

# **Grundlagen der Mathematik 1: Lineare Algebra**

Andreas Gathmann

Vorlesungsskript RPTU Kaiserslautern 2024/25

# Inhaltsverzeichnis (Grundlagen der Mathematik 1)

0. Einleitung und Motivation . . . . .	3
1. Etwas Logik und Mengenlehre . . . . .	6
1.A Logik 6   1.B Mengenlehre 11	
2. Relationen und Funktionen . . . . .	15
2.A Funktionen 15   2.B Äquivalenzrelationen 21	
3. Erste Eigenschaften der reellen Zahlen . . . . .	24
3.A Gruppen und Körper 24   3.B Vollständige Induktion 30   3.C Polynomfunktionen 31	

## Grundlagen der Mathematik 1: Lineare Algebra

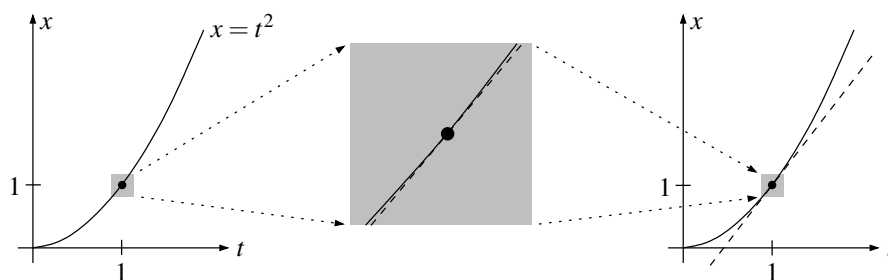
13. Vektorräume . . . . .	35
13.A Der Vektorraumbegriff 35   13.B Untervektorräume 39	
14. Basen und Dimension . . . . .	43
14.A Lineare Unabhängigkeit und Basen 44   14.B Die Dimension von Vektorräumen 49	
15. Lineare Gleichungssysteme und Matrizen . . . . .	55
15.A Matrizen 55   15.B Der Gauß-Algorithmus 62   15.C Algorithmen der linearen Algebra 66	
16. Lineare Abbildungen . . . . .	74
16.A Morphismen von Vektorräumen 74   16.B Die Klassifikation endlich-dimensionaler Vektorräume 79   16.C Abbildungsmatrizen 82	
17. Komplemente und Quotientenräume . . . . .	90
17.A Direkte Summen und Komplemente 90   17.B Quotientenräume 93	
18. Determinanten . . . . .	100
18.A Die Konstruktion der Determinante 100   18.B Eigenschaften der Determinante 107	
Literatur . . . . .	111
Index . . . . .	112

## 0. Einleitung und Motivation

In diesem Skript – so verspricht es der Titel – wollen wir uns die Grundlagen der Mathematik erarbeiten. Aber was ist das überhaupt, die „Grundlagen der Mathematik“? Es handelt sich hierbei um die Kombination zweier Themengebiete, die in der Tat das grundlegende Handwerkszeug für nahezu die gesamte Mathematik darstellen, nämlich

- der *Analysis*, d. h. der Untersuchung von Folgen und Grenzwerten, Stetigkeit, sowie der Differential- und Integralrechnung (zunächst in einer und im zweiten Semester dann auch in mehreren Variablen), und
- der *linearen Algebra*, d. h. der Theorie der Vektorräume, linearen Abbildungen und Gleichungssysteme.

Von beiden Gebieten habt ihr ja aus der Schule wahrscheinlich schon eine ungefähre Vorstellung. In der *Analysis* geht es grob gesagt darum, reelle Funktionen *lokal*, also in der Umgebung eines gewählten Punktes, zu untersuchen, und aus diesen Untersuchungen dann wieder Aussagen über die gesamte Funktion zurückzugewinnen. Betrachten wir z. B. ein Auto, das sich entlang einer geraden Strecke bewegt. Der Einfachheit halber nehmen wir an, dass die Position  $x$  des Autos nach der Zeit  $t$  (in geeigneten Einheiten) durch die Gleichung  $x = t^2$  beschrieben werden kann, so dass die Bewegung durch die Kurve im folgenden Bild links dargestellt wird:



Wir wollen diese Bewegung nun nur in einer kleinen Umgebung eines fest gewählten Zeitpunkts, z. B. des Zeitpunkts  $t = 1$  (und damit auch  $x = 1$ ) betrachten. Im Bild oben haben wir diese Umgebung grau markiert und in der Mitte stark vergrößert dargestellt. Wir sehen, dass die Kurve in dieser Umgebung fast wie eine *Gerade* aussieht; wir haben diese Gerade gestrichelt eingezeichnet und im Bild rechts auch außerhalb der gewählten Umgebung fortgesetzt. Geometrisch ist diese Gerade natürlich einfach die *Tangente* an die Kurve an der Stelle  $t = 1$ . Physikalisch repräsentiert die Steigung dieser Geraden die *Geschwindigkeit* des Autos zum betrachteten Zeitpunkt, denn sie gibt ja gerade an, in welchem Verhältnis sich dort die Strecke  $x$  mit der Zeit  $t$  verändert. Aus der Schule wisst ihr auch schon, wie man diese Steigung ausrechnet: Man muss dazu die gegebene Funktion  $t^2$  *differenzieren* – so dass man die Ableitung  $2t$  erhält – und dort die betrachtete Stelle  $t = 1$  einsetzen. Die Steigung ist in unserem Fall also gerade  $2 \cdot 1 = 2$ , und man rechnet sofort nach, dass  $x = 2t - 1$  die Gleichung der oben eingezeichneten Tangente ist. Man sagt, dass die Gerade  $x = 2t - 1$  eine *lineare Approximation* der ursprünglich gegebenen Funktion  $x = t^2$  im Punkt  $t = 1$  ist.

Wir können aus der Kenntnis der zurückgelegten Strecke zu jedem Zeitpunkt also die Geschwindigkeit des Autos durch Differenzieren bestimmen. Man kann sich natürlich auch die umgekehrte Frage stellen: Angenommen, der Kilometerzähler eures Autos ist kaputt, aber ihr beobachtet auf eurer Fahrt ständig eure Geschwindigkeit. Könnt ihr dann am Ende der Fahrt trotzdem ausrechnen, wie weit ihr gefahren seid? Dies ist offensichtlich die „Umkehrung“ des Differenzierens – und auch hier wisst ihr aus der Schule natürlich schon, dass dies auf die *Integralrechnung* führen wird.

In der Praxis bewegt man sich mit dem Auto aber nicht immer nur auf einer geraden Strecke, und demzufolge braucht man für die Beschreibung solch einer Bewegung (und natürlich auch vieler anderer natürlich auftretender Prozesse) mehrere Variablen. Wir werden daher auch Abbildungen betrachten, die mehrere Variablen auf mehrere andere abbilden, wie z. B. die folgende Vorschrift, die zwei reelle Zahlen  $y_1$  und  $y_2$  in Abhängigkeit von zwei anderen  $x_1$  und  $x_2$  ausdrückt:

$$\begin{aligned}y_1 &= 2x_1 + x_2 \cos x_1 \\y_2 &= x_1 e^{x_2} - x_2\end{aligned}$$

Genau wie oben werden wir uns auch hier wieder die Frage stellen, ob wir diese (in diesem Fall recht komplizierte) Funktion in der Nähe eines gegebenen Punktes nicht vielleicht durch eine einfache *lineare* Abhängigkeit annähern können. In der Tat ist dies möglich: Wir werden sehen, dass z. B. in einer kleinen Umgebung des Nullpunkts  $(x_1, x_2) = (0, 0)$  die obige Vorschrift näherungsweise die gleichen Ergebnisse liefert wie das lineare Gleichungssystem

$$\begin{aligned}y_1 &= 2x_1 + x_2 \\y_2 &= x_1 - x_2.\end{aligned}$$

Auch wenn diese Gleichungen natürlich viel einfacher als die ursprünglichen sind, sollte offensichtlich sein, dass auch solche linearen Gleichungssysteme bei wachsender Zahl von Variablen (und in der Praxis sind Hunderte oder Tausende von Variablen keine Seltenheit) recht kompliziert werden können. Wir werden daher einen wesentlichen Teil dieser Vorlesung mit der *linearen Algebra*, also dem Studium derartiger linearer Gleichungssysteme, verbringen. Um dabei überhaupt erst einmal den Notationsaufwand in Grenzen zu halten, tut man dabei gut daran, die Start- und Zielvariablen nicht alle einzeln hinzuschreiben, sondern sie zu sogenannten *Vektoren* zusammenzufassen. In der Tat kann man in dieser Sichtweise die lineare Algebra als das Studium von linearen Abbildungen zwischen Vektorräumen beschreiben.

Wenn wir dann die linearen Abbildungen zwischen mehreren Variablen gut genug verstanden haben, können wir uns im zweiten Teil der Analysis schließlich daran machen, die Differential- und Integralrechnung auf den Fall von mehreren Variablen auszuweiten.

Wir haben damit jetzt relativ kurz umrissen, welcher mathematische Stoff uns in dieser Vorlesung erwartet. Es wird in diesem Skript aber nicht nur darum gehen, mathematische Resultate kennenzulernen. Mindestens ebenso wichtig ist es, das „mathematische Denken“ zu lernen, d. h. die Fähigkeit zu entwickeln, mit abstrakten Konzepten umzugehen, exakte logische Schlüsse zu ziehen und Beweise zu führen. Anders als in der Schule oder im Studium z. B. ingenieurwissenschaftlicher Fächer werden wir genau darauf achten, eine „wasserdichte“ Theorie aufzubauen: Jeder neue Begriff bzw. jeder neue Satz wird nur unter Verwendung des bisher Bekannten exakt definiert bzw. bewiesen. So sind z. B. Formulierungen in dem Stil „eine Funktion wird durch eine Gerade angenähert“ oder „eine Funktion heißt stetig, wenn man sie zeichnen kann, ohne den Stift abzusetzen“ als Veranschaulichung unserer Ideen zwar sehr sinnvoll, als exakte mathematische Formulierung jedoch schlichtweg unbrauchbar. Wir werden daher in dieser Vorlesung viel exakter arbeiten als ihr es wahrscheinlich aus der Schule gewohnt seid, und es ist wichtig, dass ihr diese exakte Denkweise verinnerlicht und anzuwenden lernt. Die Mathematik ist ein riesiges Gebäude – viel größer und komplexer als ihr es euch wahrscheinlich im Moment vorstellen könnt – das ständig höher gebaut wird, indem schon bewiesene Sätze auf neue Fälle angewendet oder wieder für neue Beweise verwendet werden. Wir starten gerade beim Fundament dieses Gebäudes und können es uns da wirklich nicht leisten herumzupfuschen.

Diese konsequent logische und exakte Herangehensweise ist zwar am Anfang wahrscheinlich ungewohnt, hat jedoch für euch auch einen Vorteil: Etwas überspitzt formuliert erzählen wir euch während des gesamten Studiums eigentlich nur Dinge, die logisch aus dem folgen, was ihr ohnehin schon wusstet. Dadurch ist die Mathematik wahrscheinlich das Studienfach, in dem man am wenigsten auswendig lernen muss – in dem es aber im Gegenzug auch am meisten auf das *Verständnis* des Stoffes ankommt. Je besser euer Verständnis für die Mathematik wird, um so mehr Dinge werden euch letztlich einfach „klar“ werden, so dass es euch dann auch viel leichter fällt, sie zu lernen.

Dieses Verständnis für die Mathematik bekommt man aber natürlich nur durch intensiven und *aktiven* Umgang mit dem Stoff, weswegen neben dem Studium der Vorlesung auch die Bearbeitung der Übungsaufgaben besonders wichtig ist.

Heißt das alles nun, dass wir nur mit logischen Argumenten die Mathematik sozusagen „aus dem Nichts“ aufbauen können? Nein, das geht natürlich nicht ... von nichts kommt nichts. Man muss am Anfang immer gewisse Dinge als gegeben annehmen, also Aussagen als wahr voraussetzen, die man nicht mehr beweist bzw. beweisen kann, und auf denen dann die gesamte Theorie beruht. Derartige Annahmen bezeichnet man als **Axiome**. Natürlich versucht man in der Mathematik, mit möglichst wenigen und sehr elementaren Axiomen auszukommen, die hoffentlich niemand anzweifeln würde. In der modernen Mathematik ist es üblich, hierfür die grundlegenden Prinzipien der Logik und Mengenlehre zu verwenden (zu denen wir auch gleich in Kapitel 1 Genaueres sagen werden).

In dieser Vorlesung wollen wir uns das Leben allerdings etwas leichter machen und zusätzlich auch die *Existenz und elementaren Eigenschaften der reellen Zahlen* axiomatisch voraussetzen (um welche Eigenschaften es sich hierbei handelt, werden wir natürlich genau angeben). Man kann zwar nur aus den Axiomen der Logik und Mengenlehre beweisen, dass die reellen Zahlen existieren und dass sie die erwarteten Eigenschaften haben, der Beweis wäre zu diesem frühen Zeitpunkt im Studium aber sehr verwirrend und würde euch auch keine großartigen neuen Erkenntnisse bringen. In diesem Sinne starten wir also sozusagen doch nicht ganz beim Fundament unseres „Gebäudes Mathematik“, sondern bereits im ersten Stock.

Im weiteren Verlauf ist dieses Skript dann wie im Inhaltsverzeichnis angegeben in mehrere Teile gegliedert. Diese bauen der Reihe nach aufeinander auf, mit einer Ausnahme: Nach den in jedem Fall benötigten grundlegenden Anfangskapiteln 1 bis 3 sind die Teile „Grundlagen der Mathematik 1: Analysis“ (Kapitel ?? bis ??) und „Grundlagen der Mathematik 1: Lineare Algebra“ (Kapitel 13 bis 18) *unabhängig voneinander* und können somit in beliebiger Reihenfolge oder auch parallel studiert werden.

Aber jetzt genug der Vorrede ... beginnen wir nun also unser Studium der Mathematik mit den „Grundlagen der Grundlagen“, den für uns wesentlichen Prinzipien der Logik und Mengenlehre.

## 1. Etwas Logik und Mengenlehre

Bevor wir mit dem eigentlichen Inhalt der Vorlesung beginnen, müssen wir in diesem Kapitel kurz die exakte mathematische Sprache beschreiben, in der wir unsere Ergebnisse formulieren werden: die der Logik und Mengenlehre. Zentral hierbei sind die Begriffe der *Aussage* (in der Logik) und der *Menge* (in der Mengenlehre).

Da wir es hier mit den ersten beiden Begriffen überhaupt zu tun haben, die in der Mathematik vorkommen, können wir sie natürlich nicht durch bereits bekannte Dinge definieren oder mit bereits bekannten Resultaten ihre Eigenschaften herleiten. Wir müssen sie daher (wie schon in der Einleitung erwähnt) axiomatisch voraussetzen. Wir müssen *voraussetzen*, dass es sinnvoll ist, über logische Aussagen und deren Wahrheit zu reden, dass Mengen überhaupt existieren, dass man Mengen vereinigen und schneiden kann, aus ihnen Elemente auswählen kann, und noch einiges mehr. Wenn ihr euch zum Beispiel auf den Standpunkt stellt, dass ihr nicht an die Existenz von Mengen glaubt, wird euch niemand widerlegen können. Allerdings zweifelt ihr damit dann auch die Existenz der gesamten Mathematik an, wie sie heutzutage betrieben wird – und aus der Tatsache, dass ihr in dieser Vorlesung sitzt, schließe ich einmal, dass das nicht der Fall ist.

Glücklicherweise sind die Dinge, die wir benötigen, jedoch allesamt anschaulich sofort einleuchtend. Ich möchte es euch (und mir) daher ersparen, an dieser Stelle eine vollständige und präzise axiomatische Formulierung der Logik und Mengenlehre hinzuschreiben, zumal das momentan sicher mehr verwirren als helfen würde und außerdem gerade im Bereich der Logik auch zu sehr in die Philosophie abdriften würde. Stattdessen wollen wir uns in diesem (für den Rest der Vorlesung sehr untypischen) ersten Kapitel damit begnügen, die für uns wichtigsten Prinzipien und Notationen sowie beliebte Fehlerquellen in verständlicher Sprache zu erklären, auch wenn ein paar Dinge (insbesondere die Begriffsfestlegung – „Definition“ möchte ich es eigentlich gar nicht nennen – einer Aussage und einer Menge) dadurch recht schwammig klingen werden. Außerdem werden wir in Beispielen zur besseren Verdeutlichung bereits hier die reellen Zahlen und ihre einfachsten Eigenschaften (die euch sicherlich bekannt sein werden) benutzen, auch wenn wir diese erst später formalisieren werden. Da es sicher niemanden von euch verwirren wird, werden wir auch die Schreibweise „ $x \in \mathbb{R}$ “ für „ $x$  ist eine reelle Zahl“ schon verwenden, bevor sie in den Notationen 1.12 und 1.14 offiziell eingeführt wird. Ab Kapitel 2 werden wir dann mit dem strukturierten Aufbau der Grundlagen der Mathematik beginnen, also alle Definitionen mit bereits eingeführten Notationen präzise formulieren und alle Aussagen aus vorherigen exakt beweisen.

### 1.A Logik

Beginnen wir also mit der Logik. Unter einer **Aussage** verstehen wir (grob gesagt) ein sprachliches Gebilde, das entweder wahr oder falsch ist (wobei wir in der Mathematik natürlich letztlich daran interessiert sind, *wahre* Aussagen herzuleiten – wenn wir später mathematische Sätze formulieren, ist damit also stets gemeint, dass die dort gemachte Aussage wahr ist). Wichtig sind auch sprachliche Gebilde, in denen freie **Variablen**, also Platzhalter, vorkommen, und die erst beim Einsetzen von Werten für diese Variablen Aussagen liefern. Man bezeichnet sie als **Aussageformen**.

#### Beispiel 1.1.

- (a)  $1 + 1 = 2$  ist eine wahre,  $1 + 1 = 3$  eine falsche, und  $1 + 1$  überhaupt keine Aussage.
- (b)  $x + 1 = 2$  ist eine Aussageform, die beim Einsetzen von  $x = 1$  in eine wahre, beim Einsetzen jeder anderen reellen Zahl in eine falsche Aussage übergeht.

**Bemerkung 1.2.** Als Variablen in Aussageformen kann man beliebige Symbole benutzen. Üblich sind neben den normalen lateinischen Klein- und Großbuchstaben auch die griechischen Buchstaben, die wir zur Erinnerung hier auflisten:

A $\alpha$ alpha	B $\beta$ beta	$\Gamma$ $\gamma$ gamma	$\Delta$ $\delta$ delta	E $\varepsilon$ epsilon	Z $\zeta$ zeta	H $\eta$ eta	$\Theta$ $\vartheta$ theta
I $\iota$ iota	K $\kappa$ kappa	$\Lambda$ $\lambda$ lambda	M $\mu$ my	N $\nu$ ny	$\Xi$ $\xi$ xi	O $o$ omikron	$\Pi$ $\pi$ pi
P $\rho$ rho	$\Sigma$ $\sigma$ sigma	T $\tau$ tau	Y $\upsilon$ ypsilon	$\Phi$ $\varphi$ phi	X $\chi$ chi	$\Psi$ $\psi$ psi	$\Omega$ $\omega$ omega

Oft verziert man Buchstaben auch noch mit einem Symbol oder versieht sie mit einem Index, um neue Variablen zu erhalten: So sind z. B.  $x, x', \bar{x}, \bar{\bar{x}}, x_1, x_2, \dots$  alles Symbole für verschiedene Variablen, die zunächst einmal nichts miteinander zu tun haben (aber möglichst für irgendwie miteinander zusammenhängende Objekte eingesetzt werden sollten, wenn man den Leser nicht vollends verwirren will).

**Notation 1.3** (Zusammengesetzte Aussagen). Sind  $A$  und  $B$  Aussagen, so lassen sich daraus wie folgt neue bilden:

Symbol	Wahrheitstafel				Bedeutung
$A$	w	f	w	f	
$B$	w	w	f	f	
$\neg A$	f	w			nicht $A$
$A \wedge B$	w	f	f	f	$A$ und $B$
$A \vee B$	w	w	w	f	$A$ oder $B$ (oder beides): „nicht-ausschließendes Oder“
$A \Leftrightarrow B$	w	f	f	w	$A$ und $B$ sind gleichbedeutend / äquivalent, bzw. $A$ genau dann, wenn $B$
$A \Rightarrow B$	w	w	f	w	aus $A$ folgt $B$ , bzw. wenn $A$ dann $B$

Die sogenannte **Wahrheitstafel** in den mittleren vier Spalten ist dabei die eigentliche Definition der neuen zusammengesetzten Aussagen. Sie gibt in Abhängigkeit der Wahrheit von  $A$  und  $B$  (in den ersten beiden Zeilen) an, ob die zusammengesetzte Aussage wahr oder falsch ist.

Bemerkenswert ist hierbei wohl nur die Folgerungsaussage  $A \Rightarrow B$ , die keine Aussage über die Richtigkeit von  $A$  oder  $B$  separat macht, sondern nur sagt, dass  $B$  wahr ist, wenn auch  $A$  es ist. Ist hingegen  $A$  falsch, so ist die Folgerungsaussage  $A \Rightarrow B$  stets wahr („aus einer falschen Voraussetzung kann man alles folgern“). So ist z. B.  $0 = 1 \Rightarrow 2 = 3$  eine wahre Aussage. Wie schon erwähnt wollen wir uns in der Mathematik aber natürlich in der Regel mit wahren Aussagen beschäftigen, und neue wahre Aussagen aus alten herleiten. Gerade in Beweisen ist die übliche Verwendung der Notation  $A \Rightarrow B$  daher, dass  $A$  eine bereits als wahr erkannte Aussage ist, und wir damit nun schließen wollen, dass auch  $B$  wahr ist.

**Bemerkung 1.4** (Beweise mit Wahrheitstafeln). Wollen wir kompliziertere zusammengesetzte Aussagen miteinander vergleichen, so können wir dies auch mit Hilfe von Wahrheitstafeln tun. So ist für zwei Aussagen  $A$  und  $B$  z. B.

$$A \Rightarrow B \text{ äquivalent zu } (\neg A) \vee B,$$

denn wenn wir in der Wahrheitstafel

$A$	w	f	w	f
$B$	w	w	f	f
$\neg A$	f	w	f	w
$(\neg A) \vee B$	w	w	f	w

mit Hilfe der Definitionen von  $\neg$  und  $\vee$  aus Notation 1.3 zunächst  $\neg A$  und dann  $(\neg A) \vee B$  berechnen, sehen wir, dass das Ergebnis mit  $A \Rightarrow B$  übereinstimmt. Nach der Bemerkung aus Notation 1.3 ist dies auch anschaulich klar: Die Folgerungsaussage  $A \Rightarrow B$  ist ja genau dann wahr, wenn  $A$  falsch (also  $\neg A$  wahr) ist, oder wenn  $B$  wahr ist (oder beides).

Genauso zeigt man die ebenfalls einleuchtende Aussage, dass

$$A \Leftrightarrow B \text{ äquivalent zu } (A \Rightarrow B) \wedge (B \Rightarrow A)$$

ist – was auch die übliche Art ist, wie man eine Äquivalenz zeigt: Man zeigt separat die beiden Folgerungen  $A \Rightarrow B$  und  $B \Rightarrow A$ .

**Notation 1.5.** Folgerungen („ $\Rightarrow$ “) und Äquivalenzen („ $\Leftrightarrow$ “) sind natürlich zwei verschiedene Dinge, die man nicht durcheinanderwerfen darf (auch wenn das in der Schule wahrscheinlich manchmal nicht so genau genommen wird). Es hat sich jedoch in der Mathematik eingebürgert, bei *Definitionen* von Begriffen durch eine äquivalente, definierende Eigenschaft die Sprechweise „wenn“ anstatt des eigentlich korrekten „genau dann, wenn“ zu verwenden: So würde man z. B. als Definition des Begriffs einer geraden Zahl hinschreiben

„Eine ganze Zahl  $x$  heißt gerade, wenn  $\frac{x}{2}$  eine ganze Zahl ist“,

obwohl man genau genommen natürlich meint

„Eine ganze Zahl  $x$  heißt *genau dann* gerade, wenn  $\frac{x}{2}$  eine ganze Zahl ist“.

Eine gewöhnliche Folgerungsaussage wie z. B. die wahre Aussage

„Wenn eine ganze Zahl  $x$  positiv ist, dann ist auch  $x + 1$  positiv“

ist dagegen immer nur in einer Richtung zu verstehen; hier wird also nicht behauptet, dass mit  $x + 1$  auch  $x$  immer positiv sein muss (was ja auch falsch wäre).

**Notation 1.6** (Quantoren). Natürlich kann man nicht nur zwei, sondern auch mehrere Aussagen wie in Notation 1.3 mit „und“ bzw. „oder“ verknüpfen – also die neue Aussage konstruieren, dass *jede* bzw. *mindestens eine* der ursprünglichen Aussagen wahr ist. Am einfachsten notiert man dies mit einer Aussageform  $A$ , in der eine freie Variable  $x$  vorkommt. Wir schreiben dies dann auch als  $A(x)$  und setzen

Symbol	Bedeutung
$\forall x: A(x)$	für alle $x$ gilt $A(x)$ (also eine „und“-Verknüpfung aller Aussagen $A(x)$ )
$\exists x: A(x)$	es gibt ein $x$ mit $A(x)$ (also eine „oder“-Verknüpfung aller Aussagen $A(x)$ )

Die beiden Symbole  $\forall$  und  $\exists$  bezeichnet man als **Quantoren**. Beachte, dass diese beiden Quantoren *nicht* miteinander vertauschbar sind: So besagt z. B. die Aussage

$$\forall x \in \mathbb{R} \exists y \in \mathbb{R}: y > x$$

„zu jeder reellen Zahl  $x$  gibt es eine Zahl  $y$ , die größer ist“ (was offensichtlich wahr ist), während die Umkehrung der beiden Quantoren die Aussage

$$\exists y \in \mathbb{R} \forall x \in \mathbb{R}: y > x$$

„es gibt eine reelle Zahl  $y$ , die größer als jede reelle Zahl  $x$  ist“ liefern würde (was ebenso offensichtlich falsch ist). Der Unterschied besteht einfach darin, dass im ersten Fall zuerst das  $x$  gewählt werden muss und dann ein  $y$  dazu existieren muss (das von  $x$  abhängen darf), während es im zweiten Fall *dasselbe*  $y$  für alle  $x$  sein müsste.

**Bemerkung 1.7.** Jede Aussage lässt sich natürlich auf viele Arten aufschreiben, sowohl als deutscher Satz als auch als mathematische Formel. Die gerade eben betrachtete Aussage könnte man z. B. auf die folgenden (absolut gleichwertigen) Arten aufschreiben:

- (a)  $\forall x \in \mathbb{R} \exists y \in \mathbb{R}: y > x$ .
- (b) Es sei  $x \in \mathbb{R}$ . Dann gibt es ein  $y \in \mathbb{R}$  mit  $y > x$ .
- (c) Zu jeder reellen Zahl gibt es noch eine größere.

Welche Variante man beim Aufschreiben wählt, ist weitestgehend Geschmackssache. Die Formulierung einer Aussage als deutscher Satz hat den Vorteil, dass wir sie oft leichter verstehen können, weil wir die deutsche Sprache schon länger kennen als die mathematische. Wenn wir uns jedoch erst einmal an die mathematische Sprache gewöhnt haben, wird auch sie ihre Vorzüge bekommen:

Sie ist deutlich kürzer und besser logisch strukturiert. Wir werden im Folgenden beide Schreibweisen mischen und jeweils diejenige wählen, mit der unsere Aussagen (hoffentlich) am einfachsten verständlich werden.

Wenn wir mathematische Symbole verwenden, müssen wir diese aber auch stets in ihrer korrekten Notation und nicht als „Abkürzungen“ für deutsche Wörter verwenden: Man würde die Aussage „2 und 4 sind gerade Zahlen“ sicher niemals schreiben als „ $2 \wedge 4$  sind gerade Zahlen“, und analog genauso wenig „Es gilt  $x^2 \geq 0$  für alle  $x \in \mathbb{R}$ “ als „Es gilt  $x^2 \geq 0 \forall x \in \mathbb{R}$ “.

**Bemerkung 1.8** (Negationen). Es ist wichtig zu wissen, wie man von einer Aussage das Gegenteil, also die Negation bzw. „Verneinung“ bildet. Da hierbei oft Fehler gemacht werden, wollen wir die allgemeinen Regeln hierfür im Folgenden kurz auflisten. Dabei könnte man (a), (b) und (c) wieder schnell mit Wahrheitstabellen zeigen; die Regeln (d) und (e) sind wie in Notation 1.6 analog zu (b) und (c) für die Verknüpfung mehrerer Aussagen:

- (a)  $\neg(\neg A) \Leftrightarrow A$ : Ist es falsch, dass  $A$  falsch ist, so bedeutet dies genau, dass  $A$  wahr ist.
- (b)  $\neg(A \wedge B) \Leftrightarrow (\neg A) \vee (\neg B)$ : Das Gegenteil von „ $A$  und  $B$  sind richtig“ ist „ $A$  oder  $B$  ist falsch“.
- (c)  $\neg(A \vee B) \Leftrightarrow (\neg A) \wedge (\neg B)$ : Das Gegenteil von „ $A$  oder  $B$  ist richtig“ ist „ $A$  und  $B$  sind falsch“.
- (d)  $\neg(\forall x: A(x)) \Leftrightarrow \exists x: \neg A(x)$ : Das Gegenteil von „für alle  $x$  gilt  $A(x)$ “ ist „es gibt ein  $x$ , für das  $A(x)$  falsch ist“.
- (e)  $\neg(\exists x: A(x)) \Leftrightarrow \forall x: \neg A(x)$ : Das Gegenteil von „es gibt ein  $x$ , für das  $A(x)$  gilt“ ist „für alle  $x$  ist  $A(x)$  falsch“.

Man kann also sagen, dass eine Verneinung dazu führt, dass „und“ mit „oder“ sowie „für alle“ mit „es gibt“ vertauscht werden. So ist z. B. das Gegenteil der Aussage

„In Frankfurt haben *alle* Haushalte Strom *und* fließendes Wasser“

die Aussage

„In Frankfurt *gibt* es einen Haushalt, der keinen Strom *oder* kein fließendes Wasser hat“.

### Beispiel 1.9.

- (a) Wollen wir eine Folgerung  $A \Rightarrow B$  verneinen, so können wir sie zunächst mit Bemerkung 1.4 zu  $(\neg A) \vee B$  umformen, und erhalten nach Bemerkung 1.8 als Umkehrung dann  $A \wedge \neg B$ . Dies ist auch anschaulich einleuchtend: Die Folgerungsaussage „wenn  $A$  dann  $B$ “ ist genau dann falsch, wenn die Voraussetzung  $A$  zwar gilt, die Behauptung  $B$  aber nicht. Wir sehen also:

Die Verneinung einer Folgerung  $A \Rightarrow B$  ist  $A \wedge \neg B$

(und nicht etwa  $A \Rightarrow \neg B$ , wie man vielleicht denken könnte).

- (b) Eine oft vorkommende Anwendung der Regeln für die Verneinung von Aussagen ist der sogenannte **Widerspruchsbeweis** bzw. Beweis durch **Kontraposition**. Nach Bemerkung 1.4 gesehen ist die Folgerung  $A \Rightarrow B$  („aus  $A$  folgt  $B$ “) gleichbedeutend mit  $(\neg A) \vee B$ . Damit ist diese Aussage nach Bemerkung 1.8 (a) auch äquivalent zu  $(\neg(\neg B)) \vee (\neg A)$ , also zu  $\neg B \Rightarrow \neg A$ . Mit anderen Worten: Haben wir eine Schlussfolgerung  $A \Rightarrow B$  zu beweisen, so können wir genauso gut  $(\neg B) \Rightarrow (\neg A)$  zeigen, d. h. *wir können annehmen, dass die zu zeigende Aussage  $B$  falsch ist und dies dann zu einem Widerspruch führen bzw. zeigen, dass dann auch die Voraussetzung  $A$  falsch sein muss*.

**Beispiel 1.10.** Hier sind zwei Beispiele für die Anwendung der Prinzipien aus Bemerkung 1.8 und Beispiel 1.9 – und auch unsere ersten Beispiele dafür, wie man Beweise von Aussagen exakt aufschreiben kann.

- (a) Einen Beweis durch Widerspruch könnte man z. B. so aufschreiben:

*Behauptung:* Für alle  $x \in \mathbb{R}$  gilt  $2x + 1 > 0$  oder  $2x - 1 < 0$ .

*Beweis:* Angenommen, die Behauptung wäre falsch, d. h. (nach Bemerkung 1.8 (c) und (d)) es gäbe ein  $x \in \mathbb{R}$  mit

$$2x + 1 \leq 0 \quad (1) \quad \text{und} \quad 2x - 1 \geq 0 \quad (2).$$

Für dieses  $x$  würde dann folgen, dass

$$0 \stackrel{(1)}{\geq} 2x + 1 = 2x - 1 + 2 \stackrel{(2)}{\geq} 0 + 2 = 2.$$

Dies ist aber ein Widerspruch. Also war unsere Annahme falsch und somit die zu beweisende Aussage richtig.  $\square$

Das dabei verwendete Symbol „ $\square$ “ ist die übliche Art, das Ende eines Beweises zu kennzeichnen. Zur Verdeutlichung haben wir die beiden Ungleichungen mit (1) und (2) markiert, um später angeben zu können, wo sie verwendet werden.

- (b) Manchmal weiß man von einer Aussage aufgrund der Aufgabenstellung zunächst einmal noch nicht, ob sie wahr oder falsch ist. In diesem Fall muss man sich dies natürlich zuerst überlegen – und, falls die Aussage falsch ist, ihre Negation beweisen. Als Beispiel dafür betrachten wir die Aufgabe

Man beweise oder widerlege: Für alle  $x \in \mathbb{R}$  gilt  $2x + 1 < 0$  oder  $2x - 1 > 0$ .

In diesem Fall merkt man schnell, dass die Aussage falsch sein muss, weil die Ungleichungen schon für den Fall  $x = 0$  nicht stimmen. Man könnte als Lösung der Aufgabe unter Beachtung der Negationsregeln aus Bemerkung 1.8 also aufschreiben:

*Behauptung:* Die Aussage ist falsch, d. h. es gibt ein  $x \in \mathbb{R}$  mit  $2x + 1 \geq 0$  und  $2x - 1 \leq 0$ .

*Beweis:* Für  $x = 0$  ist  $2x + 1 = 1 \geq 0$  und  $2x - 1 = -1 \leq 0$ .  $\square$

Beachte, dass dies ein vollständiger Beweis ist: *Um eine allgemeine Aussage zu widerlegen, genügt es, ein Gegenbeispiel dafür anzugeben.*

**Bemerkung 1.11.** Bevor wir unsere kurze Auflistung der für uns wichtigen Prinzipien der Logik beenden, wollen wir noch kurz auf ein paar generelle Dinge eingehen, die man beim Aufschreiben mathematischer Beweise oder Rechnungen beachten muss.

Dass wir bei unseren logischen Argumenten sauber und exakt arbeiten – also z. B. nicht Folgerungen, die keine Äquivalenzen sind, in der falschen Richtung verwenden, „für alle“ mit „es gibt“ verwechseln oder ähnliches – sollte sich von selbst verstehen. Die folgende kleine Geschichte hilft vielleicht zu verstehen, was damit gemeint ist.

Ein Ingenieur, ein Physiker und ein Mathematiker fahren mit dem Zug nach Frankreich und sehen dort aus dem Fenster des Zuges ein schwarzes Schaf.

Da sagt der Ingenieur: „Oh, in Frankreich sind die Schafe schwarz!“

Darauf der Physiker: „Nein ... wir wissen jetzt nur, dass es in Frankreich mindestens ein schwarzes Schaf gibt.“

Der Mathematiker: „Nein ... wir wissen nur, dass es in Frankreich mindestens ein Schaf gibt, das auf mindestens einer Seite schwarz ist.“

Es gibt aber noch einen weiteren sehr wichtigen Punkt, der leider oft nicht beachtet wird: In der Regel werden wir beim Aufschreiben sowohl Aussagen notieren wollen, die wir erst noch zeigen wollen (um schon einmal zu sagen, worauf wir hinaus wollen), als auch solche, von denen wir bereits wissen, dass sie wahr sind (z. B. weil sie für die zu zeigende Behauptung als wahr vorausgesetzt werden oder weil sie sich logisch aus irgendetwas bereits Bekanntem ergeben haben). Es sollte offensichtlich sein, dass wir Aussagen mit derartig verschiedenen Bedeutungen für die Argumentationsstruktur nicht einfach zusammenhangslos hintereinander schreiben dürfen, wenn noch jemand in der Lage sein soll, die Argumente nachzuvollziehen. Betrachten wir z. B. noch einmal unseren

Beweis aus Beispiel 1.10 (a) oben, so wäre eine Art des Aufschreibens in folgendem Stil (wie man es leider oft sieht)

$$\begin{aligned} 2x + 1 > 0 \text{ oder } 2x - 1 < 0 \\ 2x + 1 \leq 0 \quad 2x - 1 \geq 0 \\ 0 \geq 2x + 1 = 2x - 1 + 2 \geq 0 + 2 = 2 \end{aligned}$$

völlig inakzeptabel, obwohl hier natürlich letztlich die gleichen Aussagen stehen wie oben. Kurz gesagt:

Von jeder aufgeschriebenen Aussage muss für den Leser *sofort* und *ohne eigenes Nachdenken* ersichtlich sein, welche Rolle sie in der Argumentationsstruktur spielt: Ist es z. B. eine noch zu zeigende Behauptung, eine Annahme oder eine Folgerung (und wenn ja, aus was)?

Dies bedeutet allerdings nicht, dass wir ganze Aufsätze schreiben müssen. Eine (schon recht platzoptimierte) Art, den Beweis aus Beispiel 1.10 (a) aufzuschreiben, wäre z. B.

Angenommen, es gäbe ein  $x \in \mathbb{R}$  mit  $2x + 1 \leq 0$  und  $2x - 1 \geq 0$ .

Dann wäre  $0 \geq 2x + 1 = 2x - 1 + 2 \geq 0 + 2 = 2$ , Widerspruch. □

01

## 1.B Mengenlehre

Nachdem wir die wichtigsten Regeln der Logik behandelt haben, wenden wir uns jetzt der Mengenlehre zu. Die gesamte moderne Mathematik basiert auf diesem Begriff der Menge, der ja auch schon aus der Schule hinlänglich bekannt ist. Zur Beschreibung, was eine Menge ist, zitiert man üblicherweise die folgende Charakterisierung von Georg Cantor (1845–1918):

„Eine **Menge** ist eine Zusammenfassung von bestimmten, wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens zu einem Ganzen.“

Die in einer Menge zusammengefassten Objekte bezeichnet man als ihre **Elemente**.

