a_{21} & \underline{a_{22}} & a_{23} & \cdots \\
\vdots & & & & & & & & \ddots &
\end{array}$$

Dann setzen wir $a := a_{00}, a_{11}a_{22}a_{33}\cdots \in \mathbb{R}$, d. h. a ist diejenige Zahl, die in obiger Aufzählung durch die unterstrichenen Diagonalelemente gegeben ist. Nun ändern wir jede der Ziffern von a ab (etwa $b_{ii} = 1$, falls $a_{ii} = 0$ und $b_{ii} = 0$ sonst) und erhalten eine Zahl

$$b = b_{00}, b_{11}b_{22}b_{33}\cdots \in \mathbb{R},$$

mit $a_{ii} \neq b_{ii}$ für alle $i \in \mathbb{N}$. Da φ bijektiv ist, gibt es ein $i \in \mathbb{N}$ mit $\varphi(i) = b$, also $a_{ii} = b_{ii}$, im Widerspruch zur Konstruktion von b. (Wir müssen noch berücksichtigen, daß $0,9999\cdots = 1$, was aber die einzige Zweideutigkeit der Dezimaldarstellung ist, und dieser weichen wir durch unsere Wahl der b_{ii} aus.) Also ist $\mathbb{R}$ überabzählbar.

b. *Die abzählbare Vereinigung höchstens abzählbarer Mengen ist wieder höchstens abzählbar.* Seien dazu $M_i = \{a_{i0}, a_{i1}, a_{i2}, \ldots\}$, $i \in \mathbb{N}$, (ohne Einschränkung) abzählbare Mengen, dann schreiben wir sie wie folgt auf:

$$\begin{array}{lccccccc}
M_0 : & a_{00} & a_{01} & a_{02} & a_{03} & a_{04} & \cdots \\
 & \downarrow\nearrow & \swarrow\nearrow & \swarrow\nearrow & \swarrow & & \\
M_1 : & a_{10} & a_{11} & a_{12} & a_{13} & a_{14} & \cdots \\
 & \swarrow\nearrow & \swarrow\nearrow & \swarrow & & & \\
M_2 : & a_{20} & a_{21} & a_{22} & a_{23} & a_{24} & \cdots \\
 & \downarrow\nearrow & \swarrow & & & & \\
M_3 : & a_{30} & a_{31} & a_{32} & a_{33} & a_{34} & \cdots \\
 & & & & & & \\
 \vdots & \vdots & \vdots & \vdots & \vdots & \vdots &
\end{array}$$

Abzählen der Elemente wie angedeutet, wobei man Elemente, die mehrfach vorkommen, nur beim ersten Mal berücksichtigt, liefert eine Bijektion von $\mathbb{N} \to \bigcup_{i\in\mathbb{N}} M_i$, mithin ist die Vereinigung abzählbar.

c. Es gilt also $\mathbb{Z} = \mathbb{N} \cup (-\mathbb{N})$ und

$$\mathbb{Q} = \bigcup_{0\neq q\in\mathbb{N}} \tfrac{1}{q}\mathbb{Z} = \bigcup_{0\neq q\in\mathbb{N}} \left\{ \tfrac{p}{q} \mid p \in \mathbb{Z} \right\}$$

sind abzählbar.

d. *Das kartesische Produkt zweier höchstens abzählbarer Mengen ist wieder höchstens abzählbar.* Seien dazu M und N zwei höchstens abzählbare Mengen, dann gilt

$$M \times N = \bigcup_{m\in M} \{m\} \times N,$$

wobei $N \to \{m\} \times N : n \mapsto (m, n)$ eine Bijektion ist, $\{m\} \times N$ also höchstens abzählbar ist.

e. *Ein Vektorraum V über einem höchstens abzählbaren Körper K mit höchstens abzählbarer Basis ist höchstens abzählbar.* Sei dazu (ohne Einschränkung) $B = (x_i \mid i \in \mathbb{N})$, eine abzählbare Basis von V. Für $n \in \mathbb{N}$ setzen wir

$$V_n := \langle x_1, \ldots, x_n \rangle.$$

Dann gilt $V_n \cong K^n$, also ist V_n nach d. mit Induktion über n abzählbar. Aber dann ist $V = \bigcup_{n\in\mathbb{N}} V_n$ nach b. abzählbar.

f. Da $\mathbb{Q}$ abzählbar ist, $\mathbb{R}$ aber überabzählbar, folgt die Behauptung aus e..

AUFGABEN

Aufgabe 2.27

Es sei V ein K-Vektorraum, $U \subset V$ ein Unterraum, $0 \neq u \in U$ und $v \in V \setminus U$. Zeige, v und u sind linear unabhängig.

Aufgabe 2.28

Es sei $(V, +, \cdot)$ ein $\mathbb{C}$-Vektorraum. Schränken wir die Skalarmultiplikation auf den Teilkörper $\mathbb{R}$ von $\mathbb{C}$ ein, so wird V offensichtlich auch zu einem $\mathbb{R}$-Vektorraum. Zeige, für Vektoren $v_1, \ldots, v_r \in V$ sind die folgenden Aussagen gleichwertig:

a. $(v_1, \ldots, v_r)$ ist linear unabhängig über $\mathbb{C}$,
b. $(v_1, iv_1, \ldots, v_r, iv_r)$ ist linear unabhängig über $\mathbb{R}$.

Aufgabe 2.29

Es sei $(V, +, \cdot)$ ein K-Vektorraum, und $x_1, \ldots, x_n \in V$ seien linear abhängige Vektoren mit der Eigenschaft, daß je $n - 1$ der Vektoren linear unabhängig sind. Zeige:

a. Es gibt $\lambda_1, \ldots, \lambda_n \in K \setminus \{0\}$ mit der Eigenschaft

$$\sum_{i=1}^{n} \lambda_i x_i = 0.$$

b. Gilt für $\mu_1, \ldots, \mu_n \in K$ ebenfalls $\sum_{i=1}^{n} \mu_i x_i = 0$, so gibt es ein $\nu \in K$ mit $\mu_i = \lambda_i \cdot \nu$ für alle $i = 1, \ldots, n$.

Aufgabe 2.30

Es sei $V = \mathbb{R}^{\mathbb{R}}$. Für $k \in \mathbb{N}$ definiere $f_k : \mathbb{R} \to \mathbb{R} : x \mapsto \sin(kx)$ und $g_k : \mathbb{R} \to \mathbb{R} : x \mapsto \cos(kx)$. Zeige, $\big(f_k, g_k \mid 0 \neq k \in \mathbb{N}\big)$ ist eine (unendliche) linear unabhängige Familie im $\mathbb{R}$-Vektorraum $\mathbb{R}^{\mathbb{R}}$.

Aufgabe 2.31 (Folgenraum)

Es sei $V = K^\infty := K^{\mathbb{N}} = \{a : \mathbb{N} \to K \mid a \text{ ist Abbildung}\} = \{(a_0, a_1, a_2, \ldots) \mid a_i \in K \ \forall\ i \in \mathbb{N}\}$ der K-Vektorraum aller *Folgen* über K mit komponentenweiser Addition und Skalarmultiplikation. Ferner sei $t^i \in V$ die Folge, die an i-ter Stelle eine Eins und sonst nur Nullen enthält, das heißt als Abbildung ist $t^i : \mathbb{N} \to K$ definiert durch $t^i(n) = \delta_{in}$ für $n \in \mathbb{N}$. Ist dann $\big(t^i \mid i \in \mathbb{N}\big)$ eine Basis von V?

Aufgabe 2.32
Es sei V ein K-Vektorraum, und $f, g \in \mathrm{End}_K(V)$ mit $\mathrm{Ker}(f) = 0$ sowie $0 \subsetneqq \mathrm{Ker}(g) \subsetneqq V$. Zeige, (f, g) ist eine linear unabhängige Familie im K-Vektorraum $\mathrm{End}_K(V)$.

Aufgabe 2.33
Es sei V ein K-Vektorraum und $f \in \mathrm{End}_K(V)$, so daß für alle $v \in V$ gilt $f(v) \in \langle v \rangle$. Zeige, dann gibt es ein $\lambda \in K$, so daß für alle $v \in V$ gilt $f(v) = \lambda v$.

Aufgabe 2.34
Ist R ein Ring, so heißt ein Ideal $\mathfrak{m} \subset R$ *maximal*[7] falls gilt:

(i) $\mathfrak{m} \neq R$, und
(ii) für alle Ideale $I \subset R$ mit $\mathfrak{m} \subseteq I \neq R$ gilt, $\mathfrak{m} = I$.

Zeige, ein Ring mit $1 \neq 0$ besitzt (mindestens) ein maximales Ideal.
Hinweis: Wende das Lemma von Zorn auf eine geeignete Menge $M \neq \emptyset$ von Idealen an!

3. Endlich-dimensionale Vektorräume

Wir betrachten jetzt endlich *erzeugte Vektorräume* V, d. h. Vektorräume, die ein endliches Erzeugendensystem besitzen. Nach Korollar 2.23 besitzt V dann auch eine endliche Basis. Für solche Vektorräume kann man die Sätze des vorigen Abschnitts teilweise verschärfen und vor allem kann man in diesen Vektorräumen mit Hilfe von Basen und Matrizen effizient rechnen.

Generalvoraussetzung: Auch in diesem Kapitel sei K stets ein Körper und, sofern nichts anderes vermerkt ist, V ein endlich-erzeugter K-Vektorraum.

Lemma 3.1 (Austauschlemma)
Sei $(x_1, \cdots, x_n)$ *eine Basis von* V *und* $y = \sum_{i=1}^{n} \lambda_i x_i \in V$ *mit* $\lambda_j \neq 0$ *für ein* $j \in \{1, \ldots, n\}$. *Dann ist* $(y, x_i \mid i = 1, \ldots, n, i \neq j)$ *eine Basis von* V.

Man kann also x_j *gegen* y *austauschen.*

Beweis: Wegen $\lambda_j \neq 0$ gilt

$$x_j = \frac{1}{\lambda_j} \cdot y - \sum_{i \neq j} \frac{\lambda_i}{\lambda_j} \cdot x_i,$$

und somit

$$V = \langle x_1, \ldots, x_n \rangle = \langle y, x_1, \ldots, x_n \rangle = \langle y, x_i \mid i \neq j \rangle.$$

Bleibt also zu zeigen, daß $(y, x_i \mid i \neq j)$ linear unabhängig ist. Seien dazu $\mu_i \in K$, $i = 1, \ldots, n$, gegeben mit

$$0 = \mu_j y + \sum_{i \neq j} \mu_i x_i = \sum_{i=1}^{n} \mu_j \lambda_i x_i + \sum_{i \neq j} \mu_i x_i$$

$$= \mu_j \lambda_j x_j + \sum_{i \neq j} (\mu_j \lambda_i + \mu_i) x_i.$$

[7] Das heißt, $\mathfrak{m}$ ist ein echtes Ideal, das in keinem anderen echten Ideal enthalten ist.

Dann folgt aus der linearen Unabhängigkeit von $x_1, \ldots, x_n$

$$\mu_j \lambda_j = 0 \quad \text{und} \quad \mu_i = -\mu_j \lambda_i, \text{ für } i \neq j.$$

Wegen $\lambda_j \neq 0$, ist also $\mu_j = 0$ und damit auch

$$\mu_i = 0 \quad \text{für } i \neq j.$$

Damit ist die lineare Unabhängigkeit von $(y, x_i \mid i \neq j)$ gezeigt. □

Beispiel 3.2
Ist zum Beispiel $E = (e_1, \ldots, e_n)$ die kanonische Basis des K^n und $x = (\lambda_1, \ldots, \lambda_n)^t \in K^n$ mit $\lambda_j \neq 0$, so können wir e_j gegen x austauschen und erhalten wieder eine Basis.

Konkret kann man in der Basis $E = (e_1, e_2, e_3)$ von $\mathbb{R}^3$ den Vektor $(1, 2, 0)^t$ gegen e_1 oder e_2 austauschen, nicht aber gegen e_3.

Das Austauschlemma wird benutzt, um den wichtigen Steinitzschen Austauschsatz zu beweisen.

Satz 3.3 (Austauschsatz von Steinitz)
Es sei $(x_1, \ldots, x_n)$ *eine Basis von* V *und* $(y_1, \ldots, y_r)$ *eine linear unabhängige Familie in* V.

Dann lassen sich die $x_1, \ldots, x_n$ *so umnumerieren, daß* $(y_1, \ldots, y_r, x_{r+1}, \ldots, x_n)$ *eine Basis von* V *ist.*

Insbesondere gilt: $r \leq n$.

Bemerkung 3.4
Der Austauschsatz von Steinitz besagt also, daß man - nach eventuellem Umnumerieren - die linear unabhängigen Vektoren $x_1, \ldots, x_r$ durch $y_1, \ldots, y_r$ ersetzen kann.

Nach dem schwachen Austauschsatz 2.25 wissen wir, daß die Familie $(y_1, \ldots, y_r)$ durch Elemente aus $\{x_1, \ldots, x_n\}$ zu einer Basis von V ergänzt werden kann. Wir müssen jedoch noch zeigen, daß die neue Basis genau n Elemente besitzt.

Beweis von Satz 3.3: Wir führen den Beweis mittels Induktion über r.

Für $r = 0$ ist die Behauptung offensichtlich richtig. Nehmen wir also an, daß $r > 0$ und daß die Behauptung bereits richtig ist für $r - 1$. D. h. nach evt. Umnumerieren ist $(y_1, \ldots, y_{r-1}, x_r, \ldots, x_n)$ eine Basis von V. Dann besitzt y_r eine Darstellung der Form

$$y_r = \lambda_1 y_1 + \ldots + \lambda_{r-1} y_{r-1} + \lambda_r x_r + \ldots + \lambda_n x_n,$$

mit $\lambda_i \in K$. Angenommen, $\lambda_r = \ldots = \lambda_n = 0$, dann wäre $(y_1, \ldots, y_r)$ linear abhängig, im Widerspruch zur Voraussetzung. Also gibt es ein $j \in \{r, \ldots, n\}$ mit $\lambda_j \neq 0$. Durch Umnumerieren können wir annehmen, daß $j = r$ gilt. Dann können wir aber nach dem Austauschlemma 3.1 y_r gegen x_r austauschen, und die Behauptung ist bewiesen. □

Bemerkung 3.5
Im Austauschsatz tauschen wir nacheinander x_{i_1} durch y_1, x_{i_2} durch y_2, etc. und schließlich x_{i_r} durch y_r für geeignete $i_1, \dots, i_r$ aus. Im j-ten Schritt wissen wir, daß wir eine Darstellung

$$y_j = \sum_{l=1}^{j-1} \lambda_l y_l + \sum_{l \notin \{i_1,\dots,i_{j-1}\}} \lambda_l x_l$$

haben mit $\lambda_l \neq 0$ für ein $i_j \notin \{i_1, \dots, i_{j-1}\}$, und dann können wir x_{i_j} durch y_j ersetzen.

Wie wir eine solche Darstellung von y_j mit Hilfe des Gauß'schen Algorithmus berechnen können, werden wir später sehen, siehe 6.1. Damit haben wir dann ein konstruktives Verfahren für die Anwendung des Steinitzschen Austauschsatzes.

Als Folgerung des Steinitzschen Austauschsatzes erhalten wir den folgenden Satz.

Satz 3.6 a. *Ist V endlich erzeugbar, so ist jede Basis von V endlich und alle Basen haben gleich viele Elemente.*
b. *Ist V nicht endlich erzeugbar, so hat jede Basis unendlich viele Elemente.*

Beweis: a. Daß V eine endliche Basis $B = (x_1, \dots, x_n)$ besitzt, folgt aus Proposition 2.22. Dabei können wir o. E. annehmen, daß n die minimale Mächtigkeit einer Basis ist. Sei nun B' eine weitere Basis von V. Angenommen, $|B'| > n$. Dann gibt es eine linear unabhängige Teilfamilie $(y_1, \dots, y_{n+1})$ in B', im Widerspruch zum Austauschsatz von Steinitz, der verlangt $n + 1 \leq n$.
b. Dies ist offensichtlich, da jede Basis V erzeugt.

□

Satz 3.6 rechtfertigt die folgende Definition.

Definition 3.7
Für einen (nicht notwendig endlich erzeugten) K-Vektorraum V definieren wir die *Dimension* von V durch

$$\dim_K(V) := \begin{cases} n, & \text{falls } V \text{ eine Basis mit } n < \infty \text{ Elementen besitzt,} \\ \infty, & \text{falls } V \text{ nicht endlich erzeugt ist.} \end{cases}$$

Ist $\dim_K(V) < \infty$, so nennen wir V einen *endlich-dimensionalen* K-Vektorraum.

Aus Satz 3.6 und Definition 3.7 folgt unmittelbar das folgende Korollar.

Korollar 3.8
Es sei $\dim_K(V) = n$, E ein Erzeugendensystem von V und B eine linear unabhängige Familie. Dann gilt:

a. $|E| \geq n$.
b. $|B| \leq n$.

c. *In beiden Fällen gilt Gleichheit genau dann, wenn die Familie eine Basis ist.*

Beispiel 3.9 a. Es gilt:

$$\dim_K(V) = 0 \Leftrightarrow V = \langle\emptyset\rangle \Leftrightarrow V = \{0\}.$$

b. $\dim_K\left(K^n\right) = n$, da die kanonische Basis $E = (e_1, \ldots, e_n)$ genau n Elemente enthält.

c. $\dim_{\mathbb{Q}}(\mathbb{Q}) = \dim_{\mathbb{R}}(\mathbb{R}) = \dim_{\mathbb{C}}(\mathbb{C}) = 1$, aber $\dim_{\mathbb{Q}}(\mathbb{R}) = \infty$ und $\dim_{\mathbb{R}}(\mathbb{C}) = 2$. Für letzteres zeige man, daß die Familie $(1, i)$ eine $\mathbb{R}$-Basis von $\mathbb{C}$ ist.

Als Folgerung können wir alle endlich-dimensionalen K-Vektorräume bis auf Isomorphie klassifizieren.

Satz 3.10
Es sei $\dim_K(V) = n$ *und* $E = (e_1, \ldots, e_n)$ *die kanonische Basis des* K^n. *Dann gelten:*

a. *Ist* $B = (x_1, \ldots, x_n)$ *eine (*numerierte[8]*) Basis von* V*, so bestimmt* B *einen Isomorphismus*

$$\phi_B : V \to K^n : x_i \mapsto e_i, \text{ für } i = 1, \ldots, n$$

durch lineare Fortsetzung. *Man nennt* ϕ_B *die* Karte *von* V *zur Basis* B.

b. *Ist umgekehrt* $\psi \in \mathrm{Hom}_K\left(V, K^n\right)$ *ein Isomorphismus, dann ist* $\psi^{-1}(E) := \left(\psi^{-1}(e_1), \ldots, \psi^{-1}(e_n)\right)$ *eine numerierte Basis von* V.

c. *Die Abbildung*

$$\begin{array}{ccccc} \alpha : \{B \mid B \textit{ num. Basis}\} & \to & \left\{\psi \in \mathrm{Hom}_K\left(V, K^n\right) \mid \psi \textit{ Isom.}\right\} \\ B & \mapsto & \phi_B \end{array}$$

ist bijektiv mit inverser Abbildung

$$\begin{array}{ccccc} \beta : \left\{\psi \in \mathrm{Hom}_K\left(V, K^n\right) \mid \psi \textit{ Isom.}\right\} & \to & \{B \mid B \textit{ num. Basis}\} \\ \psi & \mapsto & \psi^{-1}(E). \end{array}$$

Beweis: a. Nach Satz 2.12 bestimmen die Zuordnungen

$$x_i \mapsto e_i,\ i = 1, \ldots, n, \quad \text{und} \quad e_i \mapsto x_i,\ i = 1, \ldots, n,$$

zwei lineare Abbildungen $\phi_B : V \to K^n$ und $\phi^B : K^n \to V$. Es bleibt zu zeigen, daß

$$\phi_B \circ \phi^B = \mathrm{id}_{K^n} \quad \text{und} \quad \phi^B \circ \phi_B = \mathrm{id}_V.$$

Dazu reicht es wegen Lemma 2.11 b. nachzuweisen, daß die Abbildungen jeweils auf einer Basis übereinstimmen, was sie aber per definitionem tun.

[8] Dadurch, daß Basen Familien von Vektoren sind, hier also B eine Abbildung von $\{1, \ldots, n\}$ nach V ist, ist B bereits per definitionem mit einer Ordnung (=Numerierung) versehen. Wenn wir dennoch hin und wieder den Begriff *numerierte* Basis verwenden, dann nur, um den Umstand hervorzuheben, daß in der vorliegenden Situation die *Reihenfolge* der Basiselemente wesentlich ist.

b. Da ψ^{-1} ein Isomorphismus ist, folgt mit Lemma 2.11 d., daß $\psi^{-1}(E)$ eine Basis ist.

c. Bezeichnen wir die Menge der Basen mit $\mathcal{B}$ und die der Isomorphismen mit $\mathcal{I}$, so müssen wir zeigen, daß

$$\alpha \circ \beta = \mathrm{id}_{\mathcal{I}} \quad \text{und} \quad \beta \circ \alpha = \mathrm{id}_{\mathcal{B}}\,.$$

Sei nun $B = (x_1, \ldots, x_n) \in \mathcal{B}$, dann gilt

$$(\beta \circ \alpha)(B) = \phi_B^{-1}(E) = \big(\phi_B^{-1}(e_1), \ldots, \phi_B^{-1}(e_n)\big) = (x_1, \ldots, x_n) = B,$$

also $\beta \circ \alpha = \mathrm{id}_{\mathcal{B}}$. Ist umgekehrt $\psi \in \mathcal{I}$, so gilt mit $B := \psi^{-1}(E)$, daß die Isomorphismen ϕ_B und ψ auf B übereinstimmen, und somit wegen Lemma 2.11 gleich sind. Damit folgt also:

$$(\alpha \circ \beta)(\psi) = \alpha\big(\psi^{-1}(E)\big) = \phi_B = \psi,$$

d. h. $\alpha \circ \beta = \mathrm{id}_{\mathcal{I}}$.

□

Insbesondere haben wir das folgende Korollar gezeigt.

Korollar 3.11
Ist $\dim_K(V) = n$*, so gilt* $V \cong K^n$.

Korollar 3.12
Für zwei endlich-dimensionale K*-Vektorräume* V *und* W *sind gleichwertig:*

a. $V \cong W$.
b. $\dim_K(V) = \dim_K(W)$.

Beweis: Aus der Definition der Dimension folgt, daß eine Basis von V genau $n := \dim_K(V)$ Elemente hat und eine Basis von W genau $m := \dim_K(W)$ Stück. Ist $f : V \to W$ ein Isomorphismus, so überführt er laut Lemma 2.11 eine Basis von V in eine Basis von W. Mithin gilt $n = m$.

Ist umgekehrt $n = m$, so gibt es nach Korollar 3.11 Isomorphismen $f : V \to K^n$ und $g : K^n \to W$. Dann ist $g \circ f : V \to W$ ebenfalls ein Isomorphismus. □

Satz 3.13
Es seien V *und* W *zwei endlich-dimensionale* K*-Vektorräume gleicher Dimension und* $f \in \mathrm{Hom}_K(V, W)$*. Dann sind gleichwertig:*

a. f *ist bijektiv,*
b. f *ist injektiv,*
c. f *ist surjektiv.*

Beweis: Der Beweis ist dem Leser als Übungsaufgabe überlassen. □

AUFGABEN

Aufgabe 3.14
Es sei $U = \{(a_1, \ldots, a_5)^t \in \mathbb{R}^5 \mid a_1 - 2a_2 = 0 = 2a_4 + a_5\} \subset \mathbb{R}^5$. Bestimme die Dimension von U sowie eine Basis, die den Vektor $(2, 1, 1, -1, 2)^t$ enthält.

Aufgabe 3.15
Es sei V ein K-Vektorraum mit $\dim_K(V) = n \geq 1$. Ferner seien $f \in \mathrm{End}_K(V)$ und $x \in V$ so, daß $\big(x, f(x), \ldots, f^{n-1}(x)\big)$ eine linear unabhängige Familie ist.

Zeige, die Familie $\big(\mathrm{id}_V, f, f^2, \ldots, f^{n-1}\big)$ ist linear unabhängig im K-Vektorraum $\mathrm{End}_K(V)$.

Aufgabe 3.16
Es sei $B := \big((3,5,2)^t, (1,1,-1)^t, (2,4,1)^t\big)$.

a. Zeige, B ist eine Basis von $\mathbb{R}^3$.
b. Ersetze *mit Hilfe des Austauschsatzes von Steinitz* zwei Vektoren in B durch die Vektoren $(1,3,2)^t$ und $(-2,1,2)^t$.

Aufgabe 3.17
Es sei V ein K-Vektorräume mit $\dim_K(V) < \infty$ und $U \subseteq V$ ein Unterraum von V. Zeige, genau dann gilt $U = V$, wenn $\dim_K(U) = \dim_K(V)$.

Aufgabe 3.18
Es sei $K^\infty := \{(a_i)_{i\in\mathbb{N}} \mid a_i \in K\}$ der K-Vektorraum aller Folgen im Körper K, vgl. Aufgabe 2.31. Definiere Abbildungen $S, T : K^\infty \to K^\infty$ durch

$$S\big((a_i)_{i\in\mathbb{N}}\big) = (a_1, a_2, a_3, \ldots),$$

und

$$T\big((a_i)_{i\in\mathbb{N}}\big) = (0, a_0, a_1, \ldots).$$

a. Zeige, S und T sind K-lineare Abbildungen.
b. Bestimme $\mathrm{Ker}(S)$, $\mathrm{Im}(S)$, $\mathrm{Ker}(T)$ und $\mathrm{Im}(T)$.
c. Überprüfe die Abbildungen auf Injektivität und Surjektivität.
d. Bestimme die Komposita $S \circ T$ und $T \circ S$.

4. Direkte Summanden und Quotientenräume

In diesem Abschnitt geht es um die Darstellung eines Vektorraumes als (direkte) Summe von Unterräumen sowie um wichtige Dimensionsformeln.

Wir haben schon gesehen, daß die Vereinigung von Unterräumen eines Vektorraumes i. a. kein Vektorraum mehr ist. Aber der von der Vereinigung erzeugte Unterraum ist wichtig genug, um einen eigenen Namen zu bekommen.

Generalvoraussetzung: In diesem Abschnitt ist K wieder ein Körper und V ein beliebiger K-Vektorraum.

Definition 4.1
Es sei $(U_i)_{i\in I}$ eine Familie von Unterräumen von V. Dann heißt

$$\sum_{i\in I} U_i := \left\langle \bigcup_{i\in I} U_i \right\rangle_K$$

die *Summe* von $(U_i)_{i\in I}$.

Ist $I = \{1, \ldots, k\}$, so schreibt man für gewöhnlich

$$U_1 + \ldots + U_k.$$

Bemerkung 4.2
Es gilt:

$$\sum_{i\in I} U_i = \left\{ \sum_{i\in I}{}' u_i \;\middle|\; u_i \in U_i, i \in I \right\},$$

wie man leicht aus Proposition 2.8 sowie der Tatsache ableitet, daß die rechte Seite ein Unterraum ist, der die U_i enthält.

Definition 4.3
Ist $(U_i)_{i\in I}$ eine Familie von Unterräumen von V, dann heißt die Summe $U := \sum_{i\in I} U_i$ *direkt*, falls sich jedes $u \in U$ in eindeutiger Weise in der Form

$$u = \sum_{i\in I}{}' u_i, \quad \text{mit } u_i \in U_i$$

darstellen läßt.

Man schreibt dann

$$U = \bigoplus_{i\in I} U_i,$$

oder $U = U_1 \oplus \cdots \oplus U_k$, falls $I = \{1, \ldots, k\}$.

Beispiel 4.4
Ist $F = (x_i)_{i\in I}$ ein Erzeugendensystem von V, dann gilt

$$V = \sum_{i\in I} \langle x_i \rangle_K.$$

Ist F sogar eine Basis, dann gilt

$$V = \bigoplus_{i\in I} \langle x_i \rangle_K.$$

Der wichtigste Fall ist die Summe $U + U' = \{u + u' \mid u \in U, u' \in U'\}$ von zwei Unterräumen U und U' von V.

Lemma 4.5
Sind U und U' Unterräume von V, dann sind gleichwertig:

a. $U + U' = U \oplus U'$.
b. $U \cap U' = \{0\}$.

Beweis: Ist die Summe $U + U'$ direkt und $x \in U \cap U'$, dann gilt

$$x = x + 0 = 0 + x \in U + U',$$

und wegen der Eindeutigkeit der Darstellung in $U + U'$ muß $x = 0$ sein.

Ist umgekehrt $U \cap U' = \{0\}$ und sind $x_1 + x_1' = x_2 + x_2' \in U + U'$ mit $x_i \in U$ und $x_i' \in U'$, $i = 1, 2$, so gilt:

$$x_1 - x_2 = x_2' - x_1' \in U \cap U' = \{0\}.$$

Also ist $x_1 = x_2$ und $x_1' = x_2'$, d. h. die Darstellung ist eindeutig. □

Von besonderer Bedeutung ist die folgende Dimensionsformel.

Satz 4.6 (Dimensionsformel für Unterräume)
Ist $\dim_K(V) < \infty$ *und sind* U *und* U' *Unterräume von* V*, dann gilt:*

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

Beweis: Wir beweisen mehr, nämlich wie wir geeignete Basen von U, U' und $U \cap U'$ wählen können.

Sei $B_{U\cap U'} := (x_1, \ldots, x_r)$ eine Basis von $U \cap U'$. Wir ergänzen $B_{U\cap U'}$ zu einer Basis $B_U := (x_1, \ldots, x_r, y_1, \ldots, y_s)$ von U, und zu einer Basis $B_{U'} := (x_1, \ldots, x_r, z_1, \ldots, z_t)$ von U'. Das geht nach dem Basisergänzungssatz 2.23.
Behauptung: $B_{U+U'} := (x_1, \ldots, x_r, y_1, \ldots, y_s, z_1, \ldots, z_t)$ ist Basis von $U + U'$.

Dazu zeigen wir, daß jedes Element von $U + U'$ eine eindeutige Linearkombination von Elementen aus $B_{U+U'}$ ist. Sei also $x + x' \in U + U'$ mit $x \in U$ und $x' \in U'$. Dann gilt:

$$x = \sum_{i=1}^{r} \lambda_i x_i + \sum_{j=1}^{s} \mu_j y_j \quad \text{und} \quad x' = \sum_{i=1}^{r} \lambda'_i x_i + \sum_{k=1}^{t} \mu'_k z_k,$$

mit eindeutig bestimmten $\lambda_i, \lambda'_i, \mu_j, \mu'_k \in K$, $i = 1, \ldots, r$, $j = 1, \ldots, s$, $k = 1, \ldots, t$. Daraus folgt:

$$x + x' = \sum_{i=1}^{r} (\lambda_i + \lambda'_i) x_i + \sum_{j=1}^{s} \mu_j y_j + \sum_{k=1}^{t} \mu'_k z_k.$$

Wegen $x = (x+x') - x'$ und $x' = (x+x') - x$ würden sich aus nicht-eindeutigen Koeffizienten für $x + x'$ auch nicht-eindeutige Koeffizienten für x und x' ergeben, was die Behauptung beweist.

Aus der Behauptung folgt,

$$\dim_K(U+U') = r+s+t = (r+s)+(r+t)-r = \dim_K(U)+\dim_K(U')-\dim_K(U\cap U').$$

□

Korollar 4.7
Ist $\dim_K(V) < \infty$*, dann sind für Unterräume* U *und* U' *von* V *die folgenden Aussagen äquivalent:*

a. $V = U \oplus U'$,
b. $V = U + U'$ *und* $\dim_K(V) = \dim_K(U) + \dim_K(U')$, *sowie*
c. $U \cap U' = \{0\}$ *und* $\dim_K(V) = \dim_K(U) + \dim_K(U')$.

Beweis: Dies ist eine direkte Folgerung aus Lemma 4.5, Satz 4.6 und dem folgenden Lemma 4.8. □

Lemma 4.8
Ist $\dim_K(V) < \infty$ *und* $U \subseteq V$ *ein Unterraum. Dann sind gleichwertig:*

a. $U = V$.
b. $\dim_K(U) = \dim_K(V)$.

Beweis: Der Beweis ist Übungsaufgabe 3.17. □

Definition 4.9
Es sei $U \subseteq V$ ein Unterraum von V. Ein Unterraum U' von V mit $V = U \oplus U'$ heißt ein (*direktes*) *Komplement* von U in V.

Lemma 4.10
Jeder Unterraum U von V besitzt ein direktes Komplement.

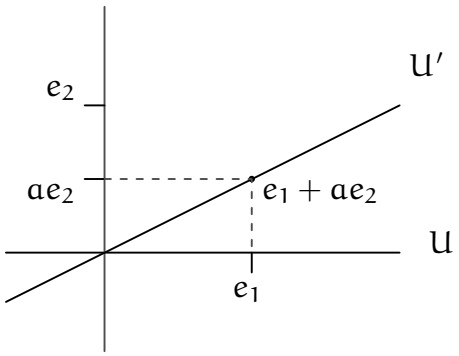
Beweis: Wähle eine Basis B von U und ergänze sie durch eine linear unabhängige Menge B' zu einer Basis $B \cup B'$ von V gemäß dem Basisergänzungssatz 2.23. Dann ist nach Korollar 4.7 $U' := \langle B' \rangle$ ein Komplement von U. □

Man beachte, daß Komplemente nicht eindeutig bestimmt sind.

Beispiel 4.11
Ist $V = \mathbb{R}^2$ und $U = \langle e_2 \rangle$ die y-Achse, dann sind

$$U_a := \langle e_1 + a e_2 \rangle$$

für $a \in \mathbb{R}$, die Geraden mit Steigung a, unendlich viele verschiedene Komplemente von U. Beachte dazu nur, daß $U \cap U_a = \{0\}$ und $\dim_{\mathbb{R}} \left(\mathbb{R}^2\right) = 2 = 1 + 1 = \dim_{\mathbb{R}}(U) + \dim_{\mathbb{R}}(U_a)$.



Wir definieren jetzt als Ersatz für die nicht-eindeutig bestimmten Komplemente den Quotientenraum, der in gewisser Weise ebenfalls U komplementiert und zudem eindeutig bestimmt ist. Er hat aber den Schönheitsfehler, daß er kein Unterraum von V ist.

Der Quotientenraum hat noch einen weiteren Vorteil gegenüber den Komplementen, er existiert auch für Untermoduln von Moduln über Ringen, während das für Komplemente i. a. nicht mehr gilt.

Definition 4.12
Es sei U ein Unterraum des K-Vektorraumes $(V, +, \cdot)$. Dann ist $(U, +)$ eine Untergruppe von $(V, +)$ und wir können die Faktorgruppe $(V/U, +)$ bilden. Auf V/U definieren wir nun eine Skalarmultiplikation

$$\cdot : K \times V/U \to V/U : \big(\lambda, [x]\big) \mapsto [\lambda x].$$

Wir nennen $(V/U, +, \cdot)$ den *Quotientenraum* oder *Faktorraum* von V nach U (bzw. von V modulo U).

Die Definition wird durch das folgende Lemma gerechtfertigt.

Lemma 4.13
Die in Definition 4.12 definierte Skalarmultiplikation ist wohldefiniert und $(V/U, +, \cdot)$ *ist ein* K*-Vektorraum.*

Beweis: Es seien $[x] = [y] \in V/U$ und $\lambda \in K$, dann ist $x - y \in U$, und mithin $\lambda x - \lambda y = \lambda(x - y) \in U$. Aber damit gilt $[\lambda x] = [\lambda y]$, und somit ist die Skalarmultiplikation wohldefiniert.

Wir haben bereits in Satz I.5.8 gesehen, daß $(V/U, +)$ eine abelsche Gruppe ist. Zudem ist V/U per definitionem abgeschlossen gegenüber der Skalarmultiplikation, und die übrigen Axiome übertragen sich von V nach V/U, wie man unmittelbar einsieht. □

Satz 4.14 (Homomorphiesatz) a. *Ist* U *ein Unterraum von* V*, so ist die Restklassenabbildung*

$$\pi : V \to V/U : x \mapsto [x]$$

ein Epimorphismus mit $\operatorname{Ker}(\pi) = U$.

b. *Sind* V *und* W K*-Vektorräume und ist* $f \in \operatorname{Hom}_K(V, W)$*, so ist die gemäß Satz I.5.12 existierende Abbildung*

$$\tilde{f} : V/\operatorname{Ker}(f) \to \operatorname{Im}(f) : [x] \mapsto f(x)$$

ein Isomorphismus.

Beweis: Unter Rückgriff auf die Sätze I.5.8 und I.5.12 bleibt jeweils nur die Verträglichkeit der Abbildungen π und $\tilde{f}$ mit der Skalarmultiplikation zu zeigen. Seien dazu $x \in V$ und $\lambda \in K$. Dann gelten:

$$\pi(\lambda x) = [\lambda x] = \lambda[x]$$

und

$$\tilde{f}\big(\lambda[x]\big) = \tilde{f}\big([\lambda x]\big) = f(\lambda x) = \lambda f(x) = \lambda \tilde{f}\big([x]\big).$$

□

Das folgende Lemma stellt die Beziehung zwischen Komplementen und Quotientenräumen her.

Lemma 4.15
Sei $U \subseteq V$ *ein Unterraum von* V *und* $U' \subseteq V$ *ein Komplement von* U *in* V*. Dann ist die Einschränkung*

$$\pi_| : U' \to V/U : x \mapsto [x]$$

von π *auf* U' *ein Isomorphismus.*

Beweis: Wir zeigen zunächst, daß $\pi_|$ surjektiv ist. Sei dazu $[x] \in V/U$ gegeben. Wegen $V = U \oplus U'$ läßt sich x als $x = y + z$ mit $y \in U$ und $z \in U'$ schreiben. Damit gilt:

$$[x] = [z] = \pi_|(z) \in \operatorname{Im}\big(\pi_|\big).$$

Also ist $\pi_|$ surjektiv.

Bleibt zu zeigen, daß $\pi_|$ injektiv ist, d. h. $\mathrm{Ker}\left(\pi_|\right) = \{0\}$. Sei dazu $z \in \mathrm{Ker}\left(\pi_|\right)$, dann gilt

$$[0] = \pi_|(z) = [z].$$

D. h. $z \in U$. Damit gilt aber $z \in U \cap U' = \{0\}$, also $z = 0$. □

Als Folgerung aus der Dimensionsformel für Unterräume und dem obigen Lemma erhalten wir den folgenden Satz.

Satz 4.16 (Dimensionsformel für lineare Abbildungen)
Es seien V *und* W K*-Vektorräume mit* $\dim_K(V) < \infty$*,* U *sei ein Unterraum von* V *und* $f \in \mathrm{Hom}_K(V, W)$*. Dann gelten:*

a. $\dim_K(V/U) = \dim_K(V) - \dim_K(U)$.
b. $\dim_K(V) = \dim_K\left(\mathrm{Ker}(f)\right) + \dim_K\left(\mathrm{Im}(f)\right)$.

Beweis: a. Ist $U' \subseteq V$ ein Komplement von U in V, so folgt aus Korollar 4.7 und Lemma 4.15

$$\dim_K(V/U) = \dim_K\left(U'\right) = \dim_K(V) - \dim_K(U).$$

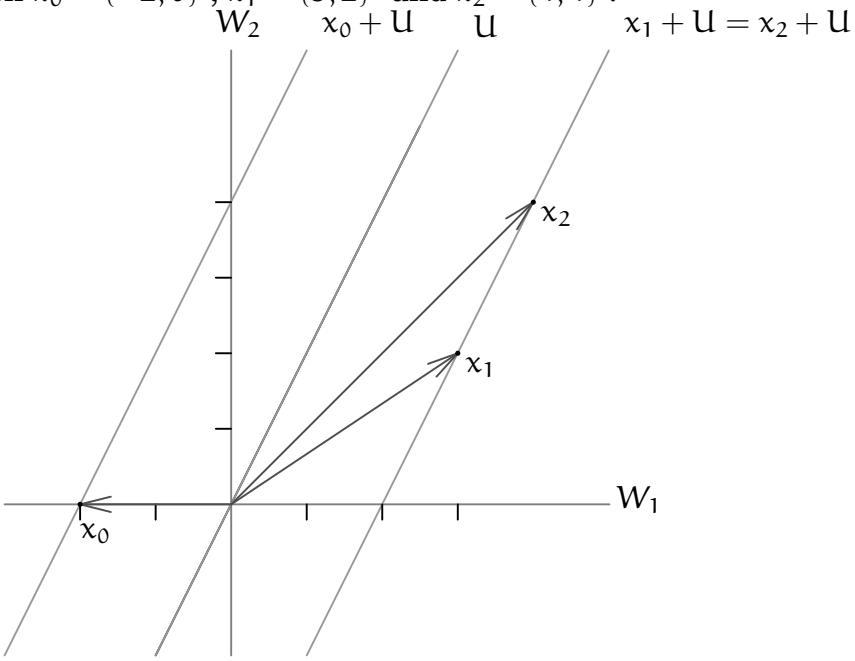
b. Die Formel folgt unmittelbar aus Satz 4.14, Korollar 3.12 sowie Teil a..

□

Beispiel 4.17
Betrachte in $V = \mathbb{R}^2$ den Unterraum

$$U = \langle (1,2)^t \rangle_{\mathbb{R}}$$

sowie die Vektoren $x_0 = (-2,0)^t$, $x_1 = (3,2)^t$ und $x_2 = (4,4)^t$.



$V/U = \left\{[x] = x + U \mid x \in \mathbb{R}^2\right\}$ mit $x + U = x' + U \Leftrightarrow x - x' \in U$. Geometrisch besteht V/U also aus allen Geraden im $\mathbb{R}^2$, die parallel zur Ursprungsgeraden U sind. Die Klasse $[0] = U$ ist die einzige Ursprungsgerade in V/U, nur sie

ist ein Unterraum. Die anderen Klassen $[x]$, $x \notin U$, sind keine Unterräume, sondern sogenannte *affine Unterräume*.

AUFGABEN

Aufgabe 4.18

Es seien $U = \langle (1,0,1,1)^t, (-1,1,0,0)^t \rangle \subset \mathbb{R}^4$ und $U' = \langle (1,0,1,0)^t, (1,1,1,1)^t \rangle \subset \mathbb{R}^4$. Zeige, $\mathbb{R}^4 = U \oplus U'$.

Aufgabe 4.19

Es sei $U = \langle (1,2,3,4)^t, (1,1,1,1)^t \rangle \subset \mathbb{R}^4$. Bestimme mit Hilfe des Austauschsatzes von Steinitz eine Basis von $\mathbb{R}^4/U$.

Aufgabe 4.20

Es sei $V = \mathbb{R}^{\mathbb{R}}$ der $\mathbb{R}$-Vektorraum aller Abbildungen von $\mathbb{R}$ nach $\mathbb{R}$, $U := \{f \in V \mid f(-x) = f(x) \ \forall\ x \in \mathbb{R}\}$ und $U' := \{f \in V \mid f(-x) = -f(x) \ \forall\ x \in \mathbb{R}\}$ – vgl. Aufgabe 1.19. Zeige, $V = U \oplus U'$.

Aufgabe 4.21

Es sei $V = \mathbb{R}^3$, $U = \{(x+y, y, y-x)^t \mid x, y \in \mathbb{R}\}$ und $U' = \{(x,y,z)^t \in \mathbb{R}^3 \mid z = 2x + y\}$. Bestimme Basen von $U + U'$, $U \cap U'$, V/U und V/U'.

Aufgabe 4.22

Es sei K ein Körper. $U := \{(a_1, \ldots, a_n) \mid a_1 = \ldots = a_n \in K\}$ und $U' := \left\{(a_1, \ldots, a_n) \;\middle|\; a_i \in K, \sum_{i=1}^n a_i = 0\right\}$ sind Unterräume des K^n. Bestimme $\dim_K(U)$, $\dim_K(U')$, $\dim_K(U \cap U')$ und $\dim_K(U + U')$.

Aufgabe 4.23

Es sei V ein K-Vektorraum mit $\dim_K(V) = 5$, und U und U' Unterräume mit $\dim_K(U) = 3$ und $\dim_K(U') = 4$.

a. Welche Werte kann $\dim_K(U \cap U')$ annehmen?
b. Gib für jeden der Werte von $\dim_K(U \cap U')$ ein Beispiel (K, V, U, U') an.

Aufgabe 4.24

Ist U ein Unterraum des K-Vektorraums V und F und G seien Familien in V.

a. Gilt $V = \langle F \rangle$, dann gilt $V/U = \langle x + U \mid x \in F \rangle$.
b. Gilt $V/U = \langle x + U \mid x \in F \rangle$ und $U = \langle G \rangle$, so gilt $V = \langle F \cup G \rangle$.

Aufgabe 4.25 (Invariante Unterräume)

Es sei $U \subseteq V$ ein Unterraum des K-Vektorraums V, und $f \in \mathrm{End}_K(V)$ mit $f(U) \subseteq U$. Wir sagen dann, U ist f-*invariant*.

Definiere $f_U : U \to U$ durch $f_U(x) = f(x)$ für alle $x \in U$, und $f_{V/U} : V/U \to V/U$ durch $f_{V/U}(x + U) = f(x) + U$ für alle $x + U \in V/U$. Zeige:

a. f_U und $f_{V/U}$ sind K-lineare Abbildungen, insbesondere ist $f_{V/U}$ wohldefiniert.
b. $\mathrm{Ker}(f_U) = \mathrm{Ker}(f) \cap U$ und $\mathrm{Im}(f_U) \subseteq \mathrm{Im}(f) \cap U$.
c. $\mathrm{Im}\left(f_{V/U}\right) = \big(\mathrm{Im}(f) + U\big)/U$.
d. $\mathrm{Ker}\left(f_{V/U}\right) \supseteq \big(\mathrm{Ker}(f) + U\big)/U$.
e. Falls $\dim(V) < \infty$, dann gilt

$$\dim_K\big(\mathrm{Ker}(f)\big) - \dim_K\big(\mathrm{Ker}(f_U)\big) \leq \dim_K\left(f_{V/U}\right) \leq \dim_K\big(\mathrm{Ker}(f)\big).$$

Aufgabe 4.26
Es sei V ein (nicht-notwendig endlich-dimensionaler) K-Vektorraum, $f \in \mathrm{End}_K(V)$ und $U \subseteq V$ ein f-invarianter Unterraum. Zeige:

a. Genau dann ist $\mathrm{Im}\left(f_U\right) = \mathrm{Im}(f)$, wenn $V = U + \mathrm{Ker}(f)$.
b. Genau dann ist $\mathrm{Ker}\left(f_{V/U}\right) = 0$, wenn $\mathrm{Ker}\left(f\right) = \mathrm{Ker}(f_U)$.

Aufgabe 4.27
Es sei $V = U_1 \oplus U_2$ ein K-Vektorraum und $f \in \mathrm{End}_K(V)$ mit $f(U_i) \subseteq U_i$ für $i = 1, 2$. Zeige:

a. $\mathrm{Im}(f) = \mathrm{Im}\left(f_{U_1}\right) \oplus \mathrm{Im}\left(f_{U_2}\right)$.
b. $\mathrm{Ker}(f) = \mathrm{Ker}\left(f_{U_1}\right) \oplus \mathrm{Ker}\left(f_{U_2}\right)$.

Aufgabe 4.28
Sei V ein endlich-dimensionaler K-Vektorraum, $f \in \mathrm{End}_K(V)$. Zeige, daß dann die folgenden Aussagen gleichwertig sind:

a. $V = \mathrm{Ker}(f) \oplus \mathrm{Im}(f)$,
b. $V = \mathrm{Ker}(f) + \mathrm{Im}(f)$,
c. $\mathrm{Ker}(f) \cap \mathrm{Im}(f) = \{0\}$,
d. $\mathrm{Ker}(f^2) = \mathrm{Ker}(f)$,
e. $\mathrm{Im}(f^2) = \mathrm{Im}(f)$.

Aufgabe 4.29
Gib Beispiele $f, g \in \mathrm{Hom}_{\mathbb{R}}\left(\mathbb{R}^2, \mathbb{R}^2\right)$ an mit $\mathrm{Im}(f) = \mathrm{Ker}(f)$ und $\mathbb{R}^2 = \mathrm{Im}(g) \oplus \mathrm{Ker}(g)$.

Aufgabe 4.30
Es sei V ein K-Vektorraum, $\mathrm{char}(K) \neq 2$, und $f \in \mathrm{End}_K(V)$ mit $f^2 = \mathrm{id}_V$. Zeige:

a. $\mathrm{Im}\left(f - \mathrm{id}_V\right) = \mathrm{Ker}\left(f + \mathrm{id}_V\right)$, und
b. $V = \mathrm{Ker}\left(f - \mathrm{id}_V\right) \oplus \mathrm{Ker}\left(f + \mathrm{id}_V\right)$.

Aufgabe 4.31
Es sei V ein K-Vektorraum, $f, g \in \mathrm{End}_K(V)$. (Vgl. Aufgabe 1.23.) Zeige:

a. Gilt $f \circ g = g \circ f$, so sind $\mathrm{Im}(f)$ und $\mathrm{Ker}(f)$ g-invariant.
b. Ist f eine Projektion und sind $\mathrm{Im}(f)$ und $\mathrm{Ker}(f)$ g-invariant, so gilt $f \circ g = g \circ f$.
c. Sind f und g Projektionen mit $f \circ g = g \circ f$, so gilt
   (i) $\mathrm{Im}(f \circ g) = \mathrm{Im}(f) \cap \mathrm{Im}(g)$, und
   (ii) $\mathrm{Ker}(f \circ g) = \mathrm{Ker}(f) + \mathrm{Ker}(g)$.
d. Ist $V = U_1 \oplus U_2 = W_1 \oplus W_2$ mit $U_1 \subseteq W_2$ und $W_1 \subseteq U_2$, so gilt

$$V = \left(U_1 + W_1\right) \oplus \left(U_2 \cap W_2\right).$$

Aufgabe 4.32 (Isomorphiesätze)
Es seien $U, U' \subseteq V$ Unterräume des K-Vektorraumes V. Zeige:

a. $(U + U')/U \cong U'/(U \cap U')$,
b. Falls $U' \subseteq U$, dann gilt $(V/U')/(U/U') \cong V/U$.

5. Lineare Abbildungen und Matrizen

Wir haben im Wesentlichen bereits in Korollar 2.14 gesehen, daß es eine Bijektion zwischen der Menge $\mathrm{Mat}(m \times n, K)$ der $(m \times n)$-Matrizen mit Einträgen in K und der Menge $\mathrm{Hom}_K\left(K^n, K^m\right)$ der linearen Abbildungen $K^n \to K^m$ gibt. Wir wollen diese Bijektion verallgemeinern zu einer Bijektion zwischen $\mathrm{Mat}(m \times n, K)$ und $\mathrm{Hom}_K(V, W)$ für beliebige K-Vektorräume mit $\dim_K(V) = n$ und $\dim_K(W) = m$. Zudem wird sich die Bijektion als Isomorphismus von Vektorräumen erweisen.

Ein wichtiger Teil dieses Abschnittes ist dem Verhalten von Matrizen einer linearen Abbildung bei Basiswechsel gewidmet.

Zum Schluß werden wir zeigen, daß der Zeilenrang einer Matrix mit dem Spaltenrang übereinstimmt und daß die reduzierte Zeilen-Stufen-Form einer Matrix eindeutig ist.

Generalvoraussetzung: In diesem Abschnitt sei K ein Körper, und V und W seien zwei endlich-dimensionale K-Vektorräume mit Basen $B = (b_1, \ldots, b_n)$ und $D = (d_1, \ldots, d_m)$. Insbesondere gelte also $\dim_K(V) = n$ und $\dim_K(W) = m$.

Ferner sei $E = (e_1, \ldots, e_n)$ die kanonische Basis von K^n und $F = (f_1, \ldots, f_m)$ die kanonische Basis von K^m.

Proposition 5.1
Die Vektorräume $\mathrm{Mat}(m \times n, K)$ *und* $\mathrm{Hom}_K\left(K^n, K^m\right)$ *sind zueinander isomorph, wobei die Abbildungen* Φ *und* Ψ *zueinander inverse Isomorphismen sind:*

$$\begin{array}{rccc} \Phi & : & \mathrm{Mat}(m \times n, K) & \to & \mathrm{Hom}_K\left(K^n, K^m\right) \\ & & A & \mapsto & f_A, \end{array}$$

und

$$\begin{array}{rccc} \Psi & : & \mathrm{Hom}_K\left(K^n, K^m\right) & \to & \mathrm{Mat}(m \times n, K) \\ & & f & \mapsto & A_f, \end{array}$$

wobei die Matrix A_f *dadurch definiert wird, daß ihr* i*-ter Spaltenvektor* a^i *gerade* $f(e_i)$ *ist.*

Beweis: Nach Proposition 1.8 ist das Bild von Φ in $\mathrm{Hom}_K\left(K^n, K^m\right)$ enthalten, und aus Lemma I.8.7 wissen wir, daß die Abbildung Φ K-linear ist. Ferner folgt aus Korollar 2.14, daß es zu *jeder* linearen Abbildung $f : K^n \to K^m$ eine *eindeutig bestimmte* $m \times n$-Matrix A mit $f = f_A$ gibt, d. h. Φ ist bijektiv und damit ein Isomorphismus.

Außerdem gilt für die Umkehrabbildung Φ^{-1}, ebenfalls nach Korollar 2.14, daß der i-te Spaltenvektor von $\Phi^{-1}(f)$ für $f \in \mathrm{Hom}_K\left(K^n, K^m\right)$ gerade $f(e_i)$ ist. Also ist $\Phi^{-1} = \Psi$, und damit ist Ψ linear und ein Isomorphismus nach Lemma 1.7. □

Bemerkung 5.2
Es seien $f \in \mathrm{Hom}_K\left(K^n, K^m\right)$ und $g \in \mathrm{Hom}_K\left(K^m, K^p\right)$. Dann gilt:

$$A_{g \circ f} = A_g \circ A_f.$$

Beachte dazu, daß wegen $f_{A_f} = f$ und $f_{A_g} = g$ für $i = 1, \ldots, n$ gilt:

$$A_{g\circ f}e_i = (g \circ f)(e_i) = g\big(f(e_i)\big) = g(A_f e_i) = A_g(A_f e_i) = (A_g \circ A_f)e_i.$$

Das heißt aber, daß die n Spalten der beiden Matrizen übereinstimmen.

Auf dem K-Vektorraum $\mathrm{End}_K(V) = \mathrm{Hom}_K(V, V)$ haben wir neben der Addition und der Skalarmultiplikation eine dritte Operation, die Komposition. Dieser Umstand motiviert die folgende Definition, die wir allgemeiner für kommutative Ringe mit Eins formulieren.

Definition 5.3
Sei R ein kommutativer Ring mit Eins 1_R. Ein R-Modul $(B, +, \cdot)$, auf dem zusätzlich eine Multiplikation

$$\circ : B \times B \to B : (x, y) \mapsto x \circ y$$

definiert ist, so daß $(B, +, \circ)$ ein Ring mit Eins 1_B ist, heißt eine R-*Algebra*, falls die Skalarmultiplikation mit der Ringmultiplikation verträglich ist, d. h. für $\lambda \in R$ und $x, y \in B$ gelten:

$$\lambda \cdot (x \circ y) = (\lambda \cdot x) \circ y = x \circ (\lambda \cdot y).$$

Ein R-*Algebrenhomomorphismus* ist ein R-Modulhomomorphismus $\varphi : A \to B$ zwischen zwei R-Algebren A und B, der zugleich ein Ringhomomorphismus ist mit $\varphi(1_A) = 1_B$.

Beispiel 5.4 a. $\big(\mathrm{End}_K(V), +, \cdot, \circ\big)$ ist eine K-Algebra.
Beachte dazu, daß aus Beispiel 1.13 folgt, daß $\mathrm{End}_K(V)$ ein K-Vektorraum ist, während mit Hilfe von Lemma 1.7 folgt, daß $\mathrm{End}_K(V)$ auch ein (i. a. nicht-kommutativer) Ring mit Eins id_V ist, in dem zudem die Verträglichkeit mit der Skalarmultiplikation mit der Ringmultiplikation gewährleistet ist.

b. $\big(\mathrm{Mat}(n, K), +, \cdot, \circ\big)$ ist ebenfalls eine K-Algebra.
Beachte dazu, daß $\mathrm{Mat}(n, K)$ nach Beispiel 1.2 ein K-Vektorraum ist und nach Korollar I.8.12 ein (i. a. nicht-kommutativer) Ring mit Eins $\mathbb{1}_n$, in dem nach Lemma I.8.7 zudem die Verträglichkeit der Skalarmultiplikation mit der Ringmultiplikation gewährleistet ist.

Proposition 5.5
Die K*-Algebren* $\mathrm{Mat}(n, K)$ *und* $\mathrm{End}_K\big(K^n\big)$ *sind zueinander isomorph und die in Proposition 5.1 definierten Abbildungen* Φ *und* Ψ *sind in diesem Falle zueinander inverse* K*-Algebrenisomorphismen.*

Beweis: Wegen Proposition 5.1 und Bemerkung I.7.5 ist nur die Verträglichkeit von Φ mit $\circ$ zu prüfen. Diese folgt aber aus Satz I.8.8. □

Wir kommen jetzt zum Zusammenhang zwischen Matrizen und linearen Abbildungen zwischen Vektorräumen mit Basen.

Definition 5.6
Ist $f \in \mathrm{Hom}_K(V,W)$ und sind ϕ_B und ϕ_D die Karten von V zur Basis B bzw. von W zur Basis D, so heißt die Komposition

$$\phi_D \circ f \circ \phi_B^{-1} : K^n \to K^m$$

die *Koordinatendarstellung* von f bezüglich der Basen B und D (oder kurz, bezüglich (B,D)).

Mit $M_D^B(f)$ bezeichnen wir die nach Proposition 5.1 zu $\phi_D \circ f \circ \phi_B^{-1}$ assoziierte Matrix $\Psi\big(\phi_D \circ f \circ \phi_B^{-1}\big) = A_{\phi_D \circ f \circ \phi_B^{-1}}$, und nennen sie auch die *Matrixdarstellung* von f bezüglich (B,D).

Für einen Vektor $x = \sum_{i=1}^n \lambda_i b_i$ nennt man den Vektor

$$M_B(x) := \phi_B(x) = (\lambda_1, \ldots, \lambda_n)^t \in K^n$$

den *Koordinatenvektor* oder die *Koordinaten* von x bezüglich B.

Bemerkung 5.7
Die Koordinatendarstellung von f schreibt man am besten in Form eines *kommutativen Diagramms*:

$$\begin{array}{ccc} V & \xrightarrow{\quad f \quad} & W \\ {\scriptstyle \phi_B}\downarrow{\scriptstyle \cong} & \circlearrowleft & {\scriptstyle \cong}\downarrow{\scriptstyle \phi_D} \\ K^n & \xrightarrow[\phi_D \circ f \circ \phi_B^{-1}]{} & K^m \end{array}$$

Die *Kommutativität eines Diagramms* von Abbildungen bedeutet, daß die Komposition von Abbildungen mit gleichem Anfangs- und Endpunkt längs Pfeilen (in Pfeilrichtung) gleich sind. Hier bedeutet dies einfach

$$\phi_D \circ f = \big(\phi_D \circ f \circ \phi_B^{-1}\big) \circ \phi_B.$$

Wie wir in Satz 3.10 gesehen haben, ist die Wahl eines Isomorphismus $V \to K^n$ äquivalent zur Wahl einer Basis von V. Die Wahl einer Basis B von V bestimmt eindeutig einen Isomorphismus $\phi_B : V \to K^n$, mit dessen Hilfe wir V und K^n identifizieren können.

Durch Wahl von Basen B in V und D in W können wir lineare Abbildungen $f : V \to W$ mit ihrer Koordinatendarstellung $\phi_D \circ f \circ \phi_B^{-1} : K^n \to K^m$ identifizieren und damit auch mit der zugehörigen Matrix $M_D^B(f)$, vermittels Proposition 5.1. Wir werden häufig nicht zwischen der Koordinatendarstellung $\phi_D \circ f \circ \phi_B^{-1}$ und der Matrixdarstellung $M_D^B(f)$ unterscheiden.

Für einen Vektor $x = \sum_{i=1}^n \lambda_i b_i$ gilt offenbar

$$M_D\big(f(x)\big) = M_D^B(f) \circ M_B(x),$$

d. h. der Koordinatenvektor $M_D\big(f(x)\big)$ von $f(x)$ bezüglich der Basis D ist das Matrixprodukt der Matrixdarstellung $M_D^B(f)$ von f bezüglich B und D mit dem Koordinatenvektor $M_B(x)$ von x bezüglich B.

Insbesondere ist die i*-te Spalte von* $M_D^B(f)$ *gerade der Koordinatenvektor von* $f(b_i)$ *bezüglich* D.

Beispiel 5.8 a. Ist $V = K^n$ und $W = K^m$, dann gilt $\phi_E = \mathrm{id}_{K^n}$ und $\phi_F = \mathrm{id}_{K^m}$. Mithin folgt $M_F^E(f) = A_f$ für $f \in \mathrm{Hom}_K\left(K^n, K^m\right)$.

b. Die Aussagen von Lemma I.8.18 lassen sich auch so formulieren, daß die Elementar- bzw. Permutationsmatrizen die Matrixdarstellungen der elementaren Zeilenoperationen sind. (Vgl. auch Bemerkung 2.16.)

c. Sei $V = K^2$ mit Basis $B = (b_1, b_2) = \left((1,2)^t, (1,1)^t\right)$ und $W = K^3$ mit Basis $D = (d_1, d_2, d_3) = \left((1,1,0)^t, (0,1,1)^t, (0,0,1)^t\right)$, und sei $f : V \to W$ die lineare Abbildung, die definiert wird durch

$$\begin{aligned} b_1 &\mapsto 3d_1 - 4d_2 + 6d_3, \\ b_2 &\mapsto 3d_1 - 3d_2 + 4d_3. \end{aligned}$$

Dann gilt:

$$M_D^B(f) = \begin{pmatrix} 3 & 3 \\ -4 & -3 \\ 6 & 4 \end{pmatrix}.$$

Unser Ziel ist es nun, in obigem Beispiel aus $M_D^B(f)$ die Matrix $A_f = M_F^E(f)$ zu bestimmen. Dazu führen wir folgende allgemeine Begriffsbildung ein.

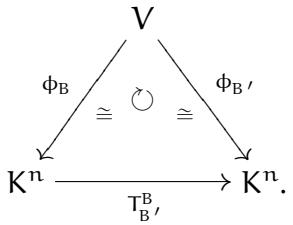
Definition 5.9

Sind $B = (b_1, \ldots, b_n)$ und $B' = \left(b'_1, \ldots, b'_n\right)$ zwei Basen von V, so heißt der Isomorphismus

$$T_{B'}^B := \phi_{B'} \circ \phi_B^{-1} : K^n \to K^n$$

Basiswechsel oder *Koordinatentransformation* bezüglich (B, B').

Wir haben also folgendes kommutatives Diagramm:



Wir schreiben für die Matrixdarstellung $M_E^E\left(T_{B'}^B\right)$ der Einfachheit halber wieder $T_{B'}^B$.

Bemerkung 5.10

Offensichtlich ist der Basiswechsel ein Spezialfall der Koordinatendarstellung, denn es gilt

$$T_{B'}^B = \phi_{B'} \circ \mathrm{id}_V \circ \phi_B^{-1},$$

und entsprechend für die Matrixdarstellung

$$T_{B'}^B = M_{B'}^B(\mathrm{id}_V).$$

Damit gilt ferner, daß *die* i*-te Spalte von* $T_{B'}^B$ *der Koordinatenvektor* $M_{B'}(b_i)$ *von* b_i *bezüglich* B' *ist.*

Außerdem folgt unmittelbar aus der Definition von $T^B_{B'}$, daß

$$\left(T^B_{B'}\right)^{-1} = T^{B'}_B.$$

Satz 5.11 (Basiswechsel)
Seien B *und* B' *Basen von* V, D *und* D' *Basen von* W *und* $f \in \mathrm{Hom}_K(V,W)$. *Dann gilt:*

$$M^{B'}_{D'}(f) = T^D_{D'} \circ M^B_D(f) \circ T^{B'}_B.$$

Ist speziell $V = W$, $B = D$, $B' = D'$, $A = M^B_D(f)$, $A' = M^{B'}_{D'}(f)$ *und* $T = T^B_{B'}$, *dann gilt:*

$$A' = T \circ A \circ T^{-1}.$$

Beweis: Es gilt:

$$\left(\phi_{D'} \circ \phi_D^{-1}\right) \circ \left(\phi_D \circ f \circ \phi_B^{-1}\right) \circ \left(\phi_B \circ \phi_{B'}^{-1}\right) = \phi_{D'} \circ f \circ \phi_{B'}^{-1}.$$

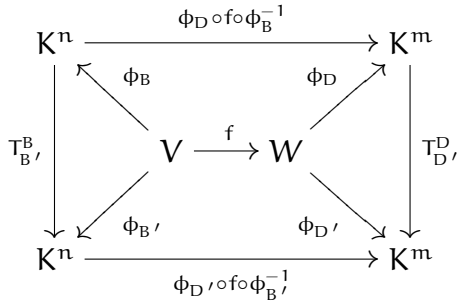
Mittels Bemerkung 5.2 gilt dann aber:

$$T^D_{D'} \circ M^B_D(f) \circ T^{B'}_B = A_{\phi_{D'} \circ \phi_D^{-1}} \circ A_{\phi_D \circ f \circ \phi_B^{-1}} \circ A_{\phi_B \circ \phi_{B'}^{-1}}$$

$$= A_{(\phi_{D'} \circ \phi_D^{-1}) \circ (\phi_D \circ f \circ \phi_B^{-1}) \circ (\phi_B \circ \phi_{B'}^{-1})} = A_{\phi_{D'} \circ f \circ \phi_{B'}^{-1}} = M^{B'}_{D'}(f).$$

□

Bemerkung 5.12
Man kann den Beweis von Satz 5.11 als *Diagrammjagd* in folgendem kommutativen Diagramm bezeichnen:



Dabei impliziert die Kommutativität der inneren Diagramme die des äußeren. Die Aussage des Satzes zusammen mit der $\left(T^B_{B'}\right)^{-1} = T^{B'}_B$ ist dann (unter Berücksichtigung von Bemerkung 5.2) gerade die Kommutativität des Diagramms.

Beispiel 5.13
Wir wollen nun für die Abbildung in Beispiel 5.8 c. die Matrixdarstellung $M^E_F(f)$ bezüglich der kanonischen Basen berechnen. Nach Satz 5.11 gilt:

$$M^E_F(f) = T^D_F \circ M^B_D(f) \circ T^E_B.$$

Um T_F^D auszurechnen, müssen wir d_1, d_2 und d_3 in der kanonischen Basis ausdrücken und die Koeffizienten als Spaltenvektoren in die Matrix T_F^D übertragen:

$$T_F^D = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

Um T_B^E zu ermitteln, müßten wir die Einheitsvektoren e_1 und e_2 als Linearkombination der Basis B darstellen, was auf das Lösen zweier Gleichungssysteme hinaus liefe. Stattdessen können wir aber auch T_E^B bestimmen und anschließend invertieren, was sich im Falle einer (2×2)-Matrix anbietet, da das Invertieren sehr einfach ist (vgl. Beispiel 10.8),

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix},$$

sofern die Matrix invertierbar ist.

Analog zum Fall von T_F^D erhalten wir

$$T_E^B = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix},$$

und somit

$$T_B^E = \left(T_E^B\right)^{-1} = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} -1 & 1 \\ 2 & -1 \end{pmatrix}.$$

Also gilt:

$$M_F^E(f) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \circ \begin{pmatrix} 3 & 3 \\ -4 & -3 \\ 6 & 4 \end{pmatrix} \circ \begin{pmatrix} -1 & 1 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

Wir können jetzt für Vektorräume mit Basen, ganz analog zu Proposition 5.1 für K^n, den Vektorraum $\mathrm{Hom}_K(V, W)$ mit dem Vektorraum $\mathrm{Mat}(m \times n, K)$ identifizieren.

Satz 5.14 a. *Die Vektorräume* $\mathrm{Hom}_K(V, W)$ *und* $\mathrm{Mat}(m \times n, K)$ *sind zueinander isomorph, wobei die Abbildungen* M_D^B *und* L_D^B *zueinander inverse Isomorphismen sind:*

$$\begin{array}{rccc} M_D^B \;:\; \mathrm{Hom}_K(V, W) & \to & \mathrm{Mat}(m \times n, K) \\ f & \mapsto & M_D^B(f), \end{array}$$

und

$$\begin{array}{rccc} L_D^B \;:\; \mathrm{Mat}(m \times n, K) & \to & \mathrm{Hom}_K(V, W) \\ A & \mapsto & L_D^B(A), \end{array}$$

mit

i-ter Spaltenvektor von $M_D^B(f)$ ist der Koordinatenvektor $M_D\big(f(b_i)\big)$ von $f(b_i)$ bezüglich der Basis D,

und die lineare Abbildung $L_D^B(A)$ ist definiert durch

$L_D^B(A)(b_i)$ *ist der Vektor in* W*, dessen Koordinatendarstellung bezüglich* D *gerade der* i*-te Spaltenvektor von* A *ist.*

b. *Für* $V = W$ *ist*

$$M_B^B : \mathrm{End}_K(V) \to \mathrm{Mat}(n, K)$$

ein K*-Algebrenisomorphismus mit* L_B^B *als inversem Isomorphismus.*

Beweis: Wegen Lemma 1.7 f. ist die Abbildung

$$\Psi_D^B : \mathrm{Hom}_K(V, W) \to \mathrm{Hom}_K\left(K^n, K^m\right) : f \mapsto \phi_D \circ f \circ \phi_B^{-1}$$

eine K-lineare Abbildung. Ferner ist die Abbildung

$$\Phi_D^B : \mathrm{Hom}_K\left(K^n, K^m\right) \to \mathrm{Hom}_K(V, W) : g \mapsto \phi_D^{-1} \circ g \circ \phi_B$$

offensichtlich die inverse Abbildung, so daß beides Isomorphismen sind. Ferner gilt, mit den Notationen aus Proposition 5.1,

$$M_D^B = \Psi \circ \Psi_D^B \quad \text{und} \quad L_D^B = \Phi_D^B \circ \Phi,$$

so daß die Behauptung aus Proposition 5.1 folgt. □

Die Koordinatentransformationen in Vektorräumen mit Basen führen auf folgende Äquivalenzbegriffe für Matrizen, die wir allgemeiner für Matrizen über Ringen formulieren.

Definition 5.15
Sei R ein kommutativer Ring mit Eins.

a. Zwei Matrizen A und A' in $\mathrm{Mat}(m \times n, R)$ heißen *äquivalent*, falls es invertierbare Matrizen $S \in \mathrm{Gl}_m(R)$ und $T \in \mathrm{Gl}_n(R)$ gibt mit

$$A' = S \circ A \circ T^{-1}.$$

b. Zwei Matrizen A und A' in $\mathrm{Mat}(n, R)$ heißen *ähnlich* oder *konjugiert*, falls es eine invertierbare Matrix $T \in \mathrm{Gl}_n(R)$ gibt mit

$$A' = T \circ A \circ T^{-1}.$$

Bemerkung 5.16
Äquivalenz und Ähnlichkeit von Matrizen sind in der Tat Äquivalenzrelationen, wie der Leser leicht nachprüfen kann.

Für einen Körper K können wir sie auch wie folgt formulieren:

a. A und A' in $\mathrm{Mat}(m \times n, K)$ sind genau dann äquivalent, wenn es Basen B und D von K^n bzw. K^m gibt, so daß

$$f_{A'} = L_D^B(A)$$

gilt. D. h. bezüglich geeigneter Basen definieren A' und A die gleiche Abbildung.

b. A und A' in $\mathrm{Mat}(n, K)$ sind genau dann ähnlich, wenn es eine Basis B des K^n gibt, so daß

$$f_{A'} = L_B^B(A)$$

gilt. In diesem Fall wird also im Bild- und Urbildraum die gleiche Basis gewählt.

Definition 5.17
Ist $f \in \mathrm{Hom}_K(V, W)$, so definieren wir den *Rang* von f als

$$\mathrm{rang}(f) := \dim_K \big(\mathrm{Im}(f)\big).$$

Ferner definieren wir für eine Matrix $A \in \mathrm{Mat}(m \times n, K)$ den *Rang* von A durch:

$$\mathrm{rang}(A) := \mathrm{rang}(f_A).$$

Satz 5.18 (Normalform bezüglich Äquivalenz)
Es sei $f \in \mathrm{Hom}_K(V, W)$ *mit* $\mathrm{rang}(f) = r$. *Dann gibt es Basen* B *von* V *und* D *von* W *mit*

$$M_D^B(f) = \left(\begin{array}{c|c} \mathbb{1}_r & 0 \\ \hline 0 & 0 \end{array}\right),$$

wobei hier 0 *jeweils die Nullmatrix der entsprechenden Größe meint.*[9]

Beweis: Wähle vermöge Lemma 4.10 ein Komplement U von $\mathrm{Ker}(f)$. Nach Satz 4.14 und Lemma 4.15 ist

$$f_{|U} : U \to \mathrm{Im}(f) : x \mapsto f(x)$$

ein Isomorphismus.

Wähle eine Basis $(d_1, \ldots, d_r)$ von $\mathrm{Im}(f)$. Dann ist $(b_1, \ldots, b_r)$ mit $b_i := (f_{|U})^{-1}(d_i)$ eine Basis von U, nach Lemma 2.11. Wähle nun eine Basis $(b_{r+1}, \ldots, b_n)$ von $\mathrm{Ker}(f)$, dann ist wegen $V = U \oplus \mathrm{Ker}(f)$ $B = (b_1, \ldots, b_n)$ eine Basis von V. Ergänze ferner $(d_1, \ldots, d_r)$ zu einer Basis $D = (d_1, \ldots, d_m)$ von W vermöge Korollar 2.23. Dann gilt:

$$f(b_i) = \begin{cases} d_i, & i = 1, \ldots, r, \\ 0, & i = r+1, \ldots, n. \end{cases}$$

Also hat $M_D^B(f)$ die gewünschte Gestalt. □

Korollar 5.19 (Normalform bezüglich Äquivalenz)
Zu $A \in \mathrm{Mat}(m \times n, K)$ *mit* $r = \mathrm{rang}(A)$ *existieren Matrizen* $S \in \mathrm{Gl}_m(K)$ *und* $T \in \mathrm{Gl}_n(K)$ *mit*

$$S \circ A \circ T^{-1} = \left(\begin{array}{c|c} \mathbb{1}_r & 0 \\ \hline 0 & 0 \end{array}\right). \tag{11}$$

[9] Man bezeichnet die vier Matrizen $\mathbb{1}_r \in \mathrm{Mat}(r, K)$, $0 \in \mathrm{Mat}\big(r \times (n-r), K\big)$, $0 \in \mathrm{Mat}\big((n-r) \times r, K\big)$ und $0 \in \mathrm{Mat}(n-r, K)$ auch als *Blöcke* von $M_D^B(f)$ und die Matrix $M_D^B(f)$ als eine *Blockmatrix*.

Beweis: Anwendung des Satzes 5.18 auf $f_A : K^n \to K^m$ liefert, es gibt Basen B und D von K^n bzw. K^m mit

$$\left(\begin{array}{c|c} \mathbb{1}_r & 0 \\ \hline 0 & 0 \end{array}\right) = M_D^B(f_A) = T_D^F \circ M_F^E(f_A) \circ T_E^B = T_D^F \circ A \circ T_E^B.$$

Die Behauptung folgt also, da $S := T_D^F$ und $T := T_B^E$ invertierbar sind. □

Bemerkung 5.20
Aus Korollar 5.19 folgt, daß zwei Matrizen genau dann äquivalent sind, wenn sie den gleichen Rang haben.

$\mathrm{Mat}(m \times n, K)$ zerfällt also in $\min\{m, n\}+1$ Äquivalenzklassen und jede Äquivalenzklasse ist durch den Rang einer ihrer Matrizen eindeutig bestimmt. Darüber hinaus besitzt jede Äquivalenzklasse $[A]$, $A \in \mathrm{Mat}(m \times n, K)$, einen besonders schönen Repräsentanten, nämlich

$$\left(\begin{array}{c|c} \mathbb{1}_r & 0 \\ \hline 0 & 0 \end{array}\right).$$

Diesen Repräsentanten von $[A]$ nennt man die *Normalform von A bezüglich Äquivalenz*.

Die Klasseneinteilung von $\mathrm{Mat}(m \times n, K)$ bezüglich Äquivalenz ist also ziemlich grob. Die Klasseneinteilung von $\mathrm{Mat}(n, K)$ bezüglich Konjugation (oder Ähnlichkeit) ist viel feiner und die Klassifikation sowie die Herleitung einer Normalform bezüglich Konjugation ist viel schwieriger. Das wird ein wichtiges Thema des nächsten Semesters sein.

Korollar 5.21
Für $A \in \mathrm{Mat}(n, K)$ sind gleichwertig:

a. *A ist invertierbar.*
b. $\mathrm{rang}(A) = n$.
c. *Es gibt eine Matrix $B \in \mathrm{Mat}(n, K)$ mit $B \circ A = \mathbb{1}_n$.*

Beweis: Für $A \in \mathrm{Mat}(n, K)$ und $S, T \in \mathrm{Gl}_n(K)$ gilt, $S \circ A \circ T$ ist genau dann invertierbar, wenn A invertierbar ist.

Denn, ist A invertierbar, so ist $T^{-1} \circ A^{-1} \circ S^{-1}$ eine Inverse von $S \circ A \circ T$. Ist umgekehrt $S \circ A \circ T$ invertierbar mit Inverse B, dann ist $T \circ B \circ S$ eine Inverse von A, da

$$A \circ (T \circ B \circ S) = S^{-1} \circ (S \circ A \circ T \circ B) \circ S = S^{-1} \circ \mathbb{1}_n \circ S = S^{-1} \circ S = \mathbb{1}_n,$$

und analog $(T \circ B \circ S) \circ A = \mathbb{1}_n$.

Da nun die Blockmatrix in (11), für $m = n$, genau dann invertierbar ist, wenn $\mathrm{rang}(A) = r = n$, folgt "a.⇔b." aus Korollar 5.19.

Lemma 5.22 impliziert, daß für eine Matrix A, für die es ein $B \in \mathrm{Mat}(n, K)$ gibt mit $B \circ A = \mathbb{1}_n$ gilt:

$$n = \mathrm{rang}(\mathbb{1}_n) = \mathrm{rang}(B \circ A) \leq \mathrm{rang}(A) \leq n,$$

also $\operatorname{rang}(A) = n$. Mithin ist eine solche Matrix A invertierbar. Die Umkehrung gilt per definitionem. □

Lemma 5.22
Für Matrizen $A \in \operatorname{Mat}(n \times p, K)$ *und* $B \in \operatorname{Mat}(m \times n, K)$ *gilt:*

$$\operatorname{rang}(B \circ A) \leq \min\{\operatorname{rang}(A), \operatorname{rang}(B)\}.$$

Beweis: Der Beweis sei dem Leser als Übungsaufgabe überlassen. □

Unser Ziel ist es als nächstes, zu zeigen, daß die maximale Anzahl linear unabhängiger Spalten einer Matrix gleich der maximalen Anzahl linear unabhängiger Zeilen ist, und daß beide Zahlen dem Rang der Matrix gleich sind.

Definition 5.23
Es sei $A \in \operatorname{Mat}(m \times n, K)$ mit Spaltenvektoren $a^1, \ldots, a^n$ und Zeilenvektoren $a_1, \ldots, a_m$.

Der von den Zeilen erzeugte Unterraum $\operatorname{ZR}(A) := \langle a_1^t, \ldots, a_m^t \rangle$ von K^n heißt der *Zeilenraum* von A, und der von den Spalten erzeugte Unterraum $\operatorname{SR}(A) := \langle a^1, \ldots, a^n \rangle$ von K^m heißt der *Spaltenraum* von A.

Wir definieren den *Zeilenrang* von A durch $\operatorname{Zrang}(A) := \dim_K\big(\operatorname{ZR}(A)\big)$ und entsprechend den *Spaltenrang* von A durch $\operatorname{Srang}(A) := \dim_K\big(\operatorname{SR}(A)\big)$.

Bemerkung 5.24
Ist A in Zeilen-Stufen-Form, so stimmt die Definition des Zeilenrangs von A mit der Definition in I.9.1 überein.

Da die Spalten einer Matrix A gerade die Zeilen der Transponierten A^t sind und umgekehrt, gilt also $\operatorname{SR}(A) = \operatorname{ZR}\big(A^t\big)$, $\operatorname{ZR}(A) = \operatorname{SR}\big(A^t\big)$, $\operatorname{Srang}(A) = \operatorname{Zrang}\big(A^t\big)$ und $\operatorname{Zrang}(A) = \operatorname{Srang}\big(A^t\big)$.

Wie man unmittelbar aus den Definitionen ablesen kann, gilt zudem

$$\operatorname{SR}(A) = \operatorname{Im}(f_A).$$

Mithin gilt insbesondere:

$$\operatorname{rang}(A) = \operatorname{Srang}(A).$$

Lemma 5.25
Es sei $A \in \operatorname{Mat}(m \times n, K)$*,* $S \in \operatorname{Gl}_m(K)$ *und* $T \in \operatorname{Gl}_n(K)$*. Dann gilt:*

$$\operatorname{Zrang}(S \circ A \circ T) = \operatorname{Zrang}(A) \quad \textit{und} \quad \operatorname{Srang}(S \circ A \circ T) = \operatorname{Srang}(A).$$

Ferner gilt:

$$\operatorname{ZR}(S \circ A) = \operatorname{ZR}(A) \quad \textit{und} \quad \operatorname{SR}(A \circ T) = \operatorname{SR}(A).$$

Beweis: Da für $S \in \operatorname{Gl}_m(K)$ und $T \in \operatorname{Gl}_n(K)$ nach Lemma I.8.23 auch $S^t \in \operatorname{Gl}_m(K)$ und $T^t \in \operatorname{Gl}_n(K)$ gilt, reicht es unter Berücksichtigung von Bemerkung 5.24, die Aussage für den Spaltenrang bzw. den Spaltenraum zu zeigen.

Es gilt $\mathrm{SR}(A) = \mathrm{Im}(f_A)$ und $\mathrm{SR}(S \circ A \circ T) = \mathrm{Im}(f_S \circ f_A \circ f_T)$. Da $f_S : K^m \to K^m$ ein Isomorphismus, also insbesondere injektiv ist, ist auch

$$(f_S)_| : \mathrm{Im}(f_A \circ f_T) \to \mathrm{Im}(f_S \circ f_A \circ f_T) = f_S\big(\mathrm{Im}(f_A \circ f_T)\big)$$

ein Isomorphismus. Mithin gilt nach Korollar 3.12, daß

$$\mathrm{Srang}(S \circ A \circ T) = \dim_K\big(\mathrm{Im}(f_S \circ f_A \circ f_T)\big) = \dim_K\big(\mathrm{Im}(f_A \circ f_T)\big) = \mathrm{Srang}(A \circ T).$$

Es reicht mithin zu zeigen, daß

$$\mathrm{SR}(A) = \mathrm{SR}(A \circ T),$$

oder anders ausgedrückt,

$$\mathrm{Im}(f_A) = \mathrm{Im}(f_A \circ f_T).$$

Sei dazu $y \in \mathrm{Im}(f_A \circ f_T)$, dann gibt es ein $x \in K^n$ mit $y = (f_A \circ f_T)(x) = f_A\big(f_T(x)\big) \in \mathrm{Im}(f_A)$.

Ist umgekehrt $y \in \mathrm{Im}(f_A)$. Dann gibt es ein $x \in K^n$ mit $y = f_A(x)$. Wir setzen $z := \big(f_T\big)^{-1}(x) \in K^n$. Dann gilt:

$$y = f_A(x) = f_A\Big(\big(f_T \circ f_T^{-1}\big)(x)\Big) = (f_A \circ f_T)(z) \in \mathrm{Im}(f_A \circ f_T).$$

□

Satz 5.26
Für $A \in \mathrm{Mat}(m \times n, K)$ *gilt:*

$$\mathrm{rang}(A) = \mathrm{Srang}(A) = \mathrm{Zrang}(A).$$

Beweis: Nach Bemerkung 5.24 wissen wir bereits, daß $\mathrm{rang}(A) = \mathrm{Srang}(A)$ gilt.

Nach Korollar 5.19 gibt es invertierbare Matrizen $S \in \mathrm{Gl}_m(K)$ und $T \in \mathrm{Gl}_n(K)$ mit

$$S \circ A \circ T = \left(\begin{array}{c|c} \mathbb{1}_r & 0 \\ \hline 0 & 0 \end{array}\right),$$

wobei $r = \mathrm{rang}(A)$. Dann gilt aber mit Hilfe von Lemma 5.25

$$\mathrm{rang}(A) = \mathrm{rang}(S \circ A \circ T) = r = \mathrm{Zrang}(S \circ A \circ T) = \mathrm{Zrang}(A).$$

□

Aus Lemma 5.25 ergibt sich nun, daß sich der Rang einer Matrix durch elementare Zeilenoperationen nicht ändert. Wir erhalten also den folgenden Algorithmus zur Bestimmung des Ranges einer Matrix.

Algorithmus 5.27 (zur Bestimmung des Rangs)
INPUT: $A \in \mathrm{Mat}(m \times n, K)$.

OUTPUT: $\mathrm{rang}(A)$

1. **Schritt:** Überführe A in reduzierte ZSF.
2. **Schritt:** Zähle die Anzahl r der Nicht-Nullzeilen in $\mathrm{RZSF}(A)$.

3. Schritt: Gib die r zurück.

Wir notieren nun noch eine einfache, aber nützliche Folgerung.

Korollar 5.28
Sei $A \in \mathrm{Mat}(m \times n, K)$. *Genau dann ist* $\mathrm{rang}(A) = r$, *wenn* A *eine invertierbare* $r \times r$-*Untermatrix*[10] *besitzt, aber keine invertierbare* $s \times s$-*Untermatrix für* $s > r$.

Beweis: $\mathrm{rang}(A)$ ist gleichzeitig die Maximalzahl linear unabhängiger Zeilen und die Maximalzahl linear unabhängiger Spalten.

Ist $\mathrm{rang}(A) = r$, so wählen wir r linear unabhängige Zeilen aus. Die entstehende Matrix hat Zeilenrang r und somit auch Spaltenrang r. Von der Untermatrix können wir also r linear unabhängige Spalten auswählen. Die resultierende $(r \times r)$-Matrix hat Rang r und ist somit gemäß Korollar 5.21 invertierbar. Es ist zudem klar, daß r maximal mit dieser Eigenschaft ist, denn sonst hätte A mehr linear unabhängige Zeilen.

Für jede invertierbare $(s \times s)$-Untermatrix von A sind die zugehörigen Zeilen linear unabhängig, also folgt auch die Umkehrung. □

Bislang sind wir dem Leser den Beweis der Eindeutigkeit der reduzierten Zeilen-Stufen-Form einer Matrix $A \in \mathrm{Mat}(m \times n, K)$ schuldig geblieben. Nun sind wir in der Lage, dieses Versäumnis nachzuholen.

Beweis der Eindeutigkeit in Satz I.9.4: Es sei also $A \in \mathrm{Mat}(m \times n, K)$ eine $m \times n$-Matrix.

Da elementare Zeilenoperationen durch Multiplikation mit Elementar- bzw. Permutationsmatrizen von links realisiert werden, gilt für eine ZSF B von A, daß es eine invertierbare Matrix $S \in \mathrm{Gl}_m(K)$ gibt mit $B = S \circ A$ (vgl. auch Korollar I.9.6). Mit Lemma 5.25 folgt dann $\mathrm{ZR}(A) = \mathrm{ZR}(B)$, insbesondere gilt mit Korollar 3.8 also, daß die Zeilen von B eine Basis von $\mathrm{ZR}(A)$ bilden, da

$$r := \dim_K \big(\mathrm{ZR}(A)\big) = \mathrm{Zrang}(A) = \mathrm{Zrang}(B). \tag{12}$$

Seien nun $B = (b_{ij})$ und $B' = \big(b'_{ij}\big)$ zwei reduzierte ZSF von A mit Zeilenvektoren $b_1, \ldots, b_m$ bzw. $b'_1, \ldots, b'_m$ und Pivotspalten $\{j_1, \ldots, j_r\}$ bzw. $\{j'_1, \ldots, j'_r\}$ - beachte, daß die Anzahl $r = \mathrm{rang}(A)$ nach (12) für beide gleich ist. Wir zeigen nun per Induktion, daß die Zeilen der Matrizen B und B' übereinstimmen.

Induktionsbehauptung: Für $i \in \mathbb{N}$ gilt entweder $i \geq r$ oder $b_{r-i} = b'_{r-i}$, insbesondere also $j_{r-i} = j'_{r-i}$.

Induktionsanfang: $i = 0$. O. E. gelte $j_r \geq j'_r$. Da $b_r \in \mathrm{ZR}(A) = \langle b'_1, \ldots, b'_r \rangle$, gibt es $\lambda_1, \ldots, \lambda_r \in K$ mit

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

[10] Dabei nennt man eine Matrix, die durch Streichen von Zeilen und Spalten aus A entsteht, eine *Untermatrix* von A.

Insbesondere gilt für $i = 1, \dots, r-1$

$$0 = b_{rj'_i} = \lambda_i \quad \text{und} \quad b_{rj'_r} = \lambda_r,$$

nach (iv) und (v) in Definition I.9.1 angewandt auf die reduzierte ZSF B' mit Pivotspalten $j'_1, \dots, j'_r$ sowie (i) angewandt auf die ZSF B. Also folgt $b_r = \lambda_r \cdot b'_r$. Da $b_r \neq 0$, muß $\lambda_r \neq 0$ gelten und somit $j'_r = j_r$ wegen (i) in I.9.1. Aber dann gilt nach (iv) in I.9.1 $1 = b_{rj_r} = \lambda_r$ und somit $b_r = b'_r$.

Induktionsschritt: $0 < i < r-1$ und die Behauptung gelte schon für $0, \dots, i-1$. O. E. gelte $j_{r-i} \geq j'_{r-i}$. Nach Induktionsvoraussetzung gilt nun $b_{r-i} \in ZR(A) = \langle b'_1, \dots, b'_{r-i}, b_{r-i+1}, \dots, b_r \rangle$ also gibt es $\lambda_1, \dots, \lambda_r \in K$ mit

$$b_{r-i} = \sum_{k=1}^{r-i} \lambda_k b'_k + \sum_{k=r-i+1}^{r} \lambda_k b_k.$$

Insbesondere gilt nach (v) in Definition I.9.1, angewandt auf die reduzierte ZSF B, für $k = r-i+1, \dots, r$

$$0 = b_{r-i\, j_k} = \lambda_k,$$

da $r-i < k$, und (i) angewandt auf B sowie (v) auf B' liefert für $k = 1, \dots, r-i-1$

$$0 = b_{r-i\, j'_k} = \lambda_k,$$

da $j'_k < j'_{r-i} \leq j_{r-i}$. Insgesamt erhalten wir also wieder

$$b_{r-i} = \lambda_{r-i} b'_{r-i}. \tag{13}$$

Wäre $j_{r-i} > j'_{r-i}$, dann wäre wieder mit (i) $0 = b_{r-i\, j'_{r-i}} = \lambda_{r-i}$ im Widerspruch zu (13) und $b_{r-i} \neq 0$. Also ist $j_{r-i} = j'_{r-i}$ und dann folgt mit (iv) aus I.9.1, daß $\lambda_{r-i} = b_{r-i\, j_{r-i}} = 1$, und damit aus (13) $b_{r-i} = b'_{r-i}$.

Also haben wir mit Induktion gezeigt, daß die Zeilen von B und B' übereinstimmen, d. h. daß die reduzierte Zeilenstufenform von A eindeutig bestimmt ist. □

Nachdem die Eindeutigkeit der reduzierten Zeilen-Stufen-Form nun gezeigt ist, erlauben wir uns aus Satz I.9.4 und Korollar I.9.6 einige Folgerungen zu ziehen.

Korollar 5.29

Es sei $A \in \mathrm{Mat}(n, K)$. *Dann sind gleichwertig:*

a. A *ist invertierbar.*
b. $\mathrm{RZSF}(A) = \mathbb{1}_n$.
c. *Es gibt Elementarmatrizen* $T_1, \dots, T_k \in \mathrm{Mat}(n, K)$ *mit:*

$$T_k \circ \dots \circ T_1 \circ A = \mathbb{1}_n.$$

d. *Es gibt Elementarmatrizen* $T'_1, \dots, T'_k \in \mathrm{Mat}(n, K)$ *mit:*

$$A = T'_1 \circ \dots \circ T'_k.$$

Beweis: Nach Korollar 5.21 gilt, daß A genau dann invertierbar ist, wenn $\operatorname{rang}(A) = n$. Also folgt die Äquivalenz von a.-d. aus Korollar I.9.6 unter Berücksichtigung von Proposition I.8.21. □

Bemerkung 5.30
Ist $(G, \cdot)$ eine Gruppe und $M \subseteq G$ eine Teilmenge von G, dann sagen wir, M *erzeugt* G oder G ist das *Erzeugnis* von M, wenn gilt:

$$G = \left\{ x_1^{\epsilon_1} \cdots x_n^{\epsilon_n} \mid n \in \mathbb{N}, x_1, \ldots, x_n \in M, \epsilon_1, \ldots, \epsilon_n \in \{-1, 1\} \right\}.$$

In diesem Sinne folgt aus Korollar 5.29, daß $\mathrm{Gl}_n(K)$ von den Elementarmatrizen erzeugt wird.

Aus Korollar 5.29 leitet sich folgendes Verfahren zur Bestimmung der Inversen einer invertierbaren Matrix ab. Hierzu beachte man, daß für Elementarmatrizen $T_1, \ldots, T_k$, für die gilt, daß $T_k \circ \ldots \circ T_1 \circ A = \mathbb{1}_n$, ferner gilt, daß

$$T_k \circ \ldots \circ T_1 \circ (A, \mathbb{1}_n) = (\mathbb{1}_n, T_k \circ \ldots \circ T_1) = \left(\mathbb{1}_n, A^{-1}\right).$$

Algorithmus 5.31 (zur Bestimmung der Inversen)
INPUT: $A \in \mathrm{Mat}(n, K)$.

OUTPUT: Inverse von A, so sie existiert, eine Fehlermeldung sonst.

1. **Schritt:** Erweitere die Matrix A um $\mathbb{1}_n$ zur Matrix $C = (A, \mathbb{1}_n) \in \mathrm{Mat}(n \times 2n, K)$.
2. **Schritt:** Überführe die C in reduzierte ZSF $C' = \left(A', B\right)$.
3. **Schritt:** Falls $\operatorname{rang}\left(A'\right) = n$, dann gib B zurück, sonst gib eine Fehlermeldung zurück.

Beispiel 5.32
Wir betrachten die 3×3-Matrix

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \in \mathrm{Mat}(3 \times 3, K)$$

und versuchen die Inverse mittels des Algorithmus 5.31 zu bestimmen.

	A			$\mathbb{1}_n$		
1	1	1	1	0	0	
0	1	1	0	1	0	
1	0	1	0	0	1	$III \mapsto III - I$
1	1	1	1	0	0	
0	1	1	0	1	0	
0	−1	0	−1	0	1	$III \mapsto III + II$
1	1	1	1	0	0	$I \mapsto I - III$
0	1	1	0	1	0	$II \mapsto II - III$
0	0	1	−1	1	1	
1	1	0	2	−1	−1	$I \mapsto I - II$
0	1	0	1	0	−1	
0	0	1	−1	1	1	
1	0	0	1	−1	0	
0	1	0	1	0	−1	
0	0	1	−1	1	1	

Hieraus ergibt sich gemäß obigem Algorithmus zunächst, daß A invertierbar ist, und ferner, daß

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

Korollar 5.33
Sei $A \in \mathrm{Mat}(m \times n, K)$ *mit* $r = \mathrm{rang}(A)$*, so läßt sich* A *durch endlich viele elementare Zeilen- und Spaltenoperationen auf die folgende Form bringen:*

$$\left(\begin{array}{c|c} \mathbb{1}_r & 0 \\ \hline 0 & 0 \end{array}\right). \tag{14}$$

Beweis: Die Aussage folgt aus Korollar 5.19 und Korollar 5.29, da elementare Operationen nach Lemma I.8.18 und Bemerkung I.8.19 durch Multiplikation mit Elementarmatrizen realisierbar sind. □

Wir wollen nun noch an einem Beispiel zeigen, wie man eine Matrix mittels des gaußschen Verfahrens auf Normalform (14) bringt.
Beispiel 5.34
Für $\lambda \in K$ sei

$$A_\lambda := \begin{pmatrix} 1 & \lambda & 0 & 0 \\ \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \end{pmatrix}.$$

Durch elementare Zeilen und Spaltenoperationen überführt man A_λ in Normalform:

$$\begin{pmatrix} 1 & \lambda & 0 & 0 \\ \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \end{pmatrix} \xrightarrow[\text{operationen}]{\text{Zeilen-}} \begin{pmatrix} 1 & \lambda & 0 & 0 \\ 0 & 1-\lambda^2 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \end{pmatrix} \xrightarrow[\text{operationen}]{\text{Spalten-}}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1-\lambda^2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow[\text{operationen}]{\text{Zeilen-}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1-\lambda^2 \end{pmatrix}.$$

Insbesondere gilt, $\operatorname{rang}(A) = 4$ für $\lambda^2 \neq 1$ und $\operatorname{rang}(A) = 3$ sonst.

Algorithmus 5.35 (Normalform-Algorithmus)
INPUT: $A \in \mathrm{Mat}(m \times n, K)$.

OUTPUT: Normalform von A bezüglich Äquivalenz

1. **Schritt:** Überführe A in reduzierte ZSF und transponiere diese.
2. **Schritt:** Überführe die Transponierte erneut in reduzierte ZSF und transponiere das Ergebnis.
3. **Schritt:** Gib die (veränderte) Matrix A zurück.

AUFGABEN

Aufgabe 5.36
Bestimme die Matrixdarstellung $M_{B'}^{B}(f)$ der $\mathbb{R}$-linearen Abbildung $f : \mathbb{R}^4 \to \mathbb{R}^3 : (x_1, x_2, x_3, x_4)^t \mapsto (x_1 - x_2, x_2 + x_4, x_1)^t$ bezüglich der Basen $B = \left((1,0,0,-1)^t, (0,1,0,-1)^t, (0,0,1,-1)^t, (0,0,0,1)^t\right)$ von $\mathbb{R}^4$ und $B' = \left((1,0,0)^t, (0,1,0)^t, (1,1,1)^t\right)$ von $\mathbb{R}^3$.

Aufgabe 5.37
Es sei $B = \left((1,1,1,1)^t, (-1,0,0,1)^t, (0,-1,0,1)^t, (0,0,-1,1)^t\right)$ und $D = \left((1,1,0)^t, (0,1,1)^t, (0,0,1)^t\right)$.

a. Zeige, B ist eine Basis des $\mathbb{R}^4$ und D eine Basis des $\mathbb{R}^3$.
b. Für $f \in \mathrm{Hom}_{\mathbb{R}}\left(\mathbb{R}^4, \mathbb{R}^3\right)$ mit $f(x_1, x_2, x_3, x_4) = (x_1 - x_2, x_3, x_2 + x_4)^t$ bestimme $M_D^B(f)$.
c. Bestimme umgekehrt die Funktionsvorschrift für $g \in \mathrm{Hom}_{\mathbb{R}}\left(\mathbb{R}^4, \mathbb{R}^3\right)$ mit

$$M_D^B(g) = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 2 & 1 & 3 & 1 \\ 0 & -1 & 2 & 0 \end{pmatrix}.$$

Aufgabe 5.38
Betrachte die Basen $B_1 = \left((1,1)^t, (0,1)^t\right)$, $B_2 = \left((2,-1)^t, (2,3)^t\right)$, $B_3 =$

$\left((4,4)^t,(1,2)^t\right)$ und $B_4 = \left((-1,1)^t,(0,2)^t\right)$ des $\mathbb{R}^2$, sowie die $\mathbb{R}$-lineare Abbildung $f : \mathbb{R}^2 \to \mathbb{R}^2$ mit Matrix-Darstellung

$$M_{B_2}^{B_1}(f) = \begin{pmatrix} 1 & 3 \\ -2 & -6 \end{pmatrix}.$$

a. Ermittle die Transformationsmatrizen $T_{B_4}^{B_2}$ und $T_{B_1}^{B_3}$ mit $T_{B_4}^{B_2} \circ M_{B_2}^{B_1}(f) \circ T_{B_1}^{B_3} = M_{B_4}^{B_3}(f)$.
b. Ermittle Basen B_5 und B_6 von $\mathbb{R}^2$ mit

$$M_{B_6}^{B_5}(f) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Aufgabe 5.39
Es sei $B = (x_1, x_2, x_3)$ eine Basis des K-Vektorraumes V, $D = (y_1, y_2)$ eine Basis des K-Vektorraumes W und $f \in \mathrm{Hom}_K(V, W =$, so daß $M_D^B(f) = (a_{i,j})_{i=1,2;j=1,2,3}$ die Matrix-Darstellung von f bezüglich (B, D) ist. Ferner setzen wir:

$$\begin{aligned} x_1' &= x_1 + x_2 + x_3 & \qquad y_1' &= y_2 \\ x_2' &= x_2 + x_3 & \qquad y_1' &= y_1 - y_2 \\ x_3' &= x_3 \end{aligned}$$

und $B' = (x_1, x_2, x_3)$ sowie $D' = (y_1', y_2')$.

a. Zeige, B' ist eine Basis von V und D' ist eine Basis von W.
b. Bestimme, die Matrix $M_{D'}^{B'}(f)$, d. h. die Matrix-Darstellung von f bezüglich der Basen $\left(B', D'\right)$.

Aufgabe 5.40
Es sei V ein $\mathbb{R}$-Vektorraum, $B = (x_1, x_2, x_3)$ eine Basis von V und $B' = (y_1, y_2, y_3)$ mit $y_1 = x_1 + x_3$, $y_2 = x_1 + x_2$ und $y_3 = x_1 + x_2 + x_3$.

a. Zeige, B' ist eine Basis von V.
b. Bestimme $M_{B'}^{B'}(f)$, wobei $f \in \mathrm{End}_{\mathbb{R}}(V)$ gegeben ist durch

$$M_B^B(f) = \begin{pmatrix} a & 0 & b \\ -b & a & a \\ a & b & b \end{pmatrix}, \text{ mit } a, b \in \mathbb{R}.$$

Aufgabe 5.41
Es sei $V \neq \{0\}$ ein n-dimensionaler K-Vektorraum, und $f \in \mathrm{End}_K(V)$. Zeige, die beiden folgenden Aussagen sind gleichwertig:

a. Für je zwei Basen B und B' von V gilt:

$$M_B^B(f) = M_{B'}^{B'}(f).$$

b. Es gibt ein $\lambda \in K$ mit $f = \lambda \cdot \mathrm{id}$.

Aufgabe 5.42
Es sei $A \in \mathrm{Mat}(n, K)$ mit $AB = BA$ für alle $B \in \mathrm{Mat}(n \times n, K)$. Zeige, es gibt ein $\lambda \in K$ mit $A = \lambda \cdot \mathbb{1}_n$.

Aufgabe 5.43

Es sei K ein Körper. Wir definieren die *Spur* einer quadratischen Matrix $A = \left(a_{ij}\right)_{i,j} \in \mathrm{Mat}(n \times n, K)$ als Summe der Diagonalelemente von A, d. h.

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

Sei nun V ein n-dimensionaler K-Vektorraum, E eine fest vorgegebene Basis von V und $f \in \mathrm{End}_K(V)$. Dann definieren wir

$$\mathrm{Spur}(f) := \mathrm{Spur}\left(M_E^E(f)\right).$$

Zeige die folgenden Aussagen:

a. $\mathrm{Spur}(AB) = \mathrm{Spur}(BA)$ für $A, B \in \mathrm{Mat}(n, K)$.
b. $\mathrm{Spur}(A) = \mathrm{Spur}(B^{-1}AB)$ für $A \in \mathrm{Mat}(n, K)$ und $B \in \mathrm{Gl}_n(K)$.
c. $\mathrm{Spur}\left(M_E^E(f)\right) = \mathrm{Spur}\left(M_F^F(f)\right)$ für eine beliebige Basis F von V. Insbesondere ist die Definition von $\mathrm{Spur}(f)$ unabhängig von der gewählten Basis.
d. Ist $(v_1, \ldots, v_l)$ eine Basis von $\mathrm{Ker}(f)$ und $G' = \left(v_{l+1} + \mathrm{Ker}(f), \ldots, v_n + \mathrm{Ker}(f)\right)$ eine Basis von $V/\mathrm{Ker}(f)$, so gilt für die Basis $G = (v_1, \ldots, v_n)$ von V:
$$M_G^G(f) = \left(\begin{array}{c|c} 0 & * \\ \hline 0 & M_{G'}^{G'}\left(f_{V/\mathrm{Ker}(f)}\right) \end{array}\right).$$
e. Gibt es ein $r \in \mathbb{N}$ mit $f^r = 0 \in \mathrm{End}_K(V)$, so gilt $\mathrm{Spur}(f) = 0$.
f. Gibt es ein $r \in \mathbb{N}$ mit $A^r = 0$, so gilt $\mathrm{Spur}(A) = 0$.

Hinweis zum Beweis von e.: Induktion über n. Zeige, zunächst $\mathrm{Ker}(f) \neq \{0\}$ und wende Induktion auf $f_{V/\mathrm{Ker}(f)}$ an, (vgl. Aufgabe 4.25). Für Teil f. beachte dann, daß $(f_A)^r = f_{A^r}$.

Aufgabe 5.44

Es sei V ein K-Vektorraum, $f \in \mathrm{End}_K(V)$ und $U \subseteq V$ ein f-invarianter Unterraum. Ferner sei nun $B' = (v_1, \ldots, v_r)$ eine Basis von U und $B = (v_1, \ldots, v_r, v_{r+1}, \ldots, v_n)$ eine Basis von V. Wir wissen bereits, daß dann $B'' = (v_{r+1} + U, \ldots, v_n + U)$ eine Basis von V/U ist.

Zeige, es gilt

$$M_B^B(f) = \left(\begin{array}{c|c} M_{B'}^{B'}(f_U) & C \\ \hline 0 & M_{B''}^{B''}\left(f_{V/U}\right) \end{array}\right),$$

wobei $0 \in \mathrm{Mat}\left((n-r) \times r, K\right)$ die Nullmatrix ist und $C \in \mathrm{Mat}\left(r \times (n-r), K\right)$ geeignet.

Aufgabe 5.45

Es sei V ein endlich-dimensionaler K-Vektorraum, $f \in \mathrm{End}_K(V)$ und es gelte $V = U_1 \oplus \ldots \oplus U_r$ für f-invariante Unterräume U_i, $i = 1, \ldots, r$, mit $\dim_K(U_i) =$

n_i. Zeige, es gibt eine Basis B von V mit

$$M_B^B(f) = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ \hline 0 & \boxed{A_2} & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & \boxed{A_r} \end{pmatrix}, \text{ wobei } A_i \in \mathrm{Mat}(n_i, K).$$

Aufgabe 5.46
U, V und W seien K-Vektorräume, $f \in \mathrm{Hom}_K(U, V)$, $g \in \mathrm{Hom}_K(V, W)$. Zeige:

a. $\mathrm{rang}(g \circ f) \leq \min\{\mathrm{rang}(f), \mathrm{rang}(g)\}$.
b. Ist f surjektiv, so gilt $\mathrm{rang}(g \circ f) = \mathrm{rang}(g)$.
c. Ist g injektiv, so gilt $\mathrm{rang}(g \circ f) = \mathrm{rang}(f)$.

Aufgabe 5.47
Es sei V ein endlich-dimensionaler K-Vektorraum, $f \in \mathrm{End}_K(V)$ und $U \subseteq V$ ein f-invarianter Unterraum von V. Zeige, dann gilt

$$\mathrm{rang}\left(f_{V/U}\right) = \dim_K\left(\mathrm{Im}(f)\right) - \dim_K\left(U \cap \mathrm{Im}(f)\right) \leq \mathrm{rang}(f) - \mathrm{rang}\left(f_U\right).$$

Aufgabe 5.48
Berechne durch elementare Umformungen den Rang von

$$\begin{pmatrix} 0 & b & b & b \\ a & 0 & b & b \\ a & a & 0 & b \end{pmatrix}.$$

Aufgabe 5.49
Schreibe eine Singular-Prozedur `rang`, die eine Matrix einliest und ihren Rang ausgibt.

Aufgabe 5.50
Bestimme die Inverse der Matrix

$$\begin{pmatrix} 1 & 3 & -1 & 4 \\ 2 & 5 & -1 & 3 \\ 0 & 4 & -3 & 1 \\ -3 & 1 & -5 & -2 \end{pmatrix}.$$

Aufgabe 5.51
Bestimme, sofern sie existiert, die Inverse der Matrix

$$A = \begin{pmatrix} 2 & -3 & 2 & 5 \\ 1 & -1 & 1 & 2 \\ -3 & 2 & 2 & 1 \\ 1 & 1 & -3 & -5 \end{pmatrix}.$$

Aufgabe 5.52
Zeige, die folgende Matrix $A \in \mathrm{Mat}(4, \mathbb{R})$ ist invertierbar und berechne die

Inverse,

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

Aufgabe 5.53
Schreibe eine Prozedur `invers`, die eine Matrix A einliest und, falls A eine invertierbare Matrix ist, die Inverse von A zurück gibt, ansonsten den String ‘‘`Die Matrix ist nicht invertierbar.`’’.

Die Prozeduren `unitmat` und `concat` aus der Bibliothek `matrix.lib` dürfen verwendet werden.

Aufgabe 5.54
Finde reguläre Matrizen X, Y mit $XAY = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$, wo

$$A = \begin{pmatrix} 1 & -2 & 3 & 0 \\ 2 & -7 & 10 & -1 \\ -2 & 4 & -7 & 2 \\ 3 & -5 & 7 & 1 \end{pmatrix} \text{ und } r = \operatorname{rang}(A).$$

Aufgabe 5.55
Schreibe eine Prozedur `normalform`, die eine Matrix A einliest und die Normalform von A bezüglich Äquivalenz ausgibt.

6. Einige Algorithmen - Teil I

Wir haben, basierend auf dem Gauß-Algorithmus I.9.8 zur Berechnung der reduzierten Zeilen-Stufen-Form einer Matrix, bereits am Ende von Abschnitt 5 Algorithmen zur Berechnung des Ranges (5.27), der Normalform (bezüglich Äquivalenz - 5.35) und ggf. der Inversen (5.31) einer Matrix entwickelt. Wir wollen in den beiden Abschnitten 6 und 8 eine Reihe weiterer Algorithmen angeben, mit deren Hilfe man Objekte, die in den vorhergehenden Abschnitten eingeführt wurden, auch rechnerisch bestimmen kann. Schließlich folgen in den Abschnitten 7 und 9 dann nochmals Algorithmen zur Berechnung der Parametrisierung eines linearen Gleichungssystems (7.16) und zur Berechnung der Determinante einer quadratischen Matrix (9.18). Sie alle beruhen ebenfalls auf dem Gauß-Algorithmus.

Generalvoraussetzung: In diesem Abschnitt sei K ein Körper.

Zunächst wollen wir das Austauschverfahren von Steinitz algorithmisch formulieren.

Algorithmus 6.1 (Austauschverfahren von Steinitz)
INPUT: Eine Basis $B = (x_1, \ldots, x_n)$ und eine linear unabhängige Familie $F = (y_1, \ldots, y_r)$ von Vektoren in $V = \langle B \rangle \subseteq K^n$.

OUTPUT: Eine Basis B' von V, die F enthält.

1. **Schritt:** Schreibe die Vektoren in B als Spalten in eine Matrix A.
2. **Schritt:** Für $i = 1, \ldots, r$ tue:
   - Bilde eine neue Matrix A' aus A und y_i (als letzter Spalte).
   - Überführe A' in reduzierte Zeilen-Stufen-Form und suche in der letzten Spalte den ersten Eintrag ungleich Null.
   - Streiche die entsprechende Spalte aus A und hänge statt dessen y_i als Spalte an A an.
3. **Schritt:** Gib die Spalten der veränderten Matrix A zurück.

Mit dem folgenden Algorithmus kann man sicherstellen, daß die Eingabedaten beim Steinitz-Algorithmus auch den gestellten Anforderungen genügen.

Algorithmus 6.2

INPUT: Zwei Familien $B = (x_1, \ldots, x_m)$ und F von Vektoren in K^n.

OUTPUT: Eins, falls B eine Basis ist und $F \subseteq \langle B \rangle$, Null sonst.

1. **Schritt:** Schreibe die Vektoren von B als Spalten in eine Matrix A und hänge die Vektoren aus F als Spalten an.
2. **Schritt:** Überführe die Matrix A in reduzierte ZSF. Gilt $\operatorname{rang}(A) = m$ und sind die ersten m Spalten Pivotspalten, so gib Eins zurück. Sonst gib Null zurück.

Der folgende Algorithmus zur Bestimmung einer Basis aus gegebenem Erzeugendensystem beruht auf der Tatsache, daß elementare Zeilenoperationen den Zeilenraum nicht verändern - vgl. Lemma 5.25.

Algorithmus 6.3 (Basisberechnung)

INPUT: Ein Erzeugendensystem F des Unterraums $U \subseteq K^n$.

OUTPUT: Eine Basis von U.

1. **Schritt:** Schreibe die Vektoren von F als Zeilen in eine Matrix A und überführe A in reduzierte Zeilen-Stufen-Form.
2. **Schritt:** Gib die ersten $\operatorname{rang}(A)$ Zeilen als Vektoren zurück.

Hieraus ergibt sich unmittelbar ein Algorithmus zur Bestimmung der Basis der Summe zweier Unterräume des K^n.

Algorithmus 6.4 (Summe zweier Unterräume)

INPUT: Erzeugendensysteme F und G von zwei Unterräumen U und U' des K^n.

OUTPUT: Eine Basis von $U + U'$.

1. **Schritt:** Bilde aus F und G ein Erzeugendensystem und berechne mittels 6.3 eine Basis von $U + U' = \langle F \cup G \rangle$.
2. **Schritt:** Gib diese Basis zurück.

Da eine endliche Familie von Vektoren genau dann linear unabhängig ist, wenn sie eine Basis ihres Erzeugnisses ist, und da die Dimension des Erzeugnisses einer solchen Familie gerade der Rang der Matrix ist (vgl. Satz 5.26),

deren Spalten die Erzeuger sind, liefert Korollar 3.8 den folgenden Algorithmus.

Algorithmus 6.5 (Test auf lineare Unabhängigkeit)
INPUT: Eine Familie $F = (x_1, \ldots, x_m)$ von Vektoren in K^n.

OUTPUT: Eins, falls F linear unabhängig ist, Null sonst.

1. **Schritt:** Ist F leer, gib Eins zurück, sonst schreibe die Vektoren in F als Spalten in eine Matrix A.
2. **Schritt:** Ist $\operatorname{rang}(A) = m$, so gib Eins zurück, sonst Null.

Ist $f = f_A$ für eine $m \times n$-Matrix A, dann wird das Bild von f von den Spalten von A erzeugt. Wir können eine Basis des Bildes also wie folgt bestimmen.

Algorithmus 6.6 (Bild von f_A)
INPUT: $A \in \operatorname{Mat}(m \times n, K)$.

OUTPUT: Eine Basis von $\operatorname{Im}(f_A)$.

1. **Schritt:** Transponiere A und überführe die Transponierte in reduzierte ZSF.
2. **Schritt:** Transponiere das Ergebnis wieder und gib die ersten $\operatorname{rang}(A)$ Spaltenvektoren zurück.

Für einige Algorithmen, etwa zur Bestimmung des Kerns einer linearen Abbildung, benötigen wir zunächst einen Algorithmus zur Berechnung einer Parametrisierung der Lösungsmenge eines linearen Gleichungssystems, was uns zum folgenden Abschnitt 7 führt.

AUFGABEN

Aufgabe 6.7
Schreibe eine Singular-Prozedur `basis`, die einen Modul `M` einliest und eine Liste `L=e,r,B` mit drei Einträgen ausgibt. `e` soll vom Typ `int` sein und den Wert `0` annehmen, wenn die Erzeuger von `M` linear abhängig sind, und `1` andernfalls. `r`, ebenfalls vom Typ `int`, soll die Dimension von `M` wiedergeben, und `B`, vom Typ `module`, soll eine Basis von `M` sein. – Ist der gegebene Modul der Null-Modul, so soll als Basis wieder der Nullmodul zurückgegeben werden.[11]

Hinweis: Ein Modul in Singular ist eine Datenstruktur, die aus einer Familie von Vektoren besteht. Mathematisch heißt das, ein Modul `M` wird dadurch gegeben, daß man Singular ein Erzeugendensystem von `M` vorgibt. Für mehr Informationen sei auf das Singular-Manual (`help module;`) verwiesen. – Ist `M` vom Typ `module`, so wird mittels `A=matrix(M);` eine Matrix `A` erzeugt, deren Spalten die Erzeuger von `M` sind. Analog kann man mit `M=module(A);` eine Matrix in einen Modul konvertieren. – Die Befehle `submat` (siehe `matrix.lib`) und `transpose` sind für die Prozedur hilfreich.

Aufgabe 6.8
Schreibe in analoger Weise Singular-Prozeduren, die die übrigen Algorithmen umsetzen.

[11] Das ist zwar mathematisch nicht korrekt, aber in Singular ist das Erzeugnis der leeren Menge nicht verfügbar.

Aufgabe 6.9
Es sei E die kanonische Basis des $\mathbb{R}^5$. Betrachte $f \in \mathrm{End}_{\mathbb{R}}\left(\mathbb{R}^5\right)$ mit

$$M_E^E(f) = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 1 & 1 \\ -1 & -1 & 0 & 1 & 1 \\ -1 & -1 & -1 & 0 & 1 \\ -1 & -1 & -1 & -1 & 0 \end{pmatrix}.$$

a. Bestimme eine Basis von $\mathrm{Ker}(f)$.
b. Bestimme eine Basis von $\mathrm{Im}(f)$.

7. Lineare Gleichungssysteme

Für nähere Ausführungen zur Bedeutung von Gleichungssystemen für Anwendungen verweisen wir auf die Einleitung des Skriptes.
Generalvoraussetzung: In diesem Abschnitt sei R stets ein kommutativer Ring mit Eins und K ein Körper. Ferner bezeichne $(e_1, \ldots, e_n)$ die kanonische Basis von K^n und $(f_1, \ldots, f_m)$ die kanonische Basis von K^m.
Definition 7.1 a. Ein *lineares Gleichungssystem* über R

$$\text{(LGS)} \qquad \begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & = & b_2 \\ & & & & \vdots & & & & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn}x_n & = & b_m \end{array}$$

besteht aus m Gleichungen in n *Unbestimmten* oder *Variablen* $x_1, \ldots, x_n$ mit $a_{ij}, b_i \in R$ für $1 \leq i \leq m$ und $1 \leq j \leq n$.
Da sich (LGS) mittels Matrixmultiplikation auch kurz schreiben läßt als

$$Ax = b,$$

sprechen wir meist von dem *linearen Gleichungssystem* $Ax = b$.
b. Die Matrix

$$A = (a_{ij}) = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \in \mathrm{Mat}(m \times n, R)$$

heißt *Koeffizientenmatrix* und der Vektor $b = (b_1, \ldots, b_m)^t \in R^m$ die *Inhomogenität* des Gleichungssystems (LGS). Ferner heißt die Matrix

$$\widetilde{A} := (A, b) := \left(\begin{array}{ccc|c} a_{11} & \ldots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \ldots & a_{mn} & b_m \end{array}\right) \in \mathrm{Mat}\left(m \times (n+1), R\right)$$

die *erweiterte Koeffizientenmatrix* von (LGS).
c. Das lineare Gleichungssystem (LGS) heißt *homogen*, falls $b = 0_{R^m}$ der Nullvektor in R^m ist. Ansonsten heißt das System *inhomogen*.

d. Ist ein lineares Gleichungssystem $Ax = b$ gegeben, so heißt das Gleichungssystem $Ax = 0$ (mit $0 = 0_{R^m}$) das *zugehörige homogenisierte Gleichungssystem*.

e. Ein Vektor $c = (c_1, \ldots, c_n)^t \in R^n$ heißt *Lösung* von (LGS), wenn die Gleichung $Ac = b$ erfüllt ist. Die Menge aller Lösungen von (LGS) wird mit

$$\text{Lös}(A, b) := \{c \in \mathbb{R}^n \mid Ac = b\}.$$

bezeichnet.

f. Sei $A \in \text{Mat}(m \times n, R)$ und seien $y_1, \ldots, y_m$ Unbestimmte, $y = (y_1, \ldots, y_m)^t$. Falls für jedes $b \in R^m$ das lineare Gleichungssystem $Ax = b$ eine Lösung besitzt, sagt man, $Ax = y$ ist *universell lösbar*.
Ist die Lösung zudem für jedes b eindeutig bestimmt, so sagt man $Ax = y$ ist *universell eindeutig lösbar*.

Bei einem linearen Gleichungssystem sind also Ringelemente a_{ij} und b_i fest vorgegeben, während für die Unbestimmten x_j Ringelemente c_j gesucht werden, die das Gleichungssystem lösen.

Falls $R = \mathbb{R}$, so kann ein lineares Gleichungssystem entweder gar keine, genau eine oder unendlich viele Lösungen haben. Wir werden im Folgenden mehrere Verfahren zur Lösung kennenlernen[12] und uns, im Fall von mehr als einer Lösung, mit der Struktur der Lösungsmenge $\text{Lös}(A, b)$ beschäftigen. Eine wichtige Rolle spielt dabei die lineare Abbildung $f_A : R^n \to R^m$.

Bemerkung 7.2
Es sei $A \in \text{Mat}(m \times n, R)$ und $b \in R^m$. Dann folgt unmittelbar aus den Definitionen:

$$\text{Lös}(A, 0) = \text{Ker}(f_A).$$

Insbesondere, $\text{Lös}(A, 0)$ ist ein R-Modul.

Ebenfalls anhand der Definitionen sieht man, daß das lineare Gleichungssystem $Ax = b$ genau dann eine Lösung besitzt, wenn $b \in \text{Im}(f_A)$.

Damit gilt dann aber, daß $Ax = y$ genau dann universell lösbar ist, wenn f_A surjektiv ist.

Und ferner, daß $Ax = y$ genau dann universell eindeutig lösbar ist, wenn f_A bijektiv ist.

Beispiel 7.3
Das lineare Gleichungssystem

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 1 \\ 2x_1 + 3x_2 &= 1 \\ x_2 - x_3 &= 0 \end{aligned}$$

[12] Siehe Gauß-Algorithmus 7.16 und Cramersche Regel 10.12.

ist inhomogen, hat als Koeffizientenmatrix

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 3 & 0 \\ 0 & 1 & -1 \end{pmatrix} \in \mathbf{Mat}(3 \times 3, \mathbb{R}),$$

und als erweiterte Matrix

$$\widetilde{A} = \left(\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 2 & 3 & 0 & 1 \\ 0 & 1 & -1 & 0 \end{array}\right) \in \mathbf{Mat}(3 \times 4, \mathbb{R}).$$

Die Lösung ist in diesem Fall ganz einfach. Wir erhalten $x_3 = x_2$ aus der 3. Gleichung, $3x_2 = 1 - 2x_1$ aus der 2. und, wenn wir das in die erste Gleichung einsetzen, $x_1 + (1 - 2x_1) = 1$, also $x_1 = 0$. Einsetzen von $x_1 = 0$ in die 2. und 3. Gleichung liefert, daß $\left(0, \frac{1}{3}, \frac{1}{3}\right)^t$ die einzige Lösung ist.

Satz 7.4

Seien $A \in \mathbf{Mat}(m \times n, K)$, $b \in K^m$ *und sei* $x^0 \in K^n$ *eine Lösung des linearen Gleichungssystems* $Ax = b$. *Dann gilt:*

$$\mathbf{Lös}(A, b) = x^0 + \mathbf{Lös}(A, 0),$$

d. h. für jede Lösung y *des zugehörigen homogenisierten Gleichungssystems* $Ax = 0$ *ist* $x^0 + y$ *eine Lösung des inhomogenen Systems* $Ax = b$, *und jede Lösung des inhomogenen Systems ist von dieser Form.*

Beweis: Sei zunächst $y \in \mathbf{Lös}(A, 0)$. Dann gilt:

$$A\left(x^0 + y\right) = Ax^0 + Ay = b + 0 = b,$$

also ist $x^0 + y \in \mathbf{Lös}(A, b)$.

Ist umgekehrt $x \in \mathbf{Lös}(A, b)$. Dann gilt für $y := x - x^0$

$$Ay = A\left(x - x^0\right) = Ax - Ax^0 = b - b = 0,$$

also ist $y \in \mathbf{Lös}(A, 0)$. Aber damit ist $x = x^0 + y \in x^0 + \mathbf{Lös}(A, 0)$. □

Von jetzt an wollen wir uns mit der Lösbarkeit von linearen Gleichungssystemen über einem Körper K beschäftigen

In diesem Fall kann man die Aussage des Satzes 7.4 auch so formulieren:

> Falls $\mathbf{Lös}(A, b) \neq \emptyset$, dann ist $\mathbf{Lös}(A, b)$ ein *affiner Unterraum* [13] von K^n, wobei der zugehörige Unterraum $\mathbf{Lös}(A, 0)$ ist.

Man beachte, daß homogene lineare Gleichungssysteme stets lösbar sind (etwa durch $x = 0$), inhomogene aber i. a. nicht. Kennt man jedoch eine Lösung des inhomogenen Systems, so kennt man alle, sobald man die Lösungen des homogenisierten Systems kennt.

[13] D. h. $\mathbf{Lös}(A, b)$ ist von der Form $x^0 + U$ mit $x^0 \in K^n$ und $U \subseteq K^n$ ein Unterraum.

Definition 7.5
Ist $A \in \mathrm{Mat}(m \times n, K)$, $b \in K^m$, $x^0 \in \mathrm{Lös}(A, b)$ und $(y_1, \ldots, y_k)$ eine Basis von $\mathrm{Lös}(A, 0)$, dann nennen wir $\left(x^0 + y_1, \ldots, x^0 + y_k\right)$ eine *Basis* von $\mathrm{Lös}(A, b)$.

Zunächst betrachten wir das homogene Gleichungssystem $Ax = 0$. Dabei folgt das folgende Lemma unmittelbar aus den Definitionen.

Satz 7.6
Sei $Ax = 0$ *mit* $A \in \mathrm{Mat}(m \times n, K)$ *ein homogenes lineares Gleichungssystem. Dann gelten:*

a. $\dim_K \left(\mathrm{Lös}(A, 0)\right) = n - \mathrm{rang}(A)$.
b. *Die folgenden Aussagen sind gleichwertig:*
   (i) $0 \in K^n$ *ist die einzige Lösung von* $Ax = 0$.
   (ii) $\mathrm{Ker}(f_A) = \{0\}$.
   (iii) f_A *ist injektiv.*
   (iv) $\mathrm{rang}(A) = n$.

Beweis: Teil b. folgt unmittelbar aus Teil a. und Bemerkung 7.2 sowie Lemma I.3.18. Für Teil a. beachte man, daß nach Satz 4.16 gilt:

$$\dim_K \left(\mathrm{Lös}(A, 0)\right) = \dim_K \left(\mathrm{Ker}(f_A)\right) = n - \dim_K \left(\mathrm{Im}(f_A)\right) = n - \mathrm{rang}(A).$$

□

Für inhomogene Gleichungssysteme $Ax = b$ betrachtet man die erweiterte Koeffizientenmatrix $\widetilde{A} = (A, b)$.

Satz 7.7
Für das inhomogene Gleichungssystem $Ax = b$ *mit* $A \in \mathrm{Mat}(m \times n, K)$ *und* $b \in K^m$ *gelten:*

a. "Existenz einer Lösung"*: Die folgenden Aussagen sind äquivalent:*
   (i) $Ax = b$ *besitzt (mindestens) eine Lösung.*
   (ii) $b \in \mathrm{Im}(f_A)$.
   (iii) $\mathrm{rang}(A) = \mathrm{rang}(A, b)$.
b. "Eindeutigkeit der Lösung"*: Besitzt* $Ax = b$ *eine Lösung* x^0*, so sind die folgenden Aussagen äquivalent:*
   (i) x^0 *ist die einzige Lösung.*
   (ii) $\mathrm{Ker}(f_A) = \{0\}$.
   (iii) $\mathrm{rang}(A) = n$.

Beweis: a. Die Äquivalenz von (i) und (ii) wurde bereits in Bemerkung 7.2 diskutiert. Es bleibt also die Gleichwertigkeit von (ii) und (iii) zu zeigen. Da $\mathrm{Im}(f_A)$ von den Vektoren $a^1, \ldots, a^n$ erzeugt wird, gilt:

$$\begin{aligned} b \in \mathrm{Im}(f_A) &\Leftrightarrow b \text{ ist Linearkombination von } a^1, \ldots, a^n \\ &\Leftrightarrow \mathrm{Im}(f_A) = \langle a^1, \ldots, a^n \rangle = \langle a^1, \ldots, a^n, b \rangle = \mathrm{Im}(f_{\widetilde{A}}). \end{aligned}$$

Da ferner $\mathrm{Im}(f_A) \subseteq \mathrm{Im}(f_{\tilde{A}})$, gilt nach Lemma 4.8 die Gleichheit der Vektorräume genau dann, wenn

$$\mathrm{rang}(f_A) = \dim_K \big(\mathrm{Im}(f_A)\big) = \dim_K \big(\mathrm{Im}(f_{\tilde{A}})\big) = \mathrm{rang}(A, b).$$

b. folgt unmittelbar aus den Sätzen 7.6 und 7.4.

□

Satz 7.8 a. *Für* $A \in \mathrm{Mat}(m \times n, K)$ *sind gleichwertig:*
(i) $Ax = y$ *ist universell lösbar.*
(ii) f_A *ist surjektiv.*
(iii) $\mathrm{rang}(A) = m$.

b. *Für* $A \in \mathrm{Mat}(n, K)$ *sind gleichwertig:*
(i) $Ax = y$ *ist universell eindeutig lösbar.*
(ii) f_A *ist bijektiv.*
(iii) $\mathrm{rang}(A) = n$.

Beweis: Die Äquivalenz von (i) und (ii) wurde in beiden Fällen bereits in Bemerkung 7.2 diskutiert. Es bleibt also jeweils die Gleichwertigkeit von (ii) und (iii) zu zeigen.

a. f_A ist surjektiv, genau dann, wenn $\mathrm{Im}(f_A) = K^m$ gilt. In Anbetracht von $\mathrm{Im}(f_A) \subseteq K^m$ ist dies mittels Lemma 4.8 gleichwertig dazu, daß

$$\mathrm{rang}(A) = \mathrm{rang}(f_A) = \dim_K \big(\mathrm{Im}(f_A)\big) = m.$$

b. Mittels Satz 3.13 folgt die Behauptung aus Teil a.

□

Aus Satz 7.6 und Satz 7.8 ergibt sich folgender Algorithmus zum Test, ob eine lineare Abbildung injektiv, surjektiv oder bijektiv ist.

Algorithmus 7.9 (Test auf Injektivität / Surjektivität)
INPUT: $A \in \mathrm{Mat}(m \times n, K)$.

OUTPUT: Meldung, ob f_A injektiv, surjektiv oder bijektiv ist.

1. Schritt: Bestimme den Rang r von A.
2. Schritt: Ist $r = m = n$, gib "f_A ist bijektiv" zurück. Ist $r = m < n$, gib "f_A ist surjektiv" zurück. Ist $r = n < m$, gib "f_A ist injektiv" zurück.

Für die folgende Definition und das anschließende Lemma, betrachten wir kurzfristig noch einmal Gleichungssysteme über einem beliebigen kommutativen Ring mit Eins.

Definition 7.10
Zwei lineare Gleichungssysteme $Ax = b$ und $A'x = b'$ mit $A, A' \in \mathrm{Mat}(m \times n, R)$ und $b, b' \in R^m$ heißen *äquivalent*, wenn es eine invertierbare Matrix $T \in \mathrm{Gl}_m(R)$ gibt mit

$$A' = T \circ A \quad \text{und} \quad b' = Tb.$$

Bemerkung 7.11
Man weist leicht nach, daß die Äquivalenz von Gleichungssystemen $Ax = b$ mit $A \in \mathrm{Mat}(m \times n, R)$ und $b \in R^m$ eine Äquivalenzrelation auf der Menge der Gleichungssysteme $\{Ax = b \mid A \in \mathrm{Mat}(m \times n, R), b \in R^m\}$ ist.

Lemma 7.12
Es seien $Ax = b$ *und* $A'x = b'$ *mit* $A, A' \in \mathrm{Mat}(m \times n, R)$ *und* $b, b' \in R^m$ *zwei äquivalente lineare Gleichungssysteme. Dann gilt:*

$$\mathrm{Lös}(A, b) = \mathrm{Lös}\left(A', b'\right).$$

Beweis: Ist $x \in \mathrm{Lös}(A, b)$, dann gilt $Ax = b$ und damit

$$b' = Tb = T(Ax) = (T \circ A)x = A'x.$$

Also ist $x \in \mathrm{Lös}\left(A', b'\right)$.

Ist andererseits $x \in \mathrm{Lös}\left(A', b'\right)$, dann gilt $A'x = b'$ und damit

$$b = T^{-1}b' = T^{-1}\left(A'x\right) = \left(T^{-1} \circ A'\right)x = Ax.$$

Also ist $x \in \mathrm{Lös}(A, b)$. □

Wir wollen nun einen Algorithmus kennenlernen, der es uns erlaubt, die Lösung eines linearen Gleichungssystems $Ax = b$ in parametrisierter Form zu bestimmen, d. h. eine Basis des Lösungsraumes zu berechnen. Der wichtigste Schritt ist hierbei die Überführung der erweiterten Koeffizientenmatrix (A, b) in reduzierte Zeilen-Stufen-Form.

Wir müssen deshalb zunächst zeigen, daß der Lösungsraum eines linearen Gleichungssystems $Ax = b$, $A \in \mathrm{Mat}(m \times n, K)$, $b \in K^m$, invariant gegenüber den folgenden *elementaren Umformungen* ist:

I Eine Gleichung von $Ax = b$ wird mit $\lambda \in K \setminus \{0\}$ multipliziert.
II Eine Gleichung von $Ax = b$ wird zu einer anderen addiert.
III Das λ-fache einer Gleichung, $\lambda \in K$, wird zu einer anderen addiert
IV Die Reihenfolge der Gleichungen wird verändert.

Beachten wir, daß die Ausführung einer elementaren Umformung vom Typ I-IV für ein Gleichungssystem gleichwertig dazu ist, daß die zugehörige Matrix mittels einer entsprechenden elementaren Zeilenoperation vom Typ I-IV umgeformt wird, dann heißt das, wir müssen den folgenden Satz zeigen.

Satz 7.13
Ist $Ax = b$ *ein Gleichungssystem mit* $A \in \mathrm{Mat}(m \times n, K)$ *und* $b \in K^m$*, und ist* $A'x = b'$ *ein Gleichungssystem, das aus* $Ax = b$ *durch eine elementare Umformung entstanden ist, dann gilt:*

$$\mathrm{Lös}(A, b) = \mathrm{Lös}\left(A', b'\right).$$

Beweis: Da $A'x = b'$ aus $Ax = b$ durch eine elementare Umformung entstanden ist, gibt es, in Anbetracht von Proposition I.8.21 und Lemma I.8.18,

eine invertierbare Matrix $T \in \mathrm{Gl}_n(K)$ mit $T \circ (A, b) = \left(A', b'\right)$. Die beiden Gleichungssysteme sind also äquivalent und die Behauptung folgt aus Lemma 7.12. □

Aus Satz 7.13 und Satz I.9.4 folgt, daß wir die erweiterte Matrix eines Gleichungssystems $Ax = b$ mit $A \in \mathrm{Mat}(m \times n, K)$ und $b \in K^m$ mittels Gauß-Algorithmus in (reduzierte) Zeilen-Stufen-Form überführen können, ohne daß sich die Lösungsmenge ändert.

Wir betrachten deshalb den Fall, daß die Matrix A in ZSF gegeben ist, näher.

Satz 7.14
Sei $A \in \mathrm{Mat}(m \times n, K)$ *eine Matrix in Zeilen-Stufen-Form und* $b \in K^m$. *Die erweiterte Koeffizientenmatrix habe die Gestalt*

$$(A, b) = \left(\begin{array}{cccccccccccc|c} 0 & \dots & 0 & a_{1j_1} & * & \dots & \dots & \dots & \dots & \dots & \dots & * & b_1 \\ 0 & \dots & \dots & 0 & \dots & 0 & a_{2j_2} & * & \dots & \dots & \dots & * & b_2 \\ \vdots & & & & & & & \ddots & & & & \vdots & \vdots \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & 0 & a_{rj_r} & * & \dots & * & b_r \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 & b_{r+1} \\ \vdots & & & & & & & & & & & \vdots & \vdots \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 & b_m \end{array}\right) \tag{15}$$

mit Pivots $a_{ij_i} \neq 0$ *für* $i = 1, \dots, r$. *Dann gilt:*

a. *Ist eines der Elemente* $b_{r+1}, \dots, b_m$ *ungleich Null, so besitzt* $Ax = b$ keine *Lösung.*
b. *Sind* $b_{r+1} = \dots = b_m = 0$ *und gilt* $r = n$, *so besitzt* $Ax = b$ genau eine *Lösung.*
c. *Sind* $b_{r+1} = \dots = b_m = 0$ *und gilt* $r < n$, *so besitzt* $Ax = b$ *mehr als eine Lösung und* $\mathrm{Lös}(A, b)$ *läßt sich durch* K^{n-r} *parametrisieren.*[14]

Beweis: Wir zeigen zunächst, daß $\mathrm{Im}(f_A) = \{y \in K^m \mid y_{r+1} = \dots = y_m = 0\}$.

Die Inklusion "$\subseteq$" folgt daraus, daß die Zeilen $a_{r+1}, \dots, a_m$ von A alle Null sind. Die Inklusion "$\supseteq$" folgt daraus, daß $\left(a^{j_1}, \dots, a^{j_r}\right)$ eine linear unabhängige Familie der Dimension r ist mit Hilfe von Lemma 4.8.

[14] Daß sich eine Menge M durch K^k parametrisieren läßt, bedeutet, daß es eine "gute" Bijektion von K^k auf die Menge gibt. Was "gut" ist, kommt dabei auf den Zusammenhang an. In der Analysis wird dies in aller Regel bedeuten, daß die Abbildung hinreichend oft differenzierbar sein soll, während Topologen sich schon damit begnügen würden, daß die Abbildung stetig ist. Wir hingegen fordern, da $M = x^0 + U$, für einen Unterraum U von V und einen Vektor $x^0 \in V$, ein affiner Raum ist, daß die Abbildung affin sein soll, d. h. sie soll von der Form sein

$$\phi : K^k \to x^0 + U : y \mapsto x^0 + f(y),$$

wobei $f : K^k \to U$ eine lineare Abbildung ist. Wir nennen die Bijektion dann auch eine *Parametrisierung* von M.

Teil a. folgt dann mit Satz 7.7 und Teil b. folgt aus Satz 7.7 zusammen mit Satz 5.26. Teil c. folgt aus den Sätzen 7.7, 7.4 und 7.6. Man beachte dabei, daß $\text{Lös}(A,0)$ nach Korollar 3.11 als Vektorraum der Dimension $n-r$ isomorph zu K^{n-r} ist. □

Bemerkung 7.15
Wir wollen nun angeben, wie man im Fall c. aus Satz 7.14 aus der Zeilen-Stufen-Form (15) von A die *Parametrisierung von* $\text{Lös}(A,b)$ als Abbildung

$$\phi : K^{n-r} \to \text{Lös}(A,b)$$

herleitet. Sei hierzu $A = \text{RZSF}(A)$ in *reduzierter* ZSF gegeben.

Die Parameter $x_{j_1},\dots,x_{j_r}$ nennen wir die *gebundenen Parameter* und die x_j mit $j \in I := \{1,\dots,n\} \setminus \{j_1,\dots,j_r\}$ die *freien Parameter*. Dies rührt daher, daß sich aus (15) für eine Lösung x ergibt

$$x_{j_i} = b_i - \sum_{j\in I} a_{ij}x_j, \quad i = 1,\dots,r. \tag{16}$$

D. h. die gebundenen Parameter hängen von den freien Parametern ab.

Identifizieren wir K^{n-r} nun mit K^I und schreiben somit $y = (y_j \mid j \in I)$ für einen Vektor $y \in K^{n-r}$, dann ergibt sich die Parametrisierung hieraus als

$$\phi : K^{n-r} \to \text{Lös}(A,b) : y \mapsto x^0 + f(y), \tag{17}$$

wobei

$$x^0_j = \begin{cases} 0, & \text{falls } j \in I, \\ b_i, & \text{falls } j = j_i, \end{cases} \tag{18}$$

und

$$f : K^{n-r} \to K^n : y \mapsto (z_1,\dots,z_n)^t, \tag{19}$$

mit

$$z_j = \begin{cases} y_j, & \text{falls } j \in I, \\ -\sum_{k\in I} a_{ik}y_k, & \text{falls } j = j_i. \end{cases} \tag{20}$$

Damit ist f eine lineare Abbildung und somit ϕ affin.

Man beachte, daß x^0 in diesem Fall eine spezielle Lösung von $Ax = b$ ist, während $\text{Im}(f) = \text{Lös}(A,0)$.

Algorithmus 7.16 (Algorithmus zur Lösung eines LGS)
INPUT: $Ax = b$, ein LGS.

OUTPUT: Eine Parametrisierung $y \mapsto x^0 + f(y)$ von $\text{Lös}(A,b)$, d. h. eine spezielle Lösung x^0 sowie eine Matrix B mit $f = f_B$, sofern eine solche existiert.

1. **Schritt:** Aufstellen der erweiterten Koeffizientenmatrix (A,b).
2. **Schritt:** Berechnen einer reduzierten Zeilen-Stufen-Form (A',b') von (A,b) mit $r = \text{Zrang}(A')$.
3. **Schritt:** Ist $b'_{r+1} \neq 0$, dann ist das LGS nicht lösbar. Andernfalls bestimmt man die Parametrisierung gemäß (17)-(20), wobei im Fall $r = n$ der Vektor x^0 die eindeutig bestimmte Lösung ist. D. h.

- Ist $b'_{r+1} \neq 0$, gib eine Meldung zurück, daß das LGS nicht lösbar ist.
- Ist $b'_{r+1} = 0$ und $r = n$, so gib $x^0 = (b'_1, \ldots, b'_r)$ als eindeutig bestimmte Lösung zurück und keine Matrix.
- Ist $b'_{r+1} = 0$ und $r < n$, so bestimme zunächst die Pivotspalten $\{j_1, \ldots, j_r\}$. Initialisiere den Vektor x^0 als Nullvektor und addiere für $i = 1, \ldots, r$ zum j_i-ten Eintrag b'_i. Ferner definiere eine Matrix $B' \in \mathrm{Mat}(n, K)$ dadurch, daß für $j = 1, \ldots, n$ die j-te Zeile gerade der Einheitsvektor e_j ist, falls $j \notin \{j_1, \ldots, j_r\}$, und $-a'_i$ falls $j = j_i$ für ein $i = 1, \ldots, r$. Sodann streiche aus B' die Spalten $j_1, \ldots, j_r$, um eine Matrix B zu erhalten. Gib (x^0, B) zurück.

Beispiel 7.17 (Fortsetzung von Beispiel 7.3)
Mit Hilfe des Gauß-Algorithmus läßt sich Beispiel 7.3 systematischer Lösen. Wir überführen die erweiterte Koeffizientenmatrix mittels elementarer Zeilenoperationen in ZSF:

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 2 & 3 & 0 & 1 \\ 0 & 1 & -1 & 0 \end{array}\right) \mapsto \left(\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 0 & -1 & -2 & -1 \\ 0 & 1 & -1 & 0 \end{array}\right) \mapsto \left(\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 0 & -1 & -2 & -1 \\ 0 & 0 & -3 & -1 \end{array}\right).$$

Das zugehörige Gleichungssystem lautet jetzt:

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 1 \\ -x_2 - 2x_3 &= 1 \\ -3x_3 &= -1. \end{aligned}$$

Es ergibt sich: $x_3 = \frac{1}{3}$, $x_2 = 1 - 2x_3 = \frac{1}{3}$, $x_1 = 1 - 2x_2 - x_3 = 0$.

Beispiel 7.18
Wir betrachten das Gleichungssystem

$$\begin{aligned} x + y + z - u &= a \\ x - y + z + u &= b \\ 3x + y + 3z - u &= c \end{aligned} \tag{21}$$

mit zunächst beliebigen $a, b, c \in K$. In Matrixschreibweise lautet das Gleichungssystem:

$$\begin{pmatrix} 1 & 1 & 1 & -1 \\ 1 & -1 & 1 & 1 \\ 3 & 1 & 3 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ u \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

Durch den Gauß-Algorithmus überführen wir die erweiterte Koeffizientenmatrix in Zeilen-Stufen-Form:

$$\left(\begin{array}{cccc|c} 1 & 1 & 1 & -1 & a \\ 1 & -1 & 1 & 1 & b \\ 3 & 1 & 3 & -1 & c \end{array}\right) \mapsto \left(\begin{array}{cccc|c} 1 & 1 & 1 & -1 & a \\ 0 & -2 & 0 & 2 & b - a \\ 0 & 0 & 0 & 0 & c - 2a - b \end{array}\right).$$

Es folgt, daß das Gleichungssystem (21) genau dann lösbar ist, wenn $c = 2a + b$ gilt, was wir im folgenden annehmen wollen.

Um die Parametrisierung zu bestimmen, überführen wir die erweiterte Koeffizientenmatrix nun in reduzierte Zeilen-Stufen-Form:

$$\left(\begin{array}{cccc|c} 1 & 1 & 1 & -1 & a \\ 1 & -1 & 1 & 1 & b \\ 3 & 1 & 3 & -1 & c \end{array}\right) \mapsto \left(\begin{array}{cccc|c} 1 & 0 & 1 & 0 & \frac{1}{2}\cdot(a+b) \\ 0 & 1 & 0 & -1 & \frac{1}{2}\cdot(a-b) \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Daraus erhalten wir als Parametrisierung:

$$\phi : K^2 \to \text{Lös}(A,b) \subset K^4 : \begin{pmatrix} z \\ u \end{pmatrix} \mapsto \begin{pmatrix} \frac{1}{2}\cdot(a+b) \\ \frac{1}{2}\cdot(a-b) \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} -z \\ u \\ z \\ u \end{pmatrix}.$$

Für große numerische Matrizen wird man, statt die reduzierte ZSF zu bestimmen, die Parametrisierung aus der einfachen ZSF durch Rücksubstitution ermitteln, da das ein rechnerisch weniger aufwendiges Verfahren ist.

AUFGABEN

Aufgabe 7.19

Prüfe, ob die folgenden linearen Gleichungssysteme lösbar sind, und bestimme ggf. sämtliche Lösungen:

$$\begin{array}{rcrcrcl} -x & + & 6y & + & 2z & = & 4 \\ 2x & - & 2y & - & z & = & 2 \\ 3x & - & 4y & - & 2z & = & 1 \end{array} \qquad \begin{array}{rcrcrcrcl} x & + & y & + & z & - & u & = & 4 \\ x & - & y & + & z & + & u & = & 8 \\ 3x & + & y & + & 3z & - & u & = & 6 \end{array}$$

Aufgabe 7.20

Für welche $a, b \in \mathbb{R}$ besitzt das lineare Gleichungssystem

$$\begin{array}{rcrcrcl} ax & + & & & z & = & ab \\ -2x & + & by & + & az & = & -b \\ & & by & + & (a+1)z & = & b \end{array}$$

außer $(b, 1, 0)$ noch weitere Lösungen. Bestimme sie.

Aufgabe 7.21

Für welche $t \in \mathbb{R}$ ist das lineare Gleichungssystem

$$\begin{array}{rcrcrcl} tx & + & y & + & z & = & 1 \\ & & ty & + & z & = & 1 \\ tx & + & y & + & tz & = & 1 \end{array}$$

lösbar? Für welche $t \in \mathbb{R}$ ist die Lösung eindeutig?

Aufgabe 7.22

Für welche $t \in \mathbb{R}$ ist das lineare Gleichungssystem

$$\begin{array}{rcrcrcl} 2x & - & ty & + & z & = & 1 \\ 3x & + & 4y & + & 2z & = & t \\ tx & - & 2y & + & 3z & = & 2 \end{array}$$

lösbar?

Aufgabe 7.23
Für welche $a, b, c, d \in \mathbb{R}$ ist das lineare Gleichungssystem

$$\begin{array}{rcrcrcl} x & + & y & + & z & = & a \\ x & & 2y & + & 3z & = & b \\ x & + & 4y & + & 7z & = & c \\ x & + & 8y & + & 15z & = & d \end{array}$$

lösbar?

Aufgabe 7.24
Diskutiere die Lösungsmenge des linearen Gleichungssystems

$$\begin{pmatrix} a & b^2 \\ a^2 & b \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a-b \\ b-a \end{pmatrix}$$

in den Fällen

a. $ab \neq 0$,
b. $ab = 0$,
c. $a = b$, und
d. $a + b = 0$.

Aufgabe 7.25
Finde ein lineares Gleichungssystem über $\mathbb{R}$ mit spezieller Lösung $(2,2,3,3,4)^t$, dessen homogenes System $(1,0,0,2,0)^t$ sowie $(1,1,0,1,0)^t$ als Lösungen besitzt und Rang zwei hat.

Aufgabe 7.26
Seien $E_1 : a_1x + a_2y + a_3z = c_1$ und $E_2 : b_1x + b_2y + b_3z = c_2$ die Gleichungen zweier Ebenen E_1 und E_2 im $\mathbb{R}^3$ – vgl. Einleitung Beispiel .1.3. Beweise die folgenden Aussagen:

a. $E_1 = E_2$ genau dann, wenn gilt:

$$\exists\, \lambda \in \mathbb{R} \setminus \{0\} \,:\, \forall\, i = 1, \ldots, 3 \;\; \lambda a_i = b_i \text{ und } \lambda c_1 = c_2.$$

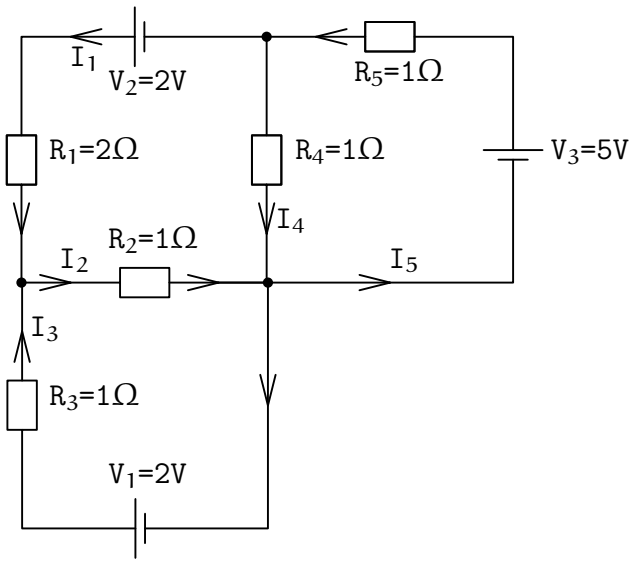
b. $E_1 \cap E_2 = \emptyset$ (d. h. E_1 und E_2 sind parallel) genau dann, wenn gilt:

$$\exists\, \lambda \in \mathbb{R} \setminus \{0\} \,:\, \forall\, i = 1, \ldots, 3 \;\; \lambda a_i = b_i, \text{ aber } \lambda c_1 \neq c_2.$$

c. $E_1 \cap E_2$ ist eine Gerade genau dann, wenn keiner der beiden obigen Fälle eintritt.

Hinweis: Wenn man geschickt vorgeht, kann man die drei Teilaufgaben mit einem gemeinsamen Ansatz bearbeiten. Zudem kann man "o. B. d. A." annehmen, daß $a_1 \neq 0$. Weshalb?

Aufgabe 7.27
Verwende das Ohmsche Gesetz sowie die beiden Kirchhoffschen Gesetze, um die in dem unten gegebenen Netzwerk fließenden Ströme $I_1, \ldots, I_5$ zu bestimmen - vgl. Einleitung Beispiel .1.7.



Aufgabe 7.28
Löse das folgende lineare Gleichungssystem mittels Gauß-Algorithmus und gib eine Parametrisierung der Lösungsmenge an:

$$\begin{array}{rcrcrcrcrcr} v & + & w & - & x & - & y & + & z & = & 0, \\ 2v & + & 2w & & & - & 2y & & & = & -4, \\ v & + & w & + & 3x & - & y & - & 3z & = & -8. \end{array}$$

Aufgabe 7.29
Prüfe, ob die folgenden Abbildungen injektiv, surjektiv oder bijektiv sind, und gib ggf. die inverse Abbildung an:

a. $f : \mathbb{R}^3 \to \mathbb{R}^3 \ : \ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} x+y \\ y+z \\ z+x \end{pmatrix}$,

b. $g : \mathbb{R}^4 \to \mathbb{R}^4 \ : \ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \mapsto \begin{pmatrix} 2x_1 + x_2 + x_4 \\ 3x_1 + x_2 + x_3 + x_4 \\ 2x_1 + x_3 + x_4 \\ 2x_1 + x_2 + x_3 \end{pmatrix}$.

8. Einige Algorithmen - Teil II

Wir wollen nun einige Algorithmen angeben, denen der Algorithmus zur Bestimmung der Parametrisierung der Lösungsmenge eines linearen Gleichungssystems zugrunde liegt.
Generalvoraussetzung: In diesem Abschnitt sei K ein Körper.

Ist $f = f_A$ für eine $m \times n$-Matrix A, dann ist der Kern von f gerade die Lösungsmenge $\text{Lös}(A, 0)$ des homogenen Gleichungssystems $Ax = 0$. Insbesondere bilden die Spalten der Matrixdarstellung der Parametrisierung eine Basis des Kerns.
Algorithmus 8.1 (Kern von f_A)
INPUT: $A \in \text{Mat}(m \times n, K)$.

OUTPUT: Eine Basis von $\mathrm{Ker}(f_A)$.

1. **Schritt:** Bestimme eine Parametrisierung (x^0, B) von $Ax = 0$ gemäß 7.16.
2. **Schritt:** Gib die Spaltenvektoren von B als Basis zurück.

Wir haben in Abschnitt 7 gelernt, daß Unterräume des K^n als Lösungsmengen von homogenen linearen Gleichungssystemen auftauchen. Um etwa den Schnitt zweier Unterräume des K^n zu bestimmen, ist es nützlich, aus dem Erzeugendensystem eines Unterraumes ein Gleichungssystem bestimmen zu können, das den Unterraum beschreibt.

Algorithmus 8.2 (Gleichungen eines Unterraumes)

INPUT: Eine Familie $F = (x_1, \ldots, x_m)$ von Vektoren im K^n.

OUTPUT: Eine Matrix $A \in \mathrm{Mat}(k \times n)$ mit $\mathrm{Lös}(A, 0) = \langle F \rangle$.

1. **Schritt:** Schreibe die Vektoren aus F als Zeilen in eine Matrix $B \in \mathrm{Mat}(m \times n, K)$ und bestimme eine Basis $(y_1, \ldots, y_k)$ von $\mathrm{Ker}(f_B) = \mathrm{Lös}(B, 0)$ mittels 8.1.
2. **Schritt:** Schreibe die Vektoren $(y_1, \ldots, y_k)$ als Zeilenvektoren in eine Matrix A.
3. **Schritt:** Gib A zurück.

Abschließend sind wir nun in der Lage, einen Algorithmus anzugeben, mittels dessen sich eine Basis des Schnitts zweier Unterräume des K^n ermitteln läßt.

Algorithmus 8.3 (Durchschnitt zweier Unterräume)

INPUT: Zwei Familien F und G von Vektoren in K^n.

OUTPUT: Eine Basis des Schnitts von $\langle F \rangle$ und $\langle G \rangle$.

1. **Schritt:** Bestimme Matrizen A und A' gemäß 8.2, so daß $\langle F \rangle = \mathrm{Lös}(A, 0)$ und $\langle G \rangle = \mathrm{Lös}\,(A', 0)$.
2. **Schritt:** Bilde aus den Zeilen von A und A' eine gemeinsame Matrix A''.
3. **Schritt:** Bestimme eine Basis B von $\mathrm{Ker}(f_{A''}) = \mathrm{Lös}\,(A'', 0)$ gemäß 8.1 und gib B zurück.

AUFGABEN

Aufgabe 8.4

Schreibe Singular-Prozeduren, die die obigen Algorithmen umsetzen.

Aufgabe 8.5

Es sei $U = \langle (2,-1,1,-1)^t, (1,-2,2,1)^t, (3,-1,0,2)^t \rangle \subset \mathbb{R}^4$ und $U' = \langle (3,-2,3,8)^t, (2,1,-5,3)^t \rangle \subset \mathbb{R}^4$. Bestimme eine Basis von $U \cap U'$.

9. Determinanten

Wir werden jetzt eine ganz neue Möglichkeit kennenlernen, um quadratische lineare Gleichungssysteme zu lösen, nämlich mit Hilfe von Determinanten. Die Determinante ordnet einer quadratischen Matrix über einem Körper ein

Element des Körpers zu, das genau dann ungleich Null ist, wenn die Matrix invertierbar ist. Die Determinante liefert aber nicht nur ein nützliches Kriterium für die Invertierbarkeit, sie ist vor allem aus theoretischen Gründen von unschätzbarem Wert. Z. B. liefert die Cramersche Regel mit Hilfe der Determinante eine geschlossene Formel für die Lösung eines linearen Gleichungssystems. Aus dieser Formel lassen sich Eigenschaften der Lösungen als Funktionen der Koeffizienten der Matrix bestimmen.

Die Determinante einer Matrix ist eine *polynomiale Funktion* in den Einträgen der Matrix. Sind diese Einträge etwa reelle oder komplexe Zahlen, so hängt die Determinante stetig von den Einträgen ab. Daraus folgt z. B. die wichtige Tatsache, daß eine invertierbare Matrix bei kleiner Störung der Einträge invertierbar bleibt. Damit wird eine Verbindung zur Analysis hergestellt. Eine weitere wichtige Bedeutung in der Analysis hat die Determinante für die Volumenberechnung.

Wir werden die Eigenschaften der Determinante soweit entwickeln, wie sie in der linearen Algebra wichtig sind. Allerdings führt uns die Determinante auch hier schon auf eine höhere Stufe: die Determinante ist nicht nur linear, sie ist *multilinear*, wie wir gleich sehen werden.

Da wir die Determinante später, im Zusammenhang mit Eigenwerten und Normalformen, auch für Ringe benötigen werden, entwickeln wir die grundlegenden Eigenschaften allgemein für Ringe.
Generalvoraussetzung: In diesem Abschnitt sei R stets ein kommutativer[15] Ring mit Eins und K ein Körper.
Definition 9.1
Wir definieren für $A \in \mathrm{Mat}(n, R)$ die *Determinante* von A durch die *Leibnizsche Formel*

$$\det(A) := |A| := \sum_{\sigma \in \mathbb{S}_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}. \tag{22}$$

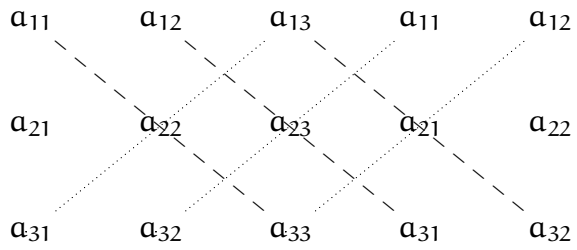
Beispiel 9.2 a. Ist $n = 1$, dann ist $A = (a) \in \mathrm{Mat}(1, K)$ und $\det(A) = a$.
b. Ist $n = 2$, dann ist $\mathbb{S}_2 = \{\mathrm{id}, (1\ 2)\}$ und damit folgt:

$$\det(A) = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21},$$

d. h. $\det(A)$ ist das Produkt der Elemente der Hauptdiagonalen minus dem Produkt der Elemente der Nebendiagonalen.

[15] Die Kommutativität ist in diesem Abschnitt von großer Bedeutung!

c. Für $n = 3$ hat $\mathbb{S}_n$ bereits sechs Elemente. Man berechnet in diesem Fall die Determinante mit der *Regel von Sarrus*:



Die Produkte der Elemente längs der gestrichelten Linien tauchen bei der Berechnung der Determinante als positive Summanden auf, die Produkte der Elemente längs der gepunkteten Linien als negative Summanden. D. h., wir erhalten:

$$\begin{aligned}\det(A) &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &\quad -a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}.\end{aligned}$$

d. Für $n = 4$ ergeben sich schon $4! = 24$ Summanden und für $n = 10$ gar $10! = 3628800$. In numerischen Anwendungen sind 1000×1000-Matrizen keine Seltenheit, so daß es sich von selbst versteht, daß dabei nicht die Definition, bei der dann für die Determinante über 10^{2567} Produkte berechnet werden müßten, zur Berechnung verwendet werden kann. In der Tat wird zur Berechnung von Determinanten über Körpern wieder der Gauß-Algorithmus eine wichtige Rolle spielen.

Proposition 9.3
Ist $A = (a_{ij}) \in \mathrm{Mat}(n, R)$ *eine obere (bzw. untere)* Dreiecksmatrix*, d. h.* $a_{ij} = 0$ *für* $i > j$ *(bzw.* $i < j$*), dann ist*

$$\det(A) = a_{11} \cdots a_{nn}$$

das Produkt der Diagonalelemente.

Beweis: Ist $\mathrm{id} \neq \sigma \in \mathbb{S}_n$, so gilt für mindestens ein $i \quad i > \sigma(i)$ (bzw. $i < \sigma(i)$). Wegen der Voraussetzung $a_{i\sigma(i)} = 0$ für $i > \sigma(i)$ (bzw. $i < \sigma(i)$) bleibt von den Summanden in (22) also nur der für id übrig. □

Bemerkung 9.4
Für die Determinante von $A \in \mathrm{Mat}(n, R)$ gilt:

$$\det(A) = \sum_{\sigma \in \mathbb{S}_n} \mathrm{sgn}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n}. \tag{23}$$

Dazu beachten wir, daß für $\sigma \in \mathbb{S}_n \quad \sigma^{-1}$ eine Permutation der Zahlen $1, \ldots, n$ ist, d. h. $\{1, \ldots, n\} = \{\sigma^{-1}(1), \ldots, \sigma^{-1}(n)\}$. Beachtet man nun noch, daß $\mathrm{sgn}(\sigma) = \mathrm{sgn}\left(\sigma^{-1}\right)$ und daß es gleich ist, ob ich über $\sigma \in \mathbb{S}_n$ summiere oder über $\sigma^{-1} \in \mathbb{S}_n$, da auf beide Weisen alle Elemente von $\mathbb{S}_n$ je einmal erreicht werden, dann

erhalten wir:

$$\begin{aligned}\det(A) &= \sum_{\sigma\in\mathbb{S}_n} \operatorname{sgn}(\sigma) a_{\sigma^{-1}(1)\sigma\left(\sigma^{-1}(1)\right)} \cdots a_{\sigma^{-1}(n)\sigma\left(\sigma^{-1}(n)\right)} \\ &= \sum_{\sigma\in\mathbb{S}_n} \operatorname{sgn}(\sigma) a_{\sigma^{-1}(1)1} \cdots a_{\sigma^{-1}(n)n} \\ &= \sum_{\sigma\in\mathbb{S}_n} \operatorname{sgn}\left(\sigma^{-1}\right) a_{\sigma^{-1}(1)1} \cdots a_{\sigma^{-1}(n)n} \\ &= \sum_{\sigma^{-1}\in\mathbb{S}_n} \operatorname{sgn}\left(\sigma^{-1}\right) a_{\sigma^{-1}(1)1} \cdots a_{\sigma^{-1}(n)n} \\ &= \sum_{\pi\in\mathbb{S}_n} \operatorname{sgn}(\pi) a_{\pi(1)1} \cdots a_{\pi(n)n}.\end{aligned}$$

Proposition 9.5
Für $A \in \mathrm{Mat}(n, R)$ *gilt:*

$$\det(A) = \det\left(A^t\right).$$

Beweis: Sei $A = (a_{ij})$ und $A^t = (a'_{ij})$, dann gilt $a'_{ij} = a_{ji}$. Mithin erhalten wir mit Hilfe von Bemerkung 9.4

$$\begin{aligned}\det(A) &= \sum_{\sigma\in\mathbb{S}_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} \\ &= \sum_{\sigma\in\mathbb{S}_n} \operatorname{sgn}(\sigma) a'_{\sigma(1)1} \cdots a'_{\sigma(n)n} \quad = \quad \det\left(A^t\right).\end{aligned}$$

□

Die Determinante läßt sich wie folgt axiomatisch charakterisieren.
Satz 9.6 a. *Die* Determinantenfunktion

$$\det : \mathrm{Mat}(n, R) \to R : A \mapsto \det(A)$$

hat folgende Eigenschaften:

(i) det *ist linear in jeder Spalte, d. h. für* $1 \le i \le n$ *und* $\lambda, \mu \in R$ *gilt:*

$$\det\left(a^1 \ldots \lambda a^i + \mu b^i \ldots a^n\right) = \lambda \det\left(a^1 \ldots a^i \ldots a^n\right) + \mu \det\left(a^1 \ldots b^i \ldots a^n\right).$$

(ii) $\det(A) = 0$*, falls* A *zwei gleiche Spalten enthält.*

(iii) $\det(\mathbb{1}_n) = 1$.

b. *Ist* $\delta : \mathrm{Mat}(n, R) \to R$ *eine Funktion mit den Eigenschaften (i) und (ii) von a., dann gilt für* $A \in \mathrm{Mat}(n, R)$

$$\delta(A) = \delta(\mathbb{1}_n) \cdot \det(A).$$

Insbesondere gilt also, daß det *durch die Eigenschaften (i)-(iii) in a. eindeutig bestimmt ist.*

Definition 9.7
Es seien V und W R-Moduln.

a. Eine Abbildung

$$f : V^n = V \times \overset{n}{\ldots} \times V \to W$$

heißt *multilinear*, falls f in jedem Argument linear ist, d. h. falls für $x_i, y_i \in V$, $i = 1, \ldots, n$ und $\lambda, \mu \in R$ gilt:

$$f(x_1, \ldots, \lambda x_i + \mu y_i, \ldots, x_n) = \lambda f(x_1, \ldots, x_i, \ldots, x_n) + \mu f(x_1, \ldots, y_i, \ldots, x_n).$$

b. Eine multilineare Abbildung $f : V^n \to W$ heißt *alternierend*, falls für $(x_1, \ldots, x_n) \in V^n$ mit $x_i = x_j$ für ein $i \neq j$, gilt:

$$f(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n) = 0.$$

Bemerkung 9.8
Die Eigenschaften (i) und (ii) in Satz 9.6 a. besagen also, daß det eine alternierende multilineare Abbildung von $\mathrm{Mat}(n, R) = R^n \times \overset{n}{\ldots} \times R^n \to R$ ist, wo die Gleichheit von $\mathrm{Mat}(n, R)$ und $R^n \times \overset{n}{\ldots} \times R^n$ so zu verstehen ist, daß wir eine Matrix als Tupel ihrer Spaltenvektoren auffassen. Bei der Eigenschaft (iii) handelt es sich um eine Normierung.

Eine bezüglich den Spalten alternierende multilineare Abbildung $\delta : \mathrm{Mat}(n, R) \to R$ heißt eine *Determinantenform*. Teil b. in Satz 9.6 besagt dann, daß es nur eine normierte Determinantenform gibt.

Beweis von Satz 9.6: Wir werden im Beweis die Formel (23) zur Berechnung der Determinante verwenden, da sie auf die Bedürfnisse der Determinante als multilineare Abbildung bezüglich der Spalten zugeschnitten ist.

a. Es seien $a^j = (a_{1j}, \ldots, a_{nj})^t$, $j = 1, \ldots, n$, und $b^i = (b_{1i}, \ldots, b_{ni})^t$. Wir setzen $A := (a^1 \ldots a^i \ldots a^n)$, $B := (a^1 \ldots b^i \ldots a^n)$ und $C := (a^1 \ldots \lambda a^i + \mu b^i \ldots a^n)$.

(i) Dann gilt:

$$\begin{aligned} \det(C) &= \textstyle\sum_{\sigma \in \mathbb{S}_n} \mathrm{sgn}(\sigma) a_{\sigma(1)1} \cdots (\lambda a_{\sigma(i)i} + \mu b_{\sigma(i)i}) \cdots a_{\sigma(n)n} \\ &= \lambda \textstyle\sum_{\sigma \in \mathbb{S}_n} \mathrm{sgn}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n} \\ &\quad + \mu \textstyle\sum_{\sigma \in \mathbb{S}_n} \mathrm{sgn}(\sigma) a_{\sigma(1)1} \cdots b_{\sigma(i)i} \cdots a_{\sigma(n)n} \\ &= \lambda \det(A) + \mu \det(B). \end{aligned}$$

(ii) Sei nun $a^i = a^j$, für ein $i \neq j$. Ist $\tau = (i\ j)$, die Transposition, die i und j vertauscht, dann besitzt $\mathbb{S}_n$ die folgende Zerlegung:

$$\mathbb{S}_n = \mathbb{A}_n \cup \mathbb{A}_n \tau.$$

Ferner gilt für $\sigma \in \mathbb{A}_n$

$$\mathrm{sgn}(\sigma) = 1 \quad \text{und} \quad \mathrm{sgn}(\sigma\tau) = -1.$$

Wir erhalten also:

$$\begin{aligned} \det(A) &= \textstyle\sum_{\sigma \in \mathbb{A}_n} a_{\sigma(1)1} \cdots a_{\sigma(i)i} \cdots a_{\sigma(j)j} \cdots a_{\sigma(n)n} \\ &\quad - \textstyle\sum_{\sigma \in \mathbb{A}_n} a_{\sigma\tau(1)1} \cdots a_{\sigma\tau(i)i} \cdots a_{\sigma\tau(j)j} \cdots a_{\sigma\tau(n)n} = 0 \end{aligned}$$

Denn wegen $a_{\sigma\tau(i)i} = a_{\sigma(j)i} = a_{\sigma(j)j}$, $a_{\sigma\tau(j)j} = a_{\sigma(i)j} = a_{\sigma(i)i}$ und da $\sigma\tau(k) = \sigma(k)$ für $k \neq i, j$ folgt mittels der Kommutativität der Multiplikation in R, daß die erste Summe gleich der zweiten ist.

(iii) $\det(\mathbb{1}_n) = \mathrm{sgn}(\mathrm{id}) \cdot 1 \cdots 1 = 1$.

b. Mit den Notationen von a. gilt $a^i = \sum_{j=1}^n a_{ji} e^j$. Aus der Multilinearität von δ folgt:

$$\delta(A) = \sum_{j_1=1}^n a_{j_1 1}\delta\big(e^{j_1}\ a^2 \dots a^n\big) = \sum_{j_1=1}^n a_{j_1 1} \sum_{j_2=1}^n a_{j_2 2}\delta\big(e^{j_1}\ e^{j_2}\ a^3 \dots a^n\big)$$

$$= \dots = \sum_{j_1,\dots,j_n=1}^n a_{j_1 1} \cdots a_{j_n n}\delta\big(e^{j_1}\ \dots\ e^{j_n}\big).$$

Genau dann, wenn $j_1, \dots, j_n$ alle paarweise verschieden sind, existiert eine Permutation $\sigma \in \mathbb{S}_n$ mit $\big(e^{j_1} \dots e^{j_n}\big) = \big(e^{\sigma(1)} \dots e^{\sigma(n)}\big)$.
Falls zwei der j_i übereinstimmen, dann gilt $\delta\big(e^{j_1}\ \dots\ e^{j_n}\big) = 0$, da δ alternierend ist.
Aus dem folgenden Lemma 9.9 folgt:

$$\delta(e^{\sigma(1)} \dots e^{\sigma(n)}) = \operatorname{sgn}(\sigma) \cdot \delta(e^1 \dots e^n) = \operatorname{sgn}(\sigma) \cdot \delta(\mathbb{1}_n).$$

Insgesamt haben wir damit gezeigt:

$$\delta(A) = \sum_{\sigma\in\mathbb{S}_n} a_{\sigma(1)1} \cdots a_{\sigma(n)n} \operatorname{sgn}(\sigma) \cdot \delta(\mathbb{1}_n) = \det(A) \cdot \delta(\mathbb{1}_n).$$

□

Lemma 9.9
Ist $f : V^n \to W$ *eine alternierende multilineare Abbildung zwischen* R*-Moduln, dann gilt für* $\sigma \in \mathbb{S}_n$

$$f(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = \operatorname{sgn}(\sigma) f(x_1, \dots, x_n).$$

Insbesondere gilt:

$$f(x_1, \dots, x_i, \dots, x_j, \dots, x_n) = -f(x_1, \dots, x_j, \dots, x_i, \dots, x_n).$$

Beweis: Da σ Produkt von Transpositionen ist und das Signum ein Gruppenhomomorphismus ist, genügt es, den Fall zu untersuchen, daß $\sigma = (i\ j)$ eine Transposition ist. Da ferner f alternierend und multilinear ist, folgt:

$$\begin{aligned} 0 &= f(x_1, \dots, x_i + x_j, \dots, x_i + x_j, \dots, x_n) \\ &= f(x_1, \dots, x_i, \dots, x_i, \dots, x_n) + f(x_1, \dots, x_i, \dots, x_j, \dots, x_n) \\ &\quad + f(x_1, \dots, x_j, \dots, x_i, \dots, x_n) + f(x_1, \dots, x_j, \dots, x_j, \dots, x_n) \\ &= f(x_1, \dots, x_i, \dots, x_j, \dots, x_n) + f(x_1, \dots, x_j, \dots, x_i, \dots, x_n). \end{aligned}$$

Daraus folgt die Behauptung. □

Aus Satz 9.6 leiten wir sofort das folgende Korollar ab.
Korollar 9.10
Es sei $A \in \operatorname{Mat}(n, R)$ *und* $\lambda \in R$.

a. *Bei Vertauschung zweier Spalten von* A *ändert sich das Vorzeichen von* $\det(A)$.
b. *Bei Multiplikation einer Spalte von* A *mit* λ *multipliziert sich* $\det(A)$ *mit* λ.

c. *Bei Addition des λ-fachen einer Spalte zu einer anderen Spalte ändert sich* $\det(A)$ *nicht.*
d. *Sind zwei Spalten von* A *linear abhängig, so ist* $\det(A) = 0$. *Insbesondere ist* $\det(A) = 0$*, falls eine Spalte von* A *gleich Null ist.*

Beweis: a. Da det nach Satz 9.6 alternierend ist, folgt die Aussage aus Lemma 9.9.
b. Dies folgt aus der Multilinearität von det, siehe Satz 9.6.
c. Für $A = \left(a^1 \dots a^n\right)$ und $A' = \left(a^1 \dots a^j + \lambda a^i \dots a^n\right)$ folgt aus der Multilinearität, und da det alternierend ist:

$$\det(A') = \det(A) + \lambda \det(a^1 \dots a^i \dots a^i \dots a^n) = \det(A) + \lambda \cdot 0 = \det(A).$$

d. Sind zwei Spalten linear abhängig, so verschwindet eine Spalte nach Addition des λ-fachen der anderen Spalte, ohne daß sich die Determinante ändert - wegen c.. Ist aber eine Spalte von A Null, so folgt $\det(A) = 0$ aus b. mit $\lambda = 0$.

□

Da die Determinante einer Matrix gleich der Determinante der Transponierten ist, sind die Begriffe Spalte und Zeile austauschbar. Eine exaktere Formulierung bietet das folgende Korollar.

Korollar 9.11
Wir können $\det : \mathrm{Mat}(n, R) \to R$ *auch als multilineare Abbildung auf den Zeilen einer Matrix* A *auffassen.*

Entsprechend gilt Korollar 9.10 auch für Zeilen statt Spalten.

Satz 9.12 (Determinantenproduktsatz)
Für Matrizen $A, B \in \mathrm{Mat}(n, R)$ *gilt:*

$$\det(A \circ B) = \det(A) \cdot \det(B).$$

Beweis: Wähle $A \in \mathrm{Mat}(n, R)$ fest und betrachte die Abbildung

$$\delta : \mathrm{Mat}(n, R) \to R : B \mapsto \det(A \circ B).$$

δ ist multilinear bezüglich der Spalten von B, da A auf jede Spalte von B linear wirkt. Außerdem ist δ alternierend, da mit B auch $A \circ B$ zwei gleiche Spalten hat. Damit folgt aus Satz 9.6:

$$\det(A \circ B) = \delta(B) = \delta(\mathbb{1}_n) \cdot \det(B) = \det(A) \cdot \det(B).$$

□

Korollar 9.13
Für $A \in \mathrm{Gl}_n(R)$ *gilt:*

$$\det\left(A^{-1}\right) = \left(\det(A)\right)^{-1}.$$

Beweis: Für $A \in \mathrm{Gl}_n(R)$ gilt:

$$1 = \det(\mathbb{1}_n) = \det\left(A \circ A^{-1}\right) = \det(A) \cdot \det\left(A^{-1}\right),$$

und analog $1 = \det\left(A^{-1}\right) \cdot \det(A)$. Also ist $\det(A)$ invertierbar und die Inverse ist $\det\left(A^{-1}\right)$. □

Korollar 9.14
Die Abbildung

$$\det : \left(\mathrm{Gl}_n(R), \circ\right) \to \left(R^*, \cdot\right)$$

ist ein Gruppenepimorphismus.

Beweis: Wegen Korollar 9.13 liegt das Bild von det in R^*. Nach Satz I.8.14 und Lemma I.7.18 sind $\left(\mathrm{Gl}_n(R), \circ\right)$ und $\left(R^*, \cdot\right)$ Gruppen, und nach Satz 9.12 ist det ein Gruppenhomomorphismus. Ferner gilt für $\lambda \in R^*$ gerade,

$$\det \left(\begin{array}{c|ccc} \lambda & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & \mathbb{1}_{n-1} & \\ 0 & & & \end{array}\right) = \lambda,$$

also ist det surjektiv. □

Korollar 9.15
Sind A und B konjugiert, d. h. es gibt ein $T \in \mathrm{Gl}_n(R)$ mit $B = T \circ A \circ T^{-1}$, dann gilt:

$$\det(A) = \det(B).$$

Beweis: Aus Satz 9.12 und Korollar 9.13 folgen:

$$\det\left(T \circ A \circ T^{-1}\right) = \det(T) \cdot \det(A) \cdot \det(T)^{-1} = \det(A).$$

□

Soweit konnten wir die Theorie der Determinanten für Matrizen über beliebigen kommutativen Ringen mit Eins entwickeln. Jetzt wollen wir spezielle Aussagen herleiten, die nur für Körper gelten.

Zunächst können wir Korollar 9.15 verwenden, um die Determinante für Endomorphismen eines endlich-dimensionalen Vektorraumes zu definieren.

Definition 9.16
Sei V ein endlich-dimensionaler K-Vektorraum mit Basis B und $f \in \mathrm{End}_K(V)$. Wir definieren die *Determinante* von f durch

$$\det(f) := \det\left(M_B^B(f)\right).$$

Bemerkung 9.17
Da die Matrixdarstellungen eines Endomorphismus f zu verschiedenen Basen nach Satz 5.11 konjugiert sind, ist diese Definition unter Berücksichtigung von Korollar 9.15 unabhängig von der Wahl der Basis B.

Wir kommen jetzt zur praktischen Berechnung von Determinanten. Wir haben schon gesehen, daß die Definition selbst nur zur Berechnung von Determinanten kleiner Matrizen geeignet ist. In der Tat treten in der Summe (22) so viele positive wie negative Vorzeichen auf. Arbeiten wir nun über $\mathbb{Z}$ oder $\mathbb{Q}$ oder $\mathbb{R}$, so wird durch das wechselnde Vorzeichen das Resultat in der Regel nicht so groß sein, als wenn man positive Zahlen aufsummieren müßte. Sind allerdings die Einträge Unbestimmte (d. h. Einträge aus dem Polynomring $R[x_{ij} \mid 1 \leq i, j \leq n]$), dann hat die Determinante wirklich $n!$ Terme, sofern sich keine Terme aufheben. In diesem Fall nützt also auch kein noch so raffinierter Algorithmus zur Berechnung, das Ergebnis ist für $n > 10$ praktisch nicht mehr nutzbar, wenn überhaupt berechenbar. Allerdings treten in der Praxis oft *dünn besetzte* Matrizen auf, also solche, die viele Nullen enthalten.

In diesem Fall, wie auch im Fall numerischer Einträge ist es extrem wichtig, daß Terme, die sich im Laufe der Rechnung wieder wegheben, gar nicht erst berechnet werden. Der Gauß-Algorithmus (und eine Variante davon für Ringe, der sogenannte *Bareiss-Algorithmus* - siehe [**Coh96**]), bewirken dies.

Algorithmus 9.18 (Algorithmus zur Berechnung der Determinante über K)
INPUT: $A \in \mathrm{Mat}(n, K)$.

OUTPUT: $\det(A)$.

1. **Schritt:** Initialisiere eine Integer z.
2. **Schritt:** Überführe A mittels Gauß-Algorithmus in nicht-reduzierte ZSF.[16] Jedesmal, wenn dabei zwei Zeilen vertauscht werden, ersetze z durch $-z$. - Tritt während der Gaußreduktion eine Nullspalte auf, gib Null zurück und breche ab.
3. **Schritt:** Gib das Produkt von z mit den Diagonalelementen der ZSF zurück.

Beispiel 9.19
Sei $A \in \mathrm{Mat}(n+1, K)$ gegeben durch

$$A = \begin{pmatrix} 0 & 1 & 2 & 3 & \dots & n \\ 1 & 0 & 1 & 2 & \dots & n-1 \\ 2 & 1 & 0 & 1 & \dots & n-2 \\ 3 & 2 & 1 & 0 & \dots & n-3 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ n & n-1 & n-2 & n-3 & \dots & 0 \end{pmatrix}.$$

[16] D. h. führe im Gauß-Algorithmus I.9.8 Schritt sechs nicht aus.

Ziehe von der i-ten Zeile die $(i+1)$-te Zeile ab für $i = 1, \dots, n$. Wir erhalten:

$$A = \begin{pmatrix} -1 & 1 & 1 & 1 & \dots & 1 & 1 \\ -1 & -1 & 1 & 1 & \dots & 1 & 1 \\ -1 & -1 & -1 & 1 & \dots & 1 & 1 \\ -1 & -1 & -1 & -1 & \dots & 1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -1 & -1 & -1 & -1 & \dots & -1 & 1 \\ n & n-1 & n-2 & n-3 & \dots & 1 & 0 \end{pmatrix}.$$

Addiere nun die erste Spalte zur i-ten Spalte für $i = 2, \dots, n+1$. Dann erhalten wir:

$$A = \begin{pmatrix} -1 & 0 & 0 & 0 & \dots & 0 & 0 \\ * & -2 & 0 & 0 & \dots & 0 & 0 \\ * & * & -2 & 0 & \dots & 0 & 0 \\ * & * & * & -2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ & * & * & * & \dots & -2 & 0 \\ * & * & * & * & \dots & * & n \end{pmatrix}.$$

Es folgt:

$$\det(A) = (-1) \cdot (-2)^{n-1} \cdot n = -n \cdot (-2)^{n-1}.$$

In diesem Beispiel haben wir durch ganz wenige Zeilen- und Spaltenoperationen die Matrix in Dreiecksgestalt überführt. Das lag aber an der speziellen Struktur der Matrix. Im allgemeinen Fall braucht der oben beschriebene Algorithmus zur Berechnung der Determinante mit Hilfe des Gauß-Algorithmus $\sim \frac{n^3}{3}$ Multiplikationen für eine $n \times n$-Matrix. In der Definition der Determinante tauchen dagegen $n!$ Summanden von je n Produkten auf, mit $n! \sim \left(\frac{n}{e}\right)^n$, wobei $e \approx 2.712\dots$ die eulersche Zahl ist. Man sagt, daß der Gauß-Algorithmus *polynomial*, die Definition aber *exponentiell* in der Größe der Matrix ist. Grundsätzlich gelten polynomiale Algorithmen als effizient, exponentielle dagegen als unakzeptabel ineffizient. Allerdings gibt es Fälle, wo keine polynomialen Algorithmen bekannt sind.

AUFGABEN

Aufgabe 9.20

Es sei R ein kommutativer Ring mit Eins und $\lambda \in R$. Bestimme die Determinante der Matrix

$$\begin{pmatrix} 1 & \lambda & \lambda^2 & \dots & \lambda^{n-1} \\ \lambda^{n-1} & 1 & \lambda & \dots & \lambda^{n-2} \\ \lambda^{n-2} & \lambda^{n-1} & 1 & \dots & \lambda^{n-3} \\ \vdots & \vdots & \vdots & & \vdots \\ \lambda & \lambda^2 & \lambda^3 & \dots & 1 \end{pmatrix} \in \mathrm{Mat}(n, R).$$

Aufgabe 9.21

Es sei K ein Körper, $A \in \mathrm{Mat}(n, K)$. Zeige:

a. Ist $\mathrm{char}(K) \neq 2$, $n \in \mathbb{N}$ ungerade und gilt $A^t = -A$, dann ist A nicht invertierbar.
b. Ist $\mathrm{char}(K) = 2$, so gibt es zu jedem $0 \neq n \in \mathbb{N}$ ein invertierbares A mit $A^t = -A$.

Aufgabe 9.22
Es sei V ein n-dimensionaler $\mathbb{C}$-Vektorraum und $f \in \mathrm{End}_{\mathbb{C}}(V)$. Mittels Einschränkung der Skalarmultiplikation können wir V als $\mathbb{R}$-Vektorraum auffassen. Zeige:

a. $f \in \mathrm{End}_{\mathbb{R}}(V)$.
b*. Bezeichnen wir mit $\det_{\mathbb{C}}(f)$ die Determinante von f als $\mathbb{C}$-lineare Abbildung und $\det_{\mathbb{R}}(f)$ die Determinante von f als $\mathbb{R}$-lineare Abbildung. Dann gilt:

$$\det_{\mathbb{R}}(f) = |\det_{\mathbb{C}}(f)|^2$$

Hinweis: Für eine $\mathbb{C}$-Basis $(v_1, \ldots, v_n)$ von V betrachte man die zugehörige $\mathbb{R}$-Basis $(v_1, \ldots, v_n, iv_1, \ldots, iv_n)$ - vgl. LAI, Aufgabe 40 - sowie jeweils die zugehörige Matrixdarstellung von f. - Wem der allgemeine Fall zu schwer ist, der beschränke sich auf die Abbildung $f : \mathbb{C} \to \mathbb{C} : z \mapsto (a+ib)\cdot z$ mit $a, b \in \mathbb{R}$ fest vorgegeben. Was ist eine Basis von $\mathbb{C}$ als $\mathbb{R}$-Vektorraum.

10. Laplacescher Entwicklungssatz und Cramersche Regel

Wir kommen jetzt zu einer alternativen Berechnung der Determinante, die allgemein für Matrizen über Ringen funktioniert.
Generalvoraussetzung: In diesem Abschnitt sei R ein kommutativer Ring mit Eins und K ein Körper.
Satz 10.1 (Kästchensatz)
Es sei $A \in \mathrm{Mat}(n, R)$ *eine* Blockmatrix *der Form*

$$A = \left(\begin{array}{c|c} B & C \\ \hline 0 & D \end{array}\right)$$

mit $B \in \mathrm{Mat}(k, R)$, $C \in \mathrm{Mat}(k \times l, R)$, $D \in \mathrm{Mat}(l, R)$, $0 \in \mathrm{Mat}(l \times k, R)$ *und* $n = k + l$. *Dann gilt:*

$$\det(A) = \det(B) \cdot \det(D).$$

Beweis: Man beachte, daß

$$A = \left(\begin{array}{c|c} B & C \\ \hline 0 & D \end{array}\right) = \left(\begin{array}{c|c} \mathbb{1}_k & 0 \\ \hline 0 & D \end{array}\right) \circ \left(\begin{array}{c|c} B & C \\ \hline 0 & \mathbb{1}_l \end{array}\right).$$

Wegen des Determinantenproduktsatzes 9.12 reicht es mithin zu zeigen:

$$\det\left(\begin{array}{c|c} \mathbb{1}_k & 0 \\ \hline 0 & D \end{array}\right) = \det(D) \tag{24}$$

und

$$\det\left(\begin{array}{c|c} B & C \\ \hline 0 & \mathbb{1}_l \end{array}\right) = \det(B). \tag{25}$$

Die Abbildung

$$\delta_1 : \mathrm{Mat}(l, R) \to R : D' \mapsto \det \left(\begin{array}{c|c} \mathbb{1}_k & 0 \\ \hline 0 & D' \end{array} \right)$$

ist offensichtlich eine Determinantenform, d. h. sie genügt den Bedingungen (i) und (ii) aus Satz 9.6 a., mithin gilt:

$$\det \left(\begin{array}{c|c} \mathbb{1}_k & 0 \\ \hline 0 & D \end{array} \right) = \delta_1(D) = \delta_1(\mathbb{1}_l) \cdot \det(D) = \det(\mathbb{1}_n) \cdot \det(D) = \det(D),$$

d. h. (24) ist erfüllt.

Analog ist die Abbildung

$$\delta_2 : \mathrm{Mat}(k, R) \to R : B' \mapsto \det \left(\begin{array}{c|c} B' & C \\ \hline 0 & \mathbb{1}_l \end{array} \right)$$

alternierend und multilinear in den Spalten von B', also eine Determinantenform. Wieder folgt aus Satz 9.6 mit Hilfe von Proposition 9.3, daß

$$\det \left(\begin{array}{c|c} B & C \\ \hline 0 & \mathbb{1}_l \end{array} \right) = \delta_2(B) = \delta_2(\mathbb{1}_k) \cdot \det(B) = \det \begin{pmatrix} \mathbb{1}_k & C \\ 0 & \mathbb{1}_l \end{pmatrix} \cdot \det(B) = \det(B),$$

womit auch (25) gezeigt ist. □

Zur Herleitung des Laplaceschen Entwicklungssatzes und der Cramerschen Regel führen wir jetzt verschiedene Hilfmatrizen ein.

Definition 10.2

Es sei $A = (a_{ij}) = (a^1 \dots a^n) \in \mathrm{Mat}(n, R)$, $n \geq 2$, und $b = (b_1, \dots, b_n)^t \in R^n$.

Wir definieren die *Ersetzungsmatrix*

$$A_i(b) := (a^1 \dots a^{i-1}\ b\ a^{i+1} \dots a^n),$$

in der die i-te Spalte von A durch b ersetzt wurde.

Ist $b = e^j$ der j-te Einheitsvektor, so gilt:

$$A_i(e^j) = \begin{pmatrix} a_{11} & \dots & 0 & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{j1} & \dots & 1 & \dots & a_{jn} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & 0 & \dots & a_{nn} \end{pmatrix}.$$

Ersetzen wir in $A_i(e^j)$ zusätzlich noch die j-te Zeile durch den i-ten Einheitsvektor, dann erhält man die Matrix

$$S_{ji}(A) = \begin{pmatrix} a_{11} & \dots & 0 & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ 0 & \dots & 1 & \dots & 0 \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & 0 & \dots & a_{nn} \end{pmatrix}.$$

Streicht man in der Matrix A die j-te Zeile und die i-te Spalte, so erhält man die *Streichungsmatrix*

$$A_{ji} = \left(\begin{array}{ccc|ccc} a_{11} & \dots & a_{1\,i-1} & a_{1\,i+1} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots & \\ a_{j-1\,1} & \dots & a_{j-1\,i-1} & a_{j-1\,i+1} & \dots & a_{j-1\,n} \\ \hline a_{j+1\,1} & \dots & a_{j+1\,i-1} & a_{j+1\,i+1} & \dots & a_{j+1\,n} \\ \vdots & & \vdots & & \vdots & \\ a_{n\,1} & \dots & a_{n\,i-1} & a_{n\,i+1} & \dots & a_{n\,n} \end{array}\right).$$

Lemma 10.3
Für $A \in \mathrm{Mat}(n, \mathbb{R})$, $n \geq 2$, $1 \leq i, j \leq n$, gilt:

$$\det\left(A_i(e^j)\right) = \det\left(S_{ji}(A)\right) = (-1)^{i+j} \det(A_{ji}).$$

Beweis: $S_{ji}(A)$ entsteht aus $A_i(e^j)$ durch Subtraktion des a_{jk}-fachen der i-ten Spalte von der k-ten Spalte, $k \in \{1, \dots, n\} \setminus \{i\}$. Also gilt nach Korollar 9.10:

$$\det\left(A_i(e^j)\right) = \det\left(S_{ji}(A)\right).$$

Durch $i - 1$ Spaltenvertauschungen und $j - 1$ Zeilenvertauschungen entsteht aus $S_{ji}(A)$ die Matrix

$$\left(\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & A_{ji} & \\ 0 & & & \end{array}\right).$$

Also folgt aus dem Kästchensatz 10.1 unter Beachtung der Korollare 9.10 und 9.11

$$\det\left(S_{ji}(A)\right) = (-1)^{i+j} \det(A_{ji}).$$

□

In der folgenden Definition beachte man die Vertauschung der Indizes!
Definition 10.4
Für $A \in \mathrm{Mat}(n, \mathbb{R})$, $n \geq 2$, $1 \leq i, j \leq n$ heißt

$$a^{\#}_{ij} := (-1)^{i+j} \det(A_{ji})$$

ein *Kofaktor* von A. Die Matrix der Kofaktoren

$$A^{\#} := \left(a^{\#}_{ij}\right) \in \mathrm{Mat}(n, \mathbb{R})$$

heißt die *Adjunkte* oder *Komplementärmatrix* von A.
Satz 10.5 (Satz über die Adjunkte)
Für $A \in \mathrm{Mat}(n, \mathbb{R})$, $n \geq 2$, gilt:

$$A^{\#} \circ A = A \circ A^{\#} = \det(A) \cdot \mathbb{1}_n.$$

Beweis: Sei $A^{\#} \circ A = (c_{ik})$. Dann gilt mit Lemma 10.3:

$$\begin{aligned} c_{ik} = \sum_{j=1}^{n} a_{ij}^{\#} a_{jk} &= \sum_{j=1}^{n} a_{jk} \det\left(a^1 \dots a^{i-1}\ e^j\ a^{i+1} \dots a^n\right) \\ &= \det\left(a^1 \dots a^{i-1}\ \sum_{j=1}^{k} a_{jk} e^j\ a^{i+1} \dots a^n\right) \\ &= \det\left(a^1 \dots a^{i-1}\ a^k\ a^{i+1} \dots a^n\right) = \delta_{ik} \cdot \det(A). \end{aligned}$$

Das dritte Gleichheitszeichen folgt aus der Multilinearität von det, das letzte, da det alternierend ist.

Der Beweis, daß $A \circ A^{\#} = \det(A) \cdot \mathbb{1}_n$ geht analog. □

Korollar 10.6
Es sei $A \in \mathrm{Mat}(n, R)$. *Dann sind gleichwertig:*

a. A *ist invertierbar.*
b. $\det(A) \in R^*$, *d. h.* $\det(A)$ *ist invertierbar in* R.
c. $f_A : R^n \to R^n : x \mapsto Ax$ *ist ein Isomorphismus.*

Zudem gilt, ist A *invertierbar, dann ist*

$$A^{-1} = \frac{1}{\det(A)} \cdot A^{\#}.$$

Beachte, ist R *ein Körper, so bedeutet b. gerade:*

$$\det(A) \neq 0.$$

Beweis: Ist A invertierbar, dann ist $\det(A)$ invertierbar nach Korollar 9.13. Ist nun umgekehrt $\det(A)$ invertierbar, dann ist nach Satz 10.5 auch A invertierbar mit Inverser $A^{-1} = \frac{1}{\det(A)} \cdot A^{\#}$.

Die Äquivalenz von a. und c. haben wir bereits in Proposition 2.15 gesehen. □

Wir wollen an dieser Stelle einmal die vielen Aussagen, die zur Invertierbarkeit einer quadratischen Matrix über einem Körper äquivalent sind, sammeln.

Korollar 10.7
Für eine Matrix $A \in \mathrm{Mat}(n, K)$ *sind gleichwertig:*

a. A *ist invertierbar.*
b. $\mathrm{rang}(A) = n$.
c. $\det(A) \neq 0$.
d. f_A *ist bijektiv.*
e. f_A *ist injektiv.*
f. f_A *ist surjektiv.*
g. $\mathrm{RZSF}(A) = \mathbb{1}_n$.
h. A *ist das Produkt endlich vieler Elementarmatrizen.*

i. *Es gibt eine Matrix* $B \in \mathrm{Mat}(n, K)$ *mit* $B \circ A = \mathbb{1}_n$.

j. *Das Gleichungssystem* $Ax = y$ *ist universell eindeutig lösbar.*

Beweis: Die unterschiedlichen Äquivalenzen sind in den Sätzen 2.15, 3.13, 5.21, 5.29, 7.2 und 10.6 gezeigt worden. □

Zur Bedeutung des Satzes 10.5 bzw. von Korollar 10.6 ist anzumerken, daß aus der Invertierbarkeit von A als Element von $\mathrm{Mat}(n, R)$ folgt, daß die Einträge von A^{-1} wieder in R sind.

Ist beispielsweise $A \in \mathrm{Mat}(n, \mathbb{Z})$, dann können wir wegen $\mathbb{Z} \subset \mathbb{Q}$ die Matrix A mittels des Gauß-Algorithmus auf Zeilen-Stufen-Form bringen. Falls dann das Produkt der Diagonalelemente (also $\det(A)$) eine Einheit in $\mathbb{Z}$ ist - und das heißt $\det(A) = \pm 1$ -, dann hat auch A^{-1} Einträge aus $\mathbb{Z}$. Ist aber $\det(A) \in \mathbb{Q} \setminus \{-1, 0, 1\}$, dann hat A^{-1} Einträge in $\mathbb{Q}$ und mindestens einer der Einträge von A^{-1} ist keine ganze Zahl - selbst, wenn $\det(A)$ eine ganze Zahl war.

Beispiel 10.8

Für eine 2×2-Matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

gilt:

$$A^{\#} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}, \quad \text{mit } \det(A) = ad - bc.$$

Ist also $ad - bc \in R^*$, so gilt:

$$A^{-1} = \frac{1}{ad - bc} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Sei nun konkret $R = \mathbb{Z}$ und

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 1 & 1 \\ 3 & 4 \end{pmatrix}.$$

Dann gilt $\det(A) = -2 \notin \{-1, 1\} = \mathbb{Z}^*$, also ist $A \notin \mathrm{Gl}_n(\mathbb{Z})$. Hingegen ist $B \in \mathrm{Gl}_2(\mathbb{Z})$, da $\det(B) = 1 \in \mathbb{Z}^*$ und $B^{-1} \in \mathrm{Mat}(2, \mathbb{Z})$ mit

$$B^{-1} = \begin{pmatrix} 4 & -1 \\ -3 & 1 \end{pmatrix}.$$

Man beachte, daß im Fall $R = \mathbb{Q}$ für die gleiche Matrix A wegen $\det(A) = -2 \in \mathbb{Q}^* = \mathbb{Q} \setminus \{0\}$ gilt, daß A über $\mathbb{Q}$ invertierbar ist, insbesondere also $A^{-1} \in \mathrm{Mat}(2, \mathbb{Q}) \setminus \mathrm{Mat}(2, \mathbb{Z})$ mit

$$A^{-1} = -\frac{1}{2} \cdot \begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix}.$$

Der Satz über die Adjunkte führt zu einer rekursiven Berechnungsformel für die Determinante, die theoretisch sehr nützlich ist. Sie ist auch als rekursive Prozedur sehr einfach zu programmieren, aber nicht sehr effizient. Sie hat die gleiche Komplexität, wie die Leibnizsche Formel (22) zur Definition der Determinante.

Satz 10.9 (Laplacescher Entwicklungssatz)
Es sei $A \in \mathrm{Mat}(n, R)$.

a. *Wir nennen die folgende Formel, die* Entwicklung nach der i-ten Zeile*:*

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}). \tag{26}$$

b. *Entsprechend nennen wir die folgende Formel, die* Entwicklung nach der j-ten Spalte*:*

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}). \tag{27}$$

Beweis: Nach Satz 10.5 gilt für $A \circ A^{\#} = (c_{ik})$

$$\det(A) = c_{ii} = \sum_{j=1}^{n} a_{ij} a^{\#}_{ji} = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}).$$

Damit folgt (26), und (27) zeigt man analog durch die Betrachtung von $A^{\#} \circ A$. □

Bemerkung 10.10
Entwickelt man $A = (a_{ij})$ nach der ersten Zeile, so gilt:

$$\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{n2} & \dots & a_{nn} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} & \dots & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n3} & \dots & a_{nn} \end{vmatrix}$$
$$+ \dots + (-1)^{n+1} a_{1n} \begin{vmatrix} a_{21} & \dots & a_{2\,n-1} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{n\,n-1} \end{vmatrix}.$$

Benutzt man dieses Verfahren, so entwickelt man am Besten nach Zeilen bzw. Spalten, die möglichst viele Nullen enthalten. Die Vorzeichen merkt man sich am Günstigsten mit der sogenannten *Schachbrettregel*:

$$\begin{vmatrix} + & - & + & - & \dots \\ - & + & - & + & \dots \\ + & - & + & - & \dots \\ - & + & - & + & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{vmatrix}$$

Für kleine Matrizen, insbesondere wenn die Matrix dünn besetzt ist, ist dies Verfahren zur Berechnung der Determinante (und zur Berechnung der Inversen) durchaus verwendbar. Für größere Matrizen ist auf jeden Fall der gaußsche Eliminationsalgorithmus (und für Ringe die Variante von Bareiss) vorzuziehen.

Zur Berechnung der Determinante ganzzahliger Matrizen ist die effizienteste Methode, die Determinante von A modulo vieler verschiedener Primzahlen p_i zu berechnen, und dann $\det(A)$ mit Hilfe des Chinesischen Restsatzes I.6.16 zu bestimmen - vgl. Beispiel I.6.17.

Algorithmus 10.11 (Laplace-Entwicklung)
INPUT: $A \in \mathrm{Mat}(n, R)$.

OUTPUT: $\det(A)$.

1. Schritt: Initialisiere det auf Null.
2. Schritt: Falls $n = 1$, setze $\det = a_{11}$ und gehe zu Schritt 3. Sonst tue für $i = 1, \ldots, n$:
- Bilde eine Hilfsmatrix B durch Streichen der ersten Spalte und der i-ten Zeile von A.
- Rufe den Algorithmus mit B auf und merke Dir das Ergebnis in einer Hilfsvariablen x.
- Addiere zu det gerade $(-1)^{i+1} \cdot a_{i1} \cdot x$.

3. Schritt: Gib det zurück.

Der Satz über die Adjunkte liefert auch eine theoretisch sehr wichtige geschlossene Formel für die Lösungen eines linearen Gleichungssystems. Dies ist die berühmte *Cramersche Regel*. Wir formulieren sie allgemein für lineare Gleichungssysteme über einem Ring.

Satz 10.12 (Cramersche Regel)
Es sei $A \in \mathrm{Mat}(n, R)$.

a. *Das lineare Gleichungssystem* $Ax = y$ *ist genau dann universell eindeutig lösbar, wenn* $\det(A) \in R^*$.
b. *Ist* $\det(A)$ *eine Einheit in* R *und ist* $b \in R^n$*, dann gilt für die eindeutig bestimmte Lösung* $x = (x_1, \ldots, x_n)^t \in R^n$ *von* $Ax = b$

$$\begin{aligned} x_i &= \tfrac{1}{\det(A)} \cdot \det\left(A_i(b)\right) \\ &= \tfrac{1}{\det(A)} \cdot \det \begin{pmatrix} a_{11} & \ldots & a_{1\,i-1} & b_1 & a_{1\,i+1} & \ldots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \ldots & a_{n\,i-1} & b_n & a_{n\,i+1} & \ldots & a_{nn} \end{pmatrix}. \end{aligned}$$

Beweis: a. Nach Satz 7.8 ist $Ax = y$ genau dann universell eindeutig lösbar, wenn f_A bijektiv ist, was nach Korollar 10.6 gleichwertig zu $\det(A) \in R^*$ ist.

b. Es sei also nun $\det(A) \in R^*$, so ist A invertierbar nach Korollar 10.6 mit

$$x = A^{-1}b = \frac{1}{\det(A)} \cdot A^{\#}b$$

als eindeutig bestimmter Lösung des linearen Gleichungssystems. Also folgt mit Lemma 10.3 und der Multilinearität der Determinante

$$\begin{aligned} x_i &= \tfrac{1}{\det(A)} \cdot \textstyle\sum_{j=1}^{n} a_{ij}^{\#} b_j = \tfrac{1}{\det(A)} \cdot \sum_{j=1}^{n} \det\left(A_i(e^j)\right) b_j \\ &= \tfrac{1}{\det(A)} \cdot \textstyle\sum_{j=1}^{n} \det\left(a^1 \dots a^{i-1}\ e^j\ a^{i+1} \dots a^n\right) b_j \\ &= \tfrac{1}{\det(A)} \cdot \det\left(a^1 \dots a^{i-1}\ b\ a^{i+1} \dots a^n\right) \\ &= \tfrac{1}{\det(A)} \cdot \det\left(A_i(b)\right). \end{aligned}$$

□

Bemerkung 10.13 a. Eine in der Analysis, aber nicht nur dort, wichtige Eigenschaft der Determinante ist, daß sie stetig von den Einträgen der Matrix $A = (a_{ij})$ abhängt. Das soll folgendes heißen:

Sind die $a_{ij} = a_{ij}(t)$ stetige Funktionen in $t \in \mathbb{R}^m$, dann ist, da

$$\det\left(A(t)\right) = \sum_{\sigma \in \mathbb{S}_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)}(t) \cdots a_{n\sigma(n)}(t)$$

Summe von Produkten der Funktionen $a_{ij}(t)$ ist, die Funktion

$$\mathbb{R}^m \to \mathbb{R} : t \mapsto \det\left(A(t)\right)$$

eine stetig von t abhängende Funktion.
Alternativ, sind die a_{ij} Variable, dann ist die Funktion

$$\mathbb{R}^{n^2} \to \mathbb{R} : (a_{ij}) \mapsto \det(a_{ij})$$

ein Polynom vom Grad n in den Variablen a_{ij}, und damit sogar unendlich oft differenzierbar.

Insbesondere folgt, sind die Funktionen $a_{ij}(t)$ stetig und gilt $\det\left(A(t_0)\right) \neq 0$, dann gilt auch $\det\left(A(t)\right) \neq 0$ für alle t in einer hinreichend kleinen Umgebung von t_0.
Aus der Cramerschen Regel folgt weiter:

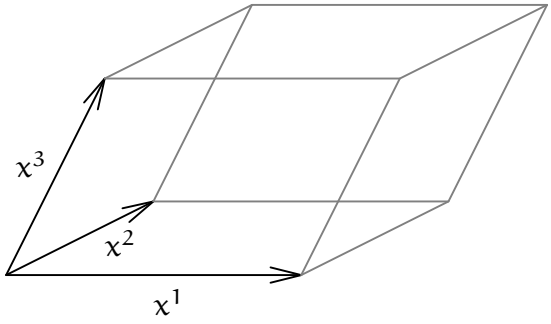
Ist das lineare Gleichungssystem $A(t)x = b$ für $t = t_0$ eindeutig lösbar, dann ist es auch für t nahe bei t_0 eindeutig lösbar, und die Lösungen $x(t) = \left(x_1(t), \dots, x_n(t)\right)^t$ hängen in einer Umgebung von t_0 stetig von t ab.

b. Eine wichtige geometrische Eigenschaft der Determinante, die man ebenfalls in der Analysis beweist, ist die Eigenschaft, eine *Volumenform* zu sein:

Seien $x^1, \dots, x^n \in \mathbb{R}^n$ und sei

$$P(x^1, \dots, x^n) := \{\lambda_1 x^1 + \dots + \lambda_n x^n \in \mathbb{R}^n \mid 0 \le \lambda_i \le 1, i = 1, \dots, n\}$$

das von den Vektoren $x^1, \dots, x^n$ aufgespannte *Parallelotop*.



Dann gilt:

$$|\det(x^1 \dots x^n)| = \text{Volumen}(P(x^1, \dots, x^n)).$$

AUFGABEN

Aufgabe 10.14

Für $t \in \mathbb{R}$ definiere $f_t \in \text{End}_{\mathbb{R}}(\mathbb{R}^3)$ durch $f(x, y, z) = (x+z, x+2y+z, tx+y-z)^t$.

a. Für welche $t \in \mathbb{R}$ ist f_t ein Automorphismus?
b. Berechne in diesen Fällen f_t^{-1} mit Hilfe der Adjungten von $M_E^E(f)$, wo E die kanonische Basis des $\mathbb{R}^3$ ist.

Aufgabe 10.15

Für $n \in \mathbb{N} \setminus \{0\}$ definieren wir $A_n \in \text{Mat}(n, \mathbb{R})$ als die Matrix, deren Einträge auf der Diagonalen sowie auf der oberen und unteren Nebendiagonalen alle eins sind, während alle anderen Einträge null sind. Ferner setzen wir $d_n = \det(A_n)$.

a. Zeige, für $n \ge 3$ gilt die Rekursionsformel $d_n = d_{n-1} - d_{n-2}$.
b. Zeige, für $k \in \mathbb{N}$ gilt

$$d_n = \begin{cases} 1, & \text{falls } n \equiv 1 (\text{mod } 6) \text{ oder } n \equiv 0 (\text{mod } 6), \\ 0, & \text{falls } n \equiv 2 (\text{mod } 6) \text{ oder } n \equiv 5 (\text{mod } 6), \\ -1, & \text{falls } n \equiv 3 (\text{mod } 6) \text{ oder } n \equiv 4 (\text{mod } 6). \end{cases}$$

Aufgabe 10.16

Es sei K ein Körper, $\lambda \in K$ und $n \in \mathbb{N} \setminus \{0\}$. Definiere

$$A_{n,\lambda} = \begin{pmatrix} \lambda & -1 & 0 & 0 & \dots & 0 \\ -1 & \lambda & -1 & 0 & \dots & 0 \\ 0 & -1 & \lambda & -1 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & -1 & \lambda & -1 \\ 0 & \dots & \dots & 0 & -1 & \lambda \end{pmatrix} \in \text{Mat}(n \times n, K).$$

Leite eine Rekursionsformel für $d_{n,\lambda} = \det(\lambda_{n,\lambda})$ her und zeige, $d_{2+3k,1} = 0$ für alle $k \in \mathbb{N}$.

Aufgabe 10.17

Es sei $V = \mathrm{Mat}(2, K)$, der Vektorraum der 2×2-Matrizen über dem Körper K. Zu $A \in V$ assoziieren wir eine K-lineare Abbildung

$$T_A : V \to V : X \mapsto A \circ X.$$

Zeige:

a. Für $A \in V$ ist $T_A \in \mathrm{End}_K(V)$.
b. Genau dann ist $\det(A) \neq 0$, wenn $\det\left(T_A\right) \neq 0$.
c. $\mathrm{Spur}\left(T_A\right) = 2\,\mathrm{Spur}(A)$.
d. $\alpha : V \to \mathrm{End}_K(V) : A \mapsto T_A$ ist K-linear, d. h. $\alpha \in \mathrm{Hom}_K\left(V, \mathrm{End}_K(V)\right)$.

Aufgabe 10.18

Es seien $a, b \in \mathbb{R}$ gegeben, V sei ein $\mathbb{R}$-Vektorraum mit Basis $B = (v_1, v_2)$ und $f \in \mathrm{End}_{\mathbb{R}}(V)$ sei gegeben durch $f(v_1) = av_1 + bv_2$ und $f(v_2) = -bv_1 + av_2$.

a. Zeige, $\left(\mathrm{id}_V, f, f^2\right)$ ist eine linear abhängige Familie im $\mathbb{R}$-Vektorraum $\mathrm{End}_{\mathbb{R}}(V)$.
b. Zeige, genau dann ist f ein Automorphismus, wenn $a^2 + b^2 \neq 0$.
c. Falls $a^2 + b^2 \neq 0$, bestimme $f^{-1}(v_1)$ und $f^{-1}(v_2)$.

KAPITEL III

Eigenwerte und Normalformen

1. Der Polynomring $K[t]$

Ziel dieses Abschnittes ist es, den Polynomring $K[t]$ über einem Körper K einzuführen und zu zeigen, daß er einige gute Eigenschaften hat, die wir bereits vom Ring $\mathbb{Z}$ der ganzen Zahlen kennen. Große Teile des Abschnittes lassen sich jedoch über einem allgemeinen kommutativen Ring mit Eins formulieren.

Generalvoraussetzung: In diesem Abschnitt bezeichne R stets einen kommutativen Ring mit Eins und K einen Körper.

Bevor wir uns dem Polynomring zuwenden, wollen wir ein Lemma beweisen, das uns sagt, daß wir nicht nur lineare Abbildungen auf einer Basis beliebig vorschreiben können, sondern auch bilineare Abbildungen.

Lemma 1.1
Sei V ein R-Modul mit Basis $B = (x_i \mid i \in I)$ und sei $F = \big(y_{ij} \mid (i,j) \in I \times I\big)$ eine Familie von Vektoren in V. Dann gibt es genau eine bilineare Abbildung $\alpha : V \times V \to V$ mit $\alpha(x_i, x_j) = y_{ij}$ für alle $(i,j) \in I \times I$.

Sind $f = \sum'_{i\in I} a_i x_i$, $g = \sum'_{i\in I} b_i x_i \in V$, so gilt

$$\alpha(f,g) = \sum_{i\in I}{}' \sum_{j\in I}{}' a_i b_j y_{ij}. \tag{28}$$

Beweis: *Existenz:* Seien $f, g \in V$ beliebig gegeben, dann lassen sich f und g eindeutig als endliche Linearkombinationen $f = \sum'_{i\in I} a_i x_i$ und $g = \sum'_{i\in I} b_i x_i$ der Basis B mit $a_i, b_i \in R$ für $i \in I$ darstellen. Definieren wir $\alpha(f,g)$ durch (28), so wird dadurch eine Abbildung

$$\alpha : V \times V \to V$$

definiert. Es bleibt zu zeigen, daß α bilinear ist.

Sei dazu ferner $h = \sum'_{i\in I} c_i x_i \in V$ und $\lambda, \mu \in R$. Dann gilt:

$$\begin{aligned}
\alpha(f, \lambda g + \mu h) &= \alpha\left(\textstyle\sum'_{i\in I} a_i x_i, \sum'_{i\in I} (\lambda b_i + \mu c_i) x_i\right) \\
&= \textstyle\sum'_{i\in I} \left(\sum'_{j\in I} a_i (\lambda b_j + \mu c_j) y_{ij}\right) \\
&= \textstyle\sum'_{i\in I} \left(\lambda \sum'_{j\in I} a_i b_j y_{ij} + \mu \sum'_{j\in I} a_i c_j y_{ij}\right) \\
&= \lambda \textstyle\sum'_{i\in I} \sum'_{j\in I} a_i b_j y_{ij} + \mu \sum'_{i\in I} \sum'_{j\in I} a_i c_j y_{ij} \\
&= \lambda\alpha(f,g) + \mu\alpha(f,h).
\end{aligned}$$

Analog sieht man:

$$\alpha(\lambda g + \mu h, f) = \lambda\alpha(g, f) + \mu\alpha(h, f).$$

Also ist α bilinear.

Eindeutigkeit: Sei $\beta : V \times V \to V$ eine weitere bilineare Abbildung mit $\beta(x_i, x_j) = y_{ij}$ für alle $(i,j) \in I \times I$. Es ist zu zeigen, daß $\alpha(f,g) = \beta(f,g)$ gilt für alle $(f,g) \in V \times V$.