### Notation 1.12.

- (a) Wir schreiben  $x \in M$ , falls  $x$  ein Element der Menge  $M$  ist, und  $x \notin M$  andernfalls.
- (b) Die einfachste Art, eine Menge konkret anzugeben, besteht darin, ihre Elemente in geschweiften Klammern aufzulisten, wobei es auf die Reihenfolge und Mehrfachnennungen nicht ankommt. So sind z. B.  $\{1, 2, 3\}$  und  $\{2, 3, 1, 3\}$  zwei Schreibweisen für dieselbe Menge mit den drei Elementen 1, 2 und 3.  
Beachte, dass die Elemente einer Menge nicht unbedingt Zahlen sein müssen – so ist z. B.  $M = \{\{2, 3\}, \{1, 3\}\}$  eine Menge mit zwei Elementen, die selbst wieder Mengen sind, nämlich  $\{2, 3\}$  und  $\{1, 3\}$ . Mit der Notation aus (a) ist also z. B.  $\{1, 3\} \in M$ . Insbesondere ist  $M$  nicht dasselbe wie die Menge  $\{1, 2, 3\}$ .
- (c) Man kann die Elemente einer Menge auch durch eine beschreibende Eigenschaft angeben:  $\{x : A(x)\}$  bezeichnet die Menge aller Objekte  $x$ , für die die Aussage  $A(x)$  wahr ist, wie z. B. in  $\{x \in \mathbb{R} : x^2 = 1\} = \{-1, 1\}$ .
- (d) Die Menge  $\{\}$  ohne Elemente, die sogenannte **leere Menge**, bezeichnen wir mit  $\emptyset$ .
- (e) Eine Menge  $M$  heißt **Teilmenge** einer Menge  $N$  (geschrieben  $M \subset N$ ), wenn jedes Element von  $M$  auch Element von  $N$  ist, bzw. in der Quantorenschreibweise von Notation 1.6 wenn

$$\forall x: x \in M \Rightarrow x \in N.$$

Man sagt in diesem Fall auch, dass  $N$  eine **Obermenge** von  $M$  ist (geschrieben  $N \supset M$ ).

Beachte, dass  $M$  und  $N$  dabei auch gleich sein können; in der Tat ist offensichtlich

$$M = N \quad \text{genau dann, wenn} \quad M \subset N \text{ und } N \subset M.$$

Oft wird man eine Gleichheit  $M = N$  von Mengen auch so beweisen, dass man separat  $M \subset N$  und  $N \subset M$  zeigt.

Wenn wir ausdrücken wollen, dass  $M$  eine Teilmenge von  $N$  und nicht gleich  $N$  ist, so schreiben wir dies als  $M \subsetneq N$  und sagen, dass  $M$  eine **echte Teilmenge** von  $N$  ist. Es ist wichtig, dies von der Aussage  $M \not\subset N$  zu unterscheiden, die bedeutet, dass  $M$  keine Teilmenge von  $N$  ist.

Achtung: Manchmal wird in der Literatur das Symbol „ $\subset$ “ für *echte* Teilmengen und „ $\subseteq$ “ für nicht notwendig echte Teilmengen verwendet.

- (f) Hat eine Menge  $M$  nur endlich viele Elemente, so nennt man  $M$  eine **endliche Menge** und schreibt die Anzahl ihrer Elemente als  $|M|$ . Andernfalls setzt man formal  $|M| = \infty$ .

**Bemerkung 1.13** (Russellsches Paradoxon). Die oben gegebene Charakterisierung von Mengen von Cantor ist aus mathematischer Sicht natürlich sehr schwammig. In der Tat hat Bertrand Russell kurz darauf bemerkt, dass sie sogar schnell zu Widersprüchen führt. Er betrachtet dazu

$$M = \{A : A \text{ ist eine Menge mit } A \notin A\}, \quad (*)$$

also „die Menge aller Mengen, die sich nicht selbst als Element enthalten“. Sicherlich ist es eine merkwürdige Vorstellung, dass eine Menge sich selbst als Element enthalten könnte – im Sinne von Cantors Charakterisierung wäre die Definition (\*) aber zulässig. Fragen wir uns nun allerdings, ob sich die so konstruierte Menge  $M$  selbst als Element enthält, so erhalten wir sofort einen Widerspruch: Wenn  $M \in M$  gilt, so würde das nach der Definition (\*) ja gerade bedeuten, dass  $M \notin M$  ist – und das wiederum, dass doch  $M \in M$  ist. Man bezeichnet dies als das *Russellsche Paradoxon*.

Die Ursache für diesen Widerspruch ist, dass die Definition (\*) rückbezüglich ist: Wir wollen eine neue Menge  $M$  konstruieren, verwenden dabei aber auf der rechten Seite der Definition *alle Mengen*, also u. a. auch die Menge  $M$ , die wir gerade erst definieren wollen. Das ist in etwa so, als würdet ihr im Beweis eines Satzes die Aussage des Satzes selbst verwenden – und das ist natürlich nicht zulässig.

Man muss bei der Festlegung, was Mengen sind und wie man sie bilden kann, also eigentlich viel genauer vorgehen, als es Cantor getan hat. Heutzutage verwendet man hierzu in der Regel das im Jahre 1930 aufgestellte Axiomensystem von Zermelo und Fraenkel, das genau angibt, wie man aus bekannten Mengen neue konstruieren darf: z. B. indem man sie schneidet oder vereint, oder aus bereits bekannten Mengen Elemente mit einer bestimmten Eigenschaft auswählt. Wir wollen dies hier in dieser Vorlesung aber nicht weiter thematisieren und uns mit der naiven Mengencharakterisierung von Cantor begnügen (sowie der Versicherung meinerseits, dass schon alles in Ordnung ist, wenn wir neue Mengen immer nur aus alten konstruieren und keine rückbezüglichen Definitionen hinschreiben). Genaueres zum Zermelo-Fraenkel-Axiomensystem könnt ihr z. B. in [E, Kapitel 13] nachlesen.

**Notation 1.14** (Reelle Zahlen). Unser wichtigstes Beispiel für eine Menge ist die Menge der **reellen Zahlen**, die wir mit  $\mathbb{R}$  bezeichnen werden. Wir wollen die Existenz der reellen Zahlen in dieser Vorlesung axiomatisch voraussetzen und begnügen uns daher an dieser Stelle damit zu sagen, dass man sie sich als die Menge der Punkte auf einer Geraden (der „Zahlengeraden“) vorstellen kann. Zusätzlich werden wir in den nächsten beiden Kapiteln die mathematischen Eigenschaften von  $\mathbb{R}$  exakt angeben (und ebenfalls axiomatisch voraussetzen) – und zwar genügend viele Eigenschaften, um  $\mathbb{R}$  dadurch eindeutig zu charakterisieren.

Ich möchte hier noch einmal betonen, dass man die Existenz und die Eigenschaften der reellen Zahlen eigentlich nicht voraussetzen müsste: Man kann das auch allein aus den Axiomen der Logik und Mengenlehre herleiten! Dies wäre jedoch relativ aufwendig und würde euch im Moment mehr verwirren als helfen, daher wollen wir hier darauf verzichten. Wer sich trotzdem dafür interessiert, kann die Einzelheiten hierzu in [E, Kapitel 1 und 2] nachlesen.

Außer den reellen Zahlen sind vor allem noch die folgenden Teilmengen von  $\mathbb{R}$  wichtig:

- (a) die Menge  $\mathbb{N} = \{0, 1, 2, \dots\}$  der **natürlichen Zahlen** (Achtung: In der Literatur wird die 0 manchmal nicht mit zu den natürlichen Zahlen gezählt!);

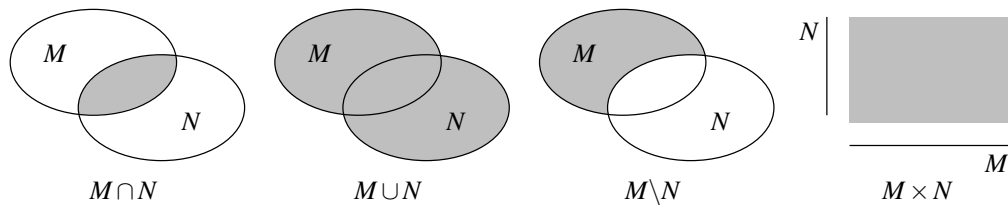
- (b) die Menge  $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$  der **ganzen Zahlen**;
- (c) die Menge  $\mathbb{Q} = \{\frac{p}{q} : p, q \in \mathbb{Z}, q \neq 0\}$  der **rationalen Zahlen**.

Offensichtlich sind diese Mengen ineinander enthalten: Es gilt  $\mathbb{N} \subsetneq \mathbb{Z} \subsetneq \mathbb{Q} \subsetneq \mathbb{R}$ . Teilmengen von  $\mathbb{R}$ , die durch Ungleichungen gegeben sind, schreiben wir in der Regel, indem wir die Ungleichungsbedingung als Index an das Symbol  $\mathbb{R}$  schreiben, z. B.  $\mathbb{R}_{\geq 0}$  für die Menge  $\{x \in \mathbb{R} : x \geq 0\}$  aller nicht-negativen Zahlen.

**Notation 1.15.** Sind  $M$  und  $N$  Mengen, so bezeichnen wir mit ...

- (a)  $M \cap N := \{x : x \in M \text{ und } x \in N\}$  die **Schnittmenge** von  $M$  und  $N$ . Gilt  $M \cap N = \emptyset$ , so sagen wir, dass  $M$  und  $N$  **disjunkt** sind.
- (b)  $M \cup N := \{x : x \in M \text{ oder } x \in N\}$  die **Vereinigungsmenge** von  $M$  und  $N$ . Im Fall einer **disjunkten Vereinigung** mit  $M \cap N = \emptyset$  schreiben wir statt  $M \cup N$  auch  $M \sqcup N$ .
- (c)  $M \setminus N := \{x : x \in M \text{ und } x \notin N\}$  die **Differenzmenge** von  $M$  und  $N$ .
- (d)  $M \times N := \{(x, y) : x \in M, y \in N\}$  die **Produktmenge** bzw. das Produkt von  $M$  und  $N$ . Die Schreibweise  $(x, y)$  steht hierbei für ein **geordnetes Paar**, d. h. einfach für die Angabe eines Elements aus  $M$  und eines aus  $N$  (wobei es auch im Fall  $M = N$  auf die Reihenfolge ankommt, d. h.  $(x, y)$  ist genau dann gleich  $(x', y')$  wenn  $x = x'$  und  $y = y'$ ). Im Fall  $M = N$  schreibt man  $M \times N = M \times M$  auch als  $M^2$ .
- (e)  $\mathcal{P}(M) := \{A : A \text{ ist Teilmenge von } M\}$  die **Potenzmenge** von  $M$ .

Das Symbol „:=“ bedeutet hierbei, dass der Ausdruck auf der linken Seite durch die rechte Seite definiert wird. Die Konstruktionen (a) bis (d) können durch die folgenden Bilder veranschaulicht werden. Natürlich sind sie auch für mehr als zwei Mengen möglich; aus der Schule kennt ihr zum Beispiel sicher den Fall  $\mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R}$ .



Die Potenzmenge  $\mathcal{P}(M)$  aller Teilmengen einer gegebenen Menge  $M$  lässt sich dagegen nicht so einfach durch ein Bild darstellen. Es ist z. B.

$$\mathcal{P}(\{0, 1\}) = \{\emptyset, \{0\}, \{1\}, \{0, 1\}\}.$$

**Aufgabe 1.16.** Wie lautet die Negation der folgenden Aussagen? Formuliere außerdem die Aussage (a) in Worten (also analog zu (b)) sowie die Aussage (b) mit Quantoren und anderen mathematischen Symbolen (also analog zu (a)).

- (a)  $\forall n \in \mathbb{N} \exists m \in \mathbb{N} : n = 2m$ .
- (b) Zwischen je zwei verschiedenen reellen Zahlen gibt es noch eine weitere reelle Zahl.
- (c) Sind  $M, N, R$  Mengen mit  $R \subset N \subset M$ , so ist  $M \setminus N \subset M \setminus R$ .

**Aufgabe 1.17.** Es seien  $A, B, C$  Aussagen und  $M, N, R$  Mengen. Man zeige:

- (a)  $A \vee (B \wedge C) \Leftrightarrow (A \vee B) \wedge (A \vee C)$  und  $A \wedge (B \vee C) \Leftrightarrow (A \wedge B) \vee (A \wedge C)$ .
- (b)  $M \cup (N \cap R) = (M \cup N) \cap (M \cup R)$  und  $M \cap (N \cup R) = (M \cap N) \cup (M \cap R)$ .

**Aufgabe 1.18.** Welche der folgenden Aussagen sind für beliebige gegebene  $x$  und  $M$  äquivalent zueinander? Zeige jeweils die Äquivalenz bzw. widerlege sie durch ein Gegenbeispiel.

- (a)  $x \in M$
- (b)  $\{x\} \subset M$
- (c)  $\{x\} \cap M \neq \emptyset$
- (d)  $\{x\} \in M$
- (e)  $\{x\} \setminus M = \emptyset$
- (f)  $M \setminus \{x\} = \emptyset$

**Aufgabe 1.19.** Man beweise oder widerlege: Für alle Mengen  $A \subset M$  und  $A' \subset M'$  gibt es Teilmengen  $B$  und  $C$  von  $M$  sowie  $B'$  und  $C'$  von  $M'$ , so dass

$$(M \times M') \setminus (A \times A') = (B \times B') \cup (C \times C').$$

Können Sie die Aussage durch eine Skizze veranschaulichen?

## 2. Relationen und Funktionen

Nachdem wir Mengen eingeführt haben, wollen wir nun auch mehrere von ihnen miteinander in Beziehung setzen können. Das Grundkonzept hierfür ist das einer Relation.

**Definition 2.1** (Relationen). Es seien  $M$  und  $N$  zwei Mengen. Eine **Relation** zwischen  $M$  und  $N$  ist eine Teilmenge  $R$  des Produkts  $M \times N$ . Für  $x \in M$  und  $y \in N$  mit  $(x, y) \in R$  sagen wir dann „ $x$  steht (bezüglich  $R$ ) in Relation zu  $y$ “. Ist  $M = N$ , so nennen wir  $R$  auch eine **Relation auf  $M$** .

**Bemerkung 2.2.** Um eine Relation  $R$  anzugeben, also eine Teilmenge  $R \subset M \times N$  zu definieren, müssen wir demzufolge einfach für alle Paare  $(x, y)$  mit  $x \in M$  und  $y \in N$  festlegen, ob  $(x, y) \in R$  gelten, also ob  $x$  in Relation zu  $y$  stehen soll.

Wie wir in diesem Kapitel sehen werden, werden Relationen in der Mathematik für sehr unterschiedliche Konzepte verwendet – z. B. um Zahlen miteinander zu vergleichen wie in Beispiel 2.3, um eine Menge auf eine andere abzubilden wie in Abschnitt 2.A, oder um die Elemente einer Menge nach bestimmten Kriterien zu Klassen zusammenzufassen wie in Abschnitt 2.B. Dementsprechend sind für die Aussage „ $x$  steht bezüglich  $R$  in Relation zu  $y$ “ auch je nach Anwendung ganz unterschiedliche Notationen üblich. Für allgemeine, nicht näher spezifizierte Relationen schreibt man hierfür oft  $xRy$ .

**Beispiel 2.3** (Kleiner-Relation). Für  $M = N = \mathbb{R}$  betrachten wir die Relation

$$R = \{(x, y) : x, y \in \mathbb{R} \text{ mit } x < y\},$$

für die  $x$  also genau dann in Relation zu  $y$  steht, wenn  $x < y$  gilt. Man nennt  $R$  deshalb auch die **Kleiner-Relation** auf  $\mathbb{R}$ . Die Notation „ $xRy$ “ aus Bemerkung 2.2 stimmt in diesem Fall also mit der Schreibweise „ $x < y$ “ überein, wenn man die Relation  $R$  direkt mit dem Symbol „ $<$ “ bezeichnet. In der Tat ist es aus diesem Grund bei manchen Relationen üblich, sie gleich mit Symbolen statt mit Buchstaben zu benennen.

### 2.A Funktionen

Die mit Abstand wichtigsten Relationen sind ohne Zweifel die Funktionen, die ihr natürlich bereits hinlänglich aus der Schule kennt. Wir wollen sie hier nun exakt einführen und ihre ersten Eigenschaften untersuchen.

**Definition 2.4** (Funktionen). Es seien  $M$  und  $N$  zwei Mengen.

- Eine **Funktion** oder **Abbildung**  $f$  von  $M$  nach  $N$ , geschrieben  $f: M \rightarrow N$ , ist eine Relation zwischen  $M$  und  $N$ , bezüglich der jedes Element  $x$  von  $M$  zu **genau einem** Element  $y$  von  $N$  in Relation steht. Wir schreiben dies dann als  $x \mapsto y$  oder  $y = f(x)$  und sagen,  $y$  ist das **Bild** von  $x$  unter  $f$  bzw. der **Wert** von  $f$  in  $x$ .
- Für eine Funktion  $f: M \rightarrow N$  bezeichnet man die Menge  $M$  als **Definitionsmenge**, **Startmenge** oder **Startraum** von  $f$ . Die Menge  $N$  heißt **Zielmenge** oder **Zielraum** von  $f$ .

**Bemerkung 2.5.**

- Um eine Funktion komplett festzulegen, müssen wir zuerst einmal den Start- und Zielraum angeben, und dann schließlich noch von jedem Element des Startraums sagen, auf welches Element des Zielraums es abgebildet wird. In welcher Form wir diese Zuordnung angeben – ob durch eine Formel, durch explizites Auflisten der Funktionswerte aller Elemente des Startraums, oder irgendwie anders – spielt dabei keine Rolle. So sind z. B.

$$f: \{0, 1\} \rightarrow \mathbb{R}, x \mapsto 2x^2, \quad g: \{0, 1\} \rightarrow \mathbb{R}, x \mapsto 2x^3, \quad h: \{0, 1\} \rightarrow \mathbb{R}, x \mapsto \begin{cases} 0 & \text{für } x = 0, \\ 2 & \text{für } x = 1 \end{cases}$$

trotz ihrer verschieden aussehenden Vorschriften dieselbe Funktion (d. h. es gilt  $f = g = h$ ), da alle drei den gleichen Start- und Zielraum haben und aus den gleichen Zuordnungen  $0 \mapsto 0$  und  $1 \mapsto 2$  bestehen. Mit anderen Worten sind zwei Funktionen  $f, g: M \rightarrow N$  also genau dann gleich, wenn sie an jedem Punkt die gleichen Werte besitzen, also wenn gilt

$$\forall x \in M: f(x) = g(x).$$

- (b) Man sieht leider oft, dass eine Funktion  $f: M \rightarrow N$  als  $f(x)$  geschrieben wird. Es ist wichtig zu verstehen, dass diese Notation gemäß Definition 2.4 falsch ist: Mit  $f(x)$  wird *der Wert der Funktion  $f$  in einem Punkt  $x \in M$*  bezeichnet. Somit ist  $f(x)$  (für gegebenes  $x$ ) ein Element von  $N$ , und damit ein ganz anderes mathematisches Objekt als die Funktion selbst, die wir nur mit  $f$  bezeichnen und die eine Relation zwischen  $M$  und  $N$  ist. Dies mag auf den ersten Blick spitzfindig erscheinen – wir werden aber später noch oft Mengen sehen, deren Elemente Funktionen sind, und dann ist es natürlich wichtig, dies von der Menge ihrer Funktionswerte zu unterscheiden.

### Beispiel 2.6.

- (a) Die Zuordnungen

$$f: \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \frac{1}{x} \quad \text{und} \quad g: \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \begin{cases} x+1 & \text{für } x \leq 0 \\ x & \text{für } x \geq 0 \end{cases}$$

sind in dieser Form keine zulässigen Funktionsdefinitionen, weil im Fall von  $f$  der Zahl 0 kein gültiger Funktionswert zugeordnet wird und im Fall  $g$  für die Zahl 0 zwei (sich widersprechende) Festlegungen des Funktionswertes gemacht werden. Dies lässt sich jedoch in beiden Fällen leicht reparieren, z. B. indem man die Festlegungen abändert in

$$f: \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}, x \mapsto \frac{1}{x} \quad \text{und} \quad g: \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \begin{cases} x+1 & \text{für } x < 0 \\ x & \text{für } x \geq 0 \end{cases}.$$

- (b) Zu jeder Menge  $M$  gibt es die **identische Abbildung**

$$\text{id}_M: M \rightarrow M, x \mapsto x,$$

die jedes Element auf sich selbst abbildet.

- (c) Ist  $f: M \rightarrow N$  eine Abbildung und  $A \subset M$  eine Teilmenge des Startraums, so erhält man durch die Einschränkung der Definitionsmenge von  $M$  auf  $A$  eine neue Abbildung, die wir mit

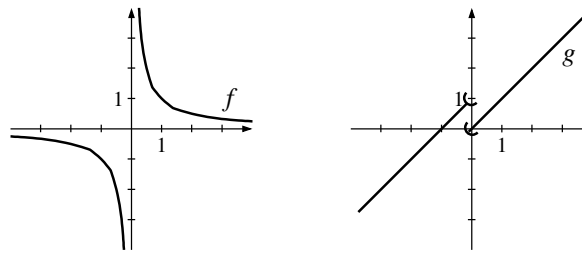
$$f|_A: A \rightarrow N, x \mapsto f(x)$$

bezeichnen und die die **Einschränkung** von  $f$  auf  $A$  genannt wird. Genauso kann man natürlich auch die Zielmenge  $N$  auf eine Teilmenge  $B$  einschränken, wenn  $f$  nur Werte in  $B$  annimmt. Es ist üblich, bei einer derartigen Einschränkung der Zielmenge immer noch den gleichen Namen für die Abbildung zu verwenden, also dann  $f: M \rightarrow B$  zu schreiben (auch wenn es sich dabei um eine andere Funktion als das ursprüngliche  $f: M \rightarrow N$  handelt).

**Bemerkung 2.7** (Graph einer Abbildung). Zu einer Abbildung  $f: M \rightarrow N$  heißt die Menge

$$\{(x, f(x)) : x \in M\} \subset M \times N$$

der **Graph** von  $f$ . Sind  $M$  und  $N$  Teilmengen von  $\mathbb{R}$ , so ist dieser Graph also eine Teilmenge von  $\mathbb{R}^2$ , und man kann ihn leicht zeichnen und dadurch die Abbildung veranschaulichen. Für die Abbildungen aus Beispiel 2.6 (a) sieht dies z. B. so aus:



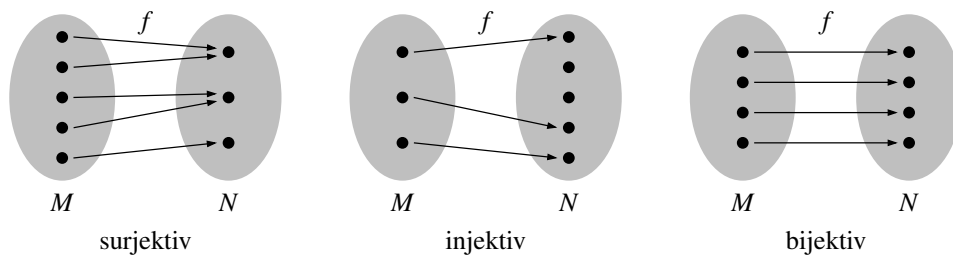
Beachte, dass dieser Graph nach den Definitionen 2.1 und 2.4 eigentlich sogar genau das gleiche ist wie die Funktion selbst, nämlich die Teilmenge des Produkts  $M \times N$ , die aus den Paaren  $(x, y)$  besteht, für die  $x$  bezüglich  $f$  in Relation zu  $y$  steht, also  $y = f(x)$  gilt. Der Begriff des Graphen soll hier also nur noch einmal deutlich machen, dass man sich die Funktion gerade wirklich als ein derart „grafisches“ Objekt vorstellt und nicht als eine „Zuordnung“ von  $M$  nach  $N$ .

In der Definition 2.4 einer Abbildung  $f: M \rightarrow N$  verlangen wir, dass jedem Element von  $M$  genau ein Element von  $N$  zugeordnet wird. Wir fordern jedoch nicht auch umgekehrt, dass jedes Element des Zielraums  $N$  das Bild von genau einem Element von  $M$  ist, oder dass es überhaupt als Bild eines Elements von  $M$  auftritt. Abbildungen, die diese Eigenschaften dennoch besitzen, haben spezielle Namen, die wir jetzt einführen wollen.

**Definition 2.8** (Eigenschaften von Abbildungen). Es sei  $f: M \rightarrow N$  eine Abbildung.

- (a) Ist  $y \in N$  und  $x \in M$  mit  $f(x) = y$ , so heißt  $x$  ein **Urbild** von  $y$  unter  $f$ .
- (b) Hat jedes  $y \in N$  ...
  - *mindestens* ein Urbild, so heißt  $f$  **surjektiv**.  
In Quantoren bedeutet dies:  $\forall y \in N \exists x \in M: f(x) = y$ .
  - *höchstens* ein Urbild, so heißt  $f$  **injektiv**.  
In Quantoren bedeutet dies:  $\forall x_1, x_2 \in M: f(x_1) = f(x_2) \Rightarrow x_1 = x_2$ . (Also: Haben zwei Elemente des Startraums das gleiche Bild, so müssen sie bereits dasselbe Element sein.)
  - *genau* ein Urbild, ist  $f$  also surjektiv und injektiv, so heißt  $f$  **bijektiv**.

Im folgenden Bild sind diese Begriffe anschaulich dargestellt (wobei  $M$  und  $N$  endlich sind und nur aus den eingezeichneten Punkten bestehen).



02

**Beispiel 2.9.** Betrachten wir noch einmal die Funktionen aus Beispiel 2.6 (a), so ist die Funktion  $f$  nicht surjektiv (und damit auch nicht bijektiv), da das Element 0 des Zielraums kein Urbild hat. Sie ist jedoch injektiv: Sind  $x_1, x_2 \in \mathbb{R} \setminus \{0\}$  mit  $f(x_1) = f(x_2)$ , also  $\frac{1}{x_1} = \frac{1}{x_2}$ , so folgt durch Multiplikation mit  $x_1 x_2$  sofort  $x_1 = x_2$ .

Die Funktion  $g$  dagegen ist surjektiv: Eine Zahl  $y \in \mathbb{R}$  hat als Urbild  $x = y$  für  $y \geq 0$ , und  $x = y - 1$  für  $y < 0$ . Sie ist allerdings nicht injektiv, denn es ist  $g(-1) = g(0) = 0$ .

Beachte, dass diese Eigenschaften auch an den Graphen in Bemerkung 2.7 ablesbar sind: Surjektivität bzw. Injektivität bedeuten gerade, dass jede horizontale Gerade auf der Höhe eines Wertes im Zielraum den Funktionsgraphen in mindestens bzw. höchstens einem Punkt schneidet. Wichtig ist

auch, dass diese Eigenschaften von der Wahl des Start- und Zielraums abhängen: So wird z. B.  $f$  bijektiv, wenn man den Zielraum  $\mathbb{R}$  durch  $\mathbb{R} \setminus \{0\}$  ersetzt, und  $g$  injektiv, wenn man den Startraum auf  $\mathbb{R}_{\geq 0}$  einschränkt (in der Notation von Beispiel 2.6 (c) also  $g|_{\mathbb{R}_{\geq 0}}$  betrachtet).

**Aufgabe 2.10.** Wie viele Abbildungen gibt es zwischen den Mengen  $\{1, 2, 3, 4\}$  und  $\{1, 2, 3, 4, 5, 6\}$ ? Wie viele von ihnen sind injektiv?

Bilder und Urbilder unter Abbildungen betrachtet man oft auch von ganzen Mengen statt nur von Punkten:

**Definition 2.11** (Bild und Urbild von Mengen). Es sei  $f: M \rightarrow N$  eine Abbildung.

(a) Für  $A \subset M$  heißt die Menge

$$f(A) := \{f(x) : x \in A\} \subset N$$

(also die Menge aller Bilder von Punkten in  $A$ ) das **Bild** von  $A$  unter  $f$ .

(b) Ist  $B \subset N$ , so heißt die Menge

$$f^{-1}(B) := \{x \in M : f(x) \in B\} \subset M$$

(also die Menge aller Urbilder von Punkten in  $B$ ) das **Urbild** von  $B$  unter  $f$ .

**Bemerkung 2.12.** Die Grundidee der Notation in Definition 2.11 (a) ist: Schreiben wir als Argument einer Funktion  $f: M \rightarrow N$  eine *Teilmenge* statt einem *Element* von  $M$ , so bedeutet dies, dass wir alle Werte  $f(x)$  für  $x \in M$  zusammen nehmen und diese wieder in einer Menge zusammenfassen. Diese Schreibweise verwendet man auch oft, wenn die Funktion aus einer Rechenverknüpfung besteht, wie z. B. in

$$\mathbb{N} + \frac{1}{2} := \{n + \frac{1}{2} : n \in \mathbb{N}\} = \{\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots\} \quad \text{oder} \quad 3\mathbb{Z} := \{3n : n \in \mathbb{Z}\} = \{\dots, -6, -3, 0, 3, 6, \dots\}.$$

**Beispiel 2.13.** Für die Funktion  $f: \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}$ ,  $x \mapsto \frac{1}{x}$  aus Beispiel 2.6 (a) ist  $f(\mathbb{R}_{>0}) = \mathbb{R}_{>0}$  und  $f^{-1}(\{0\}) = \emptyset$ .

**Beispiel 2.14.** Zwischen den Konstruktionen von Bild und Urbild aus Definition 2.11 und den Mengenoperationen aus Abschnitt 1.B gibt es sehr viele Beziehungen. Um einmal exemplarisch zu sehen, wie derartige Beziehungen aussehen und bewiesen werden können, wollen wir nun zeigen, dass für jede Abbildung  $f: M \rightarrow N$  und zwei beliebige Teilmengen  $A, B \subset M$  stets

$$f(A) \setminus f(B) \subset f(A \setminus B) \quad (*)$$

gilt.

Zum Beweis müssen wir zeigen, dass jedes Element der linken Menge auch in der rechten Menge liegt. Es sei also  $y \in f(A) \setminus f(B)$  beliebig. Insbesondere ist damit  $y \in f(A)$ , nach Definition 2.11 (a) also  $y = f(x)$  für ein  $x \in A$ . Würde nun auch  $x \in B$  gelten, so hätten wir wegen  $y = f(x)$  auch  $y \in f(B)$ , im Widerspruch zu  $y \in f(A) \setminus f(B)$ . Also ist  $x \notin B$ , und damit  $x \in A \setminus B$ . Damit besagt  $y = f(x)$  aber gerade  $y \in f(A \setminus B)$ . Insgesamt haben wir somit die behauptete Teilmengenbeziehung (\*) gezeigt.

Beachte allerdings, dass in (\*) im Allgemeinen keine Gleichheit gilt: Für  $f: \mathbb{R} \rightarrow \mathbb{R}$ ,  $x \mapsto x^2$  mit  $A = \{-1, 1\}$  und  $B = \{-1\}$  ist

$$f(A) \setminus f(B) = \{1\} \setminus \{1\} = \emptyset, \quad \text{aber} \quad f(A \setminus B) = f(\{1\}) = \{1\}.$$

**Aufgabe 2.15.** Beweise die folgenden Teilmengenbeziehungen und untersuche jeweils, ob auch die Gleichheit gilt.

(a) Für alle Mengen  $M, A, B$  gilt  $M \setminus (A \cup B) \subset (M \setminus A) \cap (M \setminus B)$ .

(b) Ist  $f: M \rightarrow N$  eine Abbildung und  $A \subset N$ , so ist  $f(f^{-1}(A)) \subset A$ .

**Aufgabe 2.16.** Es sei  $f: M \rightarrow N$  eine Abbildung. Finde für das Symbol  $\square$  jeweils eine der Mengenbeziehungen  $\subset, =, \supset$ , so dass die folgenden Aussagen wahr werden, und beweise die so entstandenen Aussagen!

(a)  $f(A) \cap f(B) \square f(A \cap B)$  für alle  $A, B \subset M$ .

(b)  $f^{-1}(A) \cap f^{-1}(B) \sqsubset f^{-1}(A \cap B)$  für alle  $A, B \subset N$ .

Als Nächstes wollen wir nun die euch sicher bereits bekannte Verkettung, also die Hintereinanderausführung von Funktionen einführen.

**Definition 2.17** (Verkettung von Funktionen). Es seien  $f: M \rightarrow N$  und  $g: N \rightarrow R$  zwei Abbildungen (also so dass die Zielmenge von  $f$  gleich der Startmenge von  $g$  ist). Dann heißt die Abbildung

$$g \circ f: M \rightarrow R, x \mapsto g(f(x))$$

die **Verkettung** von  $f$  und  $g$ .

**Bemerkung 2.18.** Bei der Verkettung zweier Funktionen kommt es natürlich auf die Reihenfolge an, allein schon weil in der Situation von Definition 2.17 in der Regel der Zielraum von  $g$  ja nicht mit dem Startraum von  $f$  übereinstimmt und die „umgekehrte Verkettung“  $f \circ g$  damit gar nicht definierbar wäre. Beachte dabei, dass die Notation  $g \circ f$  lautet, obwohl wir zuerst  $f$  (von  $M$  nach  $N$ ) und dann  $g$  (von  $N$  nach  $R$ ) anwenden – man liest  $g \circ f$  daher manchmal auch als „ $g$  nach  $f$ “. Diese vielleicht etwas merkwürdig erscheinende Notation kommt einfach daher, dass die Buchstaben in der gleichen Reihenfolge stehen sollen wie bei der Abbildungsvorschrift  $x \mapsto g(f(x))$ .

Wir werden nun unser erstes *Lemma* beweisen – „Lemma“ ist griechisch und bedeutet eigentlich „Annahme“, aber in der Mathematik wird dieser Begriff für einen *Hilfssatz* verwendet, also für ein kleines Zwischenresultat, das vielleicht für sich genommen nicht übermäßig überraschend oder interessant ist, aber das in späteren Beweisen immer wieder nützlich sein wird. In unserem momentanen Fall geht es einfach darum, dass die Verkettung von Abbildungen *assoziativ* ist (siehe auch Definition 3.1):

**Lemma 2.19** (Assoziativität der Verkettung). Sind  $f: M \rightarrow N$ ,  $g: N \rightarrow R$  und  $h: R \rightarrow S$  drei Abbildungen, so gilt  $h \circ (g \circ f) = (h \circ g) \circ f$ . (Man schreibt für diese Abbildung daher oft auch einfach  $h \circ g \circ f$ .)

*Beweis.* Nach Definition 2.4 können wir die Gleichheit zweier Funktionen zeigen, indem wir für jedes Element der Startmenge nachweisen, dass sein Bild unter beiden Funktionen übereinstimmt. Dies rechnen wir nun einfach durch wiederholtes Einsetzen von Definition 2.17 nach: Es gilt

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

und

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

Da diese beiden Ausdrücke übereinstimmen, ist das Lemma bewiesen.  $\square$

Am Ende dieses Abschnitts wollen wir schließlich noch das Konzept von Umkehrfunktionen bijektiver Funktionen einführen.

**Definition 2.20** (Umkehrfunktionen). Es sei  $f: M \rightarrow N$  eine bijektive Funktion. Dann heißt

$$f^{-1}: N \rightarrow M, y \mapsto \text{das eindeutige Urbild von } y \text{ unter } f$$

die **Umkehrfunktion** bzw. **Umkehrabbildung** von  $f$ .

**Bemerkung 2.21.** Für die Umkehrfunktion  $f^{-1}$  einer bijektiven Funktion  $f: M \rightarrow N$  gilt nach Konstruktion offensichtlich  $f^{-1} \circ f = \text{id}_M$  und  $f \circ f^{-1} = \text{id}_N$ .

Gibt es umgekehrt zu einer Funktion  $f: M \rightarrow N$  eine Abbildung  $g: N \rightarrow M$  mit  $g \circ f = \text{id}_M$  und  $f \circ g = \text{id}_N$ , so ist  $f$  bijektiv:

- $f$  ist surjektiv: Ist  $y \in N$  beliebig, so ist  $x := g(y) \in M$  ein Urbild von  $y$  unter  $f$ , denn es ist  $f(x) = f(g(y)) = y$ .
- $f$  ist injektiv: Sind  $x_1, x_2 \in M$  mit  $f(x_1) = f(x_2)$ , so folgt durch Anwenden von  $g$  sofort auch  $g(f(x_1)) = g(f(x_2))$ , und damit  $x_1 = x_2$ .

**Beispiel 2.22.** Die Funktion  $f: \mathbb{R} \rightarrow \mathbb{R}$ ,  $x \mapsto x + 1$  ist bijektiv, und ihre Umkehrabbildung ist  $f^{-1}: \mathbb{R} \rightarrow \mathbb{R}$ ,  $x \mapsto x - 1$ . In der Tat gilt nämlich für alle  $x \in \mathbb{R}$

$$(f^{-1} \circ f)(x) = (x + 1) - 1 = x \quad \text{und} \quad (f \circ f^{-1})(x) = (x - 1) + 1 = x.$$

**Bemerkung 2.23** (Urbilder und Umkehrabbildungen). Beachte, dass wir das Urbild einer Menge unter einer Abbildung  $f: M \rightarrow N$  in Definition 2.11 (b) mit dem gleichen Symbol  $f^{-1}$  bezeichnet haben wie (im Fall einer bijektiven Abbildung) die Umkehrabbildung aus Definition 2.20. Das ist vielleicht etwas unglücklich gewählt, aber in der Literatur so fest verankert, dass wir hier nicht davon abweichen wollen. Bei genauem Hinschauen kann man aber auch immer feststellen, was gemeint ist:

Ist  $f: M \rightarrow N$  eine Abbildung, so bezeichnet ...  
 ...  $f^{-1}(B)$  für eine Menge  $B \subset N$  das *Urbild* von  $B$  wie in Definition 2.11 (b); es existiert für jede Abbildung  $f$ .  
 ...  $f^{-1}(y)$  für ein Element  $y \in B$  den *Wert der Umkehrabbildung* bei  $y$  wie in Definition 2.20; er existiert nur für bijektives  $f$ .

Letztlich hängen diese beiden Notationen aber auch eng miteinander zusammen: Ist  $f$  bijektiv und ist  $x \in M$  mit  $f(x) = y$ , so ist  $f^{-1}(y) = x$  (mit  $f^{-1}$  im Sinne der Umkehrabbildung) und  $f^{-1}(\{y\}) = \{x\}$  (mit  $f^{-1}$  im Sinne des Urbildes).

**Aufgabe 2.24.**

- Untersuche die Abbildung  $f: \mathbb{Z} \rightarrow \mathbb{Z}$ ,  $x \mapsto 3x + 2$  auf Injektivität und Surjektivität.
- Untersuche die Abbildung  $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $(x, y) \mapsto (xy, x + 1)$  auf Injektivität und Surjektivität.
- Man zeige: Sind  $f: M \rightarrow N$  und  $g: N \rightarrow R$  surjektiv, so ist auch  $g \circ f: M \rightarrow R$  surjektiv.

**Aufgabe 2.25.** Es seien  $f: M \rightarrow N$  und  $g: N \rightarrow R$  bijektiv. Zeige, dass dann auch  $f^{-1}: N \rightarrow M$  und  $g \circ f: M \rightarrow R$  bijektiv sind.

**Aufgabe 2.26.** Man beweise oder widerlege:

- Sind  $f: M \rightarrow N$  und  $g: N \rightarrow R$  zwei Abbildungen und ist  $g \circ f$  injektiv, so ist auch  $f$  injektiv.
- Sind  $f: M \rightarrow N$  und  $g: N \rightarrow R$  zwei Abbildungen und ist  $g \circ f$  injektiv, so ist auch  $g$  injektiv.

**Aufgabe 2.27.** Es sei  $f: M \rightarrow N$  eine Abbildung zwischen nicht-leeren Mengen. Man zeige:

- $f$  ist genau dann surjektiv, wenn es eine Abbildung  $g: N \rightarrow M$  gibt mit  $f \circ g = \text{id}_N$ .
- $f$  ist genau dann injektiv, wenn es eine Abbildung  $g: N \rightarrow M$  gibt mit  $g \circ f = \text{id}_M$ .

**Aufgabe 2.28.** Es sei  $f: M \rightarrow N$  eine Abbildung zwischen zwei Mengen. Man beweise:

$$f \text{ ist surjektiv} \iff \text{für alle } A, B \subset N \text{ mit } f^{-1}(A) = f^{-1}(B) \text{ gilt } A = B.$$

**Aufgabe 2.29.** Für diese Aufgabe nennen wir eine Funktion  $f: \mathbb{R} \rightarrow \mathbb{R} \dots$

- *monoton wachsend*, wenn gilt:  $\forall x, y \in \mathbb{R} : x < y \Rightarrow f(x) \leq f(y)$ ;
- *irgendwann konstant*, wenn gilt:  $\exists x \in \mathbb{R} \forall y \in \mathbb{R} : x \leq y \Rightarrow f(x) = f(y)$ ;
- *interessant*, wenn gilt:  $\forall x \in \mathbb{R} \exists y \in \mathbb{R} : x < y \text{ und } f(x) < f(y)$ .

Man zeige:

- Die Funktion  $f: \mathbb{R} \rightarrow \mathbb{R}$ ,  $x \mapsto x^2$  ist interessant.
- Für die Eigenschaften „monoton wachsend“ und „irgendwann konstant“ gilt, dass mit  $f: \mathbb{R} \rightarrow \mathbb{R}$  und  $g: \mathbb{R} \rightarrow \mathbb{R}$  auch deren Summenfunktion

$$f + g: \mathbb{R} \rightarrow \mathbb{R}, x \mapsto f(x) + g(x)$$

diese Eigenschaft hat.

- Ist  $f: \mathbb{R} \rightarrow \mathbb{R}$  monoton wachsend, aber nicht interessant, so ist  $f$  irgendwann konstant.

## 2.B Äquivalenzrelationen

Am Anfang dieses Kapitels haben wir allgemeine Relationen eingeführt, als einzigen Spezialfall davon aber bisher nur die Funktionen ausführlicher betrachtet. Wir wollen daher nun noch einen ganz anderen wichtigen Typ von Relationen studieren, die sogenannten Äquivalenzrelationen.

Angenommen, wir möchten eine Menge  $M$  untersuchen, die uns zunächst einmal zu groß oder zu kompliziert erscheint. Es gibt dann zwei prinzipiell verschiedene Möglichkeiten, wie man daraus eine kleinere bzw. einfachere Menge machen kann:

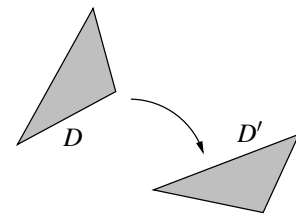
- Wir können uns auf eine Teilmenge von  $M$  beschränken – dann schließen wir allerdings manche Elemente von  $M$  von unserer Untersuchung aus.
- Wir können zwar alle Elemente von  $M$  betrachten, aber manche von ihnen miteinander identifizieren bzw. zu Klassen zusammenfassen – d. h. sie als gleich bzw. „äquivalent“ ansehen, wenn sie für das zu untersuchende Problem ähnliche Eigenschaften haben.

Diese zweite Idee der Identifizierung ähnlicher Elemente führt zum Begriff der Äquivalenzrelationen. Sie klingt vielleicht zunächst etwas abstrakt, ist euch aber sicher schon an vielen Stellen begegnet. Hier sind zwei einfache Beispiele dafür.

### Beispiel 2.30.

- (a) Eine analoge Uhr vereinfacht die recht große Menge aller (vergangenen und zukünftigen) Zeitpunkte, die wir uns als Zeitachse  $M = \mathbb{R}$  vorstellen können, indem sie nach jeweils 12 Stunden wieder dasselbe anzeigt. Sie identifiziert also zwei Zeitpunkte  $x, y \in \mathbb{R}$  (gemessen in Stunden) miteinander, wenn  $x - y$  ein ganzzahliges Vielfaches von 12 ist. Dadurch „verkleinert“ sie die ursprüngliche Zeitachse auf ein gut überschaubares Intervall von 12 Stunden – und wir alle wissen, dass uns ein Blick auf die Uhr in vielen Fällen ausreicht, wenn wir den aktuellen Zeitpunkt wissen wollen, auch wenn uns das nichts über das Datum oder die Tageszeit (vormittags oder nachmittags) sagt.
- (b) Als „mathematisches“ Beispiel können wir die Menge  $M$  aller Dreiecke in der Ebene  $\mathbb{R}^2$  betrachten.

Bekanntlich heißen zwei solche Dreiecke  $D, D' \in M$  zueinander *kongruent*, wenn sie wie im Bild rechts durch eine Drehung und / oder Verschiebung auseinander hervorgehen – wir schreiben dies im Folgenden als  $D \sim D'$ . Zueinander kongruente Dreiecke werden oft miteinander identifiziert, nämlich immer dann, wenn es uns nur auf die Form bzw. Größe der Dreiecke, aber nicht auf ihre Lage in der Ebene ankommt.



Wenn wir z. B. sagen, dass die drei Seitenlängen ein Dreieck eindeutig bestimmen, dann meinen wir damit in Wirklichkeit, dass sie das Dreieck *bis auf Kongruenz* eindeutig bestimmen, also nur die Form und Größe festlegen, aber nicht die Lage des Dreiecks in  $\mathbb{R}^2$ . Formal kann man dies so ausdrücken: zu einem Dreieck  $D$  nennt man

$$\bar{D} := \{D' \in M : D' \sim D\},$$

also die Menge aller zu  $D$  kongruenten Dreiecke, die *Kongruenzklasse* von  $D$ . Die Menge aller dieser Kongruenzklassen bezeichnen wir mit

$$M/\sim := \{\bar{D} : D \in M\}.$$

Man kann dann z. B. sagen, dass die Seitenlängen eines Dreiecks ein eindeutiges Element in  $M/\sim$  bestimmen, also eine eindeutige Kongruenzklasse von Dreiecken festlegen – nicht aber ein eindeutiges Element von  $M$ .

Mit der Idee dieser Beispiele im Kopf wollen wir nun den Begriff der Äquivalenzrelation exakt definieren.

**Definition 2.31** (Äquivalenzrelationen). Es sei  $\sim$  wie in Definition 2.1 eine Relation auf einer Menge  $M$ . Wie in Bemerkung 2.2 schreiben wir  $x \sim y$ , wenn  $x$  und  $y$  bezüglich  $\sim$  in Relation stehen.

Man nennt  $\sim$  eine **Äquivalenzrelation** auf  $M$ , wenn die folgenden Eigenschaften gelten:

- (a) (**Reflexivität**) Für alle  $x \in M$  gilt  $x \sim x$ .
- (b) (**Symmetrie**) Sind  $x, y \in M$  mit  $x \sim y$ , so gilt auch  $y \sim x$ .
- (c) (**Transitivität**) Sind  $x, y, z \in M$  mit  $x \sim y$  und  $y \sim z$ , so gilt auch  $x \sim z$ .

In diesem Fall sagt man statt  $x \sim y$  auch, dass  $x$  (bezüglich dieser Relation) zu  $y$  **äquivalent** ist. Zu  $x \in M$  heißt dann die Menge

$$\bar{x} := \{y \in M : y \sim x\}$$

aller Elemente, die zu  $x$  äquivalent sind, die **Äquivalenzklasse** bzw. einfach **Klasse** von  $x$ ; jedes Element dieser Menge nennt man einen **Repräsentanten** dieser Klasse. Die Menge aller Äquivalenzklassen schreiben wir als

$$M/\sim := \{\bar{x} : x \in M\}.$$

### Beispiel 2.32.

- (a) Das Beispiel 2.30 (a) einer analogen Uhr lässt sich mathematisch exakt wie folgt definieren: Auf  $M = \mathbb{R}$  betrachten wir die Relation

$$x \sim y :\Leftrightarrow \text{es gibt ein } k \in \mathbb{Z} \text{ mit } x - y = 12k. \quad (*)$$

Dies ist in der Tat eine Äquivalenzrelation, denn für alle  $x, y, z \in \mathbb{R}$  gilt:

- Reflexivität: Es gilt  $x - x = 12 \cdot 0 = 12k$  mit  $k = 0 \in \mathbb{Z}$ , also  $x \sim x$ .
- Symmetrie: Es gelte  $x \sim y$ , also  $x - y = 12k$  für ein  $k \in \mathbb{Z}$ . Durch Multiplikation mit  $-1$  folgt dann auch  $y - x = 12 \cdot (-k) = 12k'$  mit  $k' := -k \in \mathbb{Z}$ , und damit  $y \sim x$ .
- Transitivität: Es gelte  $x \sim y$  und  $y \sim z$ , nach Definition der Relation also  $x - y = 12k$  und  $y - z = 12k'$  für gewisse  $k, k' \in \mathbb{Z}$  (beachte, dass der Wert von  $k$  in (\*) von  $x$  und  $y$  abhängt und wir daher für die Differenz  $y - z$  eine neue Variable  $k'$  brauchen). Durch Addition dieser beiden Gleichungen erhalten wir  $x - z = 12(k + k') = 12k''$  mit  $k'' := k + k' \in \mathbb{Z}$ , und damit  $x \sim z$ .

Die Äquivalenzklasse z. B. von  $2 \in \mathbb{R}$  ist

$$\bar{2} = \{x \in \mathbb{R} : x \sim 2\} = \{x \in \mathbb{R} : \text{es gibt ein } k \in \mathbb{Z} \text{ mit } x - 2 = 12k\} = \{2 + 12k : k \in \mathbb{Z}\},$$

also die Menge aller Zeitpunkte, zu denen die Uhr auf 2 steht. Jeder beliebige Zeitpunkt  $x \in \mathbb{R}$ , zu dem die Uhr auf 2 steht, ist ein Repräsentant dieser Klasse. Die Menge  $M/\sim$  entspricht allen möglichen Ständen der Uhr; wir können sie uns anschaulich als einen „Kreis mit Umfang 12“ vorstellen.

- (b) Die Kongruenz von Dreiecken aus Beispiel 2.30 (b) ist ebenfalls eine Äquivalenzrelation (es ist offensichtlich, dass sie die Eigenschaften aus Definition 2.31 erfüllt). Die Äquivalenzklassen sind in diesem Fall genau die Kongruenzklassen.
- (c) Die Kleiner-Relation auf  $\mathbb{R}$  aus Beispiel 2.3, also die Relation, für die für  $x, y \in \mathbb{R}$  genau dann  $x \sim y$  gilt, wenn  $x < y$  ist, ist keine Äquivalenzrelation, da sie weder reflexiv noch symmetrisch ist.

Beachte, dass bei unseren Äquivalenzrelationen aus Beispiel 2.32 (a) und (b) jedes Element von  $M$  in genau einer Äquivalenzklasse liegt: Zu jedem Zeitpunkt hat eine analoge Uhr genau einen Stand, und jedes Dreieck in der Ebene liegt in genau einer Kongruenzklasse. Dies beschreibt genau unsere Idee, dass wir die Elemente von  $M$  auf eine bestimmte Art zu Klassen zusammenfassen wollen. Allgemein sind die Axiome einer Äquivalenzrelation aus Definition 2.31 anschaulich genau diejenigen, die man braucht, damit die Relation sinnvoll eine solche Identifizierung von Elementen zu Klassen beschreiben kann. Dies zeigt auch noch einmal der folgende zentrale Satz über Äquivalenzrelationen.

**Satz 2.33** (Eigenschaften von Äquivalenzrelationen). *Es sei  $\sim$  eine Äquivalenzrelation auf einer Menge  $M$ .*

- (a) *Für  $x, y \in M$  gilt  $x \sim y$  genau dann, wenn  $\bar{x} = \bar{y}$ . (Zwei Elemente sind also genau dann äquivalent zueinander, wenn sie die gleiche Äquivalenzklasse bestimmen.)*
- (b) *Jedes Element  $x \in M$  liegt in genau einer Äquivalenzklasse (nämlich in  $\bar{x}$ ). Insbesondere ist  $M$  also die disjunkte Vereinigung aller Äquivalenzklassen. Man sagt dafür auch, dass die Äquivalenzklassen eine Partition von  $M$  bilden.*

*Beweis.*

- (a) Es seien  $x, y \in M$ .

„ $\Rightarrow$ “: Es gelte  $x \sim y$ . Ist dann  $z \in M$  mit  $z \in \bar{x}$ , also  $z \sim x$ , so ist nach der Transitivität wegen  $x \sim y$  auch  $z \sim y$ , also  $z \in \bar{y}$ . Damit gilt  $\bar{x} \subset \bar{y}$ . Da mit  $x \sim y$  wegen der Symmetrie aber auch  $y \sim x$  gilt, folgt analog auch umgekehrt  $\bar{y} \subset \bar{x}$ , und somit insgesamt  $\bar{x} = \bar{y}$ .

„ $\Leftarrow$ “: Es sei nun  $\bar{x} = \bar{y}$ . Wegen der Reflexivität ist  $x \sim x$ , also  $x \in \bar{x} = \bar{y}$ , und damit  $x \sim y$ .

- (b) Wegen der Reflexivität liegt wie im Beweis von (a) jedes  $x \in M$  in seiner eigenen Äquivalenzklasse  $\bar{x}$ . Liegt  $x$  nun zusätzlich auch in  $\bar{y}$  für ein  $y \in M$ , gilt also  $x \sim y$ , so folgt nach (a) sofort  $\bar{x} = \bar{y}$ . Also waren die beiden Äquivalenzklassen schon gleich, d. h.  $x$  liegt in nur einer Äquivalenzklasse (nämlich in  $\bar{x}$ ).  $\square$

**Aufgabe 2.34.** Es sei  $M = \{n \in \mathbb{Z} : |n| \leq 100\} = \{-100, -99, \dots, 0, \dots, 99, 100\}$ . Welche der folgenden Relationen sind Äquivalenzrelationen auf  $M$ ? Gib im Fall einer Äquivalenzrelation außerdem die Äquivalenzklasse  $\overline{-34}$  explizit an.

- (a)  $x \sim y :\Leftrightarrow$  es gibt ein  $n \in \mathbb{Z}$  mit  $x = 2^n y$ .
- (b)  $x \sim y :\Leftrightarrow xy \geq 0$ .

**Aufgabe 2.35.** Welche der folgenden Relationen sind Äquivalenzrelationen auf  $\mathbb{R}^2$ ? Im Fall einer Äquivalenzrelation berechne und skizziere man außerdem die Äquivalenzklassen von  $(0, 1) \in \mathbb{R}^2$  und  $(1, 1) \in \mathbb{R}^2$ .

- (a)  $(x, y) \sim (x', y') :\Leftrightarrow$  es gibt ein  $a \in \mathbb{R} \setminus \{0\}$  mit  $x = a^2 x'$  und  $y = ay'$ ;
- (b)  $(x, y) \sim (x', y') :\Leftrightarrow$  es gibt ein  $a \in \mathbb{R} \setminus \{0\}$  mit  $x = ay'$  und  $y = ax'$ .

### 3. Erste Eigenschaften der reellen Zahlen

In Notation 1.14 haben wir bereits die reellen Zahlen  $\mathbb{R}$  als „Menge der Punkte auf einer Geraden“ eingeführt. Man kann aber natürlich noch viel mehr Dinge mit den reellen Zahlen tun als sie als eine einfache Punktmenge zu betrachten: Man kann sie addieren, multiplizieren, die Größe von zwei Zahlen miteinander vergleichen, und noch einiges mehr. Wir wollen die Eigenschaften der reellen Zahlen in diesem und dem nächsten Kapitel exakt formalisieren, damit wir danach genau wissen, welche Eigenschaften von  $\mathbb{R}$  wir in dieser Vorlesung axiomatisch voraussetzen. In der Tat werden diese Eigenschaften letztlich sogar ausreichen, um die reellen Zahlen eindeutig zu charakterisieren. Wir beginnen in diesem Kapitel aber zunächst einmal nur mit den „Grundrechenarten“, also mit der Addition und der Multiplikation sowie ihren Umkehrungen, der Subtraktion und Division.

#### 3.A Gruppen und Körper

Die Eigenschaften von Verknüpfungen wie der Addition oder Multiplikation reeller Zahlen werden mathematisch durch die Begriffe einer Gruppe bzw. eines Körpers beschrieben, die wir jetzt einführen wollen.

**Definition 3.1** (Gruppen). Eine **Gruppe** ist eine Menge  $G$  zusammen mit einer „Verknüpfung“, d. h. einer Abbildung

$$*: G \times G \rightarrow G, (x, y) \mapsto x * y,$$

so dass die folgenden Eigenschaften (auch *Gruppenaxiome* genannt) gelten:

- (a) (**Assoziativität**) Für alle  $x, y, z \in G$  gilt  $(x * y) * z = x * (y * z)$ . Man schreibt diesen Ausdruck dann in der Regel auch einfach als  $x * y * z$ , weil die Reihenfolge der Klammerung ja egal ist.
- (b) (Existenz eines neutralen Elements) Es gibt ein  $e \in G$ , für das  $e * x = x * e = x$  für alle  $x \in G$  gilt. Man nennt ein solches  $e$  ein **neutrales Element**, und verlangt davon zusätzlich:
- (c) (Existenz von inversen Elementen) Für alle  $x \in G$  gibt es ein  $x' \in G$  mit  $x' * x = x * x' = e$ . Man nennt  $x'$  dann ein **inverses Element** zu  $x$ .

Wir bezeichnen eine solche Gruppe mit  $(G, *)$ . Wenn aus dem Zusammenhang klar ist, welche Verknüpfung gemeint ist, schreiben wir oft auch einfach nur  $G$  für die Gruppe.

Gilt zusätzlich zu den obigen Eigenschaften noch

- (d) (**Kommutativität**)  $x * y = y * x$  für alle  $x, y \in G$ ,

so heißt  $(G, *)$  eine **kommutative** oder **abelsche Gruppe**.

**Bemerkung 3.2.** Manchmal wird in der Definition einer Gruppe in Teil (b) lediglich  $e * x = x$  und in Teil (c) lediglich  $x' * x = e$  gefordert (man spricht dann auch von einem **linksneutralen** bzw. **linksinversen** Element). Man kann jedoch unter Verwendung der übrigen Gruppenaxiome zeigen, dass in diesem Fall automatisch auch  $x * e = x$  und  $x * x' = e$  gelten muss, also dass linksneutrale Elemente bereits immer neutrale und linksinverse Elemente immer inverse Elemente sind [G, Satz 1.7]. Die beiden Varianten der Definition einer Gruppe stimmen also letztlich überein.

#### Beispiel 3.3.

- (a)  $(\mathbb{R}, +)$  ist eine abelsche Gruppe, denn die Addition ist (wie wir axiomatisch voraussetzen werden) eine Verknüpfung auf  $\mathbb{R}$  mit den Eigenschaften:

- $(x + y) + z = x + (y + z)$  für alle  $x, y, z \in \mathbb{R}$ ;
- $0 \in \mathbb{R}$  ist ein neutrales Element, denn  $0 + x = x + 0 = x$  für alle  $x \in \mathbb{R}$ ;
- zu jedem  $x \in \mathbb{R}$  ist  $-x \in \mathbb{R}$  ein inverses Element, denn  $(-x) + x = x + (-x) = 0$ ;

- $x + y = y + x$  für alle  $x, y \in \mathbb{R}$ .

Auf die gleiche Art sind auch  $(\mathbb{Q}, +)$  und  $(\mathbb{Z}, +)$  abelsche Gruppen, jedoch nicht  $(\mathbb{N}, +)$ : Hier existiert zwar noch ein neutrales Element 0, aber die Zahl  $1 \in \mathbb{N}$  hat kein Inverses mehr, denn es gibt kein  $x \in \mathbb{N}$  mit  $x + 1 = 0$ .

- (b)  $(\mathbb{R}, \cdot)$  ist keine Gruppe: Die Multiplikation ist zwar assoziativ und kommutativ und hat das neutrale Element 1, aber die Zahl 0 hat kein Inverses – denn dies müsste ja eine Zahl  $x \in \mathbb{R}$  sein mit  $x \cdot 0 = 1$ .

Nimmt man jedoch die 0 aus  $\mathbb{R}$  heraus, so erhält man mit  $(\mathbb{R} \setminus \{0\}, \cdot)$  wieder eine abelsche Gruppe, bei der das neutrale Element 1 und das zu einem  $x$  inverse Element  $\frac{1}{x}$  ist. Genauso funktioniert dies für  $(\mathbb{Q} \setminus \{0\}, \cdot)$ , aber z. B. nicht für  $(\mathbb{Z} \setminus \{0\}, \cdot)$ : Hier gibt es zwar noch ein neutrales Element 1, aber die Zahl  $2 \in \mathbb{Z} \setminus \{0\}$  hat kein Inverses mehr, denn es gibt kein  $x \in \mathbb{Z} \setminus \{0\}$  mit  $2 \cdot x = 1$ .

- (c) Hier ist noch ein Beispiel von einem ganz anderen Typ: Es sei  $M$  eine beliebige Menge und

$$G = \{f : f \text{ ist eine bijektive Abbildung von } M \text{ nach } M\}.$$

Da die Verkettung bijektiver Abbildungen nach Aufgabe 2.25 wieder bijektiv ist, definiert sie eine Verknüpfung auf  $G$ . In der Tat wird  $G$  damit zu einer Gruppe, denn die Verkettung ist assoziativ nach Lemma 2.19, die Identität  $\text{id}_M$  ist ein neutrales Element, und zu einem  $f \in G$  ist die Umkehrabbildung  $f^{-1}$  aus Definition 2.20 ein inverses Element: Sie ist nach Aufgabe 2.25 selbst wieder bijektiv (also in  $G$ ) und erfüllt  $f^{-1} \circ f = f \circ f^{-1} = \text{id}_M$  nach Bemerkung 2.21. Im Allgemeinen ist diese Gruppe jedoch nicht kommutativ.

Wir wollen nun ein paar einfache Eigenschaften von Gruppen beweisen, u. a. dass die in Definition 3.1 geforderten neutralen und inversen Elemente eindeutig sind und wir daher in Zukunft auch von dem neutralen und dem zu einem gegebenen Element inversen Element sprechen können.

**Lemma 3.4** (Eigenschaften von Gruppen). *Es seien  $(G, *)$  eine Gruppe und  $x, y \in G$ .*

- Es gibt genau ein neutrales Element (wie in Definition 3.1 (b)).*
- Es gibt genau ein inverses Element zu  $x$  (wie in Definition 3.1 (c)).*
- Sind  $x'$  und  $y'$  die inversen Elemente zu  $x$  bzw.  $y$ , so ist  $y' * x'$  das inverse Element zu  $x * y$ .*
- Ist  $x'$  das inverse Element zu  $x$ , so ist  $x$  das inverse Element zu  $x'$  („das Inverse des Inversen ist wieder das Ausgangselement“).*

*Beweis.*

- (a) Sind  $e$  und  $\tilde{e}$  neutrale Elemente, so folgt

$$\begin{aligned} e &= \tilde{e} * e && \text{(denn } \tilde{e} \text{ ist ein neutrales Element)} \\ &= \tilde{e} && \text{(denn } e \text{ ist ein neutrales Element).} \end{aligned}$$

- (b) Sind  $x'$  und  $\tilde{x}'$  inverse Elemente zu  $x$ , so gilt

$$\begin{aligned} x' &= e * x' && (e \text{ neutrales Element)} \\ &= (\tilde{x}' * x) * x' && (\tilde{x}' \text{ ist ein inverses Element zu } x) \\ &= \tilde{x}' * (x * x') && \text{(Assoziativität)} \\ &= \tilde{x}' * e && (x' \text{ ist ein inverses Element zu } x) \\ &= \tilde{x}' && (e \text{ neutrales Element).} \end{aligned}$$

- (c) Es gilt

$$(y' * x') * (x * y) = y' * (x' * x) * y = y' * e * y = y' * y = e$$

und analog auch  $(x * y) * (y' * x') = e$ . Damit ist  $y' * x'$  das inverse Element zu  $x * y$ .

- (d) Die Gleichung  $x' * x = x * x' = e$  besagt direkt, dass  $x$  das inverse Element zu  $x'$  ist. □

Wie wir in Beispiel 3.3 (a) und (b) gesehen haben, erlauben die reellen Zahlen zwei grundlegende Gruppenstrukturen: die Addition und (nach Herausnahme der 0) die Multiplikation. Diese beiden Strukturen sind jedoch nicht unabhängig voneinander, da sie durch das Distributivgesetz  $(x+y) \cdot z = xz + yz$  für alle  $x, y, z \in \mathbb{R}$  miteinander verbunden sind. Eine derartige Kombination zweier Gruppenstrukturen bezeichnet man als einen Körper.

**Definition 3.5** (Körper). Ein **Körper** ist eine Menge  $K$  zusammen mit zwei Verknüpfungen

$$+ : K \times K \rightarrow K \quad (\text{genannt Addition}) \quad \text{und} \quad \cdot : K \times K \rightarrow K \quad (\text{genannt Multiplikation}),$$

so dass die folgenden Eigenschaften (auch *Körperaxiome* genannt) gelten:

- (a)  $(K, +)$  ist eine abelsche Gruppe. Wir bezeichnen ihr neutrales Element mit 0 und das zu einem  $x \in K$  inverse Element mit  $-x$ .
- (b)  $(K \setminus \{0\}, \cdot)$  ist ebenfalls eine abelsche Gruppe. Wir bezeichnen ihr neutrales Element mit 1 und das zu einem  $x \in K \setminus \{0\}$  inverse Element mit  $x^{-1}$ .
- (c) (**Distributivität**) Für alle  $x, y, z \in K$  gilt  $(x+y) \cdot z = (x \cdot z) + (y \cdot z)$ .

Mit dieser Definition wollen wir nun also axiomatisch voraussetzen:

$$\boxed{\mathbb{R} \text{ ist ein Körper.}}$$

Um Verwirrungen zu vermeiden, werden wir die beiden Verknüpfungen in einem Körper immer mit den Symbolen „+“ und „·“ bezeichnen. Ebenso werden wir (wie ihr es natürlich gewohnt seid) vereinbaren, dass man den Punkt bei der Multiplikation auch weglassen darf und bei ungeklammerten Ausdrücken zuerst die Multiplikationen und dann die Additionen ausgeführt werden, so dass man also z. B. die Distributivität aus Definition 3.5 (c) auch als  $(x+y)z = xz + yz$  schreiben kann.

Es ist jedoch wichtig zu verstehen, dass wir ab jetzt *nicht* mehr voraussetzen werden, dass Addition und Multiplikation in einem Körper wie z. B.  $\mathbb{R}$  genau die Verknüpfungen sind, „an die man als Erstes denken würde“ – was auch immer das heißen mag. Stattdessen sind es einfach irgendwelche zwei Verknüpfungen, die die Eigenschaften aus Definition 3.5 haben. Unsere zukünftigen Beweise über Körper wie z. B.  $\mathbb{R}$  müssen wir also ausschließlich auf diesen Eigenschaften aufbauen.

Dieser axiomatische Zugang hat zwei Vorteile:

- Zum einen wissen wir dadurch genau, welche Eigenschaften der Grundrechenarten auf den reellen Zahlen wir eigentlich voraussetzen. Es sollte schließlich klar sein, dass wir eine *exakte* Mathematik nicht auf einer *anschaulichen* Vorstellung von  $\mathbb{R}$  aufbauen können. Solltet ihr euch also z. B. später einmal dafür interessieren, wie man die Existenz der reellen Zahlen beweisen kann, so wüsstet ihr dann genau, was eigentlich zu beweisen ist: nämlich die Existenz einer Menge mit genau den Eigenschaften, die wir jetzt axiomatisch voraussetzen.
- Zum anderen werdet ihr im Laufe eures Studiums noch viele weitere Körper kennenlernen, z. B. in Kapitel ?? den sehr wichtigen Körper der komplexen Zahlen. Alle Resultate, die nur auf den Körperaxiomen aufbauen, übertragen sich dann also sofort auf diese neuen Fälle, ohne dass man sich darüber noch einmal neu Gedanken machen muss.

### Beispiel 3.6.

- (a) Neben  $\mathbb{R}$  ist auch  $\mathbb{Q}$  (mit den gleichen Verknüpfungen wie auf  $\mathbb{R}$ ) ein Körper. Die ganzen Zahlen  $\mathbb{Z}$  bilden mit diesen Verknüpfungen jedoch keinen Körper, da  $(\mathbb{Z} \setminus \{0\}, \cdot)$  nach Beispiel 3.3 (b) keine Gruppe ist. Ebenso ist  $\mathbb{N}$  mit diesen Verknüpfungen kein Körper, da hier nach Beispiel 3.3 (a) bereits die Addition keine Gruppenstruktur liefert.
- (b) Hier ist ein Beispiel für einen Körper, der sich ganz anders verhält als  $\mathbb{R}$  und  $\mathbb{Q}$ : Wir definieren auf der Menge  $K = \{g, u\}$  zwei Verknüpfungen durch die folgenden Tabellen.

$$\begin{array}{c|cc} + & g & u \\ \hline g & g & u \\ u & u & g \end{array} \quad \text{und} \quad \begin{array}{c|cc} \cdot & g & u \\ \hline g & g & g \\ u & g & u \end{array}$$

Dabei sind  $g$  und  $u$  einfach nur Namen für die beiden Elemente von  $K$ , die für „gerade“ und „ungerade“ stehen sollen und damit auch die Verknüpfungstafeln erklären: Wir haben z. B.  $g + u$  als  $u$  definiert, weil die Addition einer geraden und einer ungeraden Zahl eine ungerade Zahl ergibt.

Man kann zeigen, dass  $K$  mit diesen beiden Verknüpfungen einen Körper bildet. Er wird in der Literatur mit  $\mathbb{Z}_2$  bezeichnet, da seine Elemente die Reste ganzer Zahlen bei Division durch 2 beschreiben. Um zu beweisen, dass  $\mathbb{Z}_2$  ein Körper ist, könnte man z. B. einfach die geforderten Eigenschaften für alle Elemente – es gibt ja nur zwei – explizit nachprüfen. In der Vorlesung „Algebraische Strukturen“ zeigt man allerdings, dass man die Körperaxiome hier auch viel eleganter direkt aus den Eigenschaften von  $\mathbb{Z}$  folgern kann [G, Satz 7.10]. Wir wollen uns hier damit begnügen, die neutralen und inversen Elemente anzugeben:

- Das additive neutrale Element ist  $g$ , wie man leicht aus der Tabelle abliest. Im Sinne der Notationen von Definition 3.5 ist also  $0 = g$ . Wegen  $g + g = u + u = g = 0$  sind die additiven inversen Elemente  $-g = g$  und  $-u = u$ . Dies stimmt natürlich auch mit der Interpretation als gerade und ungerade Zahlen überein, da das Negative von einer geraden bzw. ungeraden Zahl ebenfalls wieder gerade bzw. ungerade ist.
- Das multiplikative neutrale Element in  $\mathbb{Z}_2 \setminus \{0\}$  ist  $u$  – in der Tat ist es ja auch das einzige Element in  $\mathbb{Z}_2 \setminus \{0\}$ . Gemäß der Notation von Definition 3.5 ist also  $1 = u$ .

Beachte, dass in diesem Körper  $\mathbb{Z}_2$  die Gleichung  $1 + 1 = u + u = g = 0$  gilt. Die Körperaxiome lassen es also zu, dass man bei fortgesetzter Addition der 1 irgendwann wieder zur 0 zurück kommt. Wir werden in dieser Vorlesung nicht viel mit dem Körper  $\mathbb{Z}_2$  zu tun haben – wir haben ihn hier nur als Beispiel dafür angegeben, dass die Körperaxiome noch weit davon entfernt sind, die rationalen oder reellen Zahlen eindeutig zu charakterisieren.

Anschaulich kann man die Körperaxiome so interpretieren, dass ein Körper eine Menge ist, auf der „die vier Grundrechenarten existieren und die erwarteten Eigenschaften haben“. Wir wollen nun noch ein paar weitere dieser erwarteten Eigenschaften zeigen, die bereits aus den Körperaxiomen folgen und die wir dann beim Rechnen z. B. in  $\mathbb{R}$  natürlich ständig benutzen werden.

**Bemerkung 3.7.** Es seien  $K$  ein Körper und  $x, y \in K$ .

- (a) Wenden wir Lemma 3.4 (c) und (d) auf die (kommutative) Addition und Multiplikation an, so sehen wir sofort, dass

$$-(x + y) = (-x) + (-y) \quad \text{und} \quad -(-x) = x$$

sowie für  $x, y \neq 0$

$$(xy)^{-1} = x^{-1} \cdot y^{-1} \quad \text{und} \quad (x^{-1})^{-1} = x.$$

- (b) Etwas versteckt in Definition 3.5 steht in Teil (b) u. a. die Aussage, dass die Multiplikation überhaupt eine Verknüpfung auf  $K \setminus \{0\}$  ist, also dass für  $x, y \in K \setminus \{0\}$  auch  $xy \in K \setminus \{0\}$  gilt. Äquivalent dazu bedeutet das:

$$\text{Ist } xy = 0, \text{ so gilt } x = 0 \text{ oder } y = 0.$$

**Lemma 3.8** (Eigenschaften von Körpern). *In jedem Körper  $K$  gilt für alle  $x, y \in K$ :*

- (a)  $0 \cdot x = 0$ .  
 (b)  $x \cdot (-y) = -(xy)$ .  
 (c) Für  $x \neq 0$  ist  $-(x^{-1}) = (-x)^{-1}$ .

*Beweis.*

- (a) Es gilt

$$\begin{aligned} 0 \cdot x &= (0 + 0) \cdot x && (0 \text{ ist additives neutrales Element}) \\ &= 0 \cdot x + 0 \cdot x, && (\text{Distributivität}) \end{aligned}$$

woraus durch Addition des additiven Inversen von  $0 \cdot x$  auf beiden Seiten die gewünschte Gleichung  $0 = 0 \cdot x$  folgt.

(b) Es ist

$$\begin{aligned} x \cdot (-y) + xy &= x \cdot (-y + y) && \text{(Distributivität)} \\ &= x \cdot 0 && \text{(-y ist additives Inverses zu y)} \\ &= 0 && \text{(nach (a)),} \end{aligned}$$

daher ist  $x \cdot (-y)$  das additive Inverse zu  $xy$ , d. h. es gilt  $x \cdot (-y) = -(xy)$ .

(c) Doppelttes Anwenden von (b), einmal für den linken und einmal für den rechten Faktor, ergibt

$$(-(x^{-1})) \cdot (-x) = -(x^{-1} \cdot (-x)) = -(-(x^{-1} \cdot x)) = -(-1) \stackrel{3.7(a)}{=} 1.$$

Also ist  $-(x^{-1})$  das multiplikative Inverse zu  $-x$ , d. h. es ist  $-(x^{-1}) = (-x)^{-1}$ .  $\square$

**Notation 3.9.** In einem Körper  $K$  verwendet man üblicherweise die folgenden Notationen, von denen euch die meisten sicher bekannt sein werden:

- (a) Für  $x, y \in K$  setzt man  $x - y := x + (-y)$ . Ist  $y \neq 0$ , so setzt man  $\frac{x}{y} := x \cdot y^{-1}$ .
- (b) Für  $x \in K$  und  $n \in \mathbb{N}$  definiert man die  $n$ -te **Potenz** von  $x$  als

$$x^n := \underbrace{x \cdot \cdots \cdot x}_{n\text{-mal}},$$

wobei dieser Ausdruck für  $n = 0$  als  $x^0 := 1$  zu verstehen ist. Insbesondere legen wir also auch  $0^0 := 1$  fest. Beachte, dass aus dieser Definition (und der Kommutativität der Multiplikation) unmittelbar die Potenzrechenregeln

$$x^m \cdot x^n = x^{m+n} \quad \text{und} \quad (xy)^n = x^n \cdot y^n$$

für alle  $x, y \in K$  folgen. Ist  $x \neq 0$ , so definiert man zusätzlich Potenzen mit negativen ganzzahligen Exponenten durch  $x^{-n} := (x^{-1})^n$ .

Beachte, dass auch in einem beliebigen Körper  $K$  die Exponenten einer Potenz stets *ganze Zahlen* sind und keine Elemente aus  $K$ . Eine Potenz  $x^y$  für  $x, y \in K$  lässt sich im Allgemeinen nicht definieren (auch wenn dies für  $K = \mathbb{R}$  in vielen Fällen möglich ist, siehe Definition ??).

(c) Manchmal möchte man mehrere Elemente  $x_m, x_{m+1}, x_{m+2}, \dots, x_n$  in einem Körper (oder allgemeiner in einer additiv geschriebenen abelschen Gruppe) aufsummieren, die durch eine ganzzahlige Laufvariable indiziert werden, die von einem  $m \in \mathbb{Z}$  bis zu einem  $n \in \mathbb{Z}$  (mit  $n \geq m$ ) läuft. Man schreibt dies dann als

$$\sum_{i=m}^n x_i := x_m + x_{m+1} + x_{m+2} + \cdots + x_n$$

(also mit einem großen griechischen Sigma, das an das Wort „Summe“ erinnern soll). So steht z. B.

$$\sum_{i=1}^n i^2 = 1^2 + 2^2 + 3^2 + \cdots + n^2 \quad (*)$$

für die Summe aller Quadratzahlen bis  $n^2$ . Natürlich ist der Name der Laufvariablen dabei egal, und der Ausdruck (\*) hängt nicht von einem  $i$  ab (wie man auf der rechten Seite ja auch sieht). Außerdem kann man die Laufvariable verschieben, ohne den eigentlichen Ausdruck zu ändern: Setzt man z. B.  $i = j + 1$ , also  $j = i - 1$ , in der obigen Summe (\*), so läuft  $j$  dort von 0 bis  $n - 1$ , wenn  $i$  von 1 bis  $n$  läuft, und wir können dieselbe Summe auch schreiben als

$$\sum_{j=0}^{n-1} (j+1)^2 = 1^2 + 2^2 + 3^2 + \cdots + n^2.$$

Natürlich kann man diesen Ausdruck nun auch wieder genauso gut mit dem Buchstaben  $i$  statt  $j$  als  $\sum_{i=0}^{n-1} (i+1)^2$  schreiben, oder den Index um mehr als 1 in die eine oder andere Richtung verschieben. Also:

Der Wert einer Summe ändert sich nicht, wenn man zur Laufvariablen im zu summierenden Ausdruck eine Konstante addiert, und dafür von der Ober- und Untergrenze der Summe diese Konstante abzieht.

Wir sagen in diesem Fall, dass die neue Darstellung der Summe durch eine **Indexverschiebung** (im Beispiel oben  $i \mapsto i+1$ ) aus der alten hervorgeht.