Seien dazu $f = \sum'_{i\in I} a_i x_i$, $g = \sum'_{i\in I} b_i x_i \in V$ gegeben. Dann folgt aus der Bilinearität von β

$$\begin{aligned}\beta(f,g) &= \beta\left(\sum'_{i\in I} a_i x_i, \sum'_{i\in I} b_i x_i\right) = \sum'_{i\in I} a_i \beta\left(x_i, \sum'_{j\in I} b_j x_j\right)\\ &= \sum'_{i\in I} a_i \sum'_{j\in I} b_j \beta(x_i, x_j) = \sum'_{i\in I}\sum'_{j\in I} a_i b_j y_{ij} = \alpha(f,g).\end{aligned}$$

□

Wir kommen nun zur Definition des Polynomringes.

Definition 1.2 a. Es sei $V = R^{\mathbb{N}}$ der R-Modul der Folgen in R, vgl. Beispiel II.1.13 und Aufgabe II.2.31.
Wir definieren für $i \in \mathbb{N}$ die Folge

$$t^i : \mathbb{N} \to R : n \mapsto \delta_{in} = \begin{cases} 1, & i = n,\\ 0, & i \neq n.\end{cases}$$

Den von den t^i erzeugten Untermodul

$$R[t] := \left\langle t^i \mid i \in \mathbb{N}\right\rangle_R$$

nennen wir den *Modul der abbrechenden Folgen* in R oder den *Polynomring* in der Veränderlichen t über R. Die Elemente von $R[t]$ nennen wir *abbrechende Folgen* oder *Polynome*.

b. Wir definieren auf $R[t]$ eine bilineare Abbildung

$$\cdot : R[t] \times R[t] \to R[t]$$

durch

$$t^i \cdot t^j := t^{i+j} \quad \text{für } (i,j) \in \mathbb{N} \times \mathbb{N}. \tag{29}$$

Daß durch (29) eine bilineare Abbildung auf $R[t]$ definiert wird, folgt aus Lemma 1.1 und Teil a. der folgenden Bemerkung.

Bemerkung 1.3 a. **Behauptung:** $B := \left(t^i \mid i \in \mathbb{N}\right)$ ist eine *Basis* des R-Moduls $R[t]$.
Nach Definition gilt bereits, daß B ein Erzeugendensystem von $R[t]$ ist. Es bleibt also zu zeigen, daß B linear unabhängig ist. Seien dazu $\lambda_i \in R$, $i \in \mathbb{N}$, gegeben mit:

$$f := \sum_{i\in\mathbb{N}}{}' \lambda_i t^i = 0,$$

wobei mit 0 hier die Nullfolge, sprich die Abbildung $0 : \mathbb{N} \to R : n \mapsto 0_R$, gemeint ist. Dann gilt für $n \in \mathbb{N}$ beliebig:

$$0 = f(n) = \sum_{i\in\mathbb{N}}{}' \lambda_i t^i(n) = \sum_{i\in\mathbb{N}}{}' \lambda_i \delta_{in} = \lambda_n,$$

d. h. $\lambda_n = 0$ für alle $n \in \mathbb{N}$. Also ist die Familie B linear unabhängig.

b. Da B eine Basis von $R[t]$ ist, läßt sich also jedes Polynom $f \in R[t]$ in eindeutiger Weise als endliche Linearkombination

$$f = \sum_{i\in\mathbb{N}}{}' a_i t^i, \quad \text{mit } a_i \in R,$$

darstellen. Daraus leiten wir mit Hilfe von (28) und des Cantorschen Diagonalverfahrens die folgenden Formeln für $f = \sum'_{i\in\mathbb{N}} a_i t^i, g = \sum'_{i\in\mathbb{N}} b_i t^i \in R[t]$ und $\lambda \in R$ ab:

$$\lambda f = \sum_{i\in\mathbb{N}}{}' \lambda a_i t^i, \tag{30}$$

$$f + g = \sum_{i\in\mathbb{N}}{}' (a_i + b_i) t^i, \tag{31}$$

$$f \cdot g = \sum_{i\in\mathbb{N}}{}' \sum_{j\in\mathbb{N}}{}' a_i b_j t^{i+j} = \sum_{i\in\mathbb{N}}{}' \left(\sum_{i+j=k} a_i b_j \right) t^k = \sum_{i\in\mathbb{N}}{}' \left(\sum_{i=0}^{k} a_i b_{k-i} \right) t^k. \tag{32}$$

Daß wir $R[t]$ Polynomring nennen, rechtfertigt sich aus dem folgenden Satz.

Satz 1.4

Der R-Modul $R[t]$ wird durch die zweistellige Operation "$\cdot$" aus Definition 1.2 eine kommutative R-Algebra mit Eins t^0.

Beweis: Nach Definition II.5.3 bleibt folgendes zu zeigen:

a. $\big(R[t], +, \cdot\big)$ ist ein kommutativer Ring mit Eins t^0, und
b. für alle $f, g \in R[t]$ und $\lambda \in R$ gilt $\lambda(f \cdot g) = (\lambda f) \cdot g = f \cdot (\lambda g)$.

Teil b. ist eine Folgerung der Bilinearität von "$\cdot$" und für a. reicht es, da $R[t]$ ein R-Modul ist, zu zeigen, daß für $f, g, h \in R[t]$ gilt:

(i) $(f \cdot g) \cdot h = f \cdot (g \cdot h)$,
(ii) $f \cdot g = g \cdot f$,
(iii) $f \cdot (g + h) = f \cdot g + f \cdot h$, und
(iv) $t^0 \cdot f = f$.

Seien dazu also $f = \sum'_{i\in\mathbb{N}} a_i t^i, g = \sum'_{i\in\mathbb{N}} b_i t^i, h = \sum'_{i\in\mathbb{N}} c_i t^i \in R[t]$ gegeben. Dann folgt aus (32) mit $d_k = \sum_{i+j=k} a_i b_j$ und $d'_k = \sum_{i+j=k} b_i c_j$

$$f \cdot g = \sum_{k\in\mathbb{N}}{}' d_k t^k \quad \text{und} \quad g \cdot h = \sum_{k\in\mathbb{N}}{}' d'_k t^k. \tag{33}$$

Damit erhalten wir also:

$$(f \cdot g) \cdot h = \sum_{k\in\mathbb{N}}{}' \left(\sum_{i+j=k} d_i c_j \right) t^k$$

und

$$f \cdot (g \cdot h) = \sum_{k\in\mathbb{N}}{}' \left(\sum_{i+j=k} a_i d'_j \right) t^k.$$

Für (i) bleibt mithin zu zeigen:

$$\sum_{i+j=k} d_i c_j = \sum_{i+j=k} a_i d'_j \quad \forall\, k \in \mathbb{N}.$$

Sei dazu $k \in \mathbb{N}$ gegeben. Dann gilt mit (33)

$$\begin{aligned} \textstyle\sum_{i+j=k} d_i c_j &= \textstyle\sum_{i+j=k} \left(\sum_{s+t=i} a_s b_t \right) c_j = \sum_{s+t+j=k} a_s b_t c_j \\ &= \textstyle\sum_{s+i=k} a_s \left(\sum_{t+j=i} b_t c_j \right) = \sum_{s+i=k} a_s d'_i. \end{aligned}$$

Damit ist (i) gezeigt. (ii) folgt aus

$$f \cdot g = \sum_{i\in\mathbb{N}}{}' \left(\sum_{i+j=k} a_i b_j \right) t^k = \sum_{i\in\mathbb{N}}{}' \left(\sum_{j+i=k} b_j a_i \right) t^k = g \cdot f.$$

(iii) folgt aus der Bilinearität von "$\cdot$", und für (iv) beachte man, daß nach (29) gilt $t^0 \cdot t^i = t^i$. Damit folgt wieder aus der Bilinearität von "$\cdot$"

$$t^0 \cdot f = \sum_{i\in\mathbb{N}}{}' a_i \big(t^0 \cdot t^i \big) = f.$$

□

Bemerkung 1.5
Für ein Polynom $f = \sum'_{i\in\mathbb{N}} a_i t^i \in R[t]$ und $\lambda \in R$ definieren wir

$$f(\lambda) := \sum_{i\in\mathbb{N}}{}' a_i \lambda^i.$$

Dann liest man aus (30)-(32) für $f, g \in R[t]$ und $\lambda, \mu \in R$ unmittelbar ab:

$(\mu f)(\lambda) = \mu\big(f(\lambda)\big), (f+g)(\lambda) = f(\lambda) + g(\lambda), (f \cdot g)(\lambda) = f(\lambda) \cdot g(\lambda)$ und $t^0(\lambda) = 1_R$.

Damit wird die Abbildung

$$\varphi : R[t] \to R^R : f \mapsto \big(R \to R : \lambda \mapsto f(\lambda) \big)$$

zum R-Algebrenhomomorphismus. Wir nennen die Unteralgebra $\mathrm{Im}(\varphi)$ von R^R die R-Algebra der *Polynomfunktionen* oder der *ganzrationalen Funktionen*.

Man beachte, daß φ *nicht* notwendigerweise injektiv ist, d. h. evt. ist φ kein Isomorphismus von $R[t]$ auf $\mathrm{Im}(\varphi)$, oder anders ausgedrückt, es besteht ein wesentlicher Unterschied zwischen Polynomen und Polynomfunktionen.

Beispiel 1.6
Wir betrachten nun den Fall $R = \mathbb{Z}_2 = \{[0], [1]\}$ und das Polynom $f = t^2 - t \in R[t]$. f ist offenbar nicht das Nullpolynom, jedoch gilt $f\big([0]\big) = [0]$ und $f\big([1]\big) = [0]$, also ist die Polynomfunktion $\varphi(f) : R \to R$ die Nullfunktion, d. h. $0 \neq f \in \mathrm{Ker}(\varphi)$.

Bemerkung 1.7
Die Abbildung

$$i : R \hookrightarrow R[t] : \lambda \mapsto \lambda t^0$$

ist ein R-Algebren*mono*morphismus. Dies erlaubt es uns, R mittels i mit $\mathrm{Im}(i)$ zu identifizieren und somit R als die Teilmenge der *konstanten Polynome* von $R[t]$ aufzufassen.

Beweis: Aus den R-Modulaxiomen folgt für $\lambda, \mu \in R$:

$$i(\lambda + \mu) = (\lambda + \mu)t^0 = \lambda t^0 + \mu t^0 = i(\lambda) + i(\mu).$$

sowie

$$i(\lambda\mu) = (\lambda\mu)t^0 = \lambda(\mu t^0) = \lambda i(\mu).$$

Da zudem t^0 die Eins des Ringes $R[t]$ ist, gilt auch:

$$i(\lambda\mu) = (\lambda\mu)t^0 = \lambda\big(\mu t^0\big) = \lambda\Big(t^0\big(\mu t^0\big)\Big) = \big(\lambda t^0\big) \cdot \big(\mu t^0\big) = i(\lambda) \cdot i(\mu).$$

Ebenfalls wieder aus den R-Modulaxiomen folgt $i(1) = 1 \cdot t^0 = t^0 = 1_{R[t]}$, und somit ist gezeigt, daß i ein R-Modulhomomorphismus ist. Es bleibt also zu zeigen, daß $\mathrm{Ker}(i)$ nur aus dem Nullelement besteht.

Für $\lambda \in \mathrm{Ker}(i)$ gilt $0 = i(\lambda) = \lambda t^0$, aber da t^0 linear unabhängig ist, folgt damit $\lambda = 0$. □

Definition 1.8
Sei $f = \sum_{i=0}^{n} a_i t^i \in R[t]$ mit $a_n \neq 0$, dann heißt $\deg(f) := n$ der *Grad* von f und $\mathrm{lc}(f) := a_n$ der *Leitkoeffizient* von f. Ist $\mathrm{lc}(f) = 1$, so nennen wir f *normiert*.

Für $f = 0 \in R[t]$ setzen wir $\deg(f) := -\infty$. Damit gilt:

$$\deg : R[t] \to \mathbb{N} \cup \{-\infty\}.$$

Beachte, ein Polynom f ist genau dann konstant, wenn $\deg(f) \leq 0$.

Lemma 1.9 (Gradformeln)
Seien $f, g \in R[t] \setminus \{0\}$. *Dann gelten:*

a. $\deg(f + g) \leq \max\big\{\deg(f), \deg(g)\big\}$.
b. $\deg(f \cdot g) \leq \deg(f) + \deg(g)$.
c. *Falls* $\deg(f) \neq \deg(g)$, *dann gilt* $\deg(f + g) = \max\big\{\deg(f), \deg(g)\big\}$.
d. *Falls* $\mathrm{lc}(f) \cdot \mathrm{lc}(g) \neq 0$, *dann gilt* $\deg(f \cdot g) = \deg(f) + \deg(g)$.

Beweis: Seien $f = \sum_{i=0}^{n} a_i t^i, g = \sum_{i=0}^{m} b_i t^i \in R[t]$ mit $a_n \neq 0 \neq b_m$. Wir setzen $a_i = 0$ für $i > n$ und $b_i = 0$ für $i > m$.

a. O. E. $n \geq m$. Dann gilt

$$f + g = \sum_{i=0}^{n} (a_i + b_i)t^i, \tag{34}$$

also $\deg(f + g) \leq n = \max\big\{\deg(f), \deg(g)\big\}$.

c. O. E. $n > m$. Dann gilt $a_n + b_n = a_n \neq 0$, und aus (34) folgt $\deg(f+g) = n = \max\{\deg(f), \deg(g)\}$.

b. Nach (32) gilt $f \cdot g = \sum'_{k\in\mathbb{N}} \big(\sum_{i=0}^{k} a_i b_{k-i}\big) t^k$ mit

$$\sum_{i=0}^{k} a_i b_{k-i} = \begin{cases} a_n b_m, & \text{falls } k = n+m, \\ \sum_{i=0}^{n} a_i b_{k-i} + \sum_{i=n+1}^{k} a_i b_{k-i} = 0, & \text{falls } k \geq n+m+1. \end{cases} \tag{35}$$

Also gilt $\deg(f \cdot g) \leq n + m = \deg(f) + \deg(g)$.

d. Für $a_n b_m = \mathrm{lc}(f)\,\mathrm{lc}(g) \neq 0$ folgt aus (35) $\deg(f \cdot g) = \deg(f) + \deg(g)$.

□

Beispiel 1.10 a. Sei $f = 2t + 1, g = -2t + 1 \in \mathbb{Z}[t]$, dann gilt $f + g = 2$, also $\deg(f+g) < \max\{\deg(f), \deg(g)\}$, aber $f \cdot g = -4t^2 + 1$ und somit $\deg(f \cdot g) = \deg(f) + \deg(g)$.

b. Betrachten wir hingegen $f = [2]t + [1], g = [-2]t + [1] \in \mathbb{Z}_4[t]$, dann gilt $f \cdot g = [-4]t^2 + [1] = [1]$ und somit $\deg(f \cdot g) < \deg(f) + \deg(g)$.

Da in einem Körper das Produkt zweier Elemente ungleich Null stets wieder ungleich Null ist, erhalten wir aus Teil d. unmittelbar folgendes Korollar.

Korollar 1.11 (Gradformel)
Für $f, g \in K[t] \setminus \{0\}$ *gilt* $\deg(f \cdot g) = \deg(f) + \deg(g)$.

Korollar 1.12
$K[t]^* = K^* = K \setminus \{0\} = \{f \in K[t] \mid \deg(f) = 0\}$.

Beweis: Wir wissen bereits, daß $K^* = K\setminus\{0\}$ und per definitionem gilt $K\setminus\{0\} = \{f \in K[t] \mid \deg(f) = 0\}$. Außerdem ist klar, daß jede Einheit in K auch eine Einheit in $K[t]$ ist.

Seien also $f, g \in K[t]$ gegeben mit $f \cdot g = 1$, dann gilt $\deg(f) + \deg(g) = \deg(1) = 0$, also $\deg(f) = \deg(g) = 0$, d. h. $f, g \in K \setminus \{0\} = K^*$. □

Korollar 1.13
$K[t]$ *ist nullteilerfrei.*

Beweis: Angenommen, es seien $f, g \in K[t] \setminus \{0\}$ mit $f \cdot g = 0$. Dann gilt $-\infty = \deg(f \cdot g) = \deg(f) + \deg(g) \geq 0$, was einen Widerspruch ergibt. □

Satz 1.14 (Allgemeine Division mit Rest)
Seien $f, g \in R[t] \setminus \{0\}$ *mit* $a = \mathrm{lc}(g) \in R$. *Dann gibt es Polynome* $q, r \in R[t]$ *und eine natürliche Zahl* $k \in \mathbb{N}$ *mit*

$$a^k \cdot f = q \cdot g + r \quad \textit{und} \quad \deg(r) < \deg(g). \tag{36}$$

Beweis: Wir führen den Beweis durch Induktion über $n = \deg(f)$.

Ist $\deg(f) < \deg(g)$, so setze man $k = 0$, $q = 0$ und $r = f$, und (36) ist erfüllt, vollkommen unabhängig von n.

Sei nun also $n = \deg(f) \geq \deg(g) = m$ und seien $f = \sum_{i=0}^{n} a_i t^i$ und $g = \sum_{i=0}^{m} b_i t^i$. Insbesondere ist dann $a = b_m$.

Im Falle $n = 0$ ist auch $m = 0$, $f = a_0$ und $g = b_0 = a$. (36) folgt mithin mit $q = a_0$, $r = 0$ und $k = 1$, da $af = a_0 g + 0$.

Sei nun also $n > 0$ und für alle Polynome $h \in R[t] \setminus \{0\}$ mit $\deg(h) < n$ gebe es bereits $q', r' \in R[t]$ und $k' \in \mathbb{N}$ mit

$$a^{k'} \cdot h = q' \cdot g + r' \quad \text{und} \quad \deg\left(r'\right) < \deg(g). \tag{37}$$

Setze $h := a \cdot f - a_n \cdot t^{n-m} \cdot g \in R[t]$. Falls $h = 0$, dann sind wir fertig mit $k = 1$, $q = a_n t^{n-m}$ und $r = 0$. Andernfalls wenden wir Induktion auf h an und erhalten q', r' und k', so daß (37) erfüllt ist. Setzen wir nun $r = r'$, $q = q' + a^{k'} a_n t^{n-m}$ und $k = k' + 1$, so erhalten wir

$$a^k \cdot f = a^{k'} \cdot h + a^{k'} \cdot a_n \cdot t^{n-m} \cdot g = q \cdot g + r$$

und $\deg(r) = \deg\left(r'\right) < \deg(g)$, wie gewünscht.

Die Behauptung folgt also mittels Induktion. □

Korollar 1.15 (Division mit Rest)
Seien $f, g \in K[t] \setminus \{0\}$. *Dann gibt es* eindeutig bestimmte *Polynome* $q, r \in K[t]$ *mit*

$$f = q \cdot g + r \quad \textit{und} \quad \deg(r) < \deg(g). \tag{38}$$

Beweis: Aus Satz 1.14 folgt, daß es ein $k \in \mathbb{N}$ und $q', r' \in K[t]$ gibt mit $\mathrm{lc}(g)^k \cdot f = q' \cdot g + r'$ und $\deg\left(r'\right) < \deg(g)$. Da $\mathrm{lc}(g) \neq 0$ und da K ein Körper ist, können wir $q = \frac{q'}{\mathrm{lc}(g)^k}$ und $r = \frac{r'}{\mathrm{lc}(g)^k}$ betrachten, und (38) ist erfüllt. Es bleibt also die Eindeutigkeit einer solchen Darstellung zu zeigen.

Seien dazu $q, q', r, r' \in K[t]$ mit $f = q \cdot g + r = q' \cdot g + r'$ und $\deg(r) \leq \deg\left(r'\right) < \deg(g)$. Dann gilt

$$\left(q - q'\right) \cdot g = r' - r. \tag{39}$$

Angenommen, $r \neq r'$, dann folgt aus (39), daß auch $q - q' \neq 0$, und es gilt mit Hilfe der Gradformeln in Lemma 1.9 und Korollar 1.11

$$\deg\left(r'\right) \geq \deg\left(r' - r\right) = \deg\Big(\left(q - q'\right) \cdot g\Big) \geq \deg\left(q - q'\right) + \deg(g) \geq \deg(g),$$

im Widerspruch zu $\deg\left(r'\right) < \deg(g)$.

Mithin ist also $r = r'$ und (39) impliziert dann, daß auch $\left(q - q'\right) \cdot g = 0$. Da nun $g \neq 0$ und da $K[t]$ nullteilerfrei ist, folgt $q = q'$. □

Beispiel 1.16
Seien $f = t^7 - 2t^6 + t^4 - t^3 + 2t^2 - 1, g = t^6 - 2t^5 + 2t^2 - 1 \in \mathbb{Q}[t]$ gegeben. Wir führen Polynomdivision durch

$$\begin{array}{l} (t^7 - 2t^6 \ + t^4 - \ t^3 + 2t^2 \quad\quad - 1) \ : \ (t^6 - 2t^5 + 2t^2 - 1) = t + \frac{r}{t^6 - 2t^5 + 2t^2 - 1} \\ \underline{\ t^7 - 2t^6 \quad\quad + 2t^3 \quad\quad - t} \\ \quad\quad\quad\quad t^4 - 3t^3 + 2t^2 + t - \ 1 =: r \end{array}$$

und erhalten $f = t \cdot g + (t^4 - 3t^3 + 2t^2 + t - 1)$.

Definition 1.17
Es sei B eine R-Algebra, $b \in B$ und $f = \sum'_{i\in\mathbb{N}} a_i t^i \in R[t]$. Wir setzen $f(b) := \sum'_{i\in\mathbb{N}} a_i b^i \in B$.

a. Gilt $f(b) = 0$, so heißt b eine *Nullstelle* von f in B.
b. Gilt $B = R$, $f \neq 0$ und gilt für $m \in \mathbb{N}$ $(t-b)^m \mid f$, aber $(t-b)^{m+1} \nmid f$, so nennen wir b eine Nullstelle mit *Vielfachheit* $\mathrm{mult}(f,b) = m$.
c. Gilt $R \subseteq B$ und gibt es $b_1, \ldots, b_r \in B$ mit $f = (t-b_1)\cdots(t-b_r)$, so sagen wir, f *zerfällt über* B *in Linearfaktoren*.

Beispiel 1.18
Sei $f = t^2 - 4 \in \mathbb{R}[t]$, $B = \mathrm{Mat}(2,\mathbb{R})$ und $b = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$, dann gilt

$$f(b) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}^2 - 4 \cdot \mathbb{1}_2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Also ist b eine Nullstelle von f in $\mathrm{Mat}(2,\mathbb{R})$.

Bemerkung 1.19 a. Nach Aufgabe 1.38 gibt es genau einen R-Algebrenhomomorphismus $\phi_b : R[t] \to B : t \mapsto b$. Für diesen gilt gerade $\phi_b(f) = f(b)$.
Insbesondere gilt für $f, g \in R[t]$ und $\lambda \in R$ also:

$$\begin{aligned} (\lambda f)(b) &= \lambda \cdot f(b), \\ (f+g)(b) &= f(b) + g(b), \\ (f\cdot g)(b) &= f(b)\cdot g(b), \\ \left(1_{R[t]}\right)(b) &= 1_B. \end{aligned}$$

b. Ist $b \in R$ eine Nullstelle von $f \in R[t]$ mit Vielfachheit $m > 0$, dann gilt sicher $(t-b) \mid f$, d. h. es gibt ein Polynom $q \in R[t]$ mit $f = q\cdot(t-b)$ und somit $f(b) = q(b)\cdot(b-b) = 0$. Also ist b in diesem Fall auch eine Nullstelle von f in R.
Ist umgekehrt $b \in R$ eine Nullstelle von $f \neq 0$ in R, so gilt nach dem folgenden Lemma 1.20, daß b eine Nullstelle mit Vielfachheit $m > 0$ ist.
c. Für ein konstantes Polynom $r \in R$ gilt $r(b) = r$ für jedes $b \in R$. Also besitzt ein konstantes Polynom genau dann eine Nullstelle in R, wenn es das Nullpolynom ist.

Lemma 1.20
Ist $f \in R[t]$ *und* $\lambda \in R$ *eine Nullstelle von* f *in* R*, dann gibt es ein Polynom* $q \in R[t]$ *mit* $f = q\cdot(t-\lambda)$.

Beweis: Setze $g := t - \lambda \in R[t]\setminus\{0\}$. Da der Leitkoeffizient von g Eins ist, gibt es nach Satz 1.14 Polynome $q, r \in R[t]$ mit

$$f = q\cdot(t-\lambda) + r$$

und $\deg(r) < \deg(t-\lambda) = 1$. Also ist $r \in R$ ein konstantes Polynom. Aber damit gilt

$$0 = f(\lambda) = q(\lambda)\cdot(\lambda-\lambda) + r(\lambda) = r,$$

und somit ist $f = q \cdot (t - \lambda)$. □

Beispiel 1.21
Sei $f = t^3 - 1 \in \mathbb{R}[t]$, dann gilt offenbar $f(1) = 1^3 - 1 = 0$. Polynomdivision liefert:

$$\begin{array}{rl} (t^3 & -1) : (t-1) = t^2 + t + 1. \\ \underline{t^3 - t^2} & \\ t^2 & \\ \underline{t^2 - t} & \\ t - & 1 \\ \underline{t - & 1} \\ - & \end{array}$$

Also gilt $f = (t^2 + t + 1) \cdot (t - 1)$.

Lemma 1.22
Ist $0 \neq f \in K[t]$ *mit* $\deg(f) = n$*, so hat* f *höchstens* n *Nullstellen in* K.

Beweis: Beweis durch Induktion nach n.

Ist $n = 0$, so ist f konstant, und nach Bemerkung 1.19 hat f keine Nullstelle.

Sei nun $n > 0$. Besitzt f keine Nullstelle, so sind wir fertig. Sei also λ eine Nullstelle von f. Nach Lemma 1.20 gibt es ein Polynom $q \in K[t]$ mit $f = q \cdot (t - \lambda)$. Aus Korollar 1.11 folgt, daß dann $\deg(q) = \deg(f) - 1 < n$ gilt. Nach Induktion hat also q höchstens $n - 1$ Nullstellen. Es reicht deshalb, die folgende Behauptung zu zeigen.
Behauptung: Ist N die Menge der Nullstellen von f, M die der Nullstellen von q und $N' = N \setminus \{\lambda\}$, so ist $N' \subseteq M$.

Sei dazu $\mu \in N'$, dann gilt

$$0 = f(\mu) = q(\mu) \cdot (\mu - \lambda).$$

Da $\mu - \lambda \neq 0$ und da K als Körper nullteilerfrei ist, muß mithin $q(\mu) = 0$ gelten, d. h. $\mu \in M$. □

Korollar 1.23
Gilt $\#K = \infty$*, so ist der Algebrenhomomorphismus*

$$\varphi : K[t] \to \mathrm{Im}(\varphi) \subseteq K^K$$

aus Bemerkung 1.5 ein Isomorphismus, sprich, wir können die Polynome über K *mit den Polynomfunktionen identifizieren.*

Beweis: Es reicht zu zeigen, daß φ injektiv ist, d. h. daß $\ker(\varphi) = \{0\}$.

Sei dazu $f \in \ker(\varphi)$ gegeben. Dann gilt $\varphi(f) : K \to K$ ist die Nullfunktion, d. h. das Polynom f hat unendlich viele Nullstellen. Aber dann kann wegen Lemma 1.22 f nur das Nullpolynom sein. □

Bemerkung 1.24 a. $f = t^2 + 1 \in \mathbb{R}[t]$ hat in $\mathbb{R}$ keine Nullstelle, zerfällt aber über $\mathbb{C}$ in Linearfaktoren, $f = (t - i)(t + i)$.

b. Ein Körper K heißt *algebraisch abgeschlossen*, wenn jedes Polynom $f \in K[t] \setminus K$ über K in Linearfaktoren zerfällt.
Ein zentraler Satz der Algebra, den wir hier jedoch nicht beweisen können, und der für gewöhnlich mit Mitteln der Analysis oder der Topologie bewiesen wird, ist der folgende
Fundamentalsatz der Algebra: $\mathbb{C}$ ist algebraisch abgeschlossen.

c. In der Algebra zeigt man, daß jeder Körper K in einem algebraisch abgeschlossenen Körper $\overline{K}$, dem *algebraischen Abschluß* enthalten ist. Ein Polynom $f \in K[t] \subseteq \overline{K}[t]$ können wir also als Polynom in $\overline{K}[t]$ auffassen, und als solches zerfällt es über $\overline{K}$ in Linearfaktoren. Diese Zerlegung ist in Beweisen häufig hilfreich.

Definition 1.25
Sei R nullteilerfrei und seien $a, b, d \in R$.

a. Wir definieren
$$a \mid b \ :\Leftrightarrow\ \exists\, c \in R \ :\ a \cdot c = b.$$
Wir sagen in diesem Falle a *teilt* b oder a ist ein *Teiler* von b.

b. d heißt *größter gemeinsamer Teiler* von a und b, falls gilt:
(i) $d \mid a$ und $d \mid b$, und
(ii) für alle $d' \in R$ mit $d' \mid a$ und $d' \mid b$ gilt $d' \mid d$.
Wir schreiben kurz $d = \mathrm{ggT}(a, b)$.

c. d heißt *kleinstes gemeinsames Vielfaches* von a und b, falls gilt:
(i) $a \mid d$ und $b \mid d$, und
(ii) für alle $d' \in R$ mit $a \mid d'$ und $b \mid d'$ gilt $d \mid d'$.
Wir schreiben kurz $d = \mathrm{kgV}(a, b)$.

d. Ist $0 \neq d \notin R^*$, so sagen wir, d ist *irreduzibel*, wenn aus $d = a \cdot b$ folgt $a \in R^*$ oder $b \in R^*$.

e. Ist $0 \neq d \notin R^*$, so sagen wir, d ist *prim*, wenn aus $d \mid a \cdot b$ folgt, $d \mid a$ oder $d \mid b$.

f. d heißt *quadratfrei*, falls für alle $p \in R \setminus R^*$ gilt $p^2 \not| \, d$.

g. Ein Ideal $I \trianglelefteq R$ der Form $I = \langle d \rangle_R = dR$ heißt ein *Hauptideal*. Ist jedes Ideal von R ein Hauptideal, so heißt R ein *Hauptidealring*.

Beispiel 1.26 a. Ein Polynom $f \in K[t]$ vom Grad Eins ist stets irreduzibel.
Denn für $g, h \in K[t]$ mit $f = g \cdot h$ muß gelten $1 = \deg(f) = \deg(g) + \deg(h)$, also o. E. $\deg(g) = 0$ und damit ist g eine Einheit.

b. Das Polynom $f = x^2 + 1 \in \mathbb{R}[x]$ ist irreduzibel.
Angenommen, das wäre nicht der Fall, dann würde es zwei Polynome $g, h \in \mathbb{R}[t]$ vom Grad mindestens eins geben mit $f = g \cdot h$. Aus der Gradformel würde dann aber $\deg(g) = \deg(h) = 1$ folgen, insbesondere gäbe es $\lambda, \mu \in \mathbb{R}$, $\lambda \neq 0$, mit $g = \lambda t + \mu$. Somit wäre $-\frac{\mu}{\lambda} \in \mathbb{R}$ eine Nullstelle von f in $\mathbb{R}$, insbesondere würde gelten $\left(\frac{\mu}{\lambda}\right)^2 = -1$, was nicht sein kann.

Bemerkung 1.27
In einem beliebigen nullteilerfreien kommutativen Ring mit Eins wird nicht

notwendig ein ggT oder kgV zu gegebenen Elementen existieren. Und wenn ein ggT bzw. kgV existiert, so werden sie nur bis auf Einheiten eindeutig bestimmt sein. Siehe Aufgabe 1.41, dort finden sich auch alternative Beschreibungen des ggT bzw. kgV.

Satz 1.28
Seien $f, g \in K[t] \setminus \{0\}$*, so gibt es genau ein* $0 \neq d \in K[t]$ *mit* $lc(d) = 1$ *und* $d = ggT(f, g)$.

Ferner gilt, es gibt Polynome $p, q \in K[t]$ *mit*

$$d = p \cdot f + q \cdot g. \tag{40}$$

Beweis: Ersetzt man in den Beweisen von Algorithmus I.6.3 und Korollar I.6.4 die Ausdrücke "$0 \leq r_i < r_i - 1$" durch $\deg(r_i) < \deg(r_{i-1})$, so lassen sich die Beweise unter Verwendung der Division mit Rest, Korollar 1.15, wörtlich übertragen. □

Aus dem Satz erhalten wir den folgenden Algorithmus zur Bestimmung des größten gemeinsamen Teilers zweier Polynome.

Algorithmus 1.29 (Erweiterter Euklidischer Algorithmus)
INPUT: $f, g \in K[t] \setminus \{0\}$.

OUTPUT: $\big[ggT(f, g), p, q \big]$ mit $ggT(f, g) = p \cdot f + q \cdot g$.

1. Schritt: Falls $\deg(f) < \deg(g)$, so vertausche f und g.
2. Schritt: $r = f \pmod g$ ("Rest der Division von f durch g")
3. Schritt:
```
if (r == 0)
{
  return ( [g/lc(g), 0, 1/lc(g)] );
}
else
{
  list l = ggT(g, r);
  return ([l[1],l[3],l[2] - l[3] * (f - r)/g]);
}
```

Beispiel 1.30
Seien $f = t^7 - 2t^6 + t^4 - t^3 + 2t^2 - 1, g = t^6 - 2t^5 + 2t^2 - 1 \in \mathbb{Q}[t]$ gegeben.

Wir setzen $r_0 := g$ und nach Beispiel 1.16 gilt

$$f = t \cdot g + (t^4 - 3t^3 + 2t^2 + t - 1),$$

also $r_1 = t^4 - 3t^3 + 2t^2 + t - 1$.

Polynomdivision von r_0 durch r_1 liefert:

$$\begin{array}{rl}
(t^6 - 2t^5 \qquad\quad + 2t^2 \qquad - 1) & : \ (t^4 - 3t^3 + 2t^2 + t - 1) = t^2 + t + 1. \\
\underline{t^6 - 3t^5 + 2t^4 + \ t^3 - \ t^2} & \\
\qquad t^5 - 2t^4 - \ t^3 + 3t^2 \qquad - \ 1 & \\
\underline{\quad t^5 - 3t^4 + 2t^3 + \ t^2 - t \qquad\quad} & \\
\qquad\quad t^4 - 3t^3 + 2t^2 + t - \ 1 & \\
\underline{\quad t^4 - 3t^3 + 2t^2 + t - \ 1} & \\
\qquad\qquad\qquad\qquad\qquad 0 =: r_2 &
\end{array}$$

Also ist $\mathrm{ggT}(f, g) = r_1 = t^4 - 3t^3 + 2t^2 + t - 1$.

Außerdem gilt

$$\mathrm{ggT}(f, g) = \mathrm{ggT}(g, r_1) = 0 \cdot g + 1 \cdot r_1 = 1 \cdot f + (0 - 1 \cdot t) \cdot g = f - t \cdot g.$$

Definition 1.31
Für ein Polynom $f = \sum_{i=0}^{n} a_i t^i \in K[t]$ definieren wir die *formale Ableitung* $D(f) = \sum_{i=1}^{n} i a_i t^{i-1} \in K[t]$.

Bemerkung 1.32 a. Hat der Körper K die Charakteristik Null, so folgt aus $D(f) = 0$ unmittelbar, daß f Grad kleiner gleich Null hat, also konstant ist. Ist hingegen $\mathrm{char}(K) = p > 0$, so gilt zum Beispiel $D\left(x^p\right) = p x^{p-1} = 0$, so daß die Aussage nicht mehr gilt.

b. Wir überlassen es dem Leser, mittels Koeffizientenvergleichs zu zeigen, daß für die formale Ableitung die *Produktregel* gilt, d. h. seien $f, g \in K[t]$, dann gilt

$$D(f \cdot g) = f \cdot D(g) + D(f) \cdot g.$$

Definition 1.33
Es sei $K \subseteq L$ eine Körpererweiterung und $b \in L$.

a. Falls es ein Polynom $0 \neq f \in K[t]$ gibt mit $f(b) = 0 \in L$, dann heißt b *algebraisch* über K.
b. Ist b algebraisch über K, so betrachten wir den K-Algebrenhomomorphismus $\phi_b : K[t] \to L : g \mapsto g(b)$ aus Bemerkung 1.19. Dann besitzt das Ideal $0 \neq \mathrm{Ker}(\phi_b) \trianglelefteq K[t]$ einen eindeutig bestimmten normierten Erzeuger μ_b, den wir das *Minimalpolynom* von b nennen.

Bemerkung 1.34
Man beachte, daß für ein algebraisches Element $b \in L$ genau dann $\deg(\mu_b) = 1$ ist, wenn $b \in K$ gilt.

Lemma 1.35
Es sei $K \subseteq L$ eine Körpererweiterung und $b \in L$ sei algebraisch über K.

Ein Polynom $q \in K[t] \setminus \{0\}$ ist genau dann das Minimalpolynom von b, wenn gilt:

(i) $\mathrm{lc}(q) = 1$,
(ii) $q(b) = 0$, *und*

(iii) *es gibt* kein *Polynom* $p \in K[t] \setminus \{0\}$ *mit* $\deg(p) < \deg(q)$ *und* $p(b) = 0$.

Beweis: μ_b erfüllt die Bedingungen (i)-(iii) offenbar. Sei also umgekehrt $q \in K[t] \setminus \{0\}$ ein Polynom, das den Bedingungen (i)-(iii) genügt, dann ist q wegen (ii) und (iii) ein Element ungleich Null von minimalem Grad in $\mathrm{Ker}(\phi_b)$ und damit nach dem Hinweis zu Aufgabe 1.39 ein Erzeuger des Ideals. Nach (i) ist q aber auch normiert und somit das Minimalpolynom von b. □

Satz 1.36
Ist $\mathrm{char}(K) = 0$*, dann sind für ein Polynom* $f \in K[t] \setminus K$ *die folgenden Aussagen gleichwertig:*

a. f *ist quadratfrei.*
b. *Eins ist ein* ggT $\big(f, D(f)\big)$.
c. f *zerfällt über dem algebraischen Abschluß* $\overline{K}$ *von* K *in paarweise verschiedene Linearfaktoren, d. h.* f *hat keine mehrfache Nullstelle in* $\overline{K}$.

Beweis: a. ⇒b.: Angenommen, $p \in K[t]$ ist ein irreduzibles Polynom, das sowohl f als auch $D(f)$ teilt. Da f quadratfrei ist, gibt es ein $h \in K[t]$ mit $p \not| h$, aber $f = p \cdot h$. Aus der Produktregel folgt dann

$$D(p) \cdot h = D(f) - p \cdot D(h).$$

Da nun p ein Teiler von $D(f)$ ist, teilt p also auch $D(p) \cdot h$. Da ferner p nach Aufgabe 1.43 prim ist und nach Wahl von h kein Teiler von h ist, ist mithin p ein Teiler von $D(p)$, im Widerspruch zu $\deg(p) > \deg\big(D(p)\big) \geq 0$.

b. ⇒c.: Hat f einen mehrfachen Linearfaktor über $\overline{K}$, dann existiert ein $b \in \overline{K}$, ein $r > 1$ und ein $h \in \overline{K}[t]$ mit

$$f = (t - b)^r \cdot h,$$

wobei $h(b) \neq 0$. Mittels der Produktregel folgt dann aber, daß $D(f)(b) = 0$, und damit $D(f), f \in \mathrm{Ker}(\phi_b) = \mu_b K[t]$. Also ist μ_b ein (nicht-konstanter) gemeinsamer Teiler von f und $D(f)$, im Widerspruch zur Voraussetzung.

c. ⇒a.: Wäre f nicht quadratfrei, hätte f einen irreduziblen Faktor $p \in K[t]$ mit Vielfachheit größer gleich zwei und dessen Nullstellen in $\overline{K}$ wären mehrfache Nullstellen von f, im Widerspruch zur Voraussetzung. □

Bemerkung 1.37
Die Bedeutung des Satzes liegt darin, daß man den größten gemeinsamen Teiler zweier Polynome $f, g \in K[t]$ mit Hilfe des Euklidischen Algorithmus bestimmen kann. Man kann also, ohne die Nullstellen von f in $\overline{K}$ zu kennen, feststellen, ob f mehrfache Nullstellen hat oder nicht, und zwar exakt!

Ist z. B. $f \in \mathbb{Q}[t]$ gegeben (selbst bei Polynomen aus $\mathbb{R}[t]$ sind in der Praxis die Koeffizienten in der Regel aus $\mathbb{Q}$), dann kann man durch Berechnung des ggT von f und $D(f)$ in $\mathbb{Q}$ exakt feststellen, ob f mehrfache komplexe Nullstellen hat oder nicht. Durch numerische Nullstellenberechnung ist dies mitunter

gar nicht möglich, da mehrfache Nullstellen aufgrund von Rundungsfehlern oft als verschieden berechnet werden.

Das Kriterium für Quadratfreiheit wird im Zusammenhang mit der Diagonalisierbarkeit von Matrizen von Bedeutung sein. (Vgl. Algorithmus 3.36.)

AUFGABEN

Aufgabe 1.38

Es sei A eine R-Algebra und $a \in A$. Wir nennen A *frei in* a, falls für jede R-Algebra B und jedes $b \in B$ gilt, daß es *genau einen* R-Algebrenhomomorphismus $\phi_b : A \to B$ gibt mit $\phi_b(a) = b$. Zeige:

a. $R[t]$ ist frei in t.
b. Ist A frei in a, so gibt es genau einen *Isomorphismus* $\phi_b : R[t] \to A$ mit $\phi_b(t) = a$.

Aufgabe 1.39

$K[t]$ ist ein Hauptidealring.

Hinweis: Betrachte in I ein Element ungleich Null von minimalem Grad.

Aufgabe 1.40

Es seien $f, g \in K[t]$, $I \trianglelefteq K[t]$. Zeige:

a. $\deg(f) = 0 \quad \Leftrightarrow \quad f \in K^* \quad \Leftrightarrow \quad fK[t] = K[t]$.
b. $fK[t] \subseteq gK[t] \quad \Rightarrow \quad \deg(f) \geq \deg(g)$.
c. Falls $fK[t] \subseteq gK[t]$ und $\deg(f) > \deg(g)$, dann ist $fK[t] \subsetneqq gK[t]$.
d. $fK[t] = gK[t] \quad \Leftrightarrow \quad \exists\, c \in K^* \,:\, f = cg$.
e. Es gibt genau ein normiertes Polynom $f \in K[t]$ mit $I = fK[t]$.

Aufgabe 1.41

Es sei R nullteilerfrei und $a, b, d, d' \in R$.

a. Zeige, die folgenden Aussagen sind gleichwertig:
   (i) $a \mid b$,
   (ii) $bR \subseteq aR$.
b. Zeige, falls R ein Hauptidealring ist, sind die folgenden Aussagen gleichwertig:
   (i) $d = \mathrm{ggT}(a, b)$,
   (ii) $dR = aR + bR$.
c. Sind d und d' beides ggT's von a und b, so gibt es eine Einheit $c \in R^*$ mit $d = cd'$.
d. Zeige, die folgenden Aussagen sind gleichwertig
   (i) $d = \mathrm{kgV}(a, b)$,
   (ii) $dR = aR \cap bR$.
e. Sind d und d' beides kgV's von a und b, so gibt es eine Einheit $c \in R^*$ mit $d = cd'$.
f. Ist d prim, so ist d irreduzibel.
g. Folgere die Behauptung von Satz 1.28 aus der Tatsache, daß $R = K[t]$ ein Hauptidealring ist und den obigen Aussagen.

Aufgabe 1.42 (ggT)
Es seien $f_1, \dots, f_n \in K[t]$, $n \geq 2$. Ein Polynom $g \in K[t]$ heißt größter gemeinsamer Teiler von $f_1, \dots, f_n$ (kurz: $\mathrm{ggT}(f_1, \dots, f_n)$), falls gilt:

(i) $g \mid f_i$ für alle $i = 1, \dots, n$, und
(ii) für alle $g' \in K[t]$ mit $g' \mid f_i$ für alle $i = 1, \dots, n$ gilt $g' \mid g$.

a. Zeige, es gilt $\mathrm{ggT}(f_1, \dots, f_n) = \mathrm{ggT}\left(\mathrm{ggT}(f_1, \dots, f_{n-1}), f_n\right)$, falls $n > 2$.[1]
b. Zeige, $g = \mathrm{ggT}(f_1, \dots, f_n)$ genau dann, wenn $gK[t] = f_1K[t] + \dots + f_nK[t] := \{f_1h_1 + \dots f_nh_n \mid h_1, \dots, h_n \in K[t]\}$.
c. Genau dann besitzt die *diophantische Gleichung* $f_1x_1 + \dots + f_nx_n = h$ (mit $h \in K[t]$) eine Lösung $(x_1, \dots, x_n) \in K[t]^n$, wenn $\mathrm{ggT}(f_1, \dots, f_n) \mid h$.

Aufgabe 1.43
Es sei R ein Hauptidealring.

Zeige, daß ein Element $a \in R$ genau dann irreduzibel ist, wenn a prim ist.

Insbesondere gilt für ein Polynom $f \in K[t]$, f ist irreduzibel genau dann, wenn f prim ist.

Aufgabe 1.44 (ZPE-Ring)
Es sei R nullteilerfrei. Zeige, die folgenden Aussagen sind gleichwertig:

a. Ist $0 \neq a \in R \setminus R^*$, dann gibt es irreduzible Elemente $q_1, \dots, q_r$ mit $a = q_1 \cdots q_r$. Außerdem gilt für zwei Darstellungen von a der Form $a = q_1 \cdots q_r = p_1 \cdots p_s$ mit q_i und p_i irreduzibel für alle i, daß $r = s$ und nach Umnumerierung unterscheiden sich q_i und p_i nur noch um eine Einheit.
b. Ist $0 \neq a \in R \setminus R^*$, dann gibt es irreduzible Elemente $q_1, \dots, q_r$ mit $a = q_1 \cdots q_r$, und jedes irreduzible Element von R ist prim.
c. Ist $0 \neq a \in R \setminus R^*$, dann gibt es Primelemente $p_1, \dots, p_r$ mit $a = p_1 \cdots p_r$.

R heißt in diesem Fall ein *ZPE-Ring* oder ein *faktorieller Ring*, und man nennt eine Zerlegung der Form $a = p_1 \cdots p_r$ eine Primfaktorzerlegung von a.

Aufgabe 1.45
$K[t]$ ist ein ZPE-Ring, d. h. für jedes Polynom $f \in K[t] \setminus K$ gibt es eindeutig bestimmte normierte, paarweise verschiedene, irreduzible Polynome $p_1, \dots, p_r$ und positive natürliche Zahlen $\nu_1, \dots, \nu_r$, so daß

$$f = \mathrm{lc}(f) \cdot p_1^{\nu_1} \cdots p_r^{\nu_r}. \tag{41}$$

Man nennt die Zerlegung in (41) die eindeutige *Primfaktorzerlegung* von f.

Aufgabe 1.46
Es seien $f, g \in K[t]$ mit eindeutiger Primfaktorzerlegung $f = \mathrm{lc}(f) \cdot p_1^{\nu_1} \cdots p_r^{\nu_r}$ und $g = \mathrm{lc}(g) \cdot p_1^{\lambda_1} \cdots p_r^{\lambda_r}$, wobei wir hier zulassen wollen, daß ν_i bzw. λ_i null sind. Wir setzen $\mu_i := \min\{\nu_i, \lambda_i\}$ und $\kappa_i := \max\{\nu_i, \lambda_i\}$. Dann gilt:

[1] Beachte, ein ggT ist nur bis auf Einheiten eindeutig bestimmt, mithin soll die Aussage bedeuten, wenn $h \in K[t]$ ein ggT von $f_1, \dots, f_{n-1}$ ist, dann gilt: genau dann ist $g \in K[t]$ ein ggT von h und f_n, wenn g auch ein ggT von $f_1, \dots, f_n$ ist.

a. $p_1^{\mu_1} \cdots p_r^{\mu_r}$ ist ein $\mathrm{ggT}(f, g)$.
b. $p_1^{\kappa_1} \cdots p_r^{\kappa_r}$ ist ein $\mathrm{kgV}(f, g)$.

Aufgabe 1.47
Es $f \in K[t]$ mit eindeutiger Primfaktorzerlegung $f = \mathrm{lc}(f) \cdot p_1^{\nu_1} \cdots p_r^{\nu_r}$, $g \in K[t]$.

a. f ist genau dann quadratfrei, wenn $\nu_1 = \ldots = \nu_r = 1$.
b. f und g sind genau dann beide quadratfrei, wenn jedes $\mathrm{kgV}(f, g)$ quadratfrei ist.

Aufgabe 1.48
Es sei $K \subseteq L$ eine Körpererweiterung, und es seien $f, g \in K[t] \subseteq L[t]$ so, daß es ein $h \in L[t]$ gibt mit $f = g \cdot h$. Zeige, dann gilt bereits $h \in K[t]$.

Aufgabe 1.49
Es sei $f \in K[t]$ mit $\deg(f) \in \{2, 3\}$. Zeige, f ist genau dann irreduzibel über K, wenn f keine Nullstelle in K besitzt.

Aufgabe 1.50
Es seien $a, b \in \mathbb{Z}$ mit $a(a+2) \neq b(b+2)$. Zeige, $f = t^3 + at^2 + bt + 1 \in \mathbb{Q}[t]$ ist irreduzibel über $\mathbb{Q}$.

Aufgabe 1.51
Bestimme alle irreduziblen Polynome in $\mathbb{Z}_2[t]$ vom Grad kleiner als sechs.

Aufgabe 1.52
Es sei $0 \neq n \in \mathbb{N}$. Zeige, das Polynom $t^n - nt + n \in \mathbb{C}[t]$ besitzt in $\mathbb{C}$ keine mehrfache Nullstelle.

Aufgabe 1.53
Es sei $f \in \mathbb{R}[t]$ ein normiertes Polynom vom Grad zwei ohne Nullstelle in $\mathbb{R}$. Zeige, dann gibt es eine komplexe Zahl λ mit $f = t^2 + \left(\lambda + \overline{\lambda}\right)t + \lambda \cdot \overline{\lambda}$.

Aufgabe 1.54 (Funktionenkörper)
Es sei K ein Körper und $K[t]$ der Polynomring in einer Variablen. Wir definieren auf $K[t] \times (K[t] \setminus \{0\})$ eine Äquivalenzrelation durch

$$(f, g) \sim \left(f', g'\right) \quad :\Leftrightarrow \quad fg' = f'g$$

für $(f, g), \left(f', g'\right) \in K[t] \times (K[t] \setminus \{0\})$. Die zu (f, g) gehörende Äquivalenzklasse bezeichnen wir mit $\frac{f}{g}$ und die Menge der Äquivalenzklassen mit

$$K(t) = \left\{ \tfrac{f}{g} \;\middle|\; f, g \in K[t], g \neq 0 \right\}.$$

Zeige die folgenden Aussagen:

a. $\sim$ ist eine Äquivalenzrelation auf $K[t] \times (K[t] \setminus \{0\})$.
b. Durch
$$\tfrac{f}{g} + \tfrac{f'}{g'} := \tfrac{fg' + f'g}{gg'} \quad \text{und} \quad \tfrac{f}{g} \cdot \tfrac{f'}{g'} := \tfrac{ff'}{gg'}$$
für $\frac{f}{g}, \frac{f'}{g'} \in K(t)$ werden eine *wohldefinierte* Addition und Multiplikation auf $K(t)$ definiert.
c. $\left(K(t), +, \cdot\right)$ ist ein Körper, der sog. *Quotientenkörper* von $K[t]$ oder auch *Funktionenkörper* genannt.[2]

[2] Die Konstruktion von $\mathbb{Q}$ aus $\mathbb{Z}$ erfolgte ganz analog - vgl. Beispiel I.2.30.

d. $i : K[t] \to K(t) : f \mapsto \frac{f}{1}$ ist ein Ringmonomorphismus, insbesondere können wir $K[t]$ also als Unterring von $K(t)$ auffassen.

Aufgabe 1.55
Es sei R ein kommutativer Ring mit Eins und $n \geq 2$, dann definieren wir den Polynomring $R[x_1, \ldots, x_n]$ rekursiv als

$$R[x_1, \ldots, x_n] = \big(R[x_1, \ldots, x_{n-1}]\big)[x_n]$$

Wir nennen $R[x_1, \ldots, x_n]$ den *Polynomring in* n *Veränderlichen*, und die Elemente von $R[x_1, \ldots, x_n]$ nennen wir *Polynome in* n *Veränderlichen*.

Durch die natürlichen Inklusionen $R[x_1, \ldots, x_{n-1}] \hookrightarrow R[x_1, \ldots, x_n]$ aus Bemerkung 1.7 haben wir durch Rekursion auch $R \hookrightarrow R[x_1, \ldots, x_n]$ und $R[x_1, \ldots, x_n]$ ist ein R-Modul, mehr noch, sogar eine kommutative R-Algebra mit 1_R als Eins.

Wir setzen für $\nu = (\nu_1, \ldots, \nu_n) \in \mathbb{N}^n$ und nun

$$|\nu| = \nu_1 + \ldots + \nu_n \quad \text{und} \quad x^\nu := x_1^{\nu_1} \cdots x_n^{\nu_n},$$

wobei x ein neues Symbol ist. Zeige:

a. $\big(x^\nu \;\big|\; \nu \in \mathbb{N}^n\big)$ ist eine Basis von $R[x_1, \ldots, x_n]$ als R-Modul.
Insbesondere, jedes Polynom $f \in R[x_1, \ldots, x_n]$ hat eine eindeutige Darstellung als endliche Linearkombination der Form

$$f = \sum_{\nu \in \mathbb{N}^n}{}' a_\nu x^\nu = \sum_{|\nu|=0}^{d} a_\nu x_1^{\nu_1} \cdots x_n^{\nu_n}.$$

Gilt dabei $a_\nu \neq 0$ für ein $|\nu| = d$, so heißt $\deg(f) := d$ der *Grad* des Polynoms f. Die Basiselemente $x^\nu = x_1^{\nu_1} \cdots x_n^{\nu_n}$ nennen wir *Monome*.

b. Man leite die offensichtlichen Formeln für das Produkt und die Summe zweier Polynome sowie für das skalare Vielfache eines Polynoms her.

c. Ist $\lambda = (\lambda_1, \ldots, \lambda_n)^t \in R^n$, so ist der Einsetzhomomorphismus

$$\phi_\lambda : R[x_1, \ldots, x_n] \to R : f = \sum_{|\nu|=0}^{d} a_\nu x_1^{\nu_1} \cdots x_n^{\nu_n} \mapsto f(\lambda_1, \ldots, \lambda_n) := \sum_{|\nu|=0}^{d} a_\nu \lambda_1^{\nu_1} \cdots \lambda_n^{\nu_n}$$

ein R-Algebrenhomomorphismus.

d. Ein Polynom $f \in \mathbb{R}[x_1, \ldots, x_n]$ heißt *homogen vom Grad* d, wenn für alle $\lambda \in \mathbb{R}$ gilt $f(\lambda \cdot x) = \lambda^d \cdot f$.
Zeige, $f \in \mathbb{R}[x_1, \ldots, x_n]$ ist genau dann homogen vom Grad d, wenn $f = \sum_{|\nu|=d} a_\nu x^\nu$, d. h. wenn in der Darstellung von f nur Monome vom Grad d mit Koeffizienten ungleich Null vorkommen können.

e. Man leite Gradformeln für $R[x_1, \ldots, x_n]$ her.

f. Ist R nullteilerfrei, dann ist auch $R[x_1, \ldots, x_n]$ nullteilerfrei.

g. Für $n \geq 2$ ist $R[x_1, \ldots, x_n]$ nie ein Hauptidealring.

h. $R[x_1, \ldots, x_n]$ ist ein faktorieller Ring.

Aufgabe 1.56
Schreibe eine Singular-Prozedur `extggT`, die zwei Polynome f und g einliest und eine Liste ausgibt, die den normierten größten gemeinsamen Teiler d von f und g enthält sowie zwei Polynome p und q mit $d = pf + qg$.

2. Eigenwerte und das charakteristische Polynom

Es sei V ein n-dimensionaler K-Vektorraum über einem Körper K und $f : V \to V$ ein Endomorphismus von V. In diesem und den nächsten Abschnitten wollen wir die folgende Frage untersuchen:

> Gibt es eine Basis B von V so, daß die Matrix $M_B^B(f)$ eine besonders einfache Gestalt hat, d. h., daß man wichtige Eigenschaften der Abbildung f direkt aus der Matrix $M_B^B(f)$ ablesen kann?

Ist $A \in \mathrm{Mat}(n, K)$ die Matrix von f bezüglich irgendeiner Basis von V, dann transformiert sich A bei Übergang zu einer anderen Basis mittels $T \circ A \circ T^{-1}$ mit einer Matrix $T \in \mathrm{Gl}_n(K)$, die den Basiswechsel beschreibt - vgl. Satz II.5.11. Deshalb läßt sich obige Frage auch rein matrizentheoretisch wie folgt stellen:

> Gibt es zu gegebener Matrix $A \in \mathrm{Mat}(n, K)$ eine invertierbare Matrix $T \in \mathrm{Gl}_n(K)$ so, daß $T \circ A \circ T^{-1}$ eine besonders einfache Gestalt hat und wichtige Eigenschaften von A sofort sichtbar sind?

Da die Matrix A die Abbildung

$$f_A : K^n \to K^n : x \mapsto Ax$$

bestimmt, sind der Abbildungsgesichtspunkt und der Matrizengesichtspunkt äquivalent.

Eine ähnliche Frage wie oben haben wir übrigens schon in Abschnitt II.5 beantwortet. Nach Korollar II.5.19 existieren zu $A \in \mathrm{Mat}(n, K)$ Matrizen $S, T \in \mathrm{Gl}_n(K)$ mit

$$S \circ A \circ T^{-1} = \begin{pmatrix} \mathbb{1}_r & 0 \\ 0 & 0 \end{pmatrix},$$

wobei $r = \mathrm{rang}(A)$, d. h. A ist äquivalent der rechten Matrix. Wir erinnern uns, daß zwei Matrizen $A, B \in \mathrm{Mat}(n, K)$ *ähnlich* oder *konjugiert* heißen, wenn es eine einzige Matrix $T \in \mathrm{Gl}_n(K)$ gibt mit $B = T \circ A \circ T^{-1}$. Die Klassifikation der Matrizen bezüglich Ähnlichkeit ist offensichtlich schwieriger, als diejenige bezüglich Äquivalenz, da wir nur eine statt zwei Matrizen zur Transformation frei haben.

Wir führen jetzt einige grundlegende Begriffe ein, die nicht nur für die Klassifikation von Abbildungen von Bedeutung sind.
Generalvoraussetzung: Es sei K ein Körper, V ein K-Vektorraum und $f \in \mathrm{End}_K(V)$. A bezeichne stets eine Matrix in $\mathrm{Mat}(n, K)$.

Definition 2.1 a. $\lambda \in K$ heißt *Eigenwert* von f, falls es ein $x \in V \setminus \{0\}$ gibt, so daß $f(x) = \lambda x$ gilt.
b. $x \in V \setminus \{0\}$ mit $f(x) = \lambda x$ heißt dann ein *Eigenvektor* zum Eigenwert λ.
c. Die Menge $\mathrm{Eig}(f, \lambda) := \{x \in V \mid f(x) = \lambda x\}$ heißt der *Eigenraum* zum Eigenwert λ.
d. Die Menge $\sigma(f) := \{\lambda \in K \mid \lambda \text{ ist Eigenwert von } f\}$ der Eigenwerte von f heißt das *Spektrum* von f.

Entsprechend definieren wir für Matrizen.

Definition 2.2 a. $\lambda \in K$ heißt *Eigenwert* von A, falls es ein $x \in K^n \setminus \{0\}$ gibt, so daß $Ax = \lambda x$ gilt.
b. $x \in K^n \setminus \{0\}$ mit $Ax = \lambda x$ heißt dann ein *Eigenvektor* zum Eigenwert λ.
c. Die Menge $\mathrm{Eig}(A, \lambda) := \{x \in V \mid Ax = \lambda x\}$ heißt der *Eigenraum* zum Eigenwert λ.
d. Die Menge $\sigma(A) := \{\lambda \in K \mid \lambda \text{ ist Eigenwert von } A\}$ der Eigenwerte von A heißt das *Spektrum* von A.

Bemerkung 2.3
Da $f(x) = \lambda x$ für $x \in V$ und $\lambda \in K$ genau dann erfüllt ist, wenn x im Kern der linearen Abbildung $f - \lambda \mathrm{id}_V \in \mathrm{End}_K(V)$ liegt, gilt also

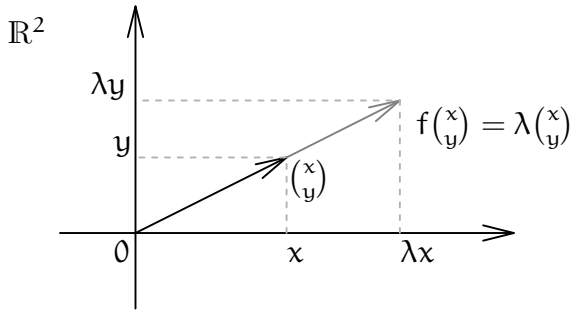
$$\mathrm{Eig}(f, \lambda) = \mathrm{Ker}(f - \lambda \mathrm{id}_V).$$

Analog erhält man:

$$\mathrm{Eig}(A, \lambda) = \mathrm{Lös}(A - \lambda \mathbb{1}_n, 0) = \mathrm{Ker}(f_A - \lambda \mathrm{id}_V).$$

Man beachte, daß der Eigenraum $\mathrm{Eig}(f, \lambda)$ von f zum Eigenwert λ per definitionem f-invariant ist.

Ist λ Eigenwert von f mit Eigenvektor x, so bedeutet das anschaulich, daß f in *Richtung* von x durch Multiplikation mit λ wirkt. Diese Anschauung liefert im Fall $V = \mathbb{R}^n$ und $\lambda > 0$, daß f den Vektor x um den Faktor λ streckt, falls $\lambda > 1$, und um den Faktor λ staucht, falls $\lambda < 1$.



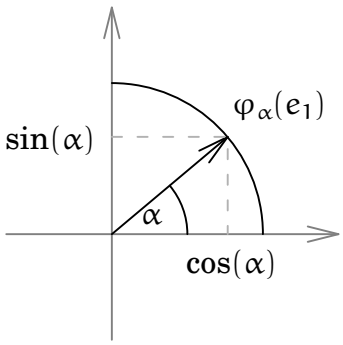
Im Falle etwa von $V = \mathbb{C}^1$ ist diese Vorstellung allerdings schon zu relativieren, denn Multiplikation von $x \in \mathbb{C}^1$ mit $\lambda = re^{i\alpha} \in \mathbb{C}$, $r > 0$, $0 \leq \alpha < 2\pi$, bedeutet eine Streckung bzw. Stauchung von x um den Faktor r und eine Drehung um den Winkel α, wenn wir $\mathbb{C}^1$ mit $\mathbb{R}^2$ identifizieren.

Dies führt uns sofort auf die Frage nach der Existenz von Eigenvektoren.

Beispiel 2.4 a. Ist $\dim_K(V) = 1$, so ist jeder Vektor ungleich Null ein Eigenvektor von f, da f schlicht die Multiplikation mit einer Konstanten ist.

b. Ist $\dim_K(V) \geq 2$, so braucht f hingegen keine Eigenvektoren zu besitzen. Dabei hängt die Frage der Existenz wesentlich vom Grundkörper K ab. Betrachte etwa die Drehung $\varphi_\alpha : \mathbb{R}^2 \to \mathbb{R}^2$ um den Winkel $\alpha \in \mathbb{R}$ aus Beispiel I.3.11 mit Matrixdarstellung bezüglich der kanonischen Basis $E = (e_1, e_2)$

$$A := M_E^E(\varphi_\alpha) = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}.$$



Aus einer rein geometrischen Betrachtung folgt unmittelbar, daß φ_α bzw. A nur dann einen Eigenvektor besitzen können, wenn α ein ganzzahliges Vielfaches von π ist.
Dies sieht man auch sofort, wenn man die Abbildung φ_α als Abbildung von $\mathbb{C}$ nach $\mathbb{C}$ auffaßt. Dann ist φ_α die Multiplikation mit $e^{i\alpha}$, und $e^{i\alpha}$ ist nur für ganzzahlige Vielfache von π reell.
Als $\mathbb{C}$-lineare Abbildung von $\mathbb{C}$ nach $\mathbb{C}$ hat A damit aber Eigenvektoren zum Eigenwert $e^{i\alpha}$, und zwar jeden Vektor $z \neq 0$. Wir werden weiter unten sehen, daß jeder Endomorphismus eines komplexen Vektorraumes Eigenvektoren besitzt.

Eine besonders angenehme Situation ist die, daß V eine Basis aus Eigenvektoren hat.

Definition 2.5
Es sei $\dim_K(V) < \infty$.

a. f heißt *diagonalisierbar* (bzw. *trigonalisierbar*), falls es eine Basis B von V gibt, so daß $M_B^B(f)$ eine Diagonalmatrix (bzw. eine obere Dreiecksmatrix) ist.

b. A heißt *diagonalisierbar* (bzw. *trigonalisierbar*), falls es eine Matrix $T \in Gl_n(K)$ gibt, so daß $T \circ A \circ T^{-1}$ eine Diagonalmatrix (bzw. eine obere Dreiecksmatrix) ist.

Lemma 2.6
Es sei $\dim_K(V) < \infty$. *Dann sind gleichwertig:*

a. V *(bzw.* K^n*) besitzt eine Basis aus Eigenvektoren von* f *(bzw.* A*).*

b. f *(bzw.* A*) ist diagonalisierbar.*

In diesem Falle sind die Diagonalelemente genau die Eigenwerte von f *(bzw.* A*).*

Beweis: Ist $B = (x_1, \ldots, x_n)$ eine beliebige Basis von V, so bestehen die Spalten von $M_B^B(f)$ aus den Koordinatenvektoren $M_B\big(f(x_i)\big)$ von $f(x_i)$ bezüglich B, $i = 1, \ldots, n$. Damit gilt:

$$f(x_i) = \lambda_i x_i \quad \Leftrightarrow \quad M_B^B(f)e_i = M_B\big(f(x_i)\big) = \lambda_i e_i.$$

Daraus folgt die Behauptung für den Fall von linearen Abbildungen. Der Fall einer Matrix A läßt sich auf die Abbildung f_A zurückführen. □

Bemerkung 2.7
Aus dem Beweis des obigen Lemmas 2.6 folgt zudem, daß im Fall $1 \leq \dim_K(V) < \infty$ für eine Basis B von V gilt:

$$\sigma(f) = \sigma\big(M_B^B(f)\big).$$

Im allgemeinen wird V keine Basis aus Eigenvektoren besitzen. Es gilt aber stets die folgende Proposition.

Proposition 2.8
Sind $x_1, \ldots, x_r \in V$ *(bzw.* K^n*) Eigenvektoren von* f *(bzw.* A*) zu paarweise verschiedenen Eigenwerten* $\lambda_1 \ldots, \lambda_r \in K$*, dann ist die Familie* $(x_1, \ldots, x_r)$ *linear unabhängig.*

Insbesondere gilt, ist $\dim_K(V) = n$ *und besitzt* f *(bzw.* A*)* n *paarweise verschiedene Eigenwerte, so ist* f *(bzw.* A*) diagonalisierbar.*

Beweis: Der Beweis sei dem Leser als Übungsaufgabe überlassen. □

Bemerkung 2.9
Kennt man einen Eigenwert $\lambda \in K$ von A, so kann man das lineare Gleichungssystem

$$(A - \lambda \mathbb{1}_n)x = 0$$

lösen und damit eine Basis des Eigenraumes $\mathrm{Eig}(A, \lambda) = \mathrm{Lös}(A - \lambda\mathbb{1}_n, 0) = \mathrm{Ker}(f_{A-\lambda\mathbb{1}_n}) = \mathrm{Ker}(f_A - \lambda \mathrm{id}_{K^n})$ bestimmen. D. h., bei Kenntnis des Eigenwertes λ lassen sich die Eigenvektoren von A zu λ durch Lösen eines linearen Gleichungssystems bestimmen.

Aber wie kommt man zu den Eigenwerten von A?

Diese Frage führt uns jetzt zum ersten Mal auf ein nicht-lineares Problem. Dazu betrachten wir die Matrix $A - \lambda\mathbb{1}_n$. Wir überlegen uns mit Hilfe von Korollar II.10.7:

$$\begin{aligned}
&\lambda \in K \text{ ist Eigenwert von } A \in \mathrm{Mat}(n, K)\\
\Leftrightarrow\ & f_{A-\lambda\mathbb{1}_n} : K^n \to K^n \text{ ist nicht injektiv}\\
\Leftrightarrow\ & A - \lambda\mathbb{1}_n \text{ ist nicht invertierbar}\\
\Leftrightarrow\ & \det(A - \lambda\mathbb{1}_n) = 0.
\end{aligned}$$

Da wir davon ausgehen, λ nicht zu kennen, sondern erst bestimmen zu wollen, machen wir einen *Ansatz*, d. h. wir ersetzen λ durch eine Unbestimmte t und betrachten die Matrix

$$A - t\mathbb{1}_n = \begin{pmatrix} a_{11} - t & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - t & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - t \end{pmatrix} \in \mathrm{Mat}\big(n, K[t]\big).$$

Hierbei sind die Koeffizienten a_{ij} in K, aber die $a_{ii} - t$ nicht mehr. Sie sind vielmehr Polynome in $K[t]$.

Da $K[t]$ ein kommutativer Ring mit Eins ist und da wir die Theorie der Determinanten für solche Ringe entwickelt haben, haben wir jetzt keine Probleme und können das Polynom $\det(A - t\mathbb{1}_n) \in \mathbb{K}[t]$ betrachten.

Definition 2.10
Für $A \in \mathrm{Mat}(n, K)$ heißt das Polynom $\chi_A := \det(A - t\mathbb{1}_n) \in K[t]$ das *charakteristische Polynom* von A.

Lemma 2.11
Konjugierte Matrizen in $\mathrm{Mat}(n, K)$ *haben das gleiche charakteristische Polynom.*

Beweis: Sind $A, B \in \mathrm{Mat}(n, K)$ zwei konjugierte Matrizen, so gibt es ein $T \in \mathrm{Gl}_n(K)$ mit $B = T \circ A \circ T^{-1}$. Damit gilt aber

$$B - t\mathbb{1}_n = T \circ A \circ T^{-1} - t\mathbb{1}_n = T \circ (A - t\mathbb{1}_n) \circ T^{-1}$$

und $T \in \mathrm{Gl}_n(K) \subset \mathrm{Gl}_n\big(K[t]\big)$, d. h. $A - t\mathbb{1}_n$ und $B - t\mathbb{1}_n$ sind konjugiert in $\mathrm{Mat}\big(n, K[t]\big)$. Aber konjugierte Matrizen haben nach Korollar II.9.15 die gleiche Determinante, und daraus folgt die Behauptung. □

Bemerkung 2.12
Ist $1 \le \dim_K(V) = n < \infty$ und sind B und D zwei Basen von V, dann sind $M_B^B(f)$ und $M_D^D(f)$ konjugiert in $\mathrm{Mat}(n, K)$. Also haben sie das gleiche charakteristische Polynom, und damit hängt die folgende Definition nicht von der gewählten Basis B ab.

Definition 2.13
Es sei $1 \le \dim_K(V) < \infty$ und B sei eine Basis von V. Dann heißt das Polynom $\chi_f := \chi_{M_B^B(f)} = \det\big(M_B^B(f) - t\mathbb{1}_n\big) \in K[t]$ *charakteristische Polynom* von f.

Die Überlegung, λ ist Eigenwert von A genau dann, wenn $\det(A - \lambda\mathbb{1}_n) = 0$, führt zu folgendem wichtigen Satz.

Satz 2.14 a. *Die Eigenwerte von* A *sind genau die Nullstellen von* χ_A *in* K.
b. *Ist* $1 \le \dim_K(V) < \infty$*, so sind die Eigenwerte von* f *genau die Nullstellen von* χ_f.

Beweis: In Definition 1.17 und Bemerkung 1.19 haben wir für $\lambda \in K$ den Einsetzhomomorphismus $\phi_\lambda : K[t] \to K : p \mapsto p(\lambda)$ kennengelernt.

Da ϕ_λ ein K-Algebrenhomomorphismus ist, folgt aus der Definition der Determinante mittels der Leibnizschen Formel

$$\det(A - t\mathbb{1}_n) = \sum_{\sigma\in\mathbb{S}_n} \operatorname{sgn}(\sigma) a'_{1\sigma(1)} \cdots a'_{n\sigma(n)}$$

mit $A = (a_{ij})$, $a'_{ij} = a_{ij}$ für $i \neq j$ und $a'_{ii} = a_{ii} - t$, daß

$$\det(A - \lambda\mathbb{1}_n) = \det\left(\left(\phi_\lambda\left(a'_{ij}\right)\right)_{i,j}\right) = \phi_\lambda\big(\det(A - t\mathbb{1}_n)\big) = \chi_A(\lambda)$$

für $\lambda \in K$.[3]

Damit gilt also, $\lambda \in K$ ist Eigenwert von A genau dann, wenn $\det(A-\lambda\mathbb{1}_n) = 0$, was wiederum genau dann der Fall ist, wenn λ eine Nullstelle von χ_A ist.

Damit ist a. gezeigt. b. folgt mit Bemerkung 2.7. □

Bevor wir das charakteristische Polynom weiter untersuchen, zunächst einige Beispiele.

Beispiel 2.15 a. Betrachten wir zunächst die folgende Matrix:

$$A = \begin{pmatrix} 0 & 1 & 1 \\ -1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix} \in \operatorname{Mat}(3, K).$$

Mit Hilfe der Regel von Sarrus oder durch den Laplaceschen Entwicklungssatz bestimmen wir das charakteristische Polynom von A als

$$\chi_A = \det\begin{pmatrix} -t & 1 & 1 \\ -1 & 2-t & 1 \\ -1 & 1 & 2-t \end{pmatrix} = -t^3 + 4t^2 - 5t + 2 = -(t-1)^2 \cdot (t-2).$$

Alternativ kann man allgemein die Matrix $A - t\mathbb{1}_n \in \operatorname{Mat}\big(n, K[t]\big) \subset \operatorname{Mat}\big(n, K(t)\big)$ auch als Matrix über dem Körper $K(t)$ auffassen. Die Definition der Determinante mittels der Leibnizschen Formel zeigt, daß die Determinante von A nicht davon abhängt, ob man A als Matrix über $K[t]$ oder über $K(t)$ auffaßt. Arbeiten wir jedoch über $K(t)$, so dürfen wir die Determinante mittels des Gaußschen Algorithmus II.9.18 bestimmen, da

[3] D. h. es kommt aufs Gleiche raus, ob ich in $A - t\mathbb{1}_n$ zunächst λ einsetze und dann die Determinante bestimme, oder ob ich zuerst die Determinante von $A-t\mathbb{1}_n$ berechne und dann λ einsetze.

$K(t)$ ein Körper ist. Insbesondere dürfen wir dabei durch Polynome (ungleich Null) dividieren!

$$\begin{pmatrix} -t & 1 & 1 \\ -1 & 2-t & 1 \\ -1 & 1 & 2-t \end{pmatrix} \xrightarrow[\text{III}\mapsto\text{III}-\frac{1}{t}\text{I}]{\text{II}\mapsto\text{II}-\frac{1}{t}\text{I}} \begin{pmatrix} -t & 1 & 1 \\ 0 & 2-t-\frac{1}{t} & 1-\frac{1}{t} \\ 0 & 1-\frac{1}{t} & 2-t-\frac{1}{t} \end{pmatrix} =$$

$$\begin{pmatrix} -t & 1 & 1 \\ 0 & -\frac{(t-1)^2}{t} & \frac{t-1}{t} \\ 0 & \frac{t-1}{t} & -\frac{(t-1)}{t} \end{pmatrix} \xrightarrow{\text{III}\mapsto\text{III}+\frac{1}{t-1}\text{II}} \begin{pmatrix} -t & 1 & 1 \\ 0 & -\frac{(t-1)^2}{t} & \frac{t-1}{t} \\ 0 & 0 & -t+2 \end{pmatrix}. \tag{42}$$

Entsprechend erhalten wir für das charakteristische Polynom

$$\chi_A = (-t)\cdot\left(-\tfrac{(t-1)^2}{t}\right)\cdot(-t+2) = -(t-1)^2\cdot(t-2).$$

Das charakteristische Polynom hat also die Nullstellen 1 und $2 = 2\cdot 1_K$, wobei 1 eine zweifache Nullstelle ist. Insbesondere ist also $\sigma(A) = \{1,2\}$. Wir können jetzt für $\lambda = 1$ und für $\lambda = 2$ jeweils den Eigenraum $\mathrm{Eig}(A,\lambda)$ bestimmen.

Hierzu bestimmen wir eine ZSF von $A - \lambda\mathbb{1}_3$.

Man beachte, daß es dabei nicht erlaubt ist, in (42) in der letzten Matrix t etwa durch $\lambda = 1$ zu ersetzen, um die ZSF zu erhalten, da wir bei den vorgenommenen Umformungen zur Ermittelung obiger Matrix durch das Polynom $t-1$ dividiert haben. Dies ist über $K(t)$ eine erlaubte Operation gewesen. Ersetzen wir jedoch t durch 1, so ist die Operation nicht mehr erlaubt! Man kann dieses Phänomen in geeigneter Weise so interpretieren, daß die Operation, die eine Matrix in ihre (reduzierte) Zeilen-Stufen-Form überführt eine *unstetige* Operation ist. (Vgl. Aufgabe 2.30.)

Als reduzierte ZSF von $A - \mathbb{1}_3$ erhalten wir:

$$\begin{pmatrix} -1 & 1 & 1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix} \mapsto \begin{pmatrix} 1 & -1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Daraus ergibt sich

$$\mathrm{Eig}(A,1) = \left\{(r+s,r,s)^t \;\middle|\; r,s\in K\right\} = \left\langle (1,1,0)^t, (1,0,1)^t\right\rangle.$$

$\mathrm{Eig}(A,1)$ ist also zweidimensional.

Analog ergibt sich $\mathrm{Eig}(A,2)$ aus

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

und damit gilt $\mathrm{Eig}(A,2) = \left\langle (1,1,1)^t\right\rangle$.

Ist $K = \mathbb{Q}$, so lassen sich diese Rechnungen mit Singular wie folgt ausführen:

```
ring r=0,t,dp;
matrix A[3][3]=0,1,1,-1,2,1,-1,1,2;
LIB "matrix.lib";
matrix B[3][3]=A-t*unitmat(3);
poly p=det(B); // Das charakteristische Polynom von A.
↦   -t3+4t2-5t+2
factorize(p); // Dies faktorisiert p über ℚ.
↦      [1]:  // Faktoren von p
             [1]=-1
             [2]=t-2
             [3]=t-1
       [2]:  // Vielfachheiten
             1,1,2
```

Dies bedeutet, daß der Leitkoeffizient von p gerade -1 ist (mit Vielfachheit eins), $t-2$ ist ein Faktor von p mit Vielfachheit eins und $t-1$ ist ein Faktor von p mit Vielfachheit zwei.

Im Allgemeinen wird ein Polynom über $\mathbb{Q}$ nicht in Linearfaktoren zerfallen. Dann kann man zu $K = \mathbb{C}$ übergehen, aber man muß die Nullstellen dann numerisch approximieren. Auch das kann Singular.

```
LIB "solve.lib";
laguerre_solve(p,5);  // Berechnet die Nullstellen auf fünf
                      // Nachkommastellen genau.
↦      [1]:
             0.99980
       [2]:
             1.002
       [3]
             2
```

Man beachte, daß bei numerischer Rechnung Rundungsfehler auftreten. Dabei wird die doppelte Nullstelle Eins als zwei dicht beieinander liegende Nullstellen interpretiert. Die Angabe der Genauigkeit (hier fünf Stellen) bedeutet, daß intern mit (etwas mehr als) fünf Stellen Genauigkeit gerechnet wird - das Ergebnis kann weniger genau sein, wie in unserem Beispiel.

Wir wollen nun noch zeigen, wie man die Gauß-Reduktion zur Bestimmung des charakteristischen Polynoms über $\mathbb{Q}(t)$ in Singular durchführen kann.

```
ring R=(0,t),x,dp; // Definiert den Polynomring ℚ(t)[x].
matrix B=map(r,B); // Bildet die Matrix B vom Ring r nach R ab.
LIB "linalg.lib";
matrix G=gauss_nf(B); // Bestimmt eine ZSF von B.
print(G);
```

```
↦    (-t),  1,                1,
     0,     (-t2+2t-1)/(t),  (t-1)/(t),
     0,     0,               (-t+2)
```

Um die Eigenräume zu bestimmen, gehen wir wieder nach `r` zurück.

```
setring r;
matrix C=subst(B,t,1); // Substituiert in B  t durch 1.
print(syz(C)); // syz(C) liefert eine Basis des Kerns von C.
↦    1,  0,
     1,  -1,
     0,  1
print(syz(subst(B,t,2)));
↦    1,
     1,
     1
```

b. Wir hatten schon durch eine geometrische Argumentation gesehen, daß die Drehung φ_α um einen Winkel α im allgemeinen keinen reellen Eigenwert besitzt. Den gleichen Sachverhalt prüfen wir nun noch einmal mit algebraischen Methoden. Die Matrixdarstellung von φ_α bezüglich der kanonischen Basis von $\mathbb{R}^2$ ist

$$A = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}.$$

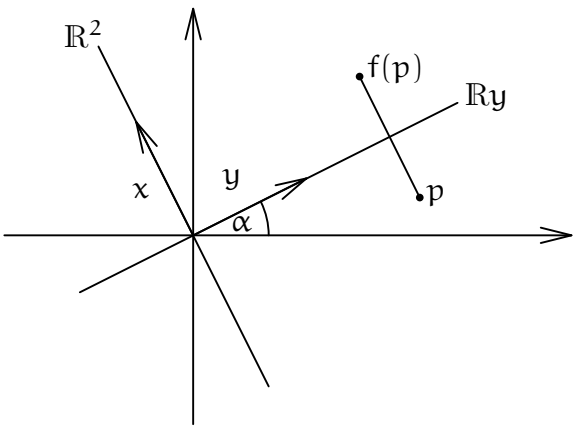
Damit gilt aber

$$\chi_{\varphi_\alpha} = \chi_A = \big(\cos(\alpha) - t\big)^2 + \sin^2(\alpha) = t^2 - 2\cos(\alpha)t + 1.$$

Die Nullstellen von χ_A sind $\cos(\alpha) + \sqrt{\cos^2(\alpha) - 1}$ und $\cos(\alpha) - \sqrt{\cos^2(\alpha) - 1}$. Für beide Terme gilt, sie sind genau dann reell, wenn α ein ganzzahliges Vielfaches von π ist.

Insbesondere hat φ_α also nur dann reelle Eigenwerte, wenn α ein ganzzahliges Vielfaches von π ist, d. h. $A = \mathbb{1}_2$ oder $A = -\mathbb{1}_2$.

c. Es sei $f \in \mathrm{End}_{\mathbb{R}}\big(\mathbb{R}^2\big)$ die *Spiegelung* an einer Geraden $\langle y\rangle = \mathbb{R}\cdot y \subset \mathbb{R}^2$ mit $0 \neq y = (y_1, y_2) \in \mathbb{R}^2$. (Vgl. Aufgabe IV.2.42.)



Wir setzen $x = (y_2, -y_1) \in \mathbb{R}^2$. Dann gilt für das Skalarprodukt von x mit y

$$\langle x, y\rangle = y_2y_1 - y_1y_2 = 0$$

und $B = (y, x)$ ist eine Basis von $\mathbb{R}^2$. Die Spiegelung f bildet mithin y auf sich selbst und x auf $-x$ ab, da x senkrecht auf $\langle y\rangle$ steht. Damit hat f die folgende Matrixdarstellung bezüglich B

$$M_B^B(f) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

und das charakteristische Polynom von f ist gerade

$$\chi_f = (t-1)\cdot(t+1).$$

Die Spiegelung von f ist also diagonalisierbar und hat das Spektrum $\sigma(f) = \{-1, 1\}$.

Beschreiben wir f in den Standardkoordinaten $E = (e_1, e_2)$ von $\mathbb{R}^2$, so ist f die Spiegelung an $\langle e_1\rangle = \mathbb{R}\cdot e_1$ gefolgt von der Drehung um den Winkel α, wenn $\frac{\alpha}{2}$ der Winkel ist, den $\langle y\rangle$ mit $\langle e_1\rangle$ einschließt. Wir erhalten also

$$M_E^E(f) = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{pmatrix} = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \circ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Das charakteristische Polynom errechnet sich aus dieser Matrixdarstellung als

$$\big(\cos(\alpha) - t\big)\cdot\big(-\cos(\alpha) - t\big) - \sin^2(\alpha) = t^2 - 1 = (t-1)\cdot(t+1).$$

Bemerkung 2.16 a. Wir haben bereits gesehen, daß elementare Zeilen-/Spaltenoperationen der Matrix $A - t\mathbb{1}_n$, die wir über $K(t)$ vornehmen nur dann mit dem Einsetzen von λ verträglich sind, wenn wir zwischendurch kein Element der Matrix durch ein Polynom $g \in K[t]$ dividiert haben, das λ als Nullstelle besitzt. (Vgl. Aufgabe 2.30.)

b. Ferner gilt festzuhalten, daß sich die Eigenräume nur dann auf die angegebene Weise bestimmen lassen, wenn die Eigenwerte exakt bekannt sind. Eine numerische Approximation der Eigenwerte ist hierfür nicht hinreichend, da die Matrix $A - (\lambda + \epsilon)\mathbb{1}_n$ schon für kleine ϵ in aller Regel vollen Rang hat, wenn $\lambda \in \sigma(A)$. In diesem Fall liefert die Numerik andere, iterative Verfahren zur gleichzeitigen Approximation von Eigenwerten und Eigenvektoren.

Algorithmus 2.17 (Charakteristisches Polynom)

INPUT: $A \in \mathrm{Mat}(n, K)$.

OUTPUT: χ_A, das charakteristische Polynom von A.

1. **Schritt:** Bilde die Matrix $B = A - t\cdot\mathbb{1}_n$.
2. **Schritt:** Bestimme die Determinante von B mittels des Algorithmus II.9.18.
3. **Schritt:** Gib die Determinante zurück.

Vgl. auch Aufgabe 2.30.

Abschließend wollen wir einen näheren Blick auf die Struktur des charakteristischen Polynoms werfen.

Lemma 2.18 a. *Es gilt:*

$$\chi_A = (-1)^n t^n + \alpha_{n-1} t^{n-1} + \ldots + \alpha_0,$$

wobei $\alpha_{n-1} = (-1)^{n-1} \cdot \mathrm{Spur}(A)$ *und* $\alpha_0 = \det(A)$.

b. *Ist* $1 \leq \dim_K(V) = n < \infty$*, so gilt:*

$$\chi_f = (-1)^n t^n + \alpha_{n-1} t^{n-1} + \ldots + \alpha_0,$$

wobei $\alpha_{n-1} = (-1)^{n-1} \cdot \mathrm{Spur}(f)$ *und* $\alpha_0 = \det(f)$.

Beweis: Ist $A = (a_{ij})$ und $A - t\mathbb{1}_n = \left(a'_{ij}\right)$, dann folgt aus der Leibnizschen Formel für die Determinante

$$\chi_A = (a_{11} - t) \cdots (a_{nn} - t) + \sum_{\mathrm{id} \neq \sigma \in \mathbb{S}_n} \mathrm{sgn}(\sigma) a'_{1\sigma(1)} \cdots a'_{n\sigma(n)}.$$

Da für $\sigma \neq \mathrm{id}$ mindestens zwei Faktoren in $a'_{1\sigma(1)} \cdots a'_{n\sigma(n)}$ konstante Polynome sind, ergibt $\sum_{\mathrm{id} \neq \sigma \in \mathbb{S}_n} \mathrm{sgn}(\sigma) a'_{1\sigma(1)} \cdots a'_{n\sigma(n)}$ ein Polynom vom Grad kleiner gleich $n-2$. Damit lassen sich die Koeffizienten von t^n und t^{n-1} in χ_A aus $(a_{11} - t) \cdots (a_{nn} - t)$ herleiten und sind wie behauptet. Ferner gilt $\alpha_0 = \chi_A(0) = \det(A)$ ist das konstante Glied im charakteristischen Polynom.

Die Aussage für Endomorphismen f ergibt sich aus der für Matrizen unter Berücksichtigung der Tatsache, daß sowohl die Determinante als auch die Spur und das charakteristische Polynom eines Endomorphismus mittels einer Matrixdarstellung zu beliebiger Basis definiert sind. □

Bemerkung 2.19

Aus dem vorigen Lemma folgt insbesondere, daß die Spur einer Matrix invariant unter Konjugation ist. Vgl. hierzu Aufgabe II.5.43.

AUFGABEN

Aufgabe 2.20

Bestimme alle Eigenwerte mit den zugehörigen Eigenräumen der beiden K-linearen Abbildungen S und T aus Aufgaben II.3.18.

Aufgabe 2.21

Man betrachte den Unterraum $P_n := \langle t^i \mid i = 0, \ldots, n \rangle$ der Polynome vom Grad kleiner gleich n von $K[t]$, und definiere zwei Abbildungen $D, T : P_n \to P_n$ durch

$$D(f) = \sum_{i=0}^{n-1} (i+1) a_{i+1} t^i \quad \text{bzw.} \quad T(f) = \sum_{i=0}^{n} a_i (t+1)^i$$

für $f = \sum_{i=0}^{n} a_i t^i$.

a. Zeige, daß D und T K-linear sind.
b. Bestimme die charakteristischen Polynome von D und T.

Aufgabe 2.22
Es seien $f, g \in \mathrm{End}_K(V)$. Zeige:

a. Ist $\dim_K(V) < \infty$, so haben $f \circ g$ und $g \circ f$ die gleichen Eigenwerte.
b. Ist $\dim_K(V) = \infty$, so gilt dies i. a. nicht mehr.

Aufgabe 2.23
Es sei V ein K-Vektorraum und $f, g \in \mathrm{End}_K(V)$ so, daß g invertierbar ist und $f \circ g - g \circ f = \lambda \mathrm{id}_V$ für ein ein $\lambda \in K$. Zeige, ist $\mu \in K$ ein Eigenwert von $g \circ f$, so ist $\mu + n \cdot \lambda \in K$ Eigenwert von $g \circ f$ für alle $n \in \mathbb{Z}$.

Aufgabe 2.24
Es sei $1 \leq \dim_K(V) < \infty$, $f \in \mathrm{End}_K(V)$.

a. Ist $U \subseteq V$ ein f-invarianter Unterraum, dann gilt:

$$\chi_f = \chi_{f_U} \cdot \chi_{f_{V/U}}.$$

b. Ist $V = \bigoplus_{i=1}^{r} U_i$, wobei die U_i f-invariant seien, dann gilt:

$$\chi_f = \prod_{i=1}^{r} \chi_{f_{U_i}}.$$

Aufgabe 2.25
Es seien $A = (a_{ij}), B = (b_{ij}) \in \mathrm{Mat}(n, K)$ zwei obere Dreiechsmatrizen. Zeige, falls A und B konjugiert sind, dann sind die beiden n-Tupel $(a_{11}, \ldots, a_{nn})$ und $(b_{11}, \ldots, b_{nn})$ bis auf die Reihenfolge gleich.

Aufgabe 2.26
Zeige, $A \in \mathrm{Mat}(n, K)$ ist genau dann diagonalisierbar, wenn A^t diagonalisierbar ist.

Aufgabe 2.27
Es sei $1 \leq \dim_K(V) < \infty$ und $f \in \mathrm{End}_K(V)$. f heißt *nilpotent*, falls es ein $r \in \mathbb{N}$ gibt mit $f^r = 0$. Zeige:

a. Ist f nilpotent, so ist 0 der einzige Eigenwert von f.
b. Ist f nilpotent, so besitzt der Eigenraum von 0 kein f-invariantes Komplement.
c. Ist $K = \mathbb{C}$ und 0 der einzige Eigenwert von f, dann ist f nilpotent und es gilt $f^{\dim_{\mathbb{C}}(V)} = 0$.

Aufgabe 2.28
Es sei V ein $\mathbb{R}$-Vektorraum und $f \in \mathrm{End}_{\mathbb{R}}(V)$.

a. Hat f^2 einen nicht-negativen Eigenwert, so hat auch f einen Eigenwert.
b. Ist $\dim_{\mathbb{R}}(V) = 2$ und hat f^2 einen negativen Eigenwert λ, so hat f keinen Eigenwert und $f^2 = \lambda \cdot \mathrm{id}_V$.
c. Finde ein Beispiel mit $\dim_{\mathbb{R}}(V) > 2$, so daß f^2 einen negativen Eigenwert besitzt und f ebenfalls einen Eigenwert hat.

Aufgabe 2.29
Es sei $A \in \mathrm{Gl}_n(K)$. Zeige:

a. Genau dann gilt $x \in \mathrm{Eig}(A,\lambda)$, wenn $x \in \mathrm{Eig}\left(A^{-1},\lambda^{-1}\right)$.
Insbesondere gilt also $\sigma(A) = \{\lambda_1,\ldots,\lambda_r\}$ impliziert $\sigma\left(A^{-1}\right) = \left\{\lambda_1^{-1},\ldots,\lambda_r^{-1}\right\}$.

b. Ist $\chi_A = (-1)^n t^n + (-1)^{n-1}\alpha_{n-1}t^{n-1} + \ldots + (-1)\alpha_1 t + \alpha_0$, dann gilt

$$\begin{aligned}\chi_{A^{-1}} &= (-1)^n t^n + (-1)^{n-1}\tfrac{\alpha_1}{\alpha_0}t^{n-1} + \ldots + (-1)\tfrac{\alpha_{n-1}}{\alpha_0}t + \tfrac{1}{\alpha_0}\\ &= \det\left(A^{-1}\right)(-t)^n\chi_A\left(\tfrac{1}{t}\right).\end{aligned}$$

Folgere, daß für $n = 2$ gilt

$$\mathrm{Spur}\left(A^{-1}\right) = \frac{\mathrm{Spur}(A)}{\det(A)}.$$

c. Falls $A^2 = A$, dann gilt $r := \mathrm{rang}(A) = \dim_K\left(\mathrm{Eig}(A,1)\right)$ und es gibt ein $T \in \mathrm{Gl}_n(K)$ mit

$$T \circ A \circ T^{-1} = \left(\begin{array}{c|c} \mathbb{1}_r & 0 \\ \hline 0 & 0 \end{array}\right).$$

Aufgabe 2.30 a. Es seien $R \subseteq S$ kommutative Ringe mit Eins und es sei $B \in \mathrm{Mat}(n,R) \subseteq \mathrm{Mat}(n,S)$. Man überzeuge sich anhand der Definition der Determinante mittels der Leibniz-Formel, daß die Determinante von B unabhängig davon ist, ob man A als Matrix über R oder als Matrix über S auffaßt.

b. Aus a. (mit $R = K[t]$ und $S = K(t)$) leite man ab, daß sich das charakteristische Polynom einer Matrix $A \in \mathrm{Mat}(n,K)$ mittels des Gauß-Algorithmus für Determinanten (vgl. Algorithmus II.9.18) berechnen läßt.

c. Schreibe eine Singular-Prozedur `char_poly`, die eine quadratische Matrix $A \in \mathrm{Mat}(n,\mathbb{Q})$ einliest und das charakteristische Polynom $\det(A - t\mathbb{1}_n)$ von A ausgibt. Dabei soll die Determinante mit Hilfe des modifizierten rekursiven Gauß-Algorithmus zur Determinantenberechnung bestimmt werden:
INPUT: $A \in \mathrm{Mat}(n,\mathbb{Q})$
OUTPUT: χ_A
1. Schritt: Falls A nicht quadratisch ist, gib 0 zurück.
2. Schritt: Beim *ersten Aufruf* von `char_poly` ersetze A durch $A - t\mathbb{1}_n$.
3. Schritt: Ist $n = 1$, gib $A[1,1]$ zurück.
4. Schritt: Für $i = 2,\ldots,n$ reduziere die i-te Zeile mit der ersten.
5. Schritt: Streiche in A die erste Zeile und die erste Spalte, um eine Matrix B zu erhalten.
6. Schritt: Rufe `char_poly` mit B auf (*stelle sicher, daß dies nicht mehr als erster Aufruf gewertet wird!*), multipliziere das Ergebnis mit $A[1,1]$ und gib es zurück.

d. Berechne das charakteristische Polynom von $A + \mathbb{1}_{10}$ mit

```
matrix A[10][10]=ideal(1..100);
```

Hinweis: Man muß über dem Ring `ring r=(0,t),x,lp;` arbeiten, was mathematisch bedeutet, daß wir unsere Matrixeinträge als Elemente des Funktionenkörpers $\mathbb{Q}(t)$ auffassen, so daß wir auch durch Polynome (ungleich Null) dividieren dürfen. – Als Prozedurkopf verwende man `proc char_poly (matrix A,list #)`. Dies erlaubt es, die Prozedur wahlweise mit einem oder mehreren Parametern aufzurufen. Zudem kann mittels `size(#);` überprüft werden, mit wieviel Parametern die Prozedur aufgerufen wurde. Damit kann man steuern, ob die Prozedur zum ersten Mal (`size(#)==0`) oder zum wiederholten Male aufgerufen wird, indem man beim internen Aufruf von `char_poly` schlicht die Prozedur mit einem zweiten Parameter aufruft. – Mit der Singular-Prozedur `charpoly` der Bibliothek `linalg.lib` kann die eigene Prozedur auf Richtigkeit getestet werden.

Aufgabe 2.31

Es sei V ein endlich-dimensionaler K-Vektorraum und $f \in \mathrm{End}_K(V)$.

a. Zeige, ist f diagonalisierbar, so besitzt jeder f-invariante Unterraum ein f-invariantes Komplement.
b. Gib ein möglichst einfaches Beispiel für einen nicht-diagonalisierbaren Endomorphismus g und einen g-invarianten Unterraum, der kein g-invariantes Komplement besitzt.

3. Der Satz von Cayley-Hamilton

Generalvoraussetzung: Es sei K ein Körper, V ein K-Vektorraum mit $1 \leq \dim_K(V) = n < \infty$ und $f \in \mathrm{End}_K(V)$. A bezeichne stets eine Matrix in $\mathrm{Mat}(n, K)$.

Wir nähern uns in diesem Abschnitt unserem Ziel, eine Matrix A durch Konjugation in möglichst einfache Form zu überführen, ohne es aber schon ganz zu erreichen. Wir stellen uns die Frage, wann man A in eine Diagonalmatrix oder wenigstens in eine obere Dreiecksmatrix überführen kann, d. h. , wann A diagonalisierbar oder wenigstens trigonalisierbar ist.

Dazu beweisen wir zunächst den folgenden Satz.

Satz 3.1

A (bzw. f) ist genau dann trigonalisierbar, wenn das charakteristische Polynom von A (bzw. von f) über K in Linearfaktoren zerfällt.

Beweis: Ist f trigonalisierbar, so gibt es eine Basis B mit

$$M_B^B(f) = \begin{pmatrix} \lambda_1 & * & \dots & \dots & * \\ 0 & \lambda_2 & * & \dots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda_{n-1} & * \\ 0 & \dots & \dots & 0 & \lambda_n \end{pmatrix}.$$

Damit folgt das charakteristische Polynom

$$\chi_f = (\lambda_1 - t) \cdots (\lambda_n - t)$$

von f zerfällt über K in Linearfaktoren.

Zerfalle nun umgekehrt das charakteristische Polynom von f in Linearfaktoren $\chi_f = (\lambda_1 - t)\cdots(\lambda_n - t)$. Wir beweisen mit Induktion über $n = \dim_K(V)$, daß dann f trigonalisierbar ist. Im Fall $n = 1$ ist f nach Beispiel 2.4 sogar diagonalisierbar. Sei also $n > 1$ und sei $0 \neq x_1 \in V$ ein Eigenvektor von f zum Eigenwert λ_1. Wir setzen $U := \langle x_1\rangle \subset V$. Wegen $f(x_1) = \lambda_1 x_1 \in U$ ist U ein f-invarianter Unterraum von V. Mithin folgt aus Aufgabe 2.24

$$\chi_{f_{V/U}} = (\lambda_2 - t)\cdots(\lambda_n - t),$$

d. h. das charakteristische Polynom von $f_{V/U}$ zerfällt über K in Linearfaktoren. Da $\dim_K(V/U) = n - 1 < n$, existiert per Induktion eine Basis $B'' = (x_2 + U, \ldots, x_n + U)$ von V/U, so daß $M_{B''}^{B''}(f_{V/U})$ eine obere Dreiecksmatrix ist. Dann ist aber $B = (x_1, \ldots, x_n)$ eine Basis von V und mit $B' = (x_1)$ gilt

$$M_B^B(f) = \left(\begin{array}{c|c} M_{B'}^{B'}(f_U) & * \\ \hline 0 & M_{B''}^{B''}(f_{V/U}) \end{array}\right) = \left(\begin{array}{c|c} \lambda_1 & * \\ \hline 0 & M_{B''}^{B''}(f_{V/U}) \end{array}\right).$$

Damit ist $M_B^B(f)$ eine obere Dreiecksmatrix und f ist trigonalisierbar.

Die Aussage für eine Matrix A erhalten wir aus der entsprechenden Aussage für f_A. □

Bemerkung 3.2
Ist K ein algebraisch abgeschlossener Körper, etwa $K = \mathbb{C}$, so sind somit jede Matrix A und jeder Endomorphismus f trigonalisierbar.

Eine vergleichbare Aussage für die Diagonalisierbarkeit gilt nicht.

Beispiel 3.3 a. Die Drehmatrix

$$A_\alpha = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

hat das charakteristische Polynom $\chi_{A_\alpha} = t^2 - 2\cos(\alpha)t + 1 = (t-\lambda)\cdot(t-\overline{\lambda})$ mit $\lambda = \cos(\alpha) + i\sin(\alpha) \in \mathbb{C}$, $\alpha \in \mathbb{R}$. Damit hat χ_{A_α} also keine reellen Nullstellen, wenn α kein ganzzahliges Vielfaches von π ist, und somit ist A_α über $\mathbb{R}$ nicht trigonalisierbar.
Hingegen zerfällt χ_{A_α} über $\mathbb{C}$ in Linearfaktoren, so daß A_α über $\mathbb{C}$ trigonalisierbar sein muß. In der Tat ist A_α sogar diagonalisierbar mit

$$A_\alpha \sim \begin{pmatrix} \lambda & 0 \\ 0 & \overline{\lambda} \end{pmatrix}.$$

Ist α kein ganzzahliges Vielfaches von π, so besitzt A_α zwei verschiedene Eigenwerte, so daß zugehörige Eigenvektoren nach Proposition 2.8 eine Basis von $\mathbb{C}^2$ bilden müssen, die nach Lemma 2.6 eine Transformation von A_α in obige Diagonalmatrix bewirkt. Ist α hingegen ein ganzzahliges Vielfaches von π, so ist $A_\alpha = \mathbb{1}_2$ oder $A_\alpha = -\mathbb{1}_2$ und hat bereits Diagonalgestalt.

b. Die Matrix

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in \mathrm{Mat}(2, \mathbb{C})$$

ist hingegen auch über $\mathbb{C}$ nicht diagonalisierbar. Denn, gäbe es eine Matrix $T \in \mathrm{Gl}_2(\mathbb{C})$ mit

$$T \circ A \circ T^{-1} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \in \mathrm{Mat}(2, \mathbb{C}),$$

dann wäre

$$\begin{pmatrix} \lambda_1^2 & 0 \\ 0 & \lambda_2^2 \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}^2 = T \circ A^2 \circ T^{-1} = 0,$$

also wären $\lambda_1 = \lambda_2 = 0$. Aber damit würde gelten:

$$0 = \mathrm{rang} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \mathrm{rang}\left(T \circ A \circ T^{-1}\right) = \mathrm{rang}(A) = 1,$$

da $T \in \mathrm{Gl}_2(\mathbb{C})$. Dies ist jedoch ein Widerspruch.

Wir wollen noch ein weiteres Kriterium für Trigonalisierbarkeit angeben.

Definition 3.4 a. Eine aufsteigende Kette von Unterräumen

$$\{0\} = V_0 \subsetneqq V_1 \subsetneqq \ldots \subsetneqq V_r = V$$

heißt eine *Fahne* von V.

b. Ist $r = \dim_K(V)$, so heißt die Fahne *vollständig*.

c. Sind die V_i, $i = 0, \ldots, r$, alle f-invariant, so heißt die Fahne f-*invariant*.

Satz 3.5

Genau dann ist f trigonalisierbar, wenn V eine vollständige f-invariante Fahne besitzt.

Beweis: Ist $B = (x_1, \ldots, x_n)$ eine Basis von V, bezüglich derer $M_B^B(f)$ eine obere Dreiecksmatrix ist, d. h. $f(x_i) \in \langle x_1, \ldots, x_i \rangle_K =: V_i$ für alle $i = 1, \ldots, n$. Dann ist

$$\{0\} =: V_0 \subsetneqq V_1 \subsetneqq \ldots \subsetneqq V_n = V$$

eine vollständige f-invariante Fahne.

Ist umgekehrt eine vollständige f-invariante Fahne

$$\{0\} =: V_0 \subsetneqq V_1 \subsetneqq \ldots \subsetneqq V_n = V$$

gegeben, dann starte man mit einer Basis von V_1, ergänze sie zu einer Basis von V_2 und diese zu einer Basis von V_3 und bilde auf diese Weise rekursiv eine Basis $B = (x_1, \ldots, x_n)$ von V, so daß $V_i = \langle x_1, \ldots, x_i \rangle_K$. Da V_i, $i = 1, \ldots, n$, f-invariant ist, folgt $f(x_i) \in \langle x_1, \ldots, x_i \rangle_K$ und mithin ist $M_B^B(f)$ eine obere Dreiecksmatrix. □

Definition 3.6
Es sei $\lambda \in K$. Wir nennen $\mathrm{mult}\left(\chi_A, \lambda\right)$ bzw. $\mathrm{mult}\left(\chi_f, \lambda\right)$ die *algebraische Vielfachheit* von λ als Eigenwert von A bzw. f, und $\dim_K\left(\mathrm{Eig}(A, \lambda)\right)$ bzw. $\dim_K\left(\mathrm{Eig}(f, \lambda)\right)$ die *geometrische Vielfachheit* von λ als Eigenwert von A bzw. f.

Bemerkung 3.7
Da konjugierte Matrizen das gleiche charakteristische Polynom haben, ist klar, daß für konjugierte Matrizen die algebraischen Vielfachheiten der Eigenwerte übereinstimmen.

In der Tat haben konjugierte Matrizen auch die gleichen geometrischen Vielfachheiten zu gleichen Eigenwerten.

Seien dazu $A, B \in \mathrm{Mat}(n, K)$ und $T \in \mathrm{Gl}_n(K)$ mit $B = T \circ A \circ T^{-1}$. Dann gilt:

$$\begin{aligned}
\dim_K\left(\mathrm{Eig}(B, \lambda)\right) &= \dim_K\left(\mathrm{Lös}(B - \lambda \mathbb{1}_n, 0)\right) \\
&= n - \mathrm{rang}(B - \lambda \mathbb{1}_n) \\
&= n - \mathrm{rang}\left(T \circ (A - \lambda \mathbb{1}_n) \circ T^{-1}\right) \\
&= n - \mathrm{rang}(A - \lambda \mathbb{1}_n) \\
&= \dim_K\left(\mathrm{Lös}(A - \lambda \mathbb{1}_n, 0)\right) \\
&= \dim_K\left(\mathrm{Eig}(A, \lambda)\right).
\end{aligned}$$

Man beachte, daß zwar die Dimensionen der Eigenräume von A und B zu λ übereinstimmen, die Eigenräume selbst aber im Allgemeinen nicht! Vielmehr gilt:

$$\mathrm{Eig}(B, \lambda) = \{Ty \mid y \in \mathrm{Eig}(A, \lambda)\},$$

denn

$$\begin{aligned}
x \in \mathrm{Eig}(B, \lambda) &\Leftrightarrow \left(T \circ A \circ T^{-1}\right)x = Bx = \lambda x \\
&\Leftrightarrow A\left(T^{-1}x\right) = \lambda\left(T^{-1}x\right) \\
&\Leftrightarrow T^{-1}x \in \mathrm{Eig}(A, \lambda) \\
&\Leftrightarrow \exists\, y \in \mathrm{Eig}(A, \lambda) \,:\, x = Ty.
\end{aligned}$$

Definition 3.8
Wir werden im Folgenden sehr häufig mit Blockmatrizen der folgenden Form arbeiten:

$$A = \left(\begin{array}{c|ccc} A_1 & 0 & \cdots & 0 \\ \hline 0 & \boxed{A_2} & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & \boxed{A_r} \end{array}\right) \in \mathrm{Mat}(n, K),$$

wobei $A_i \in \mathrm{Mat}(n_i, K)$, $i = 1, \ldots, r$ mit $n = n_1 + \ldots + n_r$. Es empfiehlt sich deshalb, eine Kurzschreibweise für solche *Diagonalblockmatrizen* einzuführen. Wir schreiben kurz:

$$A = A_1 \oplus \ldots \oplus A_r = \bigoplus_{i=1}^{r} A_i.$$

Bemerkung 3.9
Man beachte, daß es bei der obigen Schreibweise für Diagonalblockmatrizen auf die Reihenfolge der Summation ankommt, daß aber Matrizen, die durch Änderung der Summationsreihenfolge entstehen, zueinander konjugiert sind!