04

Analog schreibt man

$$\prod_{i=m}^n x_i := x_m \cdot x_{m+1} \cdot x_{m+2} \cdot \cdots \cdot x_n$$

(mit einem großen griechischen Pi für das Produkt), wenn man die Körperelemente multiplizieren statt addieren möchte. Ist schließlich die Obergrenze einer Summe oder eines Produkts kleiner als die Untergrenze (man spricht dann von der **leeren Summe** bzw. dem **leeren Produkt**), so definiert man dies als

$$\sum_{i=m}^n x_i := 0 \quad \text{und} \quad \prod_{i=m}^n x_i := 1 \quad \text{für } n < m,$$

also als das additive bzw. multiplikative neutrale Element.

(d) Ist  $n$  eine natürliche Zahl, so fasst man diese oft auch als das Element

$$\sum_{i=1}^n 1 = \underbrace{1 + \cdots + 1}_{n\text{-mal}}$$

von  $K$  auf. Im Fall  $K = \mathbb{R}$  ist dies dann einfach die natürliche Zahl  $n \in \mathbb{N} \subset \mathbb{R}$  und liefert somit keine neue Notation, aber z. B. in  $K = \mathbb{Z}_2$  aus Beispiel 3.6 (b) ist  $2 = 1 + 1 = 0$ .

**Aufgabe 3.10.** Zeige, dass in jedem Körper  $K$  die üblichen Rechenregeln

$$\frac{x}{y} + \frac{z}{w} = \frac{xw + yz}{yw} \quad \text{und} \quad \frac{x}{y} \cdot \frac{z}{w} = \frac{xz}{yw}$$

für Brüche gelten, wobei  $x, y, z, w \in K$  mit  $y, w \neq 0$ .

**Aufgabe 3.11.** Es sei  $a \in \mathbb{R}$  fest gegeben. Wir definieren auf  $\mathbb{R}^2$  eine „Addition“ und „Multiplikation“ durch

$$(x_1, x_2) + (y_1, y_2) := (x_1 + y_1, x_2 + y_2) \quad \text{und} \quad (x_1, x_2) \cdot (y_1, y_2) := (x_1 y_1 + a x_2 y_2, x_1 y_2 + x_2 y_1).$$

Man prüft leicht durch explizite Rechnung nach, dass  $\mathbb{R}^2$  mit dieser Addition eine kommutative Gruppe mit neutralem Element  $(0, 0)$  ist, dass auch die Multiplikation kommutativ ist, und dass diese beiden Operationen das Distributivgesetz erfüllen – ihr solltet euch kurz überlegen, warum das so ist, braucht das aber nicht aufzuschreiben. Man zeige stattdessen:

- (a) Die Multiplikation ist assoziativ und besitzt ein neutrales Element.
- (b) Im Fall  $a = -1$  ist  $(\mathbb{R}^2, +, \cdot)$  ein Körper, im Fall  $a = 1$  jedoch nicht.  
(Für  $a = -1$  ist  $(\mathbb{R}^2, +, \cdot)$  der sogenannte Körper der komplexen Zahlen, den wir in Kapitel ?? noch genau untersuchen werden.)

**Aufgabe 3.12.** Zu einem Körper  $K$  und einer Menge  $D$  mit  $|D| \geq 2$  sei

$$V = \{f : f \text{ ist eine Abbildung von } D \text{ nach } K\}$$

die Menge aller reellwertigen Funktionen auf  $D$ . Für  $f, g \in V$  definieren wir die Addition  $f + g$  und Multiplikation  $f \cdot g$  dieser Funktionen punktweise durch

$$f + g : D \rightarrow K, x \mapsto f(x) + g(x) \quad \text{und} \quad f \cdot g : D \rightarrow K, x \mapsto f(x) \cdot g(x).$$

- (a) Zeige, dass  $V$  mit dieser Addition eine abelsche Gruppe ist.  
 (b) Ist  $V$  mit dieser Addition und Multiplikation ein Körper?

### 3.B Vollständige Induktion

Häufig möchte man in der Mathematik Aussagen beweisen, die von einer natürlichen Zahl abhängen – z. B. bei Formeln, die Summen oder Produkte wie in Notation 3.9 mit variablen Unter- oder Obergrenzen beinhalten. Die einfachste und bekannteste solcher Aussagen ist vermutlich die folgende Formel für die Summe aller natürlichen Zahlen bis zu einer gegebenen Obergrenze.

**Satz 3.13 (Summenformel von Gauß).** Für alle  $n \in \mathbb{N}$  gilt

$$\sum_{k=1}^n k = \frac{n(n+1)}{2}.$$

**Beispiel 3.14.** Für  $n = 5$  ist z. B.

$$\sum_{k=1}^5 k = 1 + 2 + 3 + 4 + 5 = 15 = \frac{5 \cdot 6}{2}.$$

Um derartige Aussagen zu beweisen, ist oft das Beweisverfahren der (**vollständigen**) **Induktion** nützlich, das wir jetzt einführen wollen.

Angenommen, wir wollen (wie z. B. in Satz 3.13) eine Aussage  $A(n)$  für alle  $n \in \mathbb{N}$  beweisen. Dann können wir dies tun, indem wir die folgenden beiden Dinge zeigen:

- (a) (**Induktionsanfang**) Die Aussage  $A(0)$  ist wahr.  
 (b) (**Induktionsschritt**) Für alle  $n \in \mathbb{N}$  gilt  $A(n) \Rightarrow A(n+1)$ , d. h. wenn die Aussage  $A(n)$  für ein gegebenes  $n \in \mathbb{N}$  gilt (die „Induktionsannahme“ bzw. „Induktionsvoraussetzung“), dann gilt auch die Aussage  $A(n+1)$  (der „Induktionsschluss“).

Haben wir diese beiden Dinge gezeigt, so folgt daraus nämlich die Gültigkeit von  $A(n)$  für alle  $n \in \mathbb{N}$ : Die Aussage  $A(0)$  haben wir mit dem Induktionsanfang gezeigt, und durch fortgesetztes Anwenden des Induktionsschritts  $A(n) \Rightarrow A(n+1)$  für  $n = 0, 1, 2, \dots$  erhalten wir dann auch

$$A(0) \Rightarrow A(1) \Rightarrow A(2) \Rightarrow A(3) \Rightarrow \dots,$$

also die Gültigkeit von  $A(n)$  für alle  $n \in \mathbb{N} = \{0, 1, 2, 3, \dots\}$ .

Derartige Induktionsbeweise sind immer dann sinnvoll, wenn die Aussagen  $A(n)$  und  $A(n+1)$  „ähnlich genug“ sind, so dass es beim Beweis von  $A(n+1)$  hilft, die Gültigkeit von  $A(n)$  voraussetzen zu dürfen.

Mit diesem Verfahren können wir nun die Summenformel aus Satz 3.13 beweisen:

*Beweis von Satz 3.13.* Wir zeigen die Formel mit Induktion über  $n$ .

*Induktionsanfang ( $n = 0$ ):* Für  $n = 0$  stimmen die beiden Seiten der zu zeigenden Gleichung überein, denn es ist

$$\sum_{k=1}^0 k = 0 = \frac{0 \cdot (0+1)}{2}.$$

*Induktionsschritt ( $n \rightarrow n+1$ ):* Als Induktionsvoraussetzung nehmen wir an, dass die zu beweisende Formel für ein gegebenes  $n \in \mathbb{N}$  richtig ist, d. h. dass

$$\sum_{k=1}^n k = \frac{n(n+1)}{2}$$

gilt. (Beachte, dass wir diese Gleichung nicht für alle  $n \in \mathbb{N}$  voraussetzen – dies wäre ja schon die gesamte zu zeigende Aussage!) Wir müssen zeigen, dass die entsprechende Gleichung dann auch für  $n+1$  gilt, also dass

$$\sum_{k=1}^{n+1} k = \frac{(n+1)(n+2)}{2}.$$

Dies ergibt sich nun leicht aus der folgenden Rechnung:

$$\begin{aligned} \sum_{k=1}^{n+1} k &= \sum_{k=1}^n k + (n+1) && \text{(Abspalten des letzten Summanden für } k = n+1) \\ &= \frac{n(n+1)}{2} + (n+1) && \text{(nach Induktionsvoraussetzung)} \\ &= (n+1) \cdot \left(\frac{n}{2} + 1\right) \\ &= \frac{(n+1)(n+2)}{2}. \end{aligned}$$

Damit ist der Satz mit vollständiger Induktion bewiesen.  $\square$

**Bemerkung 3.15.** Offensichtlich erlaubt das Beweisverfahren der vollständigen Induktion die folgenden Abwandlungen:

- (a) Im Induktionsschritt kann man, wenn es hilfreich ist, beim Beweis der Aussage  $A(n+1)$  nicht nur die direkt vorangegangene Aussage  $A(n)$ , sondern *alle bereits gezeigten Aussagen*  $A(0), A(1), \dots, A(n)$  voraussetzen.
- (b) Möchte man die Aussage  $A(n)$  nicht für alle  $n \in \mathbb{N}$ , sondern für alle  $n \in \mathbb{Z}$  ab einem gewissen Startwert  $n_0 \in \mathbb{Z}$  zeigen, so kann man als Induktionsanfang die Aussage  $A(n_0)$  zeigen, und im Induktionsschritt dann die Folgerung  $A(n) \Rightarrow A(n+1)$  für alle  $n \geq n_0$ .

**Aufgabe 3.16.** Zeige für alle  $n \in \mathbb{N}$ :

$$(a) \sum_{k=1}^n \frac{1}{k(k+1)} = \frac{n}{n+1}, \quad (b) \prod_{k=1}^n \left(1 + \frac{1}{n+k}\right) = 2 - \frac{1}{n+1}.$$

**Aufgabe 3.17.** Zeige mit vollständiger Induktion: Ist  $a \in \mathbb{R} \setminus \{0\}$  mit  $a + \frac{1}{a} \in \mathbb{Z}$ , so gilt für alle  $n \in \mathbb{N}$  auch  $a^n + \frac{1}{a^n} \in \mathbb{Z}$ .

### 3.C Polynomfunktionen

Als erste Anwendung der Körpereigenschaften wollen wir zum Abschluss dieses Kapitels die euch sicher schon aus der Schule bekannten Polynomfunktionen behandeln – also die Funktionen, die sich aus den grundlegenden Körperoperationen Addition und Multiplikation bilden lassen.

**Definition 3.18** (Polynomfunktionen und Nullstellen). Es seien  $D$  eine Teilmenge eines Körpers  $K$  und  $f: D \rightarrow K$  eine Funktion.

- (a) Ist  $f$  von der Form

$$f(x) = \sum_{k=0}^n a_k x^k = a_n x^n + \dots + a_1 x + a_0 \quad \text{für alle } x \in D$$

mit gewissen  $a_0, \dots, a_n \in K$ , so sagt man, dass  $f$  eine **Polynomfunktion** mit **Koeffizienten**  $a_0, \dots, a_n$  ist. Ist  $n$  dabei so gewählt, dass der erste Koeffizient  $a_n$  ungleich Null ist, so heißt  $f$  eine Polynomfunktion vom **Grad**  $n$  und mit **Leitkoeffizient**  $a_n$ . Ist der Leitkoeffizient 1, so heißt  $f$  eine **normierte** Polynomfunktion.

Sind in der obigen Darstellung alle Koeffizienten  $a_0, \dots, a_n$  gleich 0 (und ist  $f$  damit die Nullfunktion), so nennen wir  $f$  formal eine Polynomfunktion vom Grad  $-\infty$ . In diesem Fall hat  $f$  keinen Leitkoeffizienten.

- (b) Ist  $x_0 \in D$  mit  $f(x_0) = 0$ , so nennt man  $x_0$  eine **Nullstelle** von  $f$ .

Das Besondere an Nullstellen von Polynomfunktionen ist, dass man sie wie im folgenden Satz als Linearfaktoren abspalten kann.

**Satz 3.19** (Abspalten von Nullstellen in Polynomfunktionen). *Es seien  $K$  ein Körper,  $D \subset K$  und  $f: D \rightarrow K$  eine Polynomfunktion vom Grad  $n \in \mathbb{N}$ .*

- (a) Ist  $x_0 \in D$  eine Nullstelle von  $f$ , so gibt es eine Polynomfunktion  $g: D \rightarrow K$  vom Grad  $n - 1$  mit  $f(x) = (x - x_0)g(x)$  für alle  $x \in D$  (d. h. man kann „den Linearfaktor  $x - x_0$  abspalten“).
- (b) Die Funktion  $f$  hat höchstens  $n$  Nullstellen.

*Beweis.* Wir zeigen die beiden Aussagen mit Induktion über  $n$ . Der Beweis von (a) ist dabei konstruktiv, d. h. er gibt auch ein Verfahren an, wie  $g$  berechnet werden kann (siehe Beispiel 3.20).

Der Induktionsanfang für  $n = 0$  ist trivial, denn  $f$  ist dann eine Konstante ungleich 0 und hat somit keine Nullstellen. Für den Induktionsschluss nehmen wir an, dass die Aussagen des Satzes bis zu einem gegebenen  $n$  gelten, und betrachten  $f: D \rightarrow K, x \mapsto a_{n+1}x^{n+1} + \dots + a_1x + a_0$  vom Grad  $n + 1$ , also mit  $a_{n+1} \neq 0$ .

- (a) Wir definieren eine Polynomfunktion  $\tilde{f}: D \rightarrow K$  durch

$$\begin{aligned} \tilde{f}(x) &:= f(x) - a_{n+1}x^n(x - x_0) \\ &= a_{n+1}x_0x^n + a_nx^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0 \end{aligned}$$

für alle  $x$ . Ist  $\tilde{f}$  die Nullfunktion, so sind wir fertig, da dann ja  $f(x) = a_{n+1}x^n(x - x_0)$  für alle  $x \in D$  gilt. Andernfalls ist  $\tilde{f}$  nach Konstruktion eine Polynomfunktion vom Grad höchstens  $n$  (der  $x^{n+1}$ -Term hebt sich ja gerade heraus), die immer noch die Nullstelle  $x_0$  hat. Nach Induktionsvoraussetzung gibt es dann also eine Polynomfunktion  $\tilde{g}: D \rightarrow K$  vom Grad höchstens  $n - 1$  mit  $\tilde{f}(x) = (x - x_0)\tilde{g}(x)$  für alle  $x \in D$ , und somit ist

$$f(x) = a_{n+1}x^n(x - x_0) + \tilde{f}(x) = (x - x_0) \cdot \underbrace{(a_{n+1}x^n + \tilde{g}(x))}_{=:g(x)}$$

für alle  $x \in D$ , wobei  $g$  offensichtlich vom Grad  $n$  ist.

- (b) Hat  $f$  keine Nullstelle, so sind wir fertig. Andernfalls wählen wir eine Nullstelle  $x_0$  von  $f$  und schreiben  $f(x) = (x - x_0)g(x)$  für alle  $x \in D$  wie in (a) mit einer Polynomfunktion  $g$  vom Grad  $n$ . Nach Induktionsvoraussetzung hat  $g$  höchstens  $n$  Nullstellen, und nach Bemerkung 3.7 (b) sind die Nullstellen von  $f$  genau  $x_0$  zusammen mit den Nullstellen von  $g$ . Also hat  $f$  höchstens  $n + 1$  Nullstellen. Damit ist die Behauptung mit Induktion bewiesen.  $\square$

**Beispiel 3.20** (Polynomdivision). Das Verfahren aus dem Beweis von Satz 3.19 (a) wird als *Polynomdivision* [G, Satz 10.19] bezeichnet: Man subtrahiert fortlaufend geeignete Vielfache von  $x - x_0$  von  $f$ , so dass sich der jeweils höchste Term von  $f$  weghebt, und sammelt die dabei verwendeten Faktoren in  $g$ . Das folgende Schema, das genauso aussieht wie eine normale schriftliche Division, verdeutlicht dieses Verfahren am Beispiel der Funktion  $f: \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x^2 + 3x - 4$  mit Nullstelle  $x_0 = 1$ , die wir als  $f(x) = (x - 1)g(x)$  für alle  $x \in \mathbb{R}$  schreiben wollen. Das Ergebnis ist in diesem Fall  $g: \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x + 4$ .

$$\begin{array}{r} f(x) \longrightarrow (x^2 + 3x - 4) : (x - 1) = x + 4 \longleftarrow g(x) \\ \quad - (x^2 - x) \longleftarrow \cdot x \\ \hline \tilde{f}(x) \longrightarrow (4x - 4) \\ \quad - (4x - 4) \longleftarrow \cdot 4 \\ \hline 0 \end{array}$$

**Bemerkung 3.21.** Satz 3.19 liefert uns zwar die neue Funktion nach dem Abspalten des Linearfaktors, er sagt uns hingegen nicht, wie wir überhaupt erst einmal eine Nullstelle von  $f$  finden können, oder ob es überhaupt Nullstellen gibt (die reelle Polynomfunktion  $f(x) = x^2 + 1$  hat ja z. B. keine Nullstellen). In der Tat gibt es im Allgemeinen kein Verfahren, wie man Nullstellen von Polynomfunktionen exakt berechnen kann! Genauer gesagt gilt:

- Für Polynomfunktionen vom Grad höchstens 4 gibt es explizite Verfahren zur exakten Bestimmung der Nullstellen (für Grad 1 ist das klar, für Grad 2 gibt es die bekannte „ $p$ - $q$ -Formel“ bzw. die quadratische Ergänzung, und für Grad 3 bzw. 4 sind die Formeln so lang, dass man im Allgemeinen nicht mehr mit ihnen arbeiten möchte).

- Für Polynomfunktionen vom Grad größer als 4 kann man beweisen(!), dass es keine derartigen Verfahren zur exakten Bestimmung der Nullstellen geben kann (das beweist man z. B. in der Vorlesung „Einführung in die Algebra“, die ihr im nächsten Studienjahr hören könnt). Aber:
- Für reelle Polynomfunktionen beliebigen Grades gibt es zumindest numerische Verfahren, die die Nullstellen (mit beliebiger Genauigkeit) näherungsweise bestimmen können.

Zum Schluss wollen wir nun noch zwei wichtige Konzepte für Polynomfunktionen untersuchen, die ihr beide im reellen Fall vielleicht schon aus der Schule kennt: den sogenannten Koeffizientenvergleich (also dass eine Polynomfunktion eindeutig ihre Koeffizienten bestimmt) und die Vielfachheit von Nullstellen. Es stellt sich jedoch heraus, dass man hierfür im allgemeinen Fall die Voraussetzung benötigt, dass die Definitionsmenge der betrachteten Funktionen unendlich viele Elemente besitzt.

**Lemma 3.22 (Koeffizientenvergleich).** *Es seien  $K$  ein Körper,  $D \subset K$  mit  $|D| = \infty$ , und  $f: D \rightarrow K$  eine Polynomfunktion mit zwei Darstellungen*

$$f(x) = a_n x^n + \cdots + a_1 x + a_0 = b_n x^n + \cdots + b_1 x + b_0 \quad \text{für alle } x \in D$$

für gewisse  $a_0, \dots, a_n, b_0, \dots, b_n \in K$ . (Beachte, dass wir dabei in beiden Darstellungen den gleichen höchsten Exponenten  $n$  wählen können, da wir nicht  $a_n \neq 0$  und  $b_n \neq 0$  vorausgesetzt haben.)

Dann gilt bereits  $a_i = b_i$  für alle  $i = 0, \dots, n$ . Es ist also nicht möglich, „eine Polynomfunktion auf zwei verschiedene Arten hinzuschreiben“.

*Beweis.* Nach Voraussetzung ist die Polynomfunktion

$$D \rightarrow K, x \mapsto (a_n - b_n)x^n + \cdots + (a_1 - b_1)x + (a_0 - b_0) = f(x) - f(x) = 0$$

die Nullfunktion auf  $D$ . Da sie damit wegen  $|D| = \infty$  unendlich viele Nullstellen besitzt, muss sie nach Satz 3.19 (b) vom Grad  $-\infty$  sein. Also sind alle Koeffizienten dieser Polynomfunktion gleich 0, d. h. es ist  $a_i = b_i$  für alle  $i = 0, \dots, n$ .  $\square$

**Bemerkung und Notation 3.23 (Polynome).** Die Voraussetzung  $|D| = \infty$  in Lemma 3.22 ist wirklich notwendig: So sind für  $D = K = \mathbb{Z}_2$  wie in Beispiel 3.6 (b) z. B.  $x \mapsto x$  und  $x \mapsto x^2$  dieselbe Funktion, da sie beide 0 auf 0 und 1 auf 1 abbilden und in  $\mathbb{Z}_2$  keine weiteren Elemente existieren.

In der Literatur bezeichnet man einen formalen Ausdruck der Form  $a_n x^n + \cdots + a_1 x + a_0$  mit  $a_0, \dots, a_n \in K$  als ein **Polynom** über  $K$  [G, Kapitel 9]. Jedes solche Polynom bestimmt natürlich eine Polynomfunktion von jeder Teilmenge  $D$  von  $K$  nach  $K$ , allerdings können verschiedene Polynome wie im eben angegebenen Beispiel durchaus dieselbe Polynomfunktion definieren: Über  $\mathbb{Z}_2$  sind  $x$  und  $x^2$  verschiedene Polynome, sie bestimmen aber dieselbe Polynomfunktion.

Mit dieser Notation ist die Aussage von Lemma 3.22 also, dass Polynome und Polynomfunktionen im Fall von unendlichen Definitionsmengen dasselbe sind. Da wir Polynomfunktionen im Folgenden in der Regel nur in diesem Fall unendlicher Definitionsmengen benötigen, werden wir die Begriffe Polynom und Polynomfunktion oft synonym verwenden. Wegen der Eindeutigkeit der Koeffizienten sind dann auch der Grad (und der Leitkoeffizient) einer Polynomfunktion  $f$  wie in Definition 3.18 (a) eindeutig bestimmt. Wir können daher eine Bezeichnung dafür einführen:

**Definition 3.24 (Grad eines Polynoms).** Wir bezeichnen den **Grad** einer Polynomfunktion  $f$  (mit unendlicher Definitionsmenge) mit  $\deg f \in \mathbb{N} \cup \{-\infty\}$  (vom englischen Wort „degree“).

In den Fällen  $\deg f = 1$  bzw.  $\deg f = 2$  nennt man  $f$  ein **lineares** bzw. **quadratisches Polynom**.

**Satz und Definition 3.25 (Vielfachheit von Nullstellen).** *Es seien  $K$  ein Körper,  $D \subset K$  mit  $|D| = \infty$  und  $f: D \rightarrow K$  eine Polynomfunktion, die nicht die Nullfunktion ist. Dann lässt sich  $f$  (bis auf die Reihenfolge der Faktoren) eindeutig als*

$$f(x) = g(x) \cdot (x - x_1)^{a_1} \cdot \cdots \cdot (x - x_k)^{a_k} \quad \text{für alle } x \in D$$

schreiben, wobei  $x_1, \dots, x_k \in D$  die verschiedenen Nullstellen von  $f$  sind,  $a_1, \dots, a_k \in \mathbb{N}_{>0}$  gilt, und  $g$  ein Polynom ohne Nullstellen in  $D$  ist. In dieser Darstellung nennt man  $a_i$  für  $i = 1, \dots, k$  die

**Vielfachheit der Nullstelle  $x_i$  von  $f$**  (in der Literatur sind auch die Bezeichnungen **Ordnung** und **Multiplizität** der Nullstelle üblich).

*Beweis.* Die Existenz einer solchen Darstellung ergibt sich sofort durch fortgesetztes Abspalten von Nullstellen gemäß Satz 3.19 (a). Wir zeigen nun die Eindeutigkeit mit Induktion über den Grad  $n := \deg f$  des Polynoms. Dabei ist der Induktionsanfang für  $n = 0$  trivial, denn dann hat  $f$  keine Nullstellen, und es ist zwangsläufig  $k = 0$  und  $g = f$ .

Für den Induktionsschritt  $n \rightarrow n + 1$  bemerken wir zuerst, dass  $x_1, \dots, x_k$  natürlich in jedem Fall als die Nullstellen von  $f$  eindeutig bestimmt sind. Wir nehmen also an, dass wir zwei Darstellungen

$$f(x) = g(x) \cdot (x - x_1)^{a_1} \cdot \dots \cdot (x - x_k)^{a_k} = h(x) \cdot (x - x_1)^{b_1} \cdot \dots \cdot (x - x_k)^{b_k}$$

eines Polynoms vom Grad  $n + 1$  wie in der Behauptung des Satzes haben. Im nullstellenfreien Fall  $k = 0$  sind wir natürlich bereits fertig. Andernfalls liefert Division durch  $x - x_1$  für alle  $x \in D \setminus \{x_1\}$  (wir müssen  $x_1$  hier herausnehmen, da wir sonst durch 0 teilen würden!)

$$\begin{aligned} &g(x) \cdot (x - x_1)^{a_1 - 1} \cdot (x - x_2)^{a_2} \cdot \dots \cdot (x - x_k)^{a_k} \\ &= h(x) \cdot (x - x_1)^{b_1 - 1} \cdot (x - x_2)^{b_2} \cdot \dots \cdot (x - x_k)^{b_k}. \end{aligned} \quad (*)$$

Wir haben also wieder zwei Darstellungen eines Polynoms auf der immer noch unendlichen Menge  $D \setminus \{x_1\}$ . Da der Grad dieses Polynoms  $n$  ist, müssen diese Darstellungen aber nach der Induktionsvoraussetzung bereits übereinstimmen. Also gilt  $g = h$ ,  $a_1 - 1 = b_1 - 1$ ,  $a_2 = b_2$ ,  $\dots$ ,  $a_k = b_k$ , und damit stimmen auch die beiden ursprünglichen Darstellungen von  $f$  überein.  $\square$

**Aufgabe 3.26.** Bestimme die Nullstellen des reellen Polynoms  $x^4 + 3x^3 - 4x$  und ihre Vielfachheiten.

**Aufgabe 3.27.** Es sei  $f$  ein reelles Polynom mit  $f(x) = x^3$  für alle  $x \in \{1, 2, 3, \dots, 8\}$ . Welchen Grad kann  $f$  haben?

**Aufgabe 3.28.** Es sei  $f$  das reelle Polynom mit  $f(x) = (x^2 - x + 1)^{2023}$  für alle  $x \in \mathbb{R}$ .

- (a) Bestimme die Summe aller Koeffizienten von  $f$ .
- (b) Bestimme die Summe aller Koeffizienten von geraden Potenzen von  $x$  in  $f$ .

**Aufgabe 3.29.**

- (a) Bestimme alle reellen Polynome  $f$  mit  $xf(x+1) = (x-1)f(x)$  für alle  $x \in \mathbb{R}$ .
- (b) Bestimme alle reellen Polynome  $f$  mit  $xf(x-1) = (x-1)f(x)$  für alle  $x \in \mathbb{R}$ .

# Grundlagen der Mathematik 1: Lineare Algebra

## 13. Vektorräume

Ausgehend von den elementaren Konzepten in den Kapiteln 1 bis 3 wollen wir in dieser Vorlesung zwei grundlegende Gebiete der Mathematik entwickeln: die Analysis und die lineare Algebra. Während sich dabei die eindimensionale Analysis in den Kapiteln ?? bis ?? hauptsächlich mit allgemeinen (in der Regel stetigen oder sogar differenzierbaren) Funktionen in einer reellen Variablen beschäftigt hat, wollen wir nun im folgenden Teil des Skripts zunächst unabhängig davon *lineare* Funktionen in *mehreren* Variablen studieren, wie sie in der Praxis z. B. in Form von linearen Gleichungssystemen auftreten. Später werden wir die erarbeiteten Resultate dann mit der Analysis kombinieren, um Funktionen in mehreren Variablen mit Hilfe von Ableitungen linear approximieren zu können.

In der Analysis arbeitet man ja hauptsächlich über den reellen Zahlen, und das ist in der linearen Algebra auch nicht viel anders. Allerdings haben wir in Kapitel 3 ja auch schon gesehen, dass viele Dinge auch über einem beliebigen Körper funktionieren. Die lineare Algebra verhält sich hier sehr „gutartig“: Da wir letztlich nur lineare Funktionen bzw. Gleichungen betrachten werden, benötigen wir gar nicht mehr Struktur der reellen Zahlen als die Körperaxiome. Wir können daher nahezu die gesamte lineare Algebra über einem beliebigen Grundkörper studieren, also z. B. auch über  $\mathbb{Q}$ , dem Körper  $\mathbb{Z}_2$  aus Beispiel 3.6 (b), oder anderen Körpern, die ihr vielleicht inzwischen kennt oder noch kennenlernen werdet. Wir vereinbaren daher:

Im Folgenden (bis zum Ende von Kapitel ??) sei  $K$  immer ein fest gewählter Grundkörper.

Beim ersten Lesen könnt ihr euch  $K$  aber auch gerne einfach als  $\mathbb{R}$  vorstellen.

### 13.A Der Vektorraumbegriff

Wie ihr ja sicher aus der Schule wisst, werden die Elemente von  $\mathbb{R}^2$  oder  $\mathbb{R}^3$  in der Regel *Vektoren* genannt. Aber was genau ist im Allgemeinen eigentlich ein Vektor? Genau wie bei Gruppen und Körpern in Abschnitt 3.A werden Vektoren über die Operationen definiert, die man mit ihnen durchführen kann: In einer Gruppe gibt es *eine* Verknüpfung, die gewisse Eigenschaften erfüllt (siehe Definition 3.1), in einem Körper *zwei* Verknüpfungen „+“ und „·“ mit den erwarteten Eigenschaften (siehe Definition 3.5). Was sind nun die analogen definierenden Verknüpfungen und Eigenschaften für Vektoren? Wir wissen alle aus der Schule, dass man Vektoren addieren und „strecken“, also mit einer reellen Zahl (bzw. mit einer Zahl des gewählten Grundkörpers  $K$ ) multiplizieren kann. Genau diese beiden Strukturen definieren einen allgemeinen Vektorraum:

**Definition 13.1** (Vektorräume). Ein **Vektorraum** über  $K$  (oder  $K$ -Vektorraum) ist eine Menge  $V$  zusammen mit zwei Verknüpfungen

$$\begin{aligned} &+ : V \times V \rightarrow V \quad (\text{Vektoraddition}) \\ \text{und} \quad &\cdot : K \times V \rightarrow V \quad (\text{Skalarmultiplikation}) \end{aligned}$$

so dass gilt:

- (a)  $(V, +)$  ist eine abelsche Gruppe (siehe Definition 3.1).
- (b) (1. Distributivität) Für alle  $\lambda, \mu \in K$  und  $x \in V$  gilt  $(\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot x$ .
- (c) (2. Distributivität) Für alle  $\lambda \in K$  und  $x, y \in V$  gilt  $\lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y$ .
- (d) (Assoziativität) Für alle  $\lambda, \mu \in K$  und  $x \in V$  gilt  $(\lambda \cdot \mu) \cdot x = \lambda \cdot (\mu \cdot x)$ .

- (e) Für alle  $x \in V$  gilt  $1 \cdot x = x$ .

Die Elemente von  $V$  heißen **Vektoren**, die Elemente von  $K$  **Skalare**.

**Bemerkung 13.2.**

- (a) Beachte, dass man einen Vektorraum nur dann definieren kann, wenn man vorher einen Körper  $K$  gewählt hat. Wenn klar ist, welcher Körper gemeint ist, werden wir jedoch auch oft nur von einem Vektorraum (statt einem  $K$ -Vektorraum) sprechen.
- (b) In Definition 13.1 haben wir mehrfach die gleichen Symbole für unterschiedliche Dinge verwendet: Es gibt z. B. zwei Additionen, die wir beide mit „+“ bezeichnet haben, nämlich die Addition  $+: K \times K \rightarrow K$  zweier Körperelemente und die Addition  $+: V \times V \rightarrow V$  der Vektoren. Da man aus der Art der verknüpften Elemente eindeutig ablesen kann, um welche Verknüpfung es sich handeln muss, können dadurch aber keine Mehrdeutigkeiten entstehen: So werden z. B. beim ersten Pluszeichen in Definition 13.1 (b) zwei Skalare, beim zweiten jedoch zwei Vektoren addiert. Nur wenn wir auch in der Notation explizit deutlich machen wollen, um welche der beiden Verknüpfungen es sich handelt, schreiben wir diese als „ $+_K$ “ bzw. „ $+_V$ “. Analog gibt es auch die Multiplikation zweimal, einmal als Multiplikation „ $\cdot_K$ “ in  $K$  und einmal als Skalarmultiplikation „ $\cdot_V$ “, und auch zweimal die Null, nämlich einmal als Null  $0_K$  im Körper  $K$  und einmal als *Nullvektor*  $0_V$ , d. h. als das neutrale Element von  $(V, +)$ . In dieser ausführlichen Notation könnte man z. B. die Bedingung aus Definition 13.1 (b) als

$$(\lambda +_K \mu) \cdot_V x = \lambda \cdot_V x +_V \mu \cdot_V x$$

schreiben. In der Regel werden wir diese Indizes  $K$  und  $V$  jedoch weglassen, genauso wie die Malzeichen sowohl für „ $\cdot_K$ “ als auch für „ $\cdot_V$ “.

- (c) Vielleicht seid ihr es aus der Schule oder anderen Studienfächern gewohnt, einen Vektor  $x$  durch eine besondere Schreibweise wie z. B.  $\vec{x}$  zu kennzeichnen. Dies ist in der Mathematik eher unüblich, da wir gleich in Beispiel 13.3 noch sehen werden, dass sehr viele Objekte der Mathematik als Vektoren aufgefasst werden können und dementsprechend mit dieser Notation versehen werden müssten. Da wir außerdem ohnehin immer von jeder Variablen angeben, um was für ein Objekt es sich handelt bzw. aus welcher Menge sie kommt, brauchen wir auch keine weitere besondere Kennzeichnung für Vektoren.

**Beispiel 13.3.**

- (a) Für jeden Körper  $K$  ist  $V = \{0\}$  (mit den trivialen Verknüpfungen) ein  $K$ -Vektorraum, der sogenannte **Nullvektorraum**.
- (b) Das mit Abstand wichtigste Beispiel ist für eine gegebene Zahl  $n \in \mathbb{N}$  die Menge

$$K^n := \underbrace{K \times \cdots \times K}_{n\text{-mal}} = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} : x_1, \dots, x_n \in K \right\}$$

aller „geordneten  $n$ -Tupel in  $K$ “ – d. h. ein Element von  $K^n$  wird dadurch angegeben, dass man  $n$  Elemente  $x_1, \dots, x_n$  von  $K$  angibt (die nicht notwendig verschieden sein müssen und auf deren Reihenfolge es ankommt). Dass wir die Elemente  $x_1, \dots, x_n$  dabei untereinander und nicht nebeneinander schreiben, ist momentan eine reine Konvention, die sich später bei der Einführung von Matrizen in Abschnitt 15.A als nützlich erweisen wird. Definiert man nun auf  $K^n$  die komponentenweisen Verknüpfungen

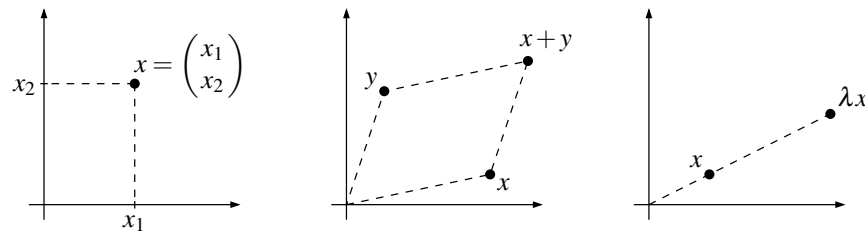
$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix} \quad \text{und} \quad \lambda \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}$$

für alle  $\lambda \in K$ , so ist  $K^n$  mit diesen Verknüpfungen ein  $K$ -Vektorraum. In der Tat folgen die Vektorraumeigenschaften alle aus den Körpereigenschaften von  $K$ ; wir zeigen hier exemplarisch Teil (b) der Definition 13.1: Für alle  $\lambda, \mu, x_1, \dots, x_n \in K$  gilt

$$(\lambda + \mu) \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} (\lambda + \mu)x_1 \\ \vdots \\ (\lambda + \mu)x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 + \mu x_1 \\ \vdots \\ \lambda x_n + \mu x_n \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \mu \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

wobei das mittlere Gleichheitszeichen genau die Distributivität in  $K$  ist und die beiden anderen aus der Definition der Vektoraddition und Skalarmultiplikation in  $K^n$  folgen.

Im Fall  $K = \mathbb{R}$  und  $n = 2$  ist  $K^n = \mathbb{R}^2$  einfach die bekannte reelle Ebene, und Addition und Skalarmultiplikation können wie im folgenden Bild veranschaulicht werden. Die geometrische Interpretation im Fall von  $\mathbb{R}^n$  für andere  $n$  ist natürlich analog.



Wenn wir im Folgenden vom Vektorraum  $K^n$  sprechen, werden wir diesen Raum immer als Vektorraum über dem Körper  $K$  mit diesen komponentenweisen Verknüpfungen betrachten (sofern wir nichts anderes angeben).

Im Fall  $n = 1$  erhält man  $K^1 = K$ , also  $K$  selbst als  $K$ -Vektorraum; der Fall  $n = 0$  wird konventionsgemäß als der Nullvektorraum  $K^0 = \{0\}$  aufgefasst.

- (c) Die axiomatische Herangehensweise an die Vektorraumtheorie hat den großen Vorteil, dass sie auch auf viele andere Fälle als  $K^n$  anwendbar ist. Als weiteres wichtiges Beispiel ist z. B. für eine beliebige Menge  $D$

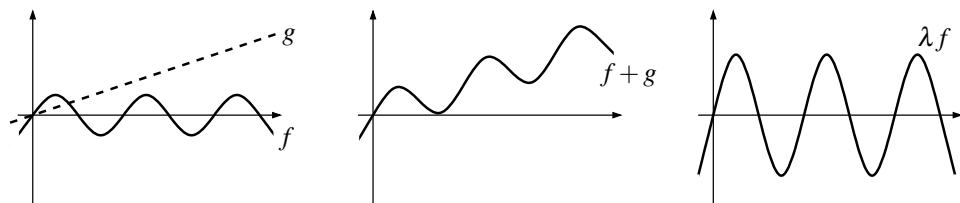
$$\text{Abb}(D, K) := \{f : f \text{ ist eine Abbildung von } D \text{ nach } K\}$$

ein  $K$ -Vektorraum, indem wir die Addition und Multiplikation solcher Abbildungen punktweise definieren als

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

für alle  $\lambda \in K, x \in D$  und  $f, g : D \rightarrow K$  (also  $f, g \in \text{Abb}(D, K)$ ). In der Tat haben wir in Aufgabe 3.12 (a) bereits gesehen, dass  $\text{Abb}(D, K)$  eine abelsche Gruppe ist (der Nullvektor ist hierbei die Funktion, die jedes Element von  $D$  auf 0 abbildet, und das zu einer Funktion  $f : D \rightarrow K$  additive Inverse die Funktion  $-f : D \rightarrow K, x \mapsto -f(x)$ ). Die anderen Vektorraumeigenschaften zeigt man wieder analog.

Beachte, dass die anschauliche Bedeutung in diesem Beispiel trotz der gleichen algebraischen Eigenschaften aus Definition 13.1 eine ganz andere ist als in (b): Zeichnen wir z. B. wie im Bild unten Abbildungen  $f, g \in \text{Abb}(\mathbb{R}, \mathbb{R})$  als ihre Graphen in  $\mathbb{R}^2$ , so entsprechen Vektoraddition und Skalarmultiplikation der Addition bzw. Streckung der Funktionswerte in vertikaler Richtung.



Ein wichtiger Spezialfall dieser Konstruktion ist der Vektorraum  $\text{Abb}(\mathbb{N}, K)$ , dessen Elemente  $f \in \text{Abb}(\mathbb{N}, K)$  wir als „unendliche Folgen“  $(f(0), f(1), f(2), \dots)$  mit Elementen in  $K$  schreiben können. Im Fall  $K = \mathbb{R}$  sind dies natürlich gerade die in der Analysis betrachteten reellen Zahlenfolgen aus Abschnitt ??.

Darüber hinaus lässt sich auf die gleiche Art auch die Menge  $\text{Abb}(D, W)$  aller Abbildungen von einer beliebigen Menge  $D$  in einen  $K$ -Vektorraum  $W$  zu einem Vektorraum machen.

- (d) Sind  $V$  und  $W$  zwei  $K$ -Vektorräume, so ist (in Verallgemeinerung von (b)) auch ihr Produkt  $V \times W$  mit komponentenweiser Addition und Skalarmultiplikation ein  $K$ -Vektorraum.
- (e)  $\mathbb{R}$  ist ein  $\mathbb{Q}$ -Vektorraum. In der Tat kann man reelle Zahlen addieren und mit einer rationalen multiplizieren, und es ist klar, dass mit diesen Definitionen alle Vektorraumeigenschaften erfüllt sind.

Wir sehen also schon, dass es ganz verschiedene Vektorräume gibt; in den nächsten Kapiteln werden wir auch noch viele weitere kennenlernen. Bei der Einführung neuer Konzepte der linearen Algebra ist es für die Anschauung aber vermutlich empfehlenswert, sich immer zuerst den Fall des Vektorraums  $\mathbb{R}^n$  mit  $n \in \mathbb{N}$  vorzustellen.

Als Erstes wollen wir jetzt ein paar elementare Eigenschaften von Vektorräumen zeigen. Sie haben einen ähnlichen Charakter wie die Axiome in Definition 13.1, folgen aber bereits aus diesen (so dass man sie nicht separat als Axiome fordern muss).

**Lemma 13.4** (Eigenschaften von Vektorräumen). *In jedem  $K$ -Vektorraum  $V$  gilt für alle  $\lambda \in K$  und  $x \in V$ :*

- (a)  $0_K \cdot x = \lambda \cdot 0_V = 0_V$ .
- (b) Ist  $\lambda \cdot x = 0_V$ , so ist  $\lambda = 0_K$  oder  $x = 0_V$ .
- (c)  $(-1) \cdot x = -x$ .

*Beweis.*

- (a) Der Beweis ist ganz analog zu dem von Lemma 3.8 (a): Wegen der 1. Distributivität aus Definition 13.1 (b) gilt  $0_K \cdot x = (0_K + 0_K) \cdot x = 0_K \cdot x + 0_K \cdot x$ , nach Subtraktion von  $0_K \cdot x$  also wie behauptet  $0_V = 0_K \cdot x$ .

Genauso erhalten wir mit der 2. Distributivität  $\lambda \cdot 0_V = \lambda \cdot (0_V + 0_V) = \lambda \cdot 0_V + \lambda \cdot 0_V$ , nach Subtraktion von  $\lambda \cdot 0_V$  also  $0_V = \lambda \cdot 0_V$ .

- (b) Ist  $\lambda x = 0_V$  und  $\lambda \neq 0_K$ , so folgt

$$\begin{aligned} x &= 1 \cdot x && \text{(Definition 13.1 (e))} \\ &= (\lambda^{-1} \cdot \lambda)x \\ &= \lambda^{-1}(\lambda x) && \text{(Definition 13.1 (d))} \\ &= 0_V && \text{(Teil (a)).} \end{aligned}$$

- (c) Es gilt

$$\begin{aligned} (-1) \cdot x + x &= (-1) \cdot x + 1 \cdot x && \text{(Definition 13.1 (e))} \\ &= (-1 + 1) \cdot x && \text{(Definition 13.1 (b))} \\ &= 0_K \cdot x \\ &= 0_V && \text{(Teil (a)),} \end{aligned}$$

also ist  $(-1) \cdot x$  das additive Inverse zu  $x$ . □

Oft möchte man mit der Vektoraddition und Skalarmultiplikation nicht nur zwei, sondern mehrere Vektoren miteinander kombinieren. Um damit in Zukunft besser arbeiten zu können, führen wir hier schon einmal die folgenden Notationen ein.

**Definition 13.5** (Familien und Linearkombinationen). Es sei  $V$  ein Vektorraum.

- (a) Eine (endliche) **Familie** von Vektoren in  $V$  ist gegeben durch  $n$  Vektoren  $x_1, \dots, x_n \in V$  für ein  $n \in \mathbb{N}$ . Wir schreiben eine solche Familie als  $B = (x_1, \dots, x_n)$  und nennen die Vektoren  $x_1, \dots, x_n$  ihre **Elemente**.
- (b) Sind  $B = (x_1, \dots, x_n)$  eine Familie von Vektoren in  $V$  und  $\lambda_1, \dots, \lambda_n \in K$ , so heißt

$$\lambda_1 x_1 + \dots + \lambda_n x_n \in V$$

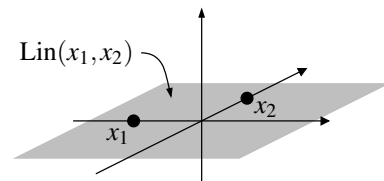
eine **Linearkombination** der Vektoren aus  $B$ . Wir bezeichnen die Menge aller dieser Linearkombinationen mit

$$\text{Lin} B := \{ \lambda_1 x_1 + \dots + \lambda_n x_n : \lambda_1, \dots, \lambda_n \in K \} \subset V.$$

Im Fall  $n = 0$  der leeren Familie fassen wir dies (im Sinne einer leeren Summe wie in Notation 3.9 (c)) als  $\text{Lin}(\cdot) := \{0\}$  auf.

### Bemerkung 13.6.

- (a) Beachte, dass  $x_1, \dots, x_n$  in Definition 13.5 im Gegensatz zu Beispiel 13.3 (b) Vektoren in  $V$ , und nicht die Komponenten *eines* Vektors in  $K^n$  sind. In der Tat ist diese Indexnotation in der Praxis für beide Bedeutungen üblich. Aus dem Zusammenhang ist aber immer offensichtlich, was gemeint ist, da  $x_1, \dots, x_n$  in Beispiel 13.3 (b) ja Elemente des Grundkörpers  $K$  sind, hier jedoch Elemente des betrachteten Vektorraums  $V$ .
- (b) Eine Familie  $B = (x_1, \dots, x_n)$  ist etwas Ähnliches wie eine Menge  $\{x_1, \dots, x_n\}$ , allerdings haben ihre Elemente  $x_1, \dots, x_n$  eine vorgegebene Reihenfolge und können in der Familie auch mehrfach auftreten. Für die Definition von  $\text{Lin} B$  ist dieser Unterschied noch irrelevant (wir hätten  $\text{Lin} B$  genauso gut auch für eine Menge  $B$  definieren können), später in Bemerkung 14.10 (b) bzw. Abschnitt 16.B wird er jedoch noch wichtig werden.
- (c) Mit der geometrischen Interpretation der Vektoraddition und Skalarmultiplikation aus Beispiel 13.3 (b) kann man sich die Mengen  $\text{Lin} B$  auch anschaulich gut vorstellen: So ist z. B.  $\text{Lin}(x_1, x_2) \subset \mathbb{R}^3$  im Bild rechts die Ebene durch den Nullpunkt und die beiden Vektoren  $x_1$  und  $x_2$ .



**Aufgabe 13.7.** Es sei  $V$  ein  $K$ -Vektorraum.

Wenn wir das Vektorraumaxiom (e) „ $1 \cdot x = x$  für alle  $x \in V$ “ in Definition 13.1 weglassen würden, könnten wir versuchen, für ein gegebenes  $a \in K$  eine geänderte Skalarmultiplikation

$$\lambda \odot x := \lambda a x \quad \text{für alle } \lambda \in K \text{ und } x \in V$$

zu definieren. Für welche  $a$  wäre  $K^n$  mit der gewöhnlichen Vektoraddition „+“ und dieser geänderten Skalarmultiplikation „ $\odot$ “ dann ein Vektorraum?

## 13.B Untervektorräume

Sehr viele weitere Beispiele von Vektorräumen können wir erhalten, indem wir in bereits bekannten Vektorräumen nach Teilmengen suchen, die mit der gegebenen Vektoraddition und Skalarmultiplikation selbst wieder die Axiome aus Definition 13.1 erfüllen. Solche Teilmengen werden als Untervektorräume bezeichnet.

**Definition 13.8** (Untervektorräume). Es sei  $V$  ein Vektorraum über einem Körper  $K$ . Eine Teilmenge  $U \subset V$  heißt **Untervektorraum** oder **Unterraum** von  $V$ , in Zeichen  $U \leq V$ , wenn gilt:

- (a)  $0 \in U$ .
- (b) (Abgeschlossenheit bzgl. Vektoraddition) Für alle  $x, y \in U$  ist  $x + y \in U$ .
- (c) (Abgeschlossenheit bzgl. Skalarmultiplikation) Für alle  $x \in U$  und  $\lambda \in K$  ist  $\lambda x \in U$ .

**Bemerkung 13.9.** Äquivalent zu Definition 13.8 kann man dort die Bedingung (a) auch durch die scheinbar schwächere Bedingung  $U \neq \emptyset$  ersetzen: Ist nämlich  $U$  nicht leer, so gibt es ein Element  $x \in U$ , und nach (c) folgt dann mit Lemma 13.4 (a) auch  $0 = 0 \cdot x \in U$ .

Die Abgeschlossenheit eines Unterraums bezüglich Vektoraddition und Skalarmultiplikation bedeutet gerade, dass sich diese beiden Verknüpfungen zu Verknüpfungen

$$+ : U \times U \rightarrow U \quad \text{und} \quad \cdot : K \times U \rightarrow U$$

auf  $U$  einschränken lassen. Wie schon angekündigt wollen wir nun sehen, dass dies die Menge  $U$  selbst wieder zu einem Vektorraum macht.

**Lemma 13.10.** *Jeder Unterraum eines  $K$ -Vektorraums ist (mit den eingeschränkten Verknüpfungen) selbst wieder ein  $K$ -Vektorraum.*

*Beweis.* Wir müssen die Eigenschaften aus Definition 13.1 für  $U$  nachweisen. Dazu beginnen wir mit (a) und zeigen zunächst, dass  $(U, +)$  eine Gruppe ist.

- (Assoziativität der Vektoraddition) Weil  $V$  ein Vektorraum ist, gilt  $(x+y)+z = x+(y+z)$  für alle  $x, y, z \in V$  und damit erst recht für alle  $x, y, z \in U$ . Die Assoziativität der Addition überträgt sich also direkt von  $V$  auf  $U$ .
- (Additives neutrales Element) Nach Definition 13.8 (a) liegt der Nullvektor in  $U$ . Dieser erfüllt  $x+0 = 0+x = x$  für alle  $x \in V$  und damit auch für alle  $x \in U$ , und ist damit ein neutrales Element für die Addition in  $U$ .
- (Additive inverse Elemente) Für jedes  $x \in U$  gilt nach Definition 13.8 (c) auch  $(-1) \cdot x \in U$ , und dies ist nach Lemma 13.4 (c) genau das additive inverse Element zu  $x$ .

Also ist  $(U, +)$  eine Gruppe. Die Kommutativität der Vektoraddition und die übrigen Vektorraumeigenschaften (b) bis (e) in Definition 13.1 sind nun alle von der Form, dass für alle Vektoren aus  $U$  eine bestimmte Gleichung gelten muss – und dies folgt jetzt genauso wie die Assoziativität oben sofort daraus, dass die betreffenden Gleichungen sogar für alle Vektoren aus  $V$  gelten.  $\square$

Das wichtigste Beispiel für Unterräume eines Vektorraums  $V$  sind wie im folgenden Lemma die Teilmengen der Form  $\text{Lin} B$  für eine Familie  $B$  in  $V$  wie in Definition 13.5. In der Tat ist  $\text{Lin} B$  sogar der kleinste Unterraum, der  $B$  enthält:

**Lemma 13.11** (Erzeugte Unterräume). *Für jede Familie  $B = (x_1, \dots, x_n)$  in einem Vektorraum  $V$  gilt:*

- $\text{Lin} B$  ist ein Unterraum von  $V$ , der die Vektoren  $x_1, \dots, x_n$  enthält.
- Ist  $U$  ein beliebiger Unterraum von  $V$ , der die Vektoren  $x_1, \dots, x_n$  enthält, so ist  $\text{Lin} B \subset U$ .

Anschaulich ist  $\text{Lin} B$  also „der kleinste Unterraum von  $V$ , der  $B$  enthält“. Man nennt ihn daher auch den von  $B$  **erzeugten** oder **aufgespannten Unterraum**.

*Beweis.*

- Natürlich sind sowohl der Nullvektor als auch jedes  $x_i$  für  $i \in \{1, \dots, n\}$  eine Linearkombination der Vektoren in  $B = (x_1, \dots, x_n)$ , denn es ist

$$0 = 0 \cdot x_1 + \dots + 0 \cdot x_n \quad \text{und} \quad x_i = 0 \cdot x_1 + \dots + 1 \cdot x_i + \dots + 0 \cdot x_n.$$

Wir müssen also nur noch die Abgeschlossenheit von  $\text{Lin} B$  überprüfen. Dies ist aber sehr einfach: Sind

$$x = \mu_1 x_1 + \dots + \mu_n x_n \quad \text{und} \quad y = \nu_1 x_1 + \dots + \nu_n x_n$$

(mit  $\mu_1, \dots, \mu_n, \nu_1, \dots, \nu_n \in K$ ) zwei beliebige Vektoren in  $\text{Lin} B$ , so gilt für alle  $\lambda \in K$  auch  $x+y = (\mu_1 + \nu_1)x_1 + \dots + (\mu_n + \nu_n)x_n \in \text{Lin} B$  und  $\lambda x = (\lambda \mu_1)x_1 + \dots + (\lambda \mu_n)x_n \in \text{Lin} B$ .

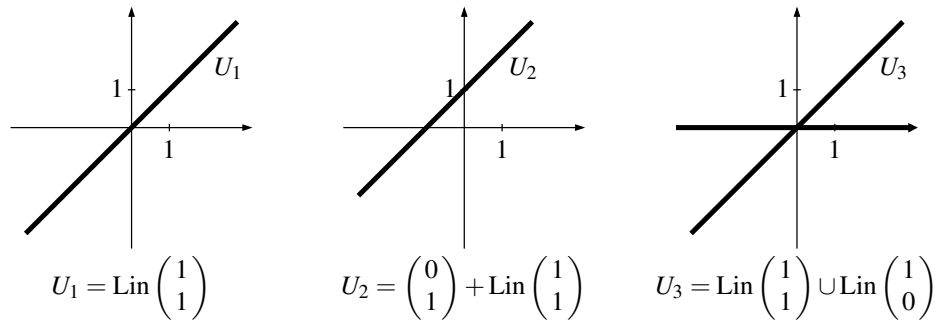
- Ist  $U$  ein solcher Unterraum, so enthält er wegen der Abgeschlossenheit mit  $x_1, \dots, x_n$  auch alle Linearkombinationen dieser Vektoren, also die Menge  $\text{Lin} B$ .  $\square$

**Beispiel 13.12.**

- (a) Für jeden Vektor  $x \in \mathbb{R}^2 \setminus \{0\}$  ist die von  $x$  aufgespannte Ursprungsgerade

$$\text{Lin}(x) = \{\lambda x : \lambda \in \mathbb{R}\} \subset \mathbb{R}^2$$

nach Lemma 13.11 ein Unterraum von  $\mathbb{R}^2$  (und damit nach Lemma 13.10 auch ein Vektorraum). Damit gilt im folgenden Bild also  $U_1 \leq \mathbb{R}^2$ .



Dagegen ist die verschobene Ursprungsgerade  $U_2$  nach Definition 13.8 (a) kein Unterraum von  $\mathbb{R}^2$ , da sie den Ursprung nicht enthält. Auch die Vereinigung  $U_3$  von zwei Ursprungsgeraden ist kein Unterraum, denn sie ist bzgl. der Vektoraddition nicht abgeschlossen: Es gilt

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} \in U_3 \text{ und } \begin{pmatrix} 1 \\ 0 \end{pmatrix} \in U_3, \text{ aber } \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \notin U_3.$$

29

- (b) Für jeden Vektorraum  $V$  sind der Nullvektorraum  $\{0\} \subset V$  und der gesamte Raum  $V \subset V$  natürlich stets Unterräume von  $V$ . Sie werden die **trivialen Unterräume** genannt.  
 (c) Für eine gegebene Teilmenge  $D \subset \mathbb{R}$  ist die Teilmenge

$$\text{Pol}(D, \mathbb{R}) := \{f \in \text{Abb}(D, \mathbb{R}) : f \text{ ist eine Polynomfunktion}\}$$

aller Polynomfunktionen  $f: D \rightarrow \mathbb{R}$  ein Unterraum des Vektorraums  $\text{Abb}(D, \mathbb{R})$  aus Beispiel 13.3 (c), denn sie enthält die Nullfunktion, und mit  $f$  und  $g$  sind auch  $f + g$  und  $\lambda f$  für alle  $\lambda \in \mathbb{R}$  Polynomfunktionen. Genauso ist für festes  $n \in \mathbb{N}$  auch die Menge

$$\text{Pol}_n(D, \mathbb{R}) := \{f \in \text{Abb}(D, \mathbb{R}) : f \text{ ist eine Polynomfunktion vom Grad höchstens } n\}$$

ein Unterraum von  $\text{Abb}(D, \mathbb{R})$ .

Wir wollen nun sehen, wie man aus mehreren Unterräumen eines Vektorraums neue konstruieren kann. Eine Möglichkeit besteht dabei einfach darin, ihren Durchschnitt zu bilden. Im Gegensatz dazu ist ihre Vereinigung zwar in der Regel kein Unterraum (wie wir an der Menge  $U_3$  in Beispiel 13.12 (a) gesehen haben); es gibt aber trotzdem eine Möglichkeit, aus ihnen einen neuen zu erzeugen, der sie enthält – die korrekte Konstruktion hierfür ist nur nicht die Vereinigung, sondern die sogenannte Summe von Unterräumen:

**Lemma 13.13** (Durchschnitte und Summen von Unterräumen). *Es seien  $U_1$  und  $U_2$  Unterräume eines Vektorraums  $V$ . Dann gilt:*

- (a) *Der Durchschnitt  $U_1 \cap U_2$  ist ebenfalls ein Unterraum von  $V$ .*  
 (b) *Die **Summe**  $U_1 + U_2 := \{x_1 + x_2 : x_1 \in U_1, x_2 \in U_2\}$  ist ebenfalls ein Unterraum von  $V$ .*

Analog gilt dies auch für Durchschnitte  $U_1 \cap \dots \cap U_n$  und Summen  $U_1 + \dots + U_n$  von mehr als zwei Unterräumen.

*Beweis.* Wir überprüfen jeweils die Bedingungen aus Definition 13.8. Dabei ist natürlich klar, dass der Nullvektor sowohl in  $U_1$  als auch in  $U_2$  liegt, und damit auch in  $U_1 \cap U_2$  und  $U_1 + U_2$ . Wir müssen also nur noch die Abgeschlossenheit zeigen.

- (a) Es seien  $\lambda \in K$  und  $x, y \in U_1 \cap U_2$ , also  $x, y \in U_1$  und  $x, y \in U_2$ . Da  $U_1$  und  $U_2$  Unterräume sind, liegen damit sowohl  $x + y$  als auch  $\lambda x$  in  $U_1$  und  $U_2$ , d. h. in  $U_1 \cap U_2$ .
- (b) Es seien  $\lambda \in K$  und  $x, y \in U_1 + U_2$ , also  $x = x_1 + x_2$  und  $y = y_1 + y_2$  mit  $x_1, y_1 \in U_1$  und  $x_2, y_2 \in U_2$ . Dann gilt wegen der Abgeschlossenheit von  $U_1$  und  $U_2$

$$x + y = x_1 + x_2 + y_1 + y_2 = \underbrace{(x_1 + y_1)}_{\in U_1} + \underbrace{(x_2 + y_2)}_{\in U_2} \in U_1 + U_2,$$

und analog auch

$$\lambda x = \lambda(x_1 + x_2) = \underbrace{\lambda x_1}_{\in U_1} + \underbrace{\lambda x_2}_{\in U_2} \in U_1 + U_2. \quad \square$$

**Beispiel 13.14** (Summen von erzeugten Unterräumen). Sind zwei Unterräume eines Vektorraums  $V$  durch erzeugende Vektoren gegeben als

$$U_1 = \text{Lin}(x_1, \dots, x_n) = \{\lambda_1 x_1 + \dots + \lambda_n x_n : \lambda_1, \dots, \lambda_n \in K\}$$

und  $U_2 = \text{Lin}(y_1, \dots, y_m) = \{\mu_1 y_1 + \dots + \mu_m y_m : \mu_1, \dots, \mu_m \in K\},$

so ist nach Definition der Summe natürlich

$$U_1 + U_2 = \{\lambda_1 x_1 + \dots + \lambda_n x_n + \mu_1 y_1 + \dots + \mu_m y_m : \lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_m \in K\}$$

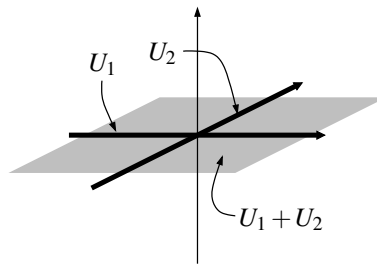
$$= \text{Lin}(x_1, \dots, x_n, y_1, \dots, y_m).$$

Als konkretes Beispiel ist für die Unterräume

$$U_1 = \text{Lin} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \text{und} \quad U_2 = \text{Lin} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

von  $\mathbb{R}^3$  ihr Durchschnitt gleich  $U_1 \cap U_2 = \{0\}$ , und ihre Summe wie im Bild rechts dargestellt die Ebene

$$U_1 + U_2 = \text{Lin} \left( \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right).$$



**Aufgabe 13.15.** Welche der folgenden Teilmengen sind Unterräume von  $\mathbb{R}^3$  (wobei  $x_1, x_2, x_3$  die Komponenten des Vektors  $x \in \mathbb{R}^3$  bezeichnen)?

- (a)  $U_1 = \left\{ \begin{pmatrix} -a \\ 0 \\ b-a \end{pmatrix} : a, b \in \mathbb{R} \right\};$
- (b)  $U_2 = \{x \in \mathbb{R}^3 : x_2 + ax_3 = 1 - a\}$  für ein festes, gegebenes  $a \in \mathbb{R};$
- (c)  $U_3 = \{x \in \mathbb{R}^3 : x_1^3 = x_2^3 = x_3^3\};$
- (d)  $U_4 = \{x \in \mathbb{R}^3 : x_1, x_2, x_3 \in \mathbb{Z}\}.$

**Aufgabe 13.16.** Es seien  $U_1$  und  $U_2$  Unterräume eines  $K$ -Vektorraums  $V$ . Zeige, dass  $U_1 \cup U_2$  genau dann ein Unterraum von  $V$  ist, wenn  $U_1 \subset U_2$  oder  $U_2 \subset U_1$  gilt.

**Aufgabe 13.17.** Im Vektorraum  $V = \text{Abb}(\mathbb{R}, \mathbb{R})$  seien

$$U_1 = \{f \in \text{Abb}(\mathbb{R}, \mathbb{R}) : f(-x) = f(x) \text{ für alle } x \in \mathbb{R}\}$$

und  $U_2 = \{f \in \text{Abb}(\mathbb{R}, \mathbb{R}) : f(-x) = -f(x) \text{ für alle } x \in \mathbb{R}\}$

die Teilmengen aller sogenannten geraden bzw. ungeraden Funktionen. Man zeige:

- (a)  $U_1$  und  $U_2$  sind Unterräume von  $V$ .
- (b)  $U_1 \cap U_2 = \{0\}$  und  $U_1 + U_2 = V$ .

## 14. Basen und Dimension

Im letzten Kapitel haben wir in Lemma 13.11 viele Beispiele von Vektorräumen konstruiert, indem wir zu einer (endlichen) Familie  $B$  in einem gegebenen Vektorraum den davon erzeugten Unterraum  $\text{Lin}B$  gebildet haben. In der Tat haben sehr viele in der Praxis auftretende Vektorräume die Eigenschaft, dass ihre Elemente als Linearkombinationen endlich vieler gegebener Vektoren geschrieben werden können – z. B. alle Unterräume von  $K^n$  für beliebige  $n \in \mathbb{N}$ , wie wir in Bemerkung 14.24 noch sehen werden. Wir wollen uns daher zur Vereinfachung oft auf solche gemäß der folgenden Definition endlich erzeugten Vektorräume beschränken. Wie man auch ohne diese Bedingung auskommen kann, werden wir kurz am Ende dieses Kapitels in Bemerkung 14.28 diskutieren.

**Definition 14.1** (Erzeugendensysteme und endlich erzeugte Vektorräume). Es sei  $V$  ein Vektorraum.

- (a) Eine (endliche) Familie  $B = (x_1, \dots, x_n)$  von Vektoren in  $V$  heißt ein **Erzeugendensystem** von  $V$ , wenn  $\text{Lin}B = V$  gilt, d. h. wenn es zu jedem  $x \in V$  Skalare  $\lambda_1, \dots, \lambda_n \in K$  gibt mit  $x = \lambda_1 x_1 + \dots + \lambda_n x_n$ .
- (b) Der Vektorraum  $V$  heißt **endlich erzeugt**, wenn er ein solches endliches Erzeugendensystem besitzt.

**Beispiel 14.2.**

- (a) Für ein gegebenes  $n \in \mathbb{N}$  bilden die sogenannten **Einheitsvektoren**

$$e_1 := \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 := \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, e_n := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \in K^n$$

ein Erzeugendensystem von  $K^n$ , denn jedes  $x \in K^n$  mit Komponenten  $x_1, \dots, x_n \in K$  hat die Form

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1 e_1 + \dots + x_n e_n.$$

Insbesondere ist  $K^n$  damit ein endlich erzeugter Vektorraum.

- (b) Für den Vektorraum  $\mathbb{R}^2$  ist auch die Familie

$$B = \left( \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right)$$

ein Erzeugendensystem: Um dies zu zeigen, müssen wir zu jedem  $x \in \mathbb{R}^2$  mit Komponenten  $x_1$  und  $x_2$  Skalare  $\lambda_1, \lambda_2 \in \mathbb{R}$  finden mit

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \lambda_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix} \Leftrightarrow \begin{array}{l} x_1 = \lambda_1 + \lambda_2 \\ \text{und } x_2 = \lambda_1 - \lambda_2. \end{array}$$

Dieses Gleichungssystem lässt sich aber leicht nach  $\lambda_1$  und  $\lambda_2$  auflösen; wir erhalten

$$\lambda_1 = \frac{x_1 + x_2}{2} \quad \text{und} \quad \lambda_2 = \frac{x_1 - x_2}{2}, \quad \text{also} \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{x_1 + x_2}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{x_1 - x_2}{2} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

- (c) Für alle  $n \in \mathbb{N}$  ist die Familie  $B = (x^0, \dots, x^n)$  aller Potenzfunktionen  $\mathbb{R} \rightarrow \mathbb{R}$ ,  $x \mapsto x^i$  für  $i = 0, \dots, n$  (die wir hier kurz als  $x^i$  schreiben) ein Erzeugendensystem des Vektorraums  $\text{Pol}_n(\mathbb{R}, \mathbb{R})$  aller reellen Polynome vom Grad höchstens  $n$  aus Beispiel 13.12 (c), denn jedes solche Polynom ist nach Definition eine Linearkombination dieser Potenzfunktionen.

Insbesondere ist  $\text{Pol}_n(\mathbb{R}, \mathbb{R})$  also endlich erzeugt. Der Vektorraum  $\text{Pol}(\mathbb{R}, \mathbb{R})$  aller reellen Polynome ohne Gradbeschränkung ist dagegen nicht endlich erzeugt: In einer (endlichen) Familie  $B$  von Polynomen in  $\text{Pol}(\mathbb{R}, \mathbb{R})$  gibt es nämlich zwangsläufig einen größten auftretenden Grad; Polynome in  $\text{Pol}(\mathbb{R}, \mathbb{R})$  von höherem Grad können dann also nicht in  $\text{Lin} B$  liegen.

- (d) Jedes Erzeugendensystem eines Vektorraums  $V$  bleibt natürlich ein Erzeugendensystem von  $V$ , wenn man beliebige Vektoren von  $V$  hinzufügt. So ist z. B. nicht nur wie in (a) die Familie der beiden Einheitsvektoren ein Erzeugendensystem von  $\mathbb{R}^2$ , sondern auch die Familie

$$B = \left( \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \end{pmatrix} \right).$$

Diese Erweiterungsmöglichkeit mit beliebigen Vektoren kann Erzeugendensysteme natürlich unnötig groß machen. Wir sollten daher vor allem nach Erzeugendensystemen suchen, die keine überflüssigen Vektoren mehr beinhalten. In diesem Fall kann man dann anschaulich erwarten, dass die Anzahl der Vektoren in einem Erzeugendensystem als „Dimension“ des Vektorraums interpretiert werden kann – so wie in Beispiel 13.12 (a) ein einzelner Vektor eine (eindimensionale) Gerade und in Bemerkung 13.6 (c) zwei Vektoren eine (zweidimensionale) Ebene aufgespannt haben. Das Ziel dieses Kapitels ist es, diese Idee genau zu untersuchen und damit insbesondere auch den sehr wichtigen Dimensionsbegriff mathematisch exakt einzuführen.

## 14.A Lineare Unabhängigkeit und Basen

In Beispiel 14.2 (d) haben wir schon einen Fall eines Erzeugendensystems  $B$  mit überflüssigen Vektoren gesehen, nämlich

$$B = \left( \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \end{pmatrix} \right) \quad \text{für den Vektorraum } V = \mathbb{R}^2.$$

Dass hier gar nicht alle drei Vektoren von  $B$  benötigt werden, um  $V$  zu erzeugen, liegt einfach daran, dass sich einer von ihnen schon als Linearkombination der beiden anderen darstellen lässt: Es ist z. B.

$$\begin{pmatrix} 2 \\ 3 \end{pmatrix} = 2 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 3 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

so dass wir diesen Vektor im Erzeugendensystem nicht benötigen. Alternativ könnten wir diese Gleichung auch nach einem der anderen Vektoren auflösen und so sehen, dass man stattdessen auch einen der anderen beiden Vektoren aus  $B$  weglassen könnte. Um hier keine Wahl treffen zu müssen, nach welchem Vektor aufgelöst werden soll, schreibt man die obige Gleichung aber in der Regel als

$$2 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 3 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} - 1 \cdot \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

also als eine Linearkombination des Nullvektors aus den gegebenen Vektoren. Um Erzeugendensysteme ohne überflüssige Vektoren zu finden, sollten wir also fordern, dass sie keine solchen Linearkombinationen des Nullvektors zulassen. Solche Erzeugendensysteme bezeichnet man als *Basen*:

**Definition 14.3** (Basen). Es sei  $B = (x_1, \dots, x_n)$  eine Familie von Vektoren in einem Vektorraum  $V$ .

- (a) Die Familie  $B$  heißt **linear abhängig**, wenn es eine Linearkombination  $\lambda_1 x_1 + \dots + \lambda_n x_n = 0$  des Nullvektors gibt, in der mindestens ein  $\lambda_i$  ungleich 0 ist (man nennt dies auch eine *nicht-triviale Linearkombination des Nullvektors*).

Ist das Gegenteil der Fall, folgt aus einer Linearkombination  $\lambda_1 x_1 + \dots + \lambda_n x_n = 0$  des Nullvektors mit zunächst beliebigen  $\lambda_1, \dots, \lambda_n \in K$  also bereits, dass  $\lambda_1 = \dots = \lambda_n = 0$  gelten muss, so heißt  $B$  **linear unabhängig**.

- (b) Die Familie  $B$  heißt eine **Basis** von  $V$ , wenn  $B$  ein linear unabhängiges Erzeugendensystem von  $V$  ist.

**Bemerkung 14.4.** Es sei  $B = (x_1, \dots, x_n)$  eine Familie von Vektoren in einem Vektorraum  $V$ .

- (a) Enthält  $B$  den Nullvektor, d. h. gilt  $x_i = 0$  für ein  $i$ , so ist  $B$  in jedem Fall linear abhängig, denn dann ist ja  $1 \cdot x_i = 0$  eine nicht-triviale Linearkombination des Nullvektors.
- (b) Ebenso ist  $B$  immer linear abhängig, wenn die Familie einen Vektor mehrfach enthält, also wenn  $x_i = x_j$  für gewisse  $i \neq j$  gilt, da dann  $1 \cdot x_i - 1 \cdot x_j = 0$  eine nicht-triviale Linearkombination des Nullvektors ist.

**Beispiel 14.5.**

- (a) Das Erzeugendensystem  $B = (e_1, \dots, e_n)$  der Einheitsvektoren von  $K^n$  aus Beispiel 14.2 (a) ist linear unabhängig, denn sind  $\lambda_1, \dots, \lambda_n \in K$  mit

$$0 = \lambda_1 e_1 + \dots + \lambda_n e_n = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix},$$

so folgt daraus natürlich sofort  $\lambda_1 = \dots = \lambda_n = 0$ . Damit ist  $B$  also eine Basis von  $K^n$ ; man nennt sie die **Standardbasis** von  $K^n$ .

Als Spezialfall davon für  $n = 0$  ist die leere Familie eine Basis des Nullvektorraums (siehe Definition 13.5 (b)).

- (b) Auch das Erzeugendensystem

$$B = \left( \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right)$$

von  $\mathbb{R}^2$  aus Beispiel 14.2 (b) ist linear unabhängig und damit eine Basis von  $\mathbb{R}^2$ : Sind nämlich  $\lambda_1, \lambda_2 \in \mathbb{R}$  mit

$$\lambda_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{array}{l} \lambda_1 + \lambda_2 = 0 \\ \text{und } \lambda_1 - \lambda_2 = 0, \end{array}$$

so folgt aus diesem Gleichungssystem natürlich sofort  $\lambda_1 = \lambda_2 = 0$ .

Neben der Standardbasis aus (a) haben wir damit also noch eine weitere Basis von  $\mathbb{R}^2$  gefunden und sehen damit schon, dass Basen von Vektorräumen nicht eindeutig bestimmt sind. Allerdings werden wir in Folgerung 14.15 noch zeigen, dass alle Basen eines endlich erzeugten Vektorraums zumindest gleich viele Elemente haben. Dass unsere gerade gefundene Basis  $B$  genau wie die Standardbasis von  $\mathbb{R}^2$  aus zwei Vektoren besteht, ist also kein Zufall.

- (c) Ebenfalls in  $\mathbb{R}^2$  ist wie in der Einleitung zu diesem Abschnitt die Familie

$$B = \left( \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \end{pmatrix} \right) \quad \text{wegen} \quad 2 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 3 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} - 1 \cdot \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

linear abhängig, und damit auch keine Basis von  $\mathbb{R}^2$ .

Beachte aber, dass mit einer Rechnung analog zu (b) je zwei der drei Vektoren dieser Familie  $B$  stets linear unabhängig sind. Die lineare Unabhängigkeit einer Familie kann also *nicht* überprüft werden, indem man immer nur zwei ihrer Vektoren miteinander vergleicht!

- (d) Wir betrachten noch einmal den Vektorraum  $\text{Pol}_n(\mathbb{R}, \mathbb{R})$  aller reellen Polynome vom Grad höchstens  $n$  mit dem Erzeugendensystem  $B = (x^0, \dots, x^n)$  der Potenzfunktionen aus Beispiel 14.2 (c). Da eine nicht-triviale Linearkombination dieser Potenzfunktionen nach dem Koeffizientenvergleich aus Lemma 3.22 nie die Nullfunktion sein kann, ist dieses Erzeugendensystem auch linear unabhängig, und damit eine Basis von  $\text{Pol}_n(\mathbb{R}, \mathbb{R})$ .

**Aufgabe 14.6.** Untersuche die Familie  $B$  in den folgenden Fällen auf lineare Unabhängigkeit im Vektorraum  $V$ :

(a)  $V = \mathbb{R}^3$ ;  $B = \left( \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}, \begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix} \right)$ .

- (b)  $V$  ein beliebiger Vektorraum;  $B = (x+y, x+z, y+z)$  für drei linear unabhängige Vektoren  $x, y, z$ .
- (c)  $V = \text{Abb}(\mathbb{R} \setminus \{0\}, \mathbb{R})$ ;  $B = (\varphi_1, \varphi_2, \varphi_3)$  mit

$$\varphi_1: \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}, x \mapsto x, \quad \varphi_2: \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}, x \mapsto \frac{1}{x}, \quad \varphi_3: \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}, x \mapsto \begin{cases} 1 & \text{falls } x > 0, \\ 0 & \text{falls } x < 0. \end{cases}$$

**Aufgabe 14.7.** Es seien  $n \in \mathbb{N}_{>0}$  und  $(x_1, \dots, x_n)$  eine Basis eines Vektorraums  $V$ . Wir setzen

$$y_k := \sum_{i=1}^k x_i \quad \text{für alle } k = 1, \dots, n.$$

Zeige, dass die Familie  $(y_1, \dots, y_n)$  dann ebenfalls eine Basis von  $V$  ist.

Oft ist die folgende äquivalente Umformulierung der Basiseigenschaft nützlich:

**Lemma und Definition 14.8** (Alternatives Kriterium für Basen). *Eine Familie  $B = (x_1, \dots, x_n)$  von Vektoren in einem Vektorraum  $V$  ist genau dann eine Basis von  $V$ , wenn es zu jedem Vektor  $x \in V$  eindeutig bestimmte Skalare  $\lambda_1, \dots, \lambda_n \in K$  gibt mit  $x = \lambda_1 x_1 + \dots + \lambda_n x_n$ .*

In diesem Fall nennt man  $\lambda_1, \dots, \lambda_n$  die **Koordinaten** von  $x$  bezüglich  $B$ .

*Beweis.* Wir zeigen beide Richtungen der behaupteten Äquivalenz:

„ $\Rightarrow$ “: Es sei  $B$  eine Basis von  $V$ . Da  $B$  dann ein Erzeugendensystem von  $V$  ist, gibt es zu jedem  $x \in V$  Skalare  $\lambda_1, \dots, \lambda_n$  mit  $x = \lambda_1 x_1 + \dots + \lambda_n x_n$ .

Diese Koordinaten sind auch eindeutig bestimmt: Sind nämlich  $\mu_1, \dots, \mu_n$  weitere Skalare mit  $x = \mu_1 x_1 + \dots + \mu_n x_n$ , so gilt

$$x = \lambda_1 x_1 + \dots + \lambda_n x_n = \mu_1 x_1 + \dots + \mu_n x_n \quad \Rightarrow \quad (\lambda_1 - \mu_1)x_1 + \dots + (\lambda_n - \mu_n)x_n = 0,$$

und damit folgt wegen der linearen Unabhängigkeit von  $B$  sofort  $\lambda_i - \mu_i = 0$ , also  $\lambda_i = \mu_i$  für alle  $i = 1, \dots, n$ .