Mit Hilfe dieser Notation gilt beispielsweise, daß eine Matrix A genau dann diagonalisierbar ist, wenn es Körperelemente $\lambda_1, \ldots, \lambda_r \in K$ und positive natürliche Zahlen $n_1, \ldots, n_r \in \mathbb{N}$ gibt sowie eine invertierbare Matrix $T \in Gl_n(K)$ mit

$$T \circ A \circ T^{-1} = \bigoplus_{i=1}^{r} \lambda_i \mathbb{1}_{n_i}.$$

Satz 3.10 (Diagonalisierbarkeit)
Die folgenden Aussagen sind äquivalent:

a. A *(bzw.* f*) ist diagonalisierbar.*
b. *Das charakteristische Polynom von* A *(bzw.* f*) zerfällt über* K *in Linearfaktoren und für jeden Eigenwert* λ *stimmen algebraische und geometrische Vielfachheit überein.*
c. *Sind* $\lambda_1, \ldots, \lambda_r$ *die paarweise verschiedenen Eigenwerte von* A *(bzw.* f*), dann gilt*

$$K^n = \bigoplus_{i=1}^{r} \mathrm{Eig}(A, \lambda_i) \quad \left(\textit{bzw.} \quad V = \bigoplus_{i=1}^{r} \mathrm{Eig}(f, \lambda_i) \right).$$

Bevor wir nun den Satz beweisen, zeigen wir, daß die algebraische Vielfachheit die geometrische Vielfachheit stets nach oben beschränkt.

Lemma 3.11
Für $\lambda \in K$ *gilt stets:*

$$\mathrm{mult}\left(\chi_A, \lambda\right) \geq \dim_K\left(\mathrm{Eig}(A, \lambda)\right) \quad \textit{und} \quad \mathrm{mult}\left(\chi_f, \lambda\right) \geq \dim_K\left(\mathrm{Eig}(f, \lambda)\right).$$

Beweis: Wähle eine Basis $B' = (x_1, \ldots, x_s)$ von $\mathrm{Eig}(f, \lambda)$ und ergänze sie durch $(x_{s+1}, \ldots, x_n)$ zu einer Basis B von V. Dann gilt mit $B'' = \big(x_{s+1} + \mathrm{Eig}(f, \lambda), \ldots, x_n + \mathrm{Eig}(f, \lambda)\big)$

$$M_B^B(f) = \begin{pmatrix} M_{B'}^{B'}\big(f_{\mathrm{Eig}(f,\lambda)}\big) & * \\ 0 & M_{B''}^{B''}\big(f_{V/\mathrm{Eig}(f,\lambda)}\big) \end{pmatrix} = \begin{pmatrix} \lambda \mathbb{1}_s & * \\ 0 & M_{B''}^{B''}\big(f_{V/\mathrm{Eig}(f,\lambda)}\big) \end{pmatrix}.$$

Es folgt $\chi_f = (\lambda - t)^s \cdot \chi_{f_{V/\mathrm{Eig}(f,\lambda)}}$ und damit $\mathrm{mult}\left(\chi_f, \lambda\right) \geq s = \dim_K\left(\mathrm{Eig}(f, \lambda)\right)$.
Die Aussage für eine Matrix A folgt aus der entsprechenden Aussage für f_A. □

Beweis von Satz 3.10: Es reicht jeweils, die Aussagen für eine Matrix A oder für einen Endomorphismus f zu zeigen.

a. ⇒ b.: Ist A diagonalisierbar, dann ist A konjugiert zu einer Matrix der Form $C = \bigoplus_{i=1}^{r} \lambda_i \mathbb{1}_{n_i}$ mit $\lambda_i \neq \lambda_j$ für $i \neq j$. Damit gilt aber:

$$\chi_A = \chi_C = (\lambda_1 - t)^{n_1} \cdots (\lambda_r - t)^{n_r}$$

und $\mathrm{Eig}(C,\lambda_i) = \langle e_{n_1+\ldots+n_{i-1}+1},\ldots,e_{n_1+\ldots+n_i}\rangle$. Insbesondere ist

$$\dim_K\big(\mathrm{Eig}(A,\lambda_i)\big) = \dim_K\big(\mathrm{Eig}(C,\lambda_i)\big) = n_i = \mathrm{mult}\left(\chi_A,\lambda_i\right)$$

und χ_A zerfällt über K.

b. $\Rightarrow$ c.: Es gelte $\chi_A = (\lambda_1 - t)^{n_1}\cdots(\lambda_r - t)^{n_r}$, mit paarweise verschiedenen λ_i, und $n_i = \dim_K\big(\mathrm{Eig}(A,\lambda_i)\big)$. Ferner sei $B_i = (x_{n_1+\ldots+n_{i-1}+1},\ldots,x_{n_1+\ldots+n_i})$ eine Basis von $\mathrm{Eig}(A,\lambda_i)$. Da $\dim_K\left(K^n\right) = n = n_1 + \ldots + n_r$, reicht es zu zeigen, daß $B = (x_1,\ldots,x_n)$ linear unabhängig ist. Seien dazu $\mu_1,\ldots,\mu_n \in K$ mit $0 = \sum_{i=1}^{n}\mu_i x_i$. Wir setzen $y_j = \sum_{i=n_1+\ldots+n_{j-1}+1}^{n_1+\ldots+n_j}\mu_i x_i \in \mathrm{Eig}(A,\lambda_j)$ und $\mathcal{M} = \{j \mid y_j \neq 0\}$.

Angenommen, $|\mathcal{M}| \geq 1$. Da die y_j Eigenvektoren zu λ_j sind, folgt mit Proposition 2.8 $(y_j \mid j \in \mathcal{M})$ ist linear unabhängig, im Widerspruch zu $\sum_{j\in\mathcal{M}} y_j = \sum_{j=1}^{r} y_j = 0$.

Also gilt für alle $j = 1,\ldots,r$

$$0 = y_j = \sum_{i=n_1+\ldots+n_{j-1}+1}^{n_1+\ldots+n_j}\mu_i x_i,$$

und da die B_j linear unabhängig sind, folgt $\mu_1,\ldots,\mu_n = 0$.

c. $\Rightarrow$ a.: Es sei $V = \bigoplus_{i=1}^{r}\mathrm{Eig}(f,\lambda_i)$ und $B_i = (x_{n_1+\ldots+n_{i-1}+1},\ldots,x_{n_1+\ldots+n_i})$ sei eine Basis von $\mathrm{Eig}(A,\lambda_i)$. Da die Summe direkt ist, ist mithin $B = B_1\cup\ldots\cup B_r$ eine Basis von V und es gilt,

$$M_B^B(f) = \bigoplus_{i=1}^{r}\lambda_i\mathbb{1}_{n_i}$$

hat Diagonalgestalt. □

Der letzte Schritt im Beweis liefert auch ein Verfahren zur Berechnung der Transformationsmatrix T, so daß $T^{-1}\circ A\circ T$ Diagonalgestalt hat.

Algorithmus 3.12 (Algorithmus zur Diagonalisierung)

INPUT: $A \in \mathrm{Mat}(n,K)$.

OUTPUT: 0, falls A über K nicht diagonalisierbar ist,
$1, D, T$, falls A diagonaliserbar ist, wobei D eine zu A konjugierte Diagonalmatrix ist, und T die zugehörige Transformationsmatrix mit $T^{-1}\circ A\circ T = D$.

1. **Schritt:** Berechne das charakteristische Polynom von A.
2. **Schritt:** Faktorisiere das charakteristische Polynom über K. Ist einer der Faktoren nicht linear, ist A nicht diagonalisierbar (nicht einmal trigonalisierbar) und man gebe 0 zurück. Sind alle Faktoren linear, so liefert die Faktorisierung die Eigenwerte $\lambda_1,\ldots,\lambda_r$ sowie ihre algebraischen Vielfachheiten $n_1,\ldots,n_r$.

3. Schritt: Bestimme für jeden Eigenwert λ_i eine Basis des Eigenraums $\mathrm{Eig}(A, \lambda_i)$ als $\mathrm{Lös}(A - \lambda_i \mathbb{1}_n, 0)$ - vgl. Algorithmus II.7.16 - sowie seine Dimension - vgl. Algorithmus II.6.3 -, d. h. die geometrische Vielfachheit von λ_i.

4. Schritt: Stimmt für jeden Eigenwert die algebraische Vielfachheit mit der geometrischen überein, so schreibe man die im 3. Schritt bestimmten Basen als Spalten in eine Matrix und erhält so T. Ferner erhält man D, indem man die Eigenwerte $\lambda_1, \ldots, \lambda_r$ entsprechend ihren algebraischen Vielfachheiten in der Diagonalen einer Nullmatrix einträgt.

Bemerkung 3.13 a. In Singular muß man die Veränderliche `t` des charakteristischen Polynoms als Parameter einführen, wenn man durch Polynome in `t` dividieren will. Dies ist gleichbedeutend dazu, daß man vom Grundkörper `K` zum Körper `K(t)` übergeht. – Z. B. `ring r=(0,t),x,dp;`.

b. Ob die Faktorisierung eines Polynoms algorithmisch möglich ist, hängt wesentlich vom Grundkörper ab. In Singular ist die Faktorisierung über $\mathbb{Q}$ sowie über endlichen Körpern möglich. Für $\mathbb{R}$ und $\mathbb{C}$ sind solche Algorithmen nicht bekannt.
In Singular erfolgt die Faktorisierung in irreduzible Faktoren mittels der Prozedur `factorize`.

c. Äquivalent zur Zerlegung des charakteristischen Polynoms in Linearfaktoren ist die exakte Bestimmung der Eigenwerte und ihrer Vielfachheiten. Die Approximation selbiger (etwa als Nullstellen des charakteristischen Polynoms über $\mathbb{R}$ oder $\mathbb{C}$) hilft im Allgemeinen nicht weiter. Denn im 3. Schritt müssen die Nullstellen genau bekannt sein, um die Eigenräume bestimmen zu können.
Selbst, wenn alle Eigenwerte verschieden sind, kann aus ihrer Approximation keine Transformationsmatrix ermittelt werden.

Beispiel 3.14
Gegeben sei die Matrix

$$A = \begin{pmatrix} 2 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 1 & -1 & -1 & 1 \end{pmatrix} \in \mathbf{Mat}(4, \mathbb{Q}).$$

Wir überprüfen mit Hilfe von Singular, ob A diagonalisierbar ist.

```
ring r=0,t,dp;
matrix A[4][4]=2,-1,0,0,0,1,0,0,0,0,2,0,1,-1,-1,1;
LIB "matrix.lib";
matrix B[4][4]=A-t*unitmat(4);
poly p=det(B);
↦   t4-6t3+13t2-12t+4
factorize(p);
```

```
↦      [1]:  // Faktoren von p
              [1]=1
              [2]=t-2
              [3]=t-1
       [2]:   // Vielfacheiten
              1,2,2
matrix E1=syz(A-unitmat(4));
print(E1); // Eig(A,1)
↦    1,  0
     1,  0
     0,  0
     0,  1
mat_rk(E1); // dim_Q (Eig(A,1))
↦   2
matrix E2=syz(A-2*unitmat(4));
print(E2); // Eig(A,2)
↦    1,  0
     0,  0
     1, -1
     0,  1
mat_rk(E2); // dim_Q (Eig(A,2))
↦   2
/* Da die algebraische und die geometrische Vielfachheit von 1 und 2
jeweils übereinstimmen, ist A diagonalisierbar. */
matrix T=concat(E1,E2);
print(T); // T ist die Transformationsmatrix, die A diagonalisiert.
↦    1,  0,  1,  0
     1,  0,  0,  0
     0,  0,  1, -1
     0,  1,  0,  1
/* Alternativ liefert die Singular-Prozedur diag_test aus linalg.lib
... */
LIB "linalg.lib";
diag_test(A);
↦   1 // D. h. A ist diagonalisierbar.
```

Beispiel 3.15

Die Matrix

$$A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \in \mathrm{Mat}(2,\mathbb{Q}) \subset \mathrm{Mat}(2,\mathbb{C})$$

hat als charakteristisches Polynom $\chi_A = t^2 - 2t + 2 = \big(t-(1+i)\big)\cdot\big(t-(1-i)\big)$, welches über $\mathbb{Q}$ irreduzibel ist, aber über $\mathbb{C}$ in zwei verschiedene Linearfaktoren zerfällt. Es folgt, daß A über $\mathbb{Q}$ nicht diagonalisierbar ist, aber über $\mathbb{C}$

gilt

$$A \sim \begin{pmatrix} 1+i & 0 \\ 0 & 1-i \end{pmatrix}.$$

Wir kommen jetzt zu einer wichtigen Beziehung zwischen Polynomen einerseits und Matrizen bzw. Endomorphismen eines endlich-dimensionalen Vektorraums andererseits. Wir erinnern uns, daß wir in Bemerkung 1.19 für beliebige K-Algebren B und $b \in B$ den Einsetzhomomorphismus $\phi_b : K[t] \to B : p \mapsto p(b)$ kennengelernt haben. Insbesondere haben wir für $B = \mathbf{Mat}(n, K)$ und $b = A$ also einen K-Algebrenhomomorphismus

$$\phi_A : K[t] \to \mathbf{Mat}(n, K) : p \mapsto p(A)$$

und für $B = \mathrm{End}_K(V)$ und $b = f$ einen K-Algebrenhomomorphismus

$$\phi_f : K[t] \to \mathrm{End}_K(V) : p \mapsto p(f).$$

Dabei gilt insbesondere $\phi_A(1) = \mathbb{1}_n$ und $\phi_f(1) = \mathrm{id}_V$.

Bemerkung 3.16

Ist $A = \bigoplus_{i=1}^{r} A_i$ eine Diagonalblockmatrix, so verifiziert man leicht, daß für $k \in \mathbb{N}$ gilt $A^k = \bigoplus_{i=1}^{r} A_i^k$, und damit, daß für ein Polynom $p \in K[t]$ gilt

$$p(A) = \bigoplus_{i=1}^{r} p(A_i).$$

Insbesondere gilt also für eine Diagonalmatrix $D = \bigoplus_{i=1}^{n} \lambda_i \mathbb{1}_1$, daß

$$p(D) = \bigoplus_{i=1}^{n} p(\lambda_i)\mathbb{1}_1 = \begin{pmatrix} p(\lambda_1) & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & p(\lambda_n) \end{pmatrix}$$

In der Tat kann man sogar zeigen, daß für eine Blockmatrix der Form

$$A = \left(\begin{array}{c|ccc} A_1 & * & \cdots & * \\ \hline 0 & \boxed{A_2} & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & \boxed{A_r} \end{array}\right) \in \mathbf{Mat}(n, K),$$

gilt, daß

$$p(A) = \left(\begin{array}{c|ccc} p(A_1) & * & \cdots & * \\ \hline 0 & \boxed{p(A_2)} & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & \boxed{p(A_r)} \end{array}\right) \in \mathbf{Mat}(n, K),$$

wobei sich die Sterne oberhalb der Blöcke verändert haben.

Damit gilt insbesondere, daß $p(A)$ eine obere Dreiecksmatrix ist, falls A eine solche war.

Nach Aufgabe 1.39 ist $K[t]$ ein Hauptidealring, so daß insbesondere die Ideale $\mathrm{Ker}(\phi_A)$ und $\mathrm{Ker}(\phi_f)$ je einen eindeutig bestimmten normierten Erzeuger besitzen - für die Eindeutigkeit siehe Aufgabe 1.40. Diese sind sicher nicht das Nullpolynom, da sonst die K-linearen Abbildungen ϕ_A bzw. ϕ_f injektiv sein müßten, was sie jedoch nicht sein können, da $K[t]$ als K-Vektorraum unendlich-dimensional ist, $\mathrm{Mat}(n, K)$ und $\mathrm{End}_K(V)$ jedoch nur endliche Dimension besitzen.

Definition 3.17

Das eindeutig bestimmte normierte Polynom μ_A (bzw. μ_f) mit $\mathrm{Ker}(\phi_A) = \langle \mu_A \rangle_{K[t]}$ (bzw. $\mathrm{Ker}(\phi_f) = \langle \mu_f \rangle_{K[t]}$) heißt das *Minimalpolynom* von A (bzw. f).

Bemerkung 3.18

Konjugierte Matrizen haben offenbar das gleiche Minimalpolynom.

Um das einzusehen, sei $B = T^{-1} \circ A \circ T$ mit $T \in \mathrm{Gl}_n(K)$. Dann gilt $B^r = \left(T^{-1} \circ A \circ T\right)^r = T^{-1} \circ A^r \circ T$ und somit folgt für ein beliebiges Polynom $p \in K[t]$

$$p(B) = T^{-1} \circ p(A) \circ T.$$

Da nun T invertierbar ist gilt somit $p(B) = 0$ genau dann, wenn $p(A) = 0$, d. h. $\mathrm{Ker}(\phi_A) = \mathrm{Ker}(\phi_B)$.

Lemma 3.19

Ein Polynom $q \in K[t] \setminus \{0\}$ ist genau dann das Minimalpolynom von A (bzw. f), wenn gilt:

(i) $\mathrm{lc}(q) = 1$,
(ii) $q(A) = 0$ *(bzw.* $q(f) = 0$*), und*
(iii) *es gibt* kein *Polynom* $p \in K[t] \setminus \{0\}$ *mit* $\deg(p) < \deg(q)$ *und* $p(A) = 0$ *(bzw.* $p(f) = 0$*).*

Beweis: Der Beweis geht analog zum Beweis von Lemma 1.35. □

Da $\dim_K \left(\mathrm{Mat}(n, K)\right) = n^2$ gilt, sind die $n^2 + 1$ Matrizen

$$\mathbb{1}_n = A^0, A^1, A^2, \ldots, A^{n^2}$$

in $\mathrm{Mat}(n, K)$ linear abhängig. D. h. es existieren $\lambda_0, \ldots, \lambda_{n^2} \in K$, nicht alle null, mit

$$\lambda_0 A^0 + \lambda_1 A^1 + \ldots + \lambda_{n^2} A^{n^2} = 0 \in \mathrm{Mat}(n, K).$$

Ein einfaches Dimensionsargument zeigt also, es gibt ein Polynom $0 \neq p = \lambda_{n^2} t^{n^2} + \ldots + \lambda_0 \in K[t]$ vom Grad kleiner gleich n^2 mit $p(A) = 0$. Der folgende wichtige Satz von Cayley-Hamilton besagt nun, daß es sogar ein Polynom vom Grad n gibt, das A annulliert.

Satz 3.20 (Cayley Hamilton)

Für das charakteristische Polynom χ_A von A (bzw. χ_f von f) gilt $\chi_A(A) = 0$ (bzw. $\chi_f(f) = 0$).

Ist explizit $\chi_A = (-1)^n t^n + \alpha_{n-1} t^{n-1} + \ldots + \alpha_0$*, dann gilt die* charakteristische Gleichung

$$(-1)^n A^n + \alpha_{n-1} A^{n-1} + \ldots + \alpha_1 A^1 + \alpha_0 \mathbb{1}_n = 0.$$

Beweis: Da für eine Basis D von V gilt $\chi_f(f) = 0$ genau dann, wenn $0 = \chi_f\big(M_D^D(f)\big) = \chi_{M_D^D(f)}\big(M_D^D(f)\big)$, reicht es, den Fall von Matrizen zu berücksichtigen.

Betrachte dazu die Matrix $B_t := A - t\mathbb{1}_n \in \mathrm{Mat}\big(n, K[t]\big) = \big(\mathrm{Mat}(n, K)\big)[t]$ sowie die Adjunkte $B_t^\# \in \mathrm{Mat}\big(n, K[t]\big) = \big(\mathrm{Mat}(n, K)\big)[t]$ von B_t, die auch *Busadjunkte* von A genannt wird. Dabei möchte die angegebene Mengengleichheit sagen, daß wir die Matrizen B_t und $B_t^\#$ einmal auffassen können als Matrizen mit Einträgen im Polynomring $K[t]$, zum anderen aber auch als Polynome über dem Matrizenring $\mathrm{Mat}(n, K)$, d. h. als Polynome, deren Koeffizienten Matrizen sind.

Nach dem Satz über die Adjunkte II.10.5 in $\mathrm{Mat}\big(n, K[t]\big)$ gilt die Adjunktengleichung

$$B_t \circ B_t^\# = (A - t\mathbb{1}_n) \circ (A - t\mathbb{1}_n)^\# = \det(A - t\mathbb{1}_n) \cdot \mathbb{1}_n = \chi_A \cdot \mathbb{1}_n. \tag{43}$$

Beachten wir nun noch, daß die Einträge von $B_t^\#$ Determinanten von gewissen $(n-1) \times (n-1)$-Matrizen von B_t sind, also Polynome vom Grad kleiner gleich $n-1$, so hat $B_t^\#$ also auch als Polynom über dem Matrizenring höchstens den Grad $n-1$, d. h. es gibt Matrizen $B_0, \ldots, B_{n-1} \in \mathrm{Mat}(n, K)$ mit

$$B_t^\# = B_{n-1} t^{n-1} + \ldots + B_1 t + B_0 \in \big(\mathrm{Mat}(n, K)\big)[t].$$

Ist nun $\chi_A = (-1)^n t^n + \alpha_{n-1} t^{n-1} + \ldots + \alpha_0$, so folgt aus der Adjunktengleichung (43)

$$\begin{aligned} q &:= (A - \mathbb{1}_n t) \circ \big(B_{n-1} t^{n-1} + \ldots + B_1 t + B_0\big) \\ &= (-1)^n \mathbb{1}_n t^n + \alpha_{n-1} \mathbb{1}_n t^{n-1} + \ldots + \alpha_0 \mathbb{1}_n \in \big(\mathrm{Mat}(n, K)\big)[t] \end{aligned} \tag{44}$$

durch Koeffizientenvergleich für die B_i, $i = 0, \ldots, n-1$:

$$\begin{aligned} -B_{n-1} &= (-1)^n \mathbb{1}_n \\ AB_{n-1} - B_{n-2} &= \alpha_{n-1} \mathbb{1}_n \\ AB_{n-2} - B_{n-3} &= \alpha_{n-2} \mathbb{1}_n \\ &\vdots \\ AB_1 - B_0 &= \alpha_1 \mathbb{1}_n \\ AB_0 &= \alpha_0 \mathbb{1}_n \end{aligned} \tag{45}$$

Multipliziert man die i-te Zeile in (45) mit A^{n-i+1} und summiert die beiden Seiten auf, so erhält man die Behauptung:

$$\begin{array}{rcl} -A^nB_{n-1} & = & (-1)^nA^n \\ A^nB_{n-1} - A^{n-1}B_{n-2} & = & \alpha_{n-1}A^{n-1} \\ A^{n-1}B_{n-2} - A^{n-2}B_{n-3} & = & \alpha_{n-2}A^{n-2} \\ \vdots & & \\ A^2B_1 - AB_0 & = & \alpha_1 A \\ AB_0 & = & \alpha_0 \mathbb{1}_n \\ \hline 0 & = & \chi_A(A). \end{array}$$

□

Aus dem Beweis ergeben sich einige interessante Bemerkungen.

Bemerkung 3.21 a. Alternativ zum Koeffizientenvergleich im letzten Beweis könnte man auch in das Polynom q in (44) die Matrix A für t einsetzen (im Sinne[4] von Definition 1.17), so daß die linke Seite die Nullmatrix ergibt, da $A - \mathbb{1}_n \circ A = 0$. Die rechte Seite ergibt hingegen gerade $\chi_A(A)$. Womit die Behauptung ebenfalls gezeigt wäre.

b. Kennt man das charakteristische Polynom $\chi_A = (-1)^n t^n + \alpha_{n-1}t^{n-1} + \ldots + \alpha_0$, so läßt sich daraus mittels (44) und der Rekursionsformel (45) die Busadjunkte

$$(A - t\mathbb{1}_n)^{\#} = B_{n-1}t^{n-1} + \ldots + B_1 t + B_0$$

von A bestimmen. Für die B_{n-k}, $k = 1, \ldots, n$, gilt dabei explizit:

$$B_{n-k} = (-1)^{n+1}A^{k-1} - \alpha_{n-1}A^{k-2} - \ldots - \alpha_{n-k+1}A^0,$$

und speziell gilt für die Adjunkte von A

$$A^{\#} = B_0 = (-1)^{n+1}A^{n-1} - \alpha_{n-1}A^{n-2} - \ldots - \alpha_1 A^0.$$

Diese Formel zur Berechnung der Adjunkten von A ist weit effizienter, als die Minoren[5] sämtlicher Streichungsmatrizen zu berechnen.

c. Man beachte, daß der folgende *offensichtliche* Beweis für $\chi_A(A) = 0$, nämlich "$\chi_A(A) = \det(A - A * \mathbb{1}_n) = \det(0) = 0$", falsch ist, da "$*$" beim Einsetzen von A in $\det(A - t\mathbb{1}_n) \in K[t]$ eben *nicht* die Matrixmultiplikation ist! Man beachte ferner, daß die Gleichung auch schon deshalb keinen Sinn machen kann, da $\chi_A(A)$ die Nullmatrix ist, während $\det(0)$ die Null in K ist.

Korollar 3.22

Ist $\lambda \cdot p_1^{\nu_1} \cdots p_r^{\nu_r}$ *die eindeutige Primfaktorzerlegung von* χ_A *(bzw.* χ_f*), so gilt* $\mu_A = p_1^{\mu_1} \cdots p_r^{\mu_r}$ *(bzw.* $\mu_f = p_1^{\mu_1} \cdots p_r^{\mu_r}$*) für geeignete* $1 \leq \mu_i \leq \nu_i$, $i = 1, \ldots, r$.

Insbesondere teilt das Minimalpolynom von A *(bzw.* f*) das charakteristische Polynom von* A *(bzw.* f*).*

[4] Für das einfache Polynom $p = B \cdot t \in \big(\mathrm{Mat}(n, K)\big)[t]$ ergibt sich dabei dann $p(A) = B \circ A$.

[5] Für eine Definition des Begriffes der Minoren einer Matrix siehe Definition IV.3.15.

Beweis: Nach Wahl einer Basis B von V können wir annehmen, daß f durch die Matrix $A = M_B^B(f)$ gegeben ist und führen den Beweis auf die entsprechende Aussage für Matrizen zurück.

Der Satz von Cayley-Hamilton besagt, daß $\chi_A \in \mathrm{Ker}(\phi_A) = \langle \mu_A \rangle_{K[t]}$ gilt. Also gibt es ein $h \in K[t]$ mit $\chi_A = \mu_A \cdot h$.

Daraus folgt $\mu_A = p_1^{\mu_1} \cdots p_r^{\mu_r}$ für $0 \leq \mu_i \leq \nu_i$, $i = 1, \ldots, r$, geeignet. Wir müssen zeigen, daß jedes p_i in μ_A auch vorkommt, d. h., daß $\mu_i \geq 1$ für alle $i = 1, \ldots, r$.

Nehmen wir an, daß es ein i mit $\mu_i = 0$ gibt. Dann sind μ_A und p_i teilerfremde Polynome, also existieren nach Satz 1.28 Polynome $p, q \in K[t]$ mit

$$1 = p\mu_A + qp_i.$$

Am leichtesten führt man dies zum Widerspruch, indem man zum algebraischen Abschluß $\overline{K}$ von K übergeht.[6] Da p_i vom Grad $\deg(p_i) \geq 1$ ist, besitzt es eine Nullstelle $\lambda \in \overline{K}$. Aus der obigen Gleichung folgt, daß dann $\mu_A(\lambda) \neq 0$ gelten muß.

Wegen $p_i(\lambda) = 0$, ist dann aber auch $\chi_A(\lambda) = 0$ und somit ist λ ein Eigenwert von $A \in \mathrm{Mat}\left(n, \overline{K}\right)$. Sei nun $0 \neq x \in \overline{K}^n$ ein Eigenvektor von A zum Eigenwert λ. Dann gilt für das Polynom $\mu_A = \sum_{i=0}^m a_i t^i \in K[t] \subseteq \overline{K}[t]$

$$\mu_A(A)x = \sum_{i=0}^{m} a_i\left(A^i x\right) = \sum_{i=0}^{m} a_i\left(\lambda^i x\right) = \mu_A(\lambda) \cdot x \neq 0,$$

im Widerspruch zu $\mu_A(A) = 0$. □

Gehen wir zum algebraischen Abschluß $\overline{K}$ von K über, so erhalten wir das folgende Korollar.

Korollar 3.23
Hat das charakteristische Polynom von A (bzw. f) über $\overline{K}$ die Gestalt $(-1)^n \cdot (t-\lambda_1)^{n_1} \cdots (t-\lambda_s)^{n_s}$, $\lambda_i \in \overline{K}$, so gilt für das Minimalpolynom $\mu = \mu_A$ (bzw. $\mu = \mu_f$)

$$\mu = (t-\lambda_1)^{m_1} \cdots (t-\lambda_s)^{m_s}, \quad 1 \leq m_i \leq n_i, \quad i = 1, \ldots, s.$$

Beispiel 3.24 a. Ist $A = \lambda \mathbb{1}_n \in \mathrm{Mat}(n, K)$ eine Diagonalmatrix mit gleichen Diagonalelementen, so gilt offenbar

$$\chi_A = (\lambda - t)^n \quad \text{und} \quad \mu_A = t - \lambda,$$

denn $A - \lambda \mathbb{1}_n = 0$. In diesem Fall ist μ_A quadratfrei.

b. Sei $\lambda \in K$ und

$$J := J_n(\lambda) := \begin{pmatrix} \lambda & 1 & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \ldots & \ldots & 0 & \lambda \end{pmatrix} \in \mathrm{Mat}(n, K),$$

[6] Wem dies nicht behagt, der mag getrost K als Unterkörper von $\mathbb{C}$ und $\overline{K} = \mathbb{C}$ annehmen.

d. h. $J_n(\lambda)$ hat auf der Hauptdiagonalen den Wert λ und auf der oberen Nebendiagonalen Einsen stehen, ansonsten nur Nullen. Wir nennen $J_n(\lambda)$ einen *Jordanblock* (oder eine *Jordanzelle* oder ein *Jordankästchen*) der Größe n zum Eigenwert λ .
Offenbar gilt wieder

$$\chi_J = (\lambda - t)^n.$$

Nach Korollar 3.22 ist mithin $\mu_J = (t-\lambda)^m$ für ein $1 \leq m \leq n$. Dabei ist m die kleinste natürliche Zahl mit $(J - \lambda \mathbb{1}_n)^m = 0$. Nun ist aber

$$J - \lambda \mathbb{1}_n = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & 0 \end{pmatrix} =: N$$

und man sieht mittels einer einfachen Induktion, daß $N^p \neq 0$ für $p < n$, aber $N^n = 0$ (vgl. Aufgabe I.8.28). Also gilt

$$\mu_J = (t-\lambda)^n.$$

c. Ist $A = A_1 \oplus \ldots \oplus A_r \in \mathrm{Mat}(n, K)$ eine Diagonalblockmatrix mit $A_i \in \mathrm{Mat}(n_i, K)$, so folgt aus der Definition des charakteristischen Polynoms unmittelbar (vgl. Aufgabe 2.24)

$$\chi_A = \prod_{i=1}^{r} \chi_{A_i}.$$

Eine entsprechend einfache Formel für das Minimalpolynom gilt nicht. Sei etwa $A_1 = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \in \mathrm{Mat}(2, K)$ und $A_2 = (1) \in \mathrm{Mat}(1, K)$, dann gilt für $A = A_1 \oplus A_2$

$$\mu_A = (t-1)^2 \neq (t-1)^3 = \mu_{A_1} \cdot \mu_{A_2}.$$

Vgl. Aufgabe 3.42 für eine korrekte Verallgemeinerung des Verfahrens.

Bemerkung 3.25

Zur praktischen Berechnung des Minimalpolynoms von $A \in \mathrm{Mat}(n, K)$ kann man wie folgt vorgehen. Aufgrund des Satzes von Cayley-Hamilton wissen wir, daß die Matrizen $A^0, \ldots, A^n$ linear abhängig sind. Fassen wir die Matrix A^i als einen *langen* Spaltenvektor in K^{n^2} auf und bezeichnen wir diesen mit x_i, dann suchen wir das minimale m, so daß $x_0, \ldots, x_m$ linear abhängig sind, und wir suchen ferner geeignete $\beta_0, \ldots, \beta_{m-1}$ mit

$$x_m + \beta_{m-1} x_{m-1} + \ldots + \beta_0 x_0 = 0.$$

Dies ist dann gleichbedeutend damit, daß

$$t^m + \beta_{m-1} t^{m-1} + \ldots + \beta_0 \in K[t]$$

das gesuchte Minimalpolynom von A ist.

Bezeichne $X = (x_0 \dots x_n) \in \mathrm{Mat}\left(n^2 \times (n+1), K\right)$ die Matrix, deren Spalten $x_0, \dots, x_n$ sind, dann suchen wir eine Lösung des linearen Gleichungssystems

$$X \cdot \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_n \end{pmatrix} = 0 \in K^{n^2} \tag{46}$$

mit $\beta_{m+1} = \ldots = \beta_n = 0$ und $\beta_m = 1$ und so, daß m minimal mit dieser Eigenschaft ist. Da $x_0, \dots, x_{m-1}$ nach Definition von m linear unabhängig sind, $x_0, \dots, x_m$ aber linear abhängig, so bedeutet dies, daß in einer ZSF von X die Zahlen $1, \dots, m$ Pivotindizes sind, während $m+1$ kein Pivotindex mehr ist.

Berechnet man eine Parametrisierung der Lösungsmenge von (46) mittels des Algorithmus II.7.16, so erhalten wir den gesuchten Koeffizientenvektor β als ersten Spaltenvektor der Parametrisierungsmatrix.

Dies führt zu folgendem Algorithmus zur Berechnung des Minimalpolynoms einer Matrix $A \in \mathrm{Mat}(n, K)$.

Algorithmus 3.26 (Algorithmus zur Berechnung des Minimalpolynoms)
INPUT: $A \in \mathrm{Mat}(n, K)$

OUTPUT: μ_A

1. **Schritt:** Falls A nicht quadratisch ist, gib 0 zurück.
2. **Schritt:** Bilde die Potenzen $A^0, \dots, A^n$ und schreibe die Matrizen in Form von Spaltenvektoren der Länge n^2 in eine Matrix $B \in \mathrm{Mat}\left(n^2 \times (n+1), K\right)$.
3. **Schritt:** Berechne eine Parametrisierung von $\mathrm{Lös}(B, 0)$.
4. **Schritt:** Verwende die Koeffizienten der ersten Spalte der Parametrisierung als Koeffizienten eines Polynoms und gib dieses zurück.

Singular besitzt bereits ein Kommando, nämlich `syz`, zur Berechnung der Parametrisierung einer Lösungsmenge. Wir demonstrieren die Verwendung des Kommandos an einem Beispiel.

Beispiel 3.27

```
LIB "matrix.lib";
ring r=0,t,lp;
matrix A[3][3]=1,1,0,0,1,0,0,0,1;
print(A);
↦    1  1  0
     0  1  0
     0  0  1
matrix X=transpose(flatten(power(A,0)));
int i;
for (i=1;i<=3;i++)
{ X=concat(X,transpose(flatten(power(A,i)))); }
print(X); // Die i-te Spalte ist A^i.
```

```
↦    1  1  1  1
     0  1  2  3
     0  0  0  0
     0  0  0  0
     1  1  1  1
     0  0  0  0
     0  0  0  0
     0  0  0  0
     1  1  1  1
module L=syz(X); // Bestimmt eine Parametrisierung von Lös(X,0).
print(L);
↦     1   0
     -2   1
      1  -2
      0   1
```

Also ist $\mu_A = t^2 - 2t + 1$ das Minimalpolynom von A.

Zur Kontrolle kann man die ZSF von X ansehen:

```
print(gauss_row(X));
↦    1,  0,  -1,  -2,
     0,  1,  2,   3,
     0,  0,  0,   0,
         ⋮
     0,  0,  0,   0
```

$m + 1 = 3$ ist der kleinste Nicht-Pivot-Index, und $(1, -2, 1, 0)^t$ ist somit in der Tat die gesuchte Lösung.

Das Minimalpolynom läßt sich also durch Lösen eines linearen Gleichungssystems bestimmen. Wir wollen jetzt zeigen, daß das Minimalpolynom ein effizientes Kriterium dafür liefert, ob eine Matrix $A \in \mathrm{Mat}(n, K)$ diagonalisierbar ist oder nicht.

Satz 3.28

A (bzw. f) ist genau dann über K diagonalisierbar, wenn das Minimalpolynom $\mu = \mu_A$ (bzw. $\mu = \mu_f$) in paarweise verschiedene Linearfaktoren zerfällt, d. h. wenn es $\lambda_1, \ldots, \lambda_r \in K$ mit $\lambda_i \neq \lambda_j$, $i \neq j$, gibt mit $\mu = (t - \lambda_1) \cdots (t - \lambda_r)$.

Beweis: Ist A diagonalisierbar, dann ist A konjugiert zu einer Diagonalmatrix B der Form

$$B = \bigoplus_{i=1}^{r} \lambda_i \mathbb{1}_{m_i},$$

mit $\lambda_i \neq \lambda_j$ für $i \neq j$. Nach Bemerkung 3.18 ist $\mu_A = \mu_B$ und es folgt aus Aufgabe 3.42 und Beispiel 3.24 mit Induktion

$$\mu_B = (t - \lambda_1) \cdots (t - \lambda_r).$$

Für den Beweis der Umkehrung greifen wir etwas vor und verwenden Proposition 3.33. Da das Minimalpolynom in paarweise verschiedene Linearfaktoren zerfällt, folgt damit, daß V die direkte Summe der Eigenräume von A ist und somit folgt aus Satz 3.10, daß A diagonalisierbar ist. □

Korollar 3.29
Genau dann ist A über dem algebraischen Abschluß $\overline{K}$ von K diagonalisierbar[7]*, wenn μ_A quadratfrei ist.*

Beweis: Dies folgt unmittelbar aus den Sätzen 1.36 und 3.28. □

Für den Beweis der Rückrichtung in Satz 3.28 brauchen wir einen neuen Begriff, der auch im folgenden Abschnitt für die Jordansche Normalform von Bedeutung sein wird. Für $\lambda \in K$ haben wir aufsteigende Folgen von Untervektorräumen von V (vgl. Lemma 3.31 und Aufgabe 3.49)

$$\operatorname{Ker}\left(f-\lambda\operatorname{id}_V\right) \subseteq \operatorname{Ker}\left((f-\lambda\operatorname{id}_V)^2\right) \subseteq \operatorname{Ker}\left((f-\lambda\operatorname{id}_V)^3\right) \subseteq \ldots \subseteq V$$

und

$$\operatorname{Lös}\left(A-\lambda\mathbb{1}_n,0\right) \subseteq \operatorname{Lös}\left((A-\lambda\mathbb{1}_n)^2,0\right) \subseteq \operatorname{Lös}\left((A-\lambda\mathbb{1}_n)^3,0\right) \subseteq \ldots \subseteq V$$

Die Vereinigung all dieser Unterräume ist offenbar wieder ein Unterraum und führt zu folgender Definition.

Definition 3.30
Für $\lambda \in K$ heißen

$$\operatorname{Hau}(A,\lambda) = \bigcup_{k\in\mathbb{N}} \operatorname{Lös}\left((A-\lambda\mathbb{1}_n)^k,0\right) \quad \text{und} \quad \operatorname{Hau}(f,\lambda) = \bigcup_{k\in\mathbb{N}} \operatorname{Ker}\left((f-\lambda\operatorname{id}_V)^k\right)$$

der *Hauptraum* oder *verallgemeinerte Eigenraum* von A bzw. f zu λ.

Lemma 3.31
Es sei $\lambda \in K$ gegeben.

a. *Dann gibt es ein $0 \le m \le n$ mit* $\operatorname{Ker}\left((f-\lambda\operatorname{id}_V)^0\right) \subsetneqq \ldots \subsetneqq \operatorname{Ker}\left((f-\lambda\operatorname{id}_V)^{m-1}\right) \subsetneqq \operatorname{Ker}\left((f-\lambda\operatorname{id}_V)^m\right) = \operatorname{Ker}\left((f-\lambda\operatorname{id}_V)^k\right)$ *für alle $k > m$.*
b. *Für jedes $k \in \mathbb{N}$ gilt,* $\operatorname{Ker}\left((f-\lambda\operatorname{id}_V)^k\right)$ *ist f-invariant.*
Insbesondere sind also Eigenräume und Haupträume von f f-invariant.
c. *Für $l = \operatorname{mult}(\mu_f,\lambda)$ gilt* $\operatorname{Hau}(f,\lambda) = \operatorname{Ker}\left((f-\lambda\operatorname{id}_V)^l\right)$.

Die entsprechenden Aussagen für eine Matrix $A \in \operatorname{Mat}(n,K)$ gelten analog.

Beweis: Durch Betrachtung von f_A ergibt sich die Aussage für eine Matrix A unmittelbar aus der entsprechenden Aussage für Endomorphismen.

a. Sei $m \in \mathbb{N}$ minimal mit der Eigenschaft, daß $\operatorname{Ker}\left((f-\lambda\operatorname{id}_V)^m\right) = \operatorname{Ker}\left((f-\lambda\operatorname{id}_V)^{m+1}\right)$; ein solches m gibt es, da $\dim_K(V) < \infty$. Es reicht, zu zeigen, daß dann $\operatorname{Ker}\left((f-\lambda\operatorname{id}_V)^m\right) = \operatorname{Ker}\left((f-\lambda\operatorname{id}_V)^{m+k}\right)$ für alle $k > 0$, wobei "$\subseteq$" offenbar gilt. Wir führen den Beweis mittels Induktion nach k, und

[7] D. h. $A \in \operatorname{Mat}(n,K) \subseteq \operatorname{Mat}\left(n,\overline{K}\right)$ aufgefaßt als Matrix mit Einträgen in $\overline{K}$ kann mittels einer invertierbaren Matrix $T \in \operatorname{Gl}_n\left(\overline{K}\right)$ auf Diagonalgestalt gebracht werden.

der Induktionsanfang für $k = 1$ ist nach Voraussetzung erfüllt. Sei nun $k > 1$ und $x \in \operatorname{Ker}\big((f-\lambda\,\mathrm{id}_V)^{m+k}\big)$ gegeben. Dann gilt, unter Verwendung der Induktionsvoraussetzung,

$$(f-\lambda\,\mathrm{id}_V)(x) \in \operatorname{Ker}\big((f-\lambda\,\mathrm{id}_V)^{m+k-1}\big) = \operatorname{Ker}\big((f-\lambda\,\mathrm{id}_V)^{m}\big).$$

Mithin ist $(f-\lambda\,\mathrm{id}_V)^{m+1}(x) = 0$ und $x \in \operatorname{Ker}\big((f-\lambda\,\mathrm{id}_V)^{m+1}\big) = \operatorname{Ker}\big((f-\lambda\,\mathrm{id}_V)^{m}\big)$, was zu zeigen war.

b. Da f mit Potenzen von f und mit der Identität vertauschbar ist, gilt für $k \in \mathbb{N}$ und $x \in \operatorname{Ker}\big((f-\lambda\,\mathrm{id}_V)^{k}\big)$

$$(f-\lambda\,\mathrm{id}_V)^k\big(f(x)\big) = f\big((f-\lambda\,\mathrm{id}_V)^k(x)\big) = f(0) = 0,$$

woraus die Behauptung folgt.

c. Angenommen, das sei nicht der Fall, dann gibt es ein $x \in V$ mit $y := (f-\lambda\,\mathrm{id}_V)^l(x) \neq 0$, aber $(f-\lambda\,\mathrm{id}_V)^k(x) = 0$ für alle $k > l$. Da λ l-fache Nullstelle von μ_f ist, gibt es ein $q \in K[t]$ mit

$$\mu_f = (t-\lambda)^l \cdot q,$$

wobei λ keine Nullstelle von q ist. Dann sind q und $(t-\lambda)^{k-l}$, für $k > l$, aber teilerfremd und nach Satz 1.28 gibt es Polynome $p_1, p_2 \in K[t]$ mit

$$(t-\lambda)^{k-l} \cdot p_1 + q \cdot p_2 = 1.$$

Es folgt

$$\big(p_1(f) \circ (f-\lambda\,\mathrm{id}_V)^{k-l} + p_2(f) \circ q(f)\big)(y) = y. \tag{47}$$

Andererseits gilt aber $(f-\lambda\,\mathrm{id}_V)^{k-l}(y) = (f-\lambda\,\mathrm{id}_V)^k(x) = 0$ sowie

$$0 = \mu_f(f)(x) = \big(q(f) \circ (f-\lambda\,\mathrm{id}_V)^l\big)(x) = q(f)(y).$$

Aus (47) folgt damit $y = 0$, im Widerspruch zur Voraussetzung.

□

Definition 3.32

Sei $\lambda \in K$ gegeben. Die natürliche Zahl $m \in \mathbb{N}$ aus Lemma 3.31 a. heißt der *Nilpotenzindex* des Endomorphismus $f-\lambda\,\mathrm{id}_V$. (Vgl. auch Aufgabe 3.49.)

Proposition 3.33

Es sei f ein Endomorphismus, dessen charakteristisches Polynom χ_f über K in Linearfaktoren zerfällt, d. h. es gibt $\lambda_i \in K$ und $0 < m_i \leq n_i$, $i = 1, \ldots, r$, so daß

$$\chi_f = (-1)^n \cdot (t-\lambda_1)^{n_1} \cdots (t-\lambda_r)^{n_r} \quad \textit{und} \quad \mu_f = (t-\lambda_1)^{m_1} \cdots (t-\lambda_r)^{m_r}.$$

Dann gelten:

a. $V = \bigoplus_{i=1}^{r} \operatorname{Hau}(f, \lambda_i)$,
b. m_i *ist der Nilpotenzindex von* $f-\lambda_i\,\mathrm{id}_V$ *und*
c. $n_i = \dim_K\big(\operatorname{Hau}(f, \lambda_i)\big)$.

Die analogen Aussagen für eine Matrix $A \in \operatorname{Mat}(n, K)$, deren charakteristisches Polynom zerfällt, gelten analog.

Beweis: Wir beschränken uns auf den Beweis im Falle eines Endomorphismus f, da sich die entsprechende Aussage für eine Matrix A dann unmittelbar auf f_A zurückführen läßt.

a. Wir setzen $V_i := \mathrm{Hau}(f, \lambda_i)$ und

$$q_i := \frac{\mu_f}{(t-\lambda_i)^{m_i}} = \prod_{j \neq i} (t-\lambda_j)^{m_j}.$$

Da $\mathrm{ggT}(q_1, \ldots, q_r) = 1$, folgt aus Aufgabe 1.42 die Existenz von Polynomen $p_1, \ldots, p_r \in K[t]$ mit

$$p_1 q_1 + \ldots + p_r q_r = 1.$$

Setzt man $Q_i := q_i p_i$, dann folgt

$$\sum_{i=1}^{r} Q_i(f) = \mathrm{id}_V .$$

Wegen $(f - \lambda_i \mathrm{id}_V)^{m_i} \circ Q_i(f) = p_i(f) \circ \mu_f(f) = 0$ folgt mit Lemma 3.31 c.

$$\mathrm{Im}\left(Q_i(f)\right) \subseteq \mathrm{Ker}(f - \lambda_i \mathrm{id}_V)^{m_i} = V_i.$$

Außerdem gilt für jedes $x \in V$

$$x = \sum_{i=1}^{r} Q_i(f)(x) \in \sum_{i=1}^{r} \mathrm{Im}\left(Q_i(f)\right) \subseteq \sum_{i=1}^{r} V_i,$$

d. h. $V = \sum_{i=1}^{r} V_i$.
Es bleibt zu zeigen, daß die Summe direkt ist. Nun gilt

$$\left(Q_i(f)\right)_{V_j} = \begin{cases} \mathrm{id}_{V_j}, & \text{für } i = j, \\ 0, & \text{für } i \neq j, \end{cases} \tag{48}$$

denn für $i \neq j$ und $x \in V_j$ gilt $Q_i(f)(x) = p_i(f) \circ \prod_{k \neq i,j} (f - \lambda_k \mathrm{id}_V)^{m_k} \circ (f - \lambda_j \mathrm{id}_V)^{m_j}(x) = 0$ und für $x \in V_i$ gilt deshalb

$$Q_i(f)(x) = \sum_{j=1}^{r} Q_j(f)(x) = \mathrm{id}_V(x).$$

(Insbesondere gilt $Q_i(f)^2 = Q_i(f)$ und damit ist $Q_i(f)$ die Projektion auf V_i mit Kern $\bigoplus_{j \neq i} V_j$ - vgl. Aufgabe II.1.23.)
Ist nun also $0 = x_1 + \ldots + x_r$ eine Darstellung des Nullvektors mit $x_i \in V_i$, $i = 1, \ldots, r$, dann gilt

$$0 = Q_i(f)(x_1 + \ldots + x_r) = Q_i(f)(x_i) = x_i.$$

Die Summe ist also direkt.

b. Sei $i \in \{1, \ldots, r\}$ gegeben und sei $m \in \mathbb{N}$ der Nilpotenzindex von $f - \lambda_i \mathrm{id}_V$, dann gilt nach Lemma 3.31 c. $m_i \geq m$.
Angenommen, $m_i > m$. Sei nun $x = x_1 + \ldots + x_r \in V$ beliebig mit $x_j \in V_j$ für $j = 1, \ldots, r$. Dann gilt $(f - \lambda_i \mathrm{id}_V)^m(x_i) = 0$ und für $j \neq i$ gilt $q_i(f)(x_j) = 0$, also folgt für $\mu := (t - \lambda_i)^m q_i \in K[t]$

$$\mu(f)(x) = 0, \text{ d. h. } \mu(f) = 0,$$

aber $0 \leq \deg(\mu) < \deg(\mu_f)$, im Widerspruch zur Definition des Minimalpolynoms.

c. Da $V = \bigoplus_{i=1}^{r} V_i$ und da die V_i nach Lemma 3.31 f-invariant sind, folgt aus Aufgabe 2.24

$$\chi_f = \prod_{i=1}^{r} \chi_{f_{V_i}}.$$

Aus b. folgt $\mu_{f_{V_i}} = (t-\lambda_i)^{m_i}$ und damit folgt aus Korollar 3.22 $\chi_{f_{V_i}} = (t-\lambda_i)^{\dim_K(V_i)}$, was zu zeigen war.

□

Aus Proposition 3.33 Teil b. und c. folgt, da die Haupträume von f f-invariant sind, unmittelbar das folgende Korollar.

Korollar 3.34

Sei f wie in Proposition 3.33, dann gilt

$$\chi_{f_{\mathrm{Hau}(f,\lambda_i)}} = (-1)^{n_i}(t-\lambda_i)^{n_i}$$

und

$$\mu_{f_{\mathrm{Hau}(f,\lambda_i)}} = (t-\lambda_i)^{m_i}.$$

Beispiel 3.35

Wir wollen nun unter Ausnutzung von Singular zeigen, daß die Matrix

$$A = \begin{pmatrix} 3 & 2 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix} \in \mathrm{Mat}(3,\mathbb{Q})$$

nicht diagonalisierbar ist, und wir wollen die zugehörigen Haupträume bestimmen.

Aus der Gestalt der Matrix ergibt sich unmittelbar $\chi_A = (t-3)^2(t-2)$. Wie im letzten Beispiel 3.27 bestimmt man das Minimalpolynom und stellt fest, daß es gleich dem charakteristischen Polynom ist. 3 ist also eine doppelte Nullstelle des Minimalpolynoms und somit ist A nicht diagonalisierbar nach Satz 3.28

Wir wollen nun die beiden Haupträume von A zu den Eigenwerten 2 und 3 bestimmen.

```
ring r= 0,t,lp;
matrix A[3][3]=3,2,1,0,2,1,0,0,3;
matrix B=A-3*unitmat(3);
syz(B);        // Eig(A,3)
↦   gen(3)+gen(1)        // gen(i) ist der i-te Einheitsvektor in Q^3,
                         // also hier (1,0,1)^t.
syz(B*B);      // Lös((A-3*1_3)^2,0).
↦    3*gen(2)+4*gen(1)        // (4,3,0)^t, (0,-3,4)^t
     4*gen(3)-3*gen(2)
```

```
syz(B*B*B);          // Lös((A − 3𝟙3)^3, 0).
↦    3 * gen(2) + 4 * gen(1)        // D. h. syz(B*B)=syz(B*B*B).
     4 * gen(3) − 3 * gen(2)
```

Daß Lös$\big((A-3\mathbb{1}_3)^2,0\big)$ = Lös$\big((A-3\mathbb{1}_3)^3,0\big)$ wußten wir allerdings schon. Es gilt also

$$\mathrm{Hau}(A,3)=\big\langle(4,3,0)^t,(0,-3,4)^t\big\rangle$$

ist zwei-dimensional, während $\mathrm{Eig}(A,3)=\big\langle(1,0,1)^t\big\rangle$ ein-dimensional ist.

Analog bestimmt man den Hauptraum von A zum Eigenwert 2, der in diesem Fall gleich dem Eigenraum ist, als

$$\mathrm{Hau}(A,2)=\mathrm{Eig}(A,2)=\big\langle(-2,1,0)^t\big\rangle.$$

Die Haupt- und Eigenräume kann man nur berechnen, wenn die Eigenwerte exakt bekannt sind. Wir wollen jetzt einen Algorithmus kennenlernen, mit dessen Hilfe wir *ohne* Kenntnis der Eigenwerte feststellen können, ob eine Matrix A diagonalisierbar ist oder nicht. Außerdem liefert dieser Algorithmus im Falle der Diagonalisierbarkeit eine zu A konjugierte Diagonalgestalt, wobei die Eigenwerte dann eventuell numerisch zu approximieren sind. Dazu greifen wir auf das Kriterium für die Quadratfreiheit eines Polynoms aus Satz 1.36 zurück.

Algorithmus 3.36 (Algorithmus zur Diagonalisierbarkeit)
INPUT: $A\in\mathrm{Mat}(n,\mathbb{Q})$.

OUTPUT: 0, falls A über $\mathbb{C}$ nicht diagonalisierbar ist,
$1,D$, falls A über $\mathbb{C}$ diagonalisierbar ist, wobei D eine zu A konjugierte Diagonalmatrix mit näherungsweise bestimmten Eigenwerten ist.

1. **Schritt:** Bestimme das Minimalpolynom μ_A von A sowie die formale Ableitung $D(\mu_A)$ von μ_A und den ggT g von μ_A und $D(\mu_A)$.
2. **Schritt:** Ist $\deg(g)>0$, dann gebe 0 zurück, sonst gehe zu Schritt 3.
3. **Schritt:** Man bestimme das charakteristische Polynom χ_A von A und faktorisiere es über $\mathbb{Q}$.
4. **Schritt:** Für jeden irreduziblen Faktor von χ_A bestimme man näherungsweise die (paarweise verschiedenen !) Nullstellen. Diese schreibe man mit der Vielfachheit, mit der der jeweilige irreduzible Faktor als Faktor von χ_A vorkam, in eine Liste.
5. **Schritt:** Die Liste wird am Ende genau n Einträge enthalten, die man auf die Diagonale einer Matrix D schreibt. Sodann gebe man ein Liste, bestehend aus der Zahl 1 und aus D, zurück.

Bemerkung 3.37 a. Der Algorithmus läßt sich in Singular problemlos implementieren und greift dabei auf die Algorithmen I.6.3 und 3.26 zurück sowie auf singularinterne Prozeduren zur Faktorisierung von Polynomen

über $\mathbb{Q}$ und zur näherungsweisen Bestimmung von Nullstellen von Polynomen. Um zu garantieren, daß die näherungsweise bestimmten Eigenwerte auf alle Fälle mit der richtigen Vielfachheit angegeben werden, machen wir uns wesentlich zu Nutze, daß verschiedene irreduzible Faktoren auch verschiedene Nullstellen haben und daß ebenso ein irreduzibles Polynom keine mehrfache Nullstelle haben kann. Hierbei geht ein, daß $\mathbb{Q}$ Charakteristik Null hat. Wer mehr dazu wissen möchte, sei auf Bücher zur Algebra und den Begriff der *Separabilität* von Polynomen verwiesen.

b. Für praktische Anwendungen und $K = \mathbb{R}$ ist es oft wichtiger, die komplexen Eigenwerte von A numerisch approximativ zu bestimmen. Dazu existieren effiziente iterative Eigenwertverfahren oder man bestimmt approximativ die Nullstellen von μ_A in $\mathbb{C}$. Der oben dargestellte Algorithmus liefert dann die exakten Vielfachheiten der Eigenwerte, d. h. die Dimension der Eigenräume. Dies ist ein effizientes und exaktes Verfahren, das in den Bereich der *Computeralgebra* fällt.

c. Man beachte jedoch, daß der Algorithmus 3.36 nicht die Transformationsmatrix liefert, die A in Diagonalgestalt überführt.

Beispiel 3.38

Wir betrachten im Folgenden die Matrix

$$A = \begin{pmatrix} -3 & -4 & 0 & 0 \\ -1 & 4 & 0 & 0 \\ 0 & 0 & -3 & -4 \\ 0 & 0 & -1 & 4 \end{pmatrix} \in \mathbf{Mat}(4, \mathbb{Q}).$$

Es ist unser Ziel, zu zeigen, daß A über $\mathbb{C}$ diagonalisierbar ist, und eine zu A konjugierte Diagonalmatrix näherungsweise zu bestimmen.

```
LIB "linalg.lib";
ring r=0,t,lp;
matrix A[4][4]=-3,-4,0,0,-1,4,0,0,0,0,-3,-4,0,0,-1,4;
poly f=min_poly(A);  // Die Prozedur muß mittels 3.26 selbst
                     // implementiert werden.
f;
↦   t2-t-16;
poly Df=diff(f,t); // Bestimmt die formale Ableitung von f.
gcd(f,Df);
↦   1     // Also ist A diagonalisierbar.
poly g=charpoly(A,"t"); // Das charakteristische Polynom von A.
factorize(g,2);
↦    [1]:                        // D. h. f ist irreduzibel und g = f^2.
          _[1] = t2 - t - 16
      [2]:
          2
LIB "solve.lib";
```

```
laguerre_solve(f);
```

$\mapsto$ [1] :

$-3.531128874149274826183306615152$

[2] :

$4.531128874149274826183306615152$

Also ist die gesuchte Diagonalmatrix bei Rundung auf zwei Nachkommastellen

$$D = \begin{pmatrix} -3.53 & 0 & 0 & 0 \\ 0 & -3.53 & 0 & 0 \\ 0 & 0 & 4.53 & 0 \\ 0 & 0 & 0 & 4.53 \end{pmatrix}.$$

AUFGABEN

Aufgabe 3.39

Bestimme die Eigenwerte und die Eigenräume der folgenden Matrix und entscheide, ob sie diagonalisierbar bzw. trigonalisierbar ist:

$$\begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 3 & 0 & 0 \\ -1 & 1 & 2 & 1 \\ -1 & 1 & 0 & 3 \end{pmatrix}.$$

Aufgabe 3.40

Es sei $V = \mathrm{Mat}(2, K)$, $E = (E_{11}, E_{12}, E_{21}, E_{22})$ die kanonische Basis von V und $T = E_{11} + E_{12} + E_{22} \in \mathrm{Gl}_2(K)$. Betrachte die Abbildung

$$f : V \to V : A \mapsto T \circ A \circ T^{-1}.$$

Zeige:

a. $f \in \mathrm{End}_K(V)$.
b. f ist trigonalisierbar, aber nicht diagonalisierbar.
c. Bestimme eine Basis B von V, bezüglich derer $M_B^B(f)$ eine obere Dreiecksmatrix ist.

Aufgabe 3.41

Sei V ein $\mathbb{R}$-Vektorraum mit Basis $E = (x_1, x_2, x_3)$ und $f \in \mathrm{End}_{\mathbb{R}}(V)$ mit

$$\begin{aligned} f(x_1) &= x_1 + 3x_2 + 6x_3, \\ f(x_2) &= -3x_1 - 5x_2 - 6x_3, \\ f(x_3) &= 3x_1 + 3x_2 + 4x_3. \end{aligned}$$

Bestimme das Minimalpolynom von f, zeige, daß f diagonalisierbar ist, und bestimme eine Basis B, bezüglich derer f Diagonalgestalt hat.

Aufgabe 3.42

Es sei $V = U_1 \oplus U_2$ ein endlich-dimensionaler K-Vektorraum und $f \in \mathrm{End}_K(V)$ mit $f(U_i) \subseteq U_i$. Zeige:

a. $\mathrm{Ker}(\phi_f) = \mathrm{Ker}\left(\phi_{f_{U_1}}\right) \cap \mathrm{Ker}\left(\phi_{f_{U_2}}\right)$.
b. $\mu_f = \mathrm{kgV}\left(\mu_{f_{U_1}}, \mu_{f_{U_2}}\right)$.

Aufgabe 3.43 a. Zeige, eine Matrix $A \in \mathrm{Mat}(n, K)$ mit $\chi_A = (\lambda - t)^n$ ist genau dann diagonalisierbar, wenn $A = \lambda \mathbb{1}_n$.

b. Zeige, eine Diagonalblockmatrix der Form $A = \bigoplus_{i=1}^{r} A_i \in \mathrm{Mat}(n, K)$ mit $A_i \in \mathrm{Mat}(n_i, K)$ und $n = n_1 + \ldots + n_r$ ist genau dann diagonalisierbar, wenn A_i diagonalisierbar ist für alle $i = 1, \ldots, r$.

c. Es sei A eine Diagonalblockmatrix wie in b. mit $A_i = J(\lambda_i, n_i)$ ein Jordanblock für $i = 1, \ldots, r$. Zeige, A ist genau dann diagonalisierbar, wenn $n_i = 1$ für alle $i = 1, \ldots, r$.

Aufgabe 3.44 a. Es sei $V \neq 0$ ein endlich-dimensionaler K-Vektorraum, $f \in \mathrm{End}_K(V)$ und $\{0\} = U_0 \subseteq U_1 \subsetneq \ldots \subsetneq U_{m-1} \subsetneq U_m = V$ eine f-invariante Fahne von V. Ferner seien die Minimalpolynome $\mu_{f_{U_i/U_{i-1}}}$ für $i = 1, \ldots, m$ paarweise teilerfremd. Zeige, dann ist

$$\mu_f = \mu_{f_{U_m/U_{m-1}}} \cdots \mu_{f_{U_1/U_0}}.$$

b. Es seien $m_1, \ldots, m_r \in \mathbb{N}$ und $\lambda_1, \ldots, \lambda_r \in K$ seien paarweise verschieden. Zeige, die folgende Matrix ist diagonalisierbar:

$$A = \left(\begin{array}{c|ccc} \lambda_1 \mathbb{1}_{m_1} & * & \cdots & * \\ \hline 0 & \boxed{\lambda_2 \mathbb{1}_{m_2}} & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & \boxed{\lambda_r \mathbb{1}_{m_r}} \end{array}\right).$$

Aufgabe 3.45
Es sei $A = (a_{ij}) \in \mathrm{Mat}(4, K)$ eine obere Dreiecksmatrix mit $a_{11} = a_{22} \neq a_{33} = a_{44}$. Zeige, A ist genau dann diagonalisierbar, wenn $a_{12} = a_{34} = 0$.

Aufgabe 3.46
Ist $A \in \mathrm{Gl}_n(K)$, so gibt es ein Polynom $g \in K[t]$ mit $A^{-1} = g(A)$.

Aufgabe 3.47
Es sei $V \neq 0$ ein K-Vektorraum.

a. Sind $\lambda \in K$ und $f, g \in \mathrm{End}_K(V)$ mit $f \circ g = g \circ f$, dann ist $\mathrm{Eig}(f, \lambda)$ g-invariant.

b. Zeige, ist $\dim_K(V) = n < \infty$, so sind für $\mathcal{A} \subseteq \mathrm{End}_K(V)$ die folgenden beiden Aussagen gleichwertig:

(i) $\mathcal{A}$ ist simultan diagonalisierbar, d. h. es gibt eine Basis B von V, so daß für alle $f \in \mathcal{A}$ gilt $M_B^B(f)$ ist eine Diagonalmatrix.

(ii) Für alle $f \in \mathcal{A}$ gilt, f ist diagonalisierbar, und für alle $f, g \in \mathcal{A}$ gilt, $f \circ g = g \circ f$.

Hinweis: In b. führe man für die Rückrichtung Induktion über n und betrachte ggf. eine Zerlegung von V in zwei geeignete invariante Unterräume kleinerer Dimension.

Aufgabe 3.48
Es sei $V \neq 0$ ein endlich-dimensionaler K-Vektorraum und $f \in \mathrm{End}_K(V)$. Zeige, genau dann ist χ_f irreduzibel, wenn V nur die trivialen f-invarianten Unterräume $\{0\}$ und V besitzt.

Aufgabe 3.49
Es sei V ein endlich-dimensionaler K-Vektorraum, $f \in \mathrm{End}_K(V)$. Zeige:

a. Es gibt ein $m \in \mathbb{N}$ mit $\mathrm{Ker}\left(f^0\right) \subsetneqq \ldots \subsetneqq \mathrm{Ker}\left(f^{m-1}\right) \subsetneqq \mathrm{Ker}\left(f^m\right) = \mathrm{Ker}\left(f^k\right)$ für alle $k > m$.
Die Zahl m heißt auch der *Nilpotenzindex* der Abbildung f.
b. Für das m aus a. gilt ferner,
(i) $V = \mathrm{Im}\left(f^0\right) \supsetneqq \mathrm{Im}\left(f^1\right) \supsetneqq \ldots \supsetneqq \mathrm{Im}\left(f^m\right)$,
(ii) $\mathrm{Im}\left(f^m\right) = \mathrm{Im}\left(f^k\right)$ für alle $k \geq m$, und
(iii) $V = \mathrm{Ker}\left(f^m\right) \oplus \mathrm{Im}\left(f^m\right)$. ("Fitting-Zerlegung von V")
(iv) $\left(f_{\mathrm{Ker}(f^m)}\right)^m = 0$ und $f_{\mathrm{Im}(f^m)}$ ist invertierbar.
c. Ist $\dim_K(V) = \infty$, so gelten die Aussagen in a. und b. im Allgemeinen nicht mehr.

Aufgabe 3.50
Schreibe eine Singular-Prozedur `min_poly`, die eine quadratische Matrix $A \in \mathrm{Mat}(n, K)$ einliest und das Minimalpolynom von A zurückgibt.
Aufgabe 3.51
Schreibe eine Singular-Prozedur `diagonal`, die eine Matrix $A \in \mathrm{Mat}(n, \mathbb{Q})$ einliest und feststellt, ob A diagonalisierbar ist. Falls ja, soll der Wert 1 ausgegeben werden, sonst 0.
Aufgabe 3.52
Schreibe Singular-Prozeduren, die die Algorithmen 3.12 und 3.36 umsetzen.

4. Die Jordansche Normalform

Generalvoraussetzung: Es sei K ein Körper, V ein K-Vektorraum mit $1 \leq \dim_K(V) = n < \infty$ und $f \in \mathrm{End}_K(V)$. A bezeichne stets eine Matrix in $\mathrm{Mat}(n, K)$. Ferner gelte in diesem Abschnitt, daß das charakteristische Polynom von f bzw. A über K in Linearfaktoren zerfällt.

Eine Matrix $A \in \mathrm{Mat}(n, K)$, deren charakteristisches Polynom in Linearfaktoren zerfällt, was etwa für einen algebraisch abgeschlossenen Körper wie $\mathbb{C}$ stets der Fall ist, ist zu einer Matrix konjugiert, die besonders einfach gebaut ist, der sog. Jordanschen Normalform von A. Aus der Jordanschen Normalform lassen sich Invarianten von A einfach ablesen und diese Invarianten bestimmen die Matrix A bis auf Konjugation eindeutig.

Der folgende Satz über die Jordansche Normalform hat große theoretische und auch praktische Bedeutung.
Satz 4.1 (Jordansche Normalform) a. *Ist* $\chi_A = (-1)^n \prod_{i=1}^r (t-\lambda_i)^{n_i}$ *und* $\mu_A = \prod_{i=1}^r (t-\lambda_i)^{m_i}$*, dann gibt es für jedes* $1 \leq j \leq m_i$, $i = 1, \ldots, r$*, je eine natürliche Zahl* t_{ij} *so, daß* $\sum_{j=1}^{m_i} j \cdot t_{ij} = n_i = \dim_K\left(\mathrm{Hau}(A, \lambda_i)\right)$,

$\sum_{j=1}^{m_i} t_{ij} = \dim_K\big(\mathrm{Eig}(A,\lambda_i)\big)$, $t_{im_i} \geq 1$ *und* A *ist konjugiert zu der Diagonalblockmatrix*

$$J_A := \bigoplus_{i=1}^{r} \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} J_j(\lambda_i).$$

J_A *heißt* Jordansche Normalform *von* A*, und die* t_{ij} *werden* Elementarteiler *von* A *zum Eigenwert* λ_i *genannt.*

b. *Ist* $\chi_f = (-1)^n \prod_{i=1}^{r}(t-\lambda_i)^{n_i}$ *und* $\mu_f = \prod_{i=1}^{r}(t-\lambda_i)^{m_i}$*, dann gibt es für jedes* $1 \leq j \leq m_i$*,* $i = 1,\ldots,r$*, je eine natürliche Zahl* t_{ij} *und es gibt eine Basis* B *so, daß* $\sum_{j=1}^{m_i} j \cdot t_{ij} = n_i = \dim_K\big(\mathrm{Hau}(f,\lambda_i)\big)$, $\sum_{j=1}^{m_i} t_{ij} = \dim_K\big(\mathrm{Eig}(f,\lambda_i)\big)$, $t_{im_i} \geq 1$ *und*

$$J_f := M_B^B(f) = \bigoplus_{i=1}^{r} \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} J_j(\lambda_i).$$

J_f *heißt* Jordansche Normalform *von* f*, und die* t_{ij} *werden* Elementarteiler *von* f *zum Eigenwert* λ_i *genannt.*

Es scheint angebracht, den Satz zunächst etwas zu erläutern, um ihn verständlicher zu machen.

Bemerkung 4.2 a. Ziel des Abschnittes ist es, zu zeigen, daß eine Matrix A, deren charakteristisches Polynom zerfällt, konjugiert zu einer Matrix von besonders einfacher Gestalt ist. Der obige Satz sagt nun, daß in der Tat A konjugiert ist zu einer Diagonalblockmatrix, deren Diagonalblöcke, die $J_j(\lambda_i)$, alle Jordanblöcke sind, also obere Dreiecksmatrizen, die auf der Diagonalen stets den gleichen Wert λ_i stehen haben, auf der oberen Nebendiagonalen nur Einsen und ansonsten nur Nullen (vgl. Beispiel 3.24).

Dabei gelten:

- Die natürlichen Zahlen t_{ij} geben an, wieviele Jordanblöcke der Größe $j \times j$ zum Eigenwert λ_i denn vorkommen.
- $j \leq m_i$ bedeutet, daß die maximale Größe eines Jordanblockes $m_i \times m_i$ ist.
- $t_{im_i} \geq 1$ besagt, daß auch mindestens ein Block der maximalen Größe $m_i \times m_i$ vorkommt. D. h. die Vielfachheit von λ_i als Nullstelle von μ_A gibt die maximale Größe eines vorkommenden Jordanblockes in J_A zum Eigenwert λ_i an.
- Die Summe $\sum_{j=1}^{m_i} j \cdot t_{ij}$ gibt gerade an, wie oft der Eigenwert λ_i auf der Diagonalen der Diagonalblockmatrix vorkommt, und da diese das gleiche charakteristische Polynom wie A besitzt, muß die Summe mithin n_i, also die algebraische Vielfachheit von λ_i als Eigenwert von A, sein.
- Und $\sum_{j=1}^{m_i} t_{ij} = \dim_K\big(\mathrm{Eig}(A,\lambda_i)\big)$ bedeutet schließlich, daß die Anzahl der Jordanblöcke zum Eigenwert λ_i, die in J_A vorkommen, der Dimension des Eigenraumes von A zum Eigenwert λ_i entspricht.

b. Schon die direkte Summenschreibweise der Jordanschen Normalform bringt zum Ausdruck, daß die Jordansche Normalform nur bis auf die Reihenfolge der Jordanblöcke eindeutig bestimmt sein kann, und in der Tat ist sie es auch, d. h.:

Zwei Jordansche Normalformen sind genau dann konjugiert, wenn die Eigenwerte sowie zu den Eigenwerten die Elementarteiler übereinstimmen.

Es ist leicht einsichtig, daß eine Vertauschung der Blöcke durch Konjugation mit einer Reihe von Permutationsmatrizen erreicht werden kann, daß mithin zwei Jordansche Normalformen, deren Eigenwerte mit zugehörigen Elementarteilern übereinstimmen, zueinander konjugiert sind.
Seien umgekehrt zwei Jordansche Normalformen zueinander konjugiert, dann stimmen zunächst die charakteristischen Polynome und damit die Eigenwerte überein. Ferner folgt aus Aufgabe 4.15, daß die Elementarteiler übereinstimmen, da für eine invertierbare Matrix $T \in \mathrm{Gl}_n(K)$ und ein $k \in \mathbb{N}$ gilt

$$\begin{aligned}\operatorname{rang}\left(\left(T^{-1} \circ A \circ T - \lambda \mathbb{1}_n\right)^k\right) &= \operatorname{rang}\left(T^{-1} \circ (A - \lambda \mathbb{1}_n)^k \circ T\right) \\ &= \operatorname{rang}\left((A - \lambda \mathbb{1}_n)^k\right).\end{aligned}$$

Damit ist natürlich auch die Jordansche Normalform eines Endomorphismus bis auf die Reihenfolge der Jordanblöcke eindeutig bestimmt.

c. Wir wollen folgende Notation einführen, die die Jordanblöcke von A (bzw. f) zu einem Eigenwert λ_i zusammenfaßt:

$$J_A(\lambda_i) := \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} J_j(\lambda_i) \quad \text{bzw.} \quad J_f(\lambda_i) := \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} J_j(\lambda_i)$$

Dann gilt

$$J_A = \bigoplus_{i=1}^{r} J_A(\lambda_i) \quad \text{bzw.} \quad J_f = \bigoplus_{i=1}^{r} J_f(\lambda_i).$$

d. Es sei $J = (a_{ij})_{i,j}$ eine Matrix in Jordanscher Normalform. $S = (s_{ij})_{i,j}$ bezeichne die Diagonalmatrix, die entsteht, wenn man in J alle Nicht-Diagonalelemente zu Null setzt, d. h. $s_{ii} = a_{ii}$ und $s_{ij} = 0$ für $i \neq j$. Ferner setzen wir $N = J - S$, d. h. N ist eine Matrix, die nur auf der oberen Nebendiagonalen Elemente ungleich Null besitzen kann.
Dann ist N nilpotent und es gelten

$$J = S + N \quad \text{mit} \quad N \circ S = S \circ N.$$

Man nennt dies auch die *Jordan-Zerlegung*[8] von J.

[8] Allgemeiner nennt man die Darstellung einer Matrix $A \in \mathrm{Mat}(n, K)$ als $A = S + N$ mit N nilpotent und S diagonalisierbar (auch *halbeinfach* genannt, engl. semi-simple, daher das S) und $S \circ N = N \circ S$ eine Jordan-Zerlegung von A. Solche Zerlegungen von Objekten

Um die Aussage einzusehen, beachte man, daß für $i = 1, \ldots, r$ und $1 \leq j \leq m_i$ gilt

$$J_j(\lambda_i) = \lambda_i \mathbb{1}_j + J_j(0).$$

Damit gilt

$$S = \bigoplus_{i=1}^{r} \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} \lambda_i \mathbb{1}_j$$

und

$$N = \bigoplus_{i=1}^{r} \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} J_j(0).$$

Aber damit folgt unmittelbar

$$N \circ S = \bigoplus_{i=1}^{r} \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} \lambda_i J_j(0) = S \circ N.$$

Es reicht natürlich, den obigen Satz für einen Endomorphismus f zu zeigen, da die Aussagen dann mittels f_A auch für A gezeigt sind.

Wir werden den Satz zunächst für nilpotente Endomorphismen zeigen, d. h. für Endomorphismen, die nur einen Eigenwert, nämlich $\lambda = 0$, besitzen, und den allgemeinen Fall dann auf diesen zurückführen.

Lemma 4.3

Es sei $f \in \mathrm{End}_K(V)$ *ein nilpotenter Endomorphismus mit* $\mu_f = t^m$. *Dann gibt es für jedes* $1 \leq j \leq m$ *je ein natürliche Zahl* t_j *und es gibt eine Basis* B *so, daß* $\sum_{j=1}^{m} j \cdot t_j = n = \dim_K\big(\mathrm{Hau}(f,0)\big) = \dim_K(V)$, $\sum_{j=1}^{m} t_j = \dim_K\big(\mathrm{Eig}(f,0)\big)$, $t_m \geq 1$ *und*

$$J_f := M_B^B(f) = \bigoplus_{j=1}^{m} \bigoplus_{k=1}^{t_j} J_j(0).$$

Beweis: Wir beachten zunächst, daß nach Proposition 3.33 bereits gilt

$$V = \mathrm{Hau}(f,0) = \mathrm{Ker}\left(f^m\right)$$

und mithin insbesondere $f^m = 0$.

Wir führen den Beweis konstruktiv, so daß sich daraus auch ein Algorithmus zur Bestimmung der Basis B herleiten läßt. Dabei gehen wir schrittweise vor und konstruieren in jedem Schritt alle Jordanblöcke gleicher Größe, beginnend mit dem größten.

Für $j = 0, \ldots, m+1$ definieren wir $V^j := \mathrm{Ker}(f^j)$. Da m der Nilpotenzindex ist, gilt mithin

$$0 = V^0 \subsetneqq V^1 \subsetneqq \ldots \subsetneqq V^m = V.$$

Außerdem gilt $f\big(V^{j+1}\big) \subseteq V^j$ für $j = 0, \ldots, m$, so daß folgende Definition Sinn macht:

$$\overline{V}^j := V^j / \big(V^{j-1} + f\big(V^{j+1}\big)\big).$$

---

in einen halbeinfachen und einen nilpotenten Anteil spielen auch in anderen Bereichen der Mathematik eine Rolle - siehe etwa Lie-Algebren oder Jordan-Algebren.

Schließlich setzen wir für $j = 1, \dots, m$

$$t_j := \dim_K \left(\overline{V}^j \right).$$

Unter Berücksichtigung des Austauschsatzes von Steinitz wählen wir rekursiv für $j = m, \dots, 1$ Vektoren $e_{j1}, \dots, e_{jt_j} \in V^j$, deren Restklassen eine Basis von $\overline{V}^j$ bilden.

Sodann bilden wir für $j = 1, \dots, m$, $k = 1, \dots, t_j$ und $r \in \mathbb{N}$ die Familien von Vektoren[9]

$$B_{jk}^r = \left(f^{j-1}(e_{jk}), f^{j-2}(e_{jk}), \dots, f^{r+1}(e_{jk}), f^r(e_{jk}) \right).$$

Wir behaupten

a. $B = \bigcup_{j=1}^m \bigcup_{k=1}^{t_j} B_{jk}^0$ ist linear unabhängig.
b. $V^{j-1} + f\left(V^{j+1}\right) = V^{j-1} + \left\langle f^l(e_{j+l,k}) \mid l = 1, \dots, m-j; k = 1, \dots, t_{j+l} \right\rangle$ für $j = 1, \dots, m$.
c. $V^j = V^{j-1} + \left\langle f^l(e_{j+l,k}) \mid l = 0, \dots, m-j; k = 1, \dots, t_{j+l} \right\rangle$ für $j = 1, \dots, m+1$.
d. Für $j = 0, \dots, m$ ist $B_j = \left(\bigcup_{l=1}^j \bigcup_{k=1}^{t_l} B_{lk}^0 \right) \cup \left(\bigcup_{l=j+1}^m \bigcup_{k=1}^{t_l} B_{lk}^{l-j} \right)$ eine Basis von V^j.[10]
e. $B = B_m$ ist eine Basis von V und $n = \sum_{j=1}^m j \cdot t_j$.
f. B_1 ist eine Basis von $\mathrm{Eig}(f, 0) = \mathrm{Ker}(f)$ und $\dim_K \left(\mathrm{Eig}(f, 0) \right) = \sum_{j=1}^m t_j$.
g. Für $j = 1, \dots, m$ und $k = 1, \dots, t_j$ ist $W_{jk} = \left\langle B_{jk}^0 \right\rangle$ ein f-invarianter, zyklischer Unterraum mit
$$M_{B_{jk}^0}^{B_{jk}^0}\left(f_{W_{jk}}\right) = J_j(0).$$
h. $M_B^B(f) = \bigoplus_{j=1}^m \bigoplus_{k=1}^{t_j} J_j(0)$.
i. $t_m \geq 1$.

Zu a.: Angenommen, B wäre linear abhängig. Dann gibt es eine nichttriviale Linearkombination

$$\sum_{j=1}^{m} \sum_{k=1}^{t_j} \sum_{l=0}^{j-1} \lambda_{jkl} f^l(e_{jk}) = 0. \tag{49}$$

Wir setzen zunächst $\mathcal{N} := \{(j, l) \mid \exists k : \lambda_{jkl} \neq 0\} \neq \emptyset$. Sodann wählen wir

$$\tau := \max\{j - l \mid (j, l) \in \mathcal{N}\} > 0$$

und setzen $\mathcal{N}_\tau := \{(j, l) \in \mathcal{N} \mid j - l = \tau\} \neq \emptyset$. Ferner definieren wir

$$j_0 := \max\{j \mid \exists\, l : (j, l) \in \mathcal{N}_\tau\}.$$

Schließlich wählen wir noch

$$j' := \min\{j \mid \exists\, l : (j, l) \in \mathcal{N}_\tau\} = \min\{j \mid \tau \leq j \leq j_0, \exists\, k : \lambda_{j,k,j-\tau} \neq 0\}.$$

[9] Beachte, für $r > j - 1$ ist die Familie $B_{jk}^r = \emptyset$ leer. Außerdem beachte man, daß stets $B_{jk}^{r+1} \subseteq B_{jk}^r$ gilt.

[10] Wobei wir für $j = 0$ die Familie $B_0 = \emptyset$ als Ergebnis der leeren Vereinigung erhalten, da es keine B_{0k} gibt.

Sei nun $1 \leq j \leq m$ gegeben, dann folgt aus der Wahl von τ

$$\lambda_{jkl} = 0 \quad \forall\, l < j - \tau \;\; \forall\, k = 1, \ldots, t_j.$$

Ist $j > j_0$ dann gilt zudem nach Wahl von j_0

$$\lambda_{jkl} = 0 \quad \forall\, l = j - \tau \;\; \forall\, k = 1, \ldots, t_j.$$

Damit reduziert sich (49) auf

$$0 = \sum_{j=1}^{j_0} \sum_{k=1}^{t_j} \sum_{l=\max\{0,j-\tau\}}^{j-1} \lambda_{jkl} f^l(e_{jk}) + \sum_{j=j_0+1}^{m} \sum_{k=1}^{t_j} \sum_{l=j-\tau+1}^{j-1} \lambda_{jkl} f^l(e_{jk}). \tag{50}$$

Beachten wir nun, daß für $\nu \in \mathbb{N}$ gilt $f^\nu(e_{jk}) = 0$ für alle $\nu \geq j$, dann verschwindet der zweite Summand in (50) nach Anwenden von $f^{\tau-1}$ und wir erhalten

$$0 = f^{\tau-1}(0) = \sum_{j=1}^{j_0} \sum_{k=1}^{t_j} \sum_{l=\max\{0,j-\tau\}}^{j-1} \lambda_{jkl} f^{l+\tau-1}(e_{jk}). \tag{51}$$

Da aber mit dem gleichen Argument $f^{l+\tau-1}(e_{jk}) = 0$ gilt für alle $l > j - \tau \geq 0$ (sowie für alle $l \geq 0$, falls $j < \tau$), reduziert sich (51) auf

$$0 = \sum_{j=\tau}^{j_0} \sum_{k=1}^{t_j} \lambda_{j,k,j-\tau} f^{j-1}(e_{jk}).$$

Nach Wahl von j' wiederum ist dies gerade

$$0 = \sum_{j=j'}^{j_0} \sum_{k=1}^{t_j} \lambda_{j,k,j-\tau} f^{j-1}(e_{jk}) = f^{j'-1}\left(\sum_{j=j'}^{j_0} \sum_{k=1}^{t_j} \lambda_{j,k,j-\tau} f^{j-j'}(e_{jk}) \right),$$

also

$$w := \sum_{j=j'}^{j_0} \sum_{k=1}^{t_j} \lambda_{j,k,j-\tau} f^{j-j'}(e_{jk}) \in \operatorname{Ker}\left(f^{j'-1}\right) = V^{j'-1}.$$

Aber damit folgt

$$\sum_{k=1}^{t_{j'}} \lambda_{j',k,j'-\tau} e_{j'k} = w - f\left(\sum_{j=j'+1}^{j_0} \sum_{k=1}^{t_j} \lambda_{j,k,j-\tau} f^{j-j'-1}(e_{jk}) \right) \in V^{j'-1} + f\big(V^{j'+1}\big).$$

Da $(e_{j'k} \mid k = 1, \ldots, t_{j'})$ aber linear unabhängig modulo $V^{j'-1} + f\big(V^{j'+1}\big)$ ist, erhalten wir

$$\sum_{k=1}^{t_{j'}} \lambda_{j',k,j'-\tau} e_{j'k} = 0,$$

und sogar $\lambda_{j',k,j'-\tau} = 0$ für alle k, im Widerspruch zur Wahl von j'.

Zu b.: Wir führen den Beweis durch absteigende Induktion nach j.

Sei also zunächst $j = m$. Man beachte dabei, daß die Familie der $e_{m+1,k}$ leer ist! Ferner gilt $V^{m+1} = V = V^m$ und $f\big(V^m\big) = f\big(\operatorname{Ker}\left(f^m\right)\big) \subseteq \operatorname{Ker}\left(f^{m-1}\right) = V^{m-1}$. Damit folgt insbesondere

$$V^{m-1} + f\big(V^{m+1}\big) = V^{m-1},$$

womit der Induktionsanfang gezeigt ist.

Sei nun $1 \leq j < m$. Es reicht zu zeigen, daß $f\big(V^{j+1}\big) \subseteq V^{j-1} + \langle f^l(e_{j+l,k}) \mid l = 1, \ldots, m-j; k = 1, \ldots, t_{j+l}\rangle$. Nach Definition der $e_{j+1,k}$ und Induktionsvoraussetzung gilt

$$\begin{aligned} V^{j+1} &= \langle e_{j+1,k} \mid k = 1, \ldots, t_{j+1}\rangle + V^j + f\big(V^{j+2}\big) \\ &= \langle e_{j+1,k} \mid k = 1, \ldots, t_{j+1}\rangle + V^j \\ &\quad + \big\langle f^l(e_{j+1+l,k}) \mid l = 1, \ldots, m-(j+1); k = 1, \ldots, t_{j+1+l}\big\rangle \\ &= V^j + \big\langle f^l(e_{j+1+l,k}) \mid l = 0, \ldots, m-(j+1); k = 1, \ldots, t_{j+1+l}\big\rangle. \end{aligned} \tag{52}$$

Dabei gilt $f\big(V^j\big) \subseteq V^{j-1}$. Dann folgt aber

$$f\big(V^{j+1}\big) \subseteq V^{j-1} + \big\langle f^l(e_{j+l,k}) \mid l = 1, \ldots, m-j; k = 1, \ldots, t_{j+1}\big\rangle.$$

Zu c.: Wir führen den Beweis durch absteigende Induktion nach j. Wegen $V^{m+1} = V = V^m$ ist für $j = m+1$ nichts zu zeigen. Sei also $2 \leq j \leq m$, dann folgt die Behauptung aus Gleichung (52), und der Fall $j = 1$ folgt analog.

Zu d.: Nach a. ist $B_j \subseteq B$ linear unabhängig für $j = 0, \ldots, m$. Es reicht also, zu zeigen, daß $V^j = \langle B_j\rangle$. Dies tun wir durch Induktion nach j. Dabei ist für den Induktionsanfang mit $j = 0$ nichts zu zeigen, da $V_0 = \{0\} = \langle\emptyset\rangle$ gilt.

Sei also $j \geq 1$. Dann gilt nach Teil c.

$$V^j = V^{j-1} + \big\langle f^l(e_{j+l,k}) \;\big|\; l = 0, \ldots, m-j; k = 1, \ldots, t_{j+l}\big\rangle$$

Da $f^l(e_{j+l,k}) \in B^l_{j+l,k}$ für $l = 0, \ldots, m-j$ und $k = 1, \ldots, t_{j+l}$ gilt und wegen $B^1_{j,k} \subset B^0_{j,k}$ für $k = 1, \ldots, t_j$, folgt mit Induktion

$$\begin{aligned} V^j &\subseteq \Big\langle \Big(\bigcup_{l=1}^{j-1}\bigcup_{k=1}^{t_l} B^0_{lk}\Big) \cup \Big(\bigcup_{l=j}^{m}\bigcup_{k=1}^{t_l} B^{l-j+1}_{lk}\Big)\Big\rangle + \Big\langle \bigcup_{l=0}^{m-j}\bigcup_{k=1}^{t_{j+l}} B^l_{j+l,k}\Big\rangle \\ &\subseteq \Big\langle \Big(\bigcup_{l=1}^{j}\bigcup_{k=1}^{t_l} B^0_{lk}\Big) \cup \Big(\bigcup_{l=j+1}^{m}\bigcup_{k=1}^{t_l} B^{l-j}_{lk}\Big)\Big\rangle. \end{aligned}$$

Die umgekehrte Inklusion "$\supseteq$" ist hingegen klar.

Zu e.: Offenbar ist $B = B_m$ und die Aussage folgt somit aus d..

Zu f.: Da $V^1 = \mathrm{Ker}(f) = \mathrm{Eig}(f, 0)$ ist, folgt aus d., daß B_1 eine Basis von $\mathrm{Eig}(f, 0)$ ist. Aus der Definition von B_1 folgt unmittelbar, daß $\dim_K\big(\mathrm{Eig}(f, 0)\big) = |B_1| = \sum_{j=1}^m t_j$, da $\big|B^{j-1}_{jk}\big| = 1$ für $j = 1, \ldots, m$ und $k = 1, \ldots, t_j$.

Zu g.: Dies folgt unmittelbar aus der Definition von B^0_{jk}.

Zu h.: Dies folgt unmittelbar aus g. und dem Umstand, daß $B = B_m$, d. h. $V = \bigoplus_{j=1}^m \bigoplus_{k=1}^{t_j} W_{jk}$.

Zu i.: Wäre $t_m = 0$, so wäre $J_f^{m-1} = 0$ und mithin $f^{m-1} = 0$, im Widerspruch zu $\mu_f = t^m$.

Damit ist das Lemma bewiesen. □

Definition 4.4 a. Für eine natürliche Zahl $n \geq 1$ heißt ein Tupel $(\nu_1, \ldots, \nu_k) \in \big(\mathbb{N} \setminus \{0\}\big)^k$ eine *Partition* von n, falls gilt

$$n = \nu_1 + \ldots + \nu_k \quad \text{und} \quad \nu_1 \geq \nu_2 \geq \ldots \geq \nu_k \geq 1.$$

b. $\pi : \mathbb{N} \setminus \{0\} \to \mathbb{N} \setminus \{0\}$ bezeichne die Funktion, die einer natürlichen Zahl n die Anzahl der Partitionen von n zuordnet.

Korollar 4.5 a. *Ist* $f \in \mathrm{End}_K(V)$ *ein nilpotenter Endomorphismus, dann besitzt* V *eine direkte Zerlegung*

$$V = W_1 \oplus \ldots \oplus W_k$$

in f*-invariante zyklische Unterräume* W_i *der Dimension* $\dim_K(W_i) = \nu_i$. *Dabei sind die* ν_i *durch* f *eindeutig bestimmt und bilden eine* Partition[11] *von* n.

b. *Ferner gilt, zwei nilpotente Endomorphismen* $f, g \in \mathrm{End}_K(V)$ *sind genau dann* konjugiert,[12] *wenn die zugeordneten Partitionen gleich sind.*

Beweis: Teil a. folgt aus dem Beweis von Lemma 4.3, wenn man dort die W_{jk} als zyklische Unterräume nimmt und nach ihrer Dimension ordnet. Die Eindeutigkeit der Partition sowie die Aussage in b. folgen aus Bemerkung 4.2 b. und damit letztlich aus Aufgabe 4.15, wenn man beachtet, daß sich aus den Elementarteilern die Partition und umgekehrt aus der Partition die Elementarteiler in offensichtlicher Weise bestimmen lassen. □

Korollar 4.6

Jede nilpotente Matrix $A \in \mathrm{Mat}(n, K)$ *ist konjugiert zu einer eindeutig bestimmten Matrix*

$$J_A = J_{\nu_1}(0) \oplus \ldots \oplus J_{\nu_k}(0) = \begin{pmatrix} J_{\nu_1}(0) & 0 & \cdots & 0 \\ 0 & \boxed{J_{\nu_2}(0)} & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & \boxed{J_{\nu_k}(0)} \end{pmatrix},$$

wobei $(\nu_1, \ldots, \nu_k)$ *eine Partition von* n *ist. Dabei gilt* $\chi_A = (-1)^n t^n$ *und* $\mu_A = t^{\nu_1}$.

Bemerkung 4.7

Ist A eine nilpotente Matrix mit $\mu_A = t^m$ und bestimmt man analog zu dem Verfahren im Beweis von Lemma 4.3 linear unabhängige Familien

$$B^0_{jk} = \left(f^{j-1}(e_{jk}), f^{j-2}(e_{jk}), \ldots, f^1(e_{jk}), e_{jk}\right) \subset K^n$$

in $\mathrm{Lös}(A^j, 0)$ für $j = 1, \ldots, m$ und $k = 1, \ldots, t_j$, dann ist die Matrix $T \in \mathrm{Gl}_n(K)$, deren Spalten gerade die Vektoren der Basis B von K^n sind, eine Transformationsmatrix, die A in Jordansche Normalform überführt.

Korollar 4.8

Es gibt für jedes $n \geq 1$ *in* $\mathrm{Mat}(n, K)$ *genau* $\pi(n)$ *Konjugationsklassen von nilpotenten Matrizen.*

---

[11] Die ν_i sind die Dimensionen der zyklischen Unterräume und damit die Größen der vorkommenden Jordankästchen. Aus ihnen lassen sich also die Elementarteiler bestimmen und umgekehrt!

[12] D. h. es gibt einen Automorphismus $\psi \in \mathrm{Gl}_K(V)$ mit $f = \psi^{-1} \circ g \circ \psi$, oder äquivalent, es gibt zwei Basen B und D von V mit $M^B_B(f) = M^D_D(g)$.

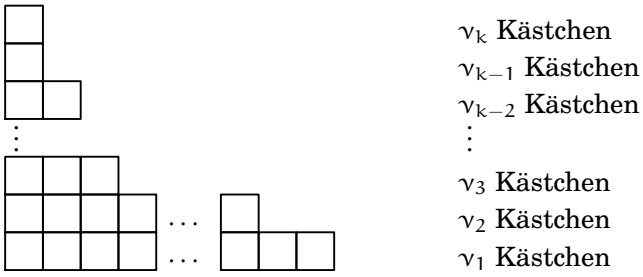
Bemerkung 4.9
Es ist offensichtlich, daß π in der Tat nur Werte in $\mathbb{N}\backslash\{0\}$ annimmt, daß also für jedes $n \geq 1$ der Wert $\pi(n)$ endlich ist. π ist jedoch eine ziemlich komplizierte zahlentheoretische Funktion. Wir wollen einige Werte von π angeben:

n	1	2	3	4	5	6	7	8	9	10	100
$\pi(n)$	1	2	3	5	7	11	15	22	30	42	190569292

Für große n gilt

$$\pi(n) \approx \frac{1}{4\sqrt{3}n} e^{\pi\sqrt{\frac{2n}{3}}}.$$

Die Partitionen veranschaulicht man sich am besten durch das sogenannte *Young-Diagramm* oder *Young-Tableau*. Dabei wird jeder Partition $(\nu_1, \ldots, \nu_k)$ von n eindeutig ein Diagramm der folgenden Form zugeordnet:



Also entsprechen die Konjugationsklassen nilpotenter Matrizen bijektiv den Young-Diagrammen. Diese zu klassifizieren ist offenbar ein kombinatorisches Problem.

Wir kommen später noch einmal auf den Algorithmus zur Berechnung der Jordanschen Normalform einer nilpotenten Matrix zurück. Will man nur die Normalform von A, aber nicht die Transformationsmatrix wissen, dann reicht es, die Elementarteiler zu bestimmen, was mit Hilfe von Aufgabe 4.15 sehr viel einfacher zu bewerkstelligen ist. Dies führt auf folgenden Algorithmus zur Bestimmung der Jordanschen Normalform einer Matrix A, deren charakteristisches Polynom zerfällt.

Algorithmus 4.10 (Jordansche Normalform - I)
INPUT: $A \in \mathrm{Mat}(n, \mathbb{Q})$ mit μ_A zerfällt in Linearfaktoren

OUTPUT: Liste mit den Eigenwerten von A und den Elementarteilern

1. **Schritt:** Bestimme das Minimalpolynom μ_A von A und faktorisiere es.
2. **Schritt:** Wenn μ_A nicht in Linearfaktoren zerfällt, gib eine Fehlermeldung zurück.
3. **Schritt:** Für jeden Eigenwert λ_i mit $\mathrm{mult}(\mu_A, \lambda_i) = m_i$ bestimme man für $j = 0, \ldots, m_i+1$ die Zahlen $\mathrm{rang}\left((A-\lambda_i \mathbb{1}_n)^j\right)$ und berechne daraus den Vektor der Elementarteiler $(t_{i1}, \ldots, t_{im_i})$. Den Eigenwert und den Vektor der Elementarteiler speichere man als i-ten Eintrag in einer Liste `nf`.
4. **Schritt:** Man gebe die Liste `nf` zurück.

Kommen wir nun zum Beweis von Satz 4.1.

Beweis von Satz 4.1: Nach Proposition 3.33 zerfällt V in die direkte Summe der Haupträume $V_i := \mathrm{Hau}(f, \lambda_i)$, $i = 1, \ldots, r$, und diese sind invariant unter f und $f - \lambda_i \mathrm{id}_V$. Betrachten wir nun die Abbildungen

$$(f - \lambda_i \mathrm{id}_V)_{V_i} : V_i \to V_i$$

für $i = 1, \ldots, r$, so sind diese nilpotent mit $\chi_{(f-\lambda_i \mathrm{id}_V)_{V_i}} = (-1)^{n_i} t^{n_i}$ und $\mu(f - \lambda_i \mathrm{id}_V)_{V_i} = t^{m_i}$ (vgl. Korollar 3.34). Nach Lemma 4.3 gibt es dann aber für jedes $i = 1, \ldots, r$ Basen B_i von V_i und natürliche Zahlen t_{ij}, $j = 1, \ldots, m_i$, so daß gilt

a. $\sum_{j=1}^{m_i} j \cdot t_{ij} = n_i = \dim_K \big(\mathrm{Hau}(f, \lambda_i)\big)$,
b. $\sum_{j=1}^{m_i} t_{ij} = \dim_K \Big(\mathrm{Eig}\big((f - \lambda_i \mathrm{id}_V)_{V_i}, 0\big)\Big) = \dim_K \big(\mathrm{Eig}(f, \lambda_i)\big)$,
c. $t_{im_i} \geq 1$ und
d. $M_{B_i}^{B_i}(f_{V_i}) = \lambda_i \mathbb{1}_{n_i} + M_{B_i}^{B_i}\big((f - \lambda_i \mathrm{id}_V)_{V_i}\big) = \lambda_i \mathbb{1}_{n_i} + \Big(\bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} J_j(0)\Big) = \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} J_j(\lambda_i)$.

Da für $B = B_1 \cup \ldots \cup B_r$ gilt

$$M_B^B(f) = \bigoplus_{i=1}^{r} M_{B_i}^{B_i}(f_{V_i}),$$

folgt die Behauptung. □

Wir wollen nun die wichtigsten Ergebnisse der beiden letzten Abschnitte für Endomorphismen $f \in \mathrm{End}_K(V)$, deren charakteristisches Polynom zerfällt, in koordinatenfreier Form zusammenfassen, d. h. ohne dabei auf die Wahl einer geeigneten Basis zu verweisen.

Satz 4.11 (Jordan-Zerlegung)
Es sei $\chi_f = (-1)^n \prod_{i=1}^{r} (t - \lambda_i)^{n_i}$, $\lambda_i \neq \lambda_j$ *für* $i \neq j$. *Dann gelten:*

a. $V = \bigoplus_{i=1}^{r} V_i$ *mit* $V_i = \mathrm{Hau}(f, \lambda_i)$ *(“Hauptraum-Zerlegung”)*
b. $\dim_K (V_i) = \mathrm{mult}(\chi_f, \lambda_i)$.
c. f *hat eine* eindeutig *bestimmte additive* Jordan-Zerlegung $f = f_s + f_n$ *mit*
   (i) f_s *ist diagonalisierbar (man sagt auch* halbeinfach*),*
   (ii) f_n *ist nilpotent, und*
   (iii) $f_s \circ f_n = f_n \circ f_s$.
   Dabei gilt, es gibt Polynome $P, Q \in K[t]$ *mit* $f_s = P(f)$ *und* $f_n = Q(f)$.
d. V_i *ist invariant unter* f, f_s *und* f_n.
e. $V_i = \mathrm{Eig}(f_s, \lambda_i)$.
f. $(f_n)_{V_i}$ *hat den Nilpotenzindex* $\mathrm{mult}(\mu_f, \lambda_i)$

Beweis: Die Aussagen in a. und b. wurden bereits in Proposition 3.33 gezeigt.

Nach Satz 4.1 gibt es eine Basis B, so daß $J_f := M_B^B(f)$ Jordansche Normalform hat, und nach Bemerkung 4.2 d. gibt es dann eine Diagonalmatrix S und eine

nilpotente Matrix N mit $J_f = S+N$ und $S \circ N = N \circ S$. Seien nun $f_s, f_n \in \mathrm{End}_K(V)$ die eindeutig bestimmten Endomorphismen mit

$$M_B^B(f_s) = S \quad \text{und} \quad M_B^B(f_n) = N.$$

Dann folgt aus $M_B^B(f) = M_B^B(f_s) + M_B^B(f_n) = M_B^B(f_s + f_n)$ und $M_B^B(f_s \circ f_n) = M_B^B(f_s) \circ M_B^B(f_n) = M_B^B(f_n) \circ M_B^B(f_s) = M_B^B(f_n \circ f_s)$, daß

$$f = f_s + f_n \quad \text{und} \quad f_s \circ f_n = f_n \circ f_s.$$

Zudem ist f_s diagonalisierbar und f_n nilpotent, da $M_B^B(f_n)$ nilpotent ist. Damit ist die Existenz einer Jordan-Zerlegung in c. gezeigt, und Lemma 4.12 liefert die Existenz der Polynome P und Q.

Ferner folgt aus der Definition von f_s unmittelbar, daß $V_i = \mathrm{Eig}(f_s, \lambda_i)$ und daß $(f - \lambda_i \mathrm{id}_V)_{V_i} = (f_n)_{V_i}$, womit aus Proposition 3.33 folgt, daß $\mathrm{mult}(\mu_f, \lambda_i)$ der Nilpotenzindex von $(f_n)_{V_i}$ ist. Damit sind e. und f. gezeigt.

Die spezielle Gestalt von $M_B^B(f_s)$ und $M_B^B(f_n)$ zeigt ferner unmittelbar, daß V_i invariant ist unter f_s und unter f_n. Die Invarianz gegenüber f wurde bereits in Lemma 3.31 gezeigt. Damit ist auch d. bewiesen.

Es bleibt schließlich noch die Eindeutigkeit der Jordan-Zerlegung in c. zu zeigen.

Sei $f = f_s' + f_n'$ eine zweite Jordan-Zerlegung. Wir wissen bereits, daß es Polynome $P, Q \in K[t]$ gibt mit

$$f_s = P(f) \quad \text{und} \quad f_n = Q(f).$$

Nach Voraussetzung kommutieren f_s' und f_n' je miteinander und mit sich selbst, so daß beide auch mit f kommutieren. Aber damit folgt

$$f_s' \circ f_s = f_s' \circ P(f) = P(f) \circ f_s' = f_s \circ f_s'$$

und analog

$$f_n' \circ f_n = f_n \circ f_n'.$$

Da zudem f_s und f_s' diagonalisierbar sind, folgt aus Aufgabe 3.47, daß f_s und f_s' simultan diagonalisierbar sind, d. h. insbesondere es gibt eine Basis B, bezüglich derer

$$M_B^B(f_s - f_s') = M_B^B(f_n' - f_n) \tag{53}$$

eine Diagonalmatrix ist. Andererseits ist die rechte Seite von (53) nilpotent. Denn seien etwa $N, N' \in \mathbb{N}$ mit $f_n^N = 0 = (f_n')^{N'}$, dann ist, da f_n und f_n' kommutieren,

$$\begin{aligned}(f_n' - f_n)^{N+N'} &= \sum_{k=0}^{N'} (-1)^{N+N'-k} (f_n')^k \circ f_n^{N+N'-k} \\ &\quad + \sum_{k=N'+1}^{N+N'} (-1)^{N+N'-k} (f_n')^k \circ f_n^{N+N'-k} = 0 + 0 = 0.\end{aligned}$$

Aber die Nullmatrix ist die einzige nilpotente Diagonalmatrix und mithin gilt $f_s = f_s'$ und $f_n = f_n'$. □

Lemma 4.12
Mit den Bezeichnungen und Voraussetzungen von Satz 4.11 sei $\mu_f = \prod_{i=1}^{r}(t-\lambda_i)^{m_i}$. *Wir setzen* $q_i := \frac{\mu_f}{(t-\lambda_i)^{m_i}}$, $i = 1,\dots,r$, *dann gibt es Polynome* $p_1,\dots,p_r \in K[t]$ *mit* $p_1q_1 + \dots + p_rq_r = 1$. *Setzen wir nun* $Q_i := q_ip_i \in K[t]$, $i = 1,\dots,r$, *sowie* $P := \sum_{i=1}^{r}\lambda_iQ_i \in K[t]$ *und* $Q := \sum_{i=1}^{r}(t-\lambda_i)\cdot Q_i \in K[t]$, *dann gilt für die im Beweis von Satz 4.11 definierten Endomorphismen* f_s *und* f_n

$$f_s = P(f) \quad \textit{und} \quad f_n = Q(f).$$

Beweis: Nach Definition gilt $\mathrm{ggT}(q_1,\dots,q_r) = 1$ und mithin folgt die Existenz der p_i, $i = 1,\dots,r$. (Vgl. Beweis von Proposition 3.33.)

Es reicht nun zu zeigen, daß $P(f)$ und $Q(f)$ auf die V_i, $i = 1,\dots,r$, eingeschränkt jeweils mit f_s respektive f_n übereinstimmen.