„ $\Leftarrow$ “: Ist jeder Vektor in  $V$  eine Linearkombination der Vektoren aus  $B$ , so bedeutet dies natürlich  $\text{Lin } B = V$ . Außerdem ist  $\lambda_1 = \dots = \lambda_n = 0$  nach Voraussetzung die einzige Möglichkeit, den Nullvektor als Linearkombination  $\lambda_1 x_1 + \dots + \lambda_n x_n$  zu schreiben. Damit ist  $B$  auch linear unabhängig.  $\square$

#### Beispiel 14.9.

- (a) Die Komponenten eines Vektors  $x \in K^n$  sind genau seine Koordinaten bezüglich der Standardbasis aus Beispiel 14.5 (a). Auch ohne Erwähnung einer Basis werden wir sie daher in Zukunft oft einfach seine Koordinaten nennen.
- (b) Analog sind die Koeffizienten eines Polynoms in  $\text{Pol}_n(\mathbb{R}, \mathbb{R})$  gerade seine Koordinaten bezüglich der Basis  $(x^0, \dots, x^n)$  aus Beispiel 14.5 (d).

#### Bemerkung 14.10.

- (a) Lemma 14.8 besagt anschaulich, dass wir einen Vektor in einem Vektorraum  $V$  bei gegebener Basis  $B$  genauso gut auch durch den Vektor in  $K^n$  seiner Koordinaten  $\lambda_1, \dots, \lambda_n$  bezüglich  $B$  beschreiben können. Wir werden dies in den Abschnitten 16.B und 16.C noch genau untersuchen.
- (b) Um einen Vektor wie in (a) durch seine Koordinaten bezüglich einer Basis  $B$  beschreiben zu können, ist es wichtig, dass wir Basen als *Familien* und nicht als *Mengen* definiert haben, da die Elemente einer Menge keine vorgegebene Reihenfolge haben und wir somit bei einer Menge  $B$  keine Möglichkeit hätten, die Koordinaten  $\lambda_1, \dots, \lambda_n$  eindeutig den Vektoren in  $B$  zuzuordnen.

30

Wie können wir nun eine Basis eines endlich erzeugten Vektorraums  $V$  finden? Gemäß der Idee in der Einleitung zu diesem Abschnitt sollten wir dazu mit einem Erzeugendensystem von  $V$  starten können, in dem wir dann fortlaufend Vektoren weglassen, die in nicht-trivialen Linearkombinationen des Nullvektors auftreten. Wir wollen jetzt zeigen, dass dieses Verfahren in der Tat immer funktioniert.

**Satz 14.11 (Basisauswahl).** *Aus jedem (endlichen) Erzeugendensystem eines Vektorraums  $V$  kann man eine Basis von  $V$  auswählen.*

*Insbesondere besitzt also jeder endlich erzeugte Vektorraum eine Basis.*

*Beweis.* Es sei  $B = (x_1, \dots, x_n)$  ein Erzeugendensystem von  $V$ . Ist  $B$  bereits linear unabhängig, so sind wir natürlich schon fertig. Andernfalls gibt es nach Definition 14.3 (a) eine nicht-triviale Linearkombination  $\lambda_1 x_1 + \dots + \lambda_n x_n = 0$  des Nullvektors. Nach evtl. Ummummerieren der Vektoren von  $B$  können wir dabei annehmen, dass  $\lambda_1 \neq 0$  gilt (also dass „ $x_1$  in der Linearkombination vorkommt“). Wir lassen dann  $x_1$  weg, setzen also  $B' := (x_2, \dots, x_n)$ , und behaupten, dass  $B'$  immer noch ein Erzeugendensystem von  $V$  ist.

In der Tat ist dies leicht einzusehen: Wegen der vorausgesetzten Linearkombination des Nullvektors und  $\lambda_1 \neq 0$  ist ja

$$x_1 = \frac{1}{\lambda_1} \cdot (-\lambda_2 x_2 - \dots - \lambda_n x_n) \in \text{Lin } B'.$$

Damit enthält der Unterraum  $\text{Lin } B'$  alle Vektoren  $x_1, \dots, x_n$ , nach Lemma 13.11 (b) also auch den von diesen Vektoren erzeugten Unterraum  $\text{Lin}(x_1, \dots, x_n) = \text{Lin } B = V$ . Es ist daher auch  $\text{Lin } B' = V$ , d. h.  $B'$  ist immer noch ein Erzeugendensystem von  $V$ .

Wir wiederholen dieses Verfahren nun einfach rekursiv mit der neuen Familie  $B'$ . Da wir am Anfang nur  $n$  Vektoren hatten, muss es nach spätestens  $n$  Schritten mit einer Basis von  $V$  abbrechen.  $\square$

Als Nächstes wollen wir verschiedene Basen eines Vektorraums miteinander vergleichen können, um zu sehen, dass sie alle gleich viele Elemente besitzen müssen. Wenn ihr jetzt denkt, dass diese Aussage doch „offensichtlich“ sei, habt ihr vermutlich schon einen anschaulichen Dimensionsbegriff im Kopf und meint, dass eine Basis eines „ $n$ -dimensionalen Vektorraums“ aus  $n$  Vektoren bestehen müsse, also z. B. eine Basis einer Geraden aus einem Vektor und eine Basis einer Ebene aus zwei Vektoren. Diese Anschauung wird sich natürlich auch gleich in Abschnitt 14.B als richtig herausstellen – aber dieser Dimensionsbegriff ist nicht Teil der Definition 13.1 eines Vektorraums, und daher sind derartige Aussagen in jedem Fall beweisbedürftig! Wir beginnen dazu mit einem Hilfsresultat, in dem wir zwei Basen miteinander vergleichen, die sich in nur einem Element unterscheiden.

**Lemma 14.12 (Austauschlemma).** *Es sei  $B = (x_1, \dots, x_n)$  eine Basis eines Vektorraums  $V$ . Weiterhin sei  $y \in V$  ein beliebiger Vektor, den wir dann also nach Lemma 14.8 eindeutig als Linearkombination  $y = \lambda_1 x_1 + \dots + \lambda_n x_n$  für gewisse  $\lambda_1, \dots, \lambda_n \in K$  schreiben können.*

*Ist dann  $i \in \{1, \dots, n\}$  mit  $\lambda_i \neq 0$ , so ist auch  $B' := (x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n)$  eine Basis von  $V$ , d. h. wir können den Vektor  $x_i$  in der Basis durch  $y$  ersetzen.*

*Beweis.* Nach evtl. Umbenennung der Vektoren können wir ohne Einschränkung wieder  $i = 1$  annehmen, also  $\lambda_1 \neq 0$  und  $B' = (y, x_2, \dots, x_n)$ . Wir zeigen, dass  $B'$  eine Basis von  $V$  ist:

(a)  $B'$  ist ein Erzeugendensystem von  $V$ : Analog zum Beweis von Satz 14.11 gilt

$$x_1 = \frac{1}{\lambda_1} \cdot (y - \lambda_2 x_2 - \dots - \lambda_n x_n) \in \text{Lin } B'.$$

Also enthält  $\text{Lin } B'$  alle Vektoren  $x_1, \dots, x_n$ , und damit nach Lemma 13.11 (b) auch den davon erzeugten Unterraum  $\text{Lin}(x_1, \dots, x_n) = \text{Lin } B = V$ . Dies bedeutet genau, dass  $B'$  ein Erzeugendensystem von  $V$  ist.

(b)  $B'$  ist linear unabhängig: Es seien  $\mu_1, \dots, \mu_n \in K$  mit  $\mu_1 y + \mu_2 x_2 + \dots + \mu_n x_n = 0$ . Daraus folgt

$$\begin{aligned} 0 &= \mu_1(\lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_n x_n) + \mu_2 x_2 + \dots + \mu_n x_n \\ &= (\mu_1 \lambda_1) x_1 + (\mu_1 \lambda_2 + \mu_2) x_2 + \dots + (\mu_1 \lambda_n + \mu_n) x_n. \end{aligned}$$

Dies ist nun eine Linearkombination des Nullvektors mit Vektoren aus  $B$ . Da  $B$  linear unabhängig ist, müssen darin alle Vorfaktoren verschwinden, d. h. es gilt  $\mu_1 \lambda_1 = 0$  (und damit  $\mu_1 = 0$  wegen  $\lambda_1 \neq 0$ ) und dann auch  $\mu_1 \lambda_i + \mu_i = \mu_i = 0$  für alle  $i = 2, \dots, n$ . Also ist  $B'$  linear unabhängig.  $\square$

**Beispiel 14.13.** Wir betrachten den Vektorraum  $V = \mathbb{R}^3$  mit der Standardbasis  $B = (e_1, e_2, e_3)$  aus Beispiel 14.5 (a). Weiterhin sei

$$y = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} = e_1 + 2e_2.$$

Da in dieser Linearkombination die Vektoren  $e_1$  und  $e_2$  mit Vorfaktoren ungleich Null vorkommen, folgt aus dem Austauschlemma 14.12 also, dass wir einen dieser Vektoren in  $B$  durch  $y$  ersetzen können, d. h. dass auch  $(y, e_2, e_3)$  und  $(e_1, y, e_3)$  Basen von  $\mathbb{R}^3$  sind. Im Gegensatz dazu ist  $(e_1, e_2, y)$  wegen der Linearkombination  $e_1 + 2e_2 - y = 0$  linear abhängig, und somit keine Basis von  $\mathbb{R}^3$ .

Der folgende Satz ergibt sich nun einfach, indem man Lemma 14.12 mehrmals nacheinander anwendet.

**Satz 14.14 (Steinitzcher Austauschatz).** *Es seien  $B = (x_1, \dots, x_n)$  eine Basis und  $(y_1, \dots, y_r)$  eine linear unabhängige Familie in einem Vektorraum  $V$ .*

*Dann ist  $r \leq n$ , und  $x_1, \dots, x_n$  lassen sich so umnummerieren, dass  $B' = (y_1, \dots, y_r, x_{r+1}, \dots, x_n)$  ebenfalls eine Basis von  $V$  ist (man kann in  $B$  also  $r$  der Vektoren  $x_1, \dots, x_n$  durch die gegebenen Vektoren  $y_1, \dots, y_r$  ersetzen).*

*Beweis.* Wir beweisen den Satz mit Induktion über  $r$ ; für  $r = 0$  ist nichts zu zeigen.

Für den Induktionsschritt  $r \rightarrow r + 1$  sei nun  $(y_1, \dots, y_{r+1})$  linear unabhängig. Da dann natürlich auch  $(y_1, \dots, y_r)$  linear unabhängig ist, gilt nach Induktionsvoraussetzung  $r \leq n$ , und nach geeigneter Umnummerierung der Vektoren in  $B$  ist  $(y_1, \dots, y_r, x_{r+1}, \dots, x_n)$  eine Basis von  $V$ . Wir können den Vektor  $y_{r+1}$  also als Linearkombination

$$y_{r+1} = \lambda_1 y_1 + \dots + \lambda_r y_r + \lambda_{r+1} x_{r+1} + \dots + \lambda_n x_n$$

schreiben. Dabei muss mindestens einer der Vorfaktoren  $\lambda_{r+1}, \dots, \lambda_n$  ungleich Null sein, denn andernfalls wäre

$$\lambda_1 y_1 + \dots + \lambda_r y_r - y_{r+1} = 0$$

im Widerspruch dazu, dass die Familie  $(y_1, \dots, y_{r+1})$  linear unabhängig ist. Insbesondere muss also  $r + 1 \leq n$  gelten, und nach evtl. Umbenennung von  $x_{r+1}, \dots, x_n$  können wir annehmen, dass  $\lambda_{r+1} \neq 0$  ist. Dann können wir aber nach dem Austauschlemma 14.12 den Vektor  $x_{r+1}$  in der Basis  $(y_1, \dots, y_r, x_{r+1}, \dots, x_n)$  durch  $y_{r+1}$  ersetzen, d. h. auch  $(y_1, \dots, y_{r+1}, x_{r+2}, \dots, x_n)$  ist eine Basis von  $V$ . Damit ist der Satz mit Induktion bewiesen.  $\square$

Der Steinitzsche Austauschatz hat zwei sehr wichtige Konsequenzen: zum einen die schon angekündigte Tatsache, dass alle Basen eines endlich erzeugten Vektorraums gleich viele Elemente haben, und zum anderen ein „Gegenstück“ zur Basisauswahl in Satz 14.11.

**Folgerung 14.15.** *Alle Basen eines endlich erzeugten Vektorraums haben gleich viele Elemente.*

*Beweis.* Es seien  $(x_1, \dots, x_n)$  und  $(y_1, \dots, y_r)$  zwei Basen eines endlich erzeugten Vektorraums. Nach Satz 14.14 folgt dann direkt  $r \leq n$ , und durch Vertauschen der Rollen der beiden Basen analog auch  $n \leq r$ . Damit ist also wie behauptet  $r = n$ .  $\square$

**Folgerung 14.16 (Basisergänzung).** *Jede linear unabhängige Familie in einem endlich erzeugten Vektorraum  $V$  kann zu einer Basis ergänzt werden.*

*Beweis.* Es sei  $(y_1, \dots, y_r)$  eine linear unabhängige Familie in  $V$ . Nach Satz 14.11 existiert ferner auch eine Basis  $B = (x_1, \dots, x_n)$  von  $V$ . Die in Satz 14.14 konstruierte Familie  $B'$  ergänzt dann  $(y_1, \dots, y_r)$  zu einer Basis von  $V$ .  $\square$

**Aufgabe 14.17.** Es sei  $U = \text{Lin}(x_1, x_2, x_3) \leq \mathbb{R}^3$  mit

$$x_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad x_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \quad x_3 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

Bestimme eine Basis von  $U$ , und ergänze sie zu einer Basis von  $\mathbb{R}^3$ .

**Aufgabe 14.18.** In  $\text{Pol}_3(\mathbb{R}, \mathbb{R})$  betrachten wir die Familie  $B = (1 + x, x + x^2, x^2 + x^3, x^3 + 1)$ .

Wähle aus  $B$  eine Basis von  $\text{Lin} B$  aus, und stelle die übrigen Elemente von  $B$  als Linearkombinationen dieser Basis dar.

## 14.B Die Dimension von Vektorräumen

Da wir jetzt gesehen haben, dass jeder endlich erzeugte Vektorraum nach Satz 14.11 eine Basis besitzt, und verschiedene Basen nach Folgerung 14.15 gleich viele Elemente haben, können wir nun wie erwartet den Dimensionsbegriff einführen:

**Definition 14.19** (Dimension von Vektorräumen). Für einen endlich erzeugten Vektorraum  $V$  ist die **Dimension**, geschrieben  $\dim V$ , definiert als die Anzahl der Elemente in einer (beliebigen) Basis von  $V$ .

Ist  $V$  nicht endlich erzeugt (und hat damit natürlich auch keine endliche Basis), so schreiben wir formal  $\dim V = \infty$ . Ein endlich erzeugter Vektorraum wird daher oft auch als **endlich-dimensionaler Vektorraum** bezeichnet.

**Beispiel 14.20.**

- (a) Nach Beispiel 14.5 (a) ist  $\dim K^n = n$  für alle  $n \in \mathbb{N}$ .
- (b) Nach Beispiel 14.5 (d) hat  $\text{Pol}_n(\mathbb{R}, \mathbb{R})$  eine Basis  $(x^0, x^1, \dots, x^n)$ , und damit Dimension  $n + 1$  für alle  $n \in \mathbb{N}$ .

Da der Raum  $\text{Pol}(\mathbb{R}, \mathbb{R})$  aller reellen Polynome dagegen nach Beispiel 14.2 (c) nicht endlich erzeugt ist, ist  $\dim \text{Pol}(\mathbb{R}, \mathbb{R}) = \infty$ .

Unsere erste Anwendung des Dimensionsbegriffs ist ein vereinfachtes Kriterium dafür, ob eine gegebene Familie  $B$  eines endlich erzeugten Vektorraums  $V$  eine Basis ist: Wenn  $B$  die richtige Anzahl von Vektoren enthält, dann brauchen wir nicht mehr zu überprüfen, dass  $B$  ein Erzeugendensystem und linear unabhängig ist, sondern es reicht bereits *eine* dieser beiden Eigenschaften.

**Folgerung 14.21** (Basiskriterium). *Es sei  $B$  eine Familie von  $n$  Vektoren in einem endlich erzeugten Vektorraum  $V$ .*

- (a) *Ist  $B$  ein Erzeugendensystem von  $V$ , so gilt  $n \geq \dim V$ .*
- (b) *Ist  $B$  linear unabhängig, so gilt  $n \leq \dim V$ .*

*Ist außerdem eine dieser beiden Voraussetzungen erfüllt, so ist  $B$  genau dann eine Basis von  $V$ , wenn sogar  $n = \dim V$  gilt.*

*Beweis.* Dies folgt unmittelbar mit Hilfe der Basisauswahl und -ergänzung:

- (a) Nach Satz 14.11 können wir aus  $B$  eine Basis  $B'$  von  $V$  auswählen. Da diese Basis  $B'$  dann aus  $\dim V$  Vektoren besteht, muss natürlich  $\dim V \leq n$  gelten.
- (b) Nach Folgerung 14.16 können wir  $B$  zu einer Basis  $B'$  von  $V$  ergänzen. Da diese Basis wieder aus  $\dim V$  Vektoren besteht, folgt  $\dim V \geq n$ .

In beiden Fällen gilt dabei natürlich genau dann  $n = \dim V$ , wenn bereits  $B' = B$  ist, d. h.  $B$  schon eine Basis von  $V$  ist.  $\square$

**Bemerkung 14.22.**

- (a) Ist  $B$  eine Familie von  $n$  Vektoren in einem Vektorraum  $V$ , so ist  $B$  natürlich ein Erzeugendensystem von  $\text{Lin} B$ . Nach Folgerung 14.21 (a) gilt dann also  $\dim \text{Lin} B \leq n$ , mit Gleichheit genau dann, wenn  $B$  linear unabhängig und damit eine Basis von  $\text{Lin} B$  ist.

Insbesondere ist also für jeden Vektor  $x \in V \setminus \{0\}$  die Ursprungsgerade  $\text{Lin}(x) = \{\lambda x : \lambda \in K\}$  ein eindimensionaler Unterraum von  $V$ .

- (b) Für die Anschauung ist auch oft die folgende Umformulierung von Folgerung 14.21 nützlich: Ist  $V$  ein  $n$ -dimensionaler Vektorraum, so haben wir dort gerade in Teil (a) gesehen, dass ein Erzeugendensystem von  $V$  immer mindestens  $n$  Vektoren enthält, und genau dann eine Basis von  $V$  ist, wenn es *genau*  $n$  Vektoren enthält. Da man größere Erzeugendensysteme mit Hilfe der Basisauswahl immer zu einer Basis verkleinern kann, bedeutet dies also: *Eine Basis ist dasselbe wie ein minimales Erzeugendensystem* – also eines, das nicht mehr weiter verkleinert werden kann.

Analog ergibt sich aus Folgerung 14.21 (b) zusammen mit der der Basisergänzung: *Eine Basis ist dasselbe wie eine maximale linear unabhängige Familie* – also eine, die durch Vergrößern in jedem Fall linear abhängig wird.

Als Nächstes wollen wir die Dimensionen von Unterräumen endlich erzeugter Vektorräume untersuchen, und dabei zunächst einmal zeigen, dass solche Unterräume auch selbst wieder endlich erzeugt sind. Dieses Resultat ist vermutlich nicht wirklich überraschend, aber dennoch auch nicht völlig offensichtlich, da es keine einfache Möglichkeit gibt, aus einem Erzeugendensystem eines Vektorraums ein Erzeugendensystem eines gegebenen Unterraums zu konstruieren.

**Lemma 14.23.** *Es sei  $U$  ein Unterraum eines endlich erzeugten  $K$ -Vektorraums  $V$ . Dann gilt:*

- (a)  *$U$  ist ebenfalls endlich erzeugt.*  
 (b)  *$\dim U \leq \dim V$ , mit Gleichheit genau falls  $U = V$ .*

*Beweis.* Es sei  $n := \dim V$ .

- (b) Angenommen, wir wissen bereits, dass  $U$  endlich erzeugt ist. Dann können wir nach Satz 14.11 eine Basis  $B$  von  $U$  mit  $\dim U$  Elementen wählen. Da  $B$  dann auch in  $V$  linear unabhängig ist, bedeutet dies nach Folgerung 14.21 (b) aber  $\dim U \leq \dim V$ , mit Gleichheit genau dann, wenn  $B$  schon eine Basis von  $V$  ist, also  $U = V$  gilt.
- (a) Wäre  $U$  nicht endlich erzeugt, so könnten wir der Reihe nach Vektoren  $x_1, x_2, x_3, \dots$  in  $U$  finden mit  $x_{k+1} \notin U_k := \text{Lin}(x_1, \dots, x_k)$  für alle  $k \in \mathbb{N}$ , denn  $U_k$  ist endlich erzeugt und kann damit nicht gleich  $U$  sein. Da die Vektorräume  $\{0\} = U_0 \subsetneq U_1 \subsetneq \dots \subsetneq U_{n+1} \subset V$  alle endlich erzeugt sind, folgt daraus aber nach der schon gezeigten Aussage (b)

$$0 = \dim U_0 < \dim U_1 < \dots < \dim U_{n+1} \leq \dim V = n.$$

Dies ist ein Widerspruch, da es von 0 bis  $n$  keine  $n+2$  natürlichen Zahlen gibt.  $\square$

**Bemerkung 14.24** (Berechnung von Unterräumen). Anders ausgedrückt besagt Lemma 14.23, dass jeder Unterraum  $U$  eines endlich erzeugten Vektorraums ein (endliches) Erzeugendensystem und damit nach Satz 14.11 auch eine Basis besitzt, also als  $U = \text{Lin}(x_1, \dots, x_n)$  für linear unabhängige Vektoren  $x_1, \dots, x_n$  (mit  $n = \dim U$ ) geschrieben werden kann. Wenn wir im Folgenden sagen, dass wir einen Unterraum eines endlich erzeugten Vektorraums *berechnen* wollen, meinen wir damit, ihn so darzustellen, also eine Basis (und damit auch seine Dimension) zu bestimmen.

Für unsere bisherigen Konstruktionen mit Unterräumen, also den Durchschnitt und die Summe aus Lemma 13.13, wollen wir nun Algorithmen (d. h. Verfahren) angeben, um sie in diesem Sinne zu berechnen. Die folgende allgemeine Dimensionsaussage ist dafür nützlich:

**Satz 14.25** (Dimensionsformel für Durchschnitte und Summen). Sind  $U_1$  und  $U_2$  Unterräume eines endlich-dimensionalen  $K$ -Vektorraums  $V$ , so gilt

$$\dim(U_1 \cap U_2) + \dim(U_1 + U_2) = \dim U_1 + \dim U_2.$$

*Beweis.* Nach Lemma 14.23 sind alle betrachteten Unterräume endlich erzeugt. Wir können nach Satz 14.11 also eine Basis  $(x_1, \dots, x_n)$  von  $U_1 \cap U_2$  wählen und sie nach Folgerung 14.16 zu Basen

$$\begin{aligned} &(x_1, \dots, x_n, y_1, \dots, y_m) \quad \text{von } U_1 \\ \text{und } &(x_1, \dots, x_n, z_1, \dots, z_p) \quad \text{von } U_2 \end{aligned}$$

ergänzen. Wir zeigen, dass  $B = (x_1, \dots, x_n, y_1, \dots, y_m, z_1, \dots, z_p)$  dann eine Basis von  $U_1 + U_2$  ist:

- (a)  $B$  ist ein Erzeugendensystem von  $U_1 + U_2$  nach Beispiel 13.14.
- (b)  $B$  ist linear unabhängig: Es seien  $\lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_m, \nu_1, \dots, \nu_p \in K$  mit

$$\lambda_1 x_1 + \dots + \lambda_n x_n + \mu_1 y_1 + \dots + \mu_m y_m + \nu_1 z_1 + \dots + \nu_p z_p = 0, \tag{*}$$

also

$$\underbrace{\lambda_1 x_1 + \dots + \lambda_n x_n + \mu_1 y_1 + \dots + \mu_m y_m}_{\in U_1} = \underbrace{-\nu_1 z_1 - \dots - \nu_p z_p}_{\in U_2} \in U_1 \cap U_2.$$

Da dieser Vektor in  $U_1 \cap U_2$  liegt und  $(x_1, \dots, x_n)$  diesen Unterraum erzeugt, gibt es also  $\lambda'_1, \dots, \lambda'_n \in K$  mit

$$-\nu_1 z_1 - \dots - \nu_p z_p = \lambda'_1 x_1 + \dots + \lambda'_n x_n \Rightarrow \lambda'_1 x_1 + \dots + \lambda'_n x_n + \nu_1 z_1 + \dots + \nu_p z_p = 0.$$

Aus der linearen Unabhängigkeit von  $(x_1, \dots, x_n, z_1, \dots, z_p)$  folgt damit  $\nu_1 = \dots = \nu_p = 0$ . Setzen wir dies schließlich noch in (\*) ein, erhalten wir daraus mit der linearen Unabhängigkeit von  $(x_1, \dots, x_n, y_1, \dots, y_m)$  auch  $\lambda_1 = \dots = \lambda_n = \mu_1 = \dots = \mu_m = 0$ .

Also ist  $B$  eine Basis von  $U_1 + U_2$ . Damit folgt nun durch Abzählen der Elemente unserer Basen

$$\dim(U_1 + U_2) + \dim(U_1 \cap U_2) = (n + m + p) + n = (n + m) + (n + p) = \dim U_1 + \dim U_2. \quad \square$$

31

**Beispiel 14.26.** Beachte, dass nur die Summe der Dimensionen von  $U_1 \cap U_2$  und  $U_1 + U_2$  durch  $\dim U_1$  und  $\dim U_2$  bestimmt sind, nicht aber die Dimensionen von  $U_1 \cap U_2$  und  $U_1 + U_2$  selbst. Als einfaches Beispiel hierfür seien  $U_1$  und  $U_2$  zwei Geraden (durch den Ursprung) in  $\mathbb{R}^2$ . Dann gibt es zwei Möglichkeiten:

- (a) Ist  $U_1 = U_2$ , so ist  $U_1 \cap U_2 = U_1 + U_2 = U_1 = U_2$ , und die Dimensionsformel ergibt

$$\dim(U_1 \cap U_2) + \dim(U_1 + U_2) = 1 + 1 = 1 + 1 = \dim U_1 + \dim U_2.$$

- (b) Ist hingegen  $U_1 \neq U_2$ , so ist  $U_1 \cap U_2 = \{0\}$  und  $U_1 + U_2 = \mathbb{R}^2$ , und damit

$$\dim(U_1 \cap U_2) + \dim(U_1 + U_2) = 0 + 2 = 1 + 1 = \dim U_1 + \dim U_2.$$

**Algorithmus 14.27** (Berechnung von Durchschnitten und Summen). Es seien  $U_1$  und  $U_2$  zwei Unterräume eines Vektorraums  $V$ , die wir nach Bemerkung 14.24 als  $U_1 = \text{Lin}(x_1, \dots, x_k)$  und  $U_2 = \text{Lin}(y_1, \dots, y_l)$  schreiben können (wobei die gewählten Erzeuger hier nicht unbedingt linear unabhängig sein müssen). Als konkretes Beispiel betrachten wir dabei im Folgenden den Fall  $V = \mathbb{R}^4$  und  $k = l = 2$  mit

$$x_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 2 \end{pmatrix}, \quad x_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad y_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad y_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \in \mathbb{R}^4.$$

In diesem Fall sind die Familien  $(x_1, x_2)$  und  $(y_1, y_2)$  natürlich linear unabhängig, und damit ist  $\dim U_1 = \dim U_2 = 2$ .

- (a) Die Vektoren im Schnitt
- $U_1 \cap U_2$
- erhält man offensichtlich durch Gleichsetzen

$$\lambda_1 x_1 + \cdots + \lambda_k x_k = -\mu_1 y_1 - \cdots - \mu_l y_l \quad (1)$$

der Elemente von  $U_1$  und  $U_2$ , und damit durch Auflösen der Gleichung

$$\lambda_1 x_1 + \cdots + \lambda_k x_k + \mu_1 y_1 + \cdots + \mu_l y_l = 0 \quad (2)$$

nach  $\lambda_1, \dots, \lambda_k, \mu_1, \dots, \mu_l$  (die Vorzeichen spielen hierbei keine Rolle, da mit  $y$  auch  $-y$  in  $U_2$  liegt, und sind daher so gewählt, dass sie sich in (2) wegheben). Die sich als Lösung ergebenden Werte für  $\lambda_1, \dots, \lambda_k$  können dann in die linke Seite von (1) eingesetzt werden und liefern die gesuchten Vektoren im Schnitt  $U_1 \cap U_2$ .

Für unser konkretes Beispiel ist (2) das Gleichungssystem

$$\lambda_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 2 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} + \mu_1 \begin{pmatrix} 1 \\ 1 \\ 0 \\ 2 \end{pmatrix} + \mu_2 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{array}{l} \lambda_1 + \mu_1 = 0, \\ \lambda_2 + \mu_1 = 0, \\ \lambda_2 + \mu_2 = 0, \\ 2\lambda_1 + \lambda_2 + 2\mu_1 + \mu_2 = 0. \end{array}$$

Es ist offensichtlich äquivalent zu  $\lambda_1 = \lambda_2 = -\mu_1 = -\mu_2$  (die letzte Gleichung ist dann automatisch mit erfüllt) und hat damit die allgemeine Lösung  $\lambda_1 = \lambda$ ,  $\lambda_2 = \lambda$ ,  $\mu_1 = -\lambda$ ,  $\mu_2 = -\lambda$  für ein beliebiges  $\lambda \in \mathbb{R}$ . Einsetzen in die linke Seite von (1) liefert also

$$U_1 \cap U_2 = \left\{ \lambda \begin{pmatrix} 1 \\ 0 \\ 0 \\ 2 \end{pmatrix} + \lambda \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} : \lambda \in \mathbb{R} \right\} = \text{Lin} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 3 \end{pmatrix},$$

und damit insbesondere  $\dim(U_1 \cap U_2) = 1$ .

- (b) Für die Summe der gegebenen Unterräume
- $U_1$
- und
- $U_2$
- wissen wir aus Beispiel 13.14 bereits, dass
- $U_1 + U_2 = \text{Lin}(x_1, \dots, x_k, y_1, \dots, y_l)$
- gilt. Um eine Basis von
- $U_1 + U_2$
- zu bestimmen, müssen wir aus diesen Vektoren also nur noch mit Satz 14.11 eine Basis auswählen.

Beachte, dass wir dazu in (a) (mit  $\lambda = 1$ ) bereits die Linearkombination  $x_1 + x_2 - y_1 - y_2 = 0$  des Nullvektors berechnet haben, in der jeder der vier Ausgangsvektoren vorkommt. Nach dem Verfahren aus dem Beweis von Satz 14.11 können wir also einen beliebigen dieser Vektoren aus dem Erzeugendensystem streichen und erhalten so z. B. auch  $U_1 + U_2 = \text{Lin}(x_1, x_2, y_1)$ . Nach der Dimensionsformel aus Satz 14.25 ist aber auch

$$\dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2) \stackrel{(a)}{=} 2 + 2 - 1 = 3,$$

und damit ist  $(x_1, x_2, y_1)$  nach Folgerung 14.21 eine Basis von  $U_1 + U_2$ .

Wir haben uns in diesem Kapitel jetzt ausführlich mit Basen endlich-dimensionaler Vektorräume beschäftigt. Da in der Praxis aber auch manchmal unendlich-dimensionale Vektorräume eine Rolle spielen, wollen wir nun zum Abschluss dieses Kapitels kurz diskutieren, welche Änderungen an unseren Konstruktionen in diesem Fall nötig sind und wie sie unsere Ergebnisse zu Basen beeinflussen.

**Bemerkung 14.28** (Unendlich-dimensionale Vektorräume). Ist  $V$  ein beliebiger Vektorraum, der also nicht unbedingt von endlich vielen Elementen erzeugt werden kann (wie z. B.  $\text{Pol}(\mathbb{R}, \mathbb{R})$  in Beispiel 14.2 (c) oder der noch größere Vektorraum  $\text{Abb}(\mathbb{R}, \mathbb{R})$  aus Beispiel 13.3 (c)), müssen wir zunächst einmal auch möglicherweise unendliche Familien von Vektoren in  $V$  betrachten. Eine solche Familie besteht aus einer evtl. unendlichen Indexmenge  $I$  zusammen mit einem Vektor  $x_i$  für alle  $i \in I$ ; wir schreiben sie (analog zu Folgen in Definition ?? ??) als  $B = (x_i)_{i \in I}$ . Im Fall einer endlichen Familie könnten wir dabei die Indexmenge  $I = \{1, \dots, n\}$  mit  $n \in \mathbb{N}$  wählen und wieder die bisherige Notation  $B = (x_1, \dots, x_n)$  verwenden; im Fall der Indexmenge  $I = \mathbb{N}$  könnten wir die Familie auch als  $B = (x_0, x_1, x_2, \dots)$  schreiben.

Wichtig ist jedoch, dass wir auch im Fall einer unendlichen Familie als Linearkombinationen daraus *keine unendlichen Summen* bilden können, da wir hierfür einen Konvergenzbegriff wie z. B. in Definition ?? bräuchten – der in der linearen Algebra (über einem beliebigen Grundkörper) aber

nicht existiert. Für eine Familie  $B = (x_i)_{i \in I}$  definiert man eine *Linearkombination* der Vektoren aus  $B$  daher als einen Vektor der Form

$$\sum_{i \in I} \lambda_i x_i \in V, \quad (*)$$

wobei  $\lambda_i \in K$  für alle  $i \in I$  gilt und *nur endlich viele dieser  $\lambda_i$  ungleich 0 sind* – so dass der Ausdruck (\*) nach Weglassen aller Summanden, die 0 sind, also in jedem Fall nur eine endliche Summe ist. Mit anderen Worten sind Linearkombinationen in der linearen Algebra also immer endlich.

Die darauf aufbauenden Definitionen sind nun wieder wie erwartet: Den Unterraum aller dieser Linearkombinationen bezeichnen wir mit  $\text{Lin} B$ . Die Familie  $B$  heißt ein *Erzeugendensystem* von  $V$ , wenn  $\text{Lin} B = V$  gilt, und *linear unabhängig*, wenn es keine nicht-triviale Linearkombination des Nullvektors mit Vektoren aus  $B$  gibt. Ist  $B$  ein linear unabhängiges Erzeugendensystem von  $V$ , so heißt  $B$  eine *Basis* von  $V$ . Hier sind zwei einfache Beispiele dafür:

- Die (unendliche) Familie aller Potenzfunktionen  $(x^i)_{i \in \mathbb{N}}$  ist eine Basis des Polynomraums  $\text{Pol}(\mathbb{R}, \mathbb{R})$ : Sie ist ein Erzeugendensystem, da jedes Polynom eine (endliche) Linearkombination dieser Potenzfunktionen ist, und linear unabhängig nach dem Koeffizientenvergleich wie in Beispiel 14.5 (d).
- Im Vektorraum  $\text{Abb}(\mathbb{N}, \mathbb{R})$  aller reellen Zahlenfolgen können wir die Familie  $B = (e_i)_{i \in \mathbb{N}}$  aller „Einheitsfolgen“  $e_i = (0, \dots, 0, 1, 0, 0, \dots)$  betrachten, wobei die 1 jeweils an der Stelle  $i$  steht. Wegen der Endlichkeitsbedingung für Linearkombinationen ist  $\text{Lin} B$  dann *nicht* der gesamte Raum aller Folgen, sondern nur die Menge aller Folgen, bei denen nur endlich viele Folgenglieder ungleich 0 sind. Dementsprechend ist  $B$  also auch kein Erzeugendensystem von  $\text{Abb}(\mathbb{N}, \mathbb{R})$ . Die Familie  $B$  ist aber linear unabhängig, denn ist  $\sum_{i \in \mathbb{N}} \lambda_i e_i = (0, 0, 0, \dots)$  (für eine gewisse Summe mit nur endlich vielen  $\lambda_i \neq 0$ ), so folgt durch Vergleich des  $i$ -ten Folgengliedes natürlich sofort  $\lambda_i = 0$  für alle  $i$ .

Dieses Beispiel (b) wirft nun natürlich die Frage auf, ob denn überhaupt eine Basis von  $\text{Abb}(\mathbb{N}, \mathbb{R})$  existiert. In der Tat ist die Antwort auf diese Frage ja: Man kann zeigen, dass *jeder beliebige Vektorraum* in obigem Sinne eine Basis besitzt. Dass wir uns beim Beweis dieser Tatsache in dieser Vorlesung in Satz 14.11 auf den endlich erzeugten Fall beschränkt haben, liegt zum einen daran, dass er für Vektorräume mit unendlichen Basen deutlich komplizierter und abstrakter ist (siehe z. B. [GK] Proposition II.2.22). Zum anderen – und das ist fast der wichtigere Grund – ist der Beweis im allgemeinen Fall *nicht konstruktiv* und daher eigentlich nur von theoretischem Interesse. So weiß man also z. B., dass der Raum  $\text{Abb}(\mathbb{N}, \mathbb{R})$  aller reellen Folgen eine Basis besitzt, aber niemand kann eine solche Basis konkret angeben! Versucht doch einmal, eine Basis dieses Vektorraums zu finden – ihr werdet sehr schnell merken, dass das aussichtslos ist. Zur Erinnerung: Ihr müsstet dazu eine (unendliche) Familie von Folgen hinschreiben, aus der sich nicht die Nullfolge kombinieren lässt, und so dass jede beliebige Folge eine (endliche) Linearkombination der Folgen ist, die ihr ausgewählt habt.

**Aufgabe 14.29.** In  $\mathbb{R}^5$  seien

$$U_1 = \text{Lin} \left( \begin{pmatrix} 2 \\ -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right) \quad \text{und} \quad U_2 = \text{Lin} \left( \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right).$$

- Bestimme jeweils die Dimension und eine Basis von  $U_1 \cap U_2$  und  $U_1 + U_2$ .
- Finde einen 2-dimensionalen Unterraum  $U \leq \mathbb{R}^5$  mit  $U + U_1 = \mathbb{R}^5$ .

**Aufgabe 14.30.** Für ein gegebenes  $n \in \mathbb{N}_{>0}$  betrachten wir die Unterräume

$$U_1 = \{x \in K^n : x_1 = \dots = x_n\} \quad \text{und} \quad U_2 = \{x \in K^n : x_1 + \dots + x_n = 0\}$$

von  $K^n$ , wobei  $x_1, \dots, x_n$  die Koordinaten von  $x$  bezeichnen.

Bestimme Basen und die Dimensionen von  $U_1$ ,  $U_2$  und  $U_1 + U_2$  in den Fällen  $K = \mathbb{R}$  und  $K = \mathbb{Z}_2$ .