Im Beweis von Proposition 3.33 wurde in (48) gezeigt, daß $Q_i(f)$ die Projektion auf V_i mit Kern $\bigoplus_{j\neq i} V_j$ ist. Mithin gilt für $x \in V_i$, unter Berücksichtigung der Definition von f_n,

$$Q(f)(x) = \sum_{j=1}^{r}(f-\lambda_j\,\mathrm{id}_V)\big(Q_j(f)(x)\big) = (f-\lambda_i\,\mathrm{id}_V)(x) = f_n(x).$$

Ferner gilt

$$\begin{aligned}\big(P(f)+Q(f)\big)(x) &= \textstyle\sum_{j=1}^{r}\Big(\lambda_jQ_j(f)(x) + \big((f-\lambda_j\,\mathrm{id}_V)\circ Q_j(f)\big)(x)\Big)\\ &= \textstyle\sum_{j=1}^{r} f\big(Q_j(f)(x)\big) = f(x).\end{aligned}$$

Es folgt, $f_n = Q(f)$ und $f = P(f) + Q(f)$, also auch $f_s = f - f_n = P(f)$. □

Abschließend wollen wir noch einen Algorithmus beschreiben, mit Hilfe dessen man die Jordansche Normalform einer Matrix A inklusive der zugehörigen Transformationsmatrix bestimmen kann.

Algorithmus 4.13 (Jordansche Normalform - II)
INPUT: $A \in \mathrm{Mat}(n,\mathbb{Q})$ mit μ_A zerfällt in Linearfaktoren.

OUTPUT: J_A und eine Transformationsmatrix $T \in \mathrm{Gl}_n(K)$ mit $T^{-1}\circ A\circ T = J_A$.

1. **Schritt:** Bestimme das Minimalpolynom μ_A von A und faktorisiere es.
2. **Schritt:** Wenn μ_A nicht in Linearfaktoren zerfällt, gebe man eine Fehlermeldung zurück, andernfalls gilt $\mu_A = \prod_{i=1}^{r}(t-\lambda_i)^{m_i}$.
3. **Schritt:** Für $i = 1,\dots,r$ bilde man die Matrix $B_i = A - \lambda_i\mathbb{1}_n$ und führe folgende Schritte aus:
   - **Schritt a.:** Bestimme eine Basis D_{m_i} von $\mathrm{Lös}\left(B_i^{m_i},0\right)$ sowie eine Basis D_{m_i-1} von $\mathrm{Lös}\left(B_i^{m_i-1},0\right)$.
   - **Schritt b.:** Tausche D_{m_i-1} mittels des Satzes von Steinitz in D_{m_i} hinein und bestimme die in D_{m_i} verbliebenen Vektoren $e^i_{m_i1},\dots,e^i_{m_it_{im_i}}$.
   - **Schritt c.:** Für $k = 1,\dots,t_{im_i}$ schreibe man die Vektoren $B_i^{m_i-1}e^i_{m_ik},\dots,B_i^0e^i_{m_ik}$ als Spaltenvektoren in eine Matrix T.
   - **Schritt d.:** Für $j = m_i-1,\dots,1$ tue man folgendes:

- bestimme eine Basis D_{j-1} von Lös $\left(B_i^{j-1}, 0\right)$;
- tausche $D_{j-1} \cup \left\{B_i^{j-l} e_{l,k}^i \mid l = j+1, \ldots, m_i,\ k = 1, \ldots, t_{il}\right\}$ mittels des Satzes von Steinitz in D_j hinein;
- bestimme die in D_j verbliebenen Vektoren $e_{j1}^i, \ldots, e_{jt_{ij}}^i$;
- für $k = 1, \ldots, t_{ij}$ schreibe die Vektoren $B_i^{j-1} e_{jk}^i, \ldots, B_i^0 e_{jk}^i$ als Spaltenvektoren in die Matrix T.

4. Schritt: Gib $T^{-1} \circ A \circ T$ und T zurück.

Beispiel 4.14

Wir wollen nun die Jordansche Normalform und die Transformationsmatrix von

$$A = \begin{pmatrix} 2 & 1 & 1 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix} \in \mathbf{Mat}(4, \mathbb{Q})$$

mit Hilfe von Singular bestimmen.

```
ring r=0,t,lp;
matrix A[4][4]=2,1,1,2,0,1,0,0,0,1,1,1,0,0,0,2;
poly f=min_poly(A);  // Die Prozedur muß nach 3.26
                     // implementiert werden.
factorize(f,2);
↦    [1]:                // D. h. f = (t-1)^2(t-2)^2.
           _[1]=t-1
           _[1]=t-2
     [2]:
           2,2
matrix B=A-unitmat(4);
print(syz(B*B));
↦    -2,   0,  // D. h. Lös((A-1_4)^2, 0) = ⟨(-2,1,0,0)^t, (0,-1,2,0)^t⟩.
      1,  -1,
      0,   2,
      0,   0
print(syz(B)); ↦    -1,  // D. h. Lös(A-1_4, 0) = ⟨(-1,0,1,0)^t⟩.
                     0,
                     1,
                     0
```

Wir können mithin $e_1 := (A - \mathbb{1}_4)(0,-1,2,0)^t = (1,0,-1,0)^t$ und $e_2 = (0,-1,2,0)^t$ als Basisvektoren für Lös $\left((A - \mathbb{1}_4)^2, 0\right)$ wählen.

```
B=A-2*unitmat(4);
print(syz(B*B));
↦    1,  0,  // D. h. Lös((A-2*1_4)^2, 0) = ⟨(1,0,0,0)^t, (0,0,1,1)^t⟩.
     0,  0,
     0,  1,
     0,  1
```

```
print(syz(B)); ↦    1,  // D. h. Lös (A − 2𝟙4, 0) = ⟨(1, 0, 0, 0)^t⟩.
                    0,
                    0,
                    0
```

Wir können mithin $e_3 := (A - \mathbb{1}_4)(0,0,1,1)^t = (3,0,0,0)^t$ und $e_4 = (0,0,1,1)^t$ als Basisvektoren für Lös $\big((A-\mathbb{1}_4)^2, 0\big)$ wählen.

Damit erhalten wir als Transformationsmatrix

$$T = (e_1, e_2, e_3, e_4) = \begin{pmatrix} 1 & 0 & 3 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 2 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \in \mathrm{Gl}_4(\mathbb{Q})$$

und als Jordansche Normalform

$$J_A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

Es gibt in Singular in der Bibliothek `jordan.lib` auch bereits eine vorgefertigte Prozedur, `jordan`, zur Berechnung der Jordanschen Normalform sowie der Transformationsmatrix.

Anwendungen findet die Jordansche Normalform zum Beispiel in der Theorie der linearen Differentialgleichungen, wo ein Fundamentalsystem mit Hilfe der Exponentialabbildung einer Matrix beschrieben wird. Diese wird mit Hilfe der Jordanschen Normalform von A beschrieben. Vgl. Aufgabe IV.1.32.

AUFGABEN

Aufgabe 4.15

Mit den Bezeichnungen aus Satz 4.1 zeige man, für $i = 1, \dots, r$ und $1 \le j \le m_i$ gilt:

$$t_{ij} = \operatorname{rang}\big((f - \lambda_i \operatorname{id}_V)^{j-1}\big) - 2\operatorname{rang}\big((f - \lambda_i \operatorname{id}_V)^{j}\big) + \operatorname{rang}\big((f - \lambda_i \operatorname{id}_V)^{j+1}\big)$$

bzw.

$$t_{ij} = \operatorname{rang}\big((A - \lambda_i \mathbb{1}_n)^{j-1}\big) - 2\operatorname{rang}\big((A - \lambda_i \mathbb{1}_n)^{j}\big) + \operatorname{rang}\big((A - \lambda_i \mathbb{1}_n)^{j+1}\big).$$

Hinweise: 1. Zeige, $J_j(0)^l = \big(\delta_{\mu+l,\nu}\big)_{\mu,\nu=1,\dots,j}$ und $\operatorname{rang}\big(J_j(0)^l\big) = \max\{0, j-l\}$ für $l \in \mathbb{N}$. 2. Man betrachte zunächst den Fall $r = 1$ und $\lambda_1 = 0$. 3. Den allgemeinen Fall führe man auf die Abbildungen $g_i := (f - \lambda_i \operatorname{id}_V)_{\mathrm{Hau}(f,\lambda_i)}$ zurück.

Aufgabe 4.16

Bestimme die Jordansche Normalform der folgenden Matrizen:

$$A = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 3 & 0 & 0 \\ -1 & 1 & 2 & 1 \\ -1 & 1 & 0 & 3 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 1 & 0 & 0 & -1 \\ -1 & 0 & -1 & 0 \\ 1 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Aufgabe 4.17
Es sei $A \in \mathrm{Mat}(5,K)$ mit $\chi_A = t(t-1)^4$, $\mu_A = t(t-1)^2$ und $\mathrm{rang}(A - \mathbb{1}_5) = 2$. Bestimme die Jordansche Normalform von A.

Aufgabe 4.18
Bestimme eine Basis B von $\mathbb{R}^5$, bezüglich derer die Matrixdarstellung der Abbildung $f \in \mathrm{End}_{\mathbb{R}}\left(\mathbb{R}^5\right)$ mit $f(x_1,x_2,x_3,x_4,x_5) = (x_1 - x_2, x_1 + 2x_2 - x_3, -x_1 + 3x_3, -x_1 - 2x_2 - 2x_3 + 2x_4 - x_5, x_1 - x_3 + 2x_5)^t$ Jordansche Normalform hat.

Aufgabe 4.19
Es sei V ein endlich-dimensionaler K-Vektorraum, $f \in \mathrm{End}_K(V)$ und $V = U_1 \oplus U_2$ mit $f(U_i) \subseteq U_i$ für $i = 1,2$. Ferner sei B_i eine Basis von U_i mit $M_{B_i}^{B_i}\left(f_{U_i}\right) = J_{f_{U_i}}$ in Jordanscher Normalform, dann ist $B = B_1 \cup B_2$ eine Basis von V, so daß die folgende Matrix in Jordanscher Normalform ist $J_f := M_B^B(f) = J_{f_{U_1}} \oplus J_{f_{U_2}}$.

Aufgabe 4.20
Sei $A = A_1 \oplus A_2 \in \mathrm{Mat}(n,K)$ mit $A_i \in \mathrm{Mat}(n_i,K)$. Zeige, ist J_{A_i} eine Jordansche Normalform von A_i, so ist $J_A = J_{A_1} \oplus J_{A_2}$ eine Jordansche Normalform von A.

Aufgabe 4.21
Es sei $f \in \mathrm{End}_K(V)$ mit $n = \dim_K(V)$ und $\mu_f = t^n$. Bestimme für $i = 1,\ldots,n-1$ die Jordansche Normalform von f^i.

Aufgabe 4.22
Es sei $A \in \mathrm{Mat}(n,K)$. Zeige, A und A^t sind konjugiert.

Aufgabe 4.23
Es seien $A, B \in \mathrm{Gl}_n(K)$. Prüfe, ob die folgenden Aussagen gleichwertig sind:

a. $\chi_A = \chi_B$, $\mu_A = \mu_B$ und $\mathrm{rang}(A) = \mathrm{rang}(B)$.
b. A und B sind konjugiert.

Aufgabe 4.24 a. Schreibe eine Singular-Prozedur `jordaninvariants`, die eine quadratische Matrix A einliest und, falls das Minimalpolynom $\mu_A = \prod_{i=1}^{r}(t-\lambda_i)^{m_i}$ über $\mathbb{Q}$ in Linearfaktoren zerfällt, eine Liste `nf` von r Listen ausgibt, so daß in der Liste `nf[i]` als Einträge gerade der Eigenwert λ_i (Typ `poly`) und der Vektor der Elementarteiler $t_{i,1},\ldots,t_{i,m_i}$ (Typ `intvec`) enthalten sind.

Man verwende folgenden Algorithmus:

INPUT: $A \in \mathrm{Mat}(n,\mathbb{Q})$

OUTPUT: Liste mit den Eigenwerten von A und den Elementarteilern

1. **Schritt:** Falls A nicht quadratisch ist, gib eine Fehlermeldung zurück.
2. **Schritt:** Bestimme das Minimalpolynom μ_A von A und faktorisiere es mittels `factorize`.
3. **Schritt:** Wenn μ_A nicht in Linearfaktoren zerfällt, gib eine Fehlermeldung zurück.

4. Schritt: Für jeden Eigenwert λ_i mit $\mathrm{mult}(\mu_A, \lambda_i) = m_i$ bestimme man für $j = 0, \ldots, m_i + 1$ die Zahlen $\mathrm{rang}\left((A - \lambda_i \mathbb{1}_n)^j\right)$ und berechne daraus den Vektor der Elementarteiler $t_{i1}, \ldots, t_{im_i}$. Den Eigenwert und den Vektor der Elementarteiler speichere man in der Liste `nf[i]`.

5. Schritt: Man gebe die Liste `nf` zurück.

b. Schreibe eine Singular-Prozedur `jordannf`, die eine quadratische Matrix A einliest, und, falls das Minimalpolynom über $\mathbb{Q}$ zerfällt, eine Jordansche Normalform von A ausgibt.

Hinweise: Ist `p` ein Polynom, dann liefert `jet(p,0)` den konstanten Anteil des Polynoms. – Die Ergenisse können mit den Prozeduren `jordan` und `jordanform` aus der Bibliothek `jordan.lib` getestet werden.

KAPITEL IV

Endlich-dimensionale Hilberträume

Bisher haben wir Vektorräume über einem Körper als abstrakte, allgemeine lineare Struktur kennengelernt, in der man Unterräume verschiedener Dimension betrachten kann und zwischen denen die strukturerhaltenden Abbildungen die linearen Abbildungen sind. Dies alles wurde sehr konkret, indem wir Basen einführten: Vektoren wurden zu Tupeln von Elementen von K und lineare Abbildungen zu Matrizen.

Zwar haben wir zur Veranschaulichung den $\mathbb{R}^n$ herangezogen, Untervektorräume als lineare (nicht gekrümmte) Teilmengen, z. B. Geraden und Ebenen, durch den Ursprung interpretiert, aber daß dabei der Grundkörper $\mathbb{R}$ war, spielte, mathematisch gesehen, überhaupt keine Rolle. Nur bei der Jordanschen Normalform gingen Eigenschaften des Körpers K, wie die algebraische Abgeschlossenheit etwa von $\mathbb{C}$, ganz wesentlich ein. Der Vektorraum war für uns bisher kaum ein geometrisches Objekt, das wichtigste Vorkommen war das als Lösungsraum eines homogenen linearen Gleichungssystems.

Mit den bisherigen Begriffen haben wir keine Möglichkeit, zu sagen, was die Länge eines Vektors oder der Winkel zwischen zwei Vektoren ist. Um diese Begriffe erklären zu können, brauchen wir *neue Strukturen* wie Normen und Skalarprodukte. Wenn wir messen, dann wollen wir aber auch ausdrücken können, wann z. B. die eine Länge größer als die andere ist, d. h. die Länge oder der Winkel soll eine reelle Zahl sein. Deshalb werden in diesem Kapitel reelle (und komplexe) Vektorräume eine besondere Rolle spielen.

1. Bilinearformen und Skalarprodukte

Generalvoraussetzung: In diesem Kapitel bezeichne K, sofern nicht anders angegeben, einen beliebigen Körper und V bezeichne einen Vektorraum über K.

Wir wollen den Abschnitt mit der reellen Ebene $\mathbb{R}^2$ als Beispiel beginnen.

Beispiel 1.1

Wir definieren uns zunächst zwei Abbildungen

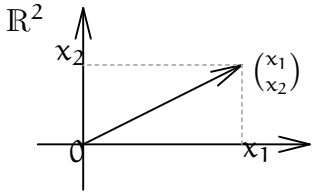
$$\|\cdot\| : \mathbb{R}^2 \to \mathbb{R}_{\geq 0},$$

die einem Vektor $x = (x_1, x_2)^t \in \mathbb{R}^2$ seine Länge $\|x\|$ zuordnet, sowie

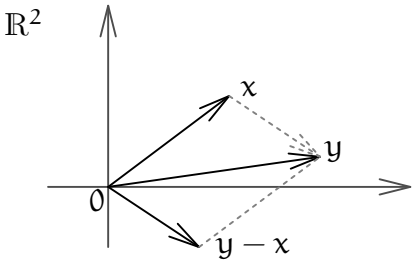
$$d : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}_{\geq 0},$$

die zwei Punkten $x \in \mathbb{R}^2$ und $y \in \mathbb{R}^2$ ihren Abstand $d(x, y)$ zuweist.

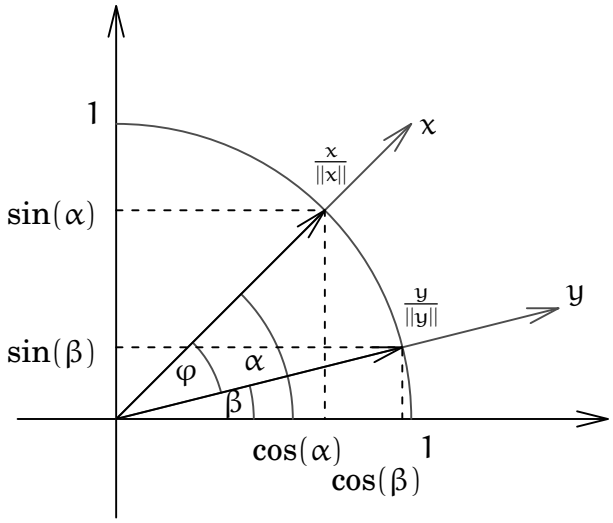
Der Satz von Pythagoras liefert dann $\|x\| = \sqrt{x_1^2 + x_2^2}$.



Wir nennen $\|x\|$ auch die *Norm* des Vektors x. Da der Abstand der Punkte $x = (x_1, x_2)^t$ und $y = (y_1, y_2)^t$ gerade die Länge des Vektors $x - y$ ist, folgt somit $d(x, y) = \|x - y\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^t}$.



Mit Hilfe der Norm können wir - nach einigen geometrischen Überlegungen - auch den Winkel $\measuredangle(x, y)$, den zwei Vektoren x und y miteinander einschließen, bestimmen.



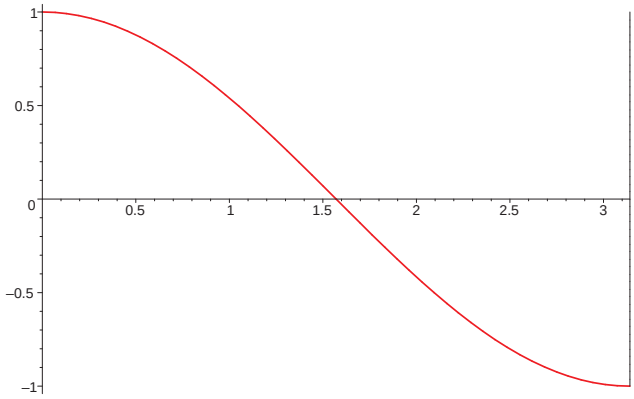
Dazu gehen wir zunächst zu den normierten Vektoren $\frac{x}{\|x\|}$ und $\frac{y}{\|y\|}$ über, die beide die Länge eins haben, wobei wir $x \neq 0$ und $y \neq 0$ voraussetzen. Mit den Bezeichnungen in der Skizze gilt dann

$$\measuredangle(x, y) = \measuredangle\left(\frac{x}{\|x\|}, \frac{y}{\|y\|}\right) = \alpha - \beta = \varphi.$$

Um φ selbst (im Bogenmaß) auszudrücken, müßte man die Länge des Kreisbogens zwischen $\frac{x}{\|x\|}$ und $\frac{y}{\|y\|}$ messen, also einer gekrümmten Linie. Dazu bedarf es gewisser Grenzprozesse, die in der Analysis betrachtet werden.

Zur anschaulichen Herleitung des Winkels φ mit $0 \leq \varphi \leq \pi$, benötigen wir nur, daß die Funktion

$$\cos : [0, \pi] \to \mathbb{R} : \varphi \mapsto \cos(\varphi)$$



injektiv ist. Also reicht es, $\cos(\varphi)$ zu kennen, um den Winkel φ eindeutig beschrieben zu haben. Unter Zuhilfenahme der obigen Skizze und des Additionstheorems für den Cosinus erhalten wir

$$\begin{aligned}\cos(\varphi) &= \cos(\alpha-\beta)\\ &= \cos(\alpha)\cos(\beta)+\sin(\alpha)\sin(\beta)\\ &= \tfrac{x_1y_1+x_2y_2}{\|x\|\cdot\|y\|}.\end{aligned}$$

Dies führt zur Definition einer weiteren Abbildung

$$\langle\cdot,\cdot\rangle : \mathbb{R}^2\times\mathbb{R}^2\to\mathbb{R} : (x,y) = \left((x_1,x_2)^t,(y_1,y_2)^t\right) \mapsto \langle x,y\rangle := x_1y_1+x_2y_2,$$

welche wir *Skalarprodukt* nennen und mit deren Hilfe wir erhalten

$$\cos(\varphi) = \frac{\langle x,y\rangle}{\|x\|\cdot\|y\|}$$

oder alternativ

$$\measuredangle(x,y) = \varphi = \arccos\left(\frac{\langle x,y\rangle}{\|x\|\cdot\|y\|}\right).$$

Wir sind also mittels recht einfacher Abbildungen in der Lage, Längen und Winkel auszudrücken.

Dieses Beispiel soll die im folgenden durchgeführten Definitionen motivieren.

Definition 1.2 a. Eine Abbildung

$$b : V\times V\to K,$$

die linear in beiden Argumenten ist, nennen wir *bilinear* oder *Bilinearform*, d. h. es muß für $x,y,z\in V$ und $\lambda,\mu\in K$ gelten (vgl. Definition II.9.7):

(i) $b(\lambda x+\mu y,z) = \lambda b(x,z)+\mu b(y,z)$, und

(ii) $b(z,\lambda x+\mu y) = \lambda b(z,x)+\mu b(z,y)$.

Wir bezeichnen mit $\mathrm{Bil}_K(V) = \{b : V\times V\to K \mid b \text{ ist bilinear}\}$ die Menge der Bilinearformen auf V.

b. Eine Bilinearform $b : V\times V\to K$ heißt *symmetrisch*, falls für $x,y\in V$ stets $b(x,y) = b(y,x)$ gilt.

c. Ist $K = \mathbb{C}$, so heißt eine Abbildung

$$b : V \times V \to \mathbb{C}$$

sesquilinear oder *Sesquilinearform*, wenn b linear im ersten und *semilinear* im zweiten Argument ist, d. h. wenn für $x, y, z \in V$ und $\lambda, \mu \in \mathbb{C}$ gilt:

(i) $b(\lambda x + \mu y, z) = \lambda b(x, z) + \mu b(y, z)$, und

(ii) $b(z, \lambda x + \mu y) = \overline{\lambda} b(z, x) + \overline{\mu} b(z, y)$.

d. Ist $K = \mathbb{C}$, dann nennen wir eine Sesquilinearform $b : V \times V \to \mathbb{C}$ *hermitesch* oder *anti-symmetrisch*, wenn für $x, y \in V$ stets $b(x, y) = \overline{b(y, x)}$ gilt.
Beachte, in diesem Fall gilt für $x \in V$ immer $b(x, x) = \overline{b(x, x)}$, also $b(x, x) \in \mathbb{R}$!

e. Ist $K = \mathbb{R}$ und b eine symmetrische Bilinearform oder ist $K = \mathbb{C}$ und b eine hermitesche Sesquilinearform, so heißt b *positiv definit* (bzw. *negativ definit*), wenn für alle $0 \neq x \in V$ gilt:

$$b(x, x) > 0 \quad \text{(bzw.} \quad b(x, x) < 0\text{)}.$$

Eine positiv definite symmetrische Bilinearform bzw. eine positiv definite hermitesche Sesquilinearform nennt man auch ein *Skalarprodukt* oder ein *inneres Produkt* auf V.
Für ein Skalarprodukt ist die Bezeichnung

$$\langle \cdot, \cdot \rangle : V \times V \to K$$

gebräuchlicher als b.

f. Ist $K = \mathbb{R}$ oder $K = \mathbb{C}$ und ist $\langle \cdot, \cdot \rangle$ ein Skalarprodukt auf V, so nennen wir das Tupel $\big(V, \langle \cdot, \cdot \rangle\big)$ im Fall $K = \mathbb{R}$ einen *euklidischen (Vektor-)Raum* und im Fall $K = \mathbb{C}$ einen *unitären (Vektor-)Raum*.
Ist zudem V endlich-dimensional, so spricht man in beiden Fällen auch von einem endlich-dimensionalen *Hilbertraum*.[1]

Bemerkung 1.3 a. Ist $K = \mathbb{R}$, dann ist eine Abbildung $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ genau dann ein Skalarprodukt, wenn für $x, y, z \in V$ und $\lambda, \mu \in \mathbb{R}$ gilt

(i) $\langle \lambda x + \mu y, z \rangle = \lambda \langle x, z \rangle + \mu \langle y, z \rangle$,

(ii) $\langle x, y \rangle = \langle y, x \rangle$,

(iii) $\langle x, x \rangle > 0$ falls $x \neq 0$.

Man beachte, daß aus der Linearität von $\langle \cdot, \cdot \rangle$ im ersten Argument sofort $\langle 0, 0 \rangle = 0$ folgt.

b. Ist $K = \mathbb{C}$, dann ist eine Abbildung $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$ genau dann ein Skalarprodukt, wenn für $x, y, z \in V$ und $\lambda, \mu \in \mathbb{C}$ gilt

[1] Im Falle $\dim_K(V) = \infty$ würde man für einen Hilbertraum zusätzlich fordern, daß jede Cauchy-Folge auch konvergiert, was im endlich-dimensionalen Fall aber ohnehin gilt. Da uns im weiteren Verlauf nur endlich-dimensionale euklidische und unitäre Räume interessieren, können wir diese Forderung also getrost vergessen und der Analysis - oder genauer der Funktionalanalysis - überlassen.

(i) $\langle \lambda x + \mu y, z\rangle = \lambda\langle x, z\rangle + \mu\langle y, z\rangle$,
(ii) $\langle x, y\rangle = \overline{\langle y, x\rangle}$,
(iii) $\langle x, x\rangle > 0$ falls $x \neq 0$.

Man beachte, daß wiederum aus der Linearität von $\langle \cdot, \cdot \rangle$ im ersten Argument sofort $\langle 0, 0\rangle = 0$ folgt.

c. Wenn wir im Folgenden den Fall eines euklidischen und eines unitären Raumes parallel behandeln wollen, dann werden wir uns häufig zunutze machen, daß für eine reelle Zahl λ gilt $\lambda = \overline{\lambda}$, daß mithin zum Beispiel für eine Bilinearform auf einem reellen Vektorraum V gleichbedeutend ist, für alle $\lambda, \mu \in \mathbb{R}$ und $x, y, z \in V$ zu zeigen

$$\langle z, \lambda x + \mu y\rangle = \lambda\langle z, x\rangle + \mu\langle z, y\rangle$$

oder

$$\langle z, \lambda x + \mu y\rangle = \overline{\lambda}\langle z, x\rangle + \overline{\mu}\langle z, y\rangle,$$

oder daß für eine reelle Matrix $A \in \mathrm{Mat}(n, \mathbb{R})$ genau dann $A = A^t$ gilt, wenn $A = \overline{A}^t$ erfüllt ist.

Dieser Umstand erlaubt es uns, viele Beweise ohne lästige Fallunterscheidungen führen zu können.

Definition 1.4 a. Ist $K = \mathbb{R}$ oder $K = \mathbb{C}$, so setzen wir $A^* := \overline{A}^t$ für $A \in \mathrm{Mat}(n, K)$. Wir nennen A^* die zu A *adjungierte Matrix*.[2] .

b. Für einen beliebigen Körper K nennen wir eine Matrix $A \in \mathrm{Mat}(n, K)$ *symmetrisch*, wenn $A = A^t$.

c. Wir nennen eine Matrix $A \in \mathrm{Mat}(n, \mathbb{C})$ *hermitesch* oder *selbstadjungiert*, wenn $A = \overline{A}^t = A^*$.

d. Wir nennen eine symmetrische Matrix $A \in \mathrm{Mat}(n, \mathbb{R})$ *positiv definit* (bzw. *negativ definit*), wenn die in Beispiel 1.5 a. definierte symmetrische Bilinearform b_A positiv definit (bzw. negativ definit) ist.

e. Wir nennen eine hermitesche Matrix $A \in \mathrm{Mat}(n, \mathbb{C})$ *positiv definit* (bzw. *negativ definit*), wenn die in Beispiel 1.5 b. definierte hermitesche Sesquilinearform b^s_A positiv definit (bzw. negativ definit) ist.

Beispiel 1.5 a. Ist $A \in \mathrm{Mat}(n, K)$ eine quadratische Matrix, dann wird durch

$$b_A : K^n \times K^n \to K : (x, y) \mapsto b_A(x, y) = x^t \circ A \circ y = x^t A y$$

eine Bilinearform auf K^n definiert, wie unmittelbar aus der Distributivität des Matrixproduktes folgt.

Zudem ist b_A genau dann symmetrisch, wenn A symmetrisch ist.

Denn $A = (a_{ij})$ symmetrisch, impliziert $b_A(x, y) = x^t A y = (x^t A y)^t = y^t A^t x = y^t A x = b_A(y, x)$. Wenn andererseit b_A symmetrisch ist, dann gilt $a_{ij} = b_A(e_i, e_j) = b_A(e_j, e_i) = a_{ji}$.

b. Analog gilt, ist $A \in \mathrm{Mat}(n, \mathbb{C})$ eine quadratische Matrix, dann wird durch

$$b^s_A : K^n \times K^n \to K : (x, y) \mapsto b^s_A(x, y) = x^t A \overline{y}$$

[2] Wie schon oben bemerkt, heißt das für $A \in \mathrm{Mat}(n, \mathbb{R})$ dann halt $A^* = A^t$.

eine Sesquilinearform auf $\mathbb{C}^n$ definiert, wie wiederum aus der Distributivität des Matrixproduktes und der Additivität der komplexen Konjugation folgt.

Zudem ist b_A^s genau dann hermitesch, wenn A hermitesch ist.

Denn $A = (a_{ij})$ hermitesch, impliziert $b_A^s(x,y) = x^t A\overline{y} = (x^t A\overline{y})^t = \overline{y^t \overline{A}^t \overline{x}} = \overline{y^t A \overline{x}} = \overline{b_A^s(y,x)}$. Wenn andererseit b_A^s hermitesch ist, dann gilt $a_{ij} = b_A^s(e_i,e_j) = \overline{b_A^s(e_j,e_i)} = \overline{a_{ji}}$.

c. Es sei $V = K^n$. Dann definiert

$$\langle\cdot,\cdot\rangle : K^n \times K^n \to K : \left((x_1,\dots,x_n)^t,(y_1,\dots,y_n)^t\right) \mapsto \sum_{i=1}^{n} x_i y_i$$

eine symmetrische Bilinearform auf K^n, wie aus a. mit $A = \mathbb{1}_n$ folgt.

d. Ist $K = \mathbb{R}$, dann ist die symmetrische Bilinearform in c. positiv definit. Wir nennen sie das *kanonische Skalarprodukt* oder das *Standardskalarprodukt* auf $\mathbb{R}^n$.

Um die Behauptung zu sehen, sei $0 \neq x = (x_1,\dots,x_n)^t \in \mathbb{R}^n$. Dann gibt es ein i, so daß $x_i \neq 0$. Aber damit gilt

$$\langle x,x\rangle = \sum_{j=1}^{n} x_j^2 \geq x_i^2 > 0,$$

da $x_j^2 \geq 0$ für alle $j \neq i$.

e. Ist $K = \mathbb{C}$ und $V = \mathbb{C}^n$, dann definieren wir

$$\langle\cdot,\cdot\rangle : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C} : \left((x_1,\dots,x_n)^t,(y_1,\dots,y_n)^t\right) \mapsto \sum_{i=1}^{n} x_i\overline{y_i}.$$

Aus b. folgt mit $A = \mathbb{1}_n$, daß $\langle\cdot,\cdot\rangle$ eine hermitesche Sesquilinearform ist. Da für $0 \neq x = (x_1,\dots,x_n)^t \in \mathbb{C}^n$ zudem $x_i\overline{x_i} = |x_i|^2$ ist, folgt wie in d., daß $\langle\cdot,\cdot\rangle$ auch positiv definit ist.

Wir nennen $\langle\cdot,\cdot\rangle$ das *kanonische Skalarprodukt* oder das *Standardskalarprodukt* auf $\mathbb{C}^n$.

f. Für

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix} \in \mathrm{Mat}(4,\mathbb{R})$$

ist nach a. die Bilinearform b_A auf $\mathbb{R}^2$ symmetrisch. Da für $x = (x_1,x_2)^t, y = (y_1,y_2)^t \in \mathbb{R}^2$ ferner gilt

$$b_A(x,x) = x_1^2 + 2x_1x_2 + 4x_2^2 = (x_1+x_2)^2 + 3x_2^2$$

ist b_A zudem positiv definit, also ein Skalarprodukt.

g. Ein wichtiges Beispiel einer symmetrischen Bilinearform wird durch die Matrix

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \in \mathrm{Mat}(4,\mathbb{R})$$

gegeben. Die zugehörige symmetrische Bilinearform auf dem $\mathbb{R}^4$ ist nicht positiv definit. In der speziellen *Relativitätstheorie* wird der $\mathbb{R}^4$ als Raum-Zeit-Kontinuum aufgefaßt, wobei die drei ersten Koordinaten die Raumkoordinaten sind und die vierte Koordinate für die Zeit steht. Die Geometrie, die zu der hier gegebenen Bilinearform gehört, ist die sogenannte *Minkowski-Geometrie*.

h. Sei $V = \mathcal{C}\big([0,1]\big)$ der $\mathbb{R}$-Vektorraum der auf dem Intervall $[0,1]$ stetigen Funktionen. Dann ist für $f, g \in V$ durch

$$\langle f, g\rangle := \int_0^1 f(x)g(x)dx \in \mathbb{R}$$

nach den Rechenregeln für Integrale eine symmetrische Bilinearform auf V definiert.
$\langle\cdot,\cdot\rangle$ ist sogar positiv definit, denn sei $f(x_0) \neq 0$ für ein $x_0 \in [0,1]$ (o. E. $x_0 \in]0,1[$), dann gibt es ein $\delta > 0$, so daß für alle $x \in [x_0 - \delta, x_0 + \delta] \subseteq [0,1]$ gilt

$$|f(x)| \geq \left|\frac{f(x_0)}{2}\right|.$$

Mithin gilt

$$\langle f, f\rangle = \int_0^1 f^2(x)dx \geq \int_{x_0-\delta}^{x_0+\delta} \frac{f^2(x_0)}{2} = \delta \cdot f^2(x_0) > 0.$$

Bemerkung 1.6 a. Im allgemeinen ist die Bedingung der positiven Definitheit durchaus nicht einfach nachzuprüfen, da man meist nicht alle Vektoren $0 \neq x \in V$ überprüfen kann. Man beachte auch, daß es nicht reicht, etwa für eine Basis $B = (x_1, \ldots, x_n)$ von V nachzuprüfen, daß $\langle x_i, x_i\rangle > 0$ für alle $i = 1, \ldots, n$ gilt.
Betrachte dazu die folgende symmetrische Bilinearform auf $\mathbb{R}^2$

$$\langle\cdot,\cdot\rangle : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R} : \big((a_1,a_2)^t, (b_1,b_2)^t\big) \mapsto a_1b_1 - a_2b_2$$

sowie die Basis $(x_1, x_2) = \big((1,0)^t, (2,1)^t\big)$. Dann gilt $\langle x_1, x_1\rangle = 1 > 0$ und $\langle x_2, x_2\rangle = 3 > 0$, aber $\langle e_2, e_2\rangle = -1 < 0$.

b. Eine symmetrische Matrix $A \in \mathrm{Mat}(n, \mathbb{R})$ ist genau dann positiv (bzw. negativ) definit, wenn für alle $x \in \mathbb{R}^n$ gilt $x^tAx > 0$ (bzw.$x^tAx < 0$).
Damit folgt aber sofort, daß A genau dann positiv definit ist, wenn $-A$ negativ definit ist, und umgekehrt.

c. Eine hermitesche Matrix $A \in \mathrm{Mat}(n, \mathbb{C})$ ist genau dann positiv (bzw. negativ) definit, wenn für alle $x \in \mathbb{C}^n$ gilt $x^tA\overline{x} > 0$ (bzw. $x^tA\overline{x} < 0$).
Und wieder gilt, A ist genau dann positiv definit, wenn $-A$ negativ definit ist, und umgekehrt.

Definition 1.7
Es sei $K = \mathbb{R}$ oder $K = \mathbb{C}$. Eine Abbildung

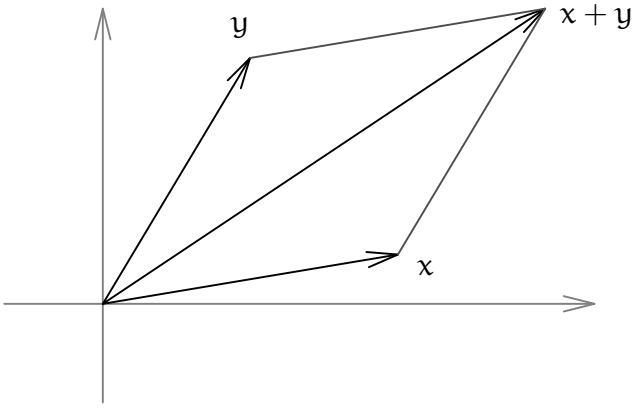
$$\|\cdot\| : V \to \mathbb{R}_{\geq 0},$$

die den folgenden drei Bedingungen genügt, heißt ein *Norm* auf V: für alle $x, y \in V$ und $\lambda \in K$ soll gelten

a. $\|x\| \geq 0$ und $\big(\|x\| = 0 \quad \Leftrightarrow \quad x = 0\big)$,
b. $\|\lambda x\| = |\lambda| \cdot \|x\|$, und
c. $\|x + y\| \leq \|x\| + \|y\|$. (“Dreiecksungleichung”)

Das Tupel $\big(V, \|\cdot\|\big)$ heißt dann ein *normierter (Vektor-)Raum*.

Bemerkung 1.8 a. Wir erinnern uns, daß eine Norm die Länge von Vektoren sowie Abstände messen soll. Bedingung a. kann dann so interpretiert werden, daß jeder Vektor eine nicht-negative Länge hat und daß nur der Nullvektor die Länge null hat. Bedingung b. bedeutet, daß die Streckung eines Vektors um den Faktor λ seine Länge um $|\lambda|$ strecken möge. Und Bedingung c. kann dahingehend interpretiert werden, daß der Weg vom Ursprung über den Punkt x hin zum Punkt $x + y$ unter gar keinen Umständen kürzer ist, als der direkte Weg vom Ursprung zum Punkt $x + y$.



Diese Forderungen scheinen allesamt für eine Funktion, die die Länge von Vektoren beziehungsweise Abstände von Punkten messen soll, nicht unbillig. Und in der Tat reichen diese Forderungen auch bereits aus, um einen vernünftigen Längenbegriff zu erhalten.

b. Um den Längen und Abstandsbegriff mittels einer Norm einzuführen, haben wir wesentlich verwendet, daß ein Vektorraum V mit dem Nullvektor einen ausgezeichneten Punkt besitzt, relativ zu dem man alle anderen Punkte betrachtet, und ferner haben wir uns zunutze gemacht, daß wir in einem Vektorraum Punkte addieren können. Man kann einen vernünftigen Abstandsbegriff jedoch auch auf einer beliebigen Menge M, die eine solche Struktur nicht besitzt, durch wenige Axiome einführen.
Eine Funktion $d : M \times M \to \mathbb{R}$ heißt *Metrik*, falls für alle $x, y, z \in M$ gilt:
(i) $d(x, y) \geq 0$, und $\big(d(x, y) = 0 \quad \Leftrightarrow \quad x = y\big)$.
(ii) $d(x, y) = d(y, x)$.
(iii) $d(x, z) \leq d(x, y) + d(y, z)$. (“Dreiecksungleichung”)
Dabei heißt a. wieder, daß der Abstand zweier Punkte, die nicht gleich sind, stets positiv sein soll, b. bedeutet, daß der Abstand von x zu y gleich dem von y zu x ist und c. sagt, daß es auf keinen Fall kürzer ist, wenn man von x nach z einen Umweg über y macht.
Das Tupel (M, d) nennt man dann einen *metrischen Raum*.

Ist $(V, \|\cdot\|)$ ein normierter Raum, dann wird durch $d(x,y) = \|x-y\|$ für $x, y \in V$ auf V eine Metrik definiert, so daß jeder normierte Raum V in kanonischer Weise ein metrischer Raum ist.

In der Analysis und Topologie ist man häufig mit Mengen beschäftigt, die keine Vektorraumstruktur tragen, auf denen man aber dennoch Abstände - und mit Hilfe der Abstände den Begriff der Konvergenz, der Stetigkeit und ähnliches - betrachten möchte. In diesem Kontext sind metrische Räume von großer Bedeutung. Da wir jedoch nur Mengen betrachten, die eine Vektorraumstruktur tragen, werden wir im Folgenden mit dem Begriff der Norm auskommen, in der Tat sogar mit ganz speziellen Normen, solchen, die von Skalarprodukten herrühren - siehe Satz 1.9.

c. Wir wollen es aber nicht versäumen, ein kurzes Beispiel zu geben, daß zeigt, daß nicht jede Metrik auf einem Vektorraum in der beschriebenen Weise von einer Norm herrührt.

Dazu definieren wir auf $\mathbb{R}$ eine Abstandsfunktion

$$d : \mathbb{R} \times \mathbb{R} \to \mathbb{R} : (x,y) \mapsto \begin{cases} 0, & x = y \\ 1, & x \neq y. \end{cases}$$

Man sieht unmittelbar, daß d in der Tat eine Metrik ist.

Angenommen, es gäbe eine Norm $\|\cdot\|$ auf $\mathbb{R}$, so, daß für $x, y \in \mathbb{R}$ gelten würde $d(x,y) = \|x-y\|$, so wäre $\|1\| = \|2-1\| = d(2,1) = 1$ und damit $1 = d(3,1) = \|3-1\| = \|2\| = |2| \cdot \|1\| = 2$, was nun nicht sein kann, weil es nicht sein darf.

d. Ein Vektorraum V über $K = \mathbb{R}$ oder $K = \mathbb{C}$ kann viele verschiedene Normen besitzen. Aus Sicht der Analysis sind zwei Normen jedoch dann *gleich gut*, wenn etwa Folgen bezüglich der einen Norm genau dann konvergieren, wenn sie das auch bezüglich der anderen Norm tun - denn dann stimmen auch Begriffe wie Stetigkeit bezüglich beider Normen überein. Dies ist nun genau dann der Fall, wenn zwei Normen im folgenden Sinne äquivalent sind:

Zwei Normen $\|\cdot\|$ und $|\cdot|$ auf V heißen *äquivalent*, kurz $\|\cdot\| \sim |\cdot|$, falls es Konstanten $m, M \in \mathbb{R}_{>0}$ gibt mit $m \cdot \|x\| \leq |x| \leq M \cdot \|x\|$ für alle $x \in V$.

In Aufgabe 1.27 wird gezeigt, daß je zwei Normen auf einem endlich-dimensionalen Vektorraum äquivalent sind, daß es also im Falle von endlich-dimensionalen Vektorräumen, etwa beim $\mathbb{R}^n$, aus der Sicht der Analysis ganz egal ist, welche Norm man auf dem Vektorraum betrachtet. Bei unendlich-dimensionalen Vektorräumen ist das nicht mehr der Fall!

Nachdem wir nun einiges über Normen im allgemeinen gehört haben, wollen wir einen speziellen Typ von Normen kennenlernen, nämlich die Normen, die von einem Skalarprodukt herrühren.

Satz 1.9
Es sei $\big(V, \langle\cdot,\cdot\rangle\big)$ *ein euklidischer oder unitärer Raum. Dann wird durch*

$$\|\cdot\| : V \to \mathbb{R}_{\geq 0} : x \mapsto \sqrt{\langle x, x\rangle}$$

eine Norm auf V *definiert.*

Beweis: Seien $x \in V$ und $\lambda \in K$. Aus der positiven Definitheit von $\langle\cdot,\cdot\rangle$ folgt, daß $\langle x, x\rangle \geq 0$ und somit $\|x\|$ definiert und stets nicht-negativ ist. Ferner folgt, daß $\|x\| = 0$ genau dann gilt, wenn x der Nullvektor ist. Aus der Bilinearität bzw. Sesquilinearität von $\langle\cdot,\cdot\rangle$ leiten wir her, daß

$$\langle \lambda x, \lambda x\rangle = \lambda\overline{\lambda}\langle x, x\rangle = |\lambda|^2\langle x, x\rangle,$$

also $\|\lambda x\| = |\lambda| \cdot \|x\|$.

Allein, die Dreiecksungleichung ist etwas schwieriger zu zeigen. Wir verwenden hierfür die Cauchy-Schwarzsche Ungleichung aus Satz 1.10 und erhalten für $x, y \in V$ - beachte dabei, für eine komplexe Zahl $c \in \mathbb{C}$ ist $\Re(c) = \frac{1}{2}\big(c + \overline{c}\big)$ der Realteil von c -

$$\begin{aligned}
\|x+y\|^2 &= \langle x+y, x+y\rangle = \langle x, x\rangle + \langle x, y\rangle + \overline{\langle x, y\rangle} + \langle y, y\rangle \\
&= \|x\|^2 + 2\cdot\Re\big(\langle x, y\rangle\big) + \|y\|^2 \\
&\leq \|x\|^2 + 2\cdot\big|\langle x, y\rangle\big| + \|y\|^2 \\
&\leq \|x\|^2 + 2\cdot\|x\|\cdot\|y\| + \|y\|^2 \\
&= \big(\|x\| + \|y\|\big)^2.
\end{aligned}$$

Hieraus folgt dann die Dreiecksungleichung. □

Es bleibt also, die Cauchy-Schwarzsche Ungleichung zu zeigen.
Satz 1.10 (Cauchy-Schwarzsche Ungleichung)
Ist $\big(V, \langle\cdot,\cdot\rangle\big)$ *ein euklidischer oder unitärer Raum, dann gilt für alle* $x, y \in V$

$$\big|\langle x, y\rangle\big| \leq \|x\| \cdot \|y\|, \tag{54}$$

zudem gilt die Gleichheit genau dann, wenn x *und* y *linear abhängig sind.*

Beweis: Es ist $K = \mathbb{R}$ oder $K = \mathbb{C}$.

Für $x = 0$ oder $y = 0$ ist die Aussage offensichtlich richtig. Wir können also $x, y \neq 0$ annehmen. Dann gilt für $\lambda \in K$

$$0 \leq \langle x - \lambda y, x - \lambda y\rangle = \langle x, x\rangle - \overline{\lambda}\langle x, y\rangle - \lambda\overline{\langle x, y\rangle} + \lambda\overline{\lambda}\langle y, y\rangle. \tag{55}$$

Wählen wir nun speziell $\lambda = \frac{\langle x,y\rangle}{\langle y,y\rangle} \in K$, dann folgt

$$\begin{aligned}
0 &\leq \langle x, x\rangle - \overline{\tfrac{\langle x,y\rangle}{\langle y,y\rangle}}\langle x, y\rangle - \tfrac{\langle x,y\rangle}{\langle y,y\rangle}\overline{\langle x, y\rangle} + \tfrac{\langle x,y\rangle}{\langle y,y\rangle}\overline{\tfrac{\langle x,y\rangle}{\langle y,y\rangle}}\langle y, y\rangle \\
&= \langle x, x\rangle - \tfrac{|\langle x,y\rangle|^2}{\langle y,y\rangle} - \tfrac{|\langle x,y\rangle|^2}{\langle y,y\rangle} + \tfrac{|\langle x,y\rangle|^2}{\langle y,y\rangle},
\end{aligned}$$

also

$$\big|\langle x, y\rangle\big|^2 \leq \langle x, x\rangle \cdot \langle y, y\rangle. \tag{56}$$

Durch Ziehen der positiven Wurzel folgt die gesuchte Ungleichung (54).

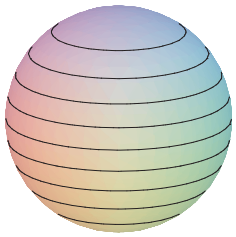
Nun sind x und y genau dann linear abhängig, wenn es ein $\lambda \in K$ gibt, für das $x = \lambda y$ gilt. Das wiederum ist wegen der positiven Definitheit von $\langle \cdot, \cdot \rangle$ gleichbedeutend dazu, daß es ein $\lambda \in K$ gibt, für das in (55) das Gleichheitszeichen gilt. Dieses λ ist eindeutig bestimmt, und aus (55) erhalten wir unmittelbar $\lambda = \frac{\langle x,y \rangle}{\langle y,y \rangle}$, und damit ist die Gleichheit in (55) gleichwertig zur Gleichheit in (56). □

Beispiel 1.11 a. Ist $\langle \cdot, \cdot \rangle$ das kanonische Skalarprodukt auf $\mathbb{R}^n$ bzw. $\mathbb{C}^n$, dann bezeichnen wir die zugehörige Norm $\| \cdot \|_2$ als die euklidische Norm auf $\mathbb{R}^n$ bzw. $\mathbb{C}^n$.

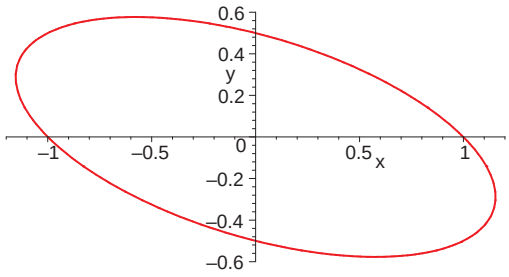
b. Es ist lehrreich, für eine Norm auf dem $\mathbb{R}^n$ die *Sphäre* $S^{n-1}_{r,\|\cdot\|}$ vom Radius $r \geq 0$ bezüglich der gegebenen Norm zu betrachten,

$$S^{n-1}_{r,\|\cdot\|} = \{x \in \mathbb{R}^n \mid \|x\| = r\}.$$

Ist $\| \cdot \|_2$ die euklidische Norm, die durch das Standardskalarprodukt auf dem $\mathbb{R}^n$ definiert wird, dann erhalten wir die Standardsphäre S^{n-1}_r:



Auf $\mathbb{R}^2$ liefert das Skalarprodukt b_A mit $A = \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix}$ aus Beispiel 1.5 etwa folgendes Bild für $S^1_{1,\|\cdot\|_A}$:



Die Cauchy-Schwarzsche Ungleichung erlaubt es uns nun, in einem beliebigen euklidischen Raum $(V, \langle \cdot, \cdot \rangle)$ Winkel zwischen zwei Vektoren zu definieren. Denn aus der Ungleichung (54) folgt für $x, y \in V$

$$-1 \leq \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|} \leq 1. \tag{57}$$

Wie schon oben bemerkt, wissen wir aus der Analysis, daß es zu jeder reellen Zahl $-1 \leq r \leq 1$ genau einen Winkel $\alpha \in [0, \pi]$ gibt mit $r = \cos(\alpha)$, nämlich $\alpha = \arccos(r)$.

Dies führt zu der folgenden Definition.

Definition 1.12

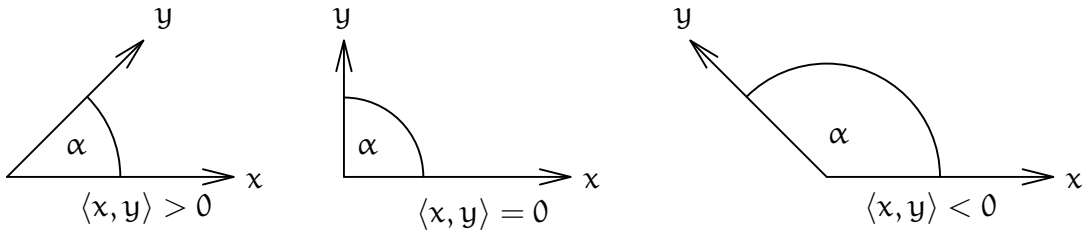
Sei $(V, \langle\cdot,\cdot\rangle)$ ein *euklidischer* Raum. Die eindeutig bestimmte Zahl $\alpha = \measuredangle(x,y) \in [0,\pi]$ mit

$$\cos(\alpha) = \frac{\langle x,y\rangle}{\|x\|\cdot\|y\|}$$

heißt *Winkel* zwischen x und y.

Ist $\langle x,y\rangle > 0$, also $\alpha \in \left[0, \frac{\pi}{2}\right[$, so spricht man von einem *spitzen Winkel*; ist $\langle x,y\rangle < 0$, also $\alpha \in \left]\frac{\pi}{2}, \pi\right]$, so spricht man von einem *stumpfen Winkel*; und im Fall von $\langle x,y\rangle = 0$, also $\alpha = \frac{\pi}{2}$ von einem *rechten Winkel*.

In letzterem Fall, wenn also $\langle x,y\rangle = 0$, so sagen wir auch x und y sind *orthogonal* zueinander und schreiben $x \perp y$.



"spitzer Winkel" "rechter Winkel" "stumpfer Winkel"

Wir wollen nun noch den Zusammenhang zwischen symmetrischen Bilinearformen bzw. hermiteschen Sesquilinearformen und Matrizen klären.

Definition 1.13

Es sei $\dim_K(V) = n < \infty$ und $b : V \times V \to K$ eine Bilinearform oder eine Sesquilinearform[3] auf V. Ist $B = (x_1, \ldots, x_n)$ eine Basis von V, dann ordnen wir b und B eine Matrix

$$M_B(b) = \big(b(x_i, x_j)\big)_{1\le i,j\le n} \in \mathrm{Mat}(n, K)$$

zu. Sie heißt die *darstellende Matrix* von b bezüglich der Basis B.

Wegen der Bilinearität bzw. Sesquilinearität von b ist b durch $M_B(b)$ eindeutig bestimmt, denn für $x = \sum_{i=1}^n \lambda_i x_i, y = \sum_{i=1}^n \mu_i x_i \in V$ gilt

$$b(x,y) = \sum_{i=1}^n \lambda_i b(x_i, y) = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \mu_j b(x_i, x_j)$$

bzw.

$$b(x,y) = \sum_{i=1}^n \lambda_i b(x_i, y) = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \overline{\mu_j} b(x_i, x_j).$$

Ferner ist b offenbar genau dann symmetrisch, wenn $b(x_i, x_j) = b(x_j, x_i)$ für alle i, j gilt, d. h. wenn $M_B(b)$ symmetrisch ist. Und b ist genau dann hermitesch, wenn $b(x_i, x_j) = \overline{b(x_j, x_i)}$ für alle i, j gilt, d. h. wenn $M_B(b)$ hermitesch ist.

[3] Was natürlich nur dann möglich ist, wenn $K = \mathbb{C}$!

Ist nun $K = \mathbb{R}$ und b symmetrisch bzw. $K = \mathbb{C}$ und b hermitesch, so ist nach Definition 1.4 und unter Berücksichtigung von Bemerkung 1.15 b genau dann positiv (bzw. negativ) definit, wenn $M_B(b)$ positiv (bzw. negativ) definit ist.

Beachten wir, daß für einen Vektor $x = \sum_{i=1}^{n} \lambda_i x_i$ mit $M_B(x) = (\lambda_1, \ldots, \lambda_n)^t$ der Koordinatenvektor von x bezüglich B bezeichnet wird, dann haben wir folgendes Lemma gezeigt.

Lemma 1.14 a. *Ist* $B = (x_1, \ldots, x_n)$ *eine Basis von* V *und* $b : V \times V \to K$ *eine Bilinearform auf* V*, dann gilt für* $x, y \in V$

$$b(x, y) = M_B(x)^t \circ M_B(b) \circ M_B(y).$$

b *ist genau dann symmetrisch, wenn* $M_B(b)$ *eine symmetrische Matrix ist, d. h.* $M_B(b)^t = M_B(b)$.
Ist $K = \mathbb{R}$ *und* b *symmetrisch, so ist* b *genau dann positiv (bzw. negativ) definit, wenn* $M_B(b)$ *positiv (bzw. negativ) definit ist.*

b. *Ist* $B = (x_1, \ldots, x_n)$ *eine Basis des* $\mathbb{C}$*-Vektorraumes* V *und* $b : V \times V \to \mathbb{C}$ *eine Sesquilinearform auf* V*, dann gilt für* $x, y \in V$

$$b(x, y) = M_B(x)^t \circ M_B(b) \circ \overline{M_B(y)}.$$

b *ist genau dann hermitesch, wenn* $M_B(b)$ *eine hermitesche Matrix ist, d. h.* $M_B(b)^* = M_B(b)$.
Ist b *hermitesch, so ist* b *genau dann positiv (bzw. negativ) definit, wenn* $M_B(b)$ *positiv (bzw. negativ) definit ist.*

Bemerkung 1.15 a. Ist $V = K^n$, E die kanonische Basis von K^n, $b : K^n \times K^n \to K$ eine Bilinearform und $A \in \mathrm{Mat}(n, K)$. Dann gilt offenbar

$$b = b_{M_E(b)} \quad \text{und} \quad A = M_E(b_A).$$

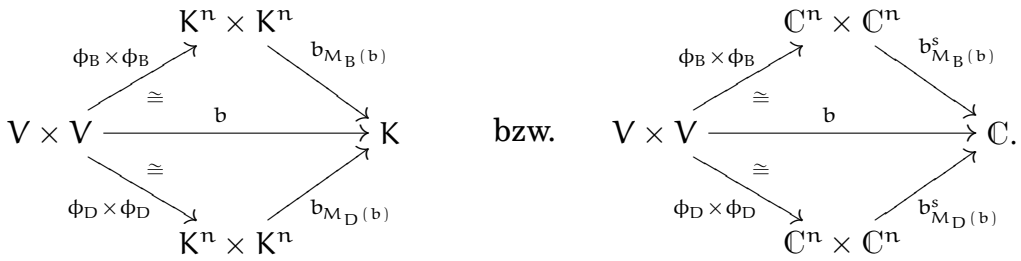
b. Entsprechend gilt, ist $V = \mathbb{C}^n$, E die kanonische Basis von $\mathbb{C}^n$, $b : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ eine Sesquilinearform und $A \in \mathrm{Mat}(n, \mathbb{C})$. Dann gilt offenbar

$$b = b^s_{M_E(b)} \quad \text{und} \quad A = M_E\big(b^s_A\big).$$

c. Die Aussage des Lemmas läßt sich mit Hilfe der Karte $\phi_B : V \to K^n : x_i \mapsto e_i$ auch durch folgende kommutative Diagramme darstellen:

Wir wollen jetzt untersuchen, wie sich die darstellende Matrix einer Bilinearform bzw. einer Sesquilinearform b unter Basiswechsel verhält.

Dazu betrachten wir folgende kommutative Diagramme, wobei B und D Basen von V seien, $\dim_K(V) = n < \infty$, und im Falle von Sesquilinearformen $K = \mathbb{C}$:



Sei nun $T_D^B = M_D^B(\mathrm{id}_V) = M_E^E\big(\phi_D \circ \phi_B^{-1}\big)$, wobei E die kanonische Basis von K^n ist, die Matrix, die den Basiswechsel in V beschreibt, dann folgt für $x, y \in K^n$ aus der Kommutativität des Diagramms

$$x^t M_B(b) y = \big(T_D^B x\big)^t M_D(b)\big(T_D^B y\big) = x^t\Big(\big(T_D^B\big)^t \circ M_D(b) \circ T_D^B\Big) y$$

bzw.

$$x^t M_B(b)\overline{y} = \big(T_D^B x\big)^t M_D(b)\overline{\big(T_D^B y\big)} = x^t\Big(\big(T_D^B\big)^t \circ M_D(b) \circ \overline{T_D^B}\Big)\overline{y},$$

also

$$M_B(b) = \big(T_D^B\big)^t \circ M_D(b) \circ T_D^B$$

bzw.

$$M_B(b) = \big(T_D^B\big)^t \circ M_D(b) \circ \overline{T_D^B},$$

wie man durch Einsetzen von e_i für x und e_j für y, $i, j = 1, \ldots, n$, sofort sieht.

Wir erhalten also den folgenden Satz.

Satz 1.16 a. *Sei* $\dim_K(V) = n < \infty$, B *und* D *zwei Basen von* V *und sei* b *eine Bilinearform auf* V, *dann gilt mit* $T_D^B \in \mathrm{Gl}_n(K)$

$$M_B(b) = \big(T_D^B\big)^t \circ M_D(b) \circ T_D^B.$$

b. *Sei* $\dim_{\mathbb{C}}(V) = n < \infty$, B *und* D *zwei Basen von* V *und sei* b *eine Sesquilinearform auf* V, *dann gilt mit* $T_D^B \in \mathrm{Gl}_n(\mathbb{C})$

$$M_B(b) = \big(T_D^B\big)^t \circ M_D(b) \circ \overline{T_D^B}.$$

Bemerkung 1.17 a. Es bleibt festzuhalten, daß sowohl Endomorphismen $f : V \to V$ als auch Bilinearformen $b : V \times V \to K$ sich nach Wahl einer Basis B durch Matrizen $M_B^B(f)$ bzw. $M_B(b)$ beschreiben lassen. Bei Basiswechsel, der durch die Matrix $T = T_D^B$ beschrieben wird, haben Endomorphismen und Bilinearformen aber ein unterschiedliches Transformationsverhalten. Es gilt:

$$M_B^B(f) = T^{-1} \circ M_D^D(f) \circ T \quad \text{und} \quad M_B(b) = T^t \circ M_D(b) \circ T.$$

b. Wir halten ferner fest, daß für $\dim_K(V) = n < \infty$ nach Wahl einer Basis B von V die Abbildung

$$\mathrm{Bil}_K(V) \to \mathrm{Mat}(n, K) : b \mapsto M_B(b)$$

eine Bijektion ist.
Dabei werden die symmetrischen Bilinearformen auf die symmetrischen Matrizen abgebildet.

Ist $K = \mathbb{R}$, dann gilt sogar, daß die Bijektion eine Bijektion zwischen der Menge der Skalarprodukte und der Menge der positiv definiten symmetrischen Matrizen liefert.
Die analoge Aussage für Sesquilinearformen gilt ebenfalls.

Wir wollen abschließend im euklidischen Fall noch einmal das Verhältnis zwischen einem Skalarprodukt und der zugehörigen Norm betrachten, oder allgemeiner zwischen einer symmetrischen Bilinearform und der zugehörigen *quadratischen Form*.

Definition 1.18
Ist $b \in \mathrm{Bil}_K(V)$ eine symmetrische Bilinearform auf V, dann nennen wir

$$q : V \to K : x \mapsto b(x, x)$$

die zu b gehörige *quadratische Form*.

Beispiel 1.19
Ist $A = (a_{ij}) \in \mathrm{Mat}(n, K)$ symmetrisch und $b = b_A$, dann gilt für $x = (x_1, \ldots, x_n)^t \in K^n$

$$q(x) = x^t A x = \sum_{i,j} a_{ij} x_i x_j = \sum_{i=1}^{n} a_{ii} x_i^2 + 2 \sum_{1 \le i < j \le n} a_{ij} x_i x_j.$$

Damit können wir q als ein homogenes quadratisches Polynom in den Unbestimmten $x_1, \ldots, x_n$ auffassen - vgl. Aufgabe III.1.55. Es gilt $q(\lambda x) = \lambda^2 q(x)$.

Auf den ersten Blick scheint es, daß die quadratische Form q weit weniger Information enthält, als die symmetrische Bilinearform b. Erstaunlicherweise kann man b jedoch aus q zurückgewinnen, wenn $\mathrm{char}(K) \neq 2$ ist.

Lemma 1.20
K sei ein Körper mit $\mathrm{char}(K) \neq 2$. *Ferner sei* $b \in \mathrm{Bil}_K(V)$ *eine symmetrische Bilinearform und* q *die zugehörige quadratische Form. Dann gilt für* $x, y \in V$

$$b(x, y) = \tfrac{1}{2}\big(q(x + y) - q(x) - q(y)\big).$$

Beweis: Die Aussage folgt durch einfaches Einsetzen der Definition von q in die rechte Seite. □

Analog führen wir die quadratische Form zu einer hermiteschen Sesquilinearform ein.

Definition 1.21
Ist $b : V \times V \to \mathbb{C}$ eine hermitesche Sesquilinearform auf V, dann nennen wir

$$q : V \to \mathbb{R} : x \mapsto b(x, x)$$

die zu b gehörige *quadratische Form*.

Und auch hier gilt, daß die quadratische Form die hermitesche Sesquilinearform bereits bestimmt.

Lemma 1.22
Sei b *eine hermitesche Sesquilinearform auf* V *und* q *die zugehörige quadratische Form. Dann gilt für* $x, y \in V$

$$b(x,y) = \tfrac{1}{4}\big(q(x+y) - q(x-y) + iq(x+iy) - iq(x-iy)\big).$$

Beweis: Die Aussage folgt wiederum durch einfaches Einsetzen der Definition von q in die rechte Seite. □

Die Tatsache, daß eine symmetrische Bilinearform bzw. eine hermitesche Sesquilinearform durch die quadratische Form schon eindeutig bestimmt ist, nennt man *Polarisierung*.

AUFGABEN

Aufgabe 1.23
Es sei $\big(V, \langle\cdot,\cdot\rangle\big)$ ein euklidischer oder unitärer Raum. Ferner sei $\|\cdot\|$ die durch das Skalarprodukt definierte Norm. Zeige für $x, y \in V$ gilt:

a. $\|x+y\|^2 + \|x-y\|^2 = 2\big(\|x\|^2 + \|y\|^2\big)$ ("Parallelogramm-Gleichung")
b. Aus $x \perp y$ folgt $\|x\|^2 + \|y\|^2 = \|x+y\|^2$. ("Satz des Pythagoras")

Aufgabe 1.24
Es sei $V = U \oplus U'$ ein K-Vektorraum mit $K = \mathbb{R}$ oder $K = \mathbb{C}$ und $\|\cdot\|_U$ bzw. $\|\cdot\|_{U'}$ seien Normen auf U bzw. U'. Zeige, durch $\|\cdot\| : V \to \mathbb{R}_{\geq 0} : v = u + u' \mapsto \|u\|_U + \|u'\|_{U'}$ wird V zu einem normierten Raum.

Aufgabe 1.25
Es sei $V = U \oplus U'$ ein K-Vektorraum mit $K = \mathbb{R}$ oder $K = \mathbb{C}$ und $\|\cdot\|$ eine Norm auf V. Wird dann durch $|\cdot| : V \to \mathbb{R}_{\geq 0} : v = u + u' \mapsto \|u\| + \|u'\|$ stets erneut eine Norm auf V definiert?

Aufgabe 1.26 (Halbnorm)
Es sei V ein K-Vektorraum mit $K = \mathbb{R}$ oder $K = \mathbb{C}$ und $\alpha : V \to \mathbb{R}$ mit

(i) $\alpha(\lambda x) = |\lambda|\alpha(x)$ für alle $\lambda \in K$ und $x \in V$.
(ii) $\alpha(x_1 + x_2) \leq \alpha(x_1) + \alpha(x_2)$ für alle $x_1, x_2 \in V$.

Zeige, dann gelten außerdem:

a. $\alpha(x) \geq 0$ für alle $x \in V$.
b. $U := \{x \in V \mid \alpha(x) = 0\}$ ist ein Unterraum von V.
c. Durch $\|\cdot\| : V/U \to \mathbb{R}_{\geq 0} : [x] \mapsto \alpha(x)$ wird eine Norm auf V/U definiert.

Aufgabe 1.27 (Äquivalenz von Normen)
Es sei $K = \mathbb{R}$ oder $K = \mathbb{C}$ und V ein K-Vektorraum. Zeige:

a. Die Äquivalenz von Normen ist eine Äquivalenzrelation auf der Menge der Normen auf V.

b. Je zwei Normen $\|\cdot\|$ und $|\cdot|$ auf K^n sind äquivalent.
c. Ist $\big(W, \|\cdot\|\big)$ ein normierter K-Vektorraum und $f \in \mathrm{Hom}_K(V, W)$ injektiv, dann wird durch $|\cdot|_f : V \to \mathbb{R}_{\geq 0} : v \mapsto \|f(v)\|$ eine Norm auf V definiert.
d. Folgere aus b. und c., daß je zwei Normen auf V äquivalent sind, falls $\dim_K(V) < \infty$.
e. Ist nun $V = W$, $\dim_K(V) < \infty$ und f bijektiv, dann zeige $\|\cdot\| \sim |\cdot|_f$ und bestimme optimale Konstanten $m, M \in \mathbb{R}_{>0}$.

Hinweis: In b. reicht es, zu zeigen, daß jede Norm $\|\cdot\|$ äquivalent zu $\|\cdot\|_2$ ist. Man zeige zunächst die Existenz der oberen Schranke M mit $\|\cdot\| \leq M \cdot \|\cdot\|_2$ und folgere daraus, daß die Abbildung $\|\cdot\| : \big(K^n, \|\cdot\|_2\big) \to \big(\mathbb{R}, |\cdot|\big)$ stetig ist. Um die untere Schranke m mit $m \cdot \|\cdot\|_2 \leq \|\cdot\|$ zu finden, verwende man sodann aus der Analysis, daß $S^{n-1} = \big\{x \in K^n \;\big|\; \|x\|_2 = 1\big\}$ kompakt in $\big(K^n, \|\cdot\|_2\big)$ ist und daß stetige Funktionen auf einem Kompaktum ihr Minimum annehmen.

Aufgabe 1.28 (p-Norm und Maximumsnorm)
Es sei $K = \mathbb{R}$ oder $K = \mathbb{C}$. Zeige:

a. Für $0 \neq p \in \mathbb{N}$ wird durch $|x|_p := \big(\sum_{i=1}^n |x_i|^p\big)^{\frac{1}{p}}$ für $x \in K^n$ auf K^n eine Norm definiert.
b. Durch $|x|_\infty := \max\{|x_i| \;|\; i = 1, \ldots, n\}$ für $x \in K^n$ wird auf K^n eine Norm definiert.
c. Finde für $\|\cdot\|_1$, $\|\cdot\|_2$ und $\|\cdot\|_\infty$ paarweise Äquivalenzkonstanten.

Aufgabe 1.29 (Euklidische Norm)
Es sei $K = \mathbb{R}$ oder $K = \mathbb{C}$. Zeige:

a. Durch $\|A\|_2 := \sum_{i,j=1}^n \sqrt{|a_{ij}|^2}$ für $A = (a_{ij}) \in \mathrm{Mat}(n, K)$ wird $\mathrm{Mat}(n, K)$ eine Norm definiert, die sogenannte *euklidische Norm*.
b. Für $A, B \in \mathrm{Mat}(n, K)$ und $x \in K^n$ gelten:
   (i) $\|Ax\|_2 \leq \|A\|_2 \cdot \|x\|_2$.
   (ii) $\|A \circ B\|_2 \leq \|A\|_2 \cdot \|B\|_2$.
c. Durch $\|A\|_z = \max\big\{\sum_{j=1}^n |a_{ij}| \;\big|\; i = 1, \ldots n\big\}$ für $A = (a_{ij}) \in \mathrm{Mat}(n, K)$ wird $\mathrm{Mat}(n, K)$ eine Norm definiert, die sogenannte *Zeilensummennorm*.
d. Durch $\|A\|_s = \max\big\{\sum_{i=1}^n |a_{ij}| \;\big|\; j = 1, \ldots n\big\}$ für $A = (a_{ij}) \in \mathrm{Mat}(n, K)$ wird $\mathrm{Mat}(n, K)$ eine Norm definiert, die sogenannte *Spaltensummennorm*.

Hinweis: Man verwende die Ungleichung von Cauchy-Schwarz.

Aufgabe 1.30 (l^1-Norm)
Es sei $V = \big\{(a_i)_{i \in \mathbb{N}} \in \mathbb{K}^{\mathbb{N}} \;\big|\; \sum_{i=0}^\infty |a_i| < \infty\big\}$.

a. Zeige, daß durch $\|\cdot\| : V \to \mathbb{R}_{\geq 0} : (a_i) \mapsto \sum_{i=0}^\infty |a_i|$ eine Norm auf V definiert wird.
b. Berechne die Operatornorm der linearen Abbildung $T : \big(V, \|\cdot\|\big) \to \big(V, \|\cdot\|\big) : (a_0, a_1, a_2, \ldots) \to (0, a_0, a_1, \ldots)$.

Aufgabe 1.31 (L^1-Norm)
Es sei $V = \mathcal{C}^0\big([0, 1]\big)$ der Vektorraum der auf dem Intervall $[0, 1]$ stetigen Funktionen. Zeige, daß durch $\|\cdot\| : V \to \mathbb{R}_{\geq 0} : f \mapsto \int_0^1 |f(t)| dt$ eine Norm auf V definiert wird.

Aufgabe 1.32 (Exponentialabbildung)
Es sei $|\cdot| : \mathbb{R}^n \to \mathbb{R}$ eine beliebige Norm und es sei

$$\|\cdot\| : \mathrm{Mat}(n,\mathbb{R}) \to \mathbb{R}_{\geq 0} : A \mapsto \sup_{0\neq x\in\mathbb{R}^n} \frac{|Ax|}{|x|}$$

die *Operatornorm* auf $\mathrm{Mat}(n,\mathbb{R})$. Es seien $A, B \in \mathrm{Mat}(n,\mathbb{R})$ und $T \in \mathrm{Gl}_n(\mathbb{R})$. Zeige:

a. $\|\cdot\|$ ist eine Norm auf $\mathrm{Mat}(n,\mathbb{R})$.
b. $\|A \circ B\| \leq \|A\| \cdot \|B\|$.
c. Die Reihe $\sum_{n=0}^{\infty} \frac{1}{n!} A^n$ ist absolut konvergent bezüglich $\|\cdot\|$.
   Insbesondere ist $e^A := \sum_{n=0}^{\infty} \frac{1}{n!} A^n \in \mathrm{Mat}(n,\mathbb{R})$.
d. Die Abbildung $\alpha : \mathrm{Mat}(n,\mathbb{R}) \to \mathrm{Mat}(n,\mathbb{R}) : X \mapsto T \circ X \circ T^{-1}$ ist stetig bez. $\|\cdot\|$.
e. $T \circ e^A \circ T^{-1} = e^{T\circ A\circ T^{-1}}$.
f. Ist $A = J(\lambda, n)$ mit $\lambda \in \mathbb{R}$, dann ist $e^A = N(\lambda, n)$ und $J_{e^A} = J\big(e^\lambda, n\big)$, wobei
   $N(\lambda, n) = (n_{ij})$ mit $n_{ij} = \begin{cases} 0, & i > j \\ \frac{e^\lambda}{k!}, & i = j-k, k = 0, \ldots, n-1. \end{cases}$
g. Ist $J_A = \bigoplus_{i=1}^{r} \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} J(\lambda_i, j)$, so ist $J_{e^A} = \bigoplus_{i=1}^{r} \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} J\big(e^{\lambda_i}, j\big)$.

Hinweise: Man beachte, daß in einem Hilbertraum aus absoluter Konvergenz einer Reihe die Konvergenz der Reihe folgt. – In Teil e. betrachte man die Folgen von Partialsummen, die die Linke bzw. rechte Seite definieren, und verwende Teil d. Für f. zerlege man A in den nilpotenten Anteil und den Diagonalanteil und verwende im Beweis den Umordnungssatz für absolut konvergente Reihen. Der Beweis von Teil g. ist dann ein Zweizeiler.

Aufgabe 1.33
Es seien $\big(V, \|\cdot\|_V\big)$ und $\big(W, \|\cdot\|_W\big)$ zwei normierte $\mathbb{R}$-Vektorräume. Wir definieren eine Abbildung

$$\|\cdot\| : \mathrm{Hom}_{\mathbb{R}}(V,W) \to \mathbb{R}_{\geq 0} \cup \{\infty\} : f \mapsto \sup_{0\neq x\in V} \frac{\|f(x)\|_W}{\|x\|_V}.$$

a. Zeige, $\|\cdot\|$ ist eine Norm auf $H := \big\{ f \in \mathrm{Hom}_{\mathbb{R}}(V,W) \;\big|\; \|f\| < \infty \big\} \subseteq \mathrm{Hom}_{\mathbb{R}}(V,W)$.
b. Zeige, jede lineare Abbildung $f \in H$ ist stetig (bezüglich der Normen $\|\cdot\|_V$ und $\|\cdot\|_W$).
c. Es sei nun $\big(W, \|\cdot\|_W\big) = \big(V, \|\cdot\|_V\big)$. Zeige:
   (i) Ist V vollständig bezüglich $\|\cdot\|_V$, so ist die Reihe $e^f := \sum_{n=0}^{\infty} \frac{1}{n!} f^n$ konvergent in $\mathrm{End}_{\mathbb{R}}(V)$.
   (ii) Für $g \in \mathrm{End}_{\mathbb{R}}(V)$ sind die Abbildungen

   $$L_g : \mathrm{End}_{\mathbb{R}}(V) \to \mathrm{End}_{\mathbb{R}}(V) : f \mapsto g \circ f$$

   und

   $$R_g : \mathrm{End}_{\mathbb{R}}(V) \to \mathrm{End}_{\mathbb{R}}(V) : f \mapsto f \circ g$$

   stetig bezüglich der obigen Norm.

2. Orthogonale und unitäre Abbildungen

Mit Hilfe eines Skalarproduktes, das wir jetzt für reelle und komplexe Vektorräume V definiert haben, können wir den Begriff der Orthogonalität erklären. Zunächst betrachten wir die Orthogonalität von Vektoren und wollen zeigen, daß es zu jedem Skalarprodukt auf V eine orthogonale Basis gibt, d. h. eine Basis, so daß je zwei verschiedene Vektoren darin aufeinander senkrecht stehen. Normiert man die einzelnen Vektoren der Basis, so erhält man eine orthonormale Basis.

Die Herleitung der Existenz einer solchen Basis ist konstruktiv, und wir werden einen einfachen Algorithmus, nach Gram und Schmidt, dafür angeben.

Im Anschluß daran werden wir Abbildungen zwischen Vektorräumen mit Skalarprodukt untersuchen, die mit dem Skalarprodukt verträglich sind, die zusätzliche Struktur also ebenfalls erhalten, sogenannte orthogonale Abbildungen. Diese sind längen- und winkeltreu.

Generalvoraussetzung: In diesem Abschnitt verwenden wir $\mathbb{K}$ stets dann, wenn Aussagen sowohl für den Fall des Körpers $\mathbb{K} = \mathbb{R}$ der reellen Zahlen, als auch für den Fall des Körpers $\mathbb{K} = \mathbb{C}$ der komplexen Zahlen gelten sollen. Ferner sei $\big(V, \langle\cdot,\cdot\rangle\big)$ ein euklidischer Raum, falls $\mathbb{K} = \mathbb{R}$, und ein unitärer Raum, falls $\mathbb{K} = \mathbb{C}$. Zugleich versehen wir V stets mit der durch $\langle\cdot,\cdot\rangle$ definierten Norm, d. h. für $x \in V$ ist stets $\|x\| = \sqrt{\langle x, x\rangle}$. – V wird nicht notwendigerweise als endlich-dimensional vorausgesetzt.

Definition 2.1
Es seien $M, N \subseteq V$ zwei Teilmengen.

a. Wir sagen, M und N sind *orthogonal* zueinander, wenn $x \perp y$, d. h. $\langle x, y\rangle = 0$, für alle $x \in M$ und $y \in N$ gilt.
   Wir schreiben dann $M \perp N$. Ist $M = \{x\}$, so schreiben wir meist kurz $x \perp N$.
b. Wir nennen $M^\perp := \{x \in V \mid x \perp M\}$ das *orthogonale Komplement* von M.

Wir werden weiter unten sehen, daß in der Tat $U^\perp$ im endlich-dimensionalen Fall ein Komplement von U ist.

Lemma 2.2
Es sei $M \subseteq V$ eine Teilmenge von V und $U \subseteq V$ ein Unterraum von V.

a. *$M^\perp = \langle M\rangle^\perp$ ist ein Unterraum von V.*
b. *Sind $0 \neq x \in M$ und $0 \neq y \in M^\perp$, so sind x und y linear unabhängig.*

Beweis: a. Wegen $0 \in M^\perp$ ist $M^\perp \neq \emptyset$. Sind $x, y \in M^\perp$ und $\lambda, \mu \in \mathbb{K}$, so gilt für $z \in M$

$$\langle \lambda x + \mu y, z\rangle = \lambda\langle x, z\rangle + \mu\langle y, z\rangle = 0,$$

Also $\lambda x + \mu y \in M^\perp$. Damit ist $M^\perp$ ein Unterraum von V.

Da $M \subseteq \langle M\rangle$, gilt offenbar $M^\perp \supseteq \langle M\rangle^\perp$. Seien deshalb $x \in M^\perp$ und $y = \sum_{i=1}^n \lambda_i x_i \in \langle M\rangle$ mit $x_i \in M$ und $\lambda_i \in \mathbb{K}$ beliebig. Dann gilt

$$\langle x, y\rangle = \sum_{i=1}^{n} \lambda_i \langle x, x_i\rangle = 0,$$

also auch $x \in \langle M\rangle^\perp$.

b. Seien $0 \neq x \in M$ und $0 \neq y \in M^\perp$ gegeben, und seien $\lambda, \mu \in \mathbb{K}$ mit $\lambda x + \mu y = 0$. Dann folgt

$$0 = \langle \lambda x + \mu y, x\rangle = \lambda \langle x, x\rangle.$$

Wegen $\langle x, x\rangle \neq 0$ ist also $\lambda = 0$. Dann ist aber $0 = \lambda x + \mu y = \mu y$ und, wegen $y \neq 0$, ist auch $\mu = 0$.

□

Definition 2.3 a. Eine Familie $F = (x_i \mid i \in I)$ von Vektoren aus V heißt *orthogonal*, falls $x_i \perp x_j$ für alle $i, j \in I$ mit $i \neq j$ gilt.
Falls zusätzlich $\|x_i\| = 1$ für alle $i \in I$ gilt, so heißt F *orthonormal*.

b. Es sei B eine Basis von V. Ist B eine orthogonale Familie, so nennen wir B eine *Orthogonalbasis*, kurz *OGB*.
Entsprechend heißt B eine *Orthonormalbasis*, kurz *ONB*, von V, wenn die Familie B orthonormal ist.

c. Eine direkte Summe $V = \bigoplus_{i\in I} U_i$ heißt *orthogonal*, wenn $U_i \perp U_j$ für alle $i, j \in I$ mit $i \neq j$ gilt.
Wir schreiben dann auch $V = \perp_{i\in I} U_i$.

Beispiel 2.4
Betrachten wir $\mathbb{K}^n$ mit dem kanonischen Skalarprodukt, dann ist die kanonische Basis $E = (e_1, \ldots, e_n)$ offenbar eine Orthonormalbasis von $\mathbb{K}^n$, da $\langle e_i, e_j\rangle = \delta_{ij}$ für $i, j \in \{1, \ldots, n\}$.

Lemma 2.5
Es sei $F = (x_i \mid i \in I)$ *eine orthogonale Familie in* $V \setminus \{0\}$.

a. $\left(\frac{1}{\|x_i\|} \cdot x_i \;\middle|\; i \in I\right)$ *ist orthonormal.*

b. F *ist linear unabhängig.*

c. *Ist* F *eine ONB von* V*, so gilt für* $x \in V$

$$x = \sum_{i\in I}{}' \langle x, x_i\rangle x_i, \tag{58}$$

insbesondere sind nur endlich viele $\langle x, x_i\rangle$*,* $i \in I$*, ungleich null.*

Beweis: a. Hier ist nichts zu zeigen.

b. Aus $\sum_{i\in I}{}' \lambda_i x_i = 0$ folgt für jedes $i \in I$

$$0 = \langle 0, x_i\rangle = \sum_{j\in I}{}' \lambda_j \langle x_j, x_i\rangle = \lambda_i \langle x_i, x_i\rangle.$$

Da $\langle x_i, x_i\rangle \neq 0$, muß also $\lambda_i = 0$ gelten.

c. Da die Darstellung $x = \sum'_{i\in I} \lambda_i x_i$ von x als endliche Linearkombination von F eindeutig ist, folgt die Behauptung aus

$$\langle x, x_i\rangle = \sum_{j\in I}{}' \lambda_j \langle x_j, x_i\rangle = \lambda_i \langle x_i, x_i\rangle = \lambda_i, \quad i \in I.$$

□

Bemerkung 2.6
Ist B eine ONB von V, so erlaubt es die Gleichung (58), einen Vektor aus V als Linearkombination von B darzustellen, ohne hierzu eigens ein LGS lösen zu müssen, durch simples Einsetzen der Vektoren in das Skalarprodukt. Dieses Verfahren ist sehr effizient und von hoher praktischer Bedeutung.

Die Tatsache, daß sich die Koordinaten eines Vektors bezüglich einer ONB mit Hilfe des Skalarproduktes so einfach ausdrücken lassen spielt aber auch in vielen Beweisen eine Rolle, und ist somit ebenfalls für die Theorie von Bedeutung.

Wir beweisen jetzt, daß jeder endlich-dimensionale Hilbertraum eine ONB besitzt. Etwas allgemeiner gilt der folgende Satz.
Satz 2.7 (Gram-Schmidt)
Sei V ein endlich-dimensionaler Hilbertraum und $U \subseteq V$ ein Unterraum von V, dann läßt sich jede ONB von U zu einer ONB von V ergänzen.

Mit $U = \{0\}$ folgt das folgende Korollar.
Korollar 2.8
Jeder endlich-dimensionale Hilbertraum besitzt eine ONB.

Der Beweis von Satz 2.7 ist konstruktiv und wird auch das *Gram-Schmidtsche Orthonormalisierungsverfahren* genannt.

Beweis von Satz 2.7: Wir beweisen den Satz durch Induktion über $m := \dim_{\mathbb{K}}(V) - \dim_{\mathbb{K}}(U)$. Sei $n = \dim_{\mathbb{K}}(V)$, $r = \dim_{\mathbb{K}}(U)$ und $(y_1, \ldots, y_r)$ eine ONB von U.

Falls $m = 0$, so ist nichts zu zeigen. Sei also $m > 0$, d. h. $n > r$. Ergänze B' durch $(x_{r+1}, \ldots, x_n)$ zu einer Basis von V. Wir setzen nun

$$\tilde{y}_{r+1} := x_{r+1} - \sum_{i=1}^{r} \langle x_{r+1}, y_i\rangle y_i. \tag{59}$$

Da $x_{r+1} \notin U$, ist $\tilde{y}_{r+1} \neq 0$, und wir können deshalb

$$y_{r+1} := \tfrac{1}{\|\tilde{y}_{r+1}\|} \cdot \tilde{y}_{r+1} \tag{60}$$

setzen. Dann ist $\|y_{r+1}\| = 1$ und außerdem gilt für $i = 1, \ldots, r$

$$\begin{aligned}
\langle y_{r+1}, y_i\rangle &= \tfrac{1}{\|\tilde{y}_{r+1}\|}\langle \tilde{y}_{r+1}, y_i\rangle \\
&= \tfrac{1}{\|\tilde{y}_{r+1}\|}\Big(\langle x_{r+1}, y_i\rangle - \textstyle\sum_{j=1}^{r}\langle x_{r+1}, y_j\rangle\langle y_j, y_i\rangle\Big) \\
&= \tfrac{1}{\|\tilde{y}_{r+1}\|}\big(\langle x_{r+1}, y_i\rangle - \langle x_{r+1}, y_i\rangle\big) = 0.
\end{aligned}$$

Dann ist aber $(y_1, \dots, y_{r+1})$ eine ONB von $U' := \langle U, x_{r+1}\rangle = \langle U, y_{r+1}\rangle$, und wir sind fertig mit Induktion. □

Algorithmus 2.9 (Gram-Schmidtsches Orthonormalisierungsverfahren)
INPUT: $M \subseteq \mathbb{K}^n$ und ein Skalarprodukt $\langle\cdot,\cdot\rangle$ auf $\mathbb{K}^2$

OUTPUT: ONB B von $\langle M\rangle$

1. **Schritt:** Bestimme eine Basis $B' = (x_1, \dots, x_r)$ von $\langle M\rangle$, z. B. mittels Algorithmus II.6.3.
2. **Schritt:** Für $i = 1, \dots, r$ führe man folgende Schritte aus:
   Schritt a.: berechne die Summe $x_i' = x_i - \sum_{j=1}^{i-1}\langle x_i, x_j\rangle x_j$;
   Schritt b.: berechne $x_i'' = \frac{1}{\|x_i'\|}\cdot x_i'$;
   Schritt c.: und ersetze x_i in B durch x_i''.
3. **Schritt:** Gib die veränderte Basis B zurück.

Bemerkung 2.10 a. Um zu normieren, ist in obigem Algorithmus das Ziehen von Wurzeln notwendig. Verzichtet man jedoch auf die Normierung der Vektoren, so kommt man ohne Wurzelziehen aus. Läßt man im obigen Algorithmus Schritt 2.b. weg und ersetzt dafür in Schritt 2.a. die rechte Seite der Gleichung durch

$$x_i - \sum_{j=1}^{i-1}\frac{\langle x_i, x_j'\rangle}{\langle x_j', x_j'\rangle}x_j',$$

dann liefert der obige Algorithmus eine Orthogonalbasis von $\langle M\rangle$. Das hat den Vorteil, daß man exakt rechnen kann - etwa in Singular, wenn die Eingabedaten rationale Zahlen waren.

b. Will man in der Praxis ein Skalarprodukt übergeben, so wird man im reellen Fall eine symmetrische Matrix übergeben und im komplexen Fall eine hermitesche. Das Skalarprodukt wird dann gemäß Beispiel 1.5 a. bzw. b. gebildet.

Beispiel 2.11
Es sei $B = (x_1, x_2, x_3) = \{(1,0,1)^t, (1,1,1)^t, (0,0,4)^t\} \subseteq \mathbb{R}^3$, wobei wir $\mathbb{R}^3$ mit dem kanonischen Skalarprodukt versehen betrachten. Man sieht leicht, daß B bereits eine Basis von $\mathbb{R}^3$ ist.

Wir wollen hier B zunächst in eine OGB von $\mathbb{R}^3$ überführen und anschließend normieren.

Wir setzen nun $x_1' := (1,0,1)^t$, dann ist $\langle x_1', x_1'\rangle = 2$.

Im nächsten Schritt setzen wir

$$x_2' = x_2 - \frac{\langle x_2, x_1'\rangle}{\langle x_1', x_1'\rangle}x_1' = (1,1,1)^t - \tfrac{2}{2}(1,0,1)^t = (0,1,0)^t,$$

dann ist $\langle x_2', x_2'\rangle = 1$.

Schließlich bilden wir

$$x_3' = x_3 - \frac{\langle x_3, x_1'\rangle}{\langle x_1', x_1'\rangle}x_1' - \frac{\langle x_3, x_2'\rangle}{\langle x_2', x_2'\rangle}x_2' = (-2,0,2)^t,$$

und erhalten $\langle x_3', x_3' \rangle = 8$.

Damit ist $B' = \big((1,0,1)^t, (0,1,0)^t, (-2,0,2)^t\big)$ eine OGB von $\mathbb{R}^3$, und nach Normierung ist

$$B'' = \Big(\big(\tfrac{1}{\sqrt{2}}, 0, \tfrac{1}{\sqrt{2}}\big)^t, (0,1,0)^t, \big(-\tfrac{1}{\sqrt{2}}, 0, \tfrac{1}{\sqrt{2}}\big)^t\Big)$$

die ONB, die aus B durch das Gram-Schmidtsche Orthonormalisierungsverfahren entsteht.

Definition 2.12

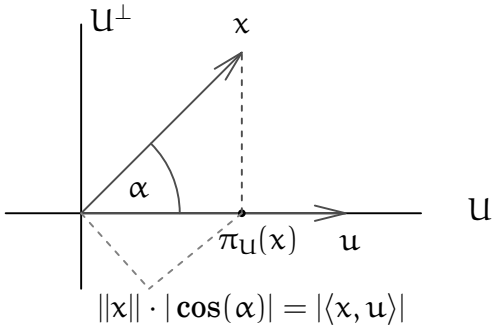
Es gelte $V = U \perp U^\perp$. Da sich jeder Vektor $x \in V$ in eindeutiger Weise darstellen läßt als $x = u + u'$ mit $u \in U$ und $u' \in U^\perp$, können wir eine Abbildung

$$\pi_U : V \to V$$

definieren durch $\pi\big(u + u'\big) = u$ für $u \in U$ und $u' \in U^\perp$.

Diese nennen wir die *orthogonale Projektion* von V auf U.

Ist $u \in U$ mit $\|u\| = 1$, so nennen wir $\big|\langle x, u\rangle\big|$ den *Betrag der Projektion von* x *auf* $\langle u\rangle$.



Lemma 2.13

Es sei $U \subseteq V$ *ein Unterraum mit* $V = U \perp U^\perp$.

a. $\pi_U \in \mathrm{End}_K(V)$ *ist eine Projektion mit* $\mathrm{Ker}(\pi_U) = U^\perp$ *und* $\mathrm{Im}(\pi_U) = U$.
b. *Ist* $\pi \in \mathrm{End}_K(V)$ *eine Projektion mit* $\mathrm{Ker}(\pi) = U^\perp$ *und* $\mathrm{Im}(\pi) = U$*, dann ist* $\pi = \pi_U$.

Beweis: a. Die Aussagen sind aus der Definition unmittelbar klar.

b. Sei also ein solches π gegeben. Dann gilt für $x = u + u' \in V$ mit $u \in U = \mathrm{Im}(\pi)$ und $u' \in U^\perp = \mathrm{Ker}(\pi)$ sowie $y \in V$ mit $u = \pi(y)$

$$\pi(x) = \pi(u) + \pi\big(u'\big) = \pi(u) = \pi^2(y) = \pi(y) = u = \pi_U(x).$$

□

Proposition 2.14

Es sei $\dim_K(V) < \infty$ *und* $U \subseteq V$ *ein Unterraum.*

a. $V = U \perp U^\perp$*, d. h.* $U \cap U^\perp = \{0\}$ *und* $V = U + U^\perp$.
b. *Ist* $(x_1, \ldots, x_r)$ *eine ONB von* U*, dann gilt*

$$\pi_U(x) = \sum_{i=1}^{r} \langle x, x_i\rangle x_i.$$

Beweis: a. Für $x \in U \cap U^\perp$ folgt aus Lemma 2.2 b., daß $x = 0$. Es bleibt also zu zeigen, daß $V = U + U^\perp$.
Aber, ist $(x_1, \dots, x_r)$ eine ONB von U, so läßt sich diese nach Satz 2.7 zu einer ONB $(x_1, \dots, x_n)$ von V ergänzen. Insbesondere gilt dann aber

$$V = \langle x_1, \dots, x_r \rangle + \langle x_{r+1}, \dots, x_n \rangle \subseteq U + U^\perp.$$

b. Ergänzen wir wie in a. $(x_1, \dots, x_r)$ zu einer ONB $(x_1, \dots, x_n)$ von V, so gilt für $x = \sum_{i=1}^{n} \lambda_i x_i \in V$ einerseits nach der Definition von π_U

$$\pi_U(x) = \sum_{i=1}^{n} \lambda_i \pi_U(x_i) = \sum_{i=1}^{r} \lambda_i x_i.$$

Andererseits folgt aus der Gleichung (58) in Lemma 2.5 $\lambda_i = \langle x, x_i \rangle$ für alle $i = 1, \dots, n$.

□

Wir kommen jetzt zu den strukturerhaltenden Abbildungen, d. h. zu solchen, die mit dem Skalarprodukt verträglich sind. Diese haben einen speziellen Namen.

Definition 2.15 a. Ist $\mathbb{K} = \mathbb{R}$, so heißt $f \in \mathrm{End}_{\mathbb{R}}(V)$ *orthogonal*, falls für alle $x, y \in V$ gilt

$$\langle x, y \rangle = \big\langle f(x), f(y) \big\rangle.$$

Wir nennen $O(V) := \{f \in \mathrm{End}_{\mathbb{R}}(V) \mid f \text{ ist orthogonal}\}$ die *orthogonale Gruppe* von V.

b. Ist $\mathbb{K} = \mathbb{C}$, so heißt $f \in \mathrm{End}_{\mathbb{C}}(V)$ *unitär*, falls für alle $x, y \in V$ gilt

$$\langle x, y \rangle = \big\langle f(x), f(y) \big\rangle.$$

Wir nennen $U(V) := \{f \in \mathrm{End}_{\mathbb{C}}(V) \mid f \text{ ist unitär}\}$ die *unitäre Gruppe* von V.

c. Wir nennen eine Matrix $A \in \mathrm{Mat}(n, \mathbb{R})$ *orthogonal*,[4] wenn $A \circ A^* = \mathbb{1}_n$.
$O(n) := \{A \in \mathrm{Mat}(n, \mathbb{R}) \mid A \text{ ist orthogonal}\}$ bezeichnen wir als *orthogonale Gruppe* vom Grad n.
$SO(n) := \{A \in O(n) \mid \det(A) = 1\}$ bezeichnen wir als *spezielle orthogonale Gruppe* vom Grad n.

d. Wir nennen eine Matrix $A \in \mathrm{Mat}(n, \mathbb{C})$ *unitär*, wenn $A \circ A^* = \mathbb{1}_n$.
$U(n) := \{A \in \mathrm{Mat}(n, \mathbb{C}) \mid A \text{ ist unitär}\}$ bezeichnen wir als *unitäre Gruppe* vom Grad n.
$SU(n) := \{A \in U(n) \mid \det(A) = 1\}$ bezeichnen wir als *spezielle unitäre Gruppe* vom Grad n.

[4] Da $A^* = A^t$ ist, kann man für beliebige Körper K definieren, daß $A \in \mathrm{Mat}(n, K)$ *orthogonal* heißt, falls $A \circ A^t = \mathbb{1}_n$ gilt. Wir setzen dann $O(n, K)$ als die Menge der orthogonalen Matrizen in $\mathrm{Mat}(n, K)$ und nennen sie die *orthogonale Gruppe* vom Grad n über K. Wie im reellen Fall, zeigt man auch im allgemeinen Fall, daß $O(n, K)$ eine Gruppe bezüglich der Matrixmultiplikation ist.

e. Wir nennen eine Matrix $A \in \mathrm{Mat}(n, \mathbb{K})$ *normal*, falls $A \circ A^* = A^* \circ A$. (Vgl. Aufgabe 2.36.)

Bemerkung 2.16
Man sieht unmittelbar aus der Definition, daß symmetrische, hermitesche, orthogonale und unitäre Matrizen auch normal sind.

Wir werden im Folgenden zeigen, daß die oben definierten *Gruppen* in der Tat Gruppen bezüglich der Matrixmultiplikation als Gruppenstruktur sind. Vgl. Korollar 2.21.

Lemma 2.17
Es seien $f, g \in \mathrm{End}_{\mathbb{K}}(V)$ *orthogonal bzw. unitär.*

a. f *erhält Längen, Abstände und (im orthogonalen Fall) Winkel, d. h. für* $x, y \in V$ *gilt*

$$\|f(x)\| = \|x\|, \quad \|f(x) - f(y)\| = \|x - y\| \quad \textit{und} \quad \measuredangle\big(f(x), f(y)\big) = \measuredangle(x, y).$$

b. *Für alle* $x, y \in V$ *gilt:*

$$x \perp y \quad \Leftrightarrow \quad f(x) \perp f(y).$$

c. *Ist* $\lambda \in \sigma(f)$*, dann ist* $|\lambda| = 1$.
d. f *ist injektiv.*
e. *Ist* f *bijektiv, so ist* f^{-1} *ebenfalls orthogonal bzw. unitär.*
f. $f \circ g$ *ist orthogonal bzw. unitär.*
g. *Ist* $\dim_{\mathbb{K}}(V) < \infty$*, dann ist* f *bijektiv.*

Beweis: a. Dies folgt unmittelbar aus der Definition der Norm und des Winkels mittels des Skalarproduktes.

b. Folgt aus a..

c. Ist $x \in V$ ein Eigenvektor zum Eigenwert $\lambda \in \mathbb{K}$, so gilt nach a.

$$\|x\| = \|f(x)\| = \|\lambda x\| = |\lambda| \cdot \|x\|,$$

also $|\lambda| = 1$, da $x \neq 0$.

d. Ist $x \in \mathrm{Ker}(f)$, so gilt nach a. $0 = \|f(x)\| = \|x\|$, und somit $x = 0$.

e. Ist f invertierbar, dann gilt für $x, y \in V$

$$\big\langle f^{-1}(x), f^{-1}(y)\big\rangle = \Big\langle f\big(f^{-1}(x)\big), f\big(f^{-1}(y)\big)\Big\rangle = \langle x, y\rangle.$$

f. Für $x, y \in V$ gilt

$$\big\langle (f \circ g)(x), (f \circ g)(y)\big\rangle = \Big\langle f\big(g(x)\big), f\big(g(y)\big)\Big\rangle = \big\langle g(x), g(y)\big\rangle = \langle x, y\rangle.$$

g. Ist $\dim_{\mathbb{K}}(V) < \infty$, dann folgt aus d., daß f auch bijektiv ist.

□

Satz 2.18
Ist $\dim_{\mathbb{K}}(V) < \infty$ *und* $f \in \mathrm{End}_{\mathbb{K}}(V)$*, dann sind gleichwertig:*

a. f *ist orthogonal bzw. unitär.*
b. *Für jede ONB* B *von* V *gilt* $M_B^B(f)$ *ist orthogonal bzw. unitär.*

c. *Es gibt eine ONB* B *von* V*, so daß* $M_B^B(f)$ *orthogonal bzw. unitär ist.*
d. *Für alle* $x \in V$ *gilt* $\|f(x)\| = \|x\|$.

Beweis: a. $\Rightarrow$ b.: Ist $B = (x_1, \ldots, x_n)$ eine ONB von V, dann sind die Spalten von $M_B^B(f)$ gerade die Vektoren $M_B\big(f(x_i)\big)$, $i = 1, \ldots, n$. Sei nun $M_B^B(f) \circ M_B^B(f)^* = (a_{ij})$ und bezeichne $\langle \cdot, \cdot \rangle_{\mathbb{K}^n}$ das kanonische Skalarprodukt auf $\mathbb{K}^n$, dann gilt unter Berücksichtigung der Definition der Adjungierten und unter Beachtung von $M_B\big(\langle \cdot, \cdot \rangle\big) = \mathbb{1}_n$

$$a_{ij} = \big\langle M_B\big(f(x_i)\big), M_B\big(f(x_j)\big)\big\rangle_{\mathbb{K}^n} = \big\langle f(x_i), f(x_j)\big\rangle = \langle x_i, x_j\rangle = \delta_{ij}.$$

D. h. $M_B^B(f) \circ M_B^B(f)^* = \mathbb{1}_n$.

b. $\Rightarrow$ c.: Klar.

c. $\Rightarrow$ a.: Es sei also $B = (x_1, \ldots, x_n)$ eine ONB bezüglich derer $M_B^B(f)$ orthogonal bzw. unitär ist. Dann folgt wie im Schritt "a. $\Rightarrow$ b".

$$\langle x_i, x_j\rangle = \delta_{ij} = \big\langle M_B\big(f(x_i)\big), M_B\big(f(x_j)\big)\big\rangle_{\mathbb{K}^n} = \langle f(x_i), f(x_j)\rangle$$

für alle $i, j \in \{1, \ldots, n\}$. Seien nun $x = \sum_{i=1}^n \lambda_i x_i, y = \sum_{i=1}^n \mu_i x_i \in V$ beliebig, dann gilt

$$\langle f(x), f(y)\rangle = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \overline{\mu_j} \langle f(x_i), f(x_j)\rangle = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \overline{\mu_j} \langle x_i, x_j\rangle = \langle x, y\rangle.$$

a. $\Rightarrow$ d.: Dies wurde bereits in Lemma 2.17 gezeigt.

d. $\Rightarrow$ a.: Dies folgt aus Lemma 1.20 im orthogonalen Fall und aus Lemma 1.22 im unitären Fall. □

Satz 2.19
Es sei $A \in \mathrm{Mat}(n, \mathbb{K})$ *und es sei* $\mathbb{K}^n$ *versehen mit dem kanonischen Skalarprodukt* $\langle \cdot, \cdot \rangle_{\mathbb{K}^n}$*. Dann sind gleichwertig:*

a. A *ist orthogonal bzw. unitär.*
b. A *ist invertierbar und* $A^{-1} = A^*$.
c. A^* *ist orthogonal bzw. unitär und* $A^{-1} = A^*$.
d. *Die Zeilen von* A *bilden eine ONB von* $\mathbb{K}^n$.
e. *Die Spalten von* A *bilden eine ONB von* $\mathbb{K}^n$.
f. $\|Ax\|_2 = \|x\|_2$ *für alle* $x \in \mathbb{K}^n$.
g. *Die lineare Abbildung* $f_A \in \mathrm{End}_{\mathbb{K}}\left(\mathbb{K}^n\right)$ *ist orthogonal bzw. unitär.*

Beweis: a. $\Leftrightarrow$ g. folgt aus Satz 2.18 unter Berücksichtigung, daß die kanonische Basis E eine ONB von $\mathbb{K}^n$ ist und daß $A = M_E^E(f_A)$.

a. $\Leftrightarrow$ f. folgt ebenfalls aus Satz 2.18, da $Ax = f_A(x)$ für $x \in V$.

a. $\Rightarrow$ b.: Dies folgt unmittelbar aus $A \circ A^* = \mathbb{1}_n$.

b. $\Rightarrow$ c.: Nach Voraussetzung gilt $A^* \circ \left(A^*\right)^* = A^* \circ A = A^{-1} \circ A = \mathbb{1}_n$, also ist A^* orthogonal bzw. unitär.

c. $\Rightarrow$ d.: Seien $a_1, \ldots, a_n$ die Zeilen von A. Dann folgt aus $\mathbb{1}_n = A \circ A^{-1} = A \circ A^* = \left(\langle a_i, a_j\rangle_{\mathbb{K}^n}\right)_{i,j}$, daß die Zeilen von A eine ONB von $\mathbb{K}^n$ bilden.

d. $\Rightarrow$ a.: Nach Voraussetzung gilt $\mathbb{1}_n = \left(\langle a_i, a_j\rangle_{\mathbb{K}^n}\right)_{i,j} = A \circ A^*$, also ist A orthogonal bzw. unitär.

a. $\Rightarrow$ e.: Ist A orthogonal bzw. unitär, dann gilt $\mathbb{1}_n = \left(A^* \circ A\right)^t = A^t \circ \left(A^t\right)^*$, also ist auch A^t orthogonal bzw. unitär. Aber dann bilden die Zeilen von A^t, sprich die Spalten von A, nach d. eine ONB.

e. $\Rightarrow$ a.: Bilden umgekehrt die Spalten von A eine ONB, dann ist nach d. A^t orthogonal bzw. unitär und somit gilt $\mathbb{1}_n = \left(A^t\right)^* \circ A^t = \left(A \circ A^*\right)^t$. Aber dann ist A orthogonal bzw. unitär. □

Lemma 2.20 a. *Sei* $A \in \mathrm{O}(n)$*, dann gilt* $\det(A) \in \{-1, 1\}$.
b. *Sei* $A \in \mathrm{U}(n)$*, dann gilt* $|\det(A)| = 1$.
c. *Sind* $A, B \in \mathrm{O}(n)$*, dann gilt* $A \circ B \in \mathrm{O}(n)$ *und* $A^{-1} \in \mathrm{O}(n)$.
d. *Sind* $A, B \in \mathrm{U}(n)$*, dann gilt* $A \circ B \in \mathrm{U}(n)$ *und* $A^{-1} \in \mathrm{U}(n)$.

Beweis: Es gilt

$$\begin{aligned} 1 &= \det(\mathbb{1}_n) = \det\left(A \circ A^*\right) = \det(A) \cdot \det\left(\overline{A}^t\right) \\ &= \det(A) \cdot \det\left(\overline{A}\right) = \det(A) \cdot \overline{\det(A)} = |\det(A)|^2. \end{aligned}$$

Daraus folgen die Aussagen a. und b..

Daß $\mathrm{O}(n)$ und $\mathrm{U}(n)$ gegen Inversion abgeschlossen sind, folgt aus Satz 2.19 c., und die Abgeschlossenheit gegenüber der Matrixmultiplikation folgt aus Lemma 2.17 f. unter Berücksichtigung von Satz 2.19. □

Korollar 2.21 a. $\mathrm{O}(n)$ *ist eine Untergruppe von* $\mathrm{Gl}_n(\mathbb{R})$*, und* $\mathrm{SO}(n)$ *ist eine Untergruppe von* $\mathrm{O}(n)$ *vom Index zwei.*
b. $\mathrm{U}(n)$ *ist eine Untergruppe von* $\mathrm{Gl}_n(\mathbb{C})$*, und* $\mathrm{SU}(n)$ *ist eine Untergruppe von* $\mathrm{U}(n)$.