**Aufgabe 14.31.**

- (a) Es seien  $V$  ein  $K$ -Vektorraum und  $U_1, U_2 \leq V$  mit  $\dim V = 6$ ,  $\dim U_1 = 5$  und  $\dim U_2 = 3$ . Welche Dimension kann  $U_1 \cap U_2$  haben? Gib für jede solche Möglichkeit ein konkretes Beispiel für  $U_1, U_2$  und  $V$  an.
- (b) Es seien  $U_1, \dots, U_k$  Unterräume eines  $n$ -dimensionalen  $K$ -Vektorraums  $V$ . Zeige, dass

$$\dim(U_1 \cap \dots \cap U_k) \geq \sum_{i=1}^k \dim U_i - (k-1)n.$$

**Aufgabe 14.32.** Es sei  $(a_n)_{n \in \mathbb{N}}$  die sogenannte *Fibonacci-Folge*, die durch  $a_0 = a_1 = 1$  sowie die Rekursionsgleichung

$$a_{n+2} = a_{n+1} + a_n \quad \text{für alle } n \in \mathbb{N} \quad (*)$$

gegeben ist, also die Folge  $(1, 1, 2, 3, 5, 8, 13, 21, \dots)$ . Wir wollen in dieser Aufgabe eine explizite Formel für das  $n$ -te Folgenglied  $a_n$  herleiten.

Es sei dazu  $V \leq \text{Abb}(\mathbb{N}, \mathbb{R})$  der Unterraum aller reellen Zahlenfolgen, die die Rekursionsgleichung  $(*)$  erfüllen (ihr braucht nicht nachzuweisen, dass dies wirklich ein Unterraum ist, solltet euch aber trotzdem kurz überlegen, warum das so ist).

- (a) Für welche  $q \in \mathbb{R}$  liegt die Folge  $(q^n)_{n \in \mathbb{N}} = (1, q, q^2, q^3, \dots)$  in  $V$ ?
- (b) Zeige, dass  $\dim V = 2$  gilt, und bestimme eine Basis von  $V$ .
- (c) Berechne mit (b) eine explizite nicht-rekursive Formel für die Glieder  $a_n$  der Fibonacci-Folge.

**Aufgabe 14.33.** Gib eine Basis des Vektorraums

$$V = \{(a_n)_n \in \text{Abb}(\mathbb{N}, \mathbb{R}) : \text{es gibt ein } N \in \mathbb{N} \text{ mit } a_N = a_{N+1} = \dots\}$$

aller ab irgendeinem Glied konstanten reellen Zahlenfolgen an.

## 15. Lineare Gleichungssysteme und Matrizen

Wie wir nun schon in einigen Fällen gesehen haben, laufen viele Fragestellungen der linearen Algebra auf rechnerischer Seite am Ende auf *lineare Gleichungssysteme* hinaus – z. B. die Überprüfung von Erzeugendensystemen und der linearen Unabhängigkeit in Beispiel 14.2 (b) bzw. Beispiel 14.5 (b), oder die Berechnung von Durchschnitten und Summen von Unterräumen in Algorithmus 14.27. Mit anderen Worten müssen wir also zu gegebenen  $m, n \in \mathbb{N}$  sowie  $a_{i,j}, b_i \in K$  für  $i = 1, \dots, m$  und  $j = 1, \dots, n$  alle  $x_1, \dots, x_n \in K$  bestimmen, die simultan die Gleichungen

$$\begin{aligned} a_{1,1}x_1 + \dots + a_{1,n}x_n &= b_1 \\ a_{2,1}x_1 + \dots + a_{2,n}x_n &= b_2 \\ &\vdots \\ a_{m,1}x_1 + \dots + a_{m,n}x_n &= b_m \end{aligned} \quad (*)$$

erfüllen. Wir wollen daher in diesem Kapitel studieren, wann solche linearen Gleichungssysteme lösbar sind, und wie im Fall der Lösbarkeit die Lösungsmengen aussehen und konkret berechnet werden können.

### 15.A Matrizen

Bevor wir mit der eigentlichen Untersuchung linearer Gleichungssysteme beginnen, sollten wir als Erstes eine effizientere Notation dafür einführen, da die Schreibweise (\*) in der Einleitung oben natürlich sehr unübersichtlich und fehleranfällig ist. Dies ist mit Hilfe von sogenannten *Matrizen* möglich. Die Idee besteht dabei einfach darin, die  $m \cdot n$  Zahlen  $a_{i,j}$  in (\*) in einem einzigen mathematischen Objekt zusammenzufassen.

**Definition 15.1** (Matrizen). Es seien  $m, n \in \mathbb{N}$ .

- (a) Eine  $m \times n$ -**Matrix** über  $K$  ist ein rechteckiges Schema

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix} \quad \text{mit } a_{i,j} \in K \text{ für alle } 1 \leq i \leq m \text{ und } 1 \leq j \leq n$$

mit  $m$  Zeilen und  $n$  Spalten. Analog zur Schreibweise für Zahlenfolgen in Definition ?? ?? bezeichnen wir eine solche Matrix auch mit

$$(a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \quad \text{oder einfach} \quad (a_{i,j})_{i,j}.$$

Es ist eine Konvention, dass wir dabei hinter den Klammern immer zuerst den Zeilenindex und dann den Spaltenindex angeben (Merkregel: „Zeile zuerst, Spalte später“).

Die Menge aller  $m \times n$ -Matrizen mit Einträgen in  $K$  wird mit  $K^{m \times n}$  bezeichnet – auch hier steht in der Bezeichnung also zuerst die Anzahl der Zeilen und dann die Anzahl der Spalten.

- (b) Eine Matrix  $A = (a_{i,j})_{i,j} \in K^{m \times n}$  mit genauso vielen Zeilen wie Spalten (also  $m = n$ ) heißt **quadratisch**. In diesem Fall bezeichnet man die Einträge  $a_{i,i}$  mit  $i = 1, \dots, n$  (also die Einträge von links oben nach rechts unten) als die **Diagonaleinträge** der Matrix.

**Definition 15.2** (Matrixoperationen). Für zwei Matrizen  $A = (a_{i,j})_{i,j}$  und  $B = (b_{i,j})_{i,j}$  in  $K^{m \times n}$  sowie  $\lambda \in K$  definieren wir

- (a) die Addition  $A + B := (a_{i,j} + b_{i,j})_{i,j} \in K^{m \times n}$ ;  
 (b) die Skalarmultiplikation  $\lambda A := \lambda \cdot A := (\lambda \cdot a_{i,j})_{i,j} \in K^{m \times n}$ ;

(c) die **transponierte Matrix**  $A^T := (a_{j,i})_{i,j} \in K^{n \times m}$ .

**Beispiel 15.3.** Die reelle Matrix (z. B. der Größe  $2 \times 3$ ), bei der in jeder Zeile  $i$  alle Einträge gleich  $i$  sind, können wir schreiben als

$$(i)_{i,j} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \end{pmatrix} \in \mathbb{R}^{2 \times 3}.$$

Addition und Skalarmultiplikation für Matrizen (und damit auch für Vektoren) sind einfach komponentenweise definiert, es ist also z. B.

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} + \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 3 & 4 \\ 6 & 7 & 8 \end{pmatrix} \quad \text{und} \quad 2 \cdot \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 2 & 4 & 6 \\ 8 & 10 & 12 \end{pmatrix}$$

in  $\mathbb{R}^{2 \times 3}$ . Die Transposition hingegen vertauscht die Rolle von Zeilen und Spalten in der Matrix, wie z. B. in

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \in \mathbb{R}^{3 \times 2}.$$

#### Bemerkung 15.4.

(a) Offensichtlich ist  $K^{m \times n}$  mit der Addition und Skalarmultiplikation aus Definition 15.2 ein  $K$ -Vektorraum. In der Tat unterscheidet sich dieser Raum von  $K^{mn}$  ja nur dadurch, dass wir die  $m \cdot n$  Einträge der Matrix nicht untereinander, sondern in einem rechteckigen Schema anordnen. Dementsprechend erhält man also auch eine Basis von  $K^{m \times n}$ , indem man alle Matrizen mit einem Eintrag 1 und allen anderen Einträgen 0 nimmt, im Fall  $K^{2 \times 2}$  also z. B.

$$\left( \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right).$$

Insbesondere gilt damit  $\dim K^{m \times n} = mn$ .

- (b) Der Nullvektor im Vektorraum  $K^{m \times n}$  ist offensichtlich die Matrix, in der alle Einträge gleich 0 sind. Diese Matrix wird dementsprechend auch die **Nullmatrix** genannt und einfach als 0 geschrieben.
- (c) Matrizen in  $K^{m \times 1}$  mit nur einer Spalte haben in ihrer Schreibweise die gleiche Form wie Vektoren in  $K^m$ . In der Tat werden wir  $m \times 1$ -Matrizen im Folgenden in der Regel mit Vektoren in  $K^m$  identifizieren.

Bisher gibt es bis auf die Art der Anordnung der Zahlen keinen nennenswerten Unterschied zwischen den Matrizen in  $K^{m \times n}$  und den Vektoren in  $K^{mn}$ . Es gibt jedoch eine sehr wichtige weitere Operation, die auf Matrizen, jedoch nicht auf Vektoren in  $K^{mn}$  definiert ist, nämlich die sogenannte **Matrixmultiplikation**:

**Definition 15.5** (Matrixmultiplikation). Für natürliche Zahlen  $m, n, p$  seien  $A = (a_{i,j})_{i,j} \in K^{m \times n}$  und  $B = (b_{j,k})_{j,k} \in K^{n \times p}$ , d. h. die Matrix  $B$  habe so viele Zeilen wie  $A$  Spalten. Dann definieren wir das Matrixprodukt  $AB$  als

$$AB := A \cdot B := \left( \sum_{j=1}^n a_{i,j} b_{j,k} \right)_{i,k} \in K^{m \times p}$$

(merke: Es wird über die „mittleren Indizes“ summiert, also über den Spaltenindex der ersten und den Zeilenindex der zweiten Matrix). Das Produkt  $AB$  hat also so viele Zeilen wie die erste Matrix und so viele Spalten wie die zweite.

#### Beispiel 15.6.

(a) Es ist z. B.

$$\begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 0 \cdot 4 + 1 \cdot 7 & 0 \cdot 5 + 1 \cdot 8 & 0 \cdot 6 + 1 \cdot 9 \\ 2 \cdot 4 + 3 \cdot 7 & 2 \cdot 5 + 3 \cdot 8 & 2 \cdot 6 + 3 \cdot 9 \end{pmatrix} = \begin{pmatrix} 7 & 8 & 9 \\ 29 & 34 & 39 \end{pmatrix}$$

(hier ist  $m = n = 2$  und  $p = 3$ ). Im Gegensatz dazu ist das Matrixprodukt

$$\begin{pmatrix} 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix}$$

nicht definiert, weil die erste Matrix drei Spalten, die zweite aber nur zwei Zeilen hat. Der Einfachheit halber werden wir in Zukunft bei einem Matrixprodukt  $AB$  stets voraussetzen, dass die zweite Matrix so viele Zeilen hat wie die erste Spalten, und dies nicht jedes Mal wieder erwähnen.

- (b) Sind  $A = (a_{i,j})_{i,j} \in K^{m \times n}$  und  $x \in K^n$  mit Koordinaten  $x_1, \dots, x_n$ , so können wir  $x$  gemäß Bemerkung 15.4 (c) als Matrix mit nur einer Spalte auffassen und erhalten das Matrixprodukt

$$Ax = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{1,1}x_1 + \cdots + a_{1,n}x_n \\ \vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n \end{pmatrix} \in K^{m \times 1} = K^m.$$

Beachte, dass dieser Ausdruck genau die linke Seite eines linearen Gleichungssystems wie in der Einleitung zu diesem Kapitel ist. Wir können lineare Gleichungssysteme in Zukunft also einfach als  $Ax = b$  schreiben, wobei  $A \in K^{m \times n}$  eine gegebene Matrix,  $b \in K^m$  ein gegebener Vektor und  $x \in K^n$  der gesuchte Vektor sind.

- (c) Ein einfacher, aber oft vorkommender und daher wichtiger Spezialfall von (b) ist, wenn  $x = e_j$  für  $j = 1, \dots, n$  der  $j$ -te Einheitsvektor ist: In diesem Fall ist  $Ae_j$  gerade die  $j$ -te Spalte von  $A$ .

Wie üblich nach dem Einführen einer neuen Struktur wollen wir auch hier zunächst einmal die grundlegenden Eigenschaften der Matrixmultiplikation angeben bzw. beweisen.

**Lemma 15.7** (Eigenschaften der Matrixmultiplikation). *Für alle Matrizen  $A, B, C$  passender Größe (d. h. so dass die betrachteten Summen und Produkte definiert sind) sowie  $\lambda \in K$  gilt:*

- (a) (Distributivität)  $(A + B)C = AC + BC$  und  $A(B + C) = AB + AC$ .  
 (b) (Verträglichkeit mit der Skalarmultiplikation)  $(\lambda A)B = A(\lambda B) = \lambda(AB)$ .  
 (c) (Assoziativität)  $(AB)C = A(BC)$ .

*Bei der Multiplikation mehrerer Matrizen werden wir die Klammern daher oft weglassen; das  $n$ -fache Matrixprodukt  $A \cdots A$  für  $n \in \mathbb{N}_{>0}$  schreiben wir als  $A^n$ .*

- (d) (Verträglichkeit mit der Transposition)  $(AB)^T = B^T A^T$ .

Das Matrixprodukt ist jedoch im Allgemeinen nicht kommutativ (aufgrund der Größenbedingung ist das Produkt  $AB$  ja auch nicht einmal genau dann definiert, wenn  $BA$  es ist).

*Beweis.* Der Beweis ergibt sich durch einfaches Nachrechnen. Wir zeigen exemplarisch den zweiten Teil von (a): Für  $A = (a_{i,j})_{i,j} \in K^{m \times n}$  und  $B = (b_{j,k})_{j,k}, C = (c_{j,k})_{j,k} \in K^{n \times p}$  gilt

$$\begin{aligned} A(B+C) &= \left( \sum_{j=1}^n a_{i,j}(b_{j,k} + c_{j,k}) \right)_{i,k} \\ &= \left( \sum_{j=1}^n a_{i,j}b_{j,k} + \sum_{j=1}^n a_{i,j}c_{j,k} \right)_{i,k} \\ &= \left( \sum_{j=1}^n a_{i,j}b_{j,k} \right)_{i,k} + \left( \sum_{j=1}^n a_{i,j}c_{j,k} \right)_{i,k} \\ &= AB + AC. \end{aligned}$$

□

**Bemerkung 15.8 (Blockmatrixmultiplikation).** Es seien  $A \in K^{m \times n}$  und  $B \in K^{n \times p}$  zwei Matrizen, die in „Blockform“

$$A = \left( \begin{array}{c|c} A^{(1,1)} & A^{(1,2)} \\ \hline A^{(2,1)} & A^{(2,2)} \end{array} \right) \quad \text{bzw.} \quad B = \left( \begin{array}{c|c} B^{(1,1)} & B^{(1,2)} \\ \hline B^{(2,1)} & B^{(2,2)} \end{array} \right)$$

mit  $A^{(1,1)} \in K^{m_1 \times n_1}$  und  $B^{(1,1)} \in K^{n_1 \times p_1}$  gegeben sind. Nach Definition der Matrixmultiplikation können wir das Produkt  $AB$  dann ebenfalls in Blockform

$$AB = \left( \begin{array}{c|c} A^{(1,1)}B^{(1,1)} + A^{(1,2)}B^{(2,1)} & A^{(1,1)}B^{(1,2)} + A^{(1,2)}B^{(2,2)} \\ \hline A^{(2,1)}B^{(1,1)} + A^{(2,2)}B^{(2,1)} & A^{(2,1)}B^{(1,2)} + A^{(2,2)}B^{(2,2)} \end{array} \right)$$

schreiben, wobei die Blöcke formal genauso multipliziert und addiert werden, als würde man das Produkt zweier Matrizen der Größe  $2 \times 2$  berechnen: Ist  $A^{(r,s)} = (a_{i,j}^{(r,s)})_{i,j}$  und  $B^{(r,s)} = (b_{j,k}^{(r,s)})_{j,k}$  für  $r, s \in \{1, 2\}$ , so ist z. B. der Eintrag von  $AB$  in Zeile  $i \leq m_1$  und Spalte  $k \leq p_1$  (also im Block links oben) gegeben durch

$$\sum_{j=1}^{n_1} a_{i,j}^{(1,1)} b_{j,k}^{(1,1)} + \sum_{j=1}^{n-n_1} a_{i,j}^{(1,2)} b_{j,k}^{(2,1)},$$

was genau der Eintrag von  $A^{(1,1)}B^{(1,1)} + A^{(1,2)}B^{(2,1)}$  an dieser Stelle ist.

Analog funktioniert diese Regel auch für eine Aufteilung in mehr oder weniger als zwei Blöcke entlang der Zeilen bzw. Spalten. Häufig kommt z. B. der Fall vor, in dem wir die Matrix  $B$  spaltenweise als  $B = (b^{(1)} | \dots | b^{(p)})$  mit  $b^{(1)}, \dots, b^{(p)} \in K^n$  schreiben; in diesem Fall ist dann

$$AB = A(b^{(1)} | \dots | b^{(p)}) = (Ab^{(1)} | \dots | Ab^{(p)}).$$

Wir werden diese Rechenregeln zur Blockmatrixmultiplikation im Folgenden oft verwenden, ohne jedesmal wieder darauf hinzuweisen.

#### Aufgabe 15.9.

- (a) Finde eine Matrix  $A \in \mathbb{R}^{2 \times 2} \setminus \{0\}$  mit  $A^2 := A \cdot A = 0$ .
- (b) Finde Matrizen  $A, B \in \mathbb{R}^{2 \times 2}$  mit  $AB = 0$  und  $BA \neq 0$ .

Mit Hilfe des Matrixprodukts können wir nun zu jeder Matrix  $A \in K^{m \times n}$  zwei Unterräume definieren; einen von  $K^m$  und einen von  $K^n$ :

**Definition 15.10** (Bild und Kern einer Matrix). Für eine Matrix  $A \in K^{m \times n}$  mit  $m, n \in \mathbb{N}$  nennen wir

- (a)  $\text{Im}A := \{Ax : x \in K^n\} \subset K^m$  das **Bild** von  $A$ ;
- (b)  $\text{Ker}A := \{x \in K^n : Ax = 0\} \subset K^n$  den **Kern** von  $A$ .

Die Bezeichnungen kommen dabei von den englischen Worten „image“ und „kernel“.

**Lemma und Definition 15.11** (Rang einer Matrix). Es sei  $A \in K^{m \times n}$  mit  $m, n \in \mathbb{N}$ .

- (a)  $\text{Im}A$  ist der von den Spalten von  $A$  erzeugte Unterraum von  $K^m$ .  
Man nennt seine Dimension  $\text{rk}A := \dim \text{Im}A$  den **Rang** von  $A$ ; die Bezeichnung kommt dabei vom englischen Wort „rank“.
- (b)  $\text{Ker}A$  ist ein Unterraum von  $K^n$ .

*Beweis.*

- (a) Schreiben wir  $A = (a_1 | \dots | a_n)$  mit  $a_1, \dots, a_n \in K^m$ , so ist nach Definition des Matrixprodukts

$$\text{Im}A = \{Ax : x \in K^n\} = \{x_1 a_1 + \dots + x_n a_n : x_1, \dots, x_n \in K\} = \text{Lin}(a_1, \dots, a_n).$$

Also ist  $\text{Im}A$  der von den Spalten von  $A$  erzeugte Unterraum.

- (b) Wir überprüfen die Unterraumkriterien aus Definition 13.8: Zunächst ist natürlich  $A \cdot 0 = 0$ , also  $0 \in \text{Ker}A$ . Für alle  $x, y \in \text{Ker}A$  und  $\lambda \in K$  gilt außerdem  $Ax = Ay = 0$ , also auch

$$A(x+y) = Ax + Ay = 0 \quad \text{und} \quad A \cdot \lambda x = \lambda \cdot Ax = 0,$$

und damit  $x+y \in \text{Ker}A$  und  $\lambda x \in \text{Ker}A$ . □

**Bemerkung 15.12.** Es sei wieder  $A \in K^{m \times n}$  mit  $m, n \in \mathbb{N}$ .

- (a) Nach Lemma 15.11 (a) bilden die Spalten von  $A$  ein Erzeugendensystem von  $\text{Im}A$ . Mit der Basisauswahl aus Satz 14.11 können wir aus diesen Erzeugern eine Basis von  $\text{Im}A$  auswählen, die dann  $\text{rk}A$  Elemente hat – während mehr als  $\text{rk}A$  Elemente nach Folgerung 14.21 (b) natürlich linear abhängig sein müssen. Wir sehen also:

Der Rang  $\text{rk}A$  ist die Maximalzahl linear unabhängiger Spalten von  $A$ .

- (b) Aus (a) folgt natürlich sofort  $\text{rk}A \leq n$ . Da  $\text{Im}A$  ein Unterraum von  $K^m$  ist, gilt nach Lemma 14.23 außerdem auch  $\text{rk}A \leq m$ .
- (c) Die Unterräume  $\text{Im}A$  und  $\text{Ker}A$  lassen sich auch mit Hilfe des linearen Gleichungssystems  $Ax = b$  interpretieren: Nach Definition ist  $\text{Im}A = \{Ax : x \in K^n\}$  die Menge aller  $b \in K^m$ , für die das Gleichungssystem lösbar ist, und  $\text{Ker}A = \{x \in K^n : Ax = 0\}$  die Lösungsmenge im Fall der rechten Seite  $b = 0$ .

**Beispiel 15.13.** Es sei

$$A = \begin{pmatrix} 1 & -1 & 0 \\ 0 & -1 & 1 \end{pmatrix} \in \mathbb{R}^{2 \times 3}.$$

Dann ist

$$\text{Im}A = \text{Lin} \left( \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) = \mathbb{R}^2 \quad \Rightarrow \quad \text{rk}A = \dim \text{Im}A = 2$$

und

$$\text{Ker}A = \{x \in \mathbb{R}^3 : Ax = 0\} = \{x \in \mathbb{R}^3 : x_1 - x_2 = -x_2 + x_3 = 0\} = \text{Lin} \left( \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right).$$

Eine wichtige Eigenschaft des Rangs ist, dass er in folgendem Sinne bei der Multiplikation mit einer anderen Matrix nicht größer werden kann.

**Lemma 15.14** (Rang eines Produkts). *Es seien  $m, n, p \in \mathbb{N}$  sowie  $A \in K^{m \times n}$  und  $B \in K^{n \times p}$ .*

*Dann gilt  $\text{rk}(AB) \leq \text{rk}A$  und  $\text{rk}(AB) \leq \text{rk}B$ .*

*Beweis.* Zunächst gilt nach Definition des Bildes

$$\text{Im}(AB) = \{ABx : x \in K^p\} = \{Ay : y \in \text{Im}B\}.$$

Zum einen folgt daraus nun wegen  $\text{Im}B \subset K^n$

$$\text{Im}(AB) \subset \{Ay : y \in K^n\} = \text{Im}A,$$

und damit  $\text{rk}(AB) \leq \text{rk}A$ . Zum anderen können wir mit  $r := \text{rk}B$  eine Basis  $(b_1, \dots, b_r)$  von  $\text{Im}B$  wählen und erhalten

$$\text{Im}(AB) = \{A(\lambda_1 b_1 + \dots + \lambda_r b_r) : \lambda_1, \dots, \lambda_r \in K\} = \text{Lin}(Ab_1, \dots, Ab_r),$$

nach Bemerkung 14.22 (a) also auch  $\text{rk}(AB) \leq r = \text{rk}B$ . □

**Aufgabe 15.15** (Rang einer Summe). Es seien  $m, n \in \mathbb{N}$  und  $A, B \in K^{m \times n}$ .

Zeige, dass dann  $\text{rk}(A + B) \leq \text{rk}A + \text{rk}B$  gilt.

Zum Ende dieses Abschnitts wollen wir nun noch die formalen Eigenschaften der Matrixmultiplikation betrachten. Nachdem sie nach Lemma 15.7 (c) schon einmal assoziativ ist, ist es dabei naheliegend zu untersuchen, ob sie auch die anderen Gruppenaxiome aus Definition 3.1 erfüllt. Damit sie überhaupt zwischen allen betrachteten Matrizen definiert ist, sollten wir uns dazu auf quadratische Matrizen einer festen Größe beschränken. Ein – wie in Definition 3.1 (b) gefordertes – neutrales Element ist dann schnell gefunden:

**Konstruktion 15.16** (Einheitsmatrix). Für  $n \in \mathbb{N}$  definieren wir die **Einheitsmatrix** der Größe  $n \times n$  als die Matrix, deren Einträge auf der Diagonale gleich 1 und sonst überall 0 sind, also

$$E_n := (e_1 \mid \cdots \mid e_n) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \in K^{n \times n}.$$

Wir schreiben sie oft auch einfach als  $E$ , wenn ihre Größe aus dem Zusammenhang klar ist. In der Literatur ist manchmal auch die Bezeichnung  $I_n$  oder  $I$  üblich. Offensichtlich gilt nun für jede Matrix  $A = (a_1 \mid \cdots \mid a_n) \in K^{m \times n}$

$$AE_n = A \cdot (e_1 \mid \cdots \mid e_n) \stackrel{15.8}{=} (Ae_1 \mid \cdots \mid Ae_n) \stackrel{15.6(c)}{=} (a_1 \mid \cdots \mid a_n) = A,$$

und damit analog auch  $A^T E_n = A^T$ , nach Transponieren mit Lemma 15.7 (d) also  $E_n A = A$ . Die Einheitsmatrizen sind in diesem Sinne also (sogar für nicht-quadratische Matrizen) neutral für die Matrixmultiplikation. Außerdem ist natürlich  $\text{rk} E_n = n$ , denn es gilt  $\text{Im} E_n = \text{Lin}(e_1, \dots, e_n) = K^n$ .

Wie wir jetzt sehen wollen, existieren inverse Matrizen jedoch nicht immer, sondern nur zu Matrizen mit maximalem Rang.

**Satz und Definition 15.17** (Inverse Matrizen). Für jede quadratische Matrix  $A \in K^{n \times n}$  mit  $n \in \mathbb{N}$  sind die folgenden Aussagen äquivalent:

- (a)  $\text{rk} A = n$ .
- (b) Es gibt eine rechtsinverse Matrix  $B$  zu  $A$ , also eine Matrix  $B \in K^{n \times n}$  mit  $AB = E$ .
- (c) Es gibt eine linksinverse Matrix  $B$  zu  $A$ , also eine Matrix  $B \in K^{n \times n}$  mit  $BA = E$ .

Eine quadratische Matrix  $A$  mit diesen Eigenschaften heißt **invertierbar**. In diesem Fall sind die rechts- und linksinverse Matrix zu  $A$  außerdem eindeutig bestimmt und gleich; man schreibt sie als  $A^{-1}$  und nennt sie die zu  $A$  **inverse Matrix**.

*Beweis.*

- (a)  $\Rightarrow$  (b): Nach Voraussetzung gilt  $\text{rk} A = \dim \text{Im} A = n$ , mit Lemma 14.23 also  $\text{Im} A = K^n$ . Für alle  $i = 1, \dots, n$  ist damit  $e_i \in \text{Im} A$ , d. h. es gibt ein  $b_i \in K^n$  mit  $Ab_i = e_i$ . Definieren wir nun  $B := (b_1 \mid \cdots \mid b_n)$ , so ergibt sich wie gewünscht

$$AB = A \cdot (b_1 \mid \cdots \mid b_n) \stackrel{15.8}{=} (Ab_1 \mid \cdots \mid Ab_n) = (e_1 \mid \cdots \mid e_n) = E.$$

- (b)  $\Rightarrow$  (c): Wegen  $\text{rk}(AB) = \text{rk} E = n$  folgt zunächst  $\text{rk} B = n$  aus Lemma 15.14. Nach der schon gezeigten Richtung „(a)  $\Rightarrow$  (b)“ gibt es also eine Matrix  $C \in K^{n \times n}$  mit  $BC = E$ . Dabei ist aber  $C = EC = ABC = AE = A$ , also wie behauptet  $BA = E$ .

- (c)  $\Rightarrow$  (a): Wegen  $\text{rk}(BA) = \text{rk} E = n$  folgt dies sofort aus Lemma 15.14.

Sind diese Bedingungen erfüllt und  $B$  eine rechts- bzw.  $C$  eine linksinverse Matrix zu  $A$ , so gilt außerdem  $C = CE = CAB = EB = B$ , d. h.  $B$  und  $C$  sind eindeutig bestimmt und gleich.  $\square$

**Lemma und Definition 15.18** (Invertierbare Matrizen als Gruppe). Es seien  $n \in \mathbb{N}$  und  $A, B \in K^{n \times n}$  zwei invertierbare Matrizen. Dann gilt:

- (a)  $AB$  ist ebenfalls invertierbar mit  $(AB)^{-1} = B^{-1}A^{-1}$ .
- (b)  $A^{-1}$  ist ebenfalls invertierbar mit  $(A^{-1})^{-1} = A$ .

Insbesondere ist die Menge aller invertierbaren Matrizen in  $K^{n \times n}$  also eine Gruppe bezüglich der Matrixmultiplikation. Man bezeichnet sie mit  $\text{GL}(n, K)$  (der Name kommt vom englischen Begriff „general linear group“).

*Beweis.* Nach Voraussetzung gilt  $A^{-1}A = B^{-1}B = E$  gemäß Satz 15.17 (c). Also folgt

$$(B^{-1}A^{-1})(AB) = B^{-1}A^{-1}AB = B^{-1}EB = B^{-1}B = E$$

und damit Teil (a); Teil (b) ergibt sich direkt aus  $A^{-1}A = E$  mit Satz 15.17 (b).

Damit ist  $GL(n, K)$  eine Gruppe: Die Matrixmultiplikation ist eine Verknüpfung auf dieser Menge nach (a) und assoziativ nach Lemma 15.7 (c), und  $GL(n, K)$  enthält das neutrale Element  $E$  sowie zu jeder Matrix  $A$  die inverse Matrix  $A^{-1}$  nach (b).  $\square$

**Beispiel 15.19.** Die reelle Matrix

$$A = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix} \quad \text{mit} \quad \text{rk}A = \dim \text{Lin} \left( \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \end{pmatrix} \right) = 2$$

ist nach Definition 15.17 invertierbar. In der Tat ist

$$\begin{pmatrix} 1 & -3 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{und damit} \quad A^{-1} = \begin{pmatrix} 1 & -3 \\ 0 & 1 \end{pmatrix}.$$

Wir werden in Satz 15.35 noch sehen, wie inverse Matrizen effizient berechnet werden können.

**Bemerkung 15.20.** Es seien  $n \in \mathbb{N}$  und  $A \in GL(n, K)$  eine invertierbare Matrix.

- (a) Die Inversenbildung vertauscht mit der Transposition, d. h. es gilt  $(A^T)^{-1} = (A^{-1})^T$ , denn es ist

$$(A^{-1})^T A^T \stackrel{15.7(d)}{=} (AA^{-1})^T = E^T = E,$$

und damit ist  $(A^{-1})^T$  die inverse Matrix zu  $A^T$ .

- (b) Für alle  $b \in K^n$  ist das lineare Gleichungssystem  $Ax = b$  äquivalent zu  $x = A^{-1}b$ , hat also die eindeutige Lösung für  $x$ . Diese Aussage ist momentan aber eher für theoretische Zwecke als für konkrete Berechnungen nützlich, zumal wir ja noch keine gute Berechnungsmöglichkeit für inverse Matrizen kennen.

**Aufgabe 15.21.** Für ein gegebenes  $a \in \mathbb{R}$  sei  $A = \begin{pmatrix} a & 2 & 0 \\ 0 & a & 1 \\ 0 & 0 & a \end{pmatrix} \in \mathbb{R}^{3 \times 3}$ .

- (a) Berechne  $A^n$  für alle  $n \in \mathbb{N}_{>0}$ .

(Hinweis: Bestimme  $A^n$  für kleine Werte von  $n$ , stelle damit eine Vermutung für die allgemeine Form von  $A^n$  auf, und beweise diese Vermutung dann mit vollständiger Induktion.)

- (b) Für welche  $a$  ist die Matrix  $A$  invertierbar? Bestimme in diesen Fällen auch die inverse Matrix  $A^{-1}$ .

**Aufgabe 15.22.** Für  $n_1, n_2 \in \mathbb{N}_{>0}$  und gegebene Matrizen  $A_1 \in K^{n_1 \times n_1}$ ,  $A_2 \in K^{n_2 \times n_2}$  und  $B \in K^{n_1 \times n_2}$  setzen wir

$$A := \begin{pmatrix} A_1 & B \\ 0 & A_2 \end{pmatrix} \in K^{n \times n}$$

mit  $n = n_1 + n_2$ .

Zeige, dass  $A$  genau dann invertierbar ist, wenn  $A_1$  und  $A_2$  invertierbar sind, und bestimme in diesem Fall die inverse Matrix  $A^{-1}$  aus  $A_1$ ,  $A_2$  und  $B$ .

**Aufgabe 15.23.** Es seien  $m, n \in \mathbb{N}$  und  $A \in K^{m \times n}$  eine Matrix vom Rang  $r$ . Für alle  $k \in \mathbb{N}$  mit  $k \leq m$  und  $k \leq n$  verstehen wir unter einer  $k \times k$ -Teilmatrix von  $A$  eine Matrix, die man erhält, indem man aus  $A$  eine beliebige Auswahl von Zeilen und Spalten herausstreicht, so dass eine quadratische Matrix der Größe  $k \times k$  übrig bleibt.

Zeige, dass es genau dann eine invertierbare  $k \times k$ -Teilmatrix von  $A$  gibt, wenn  $k \leq r$  gilt.

## 15.B Der Gauß-Algorithmus

Wie bereits angekündigt wollen wir jetzt untersuchen, wie man lineare Gleichungssysteme – die wir in Matrixschreibweise nach Beispiel 15.6 (b) nun als  $Ax = b$  mit  $A \in K^{m \times n}$  und  $b \in K^m$  schreiben werden – explizit numerisch lösen kann. Wie ihr natürlich wisst, besteht die Strategie hierbei darin, die gegebenen Gleichungen mit einem geeigneten Algorithmus z. B. durch Addition, Subtraktion oder Multiplikation mit Skalaren so umzuformen und zu kombinieren, dass ein äquivalentes Gleichungssystem entsteht, dessen Lösung man leicht ablesen kann. In der Matrixform entspricht nun jede Gleichung einer Zeile der Matrix  $A$ , und demzufolge wollen wir also die Zeilen der Matrix umformen und miteinander kombinieren können. Diese Zeilenumformungen entsprechen in unserer Schreibweise einfachen Matrixmultiplikationen, die wir jetzt einführen wollen.

**Konstruktion 15.24** (Elementarmatrizen). Es seien  $m, n \in \mathbb{N}$  und  $A = (a_{i,j})_{i,j} \in K^{m \times n}$ .

(a) Für  $k \in \{1, \dots, m\}$  und  $\lambda \in K \setminus \{0\}$  setzen wir

$$F_k(\lambda) := \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \lambda & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix} \in K^{m \times m},$$

wobei der Eintrag  $\lambda$  in Zeile und Spalte  $k$  steht. Die Matrix  $F_k(\lambda)$  ist also nichts weiter als die Einheitsmatrix, bei der der Eintrag 1 in der  $k$ -ten Zeile und Spalte durch ein  $\lambda \neq 0$  ersetzt wurde. Mit dieser Matrix ist

$$F_k(\lambda) \cdot A = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \lambda & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{k,1} & \cdots & a_{k,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix} = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ \lambda a_{k,1} & \cdots & \lambda a_{k,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix},$$

d. h. die Multiplikation von  $A$  mit  $F_k(\lambda)$  von links entspricht der Multiplikation der  $k$ -ten Zeile von  $A$  mit  $\lambda$ .

(b) Für  $k, l \in \{1, \dots, m\}$  mit  $k \neq l$  und  $\lambda \in K$  setzen wir

$$F_{k,l}(\lambda) := \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & \lambda & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix} \in K^{m \times m},$$

wobei der Eintrag  $\lambda$  in Zeile  $k$  und Spalte  $l$  steht, d. h. diesmal haben wir in der Einheitsmatrix den Eintrag 0 in Zeile  $k$  und Spalte  $l$  durch  $\lambda$  ersetzt. (Beachte, dass der Eintrag  $\lambda$  für  $k < l$  oberhalb und für  $k > l$  unterhalb der Diagonalen steht; wir haben in der Matrix oben der Einfachheit halber nur den Fall  $k < l$  dargestellt.) In diesem Fall ist

$$F_{k,l}(\lambda) \cdot A = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & \lambda & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \vdots & & \vdots \\ a_{k,1} & \cdots & a_{k,n} \\ \vdots & & \vdots \\ a_{l,1} & \cdots & a_{l,n} \\ \vdots & & \vdots \end{pmatrix} = \begin{pmatrix} \vdots & & \vdots \\ a_{k,1} + \lambda a_{l,1} & \cdots & a_{k,n} + \lambda a_{l,n} \\ \vdots & & \vdots \\ a_{l,1} & \cdots & a_{l,n} \\ \vdots & & \vdots \end{pmatrix},$$

d. h. die Multiplikation von  $A$  mit  $F_{k,l}(\lambda)$  von links entspricht der Addition des  $\lambda$ -fachen von Zeile  $l$  zu Zeile  $k$ .

Die Matrizen  $F_k(\lambda)$  für  $\lambda \in K \setminus \{0\}$  sowie  $F_{k,l}(\lambda)$  für  $k \neq l$  und  $\lambda \in K$  heißen **Elementarmatrizen**. Es gibt sie in jeder (quadratischen) Größe  $m \times m$ ; zur Vereinfachung der Schreibweise deuten wir diese Größe in der Notation  $F_k(\lambda)$  bzw.  $F_{k,l}(\lambda)$  aber nicht an.

Man sagt, dass  $F_k(\lambda) \cdot A$  und  $F_{k,l}(\lambda) \cdot A$  aus  $A$  durch eine **elementare Zeilenumformung** entstehen.

**Bemerkung 15.25.**

- (a) Die Elementarmatrizen sind invertierbar mit

$$(F_k(\lambda))^{-1} = F_k\left(\frac{1}{\lambda}\right) \quad \text{und} \quad (F_{k,l}(\lambda))^{-1} = F_{k,l}(-\lambda).$$

Dies folgt direkt aus Konstruktion 15.24: Wenn wir z. B. das Matrixprodukt

$$F_k\left(\frac{1}{\lambda}\right) \cdot F_k(\lambda) = F_k\left(\frac{1}{\lambda}\right) \cdot F_k(\lambda) \cdot E$$

bilden, multiplizieren wir die  $k$ -te Zeile in der Einheitsmatrix zuerst mit  $\lambda$  und dann mit  $\frac{1}{\lambda}$ , d. h. es kommt insgesamt wieder die Einheitsmatrix heraus – also ist  $F_k\left(\frac{1}{\lambda}\right) \cdot F_k(\lambda) = E$ , und damit  $(F_k(\lambda))^{-1} = F_k\left(\frac{1}{\lambda}\right)$ .

Analog zeigt man die Aussage über das Inverse von  $F_{k,l}(\lambda)$ .