Beweis: Beachtet man, daß $\mathbb{1}_n$ in jeder der angegebenen Mengen liegt, alle also nicht-leer sind, dann folgen die Aussagen aus Lemma 2.20. □

Bemerkung 2.22
$\left(S^1 := \{z \in \mathbb{C} \mid |z| = 1\}, \cdot\right)$ ist eine Untergruppe von $\left(\mathbb{C}^*, \cdot\right)$ und $\left(\{-1, 1\}, \cdot\right)$ eine Untergruppe von $\left(\mathbb{R}^*, \cdot\right)$. Somit induziert $\det : \mathrm{Gl}_n(\mathbb{K}) \to \mathbb{K}^*$ Gruppenhomomorphismen

$$\det : \mathrm{O}(n) \to \{-1, 1\} \quad \text{und} \quad \det : \mathrm{U}(n) \to S^1,$$

die in der Tat Epimorphismen sind, wie man leicht sieht, und deren Kerne gerade $\mathrm{SO}(n)$ bzw. $\mathrm{SU}(n)$ sind.

Insbesondere sind also $\mathrm{SO}(n)$ bzw. $\mathrm{SU}(n)$ sogar Normalteiler von $\mathrm{O}(n)$ bzw. $\mathrm{U}(n)$.

Darüberhinaus folgt, daß im orthogonalen Fall die Menge $O^-(n) := \{A \in O(n) \mid \det(A) = -1\}$ gleichmächtig zu $SO(n)$ ist und $O(n) = SO(n) \cup O^-(n)$ mit $SO(n) \cap O^-(n) = \emptyset$. $O^-(n)$ ist jedoch *keine* Gruppe!

Vom rein matrizentechnischen Standpunkt aus sind orthogonale bzw. unitäre Matrizen ausgezeichnet, weil man sie so unglaublich einfach invertieren kann, denn es gilt für $A \in O(n)$ bzw. $A \in U(n)$

$$A^{-1} = A^t \quad \text{bzw.} \quad A^{-1} = \overline{A}^t.$$

Insbesondere im Reellen haben sie aber auch eine besondere *geometrische Bedeutung*.

Wir wollen in dieser Hinsicht nun die orthogonalen Matrizen bzw. die orthogonalen Abbildungen des $\mathbb{R}^n$ für $n = 1$ und $n = 2$ untersuchen.

Bemerkung 2.23 a. $O(1) = \{-1, 1\}$, d. h. $O(\mathbb{R}) = \{-\mathrm{id}, \mathrm{id}\}$, ist recht übersichtlich. Neben der Identität haben wir nur die Spiegelung der reellen Achse am Nullpunkt.

b. Im Fall $n = 2$ gilt

$$SO(n) = \{T(\alpha) \mid \alpha \in [0, 2\pi[\} \quad \text{und} \quad O^-(n) = \{S(\alpha) \mid \alpha \in [0, 2\pi[\},$$

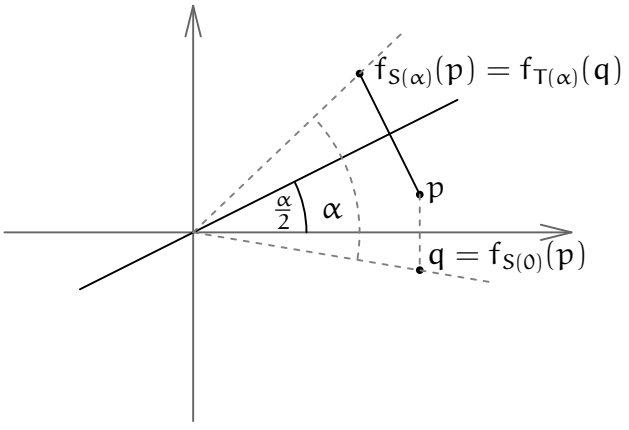
insbesondere also

$$O(n) = \{T(\alpha), S(\alpha) \mid \alpha \in [0, 2\pi[\},$$

wobei die Matrizen $T(\alpha)$ und $S(\alpha)$ für $\alpha \in \mathbb{R}$ wie folgt definiert sind:

$$T(\alpha) := \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \text{ und } S(\alpha) := \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{pmatrix}.$$

Dabei induziert $T(\alpha)$ eine Drehung um den Ursprung um den Winkel α und $S(\alpha)$ eine Spiegelung an der Geraden $\left\langle \left(\cos\left(\frac{\alpha}{2}\right), \sin\left(\frac{\alpha}{2}\right)\right)^t \right\rangle$.



Insbesondere ist im Fall $n = 2$ also jede orthogonale Abbildung eine Drehung oder eine Spieglung.

Man beachte auch daß

$$S(\alpha) = T(\alpha) \circ S(0) = T(\alpha) \circ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

d. h. die von $S(\alpha)$ induzierte Spiegelung kann auch geschrieben werden als Komposition der Spiegelung an der x-Achse gefolgt von einer Drehung um den Winkel α. Und damit gilt zugleich, daß jede Drehung im $\mathbb{R}^2$ Komposition von zwei Spiegelungen ist.
Für den Beweis der Aussagen verweisen wir den Leser auf Aufgabe 2.43.

Beachtet man, daß bei Transformation mit einer orthogonalen bzw. unitären Matrix T sowohl die darstellenden Matrizen von Bilinearformen (im Reellen) bzw. Sesquilinearformen (im Komplexen) als auch die darstellenden Matrizen von Endomorphismen das *gleiche* Transformationsverhalten zeigen, da schlicht $T^* = T^{-1}$ gilt (vgl. Satz 1.16), und beachtet man weiterhin, daß sich wichtige Eigenschaften wie die Orthogonalität bzw. Unitarität von Abbildungen (vgl. Satz 2.18, aber auch Aufgabe 2.36) nur erhalten, wenn man Matrixdarstellungen bezüglich Orthonormalbasen betrachtet, so scheint es naheliegend, sich im Folgenden für Transformationen auf orthogonale bzw. unitäre Matrizen zu beschränken oder - gleichwertig dazu - für Matrixdarstellungen nur Orthonormalbasen zuzulassen.

Ziel wird es nun zunächst sein, für orthogonale bzw. unitäre Abbildungen eine ONB von V zu finden, bezüglich derer die darstellende Matrix eine möglichst einfache Form hat.

Wie wir schon im Fall der Jordanschen Normalform gesehen haben, können wir über $\mathbb{K} = \mathbb{R}$ nicht die gleichen Ergebnisse erwarten, wie im komplexen Fall. Wir untersuchen deshalb zunächst den Fall von unitären Abbildungen.

Satz 2.24
Es sei $\dim_{\mathbb{C}}(V) < \infty$ *und sei* $f \in \mathrm{End}_{\mathbb{C}}(V)$. *Dann sind gleichwertig:*

a. f *ist unitär.*
b. V *besitzt eine ONB aus Eigenvektoren von* f *und für* $\lambda \in \sigma(f)$ *ist* $|\lambda| = 1$.
c. *Ist* $\sigma(f) = \{\lambda_1, \ldots, \lambda_r\} \subset \mathbb{C}$*, dann gilt* $|\lambda_i| = 1$ *für* $i = 1, \ldots, r$ *und*

$$V = \mathrm{Eig}(f, \lambda_1) \perp \ldots \perp \mathrm{Eig}(f, \lambda_r),$$

d. h. V *ist die orthogonale Summe der Eigenräume von* f.

Beweis: a. $\Rightarrow$ b.: Daß die Eigenwerte alle Betrag eins haben, wissen wir bereits aus Lemma 2.17. Es bleibt also zu zeigen, daß f bezüglich einer ONB diagonalisierbar ist.

Wir führen den Beweis durch Induktion nach $n = \dim_{\mathbb{C}}(V)$, wobei im Fall $n = 1$ nichts zu zeigen ist. Wir können deshalb $n > 1$ annehmen.

Da $\mathbb{K} = \mathbb{C}$, besitzt f einen Eigenwert λ sowie einen Eigenvektor $0 \neq x \in V$ zu λ. Wir setzen $U = \langle x \rangle$.
Behauptung: U und $U^\perp$ sind f-invariant.

Für U gilt die Aussage, da x ein Eigenvektor ist. Sei also $y \in U^\perp$. Wir müssen zeigen, daß $f(y) \perp x$ gilt. Da nach Lemma 2.17 $\lambda \cdot \overline{\lambda} = |\lambda| = 1$, gilt

$$\langle f(y), x\rangle = \langle f(y), \overline{\lambda}\lambda x\rangle = \lambda\langle f(y), \lambda x\rangle = \lambda\langle f(y), f(x)\rangle = \lambda\langle y, x\rangle = 0,$$

also $f(y) \perp x$, und damit $f(y) \in U^\perp$.

Da $f_{U^\perp}$ per definitionem ebenfalls unitär ist, folgt mit Induktion, daß es eine ONB B''' von $U^\perp$ aus Eigenvektoren von $f_{U^\perp}$ gibt. Damit ist dann $\tilde{B} = B' \cup B'''$ eine ONB von V aus Eigenvektoren von f.

b. $\Rightarrow$ c.: Klar.

c. $\Rightarrow$ a.: Wählen wir zu jedem der Eigenräume nach Satz 2.7 eine ONB B_i und setzen $B = B_1 \cup \ldots \cup B_r$, so ist B eine ONB von V, da die Eigenräume paarweise senkrecht aufeinander stehen. Aber dann ist $M_B^B(f) = (a_{ij})_{i,j}$ eine Diagonalmatrix und die Diagonalelemente sind gerade die Eigenwerte von f. Nach Voraussetzung gilt $a_{ii} \cdot \overline{a_{ii}} = |a_{ii}|^2 = 1$ für alle $i = 1, \ldots, r$. Damit folgt dann

$$M_B^B(f) \circ M_B^B(f)^* = \mathbb{1}_n,$$

also ist $M_B^B(f)$ - und damit f - unitär. □

Bemerkung 2.25 a. Ist $f \in O(V)$ so, daß das charakteristische Polynom von f über $\mathbb{R}$ in Linearfaktoren zerfällt, dann gilt Satz 2.24 für f analog - mit dem gleichen Beweis.

b. Der Beweis von Satz 2.7 und dann auch der von Satz 2.24 funktioniert ähnlich für abzählbar unendlich-dimensionale unitäre Vektorräume. Dann hat f höchstens abzählbar viele verschiedene Eigenwerte $\lambda_1, \lambda_2, \ldots$.

Korollar 2.26
Ist $A \in U(n)$, dann gibt es ein $T \in U(n)$ mit

$$T^* \circ A \circ T = T^{-1} \circ A \circ T = \begin{pmatrix} \lambda_1 & 0 & \ldots & \ldots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & 0 & \lambda_n \end{pmatrix}$$

wobei $\lambda_i \in \mathbb{C}$, $|\lambda_i| = 1$, $i = 1, \ldots, n$, die Eigenwerte von A sind.

Insbesondere ist jede unitäre Matrix diagonalisierbar.

Beweis: Ist A unitär, dann ist f_A unitär und wir finden eine ONB von $\mathbb{C}^n$ aus Eigenvektoren von f_A, sprich von A. Schreiben wir diese als Spalten in eine Matrix T, so ist $T \in U(n)$ und T transformiert A in eine Diagonalmatrix. □

Der Beweis ist konstruktiv, sofern man die Eigenwerte von A exakt kennt. Man leitet daraus folgenden prinzipiellen Algorithmus zur Bestimmung von T her.

Algorithmus 2.27

INPUT: $A \in \mathrm{U}(n)$.

OUTPUT: $T \in \mathrm{U}(n)$, so daß $T^* \circ A \circ T$ Diagonalgestalt hat.

1. **Schritt:** Bestimme die Eigenwerte von A.
2. **Schritt:** Bestimme für jeden Eigenwert von A eine Basis des zugehörigen Eigenraumes.
3. **Schritt:** Orthonormalisiere die Basen der Eigenräume mit dem Orthonormalisierungsverfahren von Gram-Schmidt und schreibe die Vektoren in eine Matrix T.
4. **Schritt:** Gib schließlich T zurück.

Orthogonale Abbildungen lassen sich im Allgemeinen nicht diagonalisieren, insbesondere nicht durch eine ONB. Wir haben in Beispiel III.2.4 gesehen, daß zum Beispiel die Matrix

$$T(\alpha) = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix},$$

die eine Drehung um den Ursprung um den Winkel α beschreibt, nicht diagonalisierbar ist. Der folgende Satz sagt nun, daß diese Beispiele im Wesentlichen aber auch schon die einzigen sind.

Satz 2.28

Es sei $\dim_{\mathbb{R}}(V) < \infty$ *und* $f \in \mathrm{O}(n)$. *Dann besitzt* V *eine ONB* B *und es gibt Zahlen* $r, s, t \in \mathbb{N}$ *sowie Winkel* $\alpha_1, \ldots, \alpha_t \in]0, 2\pi[\setminus\{\pi\}$, *so daß*

$$M_B^B(f) = \mathbb{1}_r \oplus -\mathbb{1}_s \oplus T(\alpha_1) \oplus \ldots \oplus T(\alpha_t).$$

Beweisskizze: Wir wollen den Beweis hier nur skizzieren, auf die wesentlichen neuen Ideen aber explizit eingehen.

1. **Schritt:** Jedes Polynom in $\mathbb{R}[t]$, insbesondere auch χ_f läßt sich über $\mathbb{R}$ in ein Produkt von Linearfaktoren und quadratischen Faktoren zerlegen. Da die Eigenwerte von f nur $+1$ und -1 sind, folgt

   $$\chi_f = (-1)^n \cdot (t-1)^r \cdot (t+1)^s \cdot g_1 \cdots g_t$$

   mit $g_i = t^2 - 2\cos(\alpha_i) + 1 \in \mathbb{R}[t]$, $\deg(g_i) = 2$ und g_i hat keine reelle Nullstelle.[5]
2. **Schritt:** Wie im unitären Fall gilt

   $$V = \mathrm{Eig}(f, 1) \perp \mathrm{Eig}(f, -1) \perp U$$

[5] Da g_i über $\mathbb{C}$ zerfällt, muß g_i somit eine Nullstelle $\lambda \in \mathbb{C} \setminus \mathbb{R}$ besitzen, und da die Koeffizienten von g_i reell sind, gilt dann $0 = \overline{0} = \overline{g_i(\lambda)} = g_i(\overline{\lambda})$, also $g_i = (t-\lambda)(t-\overline{\lambda})$. Da ferner für eine ONB B von V gilt, daß $M_B^B(f)$ als Matrix über $\mathbb{C}$ aufgefaßt unitär ist, muß λ als Eigenwert von $M_B^B(f)$ Betrag eins haben, d. h. es gibt ein $\alpha_i \in]0, 2\pi[\setminus\{\pi\}$ mit $\lambda = \cos(\alpha_i) + i\sin(\alpha_i)$ und $g_i = t^2 - 2\cos(\alpha_i)t + 1$. – Die gleiche Überlegung auf ein normiertes irreduzibles Polynom $p \in \mathbb{R}[t]$ vom Grad größer als eins angewandt, zeigt für eine Nullstelle $\lambda \in \mathbb{C} \setminus \mathbb{R}$ von p, daß $q = (t-\lambda)(t-\overline{\lambda}) = t^2 - 2\Re(\lambda)t + |\lambda|^2 \in \mathbb{R}[t]$ ein Faktor von p ist und mithin $q = p$.

mit $U = \big(\mathrm{Eig}(f,1) \perp \mathrm{Eig}(f,-1)\big)^{\perp}$ und $f(U) \subseteq U$. Es folgt

$$\chi_{f_U} = g_1 \cdots g_t.$$

Nach Cayley-Hamilton gilt dann aber für $x \in U$

$$0 = \chi_{f_U}(f)(x) = g_1(\ldots(g_t(f))\ldots)(x).$$

Für $0 \neq x \in U$ existiert dann ein i, $1 \leq i \leq t$, mit

$$y := g_{i+1}(\ldots(g_t(f))\ldots)(x) \neq 0, \text{ aber } f^2(y) - 2\cos(\alpha_i)f(y) + y = g_i(f)(y) = 0.$$

Damit folgt für y

$$f^2(y) = 2\cos(\alpha_i)f(y) - y \in \langle f(y), y\rangle =: U_i,$$

und U_i ist f-invariant mit $\chi_{f_{U_i}} = g_i$.

3. **Schritt:** Um nun U_i abspalten zu können, benötigen wir, daß das orthogonale Komplement von U_i in U f-invariant ist. Hat man dies überprüft, schließt man mittels Induktion, daß

$$V = \mathrm{Eig}(f,1) \perp \mathrm{Eig}(f,-1) \perp U_1 \perp \ldots \perp U_t$$

mit $\dim_{\mathbb{R}}(U_i) = 2$, U_i f-invariant und $\chi_{f_{U_i}} = g_i = t^2 - 2\cos(\alpha_i)t + 1$.

4. **Schritt:** Unter Verwendung von Bemerkung 2.23 schließt man dann, daß f_{U_i} bezüglich einer (beliebigen) ONB die Gestalt $T(\alpha_i)$ haben muß.

□

Bemerkung 2.29

Satz 2.24 und Satz 2.28 lösen das simultane Normalformproblem für positiv definite hermitesche (bzw. positiv definite symmetrische) Matrizen einerseits und unitäre (bzw. orthogonale) Matrizen andererseits.

Im unitären Fall gibt es zu gegebener positiv definiter hermitescher Matrix S und zu gegebenem unitärem $A \in U(n)$ *eine* Matrix $T \in U(n)$ mit

$$T^t \circ S \circ \overline{T} = \mathbb{1}_n \quad \text{und} \quad T^{-1} \circ A \circ T = \begin{pmatrix} \lambda_1 & 0 & \ldots & \ldots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & 0 & \lambda_n \end{pmatrix},$$

mit $\lambda_i \in \mathbb{C}$ und $|\lambda_i| = 1$, $i = 1 \ldots, r$.

Denn, da die Spalten von T eine ONB von $\mathbb{C}^n$ bezüglich b_S bilden, ist S bezüglich dieser Basis die Einheitsmatrix und S transformiert sich bei Basiswechsel wie angegeben, und die Aussage zu A folgt aus 2.24.

Analog gibt es zu gegebener positiv definiter symmetrischer Matrix $S \in \mathrm{Mat}(n,\mathbb{R})$ und zu gegebenem orthogonalem $A \in O(n)$ *eine* Matrix $T \in O(n)$ mit

$$T^t \circ S \circ T = \mathbb{1}_n \quad \text{und} \quad T^{-1} \circ A \circ T = \mathbb{1}_r \oplus -\mathbb{1}_s \oplus T(\alpha_1) \oplus \ldots \oplus T(\alpha_t),$$

mit $\alpha_i \in]0, 2\pi[\setminus\{\pi\}$, $i = 1, \ldots, t$.

Beispiel 2.30

Betrachten wir $\mathbb{K}^3$ mit dem kanonischen Skalarprodukt sowie die Matrix

$$A = \frac{1}{9} \cdot \begin{pmatrix} 1 & 8 & -4 \\ -4 & 4 & 7 \\ 8 & 1 & 4 \end{pmatrix} \in \mathbf{Mat}(3, \mathbb{K}).$$

Man rechnet sofort nach, daß $A \circ A^* = \mathbb{1}_3$, daß A also orthogonal bzw. unitär ist.

Sei zunächst einmal $\mathbb{K} = \mathbb{R}$. Wir rechnen nach, daß

$$\chi_A = t^3 - t^2 + t - 1 = (t-1)\big(t^2+1\big),$$

dabei ist $t^2 + 1 = t^2 - 2\cos\left(\frac{\pi}{2}\right)t + 1$. D. h., es muß eine orthogonale Matrix T geben, so daß

$$T^{-1} \circ A \circ T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Um T zu bestimmen, berechnen wir zunächst den Eigenraum $\mathrm{Eig}(A, 1)$ und finden $\left(\frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right)^t$ als ONB.

Sodann suchen wir einen Vektor $0 \neq y \in \mathbb{R}^3$ mit

$$\big(A^2 + \mathbb{1}_3\big)(y) = 0,$$

d. h. wir suchen letztlich einen Vektor in $\mathrm{Lös}\left(A^2 + \mathbb{1}_3, 0\right)$, wobei

$$A^2 + \mathbb{1}_3 = \begin{pmatrix} \frac{2}{9} & \frac{4}{9} & \frac{4}{9} \\ \frac{4}{9} & \frac{8}{9} & \frac{8}{9} \\ \frac{4}{9} & \frac{8}{9} & \frac{8}{9} \end{pmatrix} \xrightarrow{\text{Gauß}} \begin{pmatrix} 1 & 2 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Da uns eine Lösung y reicht, können wir $y = (2, 1, -2)^t$ wählen und erhalten dann als gesuchten f-invarianten Vektorraum

$$U = \langle y, Ay \rangle = \big\langle (2, 1, -2)^t, (2, -2, 1)^t \big\rangle,$$

wobei natürlich $U = \mathrm{Lös}\left(A^2 + \mathbb{1}_3, 0\right)$. Lassen wir den Orthonormalisierungsalgorithmus von Gram-Schmidt auf die Erzeuger von U los, so erhalten wir zusammen mit der ONB von $\mathrm{Eig}(A, 1)$ als Transformationsmatrix

$$T = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{pmatrix} \in \mathbf{O}(3),$$

und, wie gewünscht,

$$T^{-1} \circ A \circ T = T^t \circ A \circ T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Sei nun $\mathbb{K} = \mathbb{C}$, dann gilt $\chi_A = (t-1)(t-i)(t+i)$. Wir erhalten eine unitäre Transformationsmatrix T, die A diagonalisiert, indem wir die drei Eigenräume von T bestimmen und die Basisvektoren jeweils normieren.

Weiterhin ist $\left(\frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right)^t$ eine ONB von $\mathrm{Eig}(A, 1)$. Ferner überzeugt man sich durch Einsetzen in das Gleichungssystem, daß $(4, -1+3i, -1-3i)^t$ eine Lösung von $(A - i\mathbb{1}_3)x = 0$ ist. Durch Normierung erhalten wir dann $\left(\frac{2}{3}, -\frac{1}{6} + \frac{1}{2}i, -\frac{1}{6} - \frac{1}{2}i\right)^t$ als ONB von $\mathrm{Eig}(A, i)$. Da A eine reelle Matrix ist, muß somit $-i$ gerade den konjugiert komplexen Vektor als Eigenvektor haben, d. h. $\left(\frac{2}{3}, -\frac{1}{6} - \frac{1}{2}i, -\frac{1}{6} + \frac{1}{2}i\right)^t$ ist eine ONB von $\mathrm{Eig}(A, -i)$.

Wir erhalten also

$$T = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{6} + \frac{1}{2}i & -\frac{1}{6} - \frac{1}{2}i \\ \frac{2}{3} & -\frac{1}{6} - \frac{1}{2}i & -\frac{1}{6} + \frac{1}{2}i \end{pmatrix} \in \mathrm{U}(3)$$

als Transformationsmatrix mit

$$T^{-1} \circ A \circ T = T^* \circ A \circ T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{pmatrix}.$$

Will man sich die mühsamen Rechnungen zur Bestimmung des Eigenraumes von A zum Eigenwert i ersparen, so kann man dies näherungsweise mit Singular tun, wenn man die Prozedur `ONB` aus Aufgabe 2.45 implementiert hat:

```
ring r=complex,x,lp;
matrix A[3][3]=1/9,8/9,-4/9,-4/9,4/9,7/9,8/9,1/9,4/9;
module B=syz(A-i*unitmat(3));
print(ONB(unitmat(3),B));
↦    0.666667,
     (-0.166667 + i * 0.5),
     (-0.166667 - i * 0.5)
```

AUFGABEN

Aufgabe 2.31

Zeige, durch $\left((x_1, x_2, x_3)^t, (y_1, y_2, y_3)^t\right) := x_1y_1 + x_1y_2 + x_2y_1 + 2x_2y_2 + x_2y_3 + x_3y_2 + 2x_3y_3$ für $(x_1, x_2, x_3)^t, (y_1, y_2, y_3)^t \in \mathbb{R}^3$ wird ein Skalarprodukt auf $\mathbb{R}^3$ definiert, und bestimme eine Orthonormalbasis von $\mathbb{R}^3$ bezüglich dieses Skalarproduktes.

Aufgabe 2.32

Bestimme eine Orthonormalbasis des Unterraumes $V = \left\{f : [0,1] \to \mathbb{R} : t \mapsto at^2 + bt + c \mid a, b, c \in \mathbb{R}\right\} \subset \mathcal{C}^0\left([0,1]\right)$ bezüglich des Skalarproduktes $\langle \cdot, \cdot \rangle : V \to \mathbb{R} : (f, g) \mapsto \int_0^1 f(t)g(t)dt$.

Aufgabe 2.33

Es sei $V = \mathrm{Mat}(n, \mathbb{R})$. Wir definieren

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R} : (A, B) \mapsto \mathrm{Spur}\left(A^t \circ B\right).$$

a. Zeige, $\langle \cdot, \cdot \rangle$ ist ein Skalarprodukt auf V.
b. Zeige, für $U := \{A \in V \mid A^t = A\}$ und $W := \{A \in V \mid A^t = -A\}$ gilt $V = U \perp W$.

Aufgabe 2.34
Es sei $\big(V, \langle \cdot, \cdot \rangle\big)$ ein euklidischer oder unitärer Raum und $G \subset \mathrm{Gl}_{\mathbb{K}}(V)$ sei eine *endliche* Untergruppe der Automorphismengruppe von V. Zeige:

a. Durch $(x, y) := \sum_{f \in G} \langle f(x), f(y) \rangle$ für $x, y \in V$ wird auf V ein Skalarprodukt definiert.
b. Ist $f \in G$, so ist f unitär bezüglich des neuen Skalarproduktes $(\cdot, \cdot)$.

Aufgabe 2.35
Es sei $\big(V, \langle \cdot, \cdot \rangle\big)$ ein endlich-dimensionaler Hilbertraum und $f \in \mathrm{End}_{\mathbb{K}}(V)$. Zeige:

a. Es gibt genau ein $f^* \in \mathrm{End}_{\mathbb{K}}(V)$, die sogenannte *Adjungierte* von f, mit $\big\langle f(x), y \big\rangle = \big\langle x, f^*(y) \big\rangle$ für alle $x, y \in V$.
b. Ist $B = (x_1, \ldots, x_n)$ eine Basis von V und $G = \big(\langle x_i, x_j \rangle\big)_{i,j=1,\ldots,n} \in \mathrm{Mat}(n, \mathbb{K})$ die sogenannte *Gramsche Matrix* von B, dann gilt, G ist invertierbar und
$$M_B^B\big(f^*\big) = \overline{G}^{-1} \circ \overline{M_B^B(f)}^t \circ \overline{G}.$$
Insbesondere gilt, ist B eine Orthonormalbasis, dann ist $M_B^B\big(f^*\big)$ die adjungierte Matrix von $M_B^B(f)$.

Hinweise: In Teil a. betrachte man eine ONB $B = (x_1, \ldots, x_n)$ und definiere $f^*(x_i)$ als Linearkombination der x_j in geeigneter Weise. In Teil b. zeige man zunächst $G \circ \overline{M_B^B\big(f^*\big)} = M_B^B(f)^t \circ G$.

Aufgabe 2.36
Es sei $\big(V, \langle \cdot, \cdot \rangle\big)$ ein endlich-dimensionaler Hilbertraum und $f \in \mathrm{End}_{\mathbb{K}}(V)$. f heißt *normal*, falls $f \circ f^* = f^* \circ f$. Zeige:

a. Ist B eine Orthonormalbasis von V, dann ist f genau dann normal, $M_B^B(f)$ normal ist.
b. Ist f normal, dann gelten:
   (i) $\mathrm{Ker}(f) = \mathrm{Ker}\big(f^*\big)$.
   (ii) $V = \mathrm{Ker}(f) \perp \mathrm{Im}(f)$.
   (iii) $\mathrm{Im}(f) = \mathrm{Im}\big(f^*\big)$.
c. Genau dann ist f orthogonal (bzw. unitär), wenn $f^* = f^{-1}$.

Aufgabe 2.37
Es sei $\big(V, \langle \cdot, \cdot \rangle\big)$ ein endlich-dimensionaler Hilbertraum und $f \in \mathrm{End}_{\mathbb{K}}(V)$. Zeige:

a. Ist f normal und sind $\lambda, \mu \in \mathbb{K}$ mit $\lambda \neq \mu$, dann gilt $\mathrm{Eig}(f, \lambda) = \mathrm{Eig}\big(f^*, \overline{\lambda}\big)$ und $\mathrm{Eig}(f, \lambda) \perp \mathrm{Eig}(f, \mu)$.
b. Ist $\mathbb{K} = \mathbb{C}$, dann sind gleichwertig:
   (i) f ist normal.
   (ii) Es gibt eine Orthonormalbasis B aus Eigenvektoren von f.
   (iii) $\|f(x)\| = \|f^*(x)\|$ für alle $x \in V$.

Aufgabe 2.38
Es sei V ein endlich-dimensionaler unitärer Raum und $f \in \mathrm{End}_{\mathbb{C}}(V)$. Zeige, die folgenden Aussagen sind gleichwertig:

a. f ist normal.
b. Es gibt ein Polynom $p \in \mathbb{C}[t]$ mit $f^* = p(f)$.
c. Für alle $g \in \mathrm{End}_{\mathbb{C}}(V)$ mit $f \circ g = g \circ f$ gilt $f^* \circ g = g \circ f^*$.

Aufgabe 2.39
Es sei V ein endlich-dimensionaler euklidischer oder unitärer Raum und $f \in \mathrm{Gl}_{\mathbb{K}}(V)$. Zeige, die folgenden Aussagen sind äquivalent:

a. Für $x, y \in V$ mit $x \perp y$ gilt $f(x) \perp f(y)$.
b. Für $x, y \in V$ mit $\|x\| = \|y\|$ gilt $\|f(x)\| = \|f(y)\|$.
c. Es gibt ein $\lambda \in \mathbb{R}_{>0}$ und ein $g \in \mathrm{O}(V)$ mit $f = \lambda g$.

Aufgabe 2.40
Es sei V ein endlich-dimensionaler euklidischer Raum. Eine Abbildung $f \in \mathrm{End}_{\mathbb{R}}(V)$ heißt *Spiegelung* an $\langle x\rangle^{\perp}$ (für $0 \neq x \in V$), falls $f(x) = -x$ und $f(y) = y$ für alle $y \in \langle x\rangle^{\perp}$.

Zeige, ist $f \in \mathrm{End}_{\mathbb{R}}(V)$ eine Spiegelung an $\langle x\rangle^{\perp}$, dann gelten:

a. f ist orthogonal.
b. $\det(f) = -1$ und $f^2 = \mathrm{id}_V$.
c. Für $y \in V$ gilt: $f(y) = y - \frac{2\langle x,y\rangle}{\langle x,x\rangle}x$.

Aufgabe 2.41
Es sei V ein endlich-dimensionaler euklidischer Raum und $x, y \in V \setminus \{0\}$. Ferner seien $s \in \mathrm{End}_{\mathbb{R}}(V)$ die Spiegelung an $\langle x\rangle^{\perp}$ und $t \in \mathrm{End}_{\mathbb{R}}(V)$ die Spiegelung an $\langle y\rangle^{\perp}$. Zeige die Äquivalenz der folgenden Aussagen:

a. $s \circ t = t \circ s$.
b. $\langle x\rangle = \langle y\rangle$ oder $x \perp y$.

Aufgabe 2.42
Es sei V ein n-dimensionaler euklidischer Raum und $f \in \mathrm{O}(V)$. Dann gibt es $k \leq n - \dim_{\mathbb{R}}\big(\mathrm{Eig}(f,1)\big)$ Spiegelungen $s_1, \ldots, s_k \in \mathrm{O}(V)$ mit $f = s_1 \circ \ldots \circ s_k$.

Insbesondere wird $\mathrm{O}(V)$ von Spiegelungen erzeugt.

Hinweis: Man führe Induktion über n. Betrachte zunächst den Fall, daß $\mathrm{Eig}(f,1) \neq \{0\}$. Im Fall $\mathrm{Eig}(f,1) = \{0\}$ betrachte die Spiegelung s_1 am orthogonalen Komplement von $f(x) - x$ für ein festes $x \neq 0$ und zeige, daß $x \in \mathrm{Eig}(s_1 \circ f, 1)$. Damit ist der erste Fall auf $s_1 \circ f$ anwendbar.

Aufgabe 2.43
Zeige:

a. Die zu $S(\alpha)$ assoziierte Abbildung $f_{S(\alpha)}$ ist eine Spiegelung an der Geraden $\left\langle \left(\cos\left(\frac{\alpha}{2}\right), \sin\left(\frac{\alpha}{2}\right)\right)^t \right\rangle$.
b. $\mathrm{SO}(2) = \{T(\alpha) \mid \alpha \in \mathbb{R}\}$ und $\mathrm{O}^-(2) = \{S(\alpha) \mid \alpha \in \mathbb{R}\}$.

Insbesondere ist jede orthogonale Abbildung des $\mathbb{R}^2$ entweder eine Drehung oder eine Spiegelung.

Hinweis: In a. braucht man die Additionstheoreme für Cosinus und Sinus. In b. zeige man zunächst, daß jede Spiegelung von der Form $f_{S(\alpha)}$ ist und wende dann für $O^-(2)$ Aufgabe 2.42 an.

Aufgabe 2.44

Schreibe eine Singular-Prozedur `OGB`, die eine symmetrische Matrix $A \in \mathrm{Mat}(n, \mathbb{Q})$ sowie eine endliche Familie von Vektoren $M \subset \mathbb{Q}^n$ einliest und eine Orthogonalbasis von $\langle M \rangle_{\mathbb{R}}$ ausgibt.

Man verwende den folgenden Algorithmus:

INPUT: $A \in \mathrm{Mat}(n, \mathbb{Q})$ positiv definit symmetrisch und $M \subset \mathbb{Q}^n$.

OUTPUT: B Orthogonalbasis von $\langle M \rangle_{\mathbb{R}}$.

1. Schritt: Bestimme eine Basis B von $\langle M \rangle$ mittels des Algorithmus II.6.3.

2. Schritt: Für $i = 1, \ldots, \dim_{\mathbb{C}} \left(\langle M \rangle \right)$ führe folgende Schritte aus:

a. Berechne $v = B[i] - \sum_{j=1}^{i-1} \left(B[i]^t \circ A \circ \overline{B[j]} \right) / \left(B[j]^t \circ A \circ \overline{B[j]} \right) \cdot B[j]$.

b. Ersetze $B[i]$ durch v.

3. Schritt: Gib die veränderte Basis B zurück.

Aufgabe 2.45

Schreibe eine Singular-Prozedur `ONB`, die eine hermitesche Matrix $A \in \mathrm{Mat}(n, \mathbb{C})$ sowie eine endliche Familie von Vektoren $M \subset \mathbb{C}^n$ einliest und eine Orthonormalbasis von $\langle M \rangle$ ausgibt.

Man verwende den folgenden Algorithmus:

INPUT: $A \in \mathrm{Mat}(n, \mathbb{C})$ positiv definit hermitesch und $M \subset \mathbb{C}^n$.

OUTPUT: B Orthonormalbasis von $\langle M \rangle$.

1. Schritt: Bestimme eine Basis B von $\langle M \rangle$ mittels des Algorithmus II.6.3.

2. Schritt: Für $i = 1, \ldots, \dim_{\mathbb{C}} \left(\langle M \rangle \right)$ führe folgende Schritte aus:

a. Berechne die Summe $v = B[i] - \sum_{j=1}^{i-1} \left(B[i]^t \circ A \circ \overline{B[j]} \right) \cdot B[j]$.

b. Dividiere v durch $\sqrt{\left(v^t \circ A \circ \overline{v} \right)}$.

c. Ersetze $B[i]$ durch v.

3. Schritt: Gib die veränderte Basis B zurück.

Hinweise: M soll vom Typ `module` sein, ebenso die zurückgegebene Basis `B`. – Das Skalarprodukt, bezüglich dessen orthonormalisiert wird, ist gegeben durch $\langle x, y \rangle := x^t \circ A \circ \overline{y}$. – Man beachte, daß x und y in Singular vom Typ `matrix` sein müssen, damit `transpose(x)*A*conjugate(y)` nicht zu Problemen führt, und daß in diesem Fall das Ergebnis von `transpose(x)*A*conjugate(y)` wieder vom Typ `matrix` ist, wenn auch eine 1×1-Matrix; auf den Eintrag kann mittels `(transpose(x)*A*conjugate(y))[1,1]` zugegriffen werden. – Die Quadratwurzeln einer komplexen Zahl c kann man mittels der Prozedur `laguerre_solve` als Nullstellen von $t^2 - c$ aus `solve.lib` bestimmen. – Schließlich beachte man, daß diese Prozedur nur über dem Ring `ring r=complex,t,lp;` Sinn macht.

Aufgabe 2.46

Ziel dieser Aufgabe ist es, zu zeigen, daß die Überführung einer Matrix

in Jordansche Normalform ein unstetiger Prozeß ist, während das Gram-Schmidt-Verfahren zur Überführung einer Basis in Orthonormalform stetig ist.

Es sei $V = \mathrm{Mat}(n, \mathbb{C})$ und $\|\cdot\|_z : V \to \mathbb{R}_{\geq 0} : (a_{ij}) \mapsto \max\left\{ \sum_{j=1}^{n} |a_{ij}| \;\middle|\; i = 1, \ldots, n \right\}$ sei die Zeilen-Summen-Norm auf V. (Vgl. Aufgabe 1.29.)

a. Zeige, die Abbildung $J : \big(V, \|\cdot\|_z\big) \to \big(V, \|\cdot\|_z\big) : A \mapsto J_A$ ist nicht stetig.[6]
b. Wir betrachten nun die offene Teilmenge $U = \mathrm{Gl}_n(\mathbb{C}) = V \setminus \big\{A \in V \mid \det(A) = 0\big\}$ von V und die Abbildung $\mathrm{ONB} : U \to U : B \mapsto \mathrm{ONB}(B)$. Zeige, ONB ist stetig.[7]

3. Hauptachsentransformation

Generalvoraussetzung: In diesem Abschnitt verwenden wir $\mathbb{K}$ stets dann, wenn Aussagen sowohl für den Fall des Körpers $\mathbb{K} = \mathbb{R}$ der reellen Zahlen, als auch für den Fall des Körpers $\mathbb{K} = \mathbb{C}$ der komplexen Zahlen gelten sollen. Ferner sei $\big(V, \langle\cdot,\cdot\rangle\big)$ ein euklidischer Raum, falls $\mathbb{K} = \mathbb{R}$, und ein unitärer Raum, falls $\mathbb{K} = \mathbb{C}$. Zugleich versehen wir V stets mit der durch $\langle\cdot,\cdot\rangle$ definierten Norm, d. h. für $x \in V$ ist stets $\|x\| = \sqrt{\langle x, x\rangle}$. – V wird nicht notwendigerweise als endlich-dimensional vorausgesetzt.

Wir haben in den Sätzen 2.24 und 2.28 Normalformen für strukturerhaltende lineare Abbildungen eines endlich-dimensionalen Hilbertraumes hergeleitet.

Neben den strukturerhaltenden, d. h. den unitären bzw. orthogonalen Abbildungen gibt es andere wichtige Abbildungen, insbesondere die sogenannten selbstadjungierten Abbildungen. Diese haben ebenfalls wichtige Anwendungen und wir wollen zum Schluß dieses Abschnittes Normalformen für sie herleiten.

Definition 3.1
Ein Endomorphismus $f \in \mathrm{End}_{\mathbb{K}}(V)$ heißt *selbstadjungiert* oder *hermitesch*,

[6] Damit die Jordansche Normalform einer Matrix $A \in \mathrm{Mat}(n, \mathbb{C})$ eindeutig bestimmt ist, sprich, damit J eine Abbildung wird, müssen wir festlegen, in welcher Reihenfolge die Jordankästchen in J_A zu stehen haben. Die gewählte Reihenfolge ist für die Aussage jedoch vollkommen irrelevant. Wir einigen uns darauf, die Eigenwerte einer Matrix wie folgt zu ordnen: sind $\lambda = a + ib$ und $\mu = c + id$ zwei Eigenwerte von A, so ist $\lambda > \mu$, falls $a > c$ oder falls $a = c$ und $b > d$. Ferner einigen wir uns darauf, daß bei der Jordanschen Normalform die Jordanblöcke primär nach Größe der Eigenwerte und sekundär nach Größe der Jordankästchen sortiert werden.

[7] Beachte: Vektoren $x_1, \ldots, x_n$ bilden genau dann eine Basis von $\mathbb{C}^n$, wenn die Matrix B, deren Spalten $x_1, \ldots, x_n$ bilden, invertierbar ist. Die Abbildung ONB soll nun die Matrix B überführen in die Matrix, deren Spaltenvektoren aus den Spaltenvektoren von B durch das Orthogonalisierungsverfahren von Gram-Schmidt hervorgehen. – Daß U offen in V ist, folgt daraus, daß $\det : \big(V, \|\cdot\|\big) \to \big(\mathbb{C}, |\cdot|\big)$ eine stetige Abbildung ist. – Man zerlege ONB in eine Vielzahl von stetigen Funktionen, die durch Verkettung, Addition und Multiplikation verknüpft sind. Dabei mache man reichlich Gebrauch davon, daß auf endlich-dimensionalen Vektorräumen alle Normen äquivalent sind.

falls für alle $x, y \in V$ gilt

$$\langle f(x), y\rangle = \langle x, f(y)\rangle.$$

Bemerkung 3.2
Ist $\dim_{\mathbb{K}}(V) < \infty$, so wird in Aufgabe 2.35 für jeden Endomorphismus $f \in \mathrm{End}_{\mathbb{K}}(V)$ ein eindeutig bestimmter Endomorphismus $f^* \in \mathrm{End}_{\mathbb{K}}(V)$, der sogenannte zu f *adjungierte Endomorphismus*, definiert, der die Eigenschaft hat, daß für alle $x, y \in V$ gilt

$$\langle f(x), y\rangle = \langle x, f^*(y)\rangle.$$

Damit gilt, f ist genau dann selbstadjungiert, wenn $f = f^*$, d. h., wenn f gleich seiner Adjungierten ist.

Proposition 3.3
Für $f \in \mathrm{End}_{\mathbb{K}}(V)$ sind gleichwertig:

a. *f ist selbstadjungiert.*
b. *Für jede ONB B von V gilt $M_B^B(f)$ ist symmetrisch bzw. hermitesch.*
c. *Es gibt eine ONB B von V, so daß $M_B^B(f)$ symmetrisch bzw. hermitesch ist.*

Beweis: Es sei $B = (x_1, \ldots, x_n)$ eine ONB von V und $M_B^B(f) = A = (a_{ij})_{i,j}$. Ferner seien $v, w \in V$ und $x = M_B(v) \in \mathbb{K}^n$ sowie $y = M_B(w) \in \mathbb{K}^n$. Da B eine ONB von V ist, ist die darstellende Matrix des Skalarproduktes gerade die Einheitsmatrix und damit gilt:

$$\langle f(v), w\rangle = (Ax)^t \circ \overline{y} = x^t \circ A^t \circ \overline{y} \tag{61}$$

und

$$\langle v, f(w)\rangle = x^t \circ \left(\overline{Ay}\right) = x^t \circ \overline{A} \circ \overline{y}. \tag{62}$$

a. $\Rightarrow$ b.: Ist f selbstadjungiert und B eine beliebige ONB von V, so sind die Ausdrücke in (61) und (62) gleich für alle $v, w \in V$. Setzen wir $v = x_i$ und $w = x_j$, so erhalten wir $a_{ji} = \overline{a_{ij}}$ und mithin $A^t = \overline{A}$. Aber dann gilt $A = A^*$ und A ist symmetrisch bzw. hermitesch.

b. $\Rightarrow$ c.: Klar.

c. $\Rightarrow$ a.: Sei nun B eine ONB von V bezüglich derer $M_B^B(f) = A$ symmetrisch bzw. hermitesch ist. Dann sind die Ausdrücke in (61) und (62) wieder gleich für alle $v, w \in V$, da $A^t = \overline{A}$, und damit ist f selbstadjungiert. □

Die Bedingung $A = A^*$ ist überraschenderweise genau die Bedingung, die wir an die darstellende Matrix des Skalarproduktes im $\mathbb{K}^n$ gestellt haben. Allerdings sind zwei wichtige Unterschiede festzuhalten. Erstens verlangen wir hier nicht, daß die Matrizen positiv definit sind und zweitens werden die Matrizen hier als lineare Abbildungen aufgefaßt und nicht als symmetrische bzw. hermitesche Bilinearformen.

Lemma 3.4
Ist $n = \dim_{\mathbb{K}}(V) < \infty$ und $f \in \mathrm{End}_{\mathbb{K}}(V)$ selbstadjungiert, dann ist $\chi_f \in \mathbb{R}[t]$ und χ_f zerfällt über $\mathbb{R}$.

Insbesondere gilt, ist $\lambda \in \sigma(f)$ *ein Eigenwert von* f*, dann ist* $\lambda \in \mathbb{R}$.

Beweis: Ist $\mathcal{B}$ eine ONB, dann ist $A = M_{\mathcal{B}}^{\mathcal{B}}(f)$ symmetrisch bzw. hermitesch und es reicht zu zeigen, daß $\chi_A \in \mathbb{R}[t]$ über $\mathbb{R}$ zerfällt.

Hierfür machen wir uns zunutze, daß wir A auf alle Fälle auffassen können als eine Matrix in $\mathrm{Mat}(n, \mathbb{C})$ und daß gilt $A = A^*$. Über $\mathbb{C}$ zerfällt das charakteristische Polynom von A, d. h. es gibt $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ mit

$$\chi_A = (-1)^n(t - \lambda_1) \cdots (t - \lambda_n).$$

Es reicht nun, zu zeigen, daß $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. Nun gibt es zu jedem λ_i aber einen Vektor $0 \neq x_i \in \mathbb{C}^n$ mit $Ax_i = \lambda_i x_i$. Für diesen gilt, da $A^t = \overline{A}$,

$$\begin{aligned} \lambda_i\big(x_i^t \circ \overline{x_i}\big) &= \big(Ax_i\big)^t \circ \overline{x_i} = x_i^t \circ A^t \circ \overline{x_i} \\ &= x_i^t \circ \overline{A} \circ \overline{x_i} = x_i^t \circ \overline{Ax_i} = \overline{\lambda_i}\big(x_i^t \circ \overline{x_i}\big). \end{aligned}$$

Und da $x_i^t \circ \overline{x_i} \neq 0$, folgt $\lambda_i = \overline{\lambda_i}$, d. h. $\lambda_i \in \mathbb{R}$. □

Bemerkung 3.5
Durch Übergang V zu $V_{\mathbb{C}} = \mathbb{C} \otimes_{\mathbb{R}} V$ und von $f \in \mathrm{End}_{\mathbb{R}}(V)$ zu $\mathrm{id} \otimes f \in \mathrm{End}_{\mathbb{C}}(\mathbb{C} \otimes_{\mathbb{R}} V)$ kann man ohne Rückgriff auf Matrizen den Beweis direkt führen - der Übergang sichert uns die Existenz von Eigenwerten und Eigenvektoren. (Vgl. Beispiel V.2.19 e. und Proposition V.2.22.)

Außerdem gilt auch im unendlich-dimensionalen Fall auf alle Fälle, daß jeder Eigenwert eines selbstadjungierten Endomorphismus reell ist.

Seien nämlich $x \in V$ ein Eigenvektor zum Eigenwert $\lambda \in \mathbb{K}$, dann gilt

$$\lambda\langle x, x\rangle = \big\langle f(x), x\big\rangle = \big\langle x, f(x)\big\rangle = \overline{\lambda}\langle x, x\rangle.$$

Dann argumentiert man wie im Beweis von Lemma 3.4 weiter.

Satz 3.6
Es sei $\dim_{\mathbb{K}}(V) < \infty$ *und sei* $f \in \mathrm{End}_{\mathbb{K}}(V)$. *Dann sind gleichwertig:*

a. f *ist selbstadjungiert.*
b. V *besitzt eine ONB aus Eigenvektoren von* f *und für* $\lambda \in \sigma(f)$ *ist* $\lambda \in \mathbb{R}$.
c. *Ist* $\sigma(f) = \{\lambda_1, \ldots, \lambda_r\} \subset \mathbb{K}$*, dann gilt* $\lambda_i \in \mathbb{R}$ *für* $i = 1, \ldots, r$ *und*

$$V = \mathrm{Eig}(f, \lambda_1) \perp \ldots \perp \mathrm{Eig}(f, \lambda_r),$$

d. h. V *ist die orthogonale Summe der Eigenräume von* f.

Beweis: a. ⇒ b.: Daß die Eigenwerte alle reell sind, wissen wir bereits aus Lemma 3.4. Es bleibt also zu zeigen, daß f bezüglich einer ONB diagonlisierbar ist.

Wir führen den Beweis durch Induktion nach $n = \dim_{\mathbb{C}}(V)$, wobei im Fall $n = 1$ nichts zu zeigen ist. Wir können deshalb $n > 1$ annehmen.

Nach Lemma 3.4 zerfällt χ_f über $\mathbb{R}$ in Linearfaktoren und somit besitzt f einen Eigenwert λ sowie einen Eigenvektor $0 \neq x \in V$ zu λ. Wir setzen $U = \langle x\rangle$.

Behauptung: U und $U^\perp$ sind f-invariant.

Für U gilt die Aussage, da x ein Eigenvektor ist. Sei also $y \in U^\perp$. Wir müssen zeigen, daß $f(y) \perp x$ gilt. Nun gilt aber

$$\langle f(y), x\rangle = \langle y, f(x)\rangle = \langle y, \lambda x\rangle = \lambda\langle y, x\rangle = 0,$$

also $f(y) \perp x$, und damit $f(y) \in U^\perp$.

Da $f_{U^\perp}$ per definitionem ebenfalls selbstadjungiert sind, folgt mit Induktion, daß es eine ONB B''' von $U^\perp$ aus Eigenvektoren von $f_{U^\perp}$ gibt. Damit ist dann $\tilde{B} = B' \cup B'''$ eine ONB von V aus Eigenvektoren von f.

b. $\Rightarrow$ c.: Klar.

c. $\Rightarrow$ a.: Wählen wir zu jedem der Eigenräume nach Satz 2.7 eine ONB B_i und setzen $B = B_1 \cup \ldots \cup B_r$, so ist B eine ONB von V, da die Eigenräume paarweise senkrecht aufeinander stehen. Aber dann ist $M_B^B(f) = (a_{ij})_{i,j}$ eine Diagonalmatrix und die Diagonalelemente sind gerade die Eigenwerte von f. Nach Voraussetzung gilt $a_{ii} \in \mathbb{R}$ für alle $i = 1, \ldots, r$. Damit gilt dann

$$M_B^B(f)^* = \overline{M_B^B(f)}^t = M_B^B(f),$$

also ist $M_B^B(f)$ - und damit f - symmetrisch bzw. hermitesch. □

Korollar 3.7 (Hauptachsentransformation)
Zu jeder symmetrischen bzw. hermiteschen Matrix $A \in \mathrm{Mat}(n, \mathbb{K})$ *gibt es eine Matrix* $T \in O(n)$ *bzw.* $T \in U(n)$ *mit*

$$T^* \circ A \circ T = T^{-1} \circ A \circ T = \begin{pmatrix} \lambda_1 & 0 & \ldots & \ldots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & 0 & \lambda_n \end{pmatrix}$$

mit $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$.

Insbesondere ist jede symmetrische bzw. hermitesche Matrix diagonalisierbar und hat nur reelle Eigenwerte.

Dies ist eine wichtige Ergänzung des Satzes über die Jordansche Normalform.

In Aufgabe 2.37 zeigen wir, daß es in der Tat reicht, $A^* \circ A = A \circ A^*$ zu fordern, um zu sichern, daß A über $\mathbb{C}$ diagonalisierbar ist. Das schließt die Fälle von hermiteschen und unitären Matrizen ein, und die Eigenschaft ist algorithmisch leicht nachzuprüfen.

Wir wollen jetzt einen Algorithmus angeben, mit Hilfe dessen man eine orthogonale (bzw. unitäre) Transformationsmatrix bestimmen kann, mit Hilfe derer eine symmetrische (bzw. normale) Matrix auf Diagonalgestalt gebracht werden kann.

Algorithmus 3.8 (Diagonalisierung symmetrischer / normaler Matrizen)
INPUT: $A \in \mathrm{Mat}(n, \mathbb{R})$ symmetrisch (bzw. $A \in \mathrm{Mat}(n, \mathbb{C})$ normal).

OUTPUT: $D \in \mathrm{Mat}(n, \mathbb{R})$ (bzw. $D \in \mathrm{Mat}(n, \mathbb{C})$) Diagonalmatrix mit $D \sim A$, und $T \in \mathrm{O}(n)$ (bzw. $T \in \mathrm{U}(n)$) mit $T^{-1} \circ A \circ T = D$.

1. **Schritt:** Überprüfe, ob A symmetrisch (bzw. normal) ist.
2. **Schritt:** Bestimme die Nullstellen des charakteristischen Polynoms.
3. **Schritt:** Bestimme für jeden Eigenwert von A eine Orthonormalbasis des zugehörigen Eigenraumes. Schreibe die Basisvektoren als Spaltenvektoren in eine Matrix T.
4. **Schritt:** Gib $T^* \circ A \circ T$ und T zurück.

Bemerkung 3.9
Statt zu orthonormalisieren kann man im dritten Schritt auch nur jeweils eine OGB bestimmen. Das hat den Vorteil, daß man keine Wurzel ziehen muß, falls zum Beispiel die Einträge der Matrizen in $\mathbb{Q}$ sind. Man erhält dann eine Matrix T, deren Spalten paarweise orthogonal zueinander sind, aber nicht normiert. Es gilt dann *nicht* mehr $T^* = T^{-1}$. Aber man kann die Inverse von T berechnen, und diese hat wieder Einträge in $\mathbb{Q}$. Man kann also zumindest exakt, d. h. ohne Rundungsfehler T so bestimmen, daß $T^{-1} \circ A \circ T$ eine Diagonalmatrix ist.

Beispiel 3.10
Wir betrachten $\mathbb{C}^n$ mit dem kanonischen Skalarprodukt sowie die Matrix

$$A = \begin{pmatrix} 0 & -1 & i \\ -1 & 0 & -i \\ -i & i & 0 \end{pmatrix} \in \mathrm{Mat}(3, \mathbb{C}).$$

Da $A = A^*$ gilt, ist A hermitesch.

Wir berechnen eine Transformationsmatrix T mit paarweise orthogonalen Spalten, die A in Diagonalgestalt überführt, mit Hilfe von Singular.

```
LIB "linalg.";
ring r=(0,i),t,dp;
minpoly=i2+1; // Das bewirkt, daß i^2 = -1 gilt.
matrix A[3][3]=0,-1,i,-1,0,-i,-i,i,0;
poly p=det(A-unitmat(3)*t);
p;
↦ -t3+3x+2
```

In dem Ring r rechnen wir jetzt mit komplexen Zahlen der Form $a + ib$, wobei a und b rational sind, d. h. in $\mathbb{Q}[i]$. Singular kann in diesem Ring *nicht* faktorisieren. Wir wechseln daher in einen neuen Ring, in dem i eine Variable ist.

```
ring s=0,(i,t),dp;
poly p=imap(r,p); // Wir holen das Polynom p nach s.
factorize(p);
```

```
↦      [1]:  // Faktoren von p
             [1]=-1
             [2]=t-2
             [3]=t+1
       [2]:  // Vielfachheiten
             1,1,2
setring r;
matrix w1=syz(A-2*unitmat(3));
print(w1);
↦    1,   // Basis von Eig(f,2).
     -1,
     -i
matrix w2=syz(A+unitmat(3));
print(w2);
↦    1,  0,  // Basis von Eig(f,-1).
     1,  1,
     0,  -i
matrix w3=orthogonalize(w2);  // Prozedur aus linalg.lib, die die
                              // Spalten orthogonalisiert.
print(w3); ↦    1,  -1/2,
                1,  1/2,
                0,  -i
matrix T=concat(w1,w3);
T;
↦    1,   1,  -1/2,  // Die Transformationsmatrix.
     -1,  1,  1/2,
     -i,  0,  -i
matrix U=inverse(T);
print(U*A*T);
↦    2,  0,   0,
     0,  -1,  0,
     0,  0,   -1
```

A ist also konjugiert zu der obigen Diagonalmatrix und T ist eine Matrix mit paarweise orthogonalen Spalten, die A diagonalisiert.

Man bezeichnet Korollar 3.7 auch als reelle bzw. komplexe *Hauptachsentransformation*. Wir formulieren die Aussage nun noch einmal etwas anders und ergänzen sie etwas.

Satz 3.11 (Hauptachsentransformation)
Sei $\mathbb{K}^n$ versehen mit dem kanonischen Skalarprodukt, sei ferner $A \in \mathrm{Mat}(n,\mathbb{K})$ eine symmetrische bzw. hermitesche Matrix und $b = b_A$ die durch A beschriebene Bilinearform bzw. Sesquilinearform. Dann gelten:

a. *Ist* $B = (x_1, \dots, x_n)$ *eine ONB von* $\mathbb{K}^n$ *aus Eigenvektoren von* A*, so gilt*

$$M_B(b) = \begin{pmatrix} \lambda_1 & 0 & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & \lambda_n \end{pmatrix},$$

d. h. $b(x_i, x_j) = \lambda_i \delta_{ij}$ *für* $i, j = 1, \dots, n$*, wobei* $\lambda_i \in \mathbb{R}$ *der zu* x_i *gehörende Eigenwert von* A *ist,* $i = 1, \dots, n$.
D. h. es gibt ein $T \in \mathrm{SO}(n)$ *bzw.* $T \in \mathrm{SU}(n)$ *mit*

$$M_B(b) = T^t \circ A \circ \overline{T} = \overline{T^{-1}} \circ A \circ \overline{T}.$$

b. *Es existiert eine OGB* B' *von* $\mathbb{K}^n$ *mit*

$$M_{B'}(b) = \mathbb{1}_k \oplus -\mathbb{1}_l \oplus 0_{n-k-l},$$

wobei $0_{n-k-l} \in \mathrm{Mat}(n-k-l, \mathbb{K})$ *die Nullmatrix bezeichnet.*
Also existiert ein $T' \in \mathrm{Gl}_n(\mathbb{K})$ *mit paarweise orthogonalen Spalten und*

$$M_{B'}(b) = T'^t \circ A \circ \overline{T'} = \mathbb{1}_k \oplus -\mathbb{1}_l \oplus 0_{n-k-l}.$$

Bemerkung 3.12
Der außerordentlich wichtige Satz besagt, daß jede reelle symmetrische Matrix A mittels einer Koordinatentransformation $T \in \mathrm{SO}(n)$ diagonalisierbar ist und daß gleichzeitig mit derselben Koordinatentransformation die durch A beschriebene Bilinearform diagonalisiert wird.

Man beachte, daß die Eigenwerte $\lambda_1, \dots, \lambda_n$ nur Invarianten des Endomorphismus sind, aber *nicht* der Bilinearform. Das ergibt sich schon aus Aussage b., die besagt, daß wir die Bilinearform durch eine andere Koordinatentransformation auf noch einfachere Gestalt bringen können, nämlich so, daß in der Diagonalen nur die Werte 0, 1 und -1 auftauchen. Das geht allerdings i. a. nur, wenn wir darauf verzichten, daß die Koordinatentransformation orthogonal ist. In diesem Falle gilt *nicht* mehr $T^t = T^{-1}$ und daher beschreibt die Matrix $T^t \circ A \circ T$ auch nicht mehr den Endomorphismus A.

Die entsprechende Aussagen für hermitesche Matrizen gelten natürlich auch.

Beweis von Satz 3.11: a. Nach Korollar 3.7 wissen wir, daß A mittels einer Matrix $S \in \mathrm{O}(n)$ bzw. $S \in \mathrm{U}(n)$ diagonalisiert werden kann. Im Falle $z := \det(S) \neq 1$ multiplizieren wir die erste Spalte von S mit $\frac{1}{z} = \bar{z}$, dann ist die konjugiert Komplexe dieser Matrix eine Matrix T in $\mathrm{SO}(n)$ bzw. $\mathrm{SU}(n)$, die das Gewünschte leistet.

b. Wir numerieren die Eigenwerte so, daß $\lambda_1, \dots, \lambda_k > 0$, $\lambda_{k+1}, \dots, \lambda_{k+l} < 0$ und $\lambda_{k+l+1}, \dots, \lambda_n = 0$ ist. Dann ersetzen wir die ONB $B = (x_1, \dots, x_n)$ aus

Teil a. durch $B' = (x'_1, \ldots, x'_n)$ mit

$$x'_i := \begin{cases} \frac{1}{\sqrt{|\lambda_i|}} x_i, & i = 1, \ldots, k+l, \\ x_i, & i = k+l+1, \ldots, n. \end{cases}$$

Dann gilt $b(x'_i, x'_j) = 0$ für $i \neq j$ sowie

$$b(x'_i, x'_i) = \begin{cases} 1, & i = 1, \ldots, k, \\ -1, & i = k+1, \ldots, k+l, \\ 0, & i = k+l+1, \ldots, n. \end{cases}$$

□

Wir werden am Schluß des Abschnittes eine geometrische Anwendung der Hauptachsentransformation besprechen, die auch den Namen erklärt.

Obwohl die Eigenwerte $\lambda_1, \ldots, \lambda_n$ von A keine Invarianten der Bilinearform b_A sind, sind die in 3.11 b. auftretenden Zahlen k und l sehr wohl Invarianten von b_A. Genauer gesagt, gilt der folgende Sylvestersche Trägheitssatz.

Satz 3.13 (Sylvesterscher Trägheitssatz)
Ist $A \in \mathrm{Mat}(n, \mathbb{K})$ *eine symmetrische bzw. hermitesche Matrix und sind* $T, T' \in \mathrm{Gl}_n(\mathbb{K})$ *und* $k, k', l, l' \in \mathbb{N}$ *so, daß* $T^t \circ A \circ \overline{T} = \mathbb{1}_k \oplus -\mathbb{1}_l \oplus 0_{n-k-l}$ *und* $T'^t \circ A \circ \overline{T'} = \mathbb{1}_{k'} \oplus -\mathbb{1}_{l'} \oplus 0_{n-k'-l'}$*, dann gilt* $k = k'$*,* $l = l'$ *und* $k + l = \mathrm{rang}(A)$.

Man nennt dann k *auch den* Index *oder* Trägheitsindex *von* b_A *oder von* A*,* l *den* Morseindex *und* $k - l$ *die* Signatur.

Beweis: Da T und T' invertierbar sind, ist klar, daß

$$k + l = \mathrm{rang}\left(T^t \circ A \circ \overline{T}\right) = \mathrm{rang}(A) = \mathrm{rang}\left(T'^t \circ A \circ \overline{T'}\right) = k' + l'.$$

Es reicht also, zu zeigen, daß $k = k'$, und dazu reicht es, die folgende Behauptung zu zeigen, da die rechte Seite dort nur von A abhängt, nicht aber von T.

Behauptung: $k = \max\left\{ \dim_K(U) \mid U \subseteq \mathbb{K}^n \text{ Unterraum } : b_A(x, x) > 0 \ \forall\ 0 \neq x \in U \right\}$.

Wir bezeichnen die Spalten von T mit $x_1, \ldots, x_n$. Sie bilden eine Basis von $\mathbb{K}^n$. Da für $0 \neq x = \sum_{i=1}^{k} \lambda_i x_i \in \langle x_1, \ldots, x_k \rangle$

$$b_A(x, x) = \sum_{i,j=1}^{k} \lambda_i \overline{\lambda_j} b_A(x_i, x_j) = \sum_{i=1}^{k} |\lambda_i|^2 b_A(x_i, x_i) = \sum_{i=1}^{k} |\lambda_i|^2 > 0 \qquad (63)$$

gilt, folgt schon einmal "$\leq$".

Sei U irgendein Unterraum, so daß für alle $0 \neq x \in U$ gilt $b_A(x, x) > 0$. Wir setzen $W := \langle x_{k+1}, \ldots, x_n \rangle$. Analog zu (63) sieht man, daß für $x \in W$ gilt $b_A(x, x) \leq 0$. Mithin gilt offenbar $W \cap U = \{0\}$ und damit

$$\dim_{\mathbb{R}}(U) = \dim_{\mathbb{R}}(U + W) - \dim_{\mathbb{R}}(W) + \dim_{\mathbb{R}}(U \cap W) \leq n - (l + n - k - l) = k.$$

Damit ist auch "$\geq$" gezeigt. □

Korollar 3.14
Es sei $A \in \mathrm{Mat}(n, \mathbb{K})$ *eine symmetrische bzw. hermitesche Matrix mit Trägheitsindex* k *und Morseindex* l.

a. k *ist die maximale Dimension eines Unterraumes* U *von* $\mathbb{K}^n$*, so daß* b_A *eingeschränkt auf* U *positiv definit ist.*
b. l *ist die maximale Dimension eines Unterraumes* U *von* $\mathbb{K}^n$*, so daß* b_A *eingeschränkt auf* U *negativ definit ist.*
c. *Ist* $T \in \mathrm{Gl}_n(\mathbb{K})$ *so, daß* $T^t \circ A \circ \overline{T}$ *eine Diagonalmatrix ist, dann ist* k *genau die Anzahl der positiven Elemente auf der Diagonalen und* l *die Anzahl der negativen.*

Beweis: a. Die Aussage wurde im Beweis von Satz 3.13 als Behauptung formuliert und bewiesen.
b. Da eine Einschränkung von b_A genau dann negativ definit ist, wenn die entsprechende Einschränkung von b_{-A} positiv definit ist, folgt die Aussage aus a. und daraus, daß der Morseindex von A gerade der Trägheitsindex von $-A$ ist.
c. Wie im Beweis von Satz 3.13 läßt sich die Diagonalmatrix $T^t \circ A \circ \overline{T}$ in eine Diagonalmatrix mit den Werten 0, 1 und -1 auf der Diagonalen überführen, wobei die Anzahl der Einsen, d. h. der Trägheitsindex, genau der Anzahl der positiven Werte auf der Diagonalen von $T^t \circ A \circ \overline{T}$ entspricht. Entsprechendes gilt für l.

□

Eine symmetrische Bilinearform b auf $\mathbb{R}^n$ ist genau dann ein Skalarprodukt, wenn sie positiv definit ist. Es ist daher außerordentlich nützlich, Kriterien für die positive Definitheit zur Hand zu haben.

Definition 3.15
Ist R ein kommutativer Ring mit Eins, $A \in \mathrm{Mat}(n \times m, R)$ und $\tilde{A} \in \mathrm{Mat}(k, R)$ eine $k \times k$-Untermatrix von A, dann nennt man $\det(\tilde{A})$ einen k-*Minor* von A.

Ist $A \in \mathrm{Mat}(n, R)$ und entsteht die $k \times k$-Untermatrix $A(k)$ von A durch Streichen der letzten $n - k$ Zeilen und Spalten, so nennen wir $A(k)$ die k-te *Hauptmatrix* von A und $\det(A(k))$ den k-ten *Hauptminor* von A.

Satz 3.16
Für eine symmetrische bzw. hermitesche Matrix $A \in \mathrm{Mat}(n, \mathbb{K})$ *sind die folgenden Aussagen gleichwertig:*

a. A *ist positiv definit.*
b. *Es existiert ein* $W \in \mathrm{Gl}_n(\mathbb{K})$ *mit* $A = W \circ W^*$.
c. *Es existiert ein* $T \in \mathrm{O}(n)$ *bzw.* $T \in \mathrm{U}(n)$*, so daß* $T^t \circ A \circ \overline{T}$ *eine Diagonalmatrix mit positiven Diagonalelementen ist.*
d. *Alle Eigenwerte von* A *sind positiv.*
e. A *ist invertierbar und* A^{-1} *ist positiv definit.*

f. *Alle Hauptminoren von* A *sind positiv.*
g. *Der Trägheitsindex von* A *ist* n.
h. *Die Signatur von* A *ist* n.

Beweis: Wir schieben zwei Bemerkungen vorweg.

(i) Ist $W \in \mathrm{Gl}_n(\mathbb{K})$ beliebig, dann gilt:

$$\begin{aligned} A \text{ ist positiv definit.} &\Leftrightarrow x^t \circ A \circ \overline{x} > 0 \ \forall\, x \in \mathbb{K}^n \\ &\Leftrightarrow x^t \circ \left(W^t \circ A \circ \overline{W}\right) \circ \overline{x} = (Wx)^t \circ A \circ \left(\overline{Wx}\right) > 0 \ \forall\, x \in \mathbb{K}^n \\ &\Leftrightarrow W^t \circ A \circ \overline{W} \text{ ist positiv definit.} \end{aligned}$$

(ii) Ist $D = \bigoplus_{i=1}^n \lambda_i \mathbb{1}_1$ eine Diagonalmatrix mit $\lambda_i \in \mathbb{R}$, dann ist D positiv definit genau dann, wenn $\lambda_i > 0$ für alle $i = 1, \ldots, n$.
Denn für $x = (x_1, \ldots, x_n)^t \in \mathbb{K}^n$ gilt

$$x^t \circ D \circ \overline{x} = \sum_{i=1}^n \lambda_i |x_i|^2,$$

so daß die Rückrichtung der Behauptung klar ist und die Hinrichtung durch Einsetzen von $e_1, \ldots, e_n$ für x folgt.

Damit können wir den eigentlichen Beweis nun beginnen.

Die Äquivalenz von a. und c. folgt unter Berücksichtigung von (ii) aus Korollar 3.7. Die Äquivalenz von c. und d. folgt ebenfalls aus Korollar 3.7, da die Diagonalelemente von $T^t \circ A \circ \overline{T}$ gerade die Eigenwerte von A sind - wobei man berücksichtigt, daß diese reell sind, da es eigentlich nur die komplex Konjugierten der Eigenwerte sind.

c. $\Rightarrow$ b.: Wegen $T^t \circ A \circ \overline{T} = \bigoplus_{i=1}^n \lambda_i \mathbb{1}_1 =: D$ mit $\lambda_i \in \mathbb{R}_{>0}$ können wir die Diagonalmatrix

$$\sqrt{D} = \bigoplus_{i=1}^n \sqrt{\lambda_i}\, \mathbb{1}_1$$

definieren und dann gilt mit $W = \overline{T} \circ \sqrt{D} \in \mathrm{Gl}_n(\mathbb{K})$

$$A = \overline{T} \circ D \circ T^t = \left(\overline{T} \circ \sqrt{D}\right) \circ \left(\sqrt{D} \circ T^t\right) = \left(\overline{T} \circ \sqrt{D}\right) \circ \left(\overline{T} \circ \sqrt{D}\right)^* = W \circ W^*.$$

b. $\Rightarrow$ a.: Da $\mathbb{1}_n$ positiv definit ist, folgt die Behauptung aus (i).

a. $\Leftrightarrow$ e.: Wir können ausnutzen, daß a. und b. äquivalent sind. Für $W \in \mathrm{Gl}_n(\mathbb{K})$ gilt offenbar $\left(W^*\right)^{-1} = \left(W^{-1}\right)^*$.

Ist nun A positiv definit, dann ist $A = W \circ W^*$, für ein $W \in \mathrm{Gl}_n(\mathbb{K})$, nach b. invertierbar und $A^{-1} = \left(W^*\right)^{-1} \circ W^{-1} = \left(W^*\right)^{-1} \circ \left(\left(W^*\right)^{-1}\right)^*$ ist wiederum nach b. positiv definit.

Ist umgekehrt A^{-1} positiv definit, so folgt analog mit Hilfe von b., daß A positiv definit ist.

g. $\Leftrightarrow$ h.: Klar.

a. $\Leftrightarrow$ g.: Nach dem Hauptachsentransformationssatz 3.11 gibt es ein $W \in \mathrm{Gl}_n(\mathbb{K})$ mit $D := W^t \circ A \circ \overline{W} = \mathbb{1}_k \oplus \mathbb{1}_l \oplus 0_{n-k-l}$, wobei k der Trägheitsindex von A ist. Wegen (i) und (ii) gilt nun, daß A genau dann positiv definit ist, wenn $k = n$.

a. $\Rightarrow$ f.: Die k-te Hauptmatrix $A(k)$ beschreibt die Einschränkung der durch A definierten Bilinearform b_A auf den Unterraum $\langle e_1, \dots, e_k\rangle_{\mathbb{K}} \subseteq \mathbb{K}^n$. Da die Einschränkung einer positiv definiten Bilinearform offenbar wieder positiv definit ist, muß mithin auch die definierende Matrix $A(k)$ positiv definit sein.

f. $\Rightarrow$ a.: Wir führen den Beweis durch Induktion über n unter Ausnutzung der bereits gezeigten Äquivalenzen, wobei für $n = 1$ nichts zu zeigen ist.

Sei also $n > 1$. Nach Induktionsvoraussetzung existiert eine Matrix $T' \in \mathrm{Gl}_{n-1}(\mathbb{K})$, die $A(n-1)$ auf Diagonalgestalt mit positiven Diagonaleinträgen transformiert:

$$T'^t \circ A(n-1) \circ \overline{T'} = \bigoplus_{i=1}^{n-1} \lambda_i \mathbb{1}_1 =: D.$$

Wir setzen nun $T = T' \oplus \mathbb{1}_1 \in \mathrm{Gl}_n$. Dann gilt

$$T^t \circ A \circ \overline{T} = \left(\begin{array}{ccc|c} & D & & \overline{a_1} \\ & & & \vdots \\ & & & \overline{a_{n-1}} \\ \hline a_1 & \dots & a_{n-1} & a_n \end{array}\right) =: \tilde{A}$$

für geeignete $a_1, \dots, a_{n-1} \in \mathbb{K}$, $a_n \in \mathbb{R}$.

Setzen wir nun $c_j = -\frac{a_j}{\lambda_j}$, $j = 1, \dots, n-1$, und

$$C = \left(\begin{array}{ccc|c} & & & c_1 \\ & \mathbb{1}_{n-1} & & \vdots \\ & & & c_{n-1} \\ \hline 0 & \dots & 0 & 1 \end{array}\right) \in \mathrm{Gl}_n(\mathbb{K}),$$

dann folgt

$$(T \circ C)^t \circ A \circ \overline{T \circ C} = C^t \circ \tilde{A} \circ \overline{C} = \bigoplus_{i=1}^{n} \lambda_i \mathbb{1}_1.$$

Dabei ist

$$\lambda_1 \cdots \lambda_n = \det\left((T \circ C)^t \circ A \circ \overline{T \circ C}\right) = \det(A) \cdot |\det(T \circ C)|^2 > 0,$$

da $\det(A) > 0$ der n-te Hauptminor von A ist. Da aber $\lambda_1, \dots, \lambda_{n-1}$ nach Voraussetzung positiv waren, ist dann auch λ_n positiv.

Aus (ii) folgt somit, daß $(T \circ C)^t \circ A \circ \overline{T \circ C}$ positiv definit ist, aber dann ist mit (i) auch A positiv definit. $\square$

Wir kommen jetzt zur geometrischen Interpretation der Hauptachsentransformation. Die geometrischen Objekte sind Ellipsen und Hyperbeln, oder allgemeiner, die Nullstellengebilde eines Polynoms zweiten Grades in zwei

Veränderlichen. Der Satz über die Hauptachsentransformation besagt, daß z. B. jede Ellipse oder Hyperbel durch Translation und Drehung so bewegt werden kann, daß die Hauptachsen mit den Koordinatenachsen übereinstimmen.

Zunächst benötigen wir aber noch einige neue Begriffe.

Definition 3.17

Es sei $\big(V, \langle\cdot,\cdot\rangle\big)$ ein euklidischer Raum.

a. Eine Abbildung $f : V \to V$ heißt eine *affine Abbildung* auf V, falls es ein $y \in V$ gibt und ein $g \in \mathrm{End}_{\mathbb{R}}(V)$ mit $f(x) = y + g(x)$ für alle $x \in V$.

b. Für $y \in V$ nennen wir die affine Abbildung

$$t_y : V \to V : x \mapsto x + y$$

die *Translation* um den Vektor y.

c. Eine Abbildung $f : V \to V$ heißt eine *Ähnlichkeit*, wenn es einen Vektor $y \in V$ gibt und eine orthogonale Abbildung $g \in O(V)$ mit $f = t_y \circ g$, d. h.

$$f(x) = t_y\big(g(x)\big) = y + g(x) \quad \forall\, x \in V.$$

d. Ist $V = \mathbb{K}^n$, $\langle\cdot,\cdot\rangle$ das Standardskalarprodukt und sei $f = t_y \circ g$ mit $g \in \mathrm{Gl}\left(\mathbb{R}^n\right)$ eine bijektive affine Abbildung auf V, dann nennen wir die induzierte Abbildung

$$\mathbb{R}[x_1, \ldots, x_n] \to \mathbb{R}[x_1, \ldots, x_n] : p \mapsto p\big(f(x_1, \ldots, x_n)\big)$$

einen *affinen Koordinatenwechsel* von $\mathbb{R}[x_1, \ldots, x_n]$.

Bemerkung 3.18 a. Jede affine Abbildung $f : \mathbb{R}^n \to \mathbb{R}^n$ läßt sich offenbar in eindeutiger Weise schreiben, als $f = t_y \circ g$ mit $y = f(0) \in V$ und $g \in \mathrm{End}_{\mathbb{R}}(V)$.

b. Ist $f = t_y \circ g : \mathbb{R}^n \to \mathbb{R}^n$ eine affine Abbildung mit $y \in \mathbb{R}^n$ und $g \in \mathrm{Gl}\left(\mathbb{R}^n\right)$, dann gibt es eine eindeutig bestimmte Matrix $T \in \mathrm{Gl}_n(\mathbb{R})$ mit $g = f_T$. Damit gilt für $p \in \mathbb{R}[x_1, \ldots, x_n]$ und $x = (x_1, \ldots, x_n)^t$

$$p\big(f(x_1, \ldots, x_n)\big) = p(Tx + y).$$

Ist beispielsweise $p = x_1^2 + 3x_2 - 1 \in \mathbb{R}[x_1, x_2]$, $T = T\left(\frac{\pi}{2}\right)$ die Drehung um 90° und $y = (2, -2)$, dann ist für $f = t_y \circ f_T$

$$p\big(f(x_1, x_2)\big) = p(-x_2 + 2, x_1 - 2) = (-x_2 + 2)^2 + 3(x_1 - 2) - 1.$$

In Aufgabe III.1.55 haben wir den Polynomring in mehreren Veränderlichen eingeführt. Wir wollen nun die *Nullstellengebilde* von Polynomen in mehreren Veränderlichen definieren.

Definition 3.19

Es sei $p \in \mathbb{R}[x_1, \ldots, x_n]$ dann nennen wir die Menge

$$N(p) = \left\{\lambda = (\lambda_1, \ldots, \lambda_n)^t \in \mathbb{R}^n \;\middle|\; p(\lambda) = 0\right\}$$

eine *algebraische Hyperfläche* von $\mathbb{R}^n$.

Ist $\deg(p) = d$, so nennen wir d auch den *Grad* der Hyperfläche.

Ist $n = 2$, so sprechen wir auch von *algebraischen Kurven* statt von algebraischen Hyperflächen.

Definition 3.20

Wir definieren auf $\mathbb{R}[x_1, \ldots, x_n]$ eine Relation durch

$$p \equiv q \quad :\Leftrightarrow \quad \exists\, c \in \mathbb{R}^* : p = c \cdot q$$

für $p, q \in \mathbb{R}[x_1, \ldots, x_n]$. Wir nennen p und q mit $p \equiv q$ auch *äquivalent*.

Bemerkung 3.21

Man sieht sofort, daß $\equiv$ eine Äquivalenzrelation auf $\mathbb{R}[x_1, \ldots, x_n]$ definiert.

Ferner gilt offensichtlich, daß für zwei äquivalente Polynome $p, q \in \mathbb{R}[x_1, \ldots, x_n]$ auch $N(p) = N(q)$ gilt. Interessiert man sich also nur für das Nullstellengebilde von p, so kann man p getrost durch ein äquivalentes Polynom ersetzen und somit erreichen, daß der konstante Anteil von p entweder 0 oder -1 ist.

Im Folgenden interessieren wir uns nur noch für algebraische Kurven vom Grad zwei.

Bemerkung 3.22

Ist $p \in \mathbb{R}[x_1, x_2]$ ein allgemeines Polynom zweiten Grades, dann gibt es reelle Zahlen $\alpha_{11}, \alpha_{12} = \alpha_{21}, \alpha_{22}, \alpha_1, \alpha_2, \alpha \in \mathbb{R}$ so, daß

$$p = \alpha_{11}x_1^2 + 2\alpha_{12}x_1x_2 + \alpha_{22}x_2^2 + \alpha_1x_1 + \alpha_2x_2 + \alpha = \langle x, Sx\rangle + \langle a, x\rangle + \alpha,$$

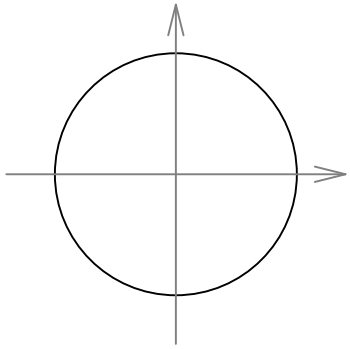
wobei $x = (x_1, x_2)^t$, $0 \neq S = (\alpha_{ij})_{i,j\in\{1,2\}} \in \mathrm{Mat}(2, \mathbb{R})$ und $a = (\alpha_1, \alpha_2)^t$.

Beispiel 3.23

Für $S = \mathbb{1}_2$, $a = (0,0)^t$ und $\alpha = -1$ erhalten wir $p = x_1^2 + x_2^2 - 1$, und die Nullstellenmenge davon,

$$N\left(x_1^2 + x_2^2 - 1\right) = \left\{\lambda = (\lambda_1, \lambda_2)^t \in \mathbb{R}^2 \;\middle|\; \lambda_1^2 + \lambda_2^2 = 1\right\},$$

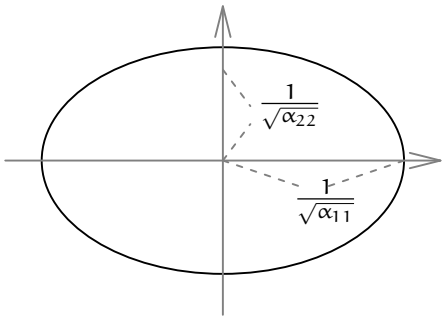
ist offenbar der Einheitskreis.



Ist S eine Diagonalmatrix mit positiven Diagonaleinträgen, d. h. $\alpha_{11}, \alpha_{22} > 0$ und $\alpha_{12} = \alpha_{21} = 0$, und ist ferner $a = (0,0)^t$ und $\alpha = -1$, dann erhalten wir als Nullstellengebilde von p

$$N\left(\left(\sqrt{\alpha_{11}}x_1\right)^2 + \left(\sqrt{\alpha_{22}}x_2\right)^2 - 1\right) = \left\{(\lambda_1, \lambda_2)^t \in \mathbb{R}^2 \;\middle|\; \left(\sqrt{\alpha_{11}}\lambda_1\right)^2 + \left(\sqrt{\alpha_{22}}\lambda_2\right)^2 = 1\right\}$$

eine Ellipse.



Satz 3.24
Es sei

$$p = \langle x, Sx\rangle + \langle a, x\rangle + \alpha \in \mathbb{R}[x_1, x_2] \tag{64}$$

ein Polynom zweiten Grades mit symmetrischer Matrix $0 \neq S = (\alpha_{ij}) \in \mathrm{Mat}(2,\mathbb{R})$. *Dann gibt es eine affine Koordinatentransformation mittels einer Ähnlichkeit* $f = t_y \circ f_T$ *von* $\mathbb{R}^2$ *mit* $T \in \mathrm{SO}(2)$, *so daß* $q := p\big(f(x_1,x_2)\big)$ *äquivalent zu einer der folgenden Normalformen ist:*

I: $\det(S) > 0$.
- I.1: $\alpha \neq 0$ *und* $\alpha_{11} > 0$. *Dann ist* $q \equiv (\lambda_1 x_1)^2 + (\lambda_2 x_2)^2 - 1$ *und* $N(q)$ *ist eine* Ellipse.
- I.2: $\alpha \neq 0$ *und* $\alpha_{11} < 0$. *Dann ist* $q \equiv (\lambda_1 x_1)^2 + (\lambda_2 x_2)^2 + 1$ *und* $N(q)$ *ist die leere Menge.*
- I.3: $\alpha = 0$. *Dann ist* $q \equiv (\lambda_1 x_1)^2 + (\lambda_2 x_2)^2$ *und* $N(q)$ *ist ein Punkt.*

II: $\det(S) < 0$.
- II.1: $\alpha \neq 0$. *Dann ist* $q \equiv (\lambda_1 x_1)^2 - (\lambda_2 x_2)^2 - 1$ *und* $N(q)$ *ist eine* Hyperbel.
- II.2: $\alpha = 0$. *Dann ist* $q \equiv (\lambda_1 x_1)^2 - (\lambda_2 x_2)^2$ *und* $N(q)$ *besteht aus zwei verschiedenen Geraden durch den Ursprung.*

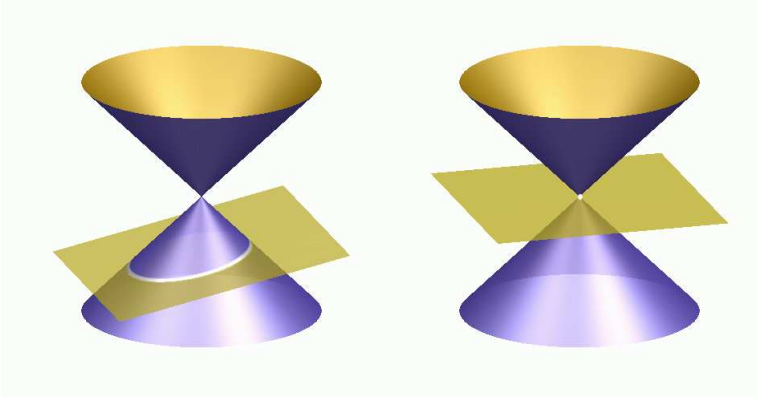
III: $\det(S) = 0$, $a \neq (0,0)^t$. *Dann ist* $q \equiv x_1^2 - \lambda x_2$ *und* $N(q)$ *ist eine* Parabel.

IV: $\det(S) = 0$, $a = (0,0)^t$.
- IV.1: $\alpha \neq 0$ *und* S *hat einen positiven Eigenwert. Dann ist* $q \equiv x_1^2 - \lambda$, $\lambda > 0$, *und* $N(q)$ *besteht aus zwei parallelen Geraden.*
- IV.2: $\alpha \neq 0$ *und* S *hat einen negativen Eigenwert. Dann ist* $q \equiv x_1^2 + \lambda$, $\lambda > 0$, *und* $N(q)$ *ist die leere Menge.*
- IV.3: $\alpha = 0$. *Dann ist* $q \equiv x_1^2$ *und* $N(q)$ *besteht aus einer* Doppelgraden, *d. h. einer Geraden, die man doppelt zählt.*

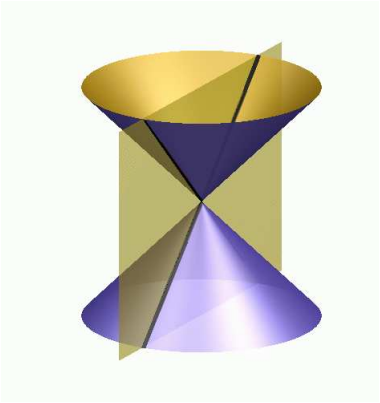
Bemerkung 3.25
Dies ist die vollständige Klassifikation der Kurven zweiten Grades. Sie heißen auch *Kegelschnitte*, da alle, bis auf die Fälle I.2, IV.1 und IV.2 als Schnitt des Kreiskegels

$$N\big(x_1^2 + x_2^2 - x_3^2\big) \subset \mathbb{R}^3$$

mit einer geeigneten Ebene im $\mathbb{R}^3$ realisierbar sind.



I.1: Ellipse I.3: Punkt II.1: Hyperbel



II.2: Zwei Geraden mit Schnitt III: Parabel IV.3: Doppelgerade

I.1 besagt, daß sich jede Ellipse durch Translation und Drehung so bewegen läßt, daß die Hauptachsen der Ellipse mit den Koordinatenachsen übereinstimmen. Daher kommt der Name Hauptachsentransformation.

Beweis von Satz 3.24: 1. Fall: $a = (0,0)^t$: Wir betrachten zunächst den Fall $a = (0,0)^t$.

Nach dem Satz über die Hauptachsentransformation 3.7 existiert ein $T \in SO(2)$, so daß

$$T^t \circ S \circ T = T^{-1} \circ S \circ T = \begin{pmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{pmatrix}.$$

Man beachte noch, daß nicht beide Eigenwerte μ_1 und μ_2 null sein können, da $S \neq 0$. Also können wir o. E. annehmen, daß $\mu_1 \neq 0$ und daß $\mu_1 \geq \mu_2$ gilt, falls $\mu_2 \neq 0$.

Die lineare Abbildung $f_T : \mathbb{R}^2 \to \mathbb{R}^2 : x \mapsto Tx$ ist eine Drehung und es gilt

$$\begin{aligned} p(Tx) &= \langle Tx, (S \circ T)x \rangle + \alpha \\ &= \langle x, (T^t \circ S \circ T)x \rangle + \alpha \\ &= \mu_1 x_1^2 + \mu_2 x_2^2 + \alpha. \end{aligned}$$

Da wir p ohnehin nur bis auf Äquivalenz klassifizieren wollen, können wir o. E. annehmen, daß $\alpha = 0$ oder $\alpha = -1$ gilt. Setzen wir nun noch $\lambda_i = \sqrt{|\mu_i|}$, dann erhalten wir folgende Fälle.

Fall 1.1: $\mu_1, \mu_2 > 0$: Dies ist gleichbedeutend dazu, daß S positiv definit ist, und nach dem Hauptminorenkriterium dazu, daß $\det(S) > 0$ und $\alpha_{11} > 0$. Ist $\alpha = -1$, so sind wir im Fall I.1, und ist $\alpha = 0$, so sind wir Fall I.3.

Fall 1.2: $\mu_1, \mu_2 < 0$: Dies ist gleichbedeutend dazu, daß $-S$ positiv definit ist, daß also $\det(S) = \det(-S) > 0$ und $-\alpha_{11} > 0$. Ist $\alpha = -1$, so sind wir im Fall I.2, und für $\alpha = 0$ wieder im Fall I.3, da wir dann das Polynom nochmals mit -1 multiplizieren können, um ein äquivalentes der gesuchten Form zu erhalten.

Fall 1.3: $\mu_1 > 0, \mu_2 < 0$: Dies ist gleichbedeutend dazu, daß $\mu_1 \cdot \mu_2 = \det(S) < 0$ ist. Im Fall $\alpha = -1$ führt dies zu Fall II.1, und im Fall $\alpha = 0$ führt es zu Fall II.2.

Fall 1.4: $\mu_1 > 0, \mu_2 = 0$ oder $\mu_1 < 0, \mu_2 = 0$: Das ist dann gleichbedeutend dazu, daß $\det(S) = 0$ ist. Für $\mu_1 > 0$ und $\alpha = -1$ erhalten wir Fall IV.1, für $\mu_1 < 0$ und $\alpha = -1$ den Fall IV.2, und für $\alpha = 0$ in den Fall IV.3.

2. Fall: $a \neq (0,0)^t$: Sind wir im Fall $a = (0,0)^t$ noch ohne Translation ausgekommen, so werden wir jetzt doch Translationen betrachten müssen.

Für $c \in \mathbb{R}^2$ bewirkt die Translation $t_c : \mathbb{R}^2 \to \mathbb{R}^2 : x \mapsto x + c$ folgende Koordinatentransformation für p

$$\begin{aligned} p(x+c) &= \langle x+c, Sx+Sc\rangle + 2\langle a, x+c\rangle + \alpha \\ &= \langle x, Sx\rangle + 2\langle a+Sc, x\rangle + \langle c, Sc\rangle + 2\langle a, c\rangle + \alpha \\ &= \langle x, Sx\rangle + 2\langle b, x\rangle + \beta, \end{aligned} \tag{65}$$

wenn wir $b = a + Sc$ und $\beta = \langle c, Sc\rangle + 2\langle a, c\rangle + \alpha$ setzen.

Fall 2.1: $\exists\, c \in \mathbb{R}^2 : b = a + Sc = (0,0)^t$: Dann haben wir p durch $p\big(t_c(x)\big)$ auf den ersten Fall "$a = (0,0)^t$" zurückgeführt. Es gibt also ein $T \in \mathrm{SO}(2)$, so daß $q = p\big((t_c \circ f_T)(x)\big)$ äquivalent zu einem der Fälle I, II oder IV ist.

Fall 2.2: $\forall\, c \in \mathbb{R}^2 : b = a + Sc \neq (0,0)^t$: Aus Lemma 3.26 folgt, daß es ein $c \in \mathbb{R}^2$ gibt mit $Sb = S^2c + Sa = 0$. Setzen wir nun noch $\delta := -\frac{\beta}{2\langle b,b\rangle}$, dann gilt für die Translation $t_{c+\delta b}$[8]

$$\begin{aligned} p(x+c+\delta b) &= \langle x, Sx\rangle + 2\big\langle a + S(c+\delta b), x\big\rangle + \big\langle c+\delta b, S(c+\delta b)\big\rangle + 2\langle a, c+\delta b\rangle + \alpha \\ &= \langle x, Sx\rangle + 2\langle b + \delta Sb, x\rangle + \delta^2\langle b, Sb\rangle + 2\delta\langle b, b\rangle + \beta \\ &= \langle x, Sx\rangle + 2\langle b, x\rangle + 2\delta\langle b, b\rangle + \beta \\ &= \langle x, Sx\rangle + 2\langle b, x\rangle. \end{aligned}$$

Beachtet man, daß, wegen $Sb = 0$, Null auf alle Fälle ein Eigenwert von S ist und daß $S \neq 0$, so folgt aus dem Satz über Hauptachsentransformation 3.7 die

[8] Man setze zunächst in der Gleichung (65) für c den Wert $c + \delta b$ ein. Dann ziehe man die Skalarprodukte auseinander und gruppiere sie neu, so daß man $b = a + Sc$, $Sb = 0$ sowie die Definition von β verwenden kann. Man beachte auch, daß S symmetrisch, also selbstadjungiert, ist.

Existenz eines $T \in SO(2)$, so daß

$$D := T^t \circ S \circ T = T^{-1} \circ S \circ T = \begin{pmatrix} \mu_1 & 0 \\ 0 & 0 \end{pmatrix},$$

wobei $\mu_1 \neq 0$. Insbesondere sind wir also in dem Fall $\det(S) = 0$.

Ferner gilt für $T^t b =: (\mu, \lambda)^t$ unter Berücksichtigung, daß $T^t = T^{-1}$,

$$(\mu_1 \mu, 0) = \big(T^t \circ S \circ T\big) \circ (T^t b) = T^t \circ (Sb) = 0,$$

und mithin ist $T^t b = (0, \lambda)^t$, wobei $\lambda \neq 0$, da T^t invertierbar und $b \neq (0,0)^t$. Aber dann überführt $x \mapsto Tx$ das Polynom $\langle x, Sx\rangle + 2\langle b, x\rangle$ in das Polynom

$$\big\langle Tx, (S \circ T)x\big\rangle + 2\big\langle b, Tx\big\rangle = \big\langle x^t, Dx\big\rangle + 2\big\langle T^t b, x\big\rangle = \mu_1 x_1^2 + 2\lambda x_2.$$

D. h. dann aber, daß

$$q := p\big((t_{c+\delta b} \circ f_T)(x)\big) = \mu_1 x_1^2 + 2\lambda x_2,$$

und damit sind wir genau im Fall III. □

Lemma 3.26
Ist $S \in \mathrm{Mat}(n, \mathbb{R})$ *symmetrisch, so gilt für die lineare Abbildung* $f_S : \mathbb{R}^n \to \mathbb{R}^n$.

a. $\mathrm{Ker}\left(f_S^2\right) = \mathrm{Ker}(f_S)$ *und* $\mathrm{Im}\left(f_S^2\right) = \mathrm{Im}(f_S)$.
b. *Zu jedem* $a \in \mathbb{R}^n$ *existiert ein* $c \in \mathbb{R}^n$*, so daß* $S^2 c + Sa = 0$.

Beweis: a. Für $x \in \mathrm{Ker}\left(f_S^2\right)$ ergibt sich aus

$$0 = \big\langle x, S^2 x\big\rangle = \langle Sx, Sx\rangle,$$

also $f_S(x) = Sx = 0$ und $x \in \mathrm{Ker}(f_S)$. Die umgekehrte Inklusion ist klar. Wir wissen bereits, daß $\mathrm{Im}(f_S) \supseteq \mathrm{Im}\left(f_S^2\right)$ gilt. Da nun ferner

$$\begin{aligned} \dim_{\mathbb{R}}\big(\mathrm{Im}(f_S)\big) &= n - \dim_{\mathbb{R}}\big(\mathrm{Ker}(f_S)\big) \\ &= n - \dim_{\mathbb{R}}\Big(\mathrm{Ker}\left(f_S^2\right)\Big) = \dim_{\mathbb{R}}\Big(\mathrm{Im}\left(f_S^2\right)\Big) \end{aligned}$$

gilt, folgt also die Gleichheit.

b. Es gilt für $a \in \mathbb{R}^n$, daß $S(-a) = f_S(-a) \in \mathrm{Im}(f_S) = \mathrm{Im}\left(f_S^2\right)$, also gibt es nach a. ein $c \in \mathbb{R}^n$ mit $S^2 c + Sa = f_S^2(c) - f_S(-a) = 0$.

□

Zum Schluß dieses Kapitels wollen wir noch einen weiteren Beweis der Existenz einer OGB eines Vektorraumes mit einer Bilinearform geben. Dieser Beweis hat den Vorteil, daß er für beliebige Körper K mit $\mathrm{char}(K) \neq 2$ gilt und gleichzeitig, im Fall $K = \mathbb{R}$, ein effizientes Verfahren zur Diagonalisierung einer symmetrischen Matrix als darstellende Matrix einer Bilinearform liefert - wohlgemerkt, nicht als darstellende Matrix eines Endomorphismus[9] - womit dann die positive Definitheit entschieden werden kann - siehe Satz 3.16.

[9] D. h. die Eigenwerte der Matrix werden sich ändern, da mit T durch $T^t \circ A \circ T$ transformiert wird und nicht durch $T^{-1} \circ A \circ T$.

In einem beliebigen Körper kann man nicht die Quadratwurzel ziehen (z. B. besitzt 2 in $\mathbb{Q}$ keine Quadratwurzel) und daher kann man nicht normieren, d. h. man kann i. a. keine ONB bestimmen. Die positive Kehrseite ist jedoch, daß alle Operationen, die man durchführt, auch wirklich im Grundkörper K verbleiben.

Satz 3.27
Es sei K ein beliebiger Körper mit $\mathrm{char}(K) \neq 2$ und V ein endlich-dimensionaler K-Vektorraum. Ist $b \in \mathrm{Bil}_K(V)$ eine symmetrische Bilinearform, dann besitzt V ein OGB $B = (x_1, \ldots, x_n)$, d. h.

$$b(x_i, x_j) = 0 \quad \textit{für alle } i \neq j.$$

Beweis: Wir führen den Beweis durch Induktion über $n = \dim_K(V)$, wobei im Fall $n = 1$ nichts zu zeigen ist.

Sei also $n > 1$. Wir bezeichnen mit $q : V \to K : x \mapsto b(x, x)$ die zu b gehörende quadratische Form. Ist q identisch Null, so ist nach Lemma 1.20 auch b identisch Null, da $\mathrm{char}(K) \neq 2$, und jede Basis ist eine OGB.

Wir können also annehmen, daß es ein $x \in V$ gibt mit $b(x, x) = q(x) \neq 0$. Setze $U := \langle x \rangle_K$ und

$$U^\perp := \{y \in V \mid b(x, y) = 0\}.$$

Aus der Bilinearität von b folgt, daß $U^\perp$ ein Unterraum von V ist. Wir wollen nun zeigen, daß in der Tat $V = U \oplus U^\perp$ gilt.

Sei dazu zunächst $y \in V$ beliebig. Dann setzen wir

$$x' := \tfrac{b(x,y)}{b(x,x)} x \in U$$

und erhalten $b\big(y - x', x\big) = b(y, x) - \frac{b(x,y)}{b(x,x)} b(x, x) = 0$, also $y - x' \in U^\perp$ und

$$y = x' + (y - x') \in U + U^\perp.$$

Also ist $V = U + U^\perp$.

Sei nun $y \in U \cap U^\perp$, dann gibt es ein $\lambda \in K$ mit $y = \lambda x$ und damit $\lambda q(x) = b(x, \lambda x) = b(x, y) = 0$. Da aber $q(x) \neq 0$, ist $\lambda = 0$ und damit $y = 0$. Also gilt $U \cap U^\perp = \{0\}$.

Schränken wir b auf $U^\perp$ ein, so erhalten wir per Induktion eine OGB $(x_2, \ldots, x_n)$ von $U^\perp$ und $B = (x, x_2, \ldots, x_n)$ ist dann die gesuchte OGB von V. $\square$

Korollar 3.28
Ist K *ein beliebiger Körper mit* $\mathrm{char}(K) \neq 2$ *und ist* $A \in \mathrm{Mat}(n, K)$ *eine symmetrische Matrix, dann gibt es eine Matrix* $T \in \mathrm{Gl}_n(K)$ *mit*

$$T^t \circ A \circ T = \begin{pmatrix} \lambda_1 & 0 & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & \lambda_n \end{pmatrix},$$

wobei $\lambda_1, \dots, \lambda_r \in K \setminus \{0\}$, $\lambda_{r+1} = \dots = \lambda_n = 0$ *und* $r = \mathrm{rang}(A)$.

Beweis: Nach Satz 3.27 gibt es zu b_A eine OGB $B = (x_1, \dots, x_n)$ von K^n. Dabei können wir durch Umnumerieren o. E. annehmen, daß $b(x_i, x_i) \neq 0$ für $i = 1, \dots, r$ und $b(x_i, x_i) = 0$ für $i = r+1, \dots, n$ für ein geeignetes r. Wähle nun T als die Matrix, deren Spalten die Vektoren in B sind, dann ist $M_B(b) = T^t \circ A \circ T$ und hat die gewünschte Gestalt, da B eine OGB ist.

Es bleibt zu zeigen, daß $r = \mathrm{rang}(A)$. Aber, da T invertierbar ist, gilt

$$r = \mathrm{rang}\left(T^t \circ A \circ T\right) = \mathrm{rang}(A).$$

□

Bemerkung 3.29
Man beachte, daß die λ_i i. a. nicht nur von A abhängen und auch nicht die Eigenwerte von A sind. Die Anzahl der Diagonalelemente ungleich Null hängt jedoch stets nur von A ab, und im Fall $K = \mathbb{R}$ haben wir bereits gesehen, daß auch die Anzahl der Elemente größer (Trägheitsindex) bzw. kleiner (Morseindex) Null nur von A abhängt.

Das Korollar 3.28 führt zu folgender Überlegung. Da $T \in \mathrm{Gl}_n(K)$ ist, ist T das Produkt von Elementarmatrizen $T = P_1 \circ \dots \circ P_k$ und somit gilt

$$D := T^t \circ A \circ T = P_k^t \circ \dots \circ P_1^t \circ A \circ P_1 \circ \dots \circ P_k.$$

Das heißt, daß die Diagonalmatrix D aus A durch gleichzeitiges Durchführen von elementaren Zeilen- und derselben Spaltenoperationen entsteht. Dabei ist es wegen $P^t \circ (A \circ P) = \left(P^t \circ A\right) \circ P$ egal, ob zuerst die Zeilenoperation oder die Spaltenoperation durchgeführt wird.

Die Überführung einer symmetrischen Matrix A in Diagonalgestalt mittels gleichzeitiger Zeilen- und Spaltenoperationen nennt man das *symmetrische Gaußsche Eliminationsverfahren* oder den *symmetrischen Gaußalgorithmus*.

Es ist klar, daß man diesen Algorithmus ebenso einfach implementieren kann, wie den Gaußschen Algorithmus. Will man zusätzlich die Transformationsmatrix T bestimmen, so startet man wie bei der Invertierung einer Matrix mit $(A \mid \mathbb{1}_n)$, führt bei A die Zeilen- und Spaltenoperationen durch, bei

$\mathbb{1}_n$ aber nur die Spaltenoperationen. Ist dann A diagonalisiert, so ist $\mathbb{1}_n$ in die Transformationsmatrix überführt.

Wir formulieren den Algorithmus nun in rekursiver Form. Die Eingabe muß dann ein Schema der Form $(A \mid \mathbb{1}_n)$ sein, damit die Ausgabe den gewünschten Erfolg hat.

Algorithmus 3.30 (Symmetrischer Gaußalgorithmus)

INPUT: A, T mit $A \in \mathrm{Mat}(n, K)$ symmetrisch und $T \in \mathrm{Gl}_r(K)$, $r \geq n$.

OUTPUT: $T \in \mathrm{Gl}_r(K)$ so, daß $\tilde{T}^t \circ A \circ \tilde{T}$ eine Diagonalmatrix ist, wobei $\tilde{T}$ durch Streichen der ersten $r - n$ Spalten und Zeilen aus T entsteht.

1. **Schritt:** Setze $m = r - n$.
2. **Schritt:** Man suche in der ersten Spalte von A den ersten Eintrag, der nicht Null ist. Existiert ein solcher, merke man sich die Zeilennummer z, sonst gehe man zu Schritt 5.
3. **Schritt:** Ist $z \neq 1$, so addiere die z-te Zeile von A zur ersten und die z-te Spalte zur ersten. Addiere ferner die $z + m$-te Spalte von T zur $m + 1$-ten Spalte.
4. **Schritt:** Für $k = 2, \ldots, n$ addiere man das $-A[1,k]/A[1,1]$-fache der $1 + m$-ten Spalte von T zur $k + m$-ten. Sodann addiere man das $-A[1,k]/A[1,1]$-fache der ersten Zeile von A zur k-ten und das $-A[1,k]/A[1,1]$-fache der ersten Spalte zur k-ten.
5. **Schritt:** Falls $n > 1$, dann erzeuge man eine Matrix B, indem man aus A die erste Zeile und die erste Spalte streicht. Sodann rufe man die Prozedur mit den Parametern B und T auf und speichere das Ergebnis in T.
6. **Schritt:** Man gebe T zurück.

Beispiel 3.31

Wir betrachten die Matrix

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \in \mathrm{Mat}(n, \mathbb{R}).$$

Sodann bilden wir das Schema $(A \mid \mathbb{1}_2)$ und wenden den symmetrischen Gaußalgorithmus an:

$$\left(\begin{array}{cc|cc} 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array}\right) \xrightarrow{\text{I=I+II}} \left(\begin{array}{cc|cc} 3 & 2 & 1 & 0 \\ 2 & 1 & 1 & 1 \end{array}\right) \xrightarrow{\text{II=II-}\frac{2}{3}\text{I}} \left(\begin{array}{cc|cc} 3 & 0 & 1 & -\frac{2}{3} \\ 0 & -\frac{1}{3} & 1 & \frac{1}{3} \end{array}\right).$$

Für $T = \begin{pmatrix} 1 & -\frac{2}{3} \\ 1 & \frac{1}{3} \end{pmatrix} \in \mathrm{Gl}_2(\mathbb{R})$ gilt also

$$T^t \circ A \circ T = \begin{pmatrix} 3 & 0 \\ 0 & -\frac{1}{3} \end{pmatrix}.$$

Alternativ können wir auch wie folgt vorgehen:

$$\left(\begin{array}{cc|cc} 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array}\right) \xrightarrow{\text{I}\leftrightarrow\text{II}} \left(\begin{array}{cc|cc} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{array}\right) \xrightarrow{\text{II=II-I}} \left(\begin{array}{cc|cc} 1 & 0 & 0 & 1 \\ 0 & -1 & 1 & -1 \end{array}\right).$$

Dann gilt für $S = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} \in \mathrm{Gl}_2(\mathbb{R})$

$$S^t \circ A \circ S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Die beiden Diagonalmatrizen besitzen nicht die gleichen Diagonaleinträge, aber wir sehen, daß ihr Trägheitsindex und ihr Morseindex übereinstimmen. Ferner sehen wir, da Trägheitsindex und Morseindex von A beide nur Eins sind, daß A weder positiv noch negativ definit ist.

AUFGABEN

Aufgabe 3.32

Bestimme eine orthogonale Matrix $T \in \mathrm{O}(3)$, die die folgende symmetrische Matrix A diagonalisiert:

$$A = \begin{pmatrix} 5 & 2 & 2 \\ 2 & 2 & -4 \\ 2 & -4 & 2 \end{pmatrix}.$$

Aufgabe 3.33

Es sei $V \neq 0$ ein endlich-dimensionaler unitärer Raum und $f \in \mathrm{End}_{\mathbb{C}}(V)$. Zeige, die folgenden Aussagen sind gleichwertig:

a. $f^* = -f$.
b. Für alle $x \in V$ gilt: $\big(f(x), x\big) \in i\mathbb{R}$.
c. Es gibt eine Orthonormalbasis von V aus Eigenvektoren von f und der Realteil aller Eigenwerte ist Null.

Aufgabe 3.34 (Spektralzerlegung selbstadjungierter Endomorphismen)

Es sei V ein endlich-dimensionaler Hilbertraum und $f \in \mathrm{End}_{\mathbb{K}}(V)$ selbstadjungiert.

Zeige, es gibt reelle Zahlen $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$ und selbstadjungierte Projektionen $\pi_1, \ldots, \pi_r \in \mathrm{End}_{\mathbb{K}}(V)$ mit den Eigenschaften

(i) $\sum_{i=1}^{r} \pi_i = \mathrm{id}_V$,
(ii) $\pi_i \circ \pi_j = \delta_{ij} \cdot \pi_i$ für $i, j \in \{1, \ldots, r\}$, und
(iii) $f = \sum_{i=1}^{r} \lambda_i \pi_i$.

Aufgabe 3.35

Schreibe eine Singular-Prozedur `normal`, die eine normale Matrix $A \in \mathrm{Mat}(n, \mathbb{Q})$ einliest und eine zu A konjugierte Diagonalmatrix $D \in \mathrm{Mat}(n, \mathbb{C})$ wieder ausgibt. Zerfällt χ_A über $\mathbb{Q}$, so soll zudem eine Matrix T bestimmt werden, deren Spalten eine OGB von $\mathbb{R}^n$ bilden und die nach Normierung eine unitäre Transformationsmatrix zur Diagonalisierung von A liefern.

Man verwende folgenden Algorithmus.

INPUT: $A \in \mathrm{Mat}(n, \mathbb{Q})$ normal.

OUTPUT: $D \in \mathrm{Mat}(n, \mathbb{C})$ Diagonalmatrix mit $D \sim A$, und $T \in \mathrm{U}(n)$ mit $T^{-1} \circ A \circ T = D$, falls μ_A über $\mathbb{Q}$ zerfällt.

1. **Schritt:** Überprüfe, ob A normal ist.
2. **Schritt:** Faktorisiere das Minimalpolynom μ_A über $\mathbb{Q}$ und prüfe, ob es zerfällt. Falls nein, gehe zu Schritt 4.
3. **Schritt:** Bestimme für jeden Eigenwert von A eine Orthogonalbasis des zugehörigen Eigenraumes mittels der Prozeduren `syz` und `OGB` - vgl. Aufgabe 2.44. Schreibe die Basisvektoren als Spaltenvektoren in eine Matrix T.
4. **Schritt:** Bestimme eine zu A konjugierte Diagonalmatrix $D \in \mathrm{Mat}(n, \mathbb{C})$ mittels der Prozedur aus Aufgabe III.3.52, und gib eine Liste bestehend aus der Matrix D sowie ggf. der Matrix T zurück.

Aufgabe 3.36
Schreibe eine rekursive Singular-Prozedur `sgauss`, die eine symmetrische Matrix $A \in \mathrm{Mat}(n, K)$ einliest und mittels des symmetrischen Gauß-Algorithmus auf Diagonalgestalt transformiert. Dabei verwende man den folgenden Algorithmus:

INPUT: $A \in \mathrm{Mat}(n, K)$ symmetrisch.

OUTPUT: $D \in \mathrm{Mat}(n, K)$ Diagonalmatrix mit $\exists\, T \in \mathrm{Gl}_n(K) \;:\; T^t \circ A \circ T$ Diagonalgestalt hat.

1. **Schritt:** Überprüfe, ob A symmetrisch ist.
2. **Schritt:** Man suche in der ersten Spalte von A den ersten Eintrag, der nicht Null ist. Existiert ein solcher, merke man sich die Zeilennummer z, sonst gehe man zu Schritt 5.
3. **Schritt:** Ist $z \neq 1$, addiere die z-te Zeile von A zur ersten und die z-te Spalte zur ersten.
4. **Schritt:** Für $k = 2, \ldots, n =$ `ncols(A)` addiere man das $-A[1,k]/A[1,1]$-fache der ersten Zeile von A zur k-ten und das $-A[1,k]/A[1,1]$-fache der ersten Spalte zur k-ten.
5. **Schritt:** Falls $n > 1$, dann erzeuge man eine Matrix B, indem man aus A die erste Zeile und die erste Spalte streicht. Ferner rufe man die Prozedur `sgauss` mit B auf und speichere das Ergebnis in `submat(A,2..n,2..n)`.
6. **Schritt:** Man gebe A zurück.

Aufgabe 3.37
Schreibe eine rekursive Singular-Prozedur `sym_gauss`, die eine symmetrische Matrix $A \in \mathrm{Mat}(n, K)$ sowie ein optionales Argument `list #` einliest und eine Matrix $T \in \mathrm{Gl}_n(K)$ mit paarweise orthogonalen Spaltenvektoren ausgibt,

so daß $T^t \circ A \circ T$ Diagonalgestalt hat. Dabei verwende man den folgenden Algorithmus:

INPUT: $A \in \mathrm{Mat}(n, K)$ symmetrisch.

OUTPUT: $T \in \mathrm{Gl}_n(K)$ so, daß $T^t \circ A \circ T$ Diagonalgestalt hat.

1. **Schritt:** Überprüfe, ob A symmetrisch ist.
2. **Schritt:** Im ersten Rekursionsschritt setze man $T = \mathbb{1}_n$, in den späteren Rekursionsschritten setze man $T =$`#[1]`.
3. **Schritt:** Setze $m =$ `ncols(T)` $-$ `ncols(A)`.
4. **Schritt:** Man suche in der ersten Spalte von A den ersten Eintrag, der nicht Null ist. Existiert ein solcher, merke man sich die Zeilennummer z, sonst gehe man zu Schritt 7.
5. **Schritt:** Ist $z \neq 1$, addiere die z-te Zeile von A zur ersten und die z-te Spalte zur ersten. Addiere ferner die $z + m$-te Spalte von T zur $m + 1$-ten Spalte.
6. **Schritt:** Für $k = 2, \ldots,$ `ncols(A)` addiere man das $-A[1,k]/A[1,1]$-fache der $1 + m$-ten Spalte von T zur $k + m$-ten. Sodann addiere man das $-A[1,k]/A[1,1]$-fache der ersten Zeile von A zur k-ten und das $-A[1,k]/A[1,1]$-fache der ersten Spalte zur k-ten.
7. **Schritt:** Falls `ncols(A)>1`, dann erzeuge man eine Matrix B, indem man aus A die erste Zeile und die erste Spalte streicht. Ferner rufe man die Prozedur `sym_gauss` mit den Parametern B und T auf und speichere das Ergebnis in T.
8. **Schritt:** Man gebe T zurück.

KAPITEL V

Dualität und multilineare Algebra

Im letzten Kapitel der Vorlesung werden noch einmal lineare, bilineare und, allgemeiner, multilineare Abbildungen behandelt. Etwas schematisch, aber doch zutreffend kann man sagen, daß es das Ziel ist, Eigenschaften von multilinearen Abbildungen zu verstehen, indem man neue Vektorräume konstruiert und auf diesen lineare Abbildungen studiert, deren Theorie man schon kennt. Die multilineare Theorie wird also, mittels der Konstruktion komplizierterer Vektorräume auf die lineare Theorie zurückgeführt. Dies erleichtert nicht nur das Verständnis, sondern erlaubt es auch, konkrete Rechnungen mit Matrizen durchzuführen.

Zunächst studieren wir jedoch lineare Abbildungen eines K-Vektorraumes V in den Körper K selbst. Diese Abbildungen bilden wieder einen Vektorraum, den Dualraum V^*, der in besonders enger Beziehung zu V steht. Dualräume spielen insbesondere bei und endlich-dimensionalen Funktionenräumen eine wichtige Rolle. Hier wird im Wesentlichen aber nur auf die endlich-dimensionale Theorie eingegangen.

1. Dualraum und transponierte Abbildung

Wir entwickeln die grundlegenden Begriffe zunächst für beliebige Moduln über einem kommutativen Ring mit Eins, da das gar keinen Unterschied macht. Für weitergehende Aussagen werden wir uns allerdings auf Vektorräume beschränken.

Generalvoraussetzung: Es sei R stets ein kommutativer Ring mit Eins und K ein Körper. V und W seien R-Moduln bzw. K-Vektorräume, je nach Kontext.

Definition 1.1 a. Der R-Modul

$$V^* = \mathrm{Hom}_R(V, R) = \{g : V \to R \mid g \text{ ist } R\text{-linear}\}$$

heißt der *Dualraum* oder *Dualmodul* von V.
Die Elemente von V^* werden *Linearformen* oder *lineare Funktionale* genannt.

b. Zu $f \in \mathrm{Hom}_R(V, W)$ definieren wir die *duale* oder *transponierte Abbildung* durch

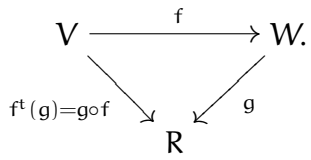
$$f^t : W^* \to V^* : g \mapsto f^t(g) = g \circ f.$$

Statt f^t schreiben wir auch f^* - vgl. Aufgabe 1.23.

c. Wir definieren eine kanonische *duale Paarung* auf V durch

$$\langle \cdot, \cdot \rangle : V^* \times V \to R : (g, x) \mapsto \langle g, x \rangle = g(x).$$

Bemerkung 1.2
Die Definition von f^t wird durch das folgende kommutative Diagramm verdeutlicht:



Lemma 1.3
Die duale Paarung $\langle\cdot,\cdot\rangle : V^* \times V \to R$ *ist bilinear.*[1]

Beweis: Sei $g \in V^*$. Dann ist die Abbildung

$$\langle g, \cdot\rangle = g : V \to R$$

per definitionem linear.

Ist nun $x \in V$, dann gilt für $\lambda, \mu \in R$ und $g, h \in V^*$

$$\langle \lambda g + \mu h, x\rangle = (\lambda g + \mu h)(x) = \lambda g(x) + \mu h(x) = \lambda\langle g, x\rangle + \mu\langle h, x\rangle.$$

Also ist auch

$$\langle \cdot, x\rangle : V^* \to R$$

linear, und mithin ist die duale Paarung bilinear. □

Proposition 1.4
Seien U, V *und* W R*-Moduln,* $\lambda \in R$*,* $f, \widetilde{f} \in \mathrm{Hom}_R(V, W)$ *und* $f' \in \mathrm{Hom}_R(W, U)$*. Dann gelten:*

a. f^t *ist* R*-linear.*
b. $(\mathrm{id}_V)^t = \mathrm{id}_{V^*}$.
c. $\left(f' \circ f\right)^t = f^t \circ f'^t$.
d. *Ist* f *ein Isomorphismus, so ist* f^t *ein Isomorphismus.*
e. $\left(f + \tilde{f}\right)^t = f^t + \tilde{f}^t$ *und* $(\lambda f)^t = \lambda \cdot f^t$.
Insbesondere haben wir eine R*-lineare Abbildung*

$$t : \mathrm{Hom}_R(V, W) \to \mathrm{Hom}_R\left(W^*, V^*\right) : f \mapsto f^t.$$

Beweis: a. Für $\lambda, \mu \in R$ und $g, h \in V^*$ gilt

$$f^t(\lambda g + \mu h) = (\lambda g + \mu h) \circ f = \lambda \cdot (g \circ f) + \mu \cdot (h \circ f) = \lambda f^t(g) + \mu f^t(h).$$

b. Sei $g \in V^*$, dann ist $\mathrm{id}_V^t(g) = g \circ \mathrm{id}_V = g = \mathrm{id}_{V^*}(g)$.
c. Sei $g \in U^*$. Dann gilt

$$\left(f' \circ f\right)^t(g) = g \circ f' \circ f = f^t\left(g \circ f'\right) = f^t\left(f'^t(g)\right) = \left(f^t \circ f'^t\right)(g).$$

[1] Für eine Definition des Begriffes bilinear sei auf Definition 2.1 verwiesen.

d. Aus Teil b. und c. folgt:

$$\left(f^{-1}\right)^t \circ f^t = \left(f \circ f^{-1}\right)^t = \mathrm{id}_W^t = \mathrm{id}_{W^*}$$

und

$$f^t \circ \left(f^{-1}\right)^t = \left(f^{-1} \circ f\right)^t = \mathrm{id}_V^t = \mathrm{id}_{V^*}\,.$$

Mithin ist f^t ein Isomorphismus mit $\left(f^{-1}\right)^t$ als Inverser.

e. Die beiden Eigenschaften folgen unmittelbar aus der Definition der dualen Abbildung.

□

Beispiel 1.5 a. Sei $b : V \times V \to R$ eine Bilinearform und $x \in V$ fest gegeben. Dann ist

$$b(x, \cdot) : V \to R : y \mapsto b(x, y)$$

linear und mithin $b(x, \cdot) \in V^*$.

b. Ist $V = R^n$ und

$$b = \langle \cdot, \cdot \rangle : R^n \times R^n \to R : \left((x_1, \ldots, x_n)^t, (y_1, \ldots, y_n)^t\right) \mapsto \sum_{i=1}^{n} x_i y_i,$$

und ist $x = (x_1, \ldots, x_n)^t \in R^n$ fest vorgegeben, dann ist nach a.

$$\langle x, \cdot \rangle : R^n \to R : y = (y_1, \ldots, y_n)^t \mapsto \langle x, y \rangle = \sum_{i=1}^{n} x_i y_i$$

eine Linearform.

c. Wählen wir in b. $x = e_k$, $k \in \{1, \ldots, n\}$, so erhalten wir,

$$\mathrm{pr}_k = \langle e_k, \cdot \rangle : R^n \to R : (y_1, \ldots, y_n)^t \mapsto y_k,$$

daß die *Projektion* auf die k-te Komponente eine Linearform ist.

d. Die Umkehrung der Aussage in Proposition 1.4 d. gilt nicht. Dazu betrachten wir das folgende Beispiel.
Es sei $R = \mathbb{Z}$ und $V = W = \mathbb{Z}_2$. Dann gilt $V^* = \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}_2, \mathbb{Z}) = \{0\}$, denn, für einen Homomorphismus $\alpha : \mathbb{Z}_2 \to \mathbb{Z}$ mit $\alpha([1]) = z \in \mathbb{Z}$ gilt

$$0 = \alpha([0]) = \alpha([1] + [1]) = \alpha([1]) + \alpha([1]) = 2z,$$

also $\alpha([1]) = z = 0$.
Dann ist aber $\mathrm{Hom}_{\mathbb{Z}}\left(W^*, V^*\right) = \mathrm{Hom}_{\mathbb{Z}}\left(\{0\}, \{0\}\right) = \left\{\mathrm{id}_{\{0\}}\right\}$, und damit ist *jedes* Element von $\mathrm{Hom}_{\mathbb{Z}}(W^*, V^*)$ ein Isomorphismus.
Hingegen ist $\mathrm{Hom}_{\mathbb{Z}}(V, W) = \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}_2, \mathbb{Z}_2) = \{\mathrm{id}_{\mathbb{Z}_2}, 0\}$, enthält also die Abbildung $0 : \mathbb{Z}_2 \to \mathbb{Z}_2$, die alles auf die Null abbildet und sicher kein Isomorphismus ist, während $0^t = \mathrm{id}_{\{0\}}$ bijektiv ist.

Bemerkung 1.6
Ist $b : V \times V \to R$ eine Bilinearform, so definiert diese eine R-lineare Abbildung

$$\phi_b : V \to V^* : x \mapsto b(x, \cdot).$$

Diese hängt jedoch von der Wahl von b ab. I. a. existiert keine *kanonische* Abbildung von V nach V^*.

Satz 1.7
Sei $B = (x_i \mid i \in I)$ *eine Basis des* R*-Moduls* V*, d. h.* V *ist ein freier Modul - vgl. Aufgabe 2.32. Dann gelten:*

a. *Für jedes* $i \in I$ *gibt es eine eindeutig bestimmte Linearform* $x_i^* \in V^*$ *mit* $\langle x_i^*, x_j \rangle = \delta_{ij}$ *für alle* $j \in I$.
b. $B^* = \left(x_i^* \mid i \in I\right)$ *ist linear unabhängig.*
c. *Ist* $B = (x_1, \ldots, x_n)$ *endlich, dann ist* $B^* = \left(x_1^*, \ldots, x_n^*\right)$ *eine Basis von* V^**, die sogenannte* zu B duale Basis.
Insbesondere gilt dann $V \cong V^*$.

Beweis: a. Lineare Abbildungen können auf einer Basis frei vorgegeben werden und sind dann eindeutig bestimmt - vgl. Aufgabe 2.32.
b. Seien $\lambda_i \in R$, $i \in I$, mit $\sum'_{i\in I} \lambda_i x_i^* = 0$. Dann folgt für $j \in I$

$$0 = \left\langle \sum_{i\in I}{}' \lambda_i x_i^*, x_j \right\rangle = \sum_{i\in I}{}' \lambda_i \langle x_i^*, x_j \rangle = \lambda_j.$$

c. Wegen b. reicht es zu zeigen, daß $V^* = \langle x_1^*, \ldots, x_n^* \rangle$.
Sei dazu $g \in V^*$. Wir setzen

$$h := \sum_{i=1}^{n} \langle g, x_i \rangle \cdot x_i^* \in \langle x_1^*, \ldots, x_n^* \rangle.$$

Dann folgt für $j \in \{1, \ldots, n\}$

$$h(x_j) = \sum_{i=1}^{n} \langle g, x_i \rangle \cdot \langle x_i^*, x_j \rangle = \langle g, x_j \rangle = g(x_j),$$

und mithin ist $g = h \in \langle x_1^*, \ldots, x_n^* \rangle$.

□

Korollar 1.8
Ist V *ein* K*-Vektorraum mit* $\dim_K(V) < \infty$*, dann gilt* $\dim_K\left(V^*\right) = \dim_K(V)$.
Korollar 1.9
Ist $B = (x_i \mid i \in I)$ *eine Basis des* R*-Moduls* V*, dann ist die* R*-lineare Abbildung*

$$\phi_B : V \to V^* : x_i \mapsto x_i^*, \quad i \in I,$$

ein Mono*morphismus.*

Ist B *endlich, dann ist* ϕ_B *sogar ein* Iso*morphismus. (Vgl. Bemerkung 1.10 c.)*
Bemerkung 1.10 a. Man beachte, zu einer endlichen Basis B hat man die duale Basis B^*. Zu einem Vektor x gibt es jedoch keinen *dualen Vektor* x^*! Bezeichnet man $\phi_B(x)$ mit x^*, so hängt x^* nicht nur von x ab, sondern auch von der gewählten Basis B.
b. Im Fall $V = R^n$ hat man die kanonische Basis $E = (e_1, \ldots, e_n)$, und - akzeptiert man, daß diese *nicht* von einer Wahl abhängt - so liefert diese

gleichsam einen kanonischen Isomorphismus

$$\phi_E : R^n \to \left(R^n\right)^* : x = (x_1, \ldots, x_n)^t \mapsto \langle x, \cdot \rangle = \sum_{i=1}^{n} x_i e_i^*.$$

Es gilt dann für $y = (y_1, \ldots, y_n)^t \in R^n$ – vgl. Bemerkung 1.16 –

$$\phi_E(x)(y) = x^t \circ y.$$

c. Ist V ein K-Vektorraum mit Basis $B = (x_i \mid i \in I)$, so gilt stets $V = \bigoplus_{i\in I} Kx_i$ und $V^* \cong \prod_{i\in I} Kx_i^*$ - vgl. Definition 2.7.
Man kann zeigen, daß $\bigoplus_{i\in I} Kx_i$ und $\prod_{i\in I} Kx_i^*$ genau dann isomorph sind, wenn I endlich ist.
Insbesondere gilt also, ist $\dim_K(V) = \infty$, so ist $V \ncong V^*$.

Beispiel 1.11

Wir wollen nun ein Beispiel für einen unendlich-dimensionalen $\mathbb{R}$-Vektorraum betrachten, bei dem es physikalisch relevante Elemente in V^* gibt, die nicht von V herkommen.

Sei $V = \mathcal{C}^0\big([0,1]\big)$. Wir betrachten auf V die Bilinearform

$$b : V \times V \to \mathbb{R} : (f, g) \mapsto \int_0^1 f(t)g(t)dt.$$

Dann erhalten wir für $f \in V$ die Linearform auf V

$$\int_0^1 f = b(f, \cdot) : V \to \mathbb{R} : g \mapsto \int_0^1 f(t)g(t)dt$$

und damit die Abbildung

$$\int_0^1 = \phi_b : V \to V^* : f \mapsto \int_0^1 f.$$

Behauptung: $\int_0^1 = \phi_b$ ist *kein* Isomorphismus.

Wir wollen zunächst einmal bemerken, daß $\int_0^1$ ein *Mono*morphismus ist, da b ein Skalarprodukt - also positiv definit - ist, so daß bestenfalls die Surjektivität schief gehen kann.

Wir betrachten zunächst für ein festes $p \in]0,1[$ das lineare Funktional

$$\delta_p : V \to \mathbb{R} : f \mapsto f(p)$$

und behaupten $\delta_p \notin \operatorname{Im}\left(\int_0^1\right)$.

Dazu nehmen wir an, es gäbe ein $\delta \in V$ mit $\delta_p = \int_0^1 \delta$, d. h., für alle $f \in V$ gilt

$$f(p) = \delta_p(f) = \int_0^1 \delta(t)f(t)dt.$$

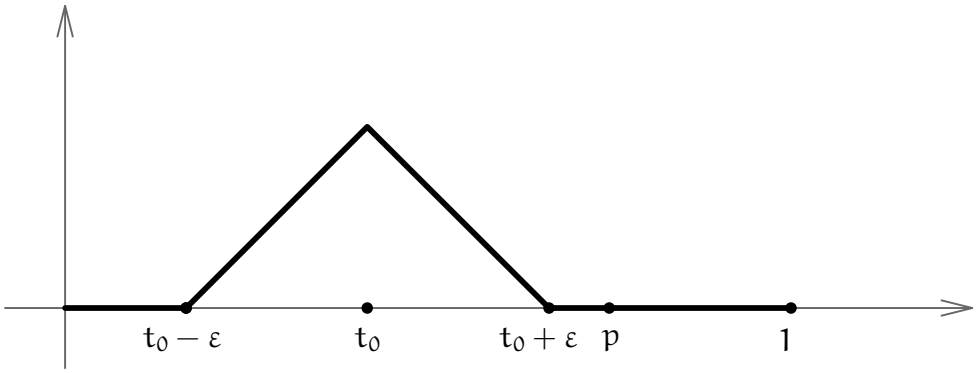
Wir nehmen ferner an, daß es ein $t_0 \neq p$ gibt mit $\delta(t_0) \neq 0$, o. E. $\delta(t_0) > 0$ und $t_0 \in]0,1[$. Da δ stetig ist, gibt es dann aber ein $0 < \varepsilon \leq \min\left\{|t_0 - p|, |t_0 - 0|, |t_0 - 1|\right\}$, so daß

$$\delta(t) > 0 \quad \forall\, t \in]t_0 - \varepsilon, t_0 + \varepsilon[. \tag{66}$$

Wir wählen nun eine stetige Funktion $f \in V$ mit

$$f(t) \begin{cases} > 0, & \forall\, t \in]t_0 - \varepsilon, t_0 + \varepsilon[, \\ = 0, & \forall\, t \notin]t_0 - \varepsilon, t_0 + \varepsilon[, \end{cases}$$

z. B. die Funktion mit folgendem Graphen



Dann führt

$$0 = f(p) = \int_0^1 \delta(t) f(t)\,dt = \int_{t_0-\varepsilon}^{t_0+\varepsilon} \delta(t) f(t)\,dt > 0$$

zu einem Widerspruch. Also muß $\delta(t) = 0$ für alle $t \neq p$ gelten, dann aber auch $\delta(p) = 0$, da δ stetig ist.

Aber damit folgt dann $\delta_p = \int_0^1 \delta = 0$, und mithin gilt für alle auf $[0,1]$ stetigen Funktionen $f \in V$, daß $f(p) = 0$ ist, was offensichtlich nicht der Fall ist.

Folglich ist $\delta_p \notin \operatorname{Im}\left(\int_0^1\right)$, und mithin ist $\int_0^1$ auch kein Isomorphismus.

Die Physiker hätten nun aber gerne, daß auch δ_p das Integral einer stetigen Funktion wäre, und sie behelfen sich dadurch, daß sie eine Funktion $\delta : [0,1] \to \mathbb{R} \cup \{\infty\}$ einführen mit $\delta(t) = 0$ für $t \neq p$ und $\delta(p) = \infty$, wobei das *Unendlich* auch noch so beschaffen ist, daß für jede Zahl $x \in \mathbb{R}$ das Integral über $x \cdot \delta$ auf dem Einheitsintervall gerade x ergibt, und leiten daraus (66) her. – Mit anderen Worten, sie fassen schlicht alle linearen Funktionale auf V als Funktionen auf $[0,1]$ auf.

Man beachte hierbei auch, daß die Injektivität von $\int_0^1$ es erlaubt, jede stetige Funktion g mit dem linearen Funktional $\int_0^1 g$ zu identifizieren, während halt nicht jedes lineare Funktional auf diese Weise gewonnen werden kann.

δ_p heißt die *Diracsche Deltafunktion*, die jedoch nicht eine Funktion im herkömmlichen Sinne ist, sondern ein lineares Funktional auf $\mathcal{C}^0\big([0,1]\big)$.

Nun kann man zu einem Dualraum V^* von V wieder den Dualraum $\big(V^*\big)^*$ bilden. Es ist bemerkenswert und auch erstaunlich, daß es eine kanonische Abbildung $V \to \big(V^*\big)^*$ gibt, obwohl keine kanonische Abbildung von $V \to V^*$ existiert.

Definition 1.12
Der $\mathbb{R}$-Modul $V^{**} = \big(V^*\big)^*$ heißt der *Bidualraum* oder *Bidualmodul* von V.

Die R-lineare Abbildung

$$** : V \to V^{**} : x \mapsto x^{**} = \langle \cdot, x \rangle$$

hängt nur von V und von keiner speziellen Wahl ab. Wir sagen deshalb, daß $**$ eine *kanonische* Abbildung ist.[2]

Lemma 1.13
Ist V ein K-Vektorraum, dann gelten:

a. $**$ *ist ein* Mono*morphismus.*
b. *Ist* $\dim_K(V) < \infty$*, dann ist* $**$ *ein* Iso*morphismus.*

Beweis: a. Angenommen, es gäbe ein $0 \neq x \in \mathrm{Ker}(**)$. Wir ergänzen die Familie (x) zu einer Basis B von V, und setzen $g := \phi_B(x) \in V^*$. Dann gilt

$$0 = x^{**}(g) = \langle g, x \rangle = \langle \phi_B(x), x \rangle = 1,$$

was ein Widerspruch ist.

b. Aus Korollar 1.8 folgt $\dim_K(V) = \dim_K(V^*) = \dim_K(V^{**})$, und mithin ist wegen a. $**$ ein Isomorphismus.

□

Im Fall von endlich-dimensionalen Vektorräumen können wir mit Hilfe der dualen Paarung Vektoren und Linearformen leicht bezüglich einer Basis bzw. ihrer dualen Basis ausdrücken. Man vergleiche hierzu auch die Gleichung (58) in Lemma IV.2.5.

Lemma 1.14
Es sei $B = (x_1, \ldots, x_n)$ eine Basis des K-Vektorraumes V und B^* die dazu duale Basis von V^*. Ferner seien $x \in V$ und $g \in V^*$. Dann gelten:

$$x = \sum_{i=1}^{n} \langle x_i^*, x \rangle \cdot x_i \tag{67}$$

und

$$g = \sum_{i=1}^{n} \langle g, x_i \rangle \cdot x_i^*. \tag{68}$$

Beweis: Sei $x = \sum_{i=1}^{n} \lambda_i x_i$. Dann gilt $\langle x_i^*, x \rangle = \sum_{j=1}^{n} \lambda_i \langle x_i^*, x_j \rangle = \lambda_i$, und (67) folgt. (68) wurde bereits im Beweis von Satz 1.7 gezeigt. □

Aus diesem Lemma leiten wir unmittelbar ab, wie sich die Matrixdarstellung einer linearen Abbildung bezüglich gegebener Basen zur Matrixdarstellung der dualen Abbildung bezüglich der dualen Basen verhält.

Proposition 1.15
Es seien V und W K-Vektorräume mit Basen $B = (x_1, \ldots, x_n)$ und $D = (y_1, \ldots, y_m)$, und $f \in \mathrm{Hom}_K(V, W)$. Dann gilt

$$M_{B^*}^{D^*}(f^t) = \left(M_D^B(f)\right)^t.$$

[2] Die R-Linearität von $**$ folgt unmittelbar aus der Bilinearität der dualen Paarung.

Insbesondere gilt für $f \in \mathrm{Hom}_K\left(K^n, K^m\right)$ *und* $A \in \mathrm{Mat}(m \times n, K)$

$$A_{f^t} = (A_f)^t \quad \textit{und} \quad (f_A)^t = f_{A^t}.$$

Beweis: Aus Lemma 1.14 folgt für $j \in \{1, \dots, n\}$

$$f(x_j) = \sum_{i=1}^{m} \left\langle y_i^*, f(x_j) \right\rangle \cdot y_i$$

und für $i \in \{1, \dots, m\}$

$$f^t\left(y_i^*\right) = \sum_{j=1}^{n} \left\langle f^t\left(y_i^*\right), x_j \right\rangle \cdot x_j^*.$$

Aber dann gilt

$$M_D^B(f) = \left(\left\langle y_i^*, f(x_j) \right\rangle\right)_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$$

und

$$M_{B^*}^{D^*}\left(f^t\right) = \left(\left\langle f^t\left(y_i^*\right), x_j \right\rangle\right)_{\substack{j=1,\dots,n \\ i=1,\dots,m}}.$$

Es bleibt also für $i = 1, \dots, m$ und $j = 1, \dots, n$ zu zeigen

$$\left\langle y_i^*, f(x_j) \right\rangle = \left\langle f^t\left(y_i^*\right), x_j \right\rangle.$$

Aber per definitionem gilt

$$\left\langle f^t\left(y_i^*\right), x_j \right\rangle = \left(y_i^* \circ f\right)(x_j) = y_i^*\left(f(x_j)\right) = \left\langle y_i^*, f(x_j) \right\rangle.$$

□

Bemerkung 1.16 a. Die Aussage in Proposition 1.15 läßt sich auch durch folgendes kommutatives Diagramm ausdrücken:

$$\begin{array}{ccccc}
 & & f & \longmapsto & f^t \\
f & & \mathrm{Hom}_K(V,W) & \longrightarrow & \mathrm{Hom}_K\left(W^*, V^*\right) & & f^t \\
\downarrow & & \downarrow \cong & & \downarrow \cong & & \downarrow \\
M_D^B(f) & & \mathrm{Mat}(m \times n, K) & \xrightarrow{\cong} & \mathrm{Mat}(n \times m, K) & & M_{B^*}^{D^*}\left(f^t\right) \\
 & & A & \longmapsto & A^t
\end{array}$$

Insbesondere folgt aus $\dim_K(V), \dim_K(W) < \infty$

$$t : \mathrm{Hom}_K(V,W) \xrightarrow{\cong} \mathrm{Hom}_K\left(W^*, V^*\right) : f \mapsto f^t.$$

Im unendlich-dimensionalen Fall gilt das nicht mehr unbedingt - vgl. Aufgabe 1.30.

b. Es sei $E = (e_1, \dots, e_n)$ die kanonische Basis des K^n und $E = \left(e_1^*, \dots, e_n^*\right)$ die dazu duale Basis von $\left(K^n\right)^*$.
Während wir im $K^n = \mathrm{Mat}(n \times 1, K)$ die Vektoren als Spaltenvektoren schreiben, wollen wir die Elemente von $\left(K^n\right)^* \cong \mathrm{Mat}(1 \times n, K)$ als

Zeilenvektoren schreiben, sprich wir identifizieren $g = \sum_{i=1}^{n} \mu_i e_i^*$ mit $(\mu_1, \ldots, \mu_n)$.
Das hat den Vorteil, daß wir die Operation von g auf $x = (\lambda_1, \ldots, \lambda_n)^t \in K^n$ als Matrixmultiplikation schreiben können:

$$\langle g, x \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_j \mu_i \langle e_i^*, e_j \rangle = \sum_{i=1}^{n} \mu_i \lambda_i = \begin{pmatrix} \mu_1 & \ldots & \mu_n \end{pmatrix} \circ \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

Wir leiten nun daraus ein *Verfahren zur Bestimmung der dualen Basis* einer gegebenen Basis ab.
Seien Vektoren $x_i = (a_{1i}, \ldots, a_{ni})^t \in K^n$ für $i = 1, \ldots, n$ gegeben und bilde $B = (x_1, \ldots, x_n)$ eine Basis von K^n.
Dann ist für $x_i^* = (b_{i1}, \ldots, b_{in}) \in \left(K^n\right)^*$, $i = 1, \ldots, n$, die Familie $\left(x_1^*, \ldots, x_n^*\right)$ genau dann die zu B duale Basis, wenn für alle $i, j = 1, \ldots, n$ gilt

$$\begin{pmatrix} b_{i1} & \ldots & b_{in} \end{pmatrix} \circ \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix} = \langle x_i^*, x_j \rangle = \delta_{ij}.$$

Aber das ist gleichwertig dazu, daß

$$\begin{pmatrix} b_{11} & \ldots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \ldots & b_{nn} \end{pmatrix} \circ \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} = \mathbb{1}_n.$$

D. h., schreibt man die Vektoren der Basis als Spalten einer Matrix, so erhält man die duale Basis als die Zeilen der Inversen.

Beispiel 1.17
Es ist $B = (x_1, x_2, x_3) = \left((0, -1, -1)^t, (-1, 0, 1,)^t, (1, -1, 0)^t\right)$ eine Basis des $\mathbb{Q}^3$. Bestimmen wir mit Singular die Inverse der Matrix

$$A = \begin{pmatrix} 0 & -1 & 1 \\ -1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix} \in Gl_3(\mathbb{Q}),$$

die die Vektoren von B als Spalten enthält, so ergibt sich

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

Mithin ist die duale Basis

$$B^* = \left(x_1^*, x_2^*, x_3^*\right) = \left(\tfrac{1}{2} \cdot (-1, -1, -1), \tfrac{1}{2} \cdot (-1, -1, 1), \tfrac{1}{2} \cdot (1, -1, 1)\right).$$

Wir wollen jetzt die Beziehung zwischen f und f^t weiter untersuchen. Dazu führen wir Orthogonalräume ein.

Definition 1.18
Es sei V ein R-Modul, $U \subseteq V$ ein Untermodul von V und $W \subseteq V^*$ ein Untermodul von V^*.

a. $U^o = \{g \in V^* \mid \langle g, x\rangle = 0 \ \forall\ x \in U\}$ heißt der *Orthogonalraum* oder *Orthogonalmodul* oder auch *Annulator* von U.
b. $W^\circledcirc = \{x \in V \mid \langle g, x\rangle = 0 \ \forall\ g \in W\}$ heißt der *Orthogonalraum* oder *Orthogonalmodul* oder auch *Annulator* von W.

Bemerkung 1.19 a. Aus der Bilinearität der dualen Paarung folgt, daß U^o sowie $W^\circledcirc$ Untermoduln von V bzw. V^* sind.
b. Betrachten wir $V = \mathbb{R}^n$ mit dem kanonischen Skalarprodukt und $\phi_E : V \to V^* : e_i \mapsto e_i^*$, dann gilt

$$\phi_{E|} : U^\perp \xrightarrow{\cong} U^o$$

ist ein Isomorphismus - vgl. Aufgabe 1.23.
Daher rührt auch die Bezeichnung *Orthogonalraum*.

Proposition 1.20
Es sei $f : V \to W$ *ein* K*-Vektorraumhomomorphismus. Dann gilt:*

a. $\mathrm{Ker}\left(f^t\right) = \left(\mathrm{Im}(f)\right)^o$.
b. $\mathrm{Im}\left(f^t\right) = \left(\mathrm{Ker}(f)\right)^o$.

Beweis: a. Es gelten die folgenden Äquivalenzen:

$$\begin{aligned} g \in \mathrm{Ker}\left(f^t\right) &\Leftrightarrow 0 = f^t(g) = g \circ f \\ &\Leftrightarrow 0 = \langle g \circ f, x\rangle = g\big(f(x)\big) \quad \forall\ x \in V \\ &\Leftrightarrow 0 = g(y) = \langle g, y\rangle \quad \forall\ y \in \mathrm{Im}(f) \\ &\Leftrightarrow g \in \left(\mathrm{Im}(f)\right)^o. \end{aligned}$$

b. Sei $g \in \mathrm{Im}\left(f^t\right)$. Dann gibt es ein $h \in W^*$ mit $g = f^t(h) = h \circ f$. Ist nun $x \in \mathrm{Ker}(f)$, dann gilt

$$\langle g, x\rangle = h\big(f(x)\big) = h(0) = 0.$$

Mithin ist $g \in \left(\mathrm{Ker}(f)\right)^o$.
Sei nun umgekehrt $g \in \left(\mathrm{Ker}(f)\right)^o$. Es ist unser Ziel, ein $h \in W^*$ zu finden, so daß $g = f^t(h)$ gilt. Wir führen den Beweis in zwei Schritten.
Schritt a.: Wir wählen eine Basis $B' = (y_i \mid i \in I)$ von $\mathrm{Im}(f)$ und ergänzen diese zu einer Basis $B = (y_i, z_j \mid i \in I, j \in J)$ von W. Sodann wählen wir zu jedem y_i, $i \in I$, ein $x_i \in f^{-1}(y_i)$, und definieren eine K-lineare Abbildung $h \in \mathrm{Hom}_K(W, K) = W^*$ durch

$$h : W \to K : \begin{cases} y_i \mapsto g(x_i) \in K, & i \in I, \\ z_j \mapsto 0 \in K, & j \in J. \end{cases}$$

Schritt b.: Wir zeigen nun, daß $f^t(h) = g$.

Sei dazu $x \in V$. Dann gibt es $\lambda_i \in K$, $i \in I$, mit $f(x) = \sum'_{i\in I} \lambda_i y_i$. Dann gilt aber für $x' := \sum'_{i\in I} \lambda_i x_i$,

$$f(x') = \sum_{i\in I}{}' \lambda_i f(x_i) = \sum_{i\in I}{}' \lambda_i y_i = f(x),$$

und damit $f(x - x') = f(x) - f(x') = 0$, d. h. $x - x' \in \mathrm{Ker}(f)$. Also gilt:

$$0 = \langle g, x - x' \rangle = g(x) - g(x'),$$

und damit

$$g(x) = g(x') = \sum_{i\in I}{}' \lambda_i g(x_i) = \sum_{i\in I}{}' \lambda_i h(y_i) = \sum_{i\in I}{}' \lambda_i h\big(f(x_i)\big)$$

$$= h\left(f\Big(\sum_{i\in I}{}' \lambda_i x_i\Big)\right) = h\Big(f(x')\Big) = h\big(f(x)\big) = f^t(h)(x).$$

Also ist $f^t(h) = g$.
Alternativer Beweis von $\big(\mathrm{Ker}(f)\big)^o \subseteq \mathrm{Im}\,(f^t)$ nach Philipp Diesinger:
Sei $g \in \big(\mathrm{Ker}(f)\big)^o$. Es ist unser Ziel, ein $h \in W^*$ zu finden, so daß $g = f^t(h)$ gilt.
Sei nun $U \subseteq V$ ein Komplement von $\mathrm{Ker}(f)$ in V. Nach Satz II.4.14 und Lemma II.4.15 besitzt $f_| : U \to \mathrm{Im}(f)$ eine Inverse, sagen wir $f' : \mathrm{Im}(f) \to U \subseteq V$. Sei nun $U' \subseteq W$ ein Komplement von $\mathrm{Im}(f)$, dann können wir die Linearform $h : W \to K$ auf $\mathrm{Im}(f)$ und auf U' unabhängig definieren durch $h_{|\mathrm{Im}(f)} = g_{|U} \circ f'$ und $h_{|U'} = 0$. Damit gilt

$$f^t(h)_{|U} = (h \circ f)_{|U} = h_{|\mathrm{Im}(f)} \circ f_{|U} = g_{|U} \circ f' \circ f_{|U} = g_{|U}.$$

Zudem gilt $f^t(h)_{|\mathrm{Ker}(f)} = 0 = g_{|\mathrm{Ker}(f)}$ wegen $\big(\mathrm{Ker}(f)\big)^o \supseteq \mathrm{Im}\,(f^t)$ und da $g_{|\mathrm{Ker}(f)} = 0$ ist. Also folgt wie oben $f^t(h) = g$.

□

Als Korollar erhalten wir einen eleganten Beweis, daß Zeilenrang und Spaltenrang einer Matrix übereinstimmen.

Korollar 1.21 a. *Sei* $f \in \mathrm{Hom}_K(V, W)$ *mit* $\dim_K(V), \dim_K(W) < \infty$*, dann gilt*

$$\mathrm{rang}(f) = \mathrm{rang}\,(f^t).$$

b. *Sei* $A \in \mathrm{Mat}(n \times m, K)$*, dann gilt* $\mathrm{Zrang}(A) = \mathrm{Srang}(A)$.

Beweis: a. Mit Hilfe von Aufgabe 1.24 erhalten wir folgende Gleichung

$$\mathrm{rang}\,(f^t) = \dim_K \mathrm{Im}\,(f^t) = \dim_K \mathrm{Ker}(f)^o = \dim_K(V) - \dim_K \mathrm{Ker}(f) = \mathrm{rang}(f).$$

b. Aus Teil a. folgt dann unter Berücksichtigung von Proposition 1.15

$$\mathrm{Srang}(A) = \mathrm{rang}(f_A) = \mathrm{rang}\,(f_A^t) = \mathrm{rang}(f_{A^t}) = \mathrm{Srang}\,(A^t) = \mathrm{Zrang}(A).$$

□

AUFGABEN

Aufgabe 1.22
Es seien V und W zwei endlich-dimensionale K-Vektorräume und $f \in \mathrm{Hom}_K(V,W)$. Zeige:

a. f^t ist genau dann ein Monomorphismus, wenn f ein Epimorphismus ist.
b. f^t ist genau dann ein Epimorphismus, wenn f ein Monomorphismus ist.
c. f^t ist genau dann ein Isomorphismus, wenn f ein Isomorphismus ist.

Aufgabe 1.23
Es sei $\big(V, \langle\cdot,\cdot\rangle\big)$ ein endlich-dimensionaler euklidischer Raum und $f \in \mathrm{End}_{\mathbb{R}}(V)$. Ferner bezeichne $f^t \in \mathrm{End}_{\mathbb{R}}\big(V^*\big)$ die zu f duale Abbildung und $f^* \in \mathrm{End}_{\mathbb{R}}(V)$ die zu f adjungierte Abbildung.

a. Die Abbildung $\phi : V \to V^* : x \mapsto \langle x, \cdot\rangle$ ist ein Isomorphismus.
b. Das folgende Diagramm ist kommutativ:

$$\begin{array}{ccc} V & \xrightarrow{f^*} & V \\ {\scriptstyle\phi}\downarrow & & \downarrow{\scriptstyle\phi} \\ V^* & \xrightarrow[f^t]{} & V^*. \end{array}$$

   Insbesondere gilt $f^* = \phi^{-1} \circ f^t \circ \phi$.
c. Ist $U \subseteq V$ ein Unterraum, dann ist $\phi\big(U^\perp\big) = U^o$.

Hinweis: In Teil b. betrachte man geeignete Matrixdarstellungen der Abbildungen.

Aufgabe 1.24
Es sei V ein endlich-dimensionaler K-Vektorraum und $U \subseteq V$ ein Unterraum von V. Zeige, die *Restriktion* $r : V^* \to U^* : f \mapsto f_{|U}$ ist ein Epimorphismus mit $\mathrm{Ker}(r) = U^o$.

Insbesondere gilt also $\dim_K\big(U^o\big) = \dim_K(V) - \dim_K(U)$.

Aufgabe 1.25
Es sei V ein K-Vektorraum und $U \subseteq V$ ein Unterraum. Wir definieren eine Abbildung

$$i : \big(V/U\big)^* \to V^* : g \mapsto \big(V \to K : v \mapsto g(v+U)\big).$$

Zeige, i ist ein Monomorphismus mit $\mathrm{Im}(i) = U^o$.

Aufgabe 1.26 (Dualität)
Sind $(M,\leq)$ und $(N,\leq)$ zwei teilgeordnete Mengen, dann heißt eine Bijektion $\alpha : M \to N$ eine *Dualität*, falls für $m, m' \in M$ gilt: $m \leq m' \Leftrightarrow \alpha\big(m'\big) \leq \alpha(m)$.

Es sei nun V ein endlich-dimensionaler K-Vektorraum. Mit $L(V)$ bzw. $L\big(V^*\big)$ bezeichnen wir die Mengen der Unterräume von V bzw. V^* - diese sind bezüglich der Inklusion "$\subseteq$" teilgeordnet. Zeige:

a. Für $U \in L(V)$ und $W \in L\big(V^*\big)$ gilt $\big(U^o\big)^{\circledcirc} = U$ und $\big(W^{\circledcirc}\big)^o = W$.
b. Die Abbildungen

$$o : L(V) \to L\big(V^*\big) : U \mapsto U^o$$

und

$$\circledcirc : \mathsf{L}\big(V^*\big) \to \mathsf{L}(V) : W \mapsto W^{\circledcirc}$$

sind zueinander inverse Dualitäten.

Aufgabe 1.27
Es sei V ein endlich-dimensionaler K-Vektorraum und $\mathsf{L}(V)$ sei die Menge der Unterräume von V, teilgeordnet durch die Inklusion. Zeige, daß die Abbildung

$$\perp : \mathsf{L}(V) \to \mathsf{L}(V) : U \mapsto U^{\perp}$$

eine selbstinverse Dualität auf $\mathsf{L}(V)$ ist.

Aufgabe 1.28
Es sei V ein endlich-dimensionaler K-Vektorraum und $U_1, U_2 \subseteq V$ seien zwei Unterräume von V. Zeige:

a. $(U_1 + U_2)^{\circ} = U_1^{\circ} \cap U_2^{\circ}$.
b. $(U_1 \cap U_2)^{\circ} = U_1^{\circ} + U_2^{\circ}$.

Aufgabe 1.29
Es sei V ein endlich-dimensionaler K-Vektorraum und $W_1, W_2 \subseteq V^*$ seien zwei Unterräume von V^*. Zeige:

a. $(W_1 + W_2)^{\circledcirc} = W_1^{\circledcirc} \cap W_2^{\circledcirc}$.
b. $(W_1 \cap W_2)^{\circledcirc} = W_1^{\circledcirc} + W_2^{\circledcirc}$.

Aufgabe 1.30
Es sei R ein kommutativer Ring mit Eins, V und W seien zwei R-Moduln und

$$t : \mathrm{Hom}_R(V, W) \to \mathrm{Hom}_R\left(W^*, V^*\right) : f \mapsto f^t.$$

Zeige:

a. Besitzen V und W Basen, dann ist t ein Monomorphismus von R-Moduln.
b. Ist t in a. auch stets ein Epimorphismus?
c. Finde $\mathbb{Z}$-Moduln V und W, für die t kein Monomorphismus ist.

2. Multilineare Abbildungen und das Tensorprodukt

In diesem Abschnitt werden multilineare Abbildungen systematisch untersucht. Wir haben bereits die Determinante als Beispiel einer Multilinearform kennengelernt, die, aufgefaßt als Abbildung von den quadratischen Matrizen in den Grundring, multilinear in ihren Zeilen und Spalten war. (Vgl. Definition II.9.7.)

Das Tensorprodukt dient unter anderem dazu, solche multilinearen Abbildungen in lineare Abbildungen zu überführen, wobei, notgedrungen, der zugrundeliegende Vektorraum komplizierter wird. Damit steht dann die ganze lineare Algebra insbesondere der Matrixkalkül auch für multilineare Abbildungen zur Verfügung.

Wir werden Tensorprodukte nur für Vektorräume untersuchen, obwohl man dies allgemeiner auch für R-Moduln über einem kommutativen Ring R mit

Eins tun kann. Der Aufwand dafür ist aber höher und die Theorie wird noch abstrakter. Die grundlegenden Definitionen kann man jedoch ohne Mehraufwand auch über Ringen geben.

Generalvoraussetzung: In diesem Kapitel bezeichne K stets einen beliebigen Körper und R einen kommutativen Ring mit Eins.

Wir wollen zunächst den Begriff der multilinearen Abbildung, den wir in Definition II.9.7 eingeführt haben, etwas verallgemeinern.

Definition 2.1

Es seien $V, V_1, \ldots, V_n$ R-Moduln. Eine Abbildung

$$f : V_1 \times \ldots \times V_n \to V$$

heißt *multilinear*, falls f in jedem Argument linear ist, d. h. falls für $x_i, y_i \in V_i$, $i = 1, \ldots, n$ und $\lambda, \mu \in R$ gilt:

$$f(x_1, \ldots, \lambda x_i + \mu y_i, \ldots, x_n) = \lambda f(x_1, \ldots, x_i, \ldots, x_n) + \mu f(x_1, \ldots, y_i, \ldots, x_n).$$

Ist $n = 2$, so nennt man f auch *bilinear*.

Die Menge der multilinearen Abbildungen von $V_1 \times \ldots \times V_n$ nach V bezeichnen wir mit $L(V_1, \ldots, V_n; V) = L_R(V_1, \ldots, V_n; V)$.

Bemerkung 2.2 a. Sind $V, V_1, \ldots, V_n$ R-Moduln, so wollen wir für $x_j \in V_j$, $j = 1, \ldots, n$, mit

$$(x_1, \ldots, \hat{x}_i, \ldots, x_n) := (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$$

und mit

$$V_1 \times \ldots \times \widehat{V}_i \times \ldots \times V_n := \{(x_1, \ldots, \hat{x}_i, \ldots, x_n) \mid x_j \in V_j, j = 1, \ldots, n\}$$

bezeichnen. Der *Hut* über einer Komponente bedeutet also, daß diese Komponente ausgelassen wird - eine sehr nützliche Notation!

Damit ist $f : V_1 \times \ldots \times V_n \to V$ genau dann multilinear, wenn für jedes $i = 1, \ldots, n$ und für alle $(x_1, \ldots, \hat{x}_i, \ldots, x_n) \in V_1 \times \ldots \times \widehat{V}_i \times \ldots \times V_n$ gilt,

$$f_| : V_i \to V : x \mapsto f(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n)$$

ist R-linear.

b. Betrachten wir den R-Modul $W := V^{V_1 \times \ldots \times V_n}$ aller Abbildungen von $V_1 \times \ldots \times V_n$ nach V aus Beispiel II.1.2 g., so ist $L(V_1 \times \ldots \times V_n; V)$ ein Untermodul von W.

Denn, da die Nullabbildung in $L(V_1 \times \ldots \times V_n; V)$ ist, ist $L(V_1 \times \ldots \times V_n; V)$ nicht-leer, und da ferner die Summe zweier multilinearer Abbildungen sowie das skalare Vielfache einer multilinearen Abbildung offenbar wieder multilinear sind, ist $L(V_1 \times \ldots \times V_n; V)$ zudem gegen Addition und Skalarmultiplikation abgeschlossen.

Beispiel 2.3 a. Sei $V_1 = \ldots = V_n = R^n$ und $V = R$, und für Vektoren $x_1, \ldots, x_n \in R^n$ bezeichne $A(x_1, \ldots, x_n) \in \mathrm{Mat}(n, R)$ die Matrix, deren Spalten die Vektoren $x_1, \ldots, x_n$ bilden, dann ist die Abbildung

$$\det : R^n \times \overset{n}{\ldots} \times R^n \to R : (x_1, \ldots, x_n) \mapsto \det\big(A(x_1, \ldots, x_n)\big)$$

multilinear nach Bemerkung II.9.8.

b. Jede Bilinearform auf einem Vektorraum V ist eine bilineare Abbildung. Insbesondere gilt also, ist $\langle\cdot,\cdot\rangle$ ein Skalarprodukt auf $\mathbb{R}^n$, so ist

$$\langle\cdot,\cdot\rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$$

bilinear.
Allerdings ist eine hermitesche Abbildung b auf einem $\mathbb{C}$-Vektorraum *keine* Bilinearform, also auch ein Skalarprodukt auf $\mathbb{C}^n$ nicht, da sie im zweiten Argument nur semilinear ist.

c. Sei $K[t]$ der Polynomring in einer Veränderlichen t, dann ist $K[t]$ unendlich-dimensional mit Basis $(t^0, t, t^2, t^3, \dots)$. Analog wissen wir aus Aufgabe III.1.55, daß der Polynomring $K[x_1, x_2]$ die Basis $\left(x_1^{\nu_1} x_2^{\nu_2} \;\middle|\; \nu_1, \nu_2 \in \mathbb{N}\right)$ besitzt. Die Abbildung

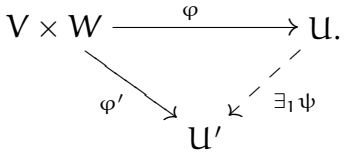
$$K[t] \times K[t] \to K[x_1, x_2] : (f, g) \mapsto f(x_1) \cdot g(x_2)$$

ist bilinear, wie aus der Distributivität sowie der Assoziativität und der Kommutativität der Multiplikation in $K[x_1, x_2]$ folgt.

Um die Notationen nicht zu unübersichtlich zu machen, behandeln wir zunächst den Fall $n = 2$, d. h. bilineare Abbildungen, eingehender.

Definition 2.4 a. Es seien V und W zwei R-Moduln. Ein Paar (U, φ) mit U ein R-Modul und $\varphi : V \times W \to U$ eine bilineare Abbildung heißt *Tensorprodukt* von V und W, wenn (U, φ) der folgenden *universellen Eigenschaft* genügt:

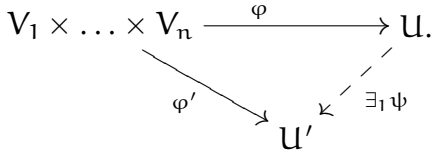
> Für jedes weitere Paar (U', φ') mit U' ein R-Modul und $\varphi' : V \times W \to U'$ eine bilineare Abbildung gilt, es *existiert genau eine lineare* Abbildung $\psi : U \to U'$ mit $\psi \circ \varphi = \varphi'$, d. h. so, daß das folgende Diagramm kommutiert:



Da nach dem folgenden Satz das Tensorprodukt, so es existiert, bis auf eindeutige Isomorphie eindeutig bestimmt ist, schreibt man $V \otimes W = V \otimes_R W$ statt U und unterschlägt - um Notation zu sparen - für gewöhnlich die bilineare Abbildung φ ganz.

b. Es seien $V_1, \dots, V_n$ R-Moduln. Ein Paar (U, φ) mit U ein R-Modul und $\varphi : V_1 \times \dots \times V_n \to U$ eine multilineare Abbildung heißt *Tensorprodukt* von $V_1, \dots, V_n$, wenn (U, φ) der folgenden *universellen Eigenschaft* genügt:

Für jedes weitere Paar (U', φ') mit U' ein R-Modul und $\varphi' : V_1 \times \ldots \times V_n \to U'$ eine multilineare Abbildung gilt, es *existiert genau eine lineare* Abbildung $\psi : U \to U'$ mit $\psi \circ \varphi = \varphi'$, d. h. so, daß das folgende Diagramm kommutiert:



Analog schreibt man wieder $V_1 \otimes \ldots \otimes V_n$ statt U und unterschlägt die multilineare Abbildung φ ganz.

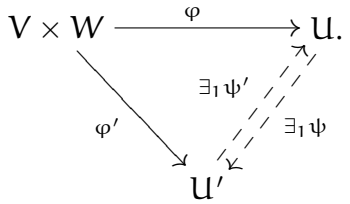
c. Wir nennen die Elemente des Tensorproduktes auch *Tensoren* und die Elemente in $\mathrm{Im}(\varphi)$ *reine Tensoren*.

Es ist - wie stets bei Objekten, die durch universelle Eigenschaften definiert werden (vgl. etwa Aufgabe III.1.38) - kein Problem, die *Eindeutigkeit* festzustellen, vorausgesetzt sie existieren. Hierbei soll *eindeutig* bedeuten, daß jedes weitere Objekt, das dieser universellen Eigenschaft genügt, isomorph zu dem gegebenen ist, und daß mehr noch, der Isomorphismus ebenfalls eindeutig bestimmt ist.

Satz 2.5 a. *Es seien V und W zwei R-Moduln und (U, φ) und (U', φ') seien zwei Tensorprodukte von V und W. Dann gibt es genau einen Isomorphismus $\psi : U \to U'$ mit $\psi \circ \varphi = \varphi'$.*

b. *Es seien $V_1, \ldots, V_n$ R-Moduln und (U, φ) und (U', φ') seien zwei Tensorprodukte von $V_1, \ldots, V_n$. Dann gibt es genau einen Isomorphismus $\psi : U \to U'$ mit $\psi \circ \varphi = \varphi'$.*

Beweis: a. Aus der universellen Eigenschaft, der sowohl (U, φ) als auch (U', φ') genügen, folgt, daß es zwei eindeutig bestimmte Abbildungen $\psi : U \to U'$ und $\psi' : U' \to U$ gibt, die das folgende Diagramm zum kommutieren bringen:



Damit gilt aber auch

$$\psi' \circ \psi \circ \varphi = \psi' \circ \varphi' = \varphi \tag{69}$$

und

$$\psi \circ \psi' \circ \varphi' = \psi \circ \varphi = \varphi'. \tag{70}$$

Und aus der Eindeutigkeit folgt unmittelbar, daß nur ψ als Kandidat für den Isomorphismus in Frage kommt!

Betrachten wir nun die bilineare Abbildung $\varphi : V \times W \to U$ selbst, so sagt die universelle Eigenschaft von (U, φ), daß es genau eine lineare

Abbildung $\pi : U \to U$ gibt mit $\pi \circ \varphi = \varphi$. Offensichtlich ist id_U eine lineare Abbildung, die diese Eigenschaft besitzt, und mithin gilt $\pi = \mathrm{id}_U$. Andererseits gilt nach (69) aber, daß $\psi' \circ \psi$ ebenfalls diese Eigenschaft hat, also

$$\psi' \circ \psi = \pi = \mathrm{id}_U .$$

Analog folgt aus (70) und der universellen Eigenschaft von $\left(U', \varphi'\right)$, daß auch

$$\psi \circ \psi' = \pi = \mathrm{id}_{U'} .$$

Mithin ist ψ ein Isomorphismus mit ψ' als Inverser.

b. Der Beweis ist wörtlich der gleiche wie in Teil a.

□

Bemerkung 2.6

War es noch kein Problem, die Eindeutigkeit des Tensorproduktes auch im Fall von beliebigen Moduln über kommutativen Ringen mit Eins zu zeigen, so bereitet die Existenz in diesem Fall doch weit größere Probleme als im Fall von Vektorräumen über einem Körper, und wir wollen uns deshalb für den Rest des Kapitels auf die Betrachtung von Vektorräumen über einem Körper K beschränken.

Wir wollen es aber nicht versäumen, eine Referenz für den Beweis zu geben, daß Tensorprodukte von beliebigen (auch unendlichen) Familien beliebiger Moduln über beliebigen kommutativen Ringen mit Eins existieren. Die Literatur hierzu ist vielfältig, und eine mögliche Referenz ist [**SS88**] §80, S. 554 sowie Satz 80.1.

Um die Existenz des Tensorproduktes zu zeigen, benötigen wir einige neue Begriffe, zum Beispiel den der äußeren direkten Summe von Vektorräumen, die ein Mittel darstellt, aus gegebenen Vektorräumen, neue zu gewinnen.

Bisher hatten wir nur die direkte Summe von Untervektorräumen $(U_i \mid i \in I)$ eines Vektorraumes V betrachtet. Diese wird auch als *innere* direkte Summe bezeichnet. Wir wollen nun die direkte Summe einer Familie von Vektorräumen $(V_i \mid i \in I)$ definieren, die zunächst nicht notwendig Unterräume eines gemeinsamen Vektorraumes V sind.

Definition 2.7

Es sei $(V_i \mid i \in I)$ eine Familie von K-Vektorräumen. Dann heißt

$$\bigoplus_{i \in I} V_i := \left\{ (x_i)_{i \in I} \in \prod_{i \in I} V_i \;\middle|\; x_i \neq 0 \text{ für höchstens endlich viele } i \in I \right\}$$

die *äußere direkte Summe* der V_i.

Für die Elemente $x = (x_i)_{i \in I}$ von $\bigoplus_{i \in I} V_i$ schreiben wir auch

$$x = (x_i)_{i \in I} = \sum_{i \in I}{}' x_i,$$

wobei der Strich an der Summe wieder darauf hinweist, daß nur endlich viele *Summanden* ungleich Null sind.[3]

Bemerkung 2.8

Es sei $(V_i \mid i \in I)$ eine Familie von K-Vektorräumen. Dann werden $V = \prod_{i\in I} V_i$ bzw. $V = \bigoplus_{i\in I} V_i$ durch die komponentenweise Addition und Skalarmultiplikation, d. h. durch

$$(x_i)_{i\in I} + (y_i)_{i\in I} := (x_i + y_i)_{i\in I}$$

und

$$\lambda \cdot (x_i)_{i\in I} := (\lambda x_i)_{i\in I}$$

für $\lambda \in K$ und $(x_i)_{i\in I}, (y_i)_{i\in I} \in V$, zu K-Vektorräumen, wie man leicht nachprüft.

Ist $I = \{1, \ldots, n\}$ eine endliche Menge, dann gilt per definitionem

$$V_1 \times \ldots \times V_n = \prod_{i=1}^{n} V_i = \bigoplus_{i=1}^{n} V_i,$$

während für $|I| = \infty$ offenbar $\bigoplus_{i\in I} V_i$ ein echter Unterraum von $\prod_{i\in I} V_i$ ist.

Die natürlichen Abbildungen

$$\epsilon_j : V_j \to \prod_{i\in I} V_i : x \mapsto (x_i \mid x_j = x, x_i = 0 \ \forall\, i \neq j)$$

und

$$\delta_j : V_j \to \bigoplus_{i\in I} V_i : x \mapsto (x_i \mid x_j = x, x_i = 0 \ \forall\, i \neq j)$$

sind Vektorraum*mono*morphismen, mittels derer wir die V_i als Unterräume von $\prod_{j\in I} V_j$ bzw. $\bigoplus_{j\in I} V_j$ auffassen können - d. h. durch Identifikation mit $\epsilon_i(V_i)$ bzw. $\delta_i(V_i)$.

Korollar 2.9

Es sei $(V_i \mid i \in I)$ *eine Familie von* K*-Vektorräumen.*

a. *Die äußere direkte Summe* $\bigoplus_{i\in I} V_i$ *ist die* innere *direkte Summe der* $\delta_i(V_i)$, $i \in I$.

b. *Ist für* $i \in I$ *die Familie* $(x_{ij} \mid j \in J_i)$ *eine Basis von* V_i *und setzen wir* $x'_{ij} := \delta_i(x_{ij}) = (y_k \mid y_i = x_{ij}, y_k = 0 \ \forall\, k \neq i)$ *für* $i \in I$ *und* $j \in J_i$*, dann ist*

$$\left(x'_{ij} \;\middle|\; i \in I, j \in J_i\right)$$

eine Basis von $\bigoplus_{i\in I} V_i$.

[3] Wir erinnern uns, daß wir bereits in Definition I.2.1 - unter Berücksichtigung von Beispiel I.2.12 e. - das kartesische Produkt $\prod_{i\in I} V_i$ eingeführt haben, als die Menge der Familien $x = (x_i)_{i\in I}$ mit $x_i \in V_i$ für $i \in I$, d. h. formal betrachtet, daß $x : I \to \coprod_{i\in I} V_i$ eine Abbildung ist mit $x(i) = x_i \in V_i$, wobei $\coprod_{i\in I} V_i := \bigcup_{i\in I} (\{i\} \times V_i) \subseteq I \times \bigcup_{i\in I} V_i$ die disjunkte Vereinigung der V_i bezeichnet. Es erscheint hier aber nicht notwendig, diesen formalen Gesichtspunkt überzustrapazieren.

c. *Ist* $I = \{1, \ldots, n\}$ *und* $\dim_K(V_i) < \infty$ *für* $i \in I$*, dann gilt*

$$\dim_K\left(\bigoplus_{i=1}^{n} V_i\right) = \sum_{i=1}^{n} \dim_K(V_i).$$

Beweis: a. Da ein Vektor $x = (x_i)_{i\in I} \in \bigoplus_{i\in I} V_i$ nur endlich viele Komponenten ungleich Null besitzt, ist die Linearkombination

$$x = \sum_{i\in I} \delta_i(x_i) \in \sum_{i\in I} \delta_i(V_i)$$

endlich, und damit $\bigoplus_{i\in I} V_i = \sum_{i\in I} \delta_i(V_i)$.
Zudem folgt aus $0 = \sum'_{i\in I} \delta_i(x_i) = (x_i)_{i\in I}$, daß $x_i = 0$ für alle $i \in I$. D. h. die Null besitzt eine eindeutige Darstellung als Linearkombination von Vektoren aus den $\delta_i(V_i)$, und somit ist die Summe der $\delta_i(V_i)$ direkt.

b./c. Dies folgt unmittelbar aus a.

□

Allgemeiner gilt die folgende Beziehung.

Lemma 2.10

Sei V *ein* K*-Vektorraum und* $(V_i \mid i \in I)$ *eine Familie von Unterräumen von* V*. Ferner bezeichne* $U = \bigoplus_{i\in I} V_i$ *die äußere direkte Summe der Vektorräume* V_i*,* $i \in I$*, und* $W = \sum_{i\in I} V_i$ *die Summe der* V_i *in* V *- vgl. Definition II.4.1.*

Dann ist W *genau dann eine innere direkte Summe, wenn die Abbildung*

$$\delta : U \to W : (x_i)_{i\in I} \mapsto \sum_{i\in I}{}' x_i$$

ein Isomorphismus ist.

Beweis: Wir wollen zunächst einmal festhalten, daß die Definition von δ Sinn macht, da für $(x_i)_{i\in I}$ nur endlich viele der x_i ungleich Null sind, und daß δ offenbar stets ein Epimorphismus ist.

Nun ist δ genau dann injektiv, wenn für $x = (x_i)_{i\in I}$ aus $0 = \delta(x) = \sum'_{i\in I} x_i$ bereits $x_i = 0$ für alle $i \in I$ folgt. Aber das ist gleichwertig dazu, daß die Null in $\sum_{i\in I} V_i$ eine eindeutige Darstellung als Linearkombination von Elementen aus den V_i besitzt, sprich, daß die Summe eine innere direkte Summe ist. □

Beispiel 2.11 a. Ist I eine beliebige Menge und ist $V_i = K$ für alle $i \in I$, dann ist

$$\bigoplus_{i\in I} K = \bigoplus_{i\in I} V_i = \{(\lambda_i)_{i\in I} \mid \lambda_i \in K, \lambda_i \neq 0 \text{ nur für endlich viele } i \in I\}.$$

In diesem Fall wäre es etwas unangenehm, statt $(\lambda_i)_{i\in I}$ kurz $\sum'_{i\in I} \lambda_i$ zu schreiben, da die Summe dann auch fälschlicherweise als Summe in K aufgefaßt werden könnte. Es macht sich von daher ganz gut, die Elemente in $V_i = K$ statt als λ als $\lambda \cdot i$ zu schreiben und somit letztlich zu

erhalten

$$\bigoplus_{i\in I} K = \left\{ \sum_{i\in I}{}' \lambda_i \cdot i \;\middle|\; \lambda_i \in K, \lambda_i \neq 0 \text{ nur für endlich viele } i \in I \right\}.$$

Aber auch dieses war mehr ein formales Problem, denn ein inhaltliches!

b. Ist in a. $I = \mathbb{N}$ oder besser $I = \{t^i \mid i \in \mathbb{N}\}$, dann ist $\bigoplus_{i\in I} K$ als K-Vektorraum in natürlicher Weise isomorph zum Polynomring $K[t]$.
In diesem Falle würde man das direkte Produkt $\prod_{i\in I} K$ als Vektorraum mit dem *Ring der formalen Potenzreihen*

$$K[[t]] := \left\{ \sum_{n=0}^{\infty} a_n t^n \;\middle|\; a_n \in K \right\}$$

identifizieren. Da wir diesen in der Vorlesung bisher nicht eingeführt haben, mag man dies auch als Definition ansehen.

Wir wollen nun wieder zu dem Fall zweier Vektorräume zurückkehren

Bemerkung 2.12
Es seien V und W K-Vektorräume. Ist $B = (x_i \mid i \in I)$ eine Basis von V und $D = (y_j \mid j \in J)$ eine Basis von W, dann ist

$$C := \big((x_i, 0), (0, y_j) \mid i \in I, j \in J\big)$$

also eine Basis des Vektorraumes $V \times W = V \oplus W$.

Wir hatten früher schon gesehen, daß eine lineare Abbildung durch Werte auf der Basis eindeutig bestimmt ist. Für multilineare Abbildungen gilt dies nicht direkt, sondern nur nach geeigneter Modifikation.

Wir können die Werte auf der Menge $\mathcal{M} := \{(x_i, y_j) \mid (i,j) \in I \times J\}$ eindeutig vorschreiben und erhalten so eine eindeutige bilineare Abbildung von $V \times W$ in einen beliebigen Vektorraum U - vgl. Lemma III.1.1.

Man beachte, daß die Menge der Tupel $\mathcal{M} := \{(x_i, y_j) \in V \times W \mid (i,j) \in I \times J\}$, im allgemeinen weder linear unabhängig noch ein Erzeugendensystem von $V \times W$ ist. – Im endlich-dimensionalen Fall sieht man schon aus Dimensionsgründen, $\dim_K(V \times W) = \dim_K(V) + \dim_K(W)$ und $|I \times J| = \dim_K(V) \cdot \dim_K(W)$, daß $\mathcal{M}$ linear abhängig sein muß.

Ferner ist offensichtlich, daß eine bilineare Abbildung f nicht eindeutig bestimmt sein kann durch ihre Werte auf C, da für $(x_i, 0), (0, y_j) \in C$ stets gilt:

$$f(x_i, 0) = f(x_i, 0 \cdot 0) = 0 \cdot f(x_i, 0) = 0$$

und

$$f(0, y_i) = f(0 \cdot 0, y_i) = 0 \cdot f(0, y_i) = 0,$$

d. h. jede bilineare Abbildung wird auf C nur den Wert Null annehmen.

Fassen wir die wesentlichen Aussagen der Bemerkung in folgendem Lemma zusammen.

Lemma 2.13
Es seien V, W *und* U *drei* K*-Vektorräume und* $B = (x_i \mid i \in I)$ *sowie* $D = (y_j \mid j \in J)$ *Basen von* V *bzw.* W*. Ferner sei* $(z_{ij} \mid i \in I, j \in J)$ *eine beliebige Familie in* U*. Dann gibt es genau eine bilineare Abbildung* $f : V \times W \to U$ *mit*

$$f\big((x_i, y_j)\big) = z_{ij} \quad \forall\, (i,j) \in I \times J.$$

Sind $x = \sum'_{i\in I} \lambda_i x_i \in V$ *und* $y = \sum'_{j\in J} \mu_j y_j \in W$*, so gilt*

$$f(x,y) = \sum_{i\in I}{}' \sum_{j\in J}{}' \lambda_i \mu_j z_{ij}.$$

Beweis: Der Beweis ist eine offensichtliche Verallgemeinerung des Beweises von Lemma III.1.1 und sei dem Leser als Übungsaufgabe überlassen. □

Nun sind wir in der Lage, die Existenz des Tensorproduktes $V \otimes W$ zu zeigen, wobei wir den Vektorraum als äußere direkte Summe von Vektorräumen so konstruieren, daß jedem Tupel (x_i, y_j) aus $B \times D$ ein Element $x_i \otimes y_j$ in $V \otimes W$ entspricht, so daß $\big(x_i \otimes y_j \;\big|\; (i,j) \in I \times J\big)$ eine Basis von $V \otimes W$ ist.
Satz 2.14
Seien V *und* W *zwei* K*-Vektorräume, dann gibt es einen Vektorraum* $V \otimes W$ *und eine bilineare Abbildung* $\varphi : V \times W \to V \otimes W$*, so daß das Tupel* $(V \otimes W, \varphi)$ *ein Tensorprodukt von* V *und* W *ist.*

Für $x \in V$ *und* $y \in W$ *definieren wir* $x \otimes y := \varphi(x,y)$*, d. h. die reinen Tensoren sind genau die Tensoren der Form* $x \otimes y$*.*

Beweis: Seien $B = (x_i \mid i \in I)$ sowie $D = (y_j \mid j \in J)$ Basen von V bzw. W. Ferner sei $\mathcal{B} = (x_i \otimes y_j \mid i \in I, j \in J)$ eine Menge von paarweise verschiedenen Symbolen. Dann setzen wir

$$V \otimes W := \bigoplus_{(i,j)\in I\times J} K = \bigoplus_{x_i\otimes y_j \in \mathcal{B}} K = \left\{ \sum\nolimits'_{(i,j)\in I\times J} \lambda_{ij} x_i \otimes y_j \;\middle|\; \lambda_{ij} \in K \right\}.$$

Man beachte, daß der Strich an der Summe bedeutet, daß nur endlich viele Koeffizienten ungleich Null sind!

Nach Korollar 2.9 ist $\mathcal{B}$ dann eine Basis von $V \otimes W$.

Ferner impliziert Lemma 2.13, daß es genau eine bilineare Abbildung

$$\varphi : V \times W \to V \otimes W : (x_i, y_j) \mapsto x_i \otimes y_j, \quad (i,j) \in I \times J,$$

gibt.[4]

Für $x = \sum'_{i\in I} \lambda_i x_i \in V$ und $y = \sum'_{j\in J} \mu_j y_j \in W$ folgt dann aus dem gleichen Lemma

$$x \otimes y := \varphi(x,y) = \sum_{i\in I}{}' \sum_{j\in J}{}' \lambda_i \mu_j x_i \otimes y_j. \tag{71}$$

[4]Eine andere Wahl blieb uns nicht, wenn wir der Definition von $x \otimes y$ aus der Behauptung des Satzes im Fall $x = x_i$ und $y = y_j$ gerecht werden wollen.

Wir haben den Vektorraum $V \otimes W$ sowie die Bilinearform φ nun definiert, und es bleibt zu zeigen, daß sie der universellen Eigenschaft genügen.

Sei dazu U' ein beliebiger K-Vektorraum und $\varphi' : V \times W \to U'$ eine bilineare Abbildung. Dann setzen wir $z_{ij} := \varphi'(x_i, y_j)$ für $(i,j) \in I \times J$. Da $\mathcal{B}$ eine Basis von $V \otimes W$ ist, gibt es genau eine lineare Abbildung

$$\psi : V \otimes W \to U' : x_i \otimes y_j \mapsto z_{ij}, \quad (i,j) \in I \times J.$$

Aber dann gilt für $x = \sum'_{i\in I} \lambda_i x_i \in V$ und $y = \sum'_{j\in J} \mu_j y_j \in W$ wegen (71)

$$\begin{aligned}(\psi \circ \varphi)(x,y) &= \psi\big(x \otimes y\big) = \textstyle\sum'_{i\in I}\sum'_{j\in J} \lambda_i\mu_j\psi(x_i \otimes y_j)\\ &= \textstyle\sum'_{i\in I}\sum'_{j\in J} \lambda_i\mu_j z_{ij} = \sum'_{i\in I}\sum'_{j\in J} \lambda_i\mu_j\varphi'(x_i, y_j)\\ &= \varphi'(x,y).\end{aligned}$$

Aber damit gilt $\psi \circ \varphi = \varphi'$. □

Aus dem Beweis des Satzes folgt unter Berücksichtigung der Eindeutigkeit des Tensorproduktes unmittelbar das folgende Korollar.

Korollar 2.15

Sind $(x_i \mid i \in I)$ *sowie* $(y_j \mid j \in J)$ *Basen von* V *bzw.* W*, so besitzt* $V \otimes W$ *eine Basis* $(x_i \otimes y_j \mid i \in I, j \in J)$.

Insbesondere, ist $\dim_K(V) = n < \infty$ *und* $\dim_K(W) = m < \infty$*, dann ist* $\dim_K(V \otimes W) = n \cdot m$.

Eine offensichtliche Verallgemeinerung von Lemma 2.13 und des Beweises von Satz 2.14 beweist die Existenz des Tensorproduktes $V_1 \otimes \ldots \otimes V_n$. Wir überlassen die Details dem Leser.

Satz 2.16

Seien $V_1, \ldots, V_n$ K*-Vektorräume, so gibt es ein Tensorprodukt* $(V_1 \otimes \ldots \otimes V_n, \varphi)$ *von* $V_1, \ldots, V_n$.

Ist dabei $(x_{ij} \mid j \in J_i)$ *eine Basis von* V_i*,* $i = 1, \ldots, n$*, dann besitzt* $V_1 \otimes \ldots \otimes V_n$ *eine Basis der Form*

$$\big(x_{1j_1} \otimes \ldots \otimes x_{nj_n} \;\big|\; (j_1, \ldots, j_n) \in J_1 \times \ldots \times J_n\big).$$

Insbesondere gilt, sind die V_i *alle endlich-dimensional, dann ist*

$$\dim_K(V_1 \otimes \ldots \otimes V_n) = \dim_K(V_1) \cdots \dim_K(V_n).$$

Für $x_i \in V_i$*,* $i = 1, \ldots, n$*, so schreiben wir*

$$x_1 \otimes \ldots \otimes x_n := \varphi(x_1, \ldots, x_n),$$

d. h. die reinen Tensoren sind genau die Tensoren der Form $x_1 \otimes \ldots \otimes x_n$.

Wir wollen nun zunächst einige Rechenregeln für Tensoren herleiten.

Lemma 2.17

Sind V, W *zwei* K*-Vektorräume, dann gelten für* $x, x' \in V$ *und* $y, y' \in W$ *und* $\lambda \in K$ *die folgenden Rechenregeln:*

a. $x \otimes \big(y + y'\big) = x \otimes y + x \otimes y'$ *und* $\big(x + x'\big) \otimes y = x \otimes y + x' \otimes y$.

b. $\lambda(x \otimes y) = (\lambda x) \otimes y = x \otimes (\lambda y)$.
Insbesondere gilt $0 \otimes y = x \otimes 0 = 0$.

c. *Jeder Tensor in* $V \otimes W$ *besitzt eine (nicht eindeutige) Darstellung als endliche Summe von reinen Tensoren, d. h. für* $z \in V \otimes W$ *gibt es Elemente* $v_i \in V$ *und* $w_i \in W$, $i = 1, \dots, r$, *mit*

$$z = \sum_{i=1}^{r} v_i \otimes w_i.$$

Beweis: Beachten wir, daß für $x \in V$ und $y \in W$ gilt $x \otimes y = \varphi(x, y)$, dann folgen die Behauptungen a. und b. aus der Bilinearität von φ.

Es bleibt also, die Aussage in c. zu zeigen. Seien dazu $(x_i \mid i \in I)$ sowie $(y_j \mid j \in J)$ Basen von V bzw. W. Dann besitzt $z \in V \otimes W$ eine Darstellung der Form

$$z = \sum_{(i,j)\in I\times J}{}' \lambda_{ij}(x_i \otimes y_j) = \sum_{(i,j)\in I\times J}{}' (\lambda_{ij}x_i) \otimes y_j = \sum_{(i,j)\in I\times J}{}' x_i \otimes (\lambda_{ij}y_j).$$

Damit ist die Aussage gezeigt, und es ist zugleich auch klar, daß die Darstellung nicht eindeutig sein kann. □

Bemerkung 2.18 a. Man sollte beachten, daß φ i. a. weder injektiv noch surjektiv ist!

b. Ferner ist es wichtig, festzuhalten, daß jedes Element des Tensorproduktes eine endliche Summe von reinen Tensoren ist, daß aber i. a. *nicht* jedes Element selbst ein reiner Tensor ist - vgl. Beispiel 2.19 c.!

c. Die wichtigste Feststellung ist aber die, daß wir eine lineare Abbildung auf dem Tensorprodukt $V \otimes W$ dadurch in eindeutiger Weise festlegen können, daß wir die Bilder der reinen Tensoren $x \otimes y$, $x \in V$, $y \in W$, beliebig so vorgeben, daß die Vorgabe linear im ersten und linear im zweiten Argument ist - formaler gesagt, indem wir eine bilineare Abbildung auf $V \times W$ vorgeben. Das ist genau die Aussage der universellen Eigenschaft!

Kommen wir nun aber zu Beispielen, die verdeutlichen, daß das Tensorprodukt eine sehr hilfreiche Konstruktion ist.

Beispiel 2.19 a. Es sei V ein K-Vektorraum. Die natürliche Abbildung

$$\psi : V \otimes K \xrightarrow{\cong} V : x \otimes \lambda \mapsto \lambda x$$

ist ein Vektorraumisomorphismus mit inverser Abbildung

$$\psi^{-1} : V \to V \otimes K : x \mapsto x \otimes 1.$$

Insbesondere gilt in diesem Fall, daß jeder Tensor ein reiner Tensor ist. Man beachte auch, daß für die bilineare Abbildung $\varphi : V \times K \to V : (x, \lambda) \mapsto \lambda x$ mithin gilt:

$$\varphi^{-1}(0) = \big(V \times \{0\}\big) \cup \big(\{0\} \times K\big),$$

da $\varphi(x, \lambda) = 0$ gleichwertig dazu ist, daß $\lambda x = 0$, also $\lambda = 0$ oder $x = 0$. Dann ist aber $\varphi^{-1}(0)$ *kein Unterraum* von $V \times K$! (φ ist halt nur bilinear, nicht aber linear!)

Um die obige Aussage zu verifizieren, beachtet man zunächst, daß φ aufgrund der Vektorraumaxiome bilinear ist und daß ψ gerade die von φ induzierte Abbildung auf dem Tensorprodukt ist - sprich, daß sie durch Vorgabe auf den reinen Tensoren bereits eindeutig beschrieben ist. Ferner ist klar, daß ψ surjektiv ist, da $x \in V$ das Bild von $x \otimes 1$ unter ψ ist.

Wir wollen nun noch zeigen, daß ψ auch injektiv ist. Dazu reicht es zu zeigen, daß jeder Tensor in $V \otimes K$ ein reiner Tensor der Form $x \otimes 1$ ist, da dann $0 = \psi(x \otimes 1) = x$ impliziert, daß $x \otimes 1 = 0 \otimes 1 = 0$. Sei dazu $z = \sum_{i=1}^{r} v_i \otimes \lambda_i \in V \otimes K$ gegeben. Dann gilt

$$z = \sum_{i=1}^{r} (v_i \otimes \lambda_i) = \sum_{i=1}^{r} (\lambda_i v_i \otimes 1) = \left(\sum_{i=1}^{r} \lambda_i v_i \right) \otimes 1.$$

Die Aussage ist also gezeigt, und zugleich sehen wir, daß die obige Beschreibung der Inversen korrekt ist.

b. Sei V ein K-Vektorraum. Die natürliche Abbildung

$$\psi : V \otimes K^n \xrightarrow{\cong} V^n = \bigoplus_{i=1}^{n} V : x \otimes (\lambda_1, \ldots, \lambda_n)^t \mapsto (\lambda_1 x, \ldots, \lambda_n x)$$

ist ein Vektorraum*iso*morphismus.

Dies leiten wir aus der universellen Eigenschaft des Tensorproduktes her.

Wir betrachten dazu die bilineare Abbildung $\varphi' : V \times K^n \to V^n : \left(x, (\lambda_1, \ldots, \lambda_n)^t\right) \mapsto (\lambda_1 x, \ldots, \lambda_n x)$. Sie induziert zunächst einmal die obige Abbildung ψ.

Wir behaupten nun aber, daß das Paar $\left(V^n, \varphi'\right)$ ebenfalls der universellen Eigenschaft des Tensorproduktes von V mit K^n genügt, und dann sind $V \otimes K^n$ und V^n isomorph, wobei die induzierte Abbildung ψ dann just dieser eindeutige Isomorphismus ist.

Sei nun dazu $\varphi'' : V \times K^n \to U$ eine bilineare Abbildung. Dann definieren wir eine Abbildung

$$\psi' : V^n \to U : (x_1, \ldots, x_n) \mapsto \varphi'' \left(\sum_{i=1}^{n} (x_i, e_i) \right).$$

Da φ'' bilinear ist, ist ψ' linear, und zudem folgt unmittelbar $\varphi'' = \psi' \circ \varphi'$.

c. Wir zeigen nun, daß i. a. nicht jedes Element ein reiner Tensor ist.

Angenommen, $e_1 \otimes e_2 + e_2 \otimes e_1 \in K^2 \otimes K^2$ wäre ein reiner Tensor. Dann gibt es Vektoren $x = \lambda_1 e_1 + \lambda_2 e_2, y = \lambda_1' e_1 + \lambda_2' e_2 \in K^2$ mit

$$e_1 \otimes e_2 + e_2 \otimes e_1 = \lambda_1 \lambda_1' e_1 \otimes e_1 + \lambda_1 \lambda_2' e_1 \otimes e_2 + \lambda_2 \lambda_1' e_2 \otimes e_1 + \lambda_2 \lambda_2' e_2 \otimes e_2.$$

Da $\left(e_i \otimes e_j \mid i, j \in \{1, 2\}\right)$ eine Basis von $K^2 \otimes K^2$ ist, folgt durch Koeffizientenvergleich

$$\lambda_1 \lambda_1' = \lambda_2 \lambda_2' = 0 \quad \text{und} \quad \lambda_1 \lambda_2' = \lambda_2 \lambda_1' = 1,$$

was aber nicht sein kann.

d. In Beispiel 2.3 haben wir gezeigt, daß die Abbildung

$$K[t] \times K[t] \to K[x_1, x_2] : (f, g) \mapsto f(x_1) \cdot g(x_2)$$

bilinear ist. Folglich existiert nach der universellen Eigenschaft des Tensorproduktes genau eine *lineare* Abbildung

$$K[t] \otimes_K K[t] \to K[x_1, x_2] : f \otimes g \mapsto f(x_1) \cdot g(x_2).$$

Diese Abbildung ist in der Tat ein *Isomorphismus*!
Um das zu sehen, reicht es zu sehen, daß eine Basis auf eine Basis abgebildet wird. Aber $\left(t^i \otimes t^j \;\middle|\; i, j \in \mathbb{N}\right)$ ist eine Basis von $K[t] \otimes K[t]$ und das Bild unter obiger Abbildung $\left(x_1^i x_2^j \;\middle|\; i, j \in \mathbb{N}\right)$ ist nach Aufgabe III.1.55 eine Basis von $K[x_1, x_2]$.
Der Polynomring in zwei Veränderlichen ist ein gutes Beispiel, um sich das Tensorprodukt allgemein zu veranschaulichen!

e. Es sei V ein $\mathbb{R}$ Vektorraum. Da $\mathbb{C}$ ebenfalls ein $\mathbb{R}$-Vektorraum ist, können wir das Tensorprodukt

$$V_{\mathbb{C}} := \mathbb{C} \otimes_{\mathbb{R}} V$$

bilden. $V_{\mathbb{C}}$ heißt *Komplexifizierung* von V und wird mittels der Skalarmultiplikation, die durch

$$\lambda \cdot (\mu \otimes x) = (\lambda\mu) \otimes x$$

für $\lambda, \mu \in \mathbb{C}$ und $x \in V$ definiert wird,[5] zu einem $\mathbb{C}$-Vektorraum.
Die Komplexifizierung kann man sich sehr schön mit Hilfe von Basen klar machen. Ist etwa $B = (x_j \mid j \in J)$ eine Basis von V als $\mathbb{R}$-Vektorraum, dann ist $B_{\mathbb{C}} = (1 \otimes x_j \mid j \in J)$ eine Basis von $V_{\mathbb{C}}$ als $\mathbb{C}$-Vektorraum, insbesondere läßt sich jedes Element $x \in \mathbb{C} \otimes_{\mathbb{R}} V$ eindeutig darstellen als

$$\sum_{j \in J}{}' \lambda_j (1 \otimes x_j) = \sum_{j \in J}{}' \lambda_j \otimes x_j,$$

mit $\lambda_j \in \mathbb{C}$, $j \in J$. Dies folgt unmittelbar daraus, daß $(1 \otimes x_j, i \otimes x_j \mid j \in J)$ eine Basis von $\mathbb{C} \otimes_{\mathbb{R}} V$ als $\mathbb{R}$-Vektorraum ist.
Man beachte auch folgendes. Ist $J = \{1, \ldots, n\}$, so liefert die Basis B zunächst einen Isomorphismus $\phi_B : V \to \mathbb{R}^n$ von $\mathbb{R}$-Vektorräumen, und dieser induziert einen Isomorphismus

$$\mathrm{id}_{\mathbb{C}} \otimes \phi_B : V_{\mathbb{C}} = \mathbb{C} \otimes_{\mathbb{R}} V \to \mathbb{C} \otimes_{\mathbb{R}} \mathbb{R}^n : \lambda \otimes x_j \mapsto \lambda \otimes e_j.$$

Nach Beispiel a. gilt aber $\mathbb{C} \otimes_{\mathbb{R}} \mathbb{R}^n \cong \mathbb{C}^n$ als $\mathbb{R}$-Vektorräume, so daß wir einen Isomorphismus

$$V_{\mathbb{C}} \to \mathbb{C}^n : 1 \otimes x_j \mapsto e_j$$

[5] Wir überlassen es dem Leser, mit Hilfe der universellen Eigenschaft des Tensorproduktes zu zeigen, daß diese Vorgabe auf den reinen Tensoren in der Tat zu einem Skalarprodukt führt. – Man beachte auch, daß der Ausdruck $\mu \otimes (\lambda x)$ nur für $\lambda \in \mathbb{R}$ Sinn macht!

erhalten. Aber, dieser Isomorphismus ist a priori nur $\mathbb{R}$-linear, daß er in der Tat auch $\mathbb{C}$-linear ist, ist noch zu zeigen - allerdings bedarf es dazu nicht mehr als des Einsatzes der Definition der Skalarmultiplikation auf $V_{\mathbb{C}}$.

Lemma 2.20
Seien U, V *und* W K*-Vektorräume. Dann existieren eindeutig bestimmte* Iso*morphismen:*

a. $V \otimes W \xrightarrow{\cong} W \otimes V : x \otimes y \mapsto y \otimes x$*;*
b. $(V \otimes W) \otimes U \xrightarrow{\cong} V \otimes (W \otimes U) \xrightarrow{\cong} V \otimes W \otimes U : (x \otimes y) \otimes z \mapsto x \otimes (y \otimes z) \mapsto x \otimes y \otimes z$*;*
c. $(V \oplus W) \otimes U \xrightarrow{\cong} (V \otimes U) \oplus (W \otimes U) : (x, y) \otimes z \mapsto (x \otimes z, y \otimes z)$*;*
d. $K \otimes V \xrightarrow{\cong} V : \lambda \otimes x \mapsto \lambda x$.

Beweis: Daß die Abbildungen existieren, eindeutig bestimmt sind und Isomorphismen sind, kann man leicht nachprüfen, indem man Basen der Vektorräume betrachtet.

Alternativ kann man jedoch auch ohne Basen nur mit der universellen Eigenschaft des Tensorproduktes arbeiten, was den Vorteil hat, daß die Beweise im Fall von Moduln über einem kommutativen Ring ebenfalls gültig bleiben. Wir wollen dies beispielhaft im Fall a. vorführen und überlassen die übrigen Fälle dem Leser als Übungsaufgabe.

Da die Abbildung $\varphi' : V \times W \to W \otimes V : (x, y) \mapsto y \otimes x$ bilinear ist, gibt es genau eine lineare Abbildung

$$\psi : V \otimes W \to W \otimes V : x \otimes y \mapsto y \otimes x.$$

Die Vertauschung der Rollen von V und W liefert eine lineare Abbildung

$$\psi' : W \otimes V \to V \otimes W : y \otimes x \mapsto x \otimes y.$$

Wir wollen nun zeigen, daß $\psi' \circ \psi = \mathrm{id}_{V \otimes W}$. Aber, wie wir im Beweis der Eindeutigkeit des Tensorproduktes bereits explizit vorgeführt haben, reicht es dazu, daß $\psi' \circ \psi$ auf den reinen Tensoren die Identität ist, und das ist der Fall. Analog gilt $\psi \circ \psi' = \mathrm{id}_{W \otimes V}$, so daß ψ ein Isomorphismus mit Inverser ψ' ist. □

Bemerkung 2.21
Wir bezeichnen mit L die Menge der K-Vektorräume.

a. Durch $V \sim W$, falls $V \cong W$ für $V, W \in L$, wird auf L eine Äquivalenzrelation definiert.
Wir setzen nun $\mathcal{L} = L/\sim$ und definieren für $[V], [W] \in \mathcal{L}$

$$[V] \oplus [W] := [V \oplus W] \quad \text{und } [V] \otimes [W] := [V \otimes W].$$

Dann folgt aus Lemma 2.20, daß $(\mathcal{L}, \oplus)$ und $(\mathcal{L}, \otimes)$ kommutative Halbgruppen sind, mit neutralen Elementen $\big[\{0\}\big]$ respektive $[K]$.

b. Analog wird durch

$$(V,V') \sim (W,W') \Leftrightarrow V \oplus W' \cong W \oplus V', \text{ für } (V,V'),(W,W') \in L \times L,$$

eine Äquivalenzrelation auf $L \times L$ definiert, und wir können die Menge $\mathcal{R} = L \times L/\sim$ betrachten.
Definieren wir nun für $\big[(V,V')\big], \big[(W,W')\big] \in \mathcal{R}$

$$\big[(V,V')\big] \oplus \big[(W,W')\big] := \big[(V \oplus W, V' \oplus W')\big]$$

und

$$\big[(V,V')\big] \otimes \big[(W,W')\big] := \big[(V \otimes W) \oplus (V' \otimes W'), (V \otimes W') \oplus (V' \otimes W)\big],$$

dann folgt aus Lemma 2.20 – im Wesentlichen –, daß $(\mathcal{R}, \oplus, \otimes)$ ein kommutativer Ring mit Eins $1_{\mathcal{R}} = \big[(K,0)\big]$ ist.
Die Abbildung

$$\mathcal{L} \to \mathcal{R} : [V] \mapsto \big[(V,0)\big]$$

ist dabei ein Homomorphismus von Halbgruppen bezüglich der beiden Operationen $\oplus$ und $\otimes$.
Man beachte, daß die Konstruktion von $\mathcal{R}$ der Konstruktion des Ringes $(\mathbb{Z}, +, \cdot)$ aus der Halbgruppe $(\mathbb{N}, +)$ nachempfunden ist.

Zum Tensorprodukt von Vektorräumen gehört auch das Tensorprodukt von linearen Abbildungen.

Proposition 2.22
Es seien V, V', V'', W, W', W'' K*-Vektorräume und* $f \in \mathrm{Hom}_K(V,V')$, $f' \in \mathrm{Hom}_K(V',V'')$, $g \in \mathrm{Hom}_K(W,W')$ *und* $g' \in \mathrm{Hom}_K(W',W'')$.

a. *Es gibt genau eine* K*-lineare Abbildung* $f \otimes g : V \otimes W \to V' \otimes W'$ *mit*

$$(f \otimes g)(x \otimes y) = f(x) \otimes g(y)$$

für alle $x \in V$ *und* $y \in W$.
b. *Es gilt* $(f' \circ f) \otimes (g' \circ g) = (f' \otimes g') \circ (f \otimes g) \in \mathrm{Hom}_K(V \otimes W, V'' \otimes W'')$.

Beweis: a. Da die Abbildung $\varphi' : V \times W \to V' \otimes W' : (x,y) \mapsto f(x) \otimes g(y)$ bilinear ist, induziert sie eine eindeutig bestimmte lineare Abbildung

$$f \otimes g : V \otimes W \to V' \otimes W' : x \otimes y \mapsto f(x) \otimes g(y).$$

b. Die beiden linearen Abbildungen $(f' \circ f) \otimes (g' \circ g)$ und $(f' \otimes g') \circ (f \otimes g)$ stimmen auf den reinen Tensoren überein. Da aber jedes Element in $V \otimes W$ eine Summe reiner Tensoren ist, sind die beiden Abbildungen gleich.

□

Bemerkung 2.23
Wie schon erwähnt, existiert das Tensorprodukt von Moduln über einem beliebigen kommutativen Ring mit Eins ebenfalls. Die Konstruktion unterscheidet sich von der, die wir gegeben haben, da nicht jeder Modul eine Basis besitzt. Es gilt aber per definitionem die universelle Eigenschaft und daher gelten alle Eigenschaften, die man nur mit Hilfe der universellen Eigenschaft beweist.

Insbesondere gelten Lemma 2.20 und Proposition 2.22. Zudem folgt mit einem kurzen Blick in die Konstruktion im allgemeinen Fall, daß auch Lemma 2.17 gilt - wobei a. und b. dort auch unabhängig davon klar sind.

Der folgende wichtige Zusammenhang zwischen dem Tensorprodukt und dem Dualraum gilt aber für Moduln, die keine Basis besitzen, nicht mehr.

Proposition 2.24
Sind V und W endlich-dimensionale K-Vektorräume, so existieren eindeutig bestimmte Isomorphismen

$$\alpha : V^* \otimes W^* \to (V \otimes W)^*$$

mit der Eigenschaft, daß für $f \in V^*$, $g \in W^*$, $x \in V$ und $y \in W$ gilt

$$\alpha(f \otimes g)(x \otimes y) = f(x) \cdot g(y) = \langle f, x\rangle \cdot \langle g, y\rangle,$$

sowie

$$\beta : V^* \otimes W \to \mathrm{Hom}_K(V, W)$$

mit der Eigenschaft, daß für $f \in V^*$, $x \in V$ und $y \in W$ gilt

$$\beta(f \otimes y)(x) = f(x) \cdot y = \langle f, x\rangle \cdot y.$$

Beweis: Sind $f \in V^* = \mathrm{Hom}_K(V, K)$ und $g \in W^* = \mathrm{Hom}_K(W, K)$ gegeben, dann ist die Abbildung

$$\varphi : V \times W \to K : (x, y) \mapsto f(x) \cdot g(y)$$

bilinear, und mithin gibt es eine eindeutig bestimmte lineare Abbildung

$$V \otimes W \to K : x \otimes y \mapsto f(x) \cdot g(y),$$

die wir dann als $\alpha(f \otimes g) \in \mathrm{Hom}_K(V \otimes W, K) = (V \otimes W)^*$ definieren.

Durch Anwenden der Eindeutigkeit sieht man unmittelbar, daß α linear ist. Es bleibt also zu zeigen, daß α bijektiv ist.

Seien dazu $B = (x_1, \ldots, x_n)$ sowie $D = (y_1, \ldots, y_m)$ Basen von V bzw. W. Dann gibt es duale Basen $B^* = (x_1^*, \ldots, x_n^*)$ und $D^* = (y_1^*, \ldots, y_m^*)$ von V^* bzw. W^*, und damit ist $\mathcal{B} = (x_i^* \otimes y_j^* \mid i = 1, \ldots, n, j = 1, \ldots, m)$ eine Basis von $V^* \otimes W^*$ und es reicht zu sehen, daß deren Bild eine Basis von $(V \otimes W)^*$ ist.

Da aber $\mathcal{D} = (x_i \otimes y_j \mid i = 1, \ldots, n, j = 1, \ldots, m)$ eine Basis von $V \otimes W$ ist, ist die duale Basis $\mathcal{D}^* = ((x_i \otimes x_j)^* \mid i = 1, \ldots, n, j = 1 \ldots, m)$ dadurch festgelegt,

daß für $x_k \otimes y_l$ mit der dualen Paarung gilt

$$\left\langle (x_i \otimes x_j)^*, x_k \otimes y_l \right\rangle = \delta_{ik} \cdot \delta_{jl}.$$

Nun gilt aber nach Definition für $\alpha\left(x_i^* \otimes x_j^*\right)$ gerade

$$\left\langle \alpha\left(x_i^* \otimes x_j^*\right), x_k \otimes y_l \right\rangle = \alpha\left(x_i^* \otimes x_j^*\right)(x_k \otimes y_l) = x_i^*(x_k) \cdot x_j^*(x_l) = \delta_{ik} \cdot \delta_{jl}.$$

Also bildet α die Basis $\mathcal{B}$ auf die Basis $\mathcal{D}^*$ ab und α ist ein Isomorphismus.

Wie im Falle von α sieht man mittels der universellen Eigenschaft des Tensorproduktes $V^* \otimes W$, daß die Abbildung β definiert und linear ist. Es bleibt wiederum zu zeigen, daß β bijektiv ist.

Wir behalten die Bezeichnungen von oben bei. Dann bilden die Abbildungen

$$\epsilon_{ij} : V \to W : x_k \mapsto \delta_{ik} y_j, \quad k = 1, \ldots, n,$$

eine Basis $\mathcal{E} = (\epsilon_{ij} \mid i = 1, \ldots, n, j = 1, \ldots, m)$ von $\mathrm{Hom}_K(V, W)$.[6] Ferner wissen wir, daß $C = \left(x_i^* \otimes y_j \;\middle|\; i = 1, \ldots, n, j = 1, \ldots, m\right)$ eine Basis von $V^* \otimes W$ ist. Nun gilt aber für $i \in \{1, \ldots, n\}$ und $j \in \{1, \ldots, m\}$

$$\beta\left(x_i^* \otimes y_j\right)(x_k) = \left\langle x_i^*, x_k \right\rangle \cdot y_j = \delta_{ik} \cdot y_j$$

für alle $k = 1, \ldots, n$. Mithin ist $\beta\left(x_i^* \otimes y_j\right) = \epsilon_{ij}$ und damit ist das Bild der Basis C unter β die Basis $\mathcal{E}$, β also ein Isomorphismus. □

Abschließend wollen wir noch einige Beispiele behandeln.

Beispiel 2.25 a. Es seien $A \in \mathrm{Mat}\left(n' \times n, K\right)$ und $B \in \mathrm{Mat}\left(m' \times m, K\right)$ gegeben, und $f = f_A : K^n \to K^{n'}$ sowie $g = f_B : K^m \to K^{m'}$ die assoziierten linearen Abbildungen. Dadurch erhalten wir eine eindeutig bestimmte lineare Abbildung

$$f \otimes g : K^n \otimes K^m \to K^{n'} \otimes K^{m'} : x \otimes y \mapsto Ax \otimes By.$$

Wir identifizieren nun den $n \cdot m$-dimensionalen K-Vektorraum $K^n \otimes K^m$ mit dem Vektorraum K^{nm} durch

$$K^n \otimes K^m \xrightarrow{\cong} K^{nm} : e_i \otimes e_j \mapsto e_{(i-1)m+j}$$

für $i = 1, \ldots, n$ und $j = 1, \ldots, m$.
Die Identifikation entspricht der Wahl einer Numerierung auf der Basis von $\mathcal{B} = (e_i \otimes e_j \mid i = 1, \ldots, n, j = 1, \ldots, m)$ sowie der anschließenden Anwendung der Karte $\phi_{\mathcal{B}}$.
Analog identifizieren wir $K^{n'} \otimes K^{m'}$ mit $K^{n'm'}$ mittels der Wahl einer Numerierung auf $\mathcal{B}' = \left(e_i \otimes e_j \;\middle|\; i = 1, \ldots, n', j = 1, \ldots, m'\right)$.
Wir wollen nun die Matrixdarstellung

$$M_{\mathcal{B}'}^{\mathcal{B}}(f \otimes g)$$

untersuchen.

[6] Wir wissen, daß die Karten ϕ_B und ϕ_D einen Isomorphismus zwischen $\mathrm{Hom}_K(V, W)$ und $\mathrm{Mat}(n \times m, K)$ induzieren und es gilt $\phi_D \circ \epsilon_{ij} \circ \phi_B^{-1} \equiv E_{ij}$, wobei die E_{ij} die kanonische Basis von $\mathrm{Mat}(n \times m, K)$ bilden. Mithin bilden die ϵ_{ij} eine Basis von $\mathrm{Hom}_K(V, W)$.

Dazu definieren wir für die Matrizen A und B das *Tensorprodukt* von A und B durch

$$A \otimes B = \begin{pmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & & \vdots \\ a_{n'1}B & \dots & a_{n'n}B \end{pmatrix} \in \mathbf{Mat}\left(n'm' \times nm, K\right),$$

und wir behaupten es gilt

$$M^{\mathcal{B}}_{\mathcal{B}'}(f \otimes g) = A \otimes B.$$

Dazu betrachten wir $f \otimes g$ angewandt auf $e_i \otimes e_j \in \mathcal{B}$. Es gilt

$$\begin{aligned} (f \otimes g)(e_i \otimes e_j) \quad &= \quad Ae_i \otimes Be_j = \textstyle\sum_{k=1}^{n'} a_{ki}e_k \otimes \sum_{l=1}^{m'} b_{lj}e_l \\ &= \quad \sum_{\substack{k=1,\dots,n' \\ l=1,\dots,m'}} a_{ki}b_{lj}e_k \otimes e_l \\ &\xmapsto{\phi_{\mathcal{B}'}} \quad \sum_{\substack{k=1,\dots,n' \\ l=1,\dots,m'}} a_{ki}b_{lj}e_{(k-1)m'+l}. \end{aligned}$$

Beachtet man nun noch, daß $\phi_{\mathcal{B}}(e_i \otimes e_j) = e_{(i-1)m+j}$ ist, so folgt die Behauptung, da dann der Eintrag in $M^{\mathcal{B}}_{\mathcal{B}'}(f \otimes g)$ zu $e_i \otimes e_j$ gerade der Spaltenvektor

$$(a_{1i}b_{1j}, a_{1i}b_{2j}, \dots, a_{1i}b_{m'j}, a_{2i}b_{1j}, \dots, a_{2i}b_{m'j}, \dots, a_{n'i}b_{1j}, \dots, a_{n'i}b_{m'j})^t$$

ist.

Man beachte, daß die Einträge von $A \otimes B$ gerade alle Produkte $a_{ki}b_{lj}$ mit $k = 1,\dots,n', i = 1,\dots,n, l = 1,\dots,m', j = 1,\dots,m$ in der *richtigen* Reihenfolge sind.

Als konkretes Beispiel betrachten wir $A \otimes B$ für

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \in \mathbf{Mat}(2, K) \quad \text{und} \quad B = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \end{pmatrix} \in \mathbf{Mat}(2 \times 3, K).$$

Damit erhalten wir

$$A \otimes B = \begin{pmatrix} 1 \cdot B & 2 \cdot B \\ 0 \cdot B & 1 \cdot B \end{pmatrix} = \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 2 & 4 & 6 \\ 0 & 1 & 2 & 0 & 2 & 4 \\ \hline 0 & 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{array}\right) \in \mathbf{Mat}(4 \times 6, K).$$

b. Wir wollen das Tensorprodukt zweier Matrizen nun benutzen, um konkret ein Erzeugendensystem des Tensorproduktes zweier Unterräume von $K^{n'}$ bzw. $K^{m'}$ anzugeben, wenn diese selbst durch Erzeugendensysteme gegeben sind.

Dazu beachten wir zunächst, daß für lineare Abbildungen $f : V \to V'$ und $g : W \to W'$ gilt

$$\mathrm{Im}(f \otimes g) = \mathrm{Im}(f) \otimes \mathrm{Im}(g).$$

Denn, sind $(x_i \mid i \in I)$ und $(y_j \mid j \in J)$ zwei Basen von V bzw. W, dann sind per definitionem $\big(f(x_i) \mid i \in I\big)$ bzw. $\big(g(y_j) \mid j \in J\big)$ Erzeugendensysteme

von $\mathrm{Im}(f)$ und $\mathrm{Im}(g)$, und ferner ist $B = \big(f(x_i) \otimes g(y_j) \mid i \in I, j \in J\big)$ ein Erzeugendensystem von $\mathrm{Im}(f \otimes g)$. Zugleich ist B aber auch ein Erzeugendensystem von $\mathrm{Im}(f) \otimes \mathrm{Im}(g)$. Somit folgt die Behauptung.
Seien nun $V \subseteq K^{n'}$ und $W \subseteq K^{m'}$ zwei Unterräume, die durch die Erzeugendensysteme $(x_1, \ldots, x_n) \subset K^{n'}$ bzw. $(y_1, \ldots, y_m) \subset K^{m'}$ gegeben sind, und seien $A \in \mathrm{Mat}\big(n' \times n, K\big)$ und $B \in \mathrm{Mat}\big(m' \times m, K\big)$ die Matrizen, deren Spalten gerade die Vektoren der beiden Erzeugendensysteme sind, dann ist

$$\mathrm{Im}(f_A) = \mathrm{SR}(A) = V \quad \text{und} \quad \mathrm{Im}(f_B) = \mathrm{SR}(B) = W.$$

Folglich ist

$$V \otimes W = \mathrm{Im}(f_A) \otimes \mathrm{Im}(f_B) = \mathrm{Im}(f_A \otimes f_B) = \mathrm{SR}(A \otimes B),$$

d. h. die Spalten der Matrix $A \otimes B$ bilden ein Erzeugendensystem von $V \otimes W$ in $K^{n'm'}$.
In dem konkreten Beispiel in a. erzeugen sowohl die Spalten von A als auch die von B ganz K^2, und somit müssen die Spalten von $A \otimes B$ ein Erzeugendensystem von K^4 liefern, wie man unmittelbar sieht, da $A \otimes B$ bereits in ZSF vorliegt.

c. In Singular kann man das Tensorprodukt von zwei Matrizen wie folgt realisieren.

```
ring r=0,x,lp;
matrix A[2][2]=1,2,0,1;
matrix B[2][3]=1,2,3,0,1,2;
matrix C[nrows(A)][1];
matrix D[1][nrows(B)];
matrix AB;
int i,j;
for (i=1;i<=ncols(A);i++)
{
   for (j=1;j<=ncols(B);j++)
   {
     C=A[1..nrows(A),i];
     D=B[1..nrows(B),j];
     AB=concat(AB,transpose(ideal(C*D)));
   }
}
AB=submat(AB,1..nrows(AB),2..ncols(AB));
print(AB);
↦    1,  2,  3,  2,  4,  6
     0,  1,  2,  0,  2,  4
     0,  0,  0,  1,  2,  3
     0,  0,  0,  0,  1,  2
```

AUFGABEN

Aufgabe 2.26

Es sei V ein K-Vektorraum und $x, y \in V$. Zeige, genau dann gilt $x \otimes y = y \otimes x$, wenn x und y linear abhängig sind.

Aufgabe 2.27

Es seien $V_1, \ldots, V_n$ K-Vektorräume und $x_i, y_i \in V_i$, $i = 1, \ldots, n$, mit $0 \neq x_1 \otimes \ldots \otimes x_n = y_1 \otimes \ldots \otimes y_n$. Dann gibt es Elemente $\lambda_1, \ldots, \lambda_n \in K$ mit $\lambda_i x_i = y_i$, $i = 1, \ldots, n$, und $\lambda_1 \cdots \lambda_n = 1$.

Aufgabe 2.28

Es seien V und W zwei K-Vektorräume, $(x_1, \ldots, x_n) \subset V$ linear unabhängig und $(y_1, \ldots, y_n) \subset W$, $(z_1, \ldots, z_n) \subset W$ mit $\sum_{i=1}^{n} x_i \otimes y_i = \sum_{i=1}^{n} x_i \otimes z_i$. Dann gilt $y_i = z_i$ für alle $i = 1, \ldots, n$.

Aufgabe 2.29

Es seien V und W zwei K-Vektorräume, $F = (x_i \mid i \in I) \subset V$ und $G = (y_j \mid j \in J) \subset W$ zwei Familien von Vektoren und $H = (x_i \otimes y_j \mid i \in I, j \in J)$. Zeige:

a. H ist linear unabhängig in $V \otimes W$ genau dann, wenn F und G linear unabhängig in V bzw. W sind.
b. H ist ein Erzeugendensystem von $V \otimes W$ genau dann, wenn F und G Erzeugendensysteme von V bzw. W sind.
c. H ist eine Basis von $V \otimes W$ genau dann, wenn F und G Basen von V bzw. W sind.

Aufgabe 2.30

Es seien V und W zwei K-Vektorräume und $0 \neq z \in V \otimes W$. Ist $r \in \mathbb{N}$ minimal, so daß es Vektoren $x_i \in V$ und $y_i \in W$, $i = 1, \ldots, r$, gibt mit $z = \sum_{i=1}^{r} x_i \otimes y_i$, so nennt man $\operatorname{rang}(z) = r$ den *Rang* des Tensors z.

Zeige, ist $0 \neq z = \sum_{i=1}^{r} x_i \otimes y_i \in V \otimes W$ beliebig, dann sind die folgenden Aussagen gleichwertig:

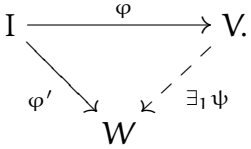
a. $r = \operatorname{rang}(z)$.
b. $(x_1, \ldots, x_r)$ und $(y_1, \ldots, y_r)$ sind linear unabhängig in V bzw. W.

Aufgabe 2.31

Beweise die Aussagen in Lemma 2.20 mit Hilfe der universellen Eigenschaft des Tensorproduktes.

Aufgabe 2.32

Es sei R ein kommutativer Ring mit Eins, V ein R-Modul und $\varphi : I \to V$ eine Abbildung. Wir nennen (V, φ) (oder auch nur V) *frei* vom *Rang* $|I|$, wenn für jedes andere Tupel (W, φ'), mit W ein R-Modul und $\varphi' : I \to W$ eine Abbildung, genau eine lineare Abbildung $\psi : V \to W$ existiert, so daß $\psi \circ \varphi = \varphi'$, d. h. , so daß das folgende Diagramm kommutiert:



Zeige:

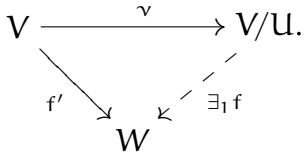
a. (V, φ) ist genau dann frei, wenn die Familie $(x_i \mid i \in I)$, mit $x_i = \varphi(i)$ für $i \in I$, eine Basis von V ist.
b. Jeder Vektorraum über einem Körper ist frei.
c. Gib ein Beispiel für einen Ring und einen Modul, der nicht frei ist.

Die obige universelle Eigenschaft beschreibt also genau das Faktum, daß eine lineare Abbildung auf einer Basis eindeutig vorgeschrieben werden kann!

Aufgabe 2.33

Es sei V ein R-Modul und U ein Untermodul. Man zeige, daß der Quotientenmodul V/U zusammen mit der Restklassenabbildung $\nu : V \to V/U$ der folgenden universellen Eigenschaft genügt:

> Es sei W ein beliebiger R-Modul und $f' \in \mathrm{Hom}_R(V, W)$ mit $f'(U) = \{0\}$, dann gibt es genau eine lineare Abbildung $f \in \mathrm{Hom}_R(V/U, W)$ mit $f' = f \circ \nu$, d. h. so, daß das folgende Diagramm kommutiert:



Aufgabe 2.34

Seien U, V, W drei K-Vektorräume. Dann gibt es genau einen Isomorphismus

$$\mathrm{Hom}_K(U, V) \otimes W \to \mathrm{Hom}_K(U, V \otimes W) : f \otimes w \mapsto \big(U \to V \otimes W : u \mapsto f(u) \otimes w\big).$$

3. Äußeres Produkt und äußere Algebra

Generalvoraussetzung: In diesem Kapitel bezeichne K stets einen beliebigen Körper und R einen kommutativen Ring mit Eins.

Seien V und W zwei R-Moduln. Wir erinnern uns, daß wir in Definition II.9.7 eine multilineare Abbildung

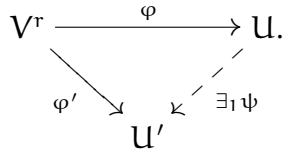
$$f : V^r = V \times \ldots \times V \to W$$

alternierend genannt haben, wenn aus $(x_1, \ldots, x_r) \in V^r$ mit $x_i = x_j$ für ein $i \neq j$ folgt, daß $f(x_1, \ldots, x_r) = 0$ ist. Man beachte, daß der Begriff *alternierend* für eine multilineare Abbildung f nur Sinn macht, wenn alle Argumente von f aus dem selben Vektorraum V stammen.

Ganz analog , wie durch das Tensorprodukt multilineare Abbildungen $V_1 \times \ldots \times V_r \to W$ eindeutig in lineare Abbildungen $V_1 \otimes \ldots \otimes V_r \to W$ transformiert werden, soll jetzt ein *äußeres Produkt* $V \wedge \ldots \wedge V$ definiert werden, so daß alternierende multilineare Abbildungen $V \times \ldots \times V \to W$ eindeutig in lineare Abbildungen $V \wedge \ldots \wedge V \to W$ überführt werden.

Definition 3.1
Es sei V ein R-Modul und $r \geq 1$. Ein Paar (U, φ) mit U ein R-Modul und $\varphi : V^r \to U$ eine alternierende multilineare Abbildung heißt r-*faches äußeres Produkt* von V, wenn (U, φ) der folgenden *universellen Eigenschaft* genügt:

> Für jedes weitere Paar (U', φ') mit U' ein R-Modul und $\varphi' : V^r \to U'$ eine alternierende multilineare Abbildung gilt, es *existiert genau eine lineare* Abbildung $\psi : U \to U'$ mit $\psi \circ \varphi = \varphi'$, d. h. so, daß das folgende Diagramm kommutiert:



Da nach dem folgenden Satz das r-fache äußere Produkt, so es existiert, bis auf eindeutige Isomorphie eindeutig bestimmt ist, schreibt man $\bigwedge^r V := V \wedge \ldots \overset{r}{\ldots} \wedge V$ statt U und unterschlägt - um Notation zu sparen - für gewöhnlich die alternierende multilineare Abbildung φ ganz.

Ferner führen wir für $x_1, \ldots, x_r \in V$ die folgende Notation ein

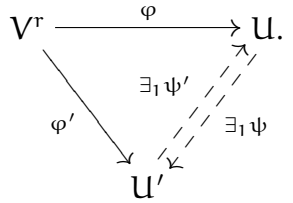
$$x_1 \wedge \ldots \wedge x_r := \varphi(x_1, \ldots, x_r),$$

und nennen Elemente dieser Form *reine Produkte* oder *zerlegbar*.

Wir werden zunächst wieder die Eindeutigkeit des äußeren Produktes zeigen, und seine Existenz dann aus der Existenz des Tensorproduktes herleiten. Da wir die Existenz von Tensorprodukten nur für Vektorräume gezeigt haben, werden wir uns dabei auch wieder auf den Fall von Vektorräumen beschränken, obwohl die Konstruktion im Falle von Moduln exakt gleich bleibt.

Satz 3.2
Es seien V ein R-Modul, $r \geq 1$ und (U, φ) und (U', φ') seien zwei r-fache äußere Produkte von V. Dann gibt es genau einen Isomorphismus $\psi : U \to U'$ mit $\psi \circ \varphi = \varphi'$.

Beweis: Aus der universellen Eigenschaft, der sowohl (U, φ) als auch (U', φ') genügen, folgt, daß es zwei eindeutig bestimmte Abbildungen $\psi : U \to U'$ und $\psi' : U' \to U$ gibt, die das folgende Diagramm zum kommutieren bringen:



Damit gilt aber auch

$$\psi' \circ \psi \circ \varphi = \psi' \circ \varphi' = \varphi \tag{72}$$

und

$$\psi \circ \psi' \circ \varphi' = \psi \circ \varphi = \varphi' \tag{73}$$

Und aus der Eindeutigkeit folgt unmittelbar, daß nur ψ als Kandidat für den Isomorphismus in Frage kommt!

Betrachten wir nun die alternierende multilineare Abbildung $\varphi : V^r \to U$ selbst, so sagt die universelle Eigenschaft von (U, φ), daß es genau eine lineare Abbildung $\pi : U \to U$ gibt mit $\pi \circ \varphi = \varphi$. Offensichtlich ist id_U eine lineare Abbildung, die diese Eigenschaft besitzt, und mithin gilt $\pi = \mathrm{id}_U$. Andererseits gilt nach (72) aber, daß $\psi' \circ \psi$ ebenfalls diese Eigenschaft hat, also

$$\psi' \circ \psi = \pi = \mathrm{id}_U .$$

Analog folgt aus (73) und der universellen Eigenschaft von $\left(U', \varphi'\right)$, daß auch

$$\psi \circ \psi' = \pi = \mathrm{id}_{U'} .$$

Mithin ist ψ ein Isomorphismus mit ψ' als Inverser. □

Vergleicht man den Beweis mit dem Beweis der Eindeutigkeit des Tensorproduktes, Satz 2.5, so stellt man fest, daß beide Beweise fast wörtlich identisch sind.

Kommen wir nun also zum Beweis der Existenz des äußeren Produktes im Falle von K-Vektorräumen. Dazu führen wir zunächst noch eine neue Notation ein.

Definition 3.3
Es sei V ein K-Vektorraum und $r \in \mathbb{N}$. Dann definieren wir $T^0(V) = K$ und rekursiv $T^r(V) = T^{r-1}(V) \otimes V$.

Damit gilt für $r \geq 1$ dann wegen der Assoziativität des Tensorproduktes

$$T^r(V) = V \otimes \overset{r}{\ldots} \otimes V.$$

Ferner setzen wir $V_0 = \{0\} \subset K$ und definieren für jedes $r \geq 1$ einen Unterraum

$$V_r := \langle x_1 \otimes \ldots \otimes x_r \mid x_i \in V \; \forall i = 1, \ldots, r; \exists\, i \neq j : x_i = x_j \rangle_K \subset T^r(V).$$

Satz 3.4
Es sei V ein K-Vektorraum und $r \geq 1$. Setzen wir $\bigwedge^r V := T^r(V)/V_r$ und

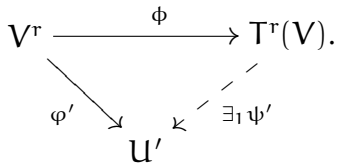
$$\varphi : V^r \to \bigwedge^r V : (x_1, \ldots, x_r) \mapsto x_1 \wedge \ldots \wedge x_r := x_1 \otimes \ldots \otimes x_r (\mathrm{mod}\ V_r),$$

dann ist das Paar $\left(\bigwedge^r V, \varphi\right)$ ein r-faches äußeres Produkt von V.

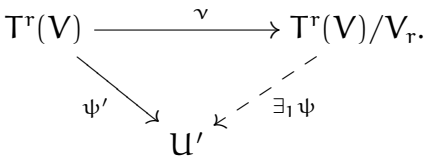
Beweis: Wir beachten zunächst, daß die Abbildung $\varphi = \nu \circ \phi$ die Komposition der zum Tensorprodukt gehörenden multilinearen Abbildung $\phi : V^r \to T^r(V)$ mit der linearen Restklassenabbildung $\nu : T^r(V) \to T^r(V)/V_r$ ist. Damit ist φ insbesondere multilinear. Sind nun $x_1, \ldots, x_r \in V$ mit $x_i = x_j$ für ein $i \neq j$, dann gilt zudem, daß $x_1 \otimes \ldots \otimes x_r \in V_r$, und mithin, $\varphi(x_1, \ldots, x_r) = 0$. D. h. φ ist auch alternierend. Es bleibt also, die universelle Eigenschaft zu überprüfen.

Sei dazu $\varphi' : V^r \to U'$ eine alternierende multilineare Abbildung. Die universelle Eigenschaft des Tensorproduktes liefert, daß es genau eine lineare

Abbildung $\psi' : T^r(V) \to U'$ gibt mit $\psi' \circ \phi = \varphi'$, d. h. so, daß folgendes Diagramm kommutiert



Da φ' alternierend ist, ist $\psi'(V_r) = 0$, und das bedeutet nach Aufgabe 2.33, daß ψ' in eindeutiger Weise durch $T^r(V)/V_r$ faktorisiert, d. h. es gibt genau eine Abbildung $\psi : T^r(V)/V_r \to U'$, so daß $\psi' = \psi \circ \nu$, d. h. so, daß das folgende Diagramm kommutiert:



Aber dann ist $\psi : T^r(V)/V_r = \bigwedge^r V \to U'$ eine Abbildung mit $\psi \circ \varphi = \psi \circ \nu \circ \phi = \psi' \circ \phi = \varphi'$.

Es bleibt noch zu zeigen, daß ψ eindeutig ist mit dieser Eigenschaft. Sei also $\psi'' : T^r(V)/V_r = \bigwedge^r V \to U'$ eine zweite Abbildung mit $\psi'' \circ \varphi = \varphi'$. Setzen wir $\psi''' := \psi'' \circ \nu : T^r(V) \to U'$, dann gilt

$$\psi''' \circ \phi = \psi'' \circ \nu \circ \phi = \psi'' \circ \varphi = \varphi',$$

also folgt mit der oben angegebenen Eindeutigkeit von ψ' beim Tensorprodukt, daß $\psi''' = \psi'$. Aber dann gilt $\psi' = \psi'' \circ \nu$ und aus der ebenfalls oben angeführten Eindeutigkeit von ψ beim Quotientenraum folgt, daß $\psi'' = \psi$. □

Lemma 3.5
Es sei V ein K-Vektorraum, dann gelten für $x, x', y, y', x_1, \ldots, x_r \in V$ und $\lambda \in K$ die folgenden Rechenregeln:

a. $x \wedge (y + y') = x \wedge y + x \wedge y'$ *und* $(x + x') \wedge y = x \wedge y + x' \wedge y$.
b. $\lambda(x \wedge y) = (\lambda x) \wedge y = x \wedge (\lambda y)$.
   Insbesondere gilt $0 \wedge y = x \wedge 0 = 0$.
c. $x \wedge y = -y \wedge x$.
d. $x_1 \wedge \ldots \wedge x_i \wedge \ldots \wedge x_j \wedge \ldots \wedge x_r = -x_1 \wedge \ldots \wedge x_j \wedge \ldots \wedge x_i \wedge \ldots \wedge x_r$.

Beweis: a. und b. folgen wieder daraus, daß φ multilinear ist.

Seien $x, y \in V$, dann gilt mit a., da φ alternierend ist

$$0 = (x + y) \wedge (x + y) = x \wedge x + x \wedge y + y \wedge x + y \wedge y = x \wedge y + y \wedge x.$$

Aber daraus folgt c., und d. folgt analog. □

Lemma 3.6
Es seien U, V, W K-Vektorräume, $f \in \mathrm{Hom}_K(V, W)$, $g \in \mathrm{Hom}_K(W, U)$ und $r \geq 1$.

a. *Es gibt genau eine lineare Abbildung*

$$\bigwedge^r f : \bigwedge^r V \to \bigwedge^r W : x_1 \wedge \ldots \wedge x_r \mapsto f(x_1) \wedge \ldots \wedge f(x_r).$$

b. *Es gilt*

$$\bigwedge^r (g \circ f) = \bigwedge^r g \circ \bigwedge^r f : \bigwedge^r V \to \bigwedge^r U.$$

Beweis: a. Nach der Verallgemeinerung von Proposition 2.22 gibt es genau eine Abbildung $T^r(f) : T^r(V) \to T^r(W)$ mit $T^r(f)(x_1 \otimes \ldots \otimes x_r) \mapsto f(x_1) \otimes \ldots \otimes f(x_r)$. Dabei gilt offenbar $T^r(f)(V_r) \subseteq W_r$. Aber dann sagt die universelle Eigenschaft des Quotientenraumes aus Aufgabe 2.33, daß $T^r(f)$ in eindeutiger Weise eine Abbildung

$$\bigwedge^r f : \bigwedge^r V \to \bigwedge^r W : x_1 \wedge \ldots \wedge x_r \mapsto f(x_1) \wedge \ldots \wedge f(x_r)$$

induziert. Wir überlassen es dem Leser, sich mit dem detaillierten Nachweis der Eindeutigkeit herumzuplagen.

b. Ebenfalls nach Proposition 2.22 gilt $T^r(g \circ f) = T^r(g) \circ T^r(f)$, und da die Quotientenbildung ebenfalls mit der Komposition verträglich ist, folgt somit die Behauptung.

□

Proposition 3.7
Sei $(x_1, \ldots, x_n)$ *eine Basis des* K*-Vektorraumes* V*. Dann gilt:*

a. $\bigwedge^r V = \{0\}$ *für alle* $r > n$.
b. $(x_{i_1} \wedge \ldots \wedge x_{i_r} \mid 1 \leq i_1 < \ldots < i_r \leq n)$ *ist eine Basis von* $\bigwedge^r V$ *für* $1 \leq r \leq n$. *Insbesondere gilt in diesem Fall:* $\dim_K \left(\bigwedge^r V \right) = \binom{n}{r}$.

Beweis: Wir beweisen zunächst die Aussage in b.

Da $(x_{i_1} \otimes \ldots \otimes x_{i_r} \mid 1 \leq i_j \leq n, j = 1, \ldots, r)$ eine Basis von $T^r(V)$ ist, ist

$$B' := (x_{i_1} \wedge \ldots \wedge x_{i_r} \mid 1 \leq i_j \leq n, j = 1, \ldots, r)$$

ein Erzeugendensystem von $\bigwedge^r V$. Bedenkt man nun noch, daß die Produkte, bei denen ein Faktor mehrfach vorkommt, Null sind, und daß die Vertauschung der Reihenfolge von Faktoren nur das Vorzeichen ändert, dann haben wir in der Tat bereits

$$B := (x_{i_1} \wedge \ldots \wedge x_{i_r} \mid 1 \leq i_1 < \ldots < i_r \leq n)$$

ist ein Erzeugendensystem von $\bigwedge^r V$.

Insbesondere gilt, $\dim_K \left(\bigwedge^r V \right) \leq \binom{n}{r} =: N$.

Es bleibt zu zeigen, daß B linear unabhängig ist.

Sei dazu $D = (e_{i_1 \ldots i_r} \mid 1 \leq i_1 < \ldots < i_r \leq n)$ eine Basis von K^N. Unser nächstes Ziel ist es, eine alternierende multilineare Abbildung $\alpha : V^r \to K^N$ zu konstruieren.

Sei dazu $y = (y_1, \ldots, y_r) \in V^r$ gegeben mit folgenden Darstellungen bezüglich der Basis $(x_1, \ldots, x_n)$ von V:

$$y_i = \sum_{j=1}^{n} a_{ij} x_j, \quad i = 1, \ldots, r.$$

Wir setzen dann $A = (a_{ij}) \in \mathrm{Mat}(r \times n, K)$, d. h., wir schreiben die Koeffizienten der Koordinatendarstellung der y_i als Zeilen in die Matrix A.

Für einen Multiindex $(j_1, \ldots, j_r)$ mit $1 \leq j_1 < \ldots < j_r \leq n$ sei

$$A(j_1, \ldots, j_r) = \det(a_{ij_l} \mid i, l = 1, \ldots, r)$$

der maximale Minor der Spalten $j_1, \ldots, j_r$ von A. Von diesen Minoren gibt es exakt $\binom{n}{r}$ Stück - beachte, $r \leq n$. Wir definieren nun

$$\alpha : V^r \to K^N : (y_1, \ldots, y_r) \mapsto \sum_{1 \leq i_1 < \ldots < i_r \leq n} A(i_1, \ldots, i_r) e_{i_1 \ldots i_r}.$$

Da die Determinante multilinear und alternierend bezüglich der Zeilen einer Matrix ist, ist α eine alternierende multilineare Abbildung mit $\alpha(x_{i_1}, \ldots, x_{i_r}) = e_{i_1 \ldots i_r}$.

Aus der universellen Eigenschaft des äußeren Produktes folgt somit die Existenz einer linearen Abbildung

$$\bigwedge^r V \to K^N : x_{i_1} \wedge \ldots \wedge x_{i_r} \mapsto e_{i_1 \ldots i_r}.$$

Da D eine Basis von K^N ist, ist diese Abbildung mithin surjektiv und somit folgt $\dim_K \left(\bigwedge^r V \right) \geq N = \binom{n}{r}$.

Insgesamt erhalten wir $\dim_K \left(\bigwedge^r V \right) = N = \binom{n}{r}$, und insbesondere, daß B eine Basis ist. Damit ist Teil b. bewiesen.

Für $r > n$ sind in $x_{j_1} \wedge \ldots \wedge x_{j_r}$, $1 \leq j_i \leq n$, mindestens zwei Faktoren gleich, also ist das Produkt Null, und es folgt Teil a., da auch in diesem Fall B' ein Erzeugendensystem von $\bigwedge^r V$ ist. □

Aus dem Beweis leiten wir das folgende Korollar her.

Korollar 3.8

Sei $(e_i \mid i = 1, \ldots, n)$ *die kanonische Basis des* K^n *und sei* $(e_{i_1 \ldots i_r} \mid 1 \leq i_1 < \ldots < i_r \leq n)$ *eine Basis von* $\bigwedge^r K^n$. *Seien ferner Vektoren* $y_1, \ldots, y_r \in V$ *gegeben mit* $y_i = \sum_{j=1}^{n} a_{ij} e_j$, $a_{ij} \in K$, *und seien die* $r \times r$*-Minoren* $A(j_1, \ldots, j_r)$ *der Koeffizientenmatrix* $A = (a_{ij}) \in \mathrm{Mat}(r \times n, K)$ *wie im Beweis von Proposition 3.7 definiert.*

Dann gilt: $y_1 \wedge \ldots \wedge y_r = \sum_{1 \leq i_1 < \ldots < i_r \leq n} A(i_1, \ldots, i_r) e_{i_1 \ldots i_r}$.

Insbesondere gilt für $r = n$, $\bigwedge^n K^n = K e_{1 \ldots n} \cong K$ *und* $y_1, \ldots, y_n \in K^n$

$$y_1 \wedge \ldots \wedge y_n = \det(A) = \det(y_1, \ldots, y_n).$$

Dieses Korollar gibt eine schöne Interpretation des äußern Produktes. Äußere Produkte spielen in der Analysis, insbesondere bei der Integration auf Mannigfaltigkeiten eine wichtige Rolle.

AUFGABEN

Aufgabe 3.9

Es sei V ein K-Vektorraum und $r \geq 1$. Zeige:

$$U_r = \langle x_1 \otimes \ldots \otimes x_r \mid x_i \in V \; \forall i = 1, \ldots, r; \; \exists \, 1 \leq i \leq r-1 : x_i = x_{i+1} \rangle.$$

Aufgabe 3.10

Es sei V ein K-Vektorraum und $(x_1, \ldots, x_r)$ eine Familie von Vektoren in V. Zeige, die folgenden beiden Aussagen sind gleichwertig:

a. $(x_1, \ldots, x_r)$ ist linear unabhängig in V.
b. $x_1 \wedge \ldots \wedge x_r \neq 0$.

Aufgabe 3.11

Es sei V ein K-Vektorraum und $U \subseteq V$ ein Unterraum von V mit Basis $(x_1, \ldots, x_r)$. Zeige, es gilt

$$U = \{x \in V \mid x \wedge x_1 \wedge \ldots \wedge x_r = 0\}.$$

Aufgabe 3.12

Es sei V ein K-Vektorraum und $U, W \subseteq V$ zwei Unterräume von V mit Basen $(x_1, \ldots, x_r)$ bzw. $(y_1, \ldots, y_r)$. Zeige, genau dann ist $U = W$, wenn $\langle x_1 \wedge \ldots \wedge x_r \rangle_K = \langle y_1 \wedge \ldots \wedge y_r \rangle_K$.

Aufgabe 3.13

Es sei V ein K-Vektorraum mit $\dim_K(V) = n < \infty$. Zeige, jedes Element von $\bigwedge^{n-1} V$ ist zerlegbar.

Aufgabe 3.14

Es seien V und W zwei K-Vektorräume, $\lambda \in K$ und $f \in \mathrm{Hom}_K(V, W)$. Zeige:

a. $\bigwedge^r(\mathrm{id}_V) = \mathrm{id}_{\bigwedge^r V}$.
b. $\bigwedge^r(\lambda f) = \lambda^r \bigwedge^r f$.

Aufgabe 3.15

Es seien V und W zwei K-Vektorräume, $r \geq 1$ und $f \in \mathrm{Hom}_K(V, W)$. Zeige:

a. Ist f injektiv, so ist $\bigwedge^r f$ injektiv.[7]
b. Ist f surjektiv, so ist $\bigwedge^r f$ surjektiv.
c. Ist f bijektiv, so ist $\bigwedge^r f$ bijektiv.

Aufgabe 3.16

Ist V ein K-Vektorraum mit $\dim_K(V) = n < \infty$ und $f \in \mathrm{End}_K(V)$. Zeige:

$$\det(f) = \mathrm{Spur}\left(\bigwedge^n f\right).$$

[7] Dies gilt bei Moduln i. a. nicht mehr!

Aufgabe 3.17

Es sei V ein endlich-dimensionaler K-Vektorraum und $r \geq 1$. Zeige, es gibt einen kanonischen Isomorphismus $\bigwedge^r (V^*) \to \left(\bigwedge^r V\right)^*$.

ANHANG A

Klausuren zur Linearen Algebra I & II

FB Mathematik — Wintersemester 1999/2000
Prof. Dr. Gert–Martin Greuel — Thomas Keilen

1. Klausur in Linearer Algebra I

Klausurtermin: Samstag, 18. Dezember 1999, 09:00-12:00 Uhr

Hinweise: Die Klausuraufgaben sind jeweils auf getrennten Blättern zu bearbeiten. *Nie zwei Aufgaben auf dem gleichen Blatt lösen!!!*

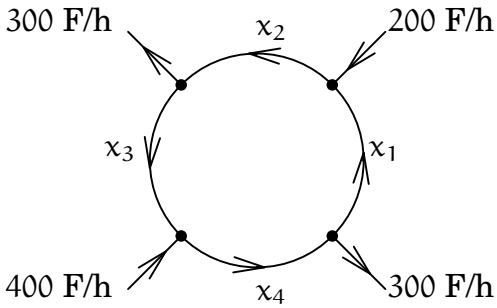
Alle nicht offensichtlichen Beweis-/Rechenschritte sind zu begründen (und ggf. durch ein Zitat zu belegen).

Die Zahlen in Klammern am rechten Seitenrand geben die Punktzahlen an, die durch Lösen der jeweiligen Aufgabe erreichbar sind. Insgesamt sind es 48 Punkte.

Jedes Blatt ist am oberen Rand der Vorderseite wie folgt zu beschriften:

eigener Name *Aufgabennummer* *Name des Übungsgruppenleiters*

Aufgabe 1: In einem Straßen-Netzwerk gibt es Straßen und Kreuzungen. Wir nehmen an, daß der totale Verkehrsfluß (gemessen in Fahrzeuge pro Stunde F/h) in eine Kreuzung hinein gleich dem totalen Verkehrsfluß aus der Kreuzung hinaus ist. In dem unten gegebenen Kreisverkehr stehen die Zahlen und Variablen für den gemessenen bzw. zu bestimmenden Verkehrsfluß.



(3)

Stelle ein lineares Gleichungssystem auf, welches das Netzwerk beschreibt, und bestimme die Lösungsmenge. Wie ist die durchschnittliche Verteilung des Verkehrs in den verschiedenen Abschnitten des Kreisverkehrs für $x_4 = 500$?

Aufgabe 2: Bestimme die *Lösungsmenge* des folgenden Gleichungssystems über $\mathbb{R}$ in Abhängigkeit vom Parameter $t \in \mathbb{R}$, und gib in dem Fall, da die Lösung nicht eindeutig ist, eine *Parametrisierung* der Lösungsmenge an: (5)

$$\begin{array}{rcrcrcl} x & + & y & + & z & = & 1 \\ & & ty & + & z & = & 1 \\ tx & + & ty & + & z & = & 1+t \end{array}$$

Aufgabe 3: Überprüfe die folgende Abbildung auf Injektivität, Surjektivität und Bijektivität: (4)

$$f : \mathbb{R}^3 \to \mathbb{R}^2 : \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} x - 2y + z \\ 2x + z \end{pmatrix}.$$

Aufgabe 4: Seien M, N nicht-leere Mengen, $f : M \to N$ eine Abbildung. Beweise, f ist genau dann surjektiv, wenn für *alle* $X, Y \in \mathcal{P}(M)$ mit $M = X \cup Y$ gilt $f(X) \cup f(Y) = N$. (4)

Aufgabe 5: Es sei $(H, \cdot)$ eine Halbgruppe, und für alle $g, h, k \in H$ gelte: $h \cdot g = k \cdot g \;\Rightarrow\; h = k$. Zeige, ist H endlich, so ist $(H, \cdot)$ eine Gruppe. (4)

Aufgabe 6: Es sei $(G, \cdot)$ eine Gruppe.

a. Für $g, h \in G$ definiere $g \sim h :\Leftrightarrow h \in \{g, g^{-1}\}$.
   Zeige, $\sim$ ist eine Äquivalenzrelation auf G mit Äquivalenzklasse $[g] = \{g, g^{-1}\}$ für $g \in G$. (4)

b. Zeige, ist G endlich und gilt $2 \mid |G|$, so gibt es mindestens ein $e \neq g \in G$ mit $g^2 = e$. (6)

Aufgabe 7: Es sei $(G, \cdot)$ eine Gruppe. Für $g \in G$ bezeichne i_g den zu g gehörenden *inneren Automorphismus* aus Beispiel 1.3.7 (4) der Vorlesung (d. h. $i_g : G \to G : x \mapsto g^{-1} \cdot x \cdot g$). Zeige, die Abbildung $\varphi : G \to \mathbb{S}(G) : g \mapsto i_{g^{-1}}$ ist ein Gruppenhomomorphismus. (4)

Aufgabe 8: Es seien $n, k, l \in \mathbb{N} \setminus \{0\}$, und $\pi = (a_1 \cdots a_k) \in \mathbb{S}_n$ sei ein k-Zyklus, $\sigma = (b_1 \cdots b_l) \in \mathbb{S}_n$ ein l-Zyklus. Zeige, genau dann gibt es eine Permutation $\zeta \in \mathbb{S}_n$ mit $\zeta^{-1} \circ \pi \circ \zeta = \sigma$, wenn $k = l$. (7)

Hinweis: Ist $k = l$, so kann man ζ konkret hinschreiben! Für die andere Richtung betrachte man σ^l und π^l.

Aufgabe 9: Schreibe eine Singular-Prozedur `euler`, die eine Variable `x` vom Typ `number` und eine ganze Zahl `n` einliest und für `n`<0 den Wert -1, ansonsten den Wert $\sum_{k=0}^{n} \frac{x^k}{k!}$ ausgibt. Die Prozedur soll einen Example-Teil mit zwei Beispielen enthalten, bei denen `x` eine reelle Zahl ist. (7)

Hinweis: Es kann u. U. von Nutzen sein, den zugrunde liegenden Algorithmus in Worten anzugeben oder die Prozedur mit Kommentaren zu versehen.

FB Mathematik
Prof. Dr. Gert–Martin Greuel

Wintersemester 1999/2000
Thomas Keilen

2. Klausur in Linearer Algebra I

Klausurtermin: Samstag, 26. Februar 2000, 09:00-12:00 Uhr

Hinweise: Die Klausuraufgaben sind jeweils auf getrennten Blättern zu bearbeiten. *Nie zwei Aufgaben auf dem gleichen Blatt lösen!!!*

Alle nicht offensichtlichen Beweis-/Rechenschritte sind zu begründen (und ggf. durch ein Zitat zu belegen).

Die Zahlen in Klammern am rechten Seitenrand geben die Punktzahlen an, die durch Lösen der jeweiligen Aufgabe erreichbar sind. Insgesamt sind es 48 Punkte.

Jedes Blatt ist am oberen Rand der Vorderseite wie folgt zu beschriften:

eigener Name *Aufgabennummer* *Name des Übungsgruppenleiters*

Aufgabe 1: Es seien $B = \left((1,1,1)^t,(1,1,0)^t,(1,0,-1)^t\right)$ und $B' = \left((2,1)^t,(1,1)^t\right)$. E bzw. E' seien die kanonischen Basen des $\mathbb{R}^3$ bzw. des $\mathbb{R}^2$. Ferner sei $f \in \mathrm{Hom}_K\left(\mathbb{R}^3,\mathbb{R}^2\right)$ gegeben durch $f\left((x,y,z)^t\right) = (x-y+z, 2x+y)^t$.

a. Zeige, B und B' sind Basen des $\mathbb{R}^3$ bzw. des $\mathbb{R}^2$. (2)
b. Bestimme $M^E_{E'}(f)$. (1)
c. Bestimme $M^B_{B'}(f)$ sowie die Transformationsmatrizen T^B_E und $T^{E'}_{B'}$ mit $T^{E'}_{B'} \cdot M^E_{E'}(f) \cdot T^B_E = M^B_{B'}(f)$. (4)

Aufgabe 2: Sei V ein endlich-dimensionaler K-Vektorraum und $f \in \mathrm{End}_K(V)$. Zeige, daß dann die folgenden Aussagen gleichwertig sind:

a. $V = \mathrm{Ker}(f) \oplus \mathrm{Im}(f)$, (6)
b. $V = \mathrm{Ker}(f) + \mathrm{Im}(f)$,
c. $\mathrm{Ker}(f) \cap \mathrm{Im}(f) = \{0\}$.
d. $\mathrm{Ker}(f^2) = \mathrm{Ker}(f)$.

Aufgabe 3: Es sei K ein Körper, V ein endlich-dimensionaler K-Vektorraum, $U, U' \subseteq V$ Unterräume mit $V = U \oplus U'$. (4)
Zeige, ist $(u_1,\ldots,u_m)$ eine Basis von U, so ist $(u_1+U',\ldots,u_m+U')$ eine Basis von V/U'.

Aufgabe 4: Seien $U_1, U_2 \subseteq V$ Unterräume des K-Vektorraumes V. (4)
Zeige: $U_2/(U_1 \cap U_2) \cong (U_1 + U_2)/U_1$.

Aufgabe 5: Es seien $a, b \in \mathbb{Z}$. Eine Zahl $k \in \mathbb{Z}$ heißt *kleinstes gemeinsames Vielfaches* von a und b (kurz $\mathrm{kgV}(a, b)$), wenn gilt:
(i) $a \mid k$ und $b \mid k$, und
(ii) für alle $k' \in \mathbb{Z}$ mit $a \mid k'$ und $b \mid k'$ gilt $k \mid k'$. (4)
Zeige, $k \in \mathbb{Z}$ ist genau dann ein $\mathrm{kgV}(a, b)$, wenn $k\mathbb{Z} = a\mathbb{Z} \cap b\mathbb{Z}$.

Aufgabe 6: Ein Ring $(R, +, \cdot)$ heißt *nullteilerfrei*, falls gilt: $\forall\, x, y \in R$ mit $x \cdot y = 0$ gilt $x = 0$ oder $y = 0$. (4)
Zeige, ein endlicher, nullteilerfreier Ring mit $1 \neq 0$ ist ein Schiefkörper (d. h. $R^* = R \setminus \{0\}$).

Aufgabe 7: Es sei V ein n-dimensionaler K-Vektorraum und $f \in \mathrm{End}_K(V)$ mit $f^{n-1} \neq 0$, aber $f^n = 0$.

a. Zeige, es gibt ein $x \in V$, so daß $B = \big(x, f(x), \ldots, f^{n-1}(x)\big)$ eine Basis von V ist. (6)
b. Stelle $M_B^B(f)$ auf.*Matrixdarstellung* (1)

Anmerkung: Man nennt einen Vektor x wie in a. auch einen *zyklischen Vektor* von V.

Aufgabe 8: Es sei K ein Körper mit $\mathrm{char}(K) \neq 2$, und $n \in \mathbb{N}$ sei ungerade. Ferner sei $A \in \mathrm{Mat}(n \times n, K)$ mit $A^t = -A$. (4)
Zeige, A ist nicht invertierbar.

Aufgabe 9: Schreibe eine rekursive Singular-Prozedur `determinante`, die eine Matrix A einliest und, falls A quadratisch ist, die Determinante von A mittels des Laplace'schen Entwicklungssatzes berechnet und ausgibt. Ist die Matrix nicht quadratisch, soll eine Fehlermeldung ausgegeben werden. (8)

Anmerkung: Ein Example-Teil ist nicht erforderlich, aber die Prozedur sollte kommentiert sein!

FB Mathematik — Wintersemester 1999/2000
Prof. Dr. Gert–Martin Greuel — Thomas Keilen

Nachklausur in Linearer Algebra I

Klausurtermin: Samstag, 01. April 2000, 09:30-12:30 Uhr

Hinweise: Die Klausuraufgaben sind jeweils auf getrennten Blättern zu bearbeiten. *Nie zwei Aufgaben auf dem gleichen Blatt lösen!!!*

Alle nicht offensichtlichen Beweis-/Rechenschritte sind zu begründen (und ggf. durch ein Zitat zu belegen).

Die Zahlen in Klammern am rechten Seitenrand geben die Punktzahlen an, die durch Lösen der jeweiligen Aufgabe erreichbar sind. Insgesamt sind es 48 Punkte.

Jedes Blatt ist am oberen Rand der Vorderseite wie folgt zu beschriften:

eigener Name *Aufgabennummer* *Name des Übungsgruppenleiters*

Aufgabe 1: Es sei $B = \left((1,1,1)^t, (1,-1,1)^t, (1,1,-1)^t\right)$ eine Familie von Vektoren im $\mathbb{R}^3$, und E bezeichne die kanonische Basis des $\mathbb{R}^3$.

a. Zeige, B ist eine Basis. (1)

b. Für $k \in \mathbb{R}$ sei $f_k \in \mathrm{Hom}_{\mathbb{R}}\left(\mathbb{R}^3, \mathbb{R}^3\right)$ gegeben durch

$$M_B^E(f_k) = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & k \\ 0 & 1 & k \end{pmatrix}. \tag{3}$$

Bestimme $\det\left(f_k\right)$.

c. Zeige, f_k ist genau dann ein Isomorphismus, wenn $k \neq 0$. (1)

Aufgabe 2: Überprüfe die folgenden Abbildungen auf Surjektivität, Injektivität und Bijektivität.

a. $f : \mathbb{R}^2 \to \mathbb{R}^2 : (x,y)^t \mapsto (x+y, x-y)^t$, (2)

b. $g : \mathbb{Z} \to \mathbb{Z} : z \mapsto -2z + 5$. (2)

Aufgabe 3: Sei $\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 4 & 5 & 1 & 7 & 3 & 6 \end{pmatrix} \in \mathbb{S}_7$. Gib die Zyklenzerlegung von π^{-1} an und berechne das Signum $\mathrm{sgn}(\pi)$. (2)

Aufgabe 4: Es sei $(G, \cdot)$ eine Gruppe und $\alpha : G \to G : g \mapsto g^2$. Zeige, genau dann ist G abelsch, wenn α ein Gruppenhomomorphismus ist. (4)

Aufgabe 5: Sei K ein Körper. Setze $Sl_n(K) := \{A \in Gl_n(K) \mid \det(A) = 1\}$. Zeige, $Sl_n(K)$ ist ein Normalteiler der Gruppe $\big(Gl_n(K), \circ\big)$. (6)

Anmerkung: Bei geschicktem Vorgehen ist der Beweis ein Dreizeiler. – $Sl_n(K)$ heißt *spezielle lineare Gruppe*.

Aufgabe 6: Es seien R und S zwei Ringe mit Eins, $\alpha : R \to S$ ein Ring-*epi*morphismus und $I \subseteq R$ ein Ideal in R. Zeige, $\alpha(I)$ ist ein Ideal in S. (4)

Hinweis: Wie zeigt man geschickt, daß $\alpha(I)$ ein Untermodul des S-Moduls S - also ein Ideal - ist?

Aufgabe 7: Es sei V ein K-Vektorraum mit $\dim_K(V) = n < \infty$. Ein echter Unterraum $U \subsetneqq V$ heißt *maximal*, falls für alle Unterräume U' mit $U \subsetneqq U' \subseteq V$ gilt: $U' = V$. Zeige:

a. Ein Unterraum U ist genau dann maximal, wenn $\dim_K(U) = n - 1$. (6)
b. Für maximale Unterräume $U_1, \ldots, U_r$ gilt $\dim_K\big(U_1 \cap \cdots \cap U_r\big) \geq n - r$. (4)

Aufgabe 8: Sei V ein K-Vektorraum und $f \in \mathrm{End}_K(V)$ mit $f^2 = f$. (6)

a. Zeige, $V = \mathrm{Im}(f) \oplus \mathrm{Ker}(f)$.
b. Zeige, ist $\dim_K(V) < \infty$ und $r = \mathrm{rang}(f)$, so gibt es eine Basis B von V mit

$$M_B^B(f) = \left(\begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array}\right). \tag{4}$$

Anmerkung: $f^2 = f$ bedeutet $f\big(f(v)\big) = f(v)$ für alle $v \in V$. Man nennt f dann auch eine *Projektion*.

Aufgabe 9: Schreibe eine Singular-Prozedur `minimum`, die zwei ganze Zahlen `a` und `b` einliest und ihr Minimum zurück gibt. (3)

Anmerkung: Ein Example-Teil ist nicht erforderlich!

FB Mathematik
Prof. Dr. Gert–Martin Greuel

Sommersemester 2000
Thomas Keilen

1. Klausur in Linearer Algebra II

Klausurtermin: Samstag, 27. Mai 2000, 10:00-13:00 Uhr

Hinweise: Die Klausuraufgaben sind jeweils auf getrennten Blättern zu bearbeiten. *Nie zwei Aufgaben auf dem gleichen Blatt lösen!!!*

Die Zahlen in Klammern am rechten Seitenrand geben die Punktzahlen an, die durch Lösen der jeweiligen Aufgabe erreichbar sind. Insgesamt sind es 48 Punkte.

Jedes Blatt ist am oberen Rand der Vorderseite wie folgt zu beschriften:

eigener Name *Aufgabennummer* *Name des Übungsgruppenleiters*

Aufgabe 1: Man zeige, die folgende Matrix $A \in \mathrm{Mat}(3, \mathbb{R})$ ist diagonalisierbar und bestimme eine Transformationsmatrix $T \in \mathrm{Gl}_3(\mathbb{R})$, so daß $T^{-1} \circ A \circ T$ Diagonalgestalt hat: (5)

$$A = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix}.$$

Aufgabe 2: Es sei K ein Körper mit $\mathrm{char}(K) \neq 2$, V ein n-dimensionaler K-Vektorraum mit Basis $B = (x_1, \ldots, x_n)$, $n \geq 3$, und $U = \langle x_3, \ldots, x_n \rangle$. Ferner sei $f \in \mathrm{End}_K(V)$ mit $f(U) \subseteq U$ und $f(x_1) = f(x_2) = x_1 + x_2$. Zeige:

a. $\chi_f = t(t-2)\chi_{f_U}$. (2)
b. f ist genau dann trigonalisierbar, wenn f_U trigonalisierbar ist. (1)
c. Sind 0 und $2 = 2 \cdot 1_K$ keine Eigenwerte von f_U, so ist $\mu_f = t(t-2)\mu_{f_U}$. (2)
d. Genau dann ist f diagonalisierbar, wenn f_U diagonalisierbar ist. (2)
e. Welche der Aussagen wird falsch, wenn $\mathrm{char}(K) = 2$? Weshalb? (2)

Aufgabe 3: Es sei $V = \langle 1, t, t^2, t^3 \rangle \subseteq \mathbb{R}[t]$ der $\mathbb{R}$-Vektorraum der Polynome vom Grad kleiner gleich drei, und $0 \neq \lambda \in \mathbb{R}$. Durch

$$f_\lambda : V \to V : p \mapsto p(t + \lambda) - \lambda \cdot D(p)$$

wird ein Endomorphismus auf V definiert (das braucht nicht gezeigt zu werden). (Hier ist $D(p)$ die formale Ableitung von p.) (5)
Bestimme das charakteristische Polynom und das Minimalpolynom von f_λ sowie eine Matrixdarstellung von f_λ in Jordanscher Normalform.

Aufgabe 4: Es sei K ein Körper und $A = (a_{ij}) \in \mathrm{Mat}(4, K)$ eine obere Dreiecksmatrix mit $a_{11} = a_{22} \neq a_{33} = a_{44}$.
Zeige, A ist genau dann diagonalisierbar, wenn $a_{12} = a_{34} = 0$. (6)

Aufgabe 5: Es sei K ein Körper mit $\mathrm{char}(K) = 0$ und $\overline{K}$ sei ein algebraisch abgeschlossener Oberkörper von K. Zeige:

a. Ist $p \in K[t]$ mit $\deg(p) \geq 1$, so gilt $\deg\big(D(p)\big) = \deg(p) - 1$. (1)
b. Ist $p \in K[t]$ irreduzibel, so hat p in $\overline{K}$ keine mehrfache Nullstelle. (4)

Aufgabe 6: Es sei K ein Körper, $\lambda, \mu \in K$ mit $\mu \notin \{n\lambda \mid n \in \mathbb{N}\}$. Ferner sei V ein K-Vektorraum und $f, g \in \mathrm{End}_K(V)$ so, daß $g \circ f - f \circ g = -\lambda\,\mathrm{id}_V$ und daß μ ein Eigenwert von $g \circ f$ ist. (7)
Zeige, dann ist auch $\mu - n \cdot \lambda$ Eigenwert von $g \circ f$ für alle $n \in \mathbb{N}$.
Was bedeutet dies für die Dimension von V, falls $\lambda \neq 0$?
Hinweis: Man führe Induktion über n.

Aufgabe 7: Sei V ein $\mathbb{R}$-Vektorraum, $\dim_{\mathbb{R}}(V) = 2$ und $f \in \mathrm{End}_{\mathbb{R}}(V)$. (7)
Zeige, hat f^2 einen negativen Eigenwert λ, so hat f keinen Eigenwert.

Aufgabe 8: Schreibe eine Singular-Prozedur `diagonal`, die eine Matrix $A \in \mathrm{Mat}(n, \mathbb{Q})$ einliest, und feststellt, ob A über $\mathbb{C}$ diagonalisierbar ist. Falls ja, soll 1 ausgegeben werden, sonst 0. (4)
Hinweise: Die in den Übungen erstellten Prozeduren dürfen verwendet werden. Für ein Polynom $p \in \mathbb{Q}[x]$ liefert `diff(p,x)` die formale Ableitung von p.

FB Mathematik Sommersemester 2000
Prof. Dr. Gert–Martin Greuel Thomas Keilen

2. Klausur in Linearer Algebra II

Klausurtermin: Samstag, 08. Juli 2000, 10:00-13:00 Uhr

Hinweise: Die Klausuraufgaben sind jeweils auf getrennten Blättern zu bearbeiten. *Nie zwei Aufgaben auf dem gleichen Blatt lösen!!!*

Alle nicht offensichtlichen Beweis-/Rechenschritte sind zu begründen (und ggf. durch ein Zitat zu belegen).

Die Zahlen in Klammern am rechten Seitenrand geben die Punktzahlen an, die durch Lösen der jeweiligen Aufgabe erreichbar sind. Insgesamt sind es 48 Punkte.

Jedes Blatt ist am oberen Rand der Vorderseite wie folgt zu beschriften:

eigener Name *Aufgabennummer* *Name des Übungsgruppenleiters*

Aufgabe 1: Bestimme eine orthogonale Matrix $T \in O(3)$, die die folgende symmetrische Matrix $A \in \mathrm{Mat}(3, \mathbb{R})$ diagonalisiert:

$$A = \begin{pmatrix} 4 & -2 & 0 \\ -2 & 3 & 2 \\ 0 & 2 & 2 \end{pmatrix}. \tag{6}$$

Aufgabe 2: Es sei $(V, \langle \cdot, \cdot \rangle)$ ein euklidischer oder unitärer Raum und $\|\cdot\|$ bezeichne die durch das Skalarprodukt definierte Norm auf V. Zeige, daß für $x, y \in V$ stets gilt:

a. $\|x+y\|^2 + \|x-y\|^2 = 2\big(\|x\|^2 + \|y\|^2\big)$ ("Parallelogrammgleichung") (1)
b. Aus $x \perp y$ folgt $\|x\|^2 + \|y\|^2 = \|x+y\|^2$. ("Satz des Pythagoras") (1)

Aufgabe 3: Es sei V ein endlich-dimensionaler K-Vektorraum und $U_1, U_2 \subseteq V$ seien zwei Unterräume von V. Zeige:

a. $(U_1 + U_2)^o = U_1^o \cap U_2^o$. (2)
b. $(U_1 \cap U_2)^o = U_1^o + U_2^o$. (3)

Aufgabe 4: Es sei $V = \mathrm{Mat}(n, \mathbb{R})$. Wir definieren

$$(\cdot,\cdot) : V \times V \to \mathbb{R} : (A, B) \mapsto \mathrm{Spur}\left(A^t \circ B\right).$$

a. Zeige, $(\cdot,\cdot)$ ist ein Skalarprodukt auf V. (4)
b. Setze $U := \{A \in V \mid A^t = A\}$ und $W := \{A \in V \mid A^t = -A\}$. Zeige, $V = U \perp W$. (4)
c. Im Fall $n = 2$ zeige man, daß $B = \left(\left(\begin{smallmatrix}1&0\\0&0\end{smallmatrix}\right), \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right), \left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right)\right)$ eine Basis von U ist und bestimme eine Orthonormalbasis von U bez. $(\cdot,\cdot)$. (4)

Aufgabe 5: Es sei V ein endlich-dimensionaler unitärer Raum und $f \in \mathrm{End}_{\mathbb{C}}(V)$ so, daß es ein $m \in \mathbb{N}$ gibt mit $f^m = \mathrm{id}_V$. Zeige, dann sind die folgenden Aussagen äquivalent: (8)

a. f ist unitär.
b. f ist normal.
c. Für Eigenwerte $\lambda \neq \mu$ von f gilt $\mathrm{Eig}(f, \lambda) \perp \mathrm{Eig}(f, \mu)$.

Aufgabe 6: Es sei $V \neq 0$ ein endlich-dimensionaler unitärer Raum und $f \in \mathrm{End}_{\mathbb{C}}(V)$. Zeige, die folgenden Aussagen sind gleichwertig: (8)

a. $f^* = -f$.
b. Für alle $x \in V$ gilt: $\big(f(x), x\big) \in i\mathbb{R}$.
c. Es gibt eine Orthonormalbasis von V aus Eigenvektoren und der Realteil aller Eigenwerte ist Null.

Hinweis: Im Tutorium wurde gezeigt, daß für $g \in \mathrm{End}_{\mathbb{C}}(V)$ aus $\big(g(x), x\big) = 0$ für alle $x \in V$ folgt, daß $g = 0$ gilt. Diese Aussage darf verwendet werden.

Aufgabe 7: Es seien V und W zwei endlich-dimensionale K-Vektorräume und $f \in \mathrm{End}_K(V)$ sowie $g \in \mathrm{End}_K(W)$ zwei Endomorphismen. Zeige, $\mathrm{Spur}(f \otimes g) = \mathrm{Spur}(f) \cdot \mathrm{Spur}(g)$. (4)

Aufgabe 8: Schreibe eine Singular-Prozedur `positiv_definit`, die eine Matrix $A \in \mathrm{Mat}(n, \mathbb{Q})$ einliest und, falls A positiv definit ist, den Wert 1 zurück gibt, sonst den Wert 0. (3)

FB Mathematik
Prof. Dr. Gert–Martin Greuel

Sommersemester 2000
Thomas Keilen

Nachklausur zur Linearen Algebra II

Klausurtermin: Samstag, 21. Oktober 2000, 10:00-13:00 Uhr

Hinweise: Die Klausuraufgaben sind jeweils auf getrennten Blättern zu bearbeiten. *Nie zwei Aufgaben auf dem gleichen Blatt lösen!!!*

Alle nicht offensichtlichen Beweis-/Rechenschritte sind zu begründen (und ggf. durch ein Zitat zu belegen).

Die Zahlen in Klammern am rechten Seitenrand geben die Punktzahlen an, die durch Lösen der jeweiligen Aufgabe erreichbar sind. Insgesamt sind es 48 Punkte.

Jedes Blatt ist am oberen Rand der Vorderseite wie folgt zu beschriften:

eigener Name *Aufgabennummer* *Name des Übungsgruppenleiters*

Aufgabe 1: Bestimme die Jordansche Normalform der folgenden Matrix $A \in \mathrm{Mat}(3, \mathbb{Q})$ und gib eine Transformationsmatrix $T \in \mathrm{Gl}_3(\mathbb{Q})$ an, die sie in Jordansche Normalform überführt: (3)

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

Aufgabe 2: Bestimme eine Basis des orthogonalen Komplementes von $U = \langle (1,0,1)^t, (1,2,1)^t \rangle$ in $\mathbb{R}^3$ bezüglich des kanonischen Skalarproduktes. (3)

Aufgabe 3: Es sei $A \in \mathrm{Mat}(5, \mathbb{Q})$ eine Matrix mit $A^3 - 4A^2 + 4A = 0$ und $\mathrm{rang}(A) = 4$. Bestimme das charakteristische Polynom von A. (4)

Aufgabe 4: Es sei V ein endlich-dimensionaler K-Vektorraum, $U \subseteq V$ ein Unterraum mit Basis $(x_1, \ldots, x_r)$ und $x \in V$. Zeige, es gilt $x \in U$ genau dann, wenn $x_1 \wedge \ldots \wedge x_r \wedge x = 0$. (4)

Aufgabe 5: Es sei V ein endlich-dimensionaler euklidischer oder unitärer Raum und $f, g \in \mathrm{End}_{\mathbb{K}}(V)$ seien *normal*. Zeige die Äquivalenz der folgenden Aussagen: (5)

a. $f \circ g = 0$.
b. $g \circ f = 0$.
c. $\mathrm{Im}(f) \perp \mathrm{Im}(g)$.

Aufgabe 6: Zeige, ist $A \in \mathrm{Mat}(n, \mathbb{R})$ eine symmetrische Matrix mit $x^t A x \geq 0$ für alle $x \in \mathbb{R}^n$, dann gibt es eine Matrix $B \in \mathrm{Mat}(n, \mathbb{R})$ mit $B^2 = A$. (5)

Aufgabe 7: Es sei $n \geq 1$. Für $a_1, \ldots, a_n \in \mathbb{R}$ definieren wir

$$A(a_1, \ldots, a_n) := \begin{pmatrix} a_1 & 1 & 0 & \ldots & 0 \\ 1 & a_2 & 1 & \ddots & \vdots \\ 0 & 1 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & a_{n-1} & 1 \\ 0 & \ldots & 0 & 1 & a_n \end{pmatrix} \in \mathrm{Mat}(n, \mathbb{R}). \tag{6}$$

Zeige, $A(a_1, \ldots, a_n)$ ist genau dann positiv definit, wenn $a_1 > 0$, $a_2 > \frac{1}{a_1}$ und $a_j > \frac{\det\left(A(a_1, \ldots, a_{j-2})\right)}{\det\left(A(a_1, \ldots, a_{j-1})\right)}$ für $3 \leq j \leq n$.

Aufgabe 8: Es sei $V \neq 0$ ein endlich-dimensionaler K-Vektorraum und $f \in \mathrm{End}_K(V)$. Zeige, genau dann ist χ_f irreduzibel, wenn V nur die trivialen f-invarianten Unterräume $\{0\}$ und V besitzt. (8)

Hinweis: Für die Rückrichtung betrachte man zunächst einen normierten Primteiler p von χ_f und zeige $p = \mu_f$. p führt dann zu einem hilfreichen zyklischen Unterraum von V.

Aufgabe 9: Es sei K ein Körper und es seien $A \in \mathrm{Mat}(n, K)$ und $B \in \mathrm{Mat}(m, K)$ zwei Matrizen. Zeige:

a. $\det(A \otimes \mathbb{1}_m) = \det(A)^m$. (4)
b. $\det(A \otimes B) = \det(A)^m \cdot \det(B)^n$. (6)

Hinweise: In a. zeige man zunächst, daß $\det(A \otimes \mathbb{1}_m) = \det\left(\begin{array}{c|c} A \otimes \mathbb{1}_{m-1} & 0 \\ \hline 0 & A \end{array}\right)$. Zum Beweis von b. kann man a. verwenden.

ANHANG B

Kurzeinführung in Singular

Thomas Keilen
Fachbereich Mathematik
Universität Kaiserslautern
67553 Kaiserslautern
keilen@mathematik.uni-kl.de

SINGULAR Version 2.0.4
Universität Kaiserslautern
Fachbereich Mathematik und Zentrum für Computeralgebra
Autoren: G.–M. Greuel, G. Pfister, H. Schönemann

Die vorliegende Kurzeinführung in das Computeralgebrasystem SINGULAR erhebt keinerlei Anspruch auf Vollständigkeit. Ziel ist es, die Teilnehmer der Übungen zu den Vorlesungen *Lineare Algebra* und *Einführung in die Computeralgebra* schrittweise mit den notwendigen Begriffen und Befehlen vertraut zu machen, die notwendig sind, das Programm im Rahmen der Übungen einzusetzen. Dies bedingt insbesondere, daß der Einführung eine strikte Systematik fehlt. Wir verweisen deshalb für eine systematische und vollständige Dokumentation von SINGULAR auf das Handbuch [**GPS03**].

Zwei Gründe haben uns bewogen, SINGULAR auch in der Linearen Algebra zur Lösung kleinerer Probleme einzusetzen: zum einen ist das Programm SINGULAR kostenlos und für alle gängigen Computerplattformen erhältlich, zum anderen ist die Programmiersprache von SINGULAR der Sprache C angelehnt, was im weiteren Verlauf des Studiums von Vorteil sein mag. Wer das Programm SINGULAR auf dem eigenen Rechner installieren möchte, findet die Sourcen sowie Installationshinweise über die Singular Home Page:

`http://www.singular.uni-kl.de/`

1. Erste Schritte

1.1. Notationen. In der vorliegenden Einführung wollen wir uns auf folgende Notationen einigen:

- SINGULAR-Ein- und Ausgaben sowie reservierte Worte werden in der Schriftart Typewriter gesetzt, etwa `exit;` oder `help`.
- Das Symbol $\mapsto$ leitet SINGULAR-Ausgaben ein, z. B.:

```
int i=5;
i;
↦ 5
```

- Eckige Klammern bezeichnen Teile der Syntax, die optional sind, also fehlen können. Z. B.

```
pmat(M,[n]);
```

  Der obige Befehl, eine Prozedur der Bibliothek `matrix.lib` dient zur Ausgabe einer Matrix `M` als formatierte Matrix. Der optionale Parameter `n` gibt die Breite der Spalten in Zeichen an. Wenn er fehlt, wird ein Standardwert genommen.
- Tasten werden ebenfalls durch die Schriftart Typewriter gekennzeichnet, etwa:
  `n` (drücke die Taste `n`),
  `RETURN` (drücke die Eingabetaste),
  `CTRL-P` (drücke die Control-Taste und die Taste `P` zugleich).

1.2. SINGULAR aufrufen und beenden. Es versteht sich von selbst, daß die erste Frage die ist, wie man das Programm startet und wie man es wieder beendet. Auf den am Fachbereich vorhanden Rechnern ist die jeweils aktuellste Version von SINGULAR durch Eingabe des Kommandos

```
Singular
```

auf der Kommandozeile des Systems zu starten. Die Eingabe von

```
Singular -v
```

sorgt dafür, daß beim Start des Programms die Versionsnummer von SINGULAR angezeigt wird.

Nach dem Start liefert SINGULAR einen Eingabeprompt, ein $>$, zurück und steht dem Nutzer fortan zur interaktiven Nutzung zur Verfügung. Sobald selbiger Nutzer von dieser Möglichkeit nicht länger Gebrauch machen möchte, empfiehlt es sich, das Programm zu beenden. Hierzu stehen ihm drei Befehle zur Verfügung: `exit;`, `quit;` oder, für die ganz Schreibfaulen, `$`.

Man beachte hierbei, daß die Semikola im letzten Satz keine deplazierten Satzzeichen darstellen, sondern zu den SINGULAR-Befehlen gehören.

Generell schließt *jeder* Befehl in SINGULAR durch ein Semikolon ab!

Das Semikolon teilt dem Rechner mit, er möge die just eingegebene Befehlfolge doch bitte *interpretieren* und, sollte er dabei erfolgreich sein, auch *ausführen*. Das Programm meldet sich entsprechend mit dem Ergebnis (bzw. einer Fehlermeldung), gefolgt von einem neuen Eingabeprompt wieder. Sollte man das Semikolon vergessen haben oder eine geschweifte Klammer geöffnet und nicht wieder geschlossen haben, so zeigt einem Singular dies dadurch an, daß als Eingabeprompt ein `.`, sprich ein Punkt, erscheint und die Möglichkeit zu weiteren Eingaben, etwa dem fehlenden Semikolon, gibt. Auf diese Weise besteht die Möglichkeit, längere Befehlsfolgen über mehrere Zeilen zu strecken.

1.3. Die Online-Hilfe `help`**.** Neben dem Starten und Beenden des Programms ist die nächstwichtigste Information die, wie man sich Hilfe besorgt, wenn man feststeckt. Hierzu steht in Singular der Befehl `help`, oder kurz `?`, zur Verfügung. Gibt man den Befehl `help` gefolgt von einem SINGULAR-Befehl, einem SINGULAR-Funktions-/Prozedurnamen oder einer SINGULAR-Bibliothek, so werden Informationen zum jeweiligen Objekt angezeigt. Bei Bibliotheken erhält man eine Auflistung der darin enthaltenen Prozeduren, bei Befehlen, Funktionen und Prozeduren erfährt man ihren Zweck und findet die allgemeine Syntax sowie, ganz wichtig, Beispiele für ihre Anwendung.

Beispiele:

```
help exit;
help standard.lib;
help printf;
```

Man kann sich die Hilfe auf verschiedenen Ausgabemedien, Browsern, anzeigen lassen. Standardmäßig wird dies bei SINGULAR 2.0.4 Netscape sein. Das heißt, daß SINGULAR nach Eingabe etwa von `help exit;` Netscape startet und den zu `exit;` gehörenden Hilfetext dort anzeigt. (Über selbsterklärende Buttons steht damit das ganze Handbuch zur Verfügung.) Neben Netscape stehen noch weitere Browser zur Verfügung, von denen hier nur `info` und `builtin` genannt sein sollen. Ersterer dürfte den Nutzern von Unix-Systemen vertraut sein, letzterer zeigt den Hilfetext schlicht auf der laufenden SINGULAR-Seite an und hat den Vorteil, auf allen Computerplattformen und ohne zusätzliche Programme (wie Netscape oder Info) zu funktionieren.

Mittels des Befehls `system("browsers");` erfährt man, welche Browser SINGULAR kennt, und durch `system("--browser","builtin");` wechselt man den Browser von Netscape zu `builtin` – für andere Browser entsprechend. Ferner besteht die Möglichkeit, bereits beim Start von SINGULAR einen Browser zu wählen, indem das Programm etwa durch den folgenden Befehl gestartet wird:

```
Singular --browser=builtin
```

Während die Bedienung der Netscape-Hilfe selbsterklärend ist, benötigt man Hinweise, wenn man mit Info noch nicht gearbeitet hat. Wer nicht mit Info arbeitet, kann zu 1.6 gehen. Um sich innerhalb von Info fortzubewegen verwende man die unten aufgeführten Kommandos, die alle aus einzelnen Buchstaben bestehen. Beachte, man benutze *niemals* RETURN oder die Pfeiltasten! Einige Befehle lesen anschließend weitere Eingaben von der Kommandozeile am unteren Rand des Bildschirms. Hierbei steht die TAB Taste zur Vervollständigung eines teilweise eingegebenen Befehls zur Verfügung.

Einige wichtige Info-Kommandos:

q	Verlassen der Online-Hilfe
n	Vorwärtsblättern zum nächsten Menüpunkt
p	Rückwärtsblättern zum vorhergehenden Menüpunkt
m	Auswählen eines durch Namen spezifizieren Menüpunktes
f	Aufrufen eines Querverweises
l	Aufrufen des zuletzt besuchten Menüpunktes
b	Zurückblättern zum Beginn des Menüpunktes
e	Vorwärtsblättern zum Ende des Menüpunktes
SPACE	Vorwärtsscrollen um eine Seite
DEL	Zurückscrollen um eine Seite
h	Aufrufen der Info-Einführung
CTRL-H	Aufruf eines Kurzüberblicks über die Online-Hilfe
s	Durchsuchen des Handbuches nach einem bestimmten String
1,...,9	Aufrufen des i-ten Unterpunktes eines Menüs

1.4. SINGULAR unterbrechen. Unter Unix-ähnlichen Systemen und unter Windows NT besteht die Möglichkeit, SINGULAR durch die Tastenkombination CTRL-C zur Unterbrechung seiner Tätigkeit zu bewegen. (Funktioniert nicht bei ESingular!) SINGULAR reagiert durch Ausgabe des derzeit auszuführenden Befehls und erwartet weitere Anweisungen. Hier stehen folgende Optionen zur Auswahl:

a	SINGULAR führt den aktuellen Befehl noch aus und kehrt dann zum Toplevel zurück,
c	SINGULAR fährt fort,
q	das Programm SINGULAR wird beendet.

1.5. Eingaben editieren. Wer sich bei einem Kommando einmal verschrieben hat, oder ein früheres Kommando noch einmal benötigt, der muß sich nicht unbedingt die Mühe machen, alles noch einmal neu einzugeben. Vorhandener SINGULAR-Text kann editiert werden. Hierzu unterhält SINGULAR eine History aller Befehle einer SINGULAR-Sitzung. Wir wollen nur eine Auswahl der zur Verfügung stehenden Tastenkombinationen zum Editieren des Textes geben:

TAB	automatische Vervollständigung von Funktions- und Dateinamen
←	
CTRL-B	bewegt den Cursor nach links
→	
CTRL-F	bewegt den Cursor nach rechts
CTRL-A	bewegt den Cursor zum Zeilenanfang
CTRL-E	bewegt den Cursor zum Zeilenende
CTRL-D	löscht das Zeichen unter dem Cursor - nie auf leere Zeile anwenden!
BACKSPACE	
DEL	
CTRL-H	löscht das Zeichen vor dem Cursor
CTRL-K	löscht alles vom Cursor bis zum Zeilenende
CTRL-U	löscht alles vom Cursor bis zum Zeilenanfang
↓	
CTRL-N	liefert die nächste Zeile aus der History
↑	
CTRL-P	liefert die vorherige Zeile aus der History
RETURN	schickt die gegenwärtig Zeile zum SINGULAR-Parser

1.6. Prozeduren. Hat man ein konkretes Problem, das mit SINGULAR zu lösen ist, ruft man das Programm auf, gibt die Befehlsfolge ein und erhält ein Ergebnis. Häufig möchte man gleiche Rechnungen aber mit unterschiedlichen Eingaben durchführen. Dann ist es sinnvoll, die Befehlsfolge als Prozedur zu schreiben, der man die gewünschten Eingaben als Argumente übergibt und die die Lösungen zurückgibt.

Die Syntax einer Prozedur ist recht einfach:

```
proc PROZEDURNAME [PARAMETERLISTE]
{
  PROZEDURKÖRPER
}
```

Als PROZEDURNAME kommt jede noch nicht anderweitig vergebene Buchstabenfolge in Frage. In der PARAMETERLISTE sind die Typen und Namen der Argumente, die der Prozedur übergeben werden, festgelegt. Dabei ist die PARAMETERLISTE durch runde Klammern einzuschließen. Der PROZEDURKÖRPER enthält eine Abfolge von zulässigem SINGULAR-Code. Soll die Prozedur ein Ergebnis zurückliefern, so sollte dieses Ergebnis in einer Variablen ergebnis gespeichert werden und die Prozedur sollte mit dem Befehl return(ergebnis); enden.

Ein Beispiel sagt meist mehr als tausend Worte:

```
proc permcol (matrix A, int c1, int c2)
{
  matrix B=A;
  B[1..nrows(B),c1]=A[1..nrows(A),c2];
  B[1..nrows(B),c2]=A[1..nrows(A),c1];
  return(B);
}
```

Die Prozedur `permcol` soll zwei Spalten einer Matrix vertauschen. Hierfür erwartet sie drei Argumente. Das erste Argument erhält den Namen `A` und ist vom Typ `matrix`, die beiden folgenden heißen `c1` und `c2` und sind vom Typ `integer`. Es folgen SINGULAR-Anweisungen und das Ergebnis wird in der Variablen `B` vom Typ `matrix` gespeichert, die dann mit `return(B);` zurückgegeben wird. Das bedeutet insbesondere, daß das Ergebnis der Prozedur vom Typ `matrix` ist (siehe Abschnitt 2).

Eine Prozedur wird aufgerufen, indem man den Prozedurnamen, gefolgt von den Argumenten in runden Klammern eingibt. Z. B.

```
LIB "matrix.lib"; LIB "inout.lib"; ring r=0,(x),lp;
matrix A[3][3]=1,2,3,4,5,6,7,8,9;
pmat(A,2);
↦ 1  2  3
  4  5  6
  7  8  9

matrix B=permcol(A,2,3);
pmat(B,2);
↦ 1  3  2
  4  6  5
  7  9  8
```

Variablen, die innerhalb einer Prozedur definiert werden, sind nur dort bekannt, und können deshalb durchaus die gleichen Namen haben, wie Objekte, die außerhalb der Prozedur definiert sind.

1.7. Bibliotheken. Um Prozeduren für mehr als eine SINGULAR-Sitzung verfügbar zu machen, ist es sinnvoll, sie in Dateien abzuspeichern, die später von Singular wieder eingelesen werden können – sogenannten Bibliotheken (= Libraries). Die Namen der Bibliotheken lassen meist Rückschlüsse auf die enthaltenen Prozeduren zu, und tragen grundsätzlich die Endung `.lib`. Bibliotheken werden in SINGULAR eingelesen durch den Befehl `LIB` gefolgt von dem in Anführungszeichen `"` eingeschachtelten Bibliotheksnamen, etwa

```
LIB "KeilenT.lib";
```

(Bibliotheksnamen sollten nach Möglichkeit nur *acht* Zeichen lang sein, um die Kompatibilität mit Betriebssystemen wie Dos zu gewährleisten!) Sofern

es sich nicht um SINGULAR-eigene Bibliotheken handelt, sollten sie sich in dem Verzeichnis befinden, von dem aus SINGULAR gestartet wurde.

Jeder Übungsteilnehmer sollte für die Prozeduren, die im Laufe des Semesters zu schreiben sind, eine Bibliothek anlegen – vorzugsweise unter dem eigenen Namen, etwa `KeilenT.lib`.

Natürlich muß auch eine Bibliothek gewissen Syntaxregeln entsprechen, und Prozeduren, die in Bibliotheken abgespeichert werden, sollten um zwei erläuternde Zusätze erweitert werden. Wir verdeutlichen dies an einem Musterbeispiel.

```
////////////////////////////////////////////////////////////////////
version="1.0";

info="
     LIBRARY:     KeilenT.lib LOESUNGEN ZU DEN LA-UEBUNGEN
     AUTHOR:      Thomas Keilen, email: keilen@mathematik.uni-kl.de
     PROCEDURES:
       permcol(matrix,int,int)  vertauscht Spalten der Matrix
       permrow(matrix,int,int)  vertauscht Zeilen der Matrix
";
////////////////////////////////////////////////////////////////////
LIB "inout.lib";
////////////////////////////////////////////////////////////////////
proc permcol (matrix A, int c1, int c2)
 "USAGE:   permcol(A,c1,c2); A matrix, c1,c2 positive integers
 RETURN:   matrix, A being modified by permuting column c1 and c2
 NOTE:     Platz für wichtige Anmerkungen,
           auch über mehrere Zeilen gestreckt
 EXAMPLE:  example permcol; shows an example"
{
  matrix B=A;
  B[1..nrows(B),c1]=A[1..nrows(A),c2];
  B[1..nrows(B),c2]=A[1..nrows(A),c1];
  return(B);
}
example
{
"EXAMPLE:";
echo = 2;
ring r=0,(x),lp;
matrix A[3][3]=1,2,3,4,5,6,7,8,9;
pmat(A);
pmat(permcol(A,2,3));
```

```
}
⋮
```

Taucht innerhalb einer Zeile ein Doppelslash // auf, so wird der Rest der Zeile als Kommentar interpretiert und ignoriert.

Der erste Abschnitt, der zwischen den beiden Kommentarzeilen steht, ist sozusagen der Kopf der Bibliothek. Die erste Zeile enthält das reservierte Wort `version`, durch das die Versionsnummer der Bibliothek festgelegt wird. Dem reservierten Wort `info` folgen allgemeine Informationen zur Bibliothek. Man sollte beachten, daß unter dem Punkt `PROCEDURES:` alle Prozedurnamen, die in der Bibliothek enthalten sind, mit einer maximal einzeiligen Beschreibung aufgeführt werden. Diesen Teil zeigt SINGULAR an, wenn die Hilfe zur entsprechenden Bibliothek angefordert wird, etwa

```
help KeilenT.lib;
```

Man beachte auch, daß sowohl `version`, als auch `info` durch das Gleichheitszeichen, `=`, Strings zugewiesen werden, so daß die Anführungszeichen `"`, die sie einschachteln ebenso erforderlich sind, wie das Semikolon am Ende der Zeile!

Abschnitt zwei dient dem Einladen von anderen Bibliotheken, deren Prozeduren bei den eigenen Prozeduren benötigt werden. Im Beispiel die Bibliothek `inout.lib`, deren Prozedur `pmat` im `example`-Teil der Prozedur `permcol` verwendet wird.

Im dritten Abschnitt folgen die Prozeduren, schlicht aneinander gereiht. (Dabei ist darauf zu achten, daß der Befehl `proc` stets am Beginn einer neuen Zeile steht!) Es empfiehlt sich, die in Abschnitt 1.6 angegebene Syntax für Prozeduren um zwei Abschnitte zu erweitern. Zwischen Prozedurkopf und -körper kann man, in Anführungszeichen `"` eingeschachtelt, einen Kommentarblock einschieben, der gewisse Schlüsselworte gefolgt von zugehöriger Information enthält. Unter `USAGE:` sollte eingegeben werden, wie der Befehl aufgerufen wird, und ggf. von welchem Typ die Argumente sind. `RETURN:` sollte Informationen darüber enthalten, von welchem Typ die Rückgabe ist, und ggf. weitere Informationen. `NOTE:` dient dazu wichtige Hinweise, zur Prozedur, ihrer Bedienung, etc. zu geben. `EXAMPLE:` gibt schließlich einen Hinweis darauf, wie man sich unter SINGULAR ein Beispiel zeigen lassen kann. Der hier erläuterte Kommentarblock enthält die Information, die angezeigt wird, wenn man unter SINGULAR Hilfe zu einer Prozedur anfordert, etwa durch

```
help permcol;
```

Der zweite Zusatzabschnitt am Ende der Prozedur wird durch das reservierte Wort `example` eingeleitet, gefolgt von einem Abschnitt in geschweiften Klammern, der SINGULAR-Code enthält. Ziel ist es, ein Beispiel für die Wirkung

der Prozedur zu geben, die dem Nutzer die Verwendung erleichtert. Der Nutzer erhält das Beispiel durch Eingabe von `example PROZEDURNAME;`.

1.8. Ausgabe in Dateien / Einlesen aus Dateien. Der Befehl `write` bietet die Möglichkeit, die Werte von Variablen oder beliebige Strings in einer Datei abzuspeichern. Hierzu werden die Variablenwerte in Strings umgewandelt. Die folgenden Zeilen speichern Variablenwerte bzw. einen String in der Datei `hallo.txt`:

```
int a=5;
int b=4;
write("hallo.txt",a,b);
write("hallo.txt","Das ist Singular.");
```

Es können also auch mehrere Variablen oder Strings durch Kommata getrennt angegeben werden. Ihre Werte werden jeweils in eine neue Zeile geschrieben.

Daten, die sich in einer Datei befinden, können mit dem Befehl `read` eingelesen werden. Dabei werden sie jedoch als Strings interpretiert, z. B.

```
read("hallo.txt");
↦ 5
  4
  Das ist Singular.
```

Soll SINGULAR-Code, der aus einer Datei eingelesen wird, auch als solcher erkannt werden, dann muß der `read`-Befehl an den Befehl `execute` übergeben werden. Enthalte etwa die Datei `hallo.txt` die folgenden Zeilen,

```
4*5-3;
6/3;
```

dann führt der Befehl

```
execute(read("hallo.txt"));
```

zu folgendem SINGULAR-Output:

```
↦ 17
  2
```

Eine Kurzform für `execute(read(...))` ist $<$, z. B.

```
< "hallo.txt";
```

Wer eine SINGULAR-Sitzung zur Sicherheit in einer Datei, etwa `hallo.txt`, dokumentieren will, kann dies mit dem Befehl `monitor` tun, z. B.

```
monitor("hallo.txt","io");
```

Die Option `"io"` bewirkt, daß sowohl Eingaben (input), als auch Ausgaben (output) gespeichert werden. Entsprechend führt das Weglassen eines der

beiden Buchstaben dazu, daß nur Eingaben oder nur Ausgaben gespeichert werden. Die Option `monitor` erweist sich dann als sehr hilfreich, wenn man unter einem Betriebssystem arbeitet, auf dem `Singular` instabil läuft bzw. auf dem man keinen gut handhabbaren Editor zur Verfügung hat. Man beachte, daß `monitor` eine Datei öffnet, aber nicht wieder schließt. Dies kann explizit durch folgende Eingabe erreicht werden:

```
monitor("");
```

2. Datentypen in SINGULAR und Ringe

SINGULAR arbeitet mit einer ganzen Reihe von unterschiedlichen Strukturen, die als verschiedene Datentypen vorliegen. Will man ein Objekt in SINGULAR definieren, sprich eine Variable einführen, so ist es notwendig, ihr von Beginn an einen Datentyp zuzuweisen.

In SINGULAR sind die Datentypen, bis auf die Ausnahmen `string`, `int`, `intvec` und `intmat`, von einer Metastruktur abhängig, dem sogenannten Ring, über dem sie leben. (Es ist Teil der Vorlesung Lineare Algebra, zu definieren, was ein Ring ist, und welche Ringe in Singular zur Verfügung stehen.) Will man eine Rechnung in SINGULAR durchführen, ist es deshalb stets unabdingbar, zunächst den Ring zu definieren, über dem man arbeitet. Für die Lineare Algebra werden wir zu Beginn mit den folgenden Ringdefinitionen auskommen:

`ring r=0,x,lp;`	Die Menge der Polynome in der Variablen x mit Koeffizienten in den rationalen Zahlen $\mathbb{Q}$.
`ring r=(0,a,b),(x,y,z),lp;`	Die Menge der Polynome in den Variablen x, y, z, wobei die Koeffizienten rationale Ausdrücke in den Variablen a und b sind. Natürlich können statt a, b bzw. x, y, z auch beliebige andere Variablen stehen. Wesentlich ist, daß die Variablen in der ersten Klammer im Nenner von Brüchen auftauchen dürfen, die in der zweiten Klammer nicht.
`ring r=(real,15),x,lp;`	Die Menge der Polynome in der Variablen x mit Koeffizienten in den rellen Zahlen $\mathbb{R}$ - bei Rechnung mit 15 Nachkommastellen.

Wir werden zunächst über den rationalen Zahlen $\mathbb{Q}$ rechnen. Reelle Zahlen als Dezimalzahlen (floating point numbers) oder gar komplexe Zahlen werden wir erst zu einem späteren Zeitpunkt benötigen.

Im folgenden geben wir eine Liste der in `Singular` verfügbaren Datentypen, und wir geben auch jeweils ein Beispiel an, indem wir eine Variable des entsprechenden Typs definieren und ihr einen Wert zuweisen, durch den Operator =. Für die Lineare Algebra werden wir zunächst mit den ersten acht Typen auskommen. Die anderen werden nur der Vollständigkeit halber aufgeführt.

`int i=1;`

Der Datentyp `integer` repräsentiert die Maschinenzahlen (= ganze Zahlen zwischen -2^{31} und $2^{31}-1$). Außerdem werden Wahrheitswerte (= `boolean`) als `integers` repräsentiert, `0 = FALSE`, `1 = TRUE`.

`string s="Hallo";`

`strings` sind beliebige Zeichenketten. Stets durch Anführungszeichen eingegrenzt.

`intvec iv=1,2,3,4;`

Ein Vektor aus `integers`.

`intmat im[2][3]=1,2,3,4,5,6;`

Eine Matrix mit 2 Zeilen und 3 Spalten mit `integer`-Einträgen, hier $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$.

`ring R=(0,a),(x,y),lp;`

Der Ring $\mathbb{Q}(a)[x,y]$ mit lexikographischer Ordnung. Für weitere Erläuterungen konsultiere man das Handbuch [**GPS03**].

`number n=4/6;`

`numbers` sind die Elemente des Körpers, der dem Ring zugrunde liegt. Bei `ring r=0,x,lp;` also die rationalen Zahlen, bei `ring r=(0,a),x,lp;` auch Brüche von Polynomen in a mit ganzzahligen Koeffizienten, etwa $\frac{a^2+1}{a-1}$.

`list l=n,iv,s;`

Eine Liste kann Objekte ganz unterschiedlicher Typen enthalten. Auf den zweiten Eintrag von `l` kann durch `l[2]` zugegriffen werden.

`matrix m[2][3]=1,2,3,4,5,6;` Eine Matrix mit 2 Zeilen und 3 Spalten, bei der die Einträge entweder vom Typ `poly` oder vom Typ `number` sind, wie hier $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$.

`vector v=[1,2,3];` Ein Vektor im Modul R^3. Sind die Einträge sämtlich vom Typ `number`, können wir ihn aber auch als Vektor über dem Grundkörper auffassen.

`proc` Der Datentyp `procedure` ist in Kapitel 1.6 ausführlich besprochen.

`poly f=x2+2x+1;` Ein Polynom in den Veränderlichen des Rings mit `numbers` als Koeffizienten, hier $f = x^2 + 2x + 1$. Beachte, daß Zahlen vor den Monomen als Koeffizienten interpretiert werden, wohingegen SINGULAR Zahlen nach einzelnen Variablen als Exponenten interpretiert.

`ideal i=f,x3;` Das von f und x^3 erzeugte Ideal in `R`.

`qring Q=i;` Der Quotientenring `R/i`.

`map g=R,x;` Die Restklassenabbildung von `R` nach `Q`, die durch $x \mapsto \overline{x}$ definiert wird.

`module mo=v,[x,x2,x+1];` Der von den Vektoren `v` und $(x, x^2, x+1)^t$ in R^3 aufgespannte Modul.

`def j;` Will man sich zum Zeitpunkt der Definition einer Variablen noch nicht festlegen, welchen Typ sie haben soll, so definiert man sie als `def`. Die erste Zuweisung, mit der der Variablen ein Wert zugewiesen wird, legt dann auch den Datentyp fest.

`link` Für den Datentyp `link` verweisen wir auf das Handbuch [**GPS03**].

`resolution` Für den Datentyp `resolution` verweisen wir auf das Handbuch [**GPS03**].

Auf den ersten Blick mag es erscheinen, als ob die Matrizen `im` und `m` identisch seien. Für SINGULAR ist das jedoch nicht der Fall, da sie von unterschiedlichem Typ sind!

Will man mit Dezimalzahlen rechnen, also gleichsam den Grundkörper $\mathbb{R}$ zur Verfügung haben, so muß man in der Definition des Rings die "Charakteristik" 0 durch `real` ersetzen (bzw. `(real,50)`, wenn man mit 50 Nachkommastellen rechnen will), z. B.

```
ring r=(real,10),x,lp;
```

Sogar die komplexen Zahlen sind verfügbar, indem man `real` durch `complex` ersetzt. `i` bezeichnet dann die imaginäre Einheit, d. h. die Quadratwurzel aus -1.

3. Einige Elemente der Programmiersprache SINGULAR

3.1. Zuweisungen. Will man in Singular Berechnungen durchführen, ist es in aller Regel unabdingbar, daß man Variablen Werte zuweist. Dies geschieht durch den Operator `=`. Man kann einer Variablen bereits bei ihrer Definition einen Wert zuweisen,

```
int i=1;
```

oder dies zu einem späteren Zeitpunkt nachholen,

```
int i;
⋮
i=2;
```

3.2. Schleifen. Es gibt zwei Typen von Schleifen, die `for`- und die `while`-Schleifen.

Die `for`-Schleife wird typischerweise verwendet, wenn man eine Befehlssequenz mehrfach ausführen möchte und die Anzahl bereits vor Eintritt in die Schleife bekannt ist. Z. B.

```
int s=0;
int i;
for (i=1; i<=10; i=i+1)
{
  s=s+i;
}
```

In gescheiften Klammern steht die Befehlssequenz, die iterativ ausgeführt werden soll; in runden befinden sich Anweisungen, wie häufig die Schleife durchlaufen werden soll. Der erste Eintrag legt die Laufvariable (vom Typ `integer`) fest; der zweite Eintrag gibt die Abbruchbedingung, d. h. die Schleife wird nur solange durchlaufen, wie der Ausdruck den Wert `TRUE` ergibt; der dritte Eintrag legt fest, wie sich die Laufvariable in jedem Durchgang ändern soll. Das Beispiel berechnet somit die Summe der ersten zehn natürlichen Zahlen.

`while`-Schleifen bieten sich an, wenn die Anzahl der Durchläufe nicht a priori klar ist. Z. B.

```
int s=10000;
int i=1;
while (s > 50)
{
  i=i*i;
  s=s-i;
}
```

Wieder folgt in geschweiften Klammern die Befehlssequenz, während in runden Klammern nur die Abbruchbedingung steht. Solange diese den Wert `TRUE` liefert, wird die Schleife ausgeführt.

Bei beiden Schleifen wird die Abbruchbedingung vor dem ersten Eintritt in die Schleife überprüft!

3.3. Verzweigungen. SINGULAR bietet als Verzweigung die `if-else`-Anweisung, wobei der `else`-Anteil fehlen kann. Z. B.

```
int i=10;
int s=7;
if (i<5 or s<10)
{
  s=5;
}
else
{
  s=0;
}
```

Wieder stehen die Befehlssequenzen als Block in geschweiften Klammern, während sich in runden Klammern die Verzweigungsbedingung findet.

3.4. Vergleichsoperatoren. In Singular gibt es die Vergleichsoperatoren `==` und `!=`, mit denen man Objekte vom gleichen Datentyp (etwa `int`, `string`, `matrix`, etc.) miteinander vergleichen kann. `==` testet auf Gleichheit und liefert mithin den Wert `1`, wenn die Objekte gleich sind, und ansonsten `0`. `!=` testet auf Ungleichheit. Den gleichen Effekt hat `<>`.

Für die Datentypen `int`, `number`, `poly` und `vector` stehen zudem die Operatoren `<`, `>`, `<=` und `>=` zur Verfügung. Ihre Bedeutung für `integers` ist klar. Für die anderen Datentypen verweisen wir auf das Handbuch [**GPS03**].

3.5. Einige weitere ausgewählte Operatoren in SINGULAR. Wie bereits gesehen, hängen die Operatoren, die zur Verfügung stehen, vom jeweiligen Datentyp ab. Wir beschränken uns auf die in der Linaren Algebra zunächst benötigten Datentypen.

3.5.1. *boolean.* Für `boolean`-Ausdrücke sind die Verknüpfungsoperatoren `and` und `or` sowie der Negierungsoperator `not` definiert.

```
not ((1==0) or (1!=0));
↦ 0
```

3.5.2. *int.* Für `integers` sind die Operationen +, - und * gänzlich unproblematisch. ^ bedeutet potenzieren

```
int i=4;
i^3;
↦ 64
```

Etwas schwieriger sind die Befehle `div` und `mod`, wobei ersterer synomym zu / ist. Führt man für zwei ganze Zahlen Division mit Rest durch, so liefert `mod` den Rest, und `div` das Ergebnis ohne Rest. Z. B. $7 = 2 * 3 + 1$, also

```
7 div 3;
↦ 2
7 mod 3;
↦ 1
```

3.5.3. *list.* Für den Datentyp `list` gibt es die folgenden Operatoren:

+	Fügt die Elemente von zwei Listen zu einer zusammen.
`delete`	Löscht ein Element aus einer Liste, `delete(L,3)` löscht das dritte Element der Liste `L`.
`insert`	Fügt einer Liste ein Element hinzu. `insert(L,4)` fügt der Liste `L` das Element `4` an der ersten Stelle hinzu, `insert(L,4,2)` an der zweiten.

3.5.4. *matrix.* Die Operatoren +, - und * stehen mit ihren offensichtlichen Bedeutungen zur Verfügung.

Wir verdeutlichen an Beispielen, wie auf einzelne Einträge einer Matrix bzw. ganze Zeilen oder Spalten einer Matrix zugegriffen werden kann:

```
matrix m[2][3]=1,2,3,4,5,6;
print(m);
↦  1,2,3,
   4,5,6
m[1,2];
↦ 2;
m[1,1..3];
↦ 1 2 3
m[1..2,3];
↦ 3 6
```

4. Einige ausgewählte Funktionen in SINGULAR

SINGULAR verfügt über ein recht beachtliches Arsenal an Funktionen, die zum Teil im Singular-Kern integriert sind, zum Teil über Bibliotheken zur

Verfügung gestellt werden. Wir wollen hier nur eine kleine Auswahl von Funktionsnamen geben, die für die Bearbeitung der Übungsblätter in Linearer Algebra von Nutzen sein können. Über ihre Syntax sollte man sich mittels `help` oder im Handbuch informieren.

4.1. Funktionen, die mit dem Datentyp `matrix` in Zusammenhang stehen. `ncols, nrows, print, size, transpose, det,` als Funktionen im Kern von SINGULAR. Ferner die Funktionen der Bibliothek `matrix.lib`, insbesondere `permrow, permcol, multrow, multcol, addrow, addcol, concat, unitmat, gauss_row, gauss_col, rowred, colred.` Auch die Funktion `pmat` aus der Bibliothek `inout.lib` ist interessant.

4.2. Funktionen, die mit dem Datentyp `int` in Zusammenhang stehen. `random, gcd, prime` als Funktionen im Kern von `Singular`.

5. `ESingular` - oder der Editor Emacs

Es gibt viele Editoren, in denen man SINGULAR-Prozeduren und Bibliotheken schreiben kann. Unter Unix-ähnlichen Systemen bietet sich der Editor Emacs (oder Xemacs) an, da er den eingegebenen Code durch entsprechende farbige Unterlegung der Schlüsselworte leichter überschaubar macht und eine Vielzahl von Optionen bietet, die das Editieren und finden von Fehlern erleichtern.

Aber noch aus einem weiteren Grund empfiehlt es sich, Emacs zu benutzen. SINGULAR kann in einem speziellen Emacs-Modus gestartet werden, nämlich als `ESingular`. Dies bedeutet, daß zunächst der Editor Emacs gestartet wird, und dann innerhalb von Emacs das Programm SINGULAR. Der Vorteil besteht darin, daß neben der vollen Funktionalität des Editors Emacs für das Editieren von Dateien eine Reihe weiterer Optionen zur Verfügung gestellt werden, die die Bedienung vereinfachen – insbesondere für den unerfahrenen Nutzer, dem Pulldown-Menüs zur Verfügung stehen. Mittels

```
ESingular --emacs=xemacs
```

besteht die Möglichkeit, die Version von Emacs, die benutzt werden soll, festzulegen, in diesem Fall der Xemacs. Alternativ kann der Standard mittels der Environment-Variablen `EMACS` verändert werden.

6. Aufgaben

Aufgabe 6.1
Schreibe eine Prozedur `binomi`, die zwei natürliche Zahlen `n` und `k` einliest und den Binomialkoeffizienten $\binom{n}{k}$ zurückgibt. (Vereinbarung: falls $k < 0$ oder $k > n$, dann $\binom{n}{k} = 0$.)

Aufgabe 6.2
Schreibe eine Prozedur `quadratsumme`, die eine natürliche Zahl `n` einliest und die Summe der Quadratzahlen $1^2, 2^2, 3^2, \ldots, n^2$ ausgibt.

Aufgabe 6.3
Schreibe eine Prozedur `minimum`, die einen Vektor von natürlichen Zahlen einliest und das Minimum der Zahlen ausgibt.

Aufgabe 6.4
Schreibe Prozeduren `zeilensummennorm`, `maximumsnorm` und `q_eukl_norm`, die eine $(m \times n)$-Matrix `A` von reellen Zahlen einlesen und

a. die Zeilensummennorm von `A` (d. h. $\max_{i=1,\ldots,m} \left(\Sigma_{j=1}^{n} |A_{ij}|\right)$),
b. die Maximumsnorm von `A` (d. h. $\max\left(|A_{ij}| \;\middle|\; i = 1,\ldots,m,\ j = 1,\ldots,n\right)$), respektive
c. das Quadrat der euklidischen Norm berechnen (d. h. $\Sigma_{i,j}|A_{ij}|^2$).

Für den Absolutbetrag verwende die Funktion `abs` aus der Bibliothek `linalg.lib`.

Aufgabe 6.5
Schreibe eine Singular-Prozedur `gauss_reduction`, die eine Matrix A einliest und die mittels Gauß-Elimination ermittelte Zeilen-Stufen-Form der Matrix ausgibt. Die Einträge der Matrizen sollen vom Typ `number` sein. Es dürfen die in `matrix.lib` definierten Prozeduren `multrow`, `addrow` und `permrow` verwendet werden. Teste Deine Ergebnisse mit der Prozedur `gnf` aus der Bibliothek `linalg.lib`.

Aufgabe 6.6
Schreibe eine Prozedur `ebenenschnitt`, die die Koeffizienten zweier Ebenengleichungen einliest und den Schnitt der Ebenen ausgibt. Verwende die Prozedur `gauss_reduction` aus Aufgabe 6.5.

7. Lösungen

Lösung zu Aufgabe 6.1

```
proc binomi (int n, int k)
"USAGE: binomi(n,k); int n, int k
RETURN: int, Binomialkoeffizient n ueber k
EXAMPLE: example binomi; zeigt ein Beispiel"
{
  if ((k < 0) or (k > n))
  {
    return(0);
  }
  else
  {
    int i;
    int nenner,zaehler1,zaehler2 = 1,1,1;
```

```
    for (i=1;i<=n;i++)
    {
     nenner = nenner * i;
    }
    for (i=1;i<=k;i++)
    {
     zaehler1 = zaehler1 * i;
    }
    for (i=1;i<=n-k;i++)
    {
     zaehler2 = zaehler2 *i;
    }
    return (nenner / (zaehler1 * zaehler2));
  }
}
example
{
  "Beispiel:";
  echo = 2;
  binomi(5,2);
  binomi(7,5);
}
```

Lösung zu Aufgabe 6.2

```
proc quadratsumme (int n)
"USAGE: quadratsumme(n); int n
RETURN: int, Summe der ersten n Quadratzahlen
EXAMPLE: example quadratsumme; zeigt ein Beispiel"
{
  if (n < 0)
    {
     return (0);
    }
    else
    {
     int i;
     int ergebnis = 0;
     for (i=1;i<=n;i++)
     {
      ergebnis = ergebnis + i*i;
     }
     return (ergebnis);
    }
}
example
{
  "Beispiel:";
```

```
  echo = 2;
  quadratsumme(3);
  quadratsumme(5);
}
```

Lösung zu Aufgabe 6.3

```
proc minimum (intvec iv)
"USAGE: minimum(iv); iv intvector
RETURN: int, the minimum of the entries in iv
EXAMPLE: example minium; shows an example"
{
  int i;
  int k=size(iv);
  int ergebnis=iv[1];
  for (i=2;i<=k;i++)
    {
     if (iv[i] < ergebnis)
     {
      ergebnis=iv[i];
     }
    }
  return(ergebnis);
}
example
{
  "EXAMPLE:";
  echo=2;
  intvec iv=3,2,5,2,1;
  print(iv);
  minimum(iv);
  iv =-3,4,5,3,-6,7;
  print(iv);
  minimum(iv);
}
```

Lösung zu Aufgabe 6.4

Wir schreiben zunächst eine eigene kurze Prozedur zum Berechnen des Absolutbetrages.

```
proc abs_val (poly r)
"USAGE: abs_val(r); poly r - eine rationale/reelle Zahl
RETURN: poly, gibt den Absolutbetrag von r wieder, falls r eine reelle
Zahl
EXAMPLE: example abs_value; zeigt ein Beispiel"
{
  if (r < 0)
  {
   return(-r);
  }
```

```
  else
  {
   return(r);
  }
}
example
{
  "Beispiel:";
  echo = 2;
  ring r=real,x,lp;
  abs_val(-5.45);
  ring s=0,x,lp;
  abs_val(-4/5);
}
proc zeilensummennorm (matrix A)
"USAGE: zeilensummennorm(A); matrix A mit rationalen/reellen Eintraegen
RETURN: poly, gibt die Zeilensummennorm von A wieder
EXAMPLE: example zeilensummennorm; zeigt ein Beispiel"
{
  int i,j;
  int n,m = ncols(A),nrows(A);
  poly r,s = 0,0;
  for (i=1;i<=m;i++)
  {
   for (j=1;j<=n;j++)
   {
    r = r + abs(A[i,j]);
   }
   if (r > s)
   {
    s = r;
   }
   r = 0;
  }
  return (s);
}
example
{
  "Beispiel:";
  echo = 2;
  ring r=real,x,lp;
  matrix A[3][2]=-3,-2,-1,3,-4,2;
  print(A);
  zeilensummennorm(A);
  ring r=0,x,lp;
  matrix B[3][2]=-7,0,0,3,-4,2;
```

```
  print(B);
  zeilensummennorm(B);
}
proc maximumsnorm (matrix A)
"USAGE: maximumsnorm(A); matrix A mit rationalen/reellen Eintraegen
RETURN: poly, gibt die Zeilensummennorm von A wieder
EXAMPLE: example maximumsnorm; zeigt ein Beispiel"
{
  int i,j;
  int n,m = ncols(A),nrows(A);
  poly r = 0;
  for (i=1;i<=m;i++)
  {
   for (j=1;j<=n;j++)
   {
    if (abs(A[i,j]) > r)
    {
     r = abs(A[i,j]);
    }
   }
  }
  return(r);
}
example
{
  "Beispiel:";
  echo = 2;
  ring r=real,x,lp;
  matrix A[3][2]=-3,-2,-1,3,-4,2;
  print(A);
  maximumsnorm(A);
  ring r=0,x,lp;
  matrix B[3][2]=-7,0,0,3,-4,2;
  print(B);
  maximumsnorm(B);
}
proc q_eukl_norm (matrix A)
"USAGE: q_eukl_norm(A); matrix A mit rationalen/reellen Eintraegen
RETURN: poly, gibt das Quadrat der euklidischen Norm von A wieder
EXAMPLE: example q_eukl_norm; zeigt ein Beispiel"
{
  int i,j;
  int n,m = ncols(A),nrows(A);
  poly r = 0;
  for (i=1;i<=m;i++)
```

```
  {
   for (j=1;j<=n;j++)
   {
    r = r + abs(A[i,j]) * abs(A[i,j]);
   }
  }
  return (r);
}
example
{
  "Beispiel:";
  echo = 2;
  ring r=real,x,lp;
  matrix A[3][2]=-3,-2,-1,3,-4,2;
  print(A);
  q_eukl_norm(A);
  ring r=0,x,lp;
  matrix B[3][2]=-7,0,0,3,-4,2;
  print(B);
  q_eukl_norm(B);
}
```

Lösung zu Aufgabe 6.6

```
proc ebenenschnitt (matrix E1, matrix E2)
"USAGE: ebenenschnitt(E1,E2); matrix E1, matrix E2 - Koeffizienten von
zwei Ebenengleichungen
RETURN: list, (string,[vc,vt]), Informationen zum Schnitt von E1 und E2,
ggf.
       eine Parametrisierung der Schnittgeraden
EXAMPLE: example ebenenschnitt; zeigt ein Beispiel"
{
  matrix m[2][4] = E1[1,1..4],E2[1,1..4];
  list l;
  matrix A[2][4] = gauss_reduction(m);
  if ((A[2,2] == 0) and (A[2,3] == 0) and (A[2,4] == 0))
  {
   l="Die beiden Ebenen sind gleich!";
   return(l);
  }
  else
  {
   if ((A[2,2] == 0) and (A[2,3] == 0) and (A[2,4] != 0))
   {
    l="Die beiden Ebenen sind parallel!";
    return(l);
   }
   else
```

```
    {
     if (A[2,2] != 0)
     {
      vector vt = [-A[1,3]/A[1,1]+(A[2,3]*A[1,2])/(A[1,1]*A[2,2]),-A[2,3]/A[2,2],1];
      vector vc = [A[1,4]/A[1,1]-A[2,4]/(A[1,1]*A[2,2]),A[2,4]/A[2,2],0];
      l="Der Schnitt der Ebenen ist die Gerade [2] + t * [3]",vc,vt;
     }
     else
     {
      if (A[1,1] != 0)
      {
       vector vt = [-A[1,2]/A[1,1],1,0];
       vector vc = [A[1,4]/A[1,1]-(A[1,3]*A[2,4])/(A[1,1]*A[2,3]),0,A[2,4]/A[2,3]];
       l="Der Schnitt der Ebenen ist die Gerade [2] + t * [3]",vc,vt;
      }
      else
      {
       vector vt = [1,0,0];
       vector vc = [0,A[1,4]/A[1,2]-(A[1,3]*A[2,4])/(A[1,2]*A[2,3]),A[2,4]/A[2,3]];
       l="Der Schnitt der Ebenen ist die Gerade [2] + t * [3]",vc,vt;
      }
     }
    }
    return (l);
  }
}
example
{
  "Beispiel:";
  echo = 2;
  ring r=0,(x,y,z),lp;
  matrix E1[1][4]=1,0,0,0;
  print(E1);
  matrix E2[1][4]=0,1,0,0;
  print(E2);
  list le=ebenenschnitt(E1,E2);
  print(le);
  matrix F1[1][4]=1,2,3,1;
  print(F1);
  matrix F2[1][4]=1,2,0,-1;
  print(F2);
  list lf=ebenenschnitt(F1,F2);
  print(lf);
  matrix I1[1][4]=0,2,3,1;
  print(I1);
  matrix I2[1][4]=0,2,0,-1;
```

```
  print(I2);
  list li=ebenenschnitt(I1,I2);
  print(li);
  matrix G1[1][4]=1,0,0,0;
  print(G1);
  matrix G2[1][4]=1,0,0,1;
  print(G2);
  list lg=ebenenschnitt(G1,G2);
  print(lg);
  matrix H1[1][4]=1,0,0,0;
  print(H1);
  matrix H2[1][4]=2,0,0,0;
  print(H2);
  list lh=ebenenschnitt(H1,H2);
  print(lh);
}
```

Index

Literaturverzeichnis

[Bri83] Egbert Brieskorn, *Lineare Algebra und analytische Geometrie*, Vieweg, 1983.

[Coh96] Henri Cohen, *A course in computational algebraic number theory*, 3 ed., Graduate Texts in Mathematics, no. 138, Springer, 1996.

[Ebb92] Heinz-Dieter Ebbinghaus (ed.), *Zahlen*, 3 ed., Springer, 1992.

[GPS03] Gert-Martin Greuel, Gerhard Pfister, and Hans Schönemann, SINGULAR *Manual, Version 2.0.4*, Fachbereich Mathematik und Zentrum für Computeralgebra, Universität Kaiserslautern, D-67653 Kaiserslautern, May 2003.

[Köc83] Max Köcher, *Lineare Algebra und analytische Geometrie*, Springer, 1983.

[Moo82] Gregory H. Moore, *Zermelo's axiom of choice: Its origins, development and influence*, Studies in the History of Mathematics and Physical Sciences, no. 8, Springer, 1982.

[SS88] Günter Scheja and Uwe Storch, *Lehrbuch der Algebra*, Mathematische Leitfäden, vol. II, Teubner, 1988.

[Sze50] Tibor Szele, *On Zorn's lemma*, Publicationes Mathematicae Debreccen **1** (1949/50), 254–57.