- (b) Führen wir mit einer Matrix  $A$  mehrere elementare Zeilenumformungen aus, so ist das Ergebnis nach Konstruktion 15.24 eine Matrix  $F_1 \cdot \dots \cdot F_r \cdot A$  mit Elementarmatrizen  $F_1, \dots, F_r$ . Da diese Elementarmatrizen nach (a) invertierbar sind, ist das Produkt  $F := F_1 \cdot \dots \cdot F_r$  nach Lemma 15.18 (a) ebenfalls wieder invertierbar. Das Ergebnis der Zeilenumformungen ist also die Matrix  $FA$ , d. h. wir haben  $A$  von links mit einer invertierbaren Matrix  $F$  multipliziert. In der Tat werden wir in Folgerung 15.36 noch sehen, dass jede invertierbare Matrix ein Produkt von Elementarmatrizen ist, also jede Multiplikation mit einer invertierbaren Matrix von links als Abfolge von elementaren Zeilenumformungen geschrieben werden kann.
- (c) Das Vertauschen von zwei Zeilen  $k, l \in \{1, \dots, m\}$  mit  $k \neq l$  in einer Matrix  $A \in K^{m \times n}$  lässt sich als Folge von elementaren Zeilenumformungen wie in (b) realisieren: Sind  $a_1, \dots, a_m$  die Zeilen von  $A$ , so erhalten wir

$$A = \begin{pmatrix} \vdots \\ a_k \\ \vdots \\ a_l \\ \vdots \end{pmatrix} \xrightarrow{F_{k,l}(1)} \begin{pmatrix} \vdots \\ a_k + a_l \\ \vdots \\ a_l \\ \vdots \end{pmatrix} \xrightarrow{F_{l,k}(-1)} \begin{pmatrix} \vdots \\ a_k + a_l \\ \vdots \\ -a_k \\ \vdots \end{pmatrix} \xrightarrow{F_{k,l}(1)} \begin{pmatrix} \vdots \\ a_l \\ \vdots \\ -a_k \\ \vdots \end{pmatrix} \xrightarrow{F_l(-1)} \begin{pmatrix} \vdots \\ a_l \\ \vdots \\ a_k \\ \vdots \end{pmatrix}.$$

- (d) Man kann die Elementarmatrizen auch verwenden, um analog **elementare Spaltenumformungen** durchzuführen: Multipliziert man eine Matrix  $A \in K^{m \times n}$  mit einer Elementarmatrix  $F_k(\lambda)$  (der Größe  $n$ ) von rechts, so entspricht dies der Multiplikation der  $k$ -ten Spalte von  $A$  mit  $\lambda$ . Analog addiert man das  $\lambda$ -fache von Spalte  $k$  zu Spalte  $l$  von  $A$ , wenn man das Produkt  $A \cdot F_{k,l}(\lambda)$  bildet. Wie in (b) lässt sich eine Abfolge von elementaren Spaltenumformungen also durch Multiplikation mit einer invertierbaren Matrix von rechts beschreiben.

Mit derartigen Zeilenumformungen wollen wir jetzt eine Matrix bzw. ein lineares Gleichungssystem auf eine möglichst einfache Form bringen. Die folgende Form wird dabei herauskommen.

**Definition 15.26 (Zeilenstufenform).** Es seien  $m, n, r \in \mathbb{N}$  und  $A = (a_{i,j})_{i,j} \in K^{m \times n}$ .



Fall 2: Es sind nicht alle Einträge in der ersten Spalte von  $A$  gleich 0.

- (a) Falls der Eintrag  $a_{1,1}$  links oben in  $A$  gleich Null ist, vertauschen wir zwei Zeilen von  $A$  so, dass dieser Eintrag nicht mehr gleich 0 ist (nach Bemerkung 15.25 (c) ist dies durch elementare Zeilenumformungen machbar).
- (b) Wir dividieren die erste Zeile durch  $a_{1,1}$  und erhalten eine Matrix der Form

$$\left( \begin{array}{c|c} 1 & * \\ a_{2,1} & * \\ \vdots & * \\ a_{m,1} & * \end{array} \right).$$

- (c) Wir subtrahieren von jeder Zeile  $k \in \{2, \dots, m\}$  das  $a_{k,1}$ -fache der ersten Zeile und bekommen dadurch

$$\left( \begin{array}{c|c} 1 & * \dots * \\ 0 & \\ \vdots & \\ 0 & \end{array} \begin{array}{c} A' \end{array} \right)$$

mit einer Matrix  $A' \in K^{(m-1) \times n}$ .

- (d) Nach Induktionsvoraussetzung können wir  $A'$  jetzt mit elementaren Zeilenumformungen in Zeilenstufenform bringen. Damit ist aber auch die gesamte Matrix bereits in Zeilenstufenform – wollen wir also nur diese normale, nicht-reduzierte Zeilenstufenform erreichen, so sind wir damit fertig.
- (e) Wollen wir eine reduzierte Zeilenstufenform erreichen, bringen wir  $A'$  gemäß der Induktionsvoraussetzung sogar in reduzierte Zeilenstufenform und bekommen eine Matrix der Form

$$\left( \begin{array}{c|cccc} 1 & * \dots * & * & * \dots * & * & * \dots * \\ 0 & & 1 & * \dots * & 0 & * \dots * \\ \vdots & & & & \vdots & \vdots \\ \vdots & & & & \vdots & \vdots \\ 0 & & & & 1 & * \dots * \end{array} \right).$$

Subtrahieren wir nun geeignete Vielfache der Zeilen  $2, \dots, m$  von der ersten Zeile, so können wir damit dann noch die Einträge in den Stufenspalten der ersten Zeile zu Null machen und erhalten so auch die reduzierte Zeilenstufenform. □

33

**Beispiel 15.28.** Wir wollen die Matrix

$$A = \begin{pmatrix} -1 & 2 & 2 & 3 & 4 \\ -1 & 2 & 3 & 4 & 7 \\ 1 & -2 & 1 & 0 & 5 \end{pmatrix} \in \mathbb{R}^{3 \times 5}$$

mit elementaren Zeilenumformungen in Zeilenstufenform bringen. Dazu können wir nach dem Verfahren im Beweis von Satz 15.27 wie folgt vorgehen (die Notation  $Z3 - 3Z2 \rightarrow Z3$  bedeutet dabei z. B., dass wir das Dreifache der zweiten Zeile von der dritten subtrahieren):

$$\begin{aligned} \begin{pmatrix} -1 & 2 & 2 & 3 & 4 \\ -1 & 2 & 3 & 4 & 7 \\ 1 & -2 & 1 & 0 & 5 \end{pmatrix} &\xrightarrow{-Z1 \rightarrow Z1} \begin{pmatrix} 1 & -2 & -2 & -3 & -4 \\ -1 & 2 & 3 & 4 & 7 \\ 1 & -2 & 1 & 0 & 5 \end{pmatrix} \xrightarrow{\begin{matrix} Z2+Z1 \rightarrow Z2 \\ Z3-Z1 \rightarrow Z3 \end{matrix}} \begin{pmatrix} 1 & -2 & -2 & -3 & -4 \\ 0 & 0 & 1 & 1 & 3 \\ 0 & 0 & 3 & 3 & 9 \end{pmatrix} \\ &\xrightarrow{Z3-3Z2 \rightarrow Z3} \begin{pmatrix} 1 & -2 & -2 & -3 & -4 \\ 0 & 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

Möchten wir eine Matrix in reduzierter Zeilenstufenform erreichen, so addieren wir schließlich noch das Doppelte der zweiten Zeile zur ersten und erhalten

$$\left( \begin{array}{ccc|cc} 1 & -2 & 0 & -1 & 2 \\ 0 & 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right).$$

**Aufgabe 15.29.** Bringe die reelle Matrix

$$A = \begin{pmatrix} 0 & 1 & 2 & 1 & 0 & 5 \\ 1 & 0 & 3 & 0 & 0 & 1 \\ -1 & 1 & -1 & 3 & 1 & 9 \\ 1 & 0 & 3 & 1 & 1 & 4 \end{pmatrix}$$

mit elementaren Zeilenumformungen auf reduzierte Zeilenstufenform.

## 15.C Algorithmen der linearen Algebra

Nachdem wir jetzt gesehen haben, wie man eine Matrix  $A \in K^{m \times n}$  mit dem Gauß-Algorithmus durch elementare Zeilenumformungen in (evtl. reduzierte) Zeilenstufenform bringen kann, wollen wir nun untersuchen, wie man damit die in der linearen Algebra auftretenden Fragestellungen numerisch lösen kann. Für die Beweise dieser Aussagen ist es dabei nützlich, die Zeilenstufenform von  $A$  wie in Satz 15.27 als  $FA$  für eine invertierbare Matrix  $F \in \text{GL}(m, K)$  zu schreiben – beachte jedoch, dass wir diese Matrix  $F$  für konkrete Anwendungen oft gar nicht explizit kennen müssen, sondern es genügt, wenn wir wie in Beispiel 15.28 direkt die Zeilenstufenform  $FA$  bestimmen.

Wir beginnen dabei mit dem Rang von  $A$ . Dieser ist sehr einfach zu bestimmen, er ist nämlich wie im folgenden Satz immer die Anzahl der Stufen in einer zugehörigen Zeilenstufenform. Gleichzeitig liefert dieses Argument nicht nur den Rang – also die Dimension von  $\text{Im}A$  – sondern auch eine Basis von  $\text{Im}A$ , nämlich indem wir die Spalten von  $A$  nehmen, die den Stufenspalten entsprechen.

**Satz 15.30** (Bild und Rang einer Matrix). *Es seien  $m, n \in \mathbb{N}$  und  $A = (a_1 | \dots | a_n) \in K^{m \times n}$ . Wir bringen die Matrix  $A$  in Zeilenstufenform  $FA$  mit  $F \in \text{GL}(m, K)$ ; dabei seien  $r$  die Anzahl der Stufen und  $k_1 < \dots < k_r$  die Stufenspalten von  $FA$ .*

*Dann ist die Familie  $(a_{k_1}, \dots, a_{k_r})$  eine Basis von  $\text{Im}A$ . Insbesondere gilt also  $\text{rk}A = r$ .*

*Beweis.* Da wir jede Matrix in Zeilenstufenform mit dem Gauß-Algorithmus aus Satz 15.27 in reduzierte Zeilenstufenform mit denselben Stufenspalten überführen können, können wir annehmen, dass  $FA = (Fa_1 | \dots | Fa_n)$  sogar in reduzierter Zeilenstufenform ist. Für alle  $i = 1, \dots, r$  ist die  $i$ -te Stufenspalte der Matrix  $FA$  dann gleich  $Fa_{k_i} = e_i$ . Wir müssen nun zwei Dinge zeigen:

- (a)  $B := (a_{k_1}, \dots, a_{k_r})$  ist ein Erzeugendensystem von  $\text{Im}A$ : Da  $FA$  in Zeilenstufenform mit  $r$  Stufen ist, haben die Spalten von  $FA$  höchstens in den ersten  $r$  Zeilen Einträge ungleich 0. Es gilt also nicht nur  $\text{Lin}(Fa_{k_1}, \dots, Fa_{k_r}) = \text{Lin}(e_1, \dots, e_r)$ , sondern sogar

$$\text{Lin}(Fa_{k_1}, \dots, Fa_{k_r}) = \text{Lin}(e_1, \dots, e_r) = \text{Lin}(Fa_1, \dots, Fa_n).$$

Multiplikation mit  $F^{-1}$  von links ergibt also wie behauptet

$$\text{Lin}B = \text{Lin}(a_{k_1}, \dots, a_{k_r}) = \text{Lin}(a_1, \dots, a_n) = \text{Im}A.$$

- (b)  $B$  ist linear unabhängig: Es seien  $\lambda_1, \dots, \lambda_r \in K$  mit  $\lambda_1 a_{k_1} + \dots + \lambda_r a_{k_r} = 0$ . Multiplikation mit  $F$  von links liefert dann sofort

$$\lambda_1 Fa_{k_1} + \dots + \lambda_r Fa_{k_r} = 0 \quad \Rightarrow \quad \lambda_1 e_1 + \dots + \lambda_r e_r = 0,$$

und damit  $\lambda_1 = \dots = \lambda_r = 0$ . □



Alternativ hätten wir  $(A|b)$  auch nur in nicht notwendig reduzierte Zeilenstufenform bringen können (was natürlich weniger Arbeit ist). Dann drückt Zeile  $i$  des umgeformten Gleichungssystems die  $i$ -te Stufenvariable  $x_{k_i}$  durch alle freien Variablen  $x_{j_1}, \dots, x_{j_{n-r}}$  und die weiter rechts stehenden Stufenvariablen  $x_{k_{i+1}}, \dots, x_{k_r}$  aus. Wir müssen diese Zeilen dann also von unten nach oben durchgehen und so der Reihe nach  $x_{k_r}, \dots, x_{k_1}$  aus den freien Variablen und den bereits bestimmten Stufenvariablen berechnen (was dann wieder etwas mehr Arbeit ist).

Im Fall (b) der Lösbarkeit haben wir oben in (1) den Wert jeder Variablen einzeln angegeben. Fassen wir nun noch alle diese Komponenten von  $x$  zu einem Vektor zusammen, so können wir die Lösung auch schreiben als

$$x = w - \sum_{l=1}^{n-r} x_{j_l} w_{j_l} \in K^n, \quad (2)$$

wobei die Koordinaten der Vektoren  $w_{j_1}, \dots, w_{j_{n-r}}, w$  gegeben sind durch:

- die  $k_i$ -te Koordinate von  $w_{j_l}$  bzw.  $w$  ist  $a_{i,j_l}$  bzw.  $b_i$  für  $i = 1, \dots, r$  und  $j = 1, \dots, n-r$  (die Koordinate  $k_i$  von (2) wird damit genau (1));
- die  $j_l$ -te Koordinate von  $w_{j_l}$  ist  $-1$  für  $j = 1, \dots, n-r$ , und alle anderen Koordinaten zu Nichtstufenvariablen in  $w_{j_l}$  und  $w$  sind 0 (die Koordinate  $j_l$  von (2) wird damit die triviale Gleichung  $x_{j_l} = x_{j_l}$ ).

Als Merkregel können wir die Vektoren  $w_{j_1}, \dots, w_{j_{n-r}}, w$  damit als die entsprechenden Spalten der folgenden Matrix  $W = (w_1 | \dots | w_n | w) \in K^{n \times (n+1)}$  erhalten:

Zu einer gegebenen Matrix  $(FA|Fb) \in K^{m \times (n+1)}$  in reduzierter Zeilenstufenform mit Stufenspalten  $k_1 < \dots < k_r$  konstruieren wir eine Matrix  $W \in K^{n \times (n+1)}$ , indem wir für  $i = 1, \dots, r$  Zeile  $i$  von  $(FA|Fb)$  in Zeile  $k_i$  von  $W$  schreiben; die restlichen Einträge von  $W$  sind  $-1$  auf der Diagonalen und 0 sonst.

Die Lösungsmenge des Gleichungssystems  $Ax = b$  ist mit diesen Bezeichnungen nach (2) dann also  $w + \text{Lin}(w_{j_1}, \dots, w_{j_{n-r}})$ ; im Fall  $b = 0$  für die Berechnung von  $\text{Ker} A$  ist dabei natürlich auch  $w = 0$ .

Die hierbei auftretenden Vektoren  $w_{j_1}, \dots, w_{j_{n-r}}$  sind darüber hinaus linear unabhängig: Sind nämlich  $\lambda_1, \dots, \lambda_{n-r} \in K$  mit  $\lambda_1 w_{j_1} + \dots + \lambda_{n-r} w_{j_{n-r}} = 0$ , so besagt die  $j_l$ -te Koordinate dieser Gleichung genau  $\lambda_l \cdot (-1) = 0$  für alle  $l = 1, \dots, n-r$ .

Fassen wir die Ergebnisse unserer gesamten Konstruktion zusammen, haben wir damit insgesamt also gezeigt:

**Satz 15.33** (Kern einer Matrix und lineare Gleichungssysteme). *Es seien  $m, n \in \mathbb{N}$  sowie  $A \in K^{m \times n}$  und  $b \in K^m$  mit  $r := \text{rk} A$ . Wir bringen die erweiterte Koeffizientenmatrix  $(A|b)$  in reduzierte Zeilenstufenform mit Nichtstufenspalten  $1 \leq j_1 < \dots < j_{n-r} \leq n$  von  $A$ , und bilden dazu die Matrix  $W = (w_1 | \dots | w_n | w) \in K^{n \times (n+1)}$  aus Algorithmus 15.32. Dann gilt:*

- (a) Die Familie  $(w_{j_1}, \dots, w_{j_{n-r}})$  ist eine Basis von  $\text{Ker} A$ . Insbesondere ist die Dimension des Kerns also  $\dim \text{Ker} A = n - r = n - \text{rk} A$ , und es folgt die **Dimensionsformel für Matrizen**

$$\dim \text{Im} A + \dim \text{Ker} A = n.$$

- (b) Das lineare Gleichungssystem  $Ax = b$  ist genau dann lösbar, wenn  $\text{rk}(A|b) = \text{rk} A$  gilt. In diesem Fall ist die Lösungsmenge gegeben durch

$$\{x \in K^n : Ax = b\} = w + \text{Ker} A$$

und damit ein verschobener Unterraum der Dimension  $n - r$ .

**Beispiel 15.34.** Über dem Körper  $\mathbb{R}$  seien

$$A = \begin{pmatrix} -1 & 2 & 2 & 3 \\ -1 & 2 & 3 & 4 \\ 1 & -2 & 1 & 0 \end{pmatrix} \quad \text{und} \quad b = \begin{pmatrix} 4 \\ 7 \\ 5 \end{pmatrix}.$$

Die erweiterte Koeffizientenmatrix  $(A|b)$  ist dann genau die Matrix, die wir in Beispiel 15.28 in reduzierte Zeilenstufenform

$$\left( \begin{array}{cccc|c} 1 & -2 & 0 & -1 & 2 \\ 0 & 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) \quad (*)$$

gebracht haben; im linken Teil der Matrix haben wir dabei auch  $A$  in reduzierte Zeilenstufenform gebracht. Nach dieser Umformung besagt das Gleichungssystem  $Ax = b$  nun genau

$$x_1 - 2x_2 - x_4 = 2 \quad \text{und} \quad x_3 + x_4 = 3,$$

was sich natürlich auch ohne einen weiteren Algorithmus leicht lösen lässt: Die Nichtstufenvariablen  $x_2$  und  $x_4$  sind frei wählbar, und die anderen beiden Variablen gegeben durch  $x_1 = 2x_2 + x_4 + 2$  und  $x_3 = -x_4 + 3$ .

Wollen wir trotzdem das obige „Kochrezept“ aus Algorithmus 15.32 bzw. Satz 15.33 verwenden, lesen wir an (\*) zunächst ab, dass  $\text{rk}(A|b) = 2 = \text{rk}A$  gilt und das Gleichungssystem damit lösbar ist. Wir bilden die Matrix  $W \in \mathbb{R}^{4 \times 5}$ , indem wir die Zeilen 1 und 2 von (\*) in die Zeilen 1 und 3 von  $W$  schreiben (unten eingekreist, die Stufeneinsen stehen nun also genau auf der Diagonalen); die übrigen Einträge von  $W$  sind  $-1$  auf der Diagonalen und  $0$  sonst:

$$W = (w_1 | w_2 | w_3 | w_4 | w) = \left( \begin{array}{ccccc} 1 & -2 & 0 & -1 & 2 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & -1 & 0 \end{array} \right).$$

Eine Basis von  $\text{Ker}A$  ist nun nach Satz 15.33 (a) gegeben durch die oben grau markierten Spalten  $(w_2, w_4)$  von  $W$  (also die Spalten der Zusatzeinträge  $-1$ ), und die Lösungsmenge des Gleichungssystems  $Ax = b$  ist  $w + \text{Lin}(w_2, w_4)$  nach Satz 15.33 (b). Wollen wir nur den Kern von  $A$  bestimmen, können wir natürlich komplett mit  $A$  statt mit  $(A|b)$  rechnen und haben demzufolge sowohl in (\*) als auch in  $W$  keine separate rechte Spalte.

Auch die Berechnung inverser Matrizen ist mit Zeilenstufenformen einfach:

**Satz 15.35** (Inverse Matrix). *Es seien  $n \in \mathbb{N}$  und  $A \in K^{n \times n}$ . Wir bringen  $A$  mit elementaren Zeilenumformungen in reduzierte Zeilenstufenform  $FA$  mit  $F \in \text{GL}(n, K)$ , führen aber alle Umformungen mit der erweiterten Matrix  $(A|E) \in K^{n \times (2n)}$  durch, so dass wir am Ende die Matrix  $F(A|E) = (FA|FE)$  erhalten.*

*Dann ist  $A$  genau dann invertierbar, wenn nach der Umformung in der linken Hälfte der erweiterten Matrix die Einheitsmatrix  $FA = E$  steht. In diesem Fall steht in der rechten Hälfte die inverse Matrix  $FE = F = A^{-1}$ .*

*Beweis.* Nach Satz 15.17 ist  $A$  genau dann invertierbar, wenn  $\text{rk}A = n$  gilt, also  $FA$  in reduzierter Zeilenstufenform mit  $n$  Stufen ist. Die einzige  $n \times n$ -Matrix in reduzierter Zeilenstufenform mit  $n$  Stufen ist aber die Einheitsmatrix. Also ist  $A$  genau dann invertierbar, wenn  $FA = E$  gilt, und die inverse Matrix  $F = A^{-1}$  steht in diesem Fall in der rechten Hälfte  $FE$  der umgeformten erweiterten Matrix.  $\square$

**Folgerung 15.36.** *Jede invertierbare Matrix ist ein Produkt von Elementarmatrizen.*

*Beweis.* Zu jedem  $A \in \text{GL}(n, K)$  gibt es nach Satz 15.35 ein Produkt  $F = F_1 \cdot \dots \cdot F_k$  von Elementarmatrizen  $F_1, \dots, F_k$  mit  $F = A^{-1}$ . Da die inversen Matrizen von Elementarmatrizen nach Bemerkung 15.25 (a) wieder Elementarmatrizen sind, ist damit aber auch  $A = F^{-1} = F_k^{-1} \cdot \dots \cdot F_1^{-1}$  ein Produkt von Elementarmatrizen.  $\square$

**Beispiel 15.37.** Es sei

$$A = \begin{pmatrix} 1 & 1 \\ 3 & 4 \end{pmatrix} \in \mathbb{R}^{2 \times 2}.$$



aus Beispiel 15.28, und bringen sie diesmal mit elementaren Spaltenumformungen in Spaltenstufenform:

$$\begin{aligned} \begin{pmatrix} -1 & 2 & 2 & 3 & 4 \\ -1 & 2 & 3 & 4 & 7 \\ 1 & -2 & 1 & 0 & 5 \end{pmatrix} &\xrightarrow{-S_1 \rightarrow S_1} \begin{pmatrix} 1 & 2 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 & 7 \\ -1 & -2 & 1 & 0 & 5 \end{pmatrix} \xrightarrow{\substack{S_2 - 2S_1 \rightarrow S_2 \\ S_3 - 2S_1 \rightarrow S_3 \\ S_4 - 3S_1 \rightarrow S_4 \\ S_5 - 4S_1 \rightarrow S_5}} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 3 \\ -1 & 0 & 3 & 3 & 9 \end{pmatrix} \\ &\xrightarrow{S_2 \leftrightarrow S_3} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 3 \\ -1 & 3 & 0 & 3 & 9 \end{pmatrix} \xrightarrow{\substack{S_4 - S_2 \rightarrow S_4 \\ S_5 - 3S_2 \rightarrow S_5}} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 3 \\ -1 & 3 & 0 & 3 & 9 \end{pmatrix}. \end{aligned}$$

Nach Satz 15.38 ist also

$$\left( \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix} \right)$$

eine Basis von  $\text{Im} A$  (die sich von der aus Beispiel 15.31 unterscheidet). Beachte, dass wir diesmal im Gegensatz zu Beispiel 15.31 die Spalten der umgeformten Matrix und nicht die von  $A$  nehmen müssen!

**Bemerkung 15.40** (Invarianz des Rangs einer Matrix unter Transposition). Als Folgerung aus unseren beiden Algorithmen zur Berechnung des Rangs einer Matrix haben wir in Satz 15.38 gesehen, dass  $\text{rk} A = \text{rk}(A^T)$  für jede Matrix  $A \in K^{m \times n}$  gilt. Diese wichtige Eigenschaft des Rangs ist ganz und gar nicht aus der Definition  $\text{rk} A = \dim \text{Im} A$  bzw.  $\text{rk}(A^T) = \dim \text{Im}(A^T)$  offensichtlich – in der Tat sind  $\text{Im} A$  und  $\text{Im}(A^T)$  komplett verschiedene Unterräume, die ja sogar in unterschiedlichen Vektorräumen  $K^m$  bzw.  $K^n$  liegen. In der Praxis bedeutet sie, dass jede Eigenschaft des Rangs, die für die Spalten einer Matrix gilt, auch für die Zeilen gilt: So ist der Rang z. B. nicht nur wie in Bemerkung 15.12 (a) die Maximalzahl linear unabhängiger Spalten von  $A$ , sondern auch die Maximalzahl linear unabhängiger Zeilen von  $A$ .

34

**Algorithmus 15.41** (Weitere Algorithmen der linearen Algebra). Die numerische Berechnung unserer anderen bisher aufgetretenen Fragestellungen der linearen Algebra lässt sich leicht auf die in diesem Abschnitt behandelten Sätze bzw. Algorithmen zurückführen:

- (a) *Eindeutige Lösbarkeit von linearen Gleichungssystemen*: Es seien  $m, n \in \mathbb{N}$  sowie  $A \in K^{m \times n}$  und  $b \in K^m$ . Das lineare Gleichungssystem  $Ax = b$  ist nach Satz 15.33 genau dann eindeutig lösbar, wenn  $\text{rk}(A|b) = \text{rk} A$  und  $\dim \text{Ker} A = 0$  gilt, was nach der Dimensionsformel für Matrizen äquivalent ist zu  $\text{rk}(A|b) = \text{rk} A = n$ .

Ist  $A$  sogar quadratisch (also  $m = n$ ), so folgt dabei aus  $\text{rk} A = n$  bereits  $\text{rk}(A|b) = n$ , denn wegen  $(A|b) \in K^{n \times (n+1)}$  kann der Rang dieser Matrix nicht größer werden als  $n$ . Damit ist das Gleichungssystem  $Ax = b$  im quadratischen Fall genau dann eindeutig lösbar, wenn  $\text{rk} A = n$  gilt, also  $A$  invertierbar ist. Die eindeutige Lösung ist dann natürlich  $x = A^{-1}b$  wie in Bemerkung 15.20 (b).

- (b) *Lineare Abhängigkeit*: Es seien  $r, n \in \mathbb{N}$  und  $x_1, \dots, x_r \in K^n$ . Die Familie  $(x_1, \dots, x_r)$  ist genau dann linear unabhängig, wenn sie eine Basis von  $\text{Lin}(x_1, \dots, x_r) = \text{Im} A$  mit  $A := (x_1 | \dots | x_r)$  ist. Mit Folgerung 14.21 ist dies äquivalent zu  $\text{rk} A = r$ .
- (c) *Basisauswahl*: Es seien wieder  $r, n \in \mathbb{N}$  und  $x_1, \dots, x_r \in K^n$  Vektoren, die damit ein Erzeugendensystem von  $U := \text{Lin}(x_1, \dots, x_r) = \text{Im} A$  mit  $A := (x_1 | \dots | x_r)$  bilden. Um aus diesen Vektoren eine Basis von  $U$  auszuwählen, können wir einfach mit Satz 15.30 eine Basis von  $U$  bestimmen, die aus geeigneten Spalten von  $A$  besteht.
- (d) *Basisergänzung*: Nun seien die Vektoren  $x_1, \dots, x_r \in K^n$  linear unabhängig; wir wollen sie zu einer Basis von  $K^n$  ergänzen. Auch dazu bilden wir wieder die Matrix  $A = (x_1 | \dots | x_r)$ , die nach (b) Rang  $r$  hat. Diesmal bringen wir sie aber wie in Satz 15.38 mit elementaren Spaltenumformungen in eine Spaltenstufenform mit  $r$  Stufen.

Sind dabei  $j_1, \dots, j_{n-r}$  die Nichtstufenzeilen, so ergänzen die entsprechenden Einheitsvektoren  $e_{j_1}, \dots, e_{j_{n-r}}$  die Vektoren  $x_1, \dots, x_r$  zu einer Basis von  $K^n$ : Diese Vektoren liefern nämlich die fehlenden Stufen in der Spaltenstufenform, d.h. wir können die erweiterte Matrix  $(x_1 | \dots | x_r | e_{j_1} | \dots | e_{j_{n-r}})$  mit elementaren Spaltenumformungen in eine Spaltenstufenform mit  $n$  Stufen bringen – was genau bedeutet, dass ihr Rang gleich  $n$  ist und die Vektoren in ihren Spalten damit eine Basis von  $K^n$  sind.

- (e) *Durchschnitt und Summe von Unterräumen:* Wie wir in Algorithmus 14.27 schon gesehen haben, führt die Berechnung des Durchschnitts von Unterräumen auf ein lineares Gleichungssystem, und die Berechnung der Summe auf eine Basisauswahl. Beides können wir nun mit Hilfe der Methoden aus diesem Kapitel behandeln.

**Aufgabe 15.42.**

- (a) Berechne Basen von  $\text{Im}A$  und  $\text{Ker}A$  für die reelle Matrix  $A = \begin{pmatrix} 2 & 1 & -1 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \\ 1 & 2 & -1 \end{pmatrix}$ .
- (b) Berechne den Rang der reellen Matrix  $A = \begin{pmatrix} 0 & b & b & b \\ a & 0 & b & b \\ a & a & 0 & b \end{pmatrix}$  in Abhängigkeit von  $a, b \in \mathbb{R}$ .
- (c) Zeige, dass die reelle Matrix  $A = \begin{pmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 8 \end{pmatrix}$  invertierbar ist, und berechne  $A^{-1}$ .

**Aufgabe 15.43.** Gegeben seien in  $\mathbb{R}^5$  die Vektoren

$$x_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad x_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 2 \\ 0 \end{pmatrix}, \quad x_3 = \begin{pmatrix} 0 \\ 2 \\ 2 \\ 2 \\ 1 \end{pmatrix}, \quad x_4 = \begin{pmatrix} -3 \\ 0 \\ 3 \\ 2 \\ 0 \end{pmatrix}, \quad x_5 = \begin{pmatrix} -1 \\ 2 \\ 1 \\ 1 \\ 3 \end{pmatrix};$$

wir setzen  $A = (x_1 | x_2 | x_3 | x_4 | x_5) \in \mathbb{R}^{5 \times 5}$ .

- (a) Wähle aus  $(x_1, \dots, x_5)$  eine Basis von  $\text{Im}A$  aus und berechne eine Basis von  $\text{Im}(A^T)$ .
- (b) Bestimme die Dimension von  $\text{Lin}(x_1, x_2, x_3, x_4)$ .
- (c) Gilt  $\text{Lin}(x_1, x_4) \cap \text{Lin}(x_1, x_5) = \{0\}$ ?

(Hinweis: Die Lösung zu allen Aufgabenteilen lässt sich aus derselben Rechnung ableiten.)

**Aufgabe 15.44.** In  $\mathbb{R}^4$  betrachten wir die Unterräume

$$U_1 = \text{Lin} \left( \begin{pmatrix} 1 \\ 2 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ -1 \\ -1 \\ 3 \end{pmatrix}, \begin{pmatrix} -1 \\ 4 \\ -5 \\ 6 \end{pmatrix} \right) \quad \text{und} \quad U_2 = \{x \in \mathbb{R}^4 : x_1 - x_2 - x_3 = x_1 + x_2 - 3x_4 = 0\}.$$

- (a) Berechne Basen von  $U_1, U_2, U_1 + U_2$  und  $U_1 \cap U_2$ .
- (b) Ergänze die in (a) gefundene Basis von  $U_2$  zu einer Basis von  $U_1 + U_2$ .

**Aufgabe 15.45** (Elementare Codierungstheorie). In dieser Aufgabe seien  $K = \mathbb{Z}_2$  der Körper mit 2 Elementen aus Beispiel 3.6 (b), sowie

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{pmatrix} \in K^{7 \times 4}.$$

- (a) Berechne eine Matrix  $B \in K^{3 \times 7}$  mit  $\text{Im}A = \text{Ker}B$ .
- (b) Zeige, dass es zu jedem  $y \in K^7$  genau ein  $x \in K^4$  gibt, für das sich  $y$  von  $Ax$  in höchstens einer Koordinate unterscheidet.

(Hinweis: der Vektor  $By \in K^3$  sollte erkennen lassen, wo dieser Unterschied liegt.)

Diese Aufgabe beinhaltet die Grundidee der sogenannten Codierungstheorie. Angenommen, wir wollen einen Vektor  $x \in K^4$  (also 4 Bits) über eine gestörte Leitung an einen Empfänger übertragen. Wenn wir dann statt der 4 Bits  $x$  die 7 Bits  $Ax$  übertragen und beim Empfänger statt  $Ax$  fälschlicherweise  $y$  ankommt, kann dieser nach (b) die ursprüngliche Information  $x$  auch dann noch aus  $y$  rekonstruieren, wenn bei ihm ein Bit falsch angekommen ist. Auf vielen Datenträgern, Strichcodes oder QR-Codes sind die Daten in einer solchen Art codiert, damit Lesefehler automatisch erkannt und korrigiert werden können.

**Aufgabe 15.46.** Es sei  $A \in K^{m \times n}$  eine Matrix vom Rang  $r$ .

Beweise, dass es dann Matrizen  $B \in K^{m \times r}$  und  $C \in K^{r \times n}$  vom Rang  $r$  gibt mit  $A = BC$ .

(Hinweis: Zeige die Aussage zunächst für den Fall, dass  $A$  in Zeilenstufenform ist.)

**Aufgabe 15.47.** Es seien  $m, n \in \mathbb{N}_{>0}$  und  $A \in K^{m \times n}$ . Wir wählen eine beliebige Basis  $(x_1, \dots, x_k)$  von  $\text{Ker}A$  und ergänzen sie zu einer Basis  $(x_1, \dots, x_k, y_1, \dots, y_l)$  von  $K^n$ .

Zeige durch direktes Nachprüfen der Definitionen, dass  $(Ay_1, \dots, Ay_l)$  dann eine Basis von  $\text{Im}A$  ist.

Insbesondere erhalten wir auf diese Art also einen alternativen Beweis der Dimensionsformel für Matrizen (siehe Satz 15.33 (a))

$$\dim \text{Ker}A + \dim \text{Im}A = n.$$

## 16. Lineare Abbildungen

In den letzten Kapiteln haben wir ausführlich Vektorräume untersucht – also die grundlegende Struktur, mit der sich die lineare Algebra befasst – und dabei auch gesehen, wie man konkret mit ihnen rechnen kann. Um nun verschiedene Vektorräume miteinander in Verbindung setzen zu können, müssen wir als Nächstes Abbildungen zwischen Vektorräumen untersuchen, die mit den gegebenen Vektorraumstrukturen verträglich sind. Dies wollen wir in diesem Kapitel tun.

In der Tat ist dies eine sehr allgemeine Vorgehensweise der Mathematik: Immer wenn man eine neue mathematische Struktur (wie z. B. Gruppen, Körper, oder jetzt hier die Vektorräume) einführt, wird man als Erstes zwei Dinge untersuchen:

- die sogenannten *Unterstrukturen*, d. h. Teilmengen, die selbst wieder die betrachtete Struktur haben (in der linearen Algebra also die Untervektorräume aus Abschnitt 13.B), und
- die sogenannten *Morphismen*, d. h. Abbildungen, die diese Struktur erhalten.

Wir haben dies in Abschnitt 3.A nur deswegen für Gruppen und Körper nicht getan, weil wir in dieser Vorlesung nur Vektorräume, aber nicht Gruppen und Körper ausführlich studieren wollen. Diejenigen von euch, die auch die Parallelvorlesung „Algebraische Strukturen“ hören, haben dort aber z. B. auch schon Untergruppen und Morphismen von Gruppen kennengelernt – und werden im Vergleich zur linearen Algebra sicher feststellen, dass sich Untervektorräume und Morphismen von Vektorräumen in vielerlei Hinsicht sehr ähnlich verhalten.

### 16.A Morphismen von Vektorräumen

In der linearen Algebra besteht die „Struktur“ eines Vektorraums gemäß Definition 13.1 in der Existenz einer Vektoraddition und einer Skalarmultiplikation mit gewissen Eigenschaften. Die mit dieser Struktur verträglichen Abbildungen sind also genau die im Sinne der folgenden Definition *linearen* Abbildungen.

**Definition 16.1** (Lineare Abbildungen bzw. Morphismen). Es seien  $V$  und  $W$  zwei Vektorräume über demselben Grundkörper  $K$ . Man nennt eine Abbildung  $f: V \rightarrow W$  eine **lineare Abbildung** oder **Morphismus** (oder **(Vektorraum-)Homomorphismus**), wenn für alle  $x, y \in V$  und  $\lambda \in K$  gilt, dass

$$f(x+y) = f(x) + f(y) \quad (\text{„}f \text{ ist verträglich mit der Vektoraddition“})$$

und

$$f(\lambda x) = \lambda f(x) \quad (\text{„}f \text{ ist verträglich mit der Skalarmultiplikation“}).$$

Die Menge aller solchen Morphismen mit Startraum  $V$  und Zielraum  $W$  wird mit  $\text{Hom}_K(V, W)$  bezeichnet (oder auch nur mit  $\text{Hom}(V, W)$ , wenn der Grundkörper aus dem Zusammenhang klar ist).

Ist  $V = K^n$  für ein  $n \in \mathbb{N}$ , so schreiben wir statt  $f \left( \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right)$  der Einfachheit halber oft nur  $f \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ .

**Bemerkung 16.2.** Setzt man  $\lambda = 0$  in Definition 16.1 ein, so erhält man sofort, dass  $f(0) = 0$  für jeden Morphismus  $f: V \rightarrow W$  gelten muss. Für ein festes  $b \in V \setminus \{0\}$  ist also z. B. die Verschiebeabbildung  $f: V \rightarrow V$ ,  $x \mapsto x + b$  wegen  $f(0) = b \neq 0$  nie ein Morphismus.

Beachte, dass es hier also (leider) unterschiedliche Notationen in der Analysis und der linearen Algebra gibt: Eine Funktion der Form  $f: \mathbb{R} \rightarrow \mathbb{R}$ ,  $x \mapsto ax + b$  mit  $a, b \in \mathbb{R} \setminus \{0\}$  würde man in der Analysis im Sinne von Definition 3.24 als linear bezeichnen, sie ist in der linearen Algebra gemäß Definition 16.1 aber wegen  $f(0) \neq 0$  keine lineare Abbildung!

**Beispiel 16.3.**

- (a) Für beliebige  $K$ -Vektorräume  $V$  und  $W$  ist die Nullabbildung  $f: V \rightarrow W, x \mapsto 0$  natürlich immer ein Morphismus, ebenso für gleichen Start- und Zielraum die identische Abbildung  $\text{id}_V: V \rightarrow V, x \mapsto x$ .
- (b) Für eine Matrix  $A \in K^{m \times n}$  bezeichnen wir mit  $f_A$  die Abbildung

$$f_A: K^n \rightarrow K^m, x \mapsto Ax.$$

Sie ist ein Morphismus, denn für alle  $x, y \in K^n$  und  $\lambda \in K$  gilt

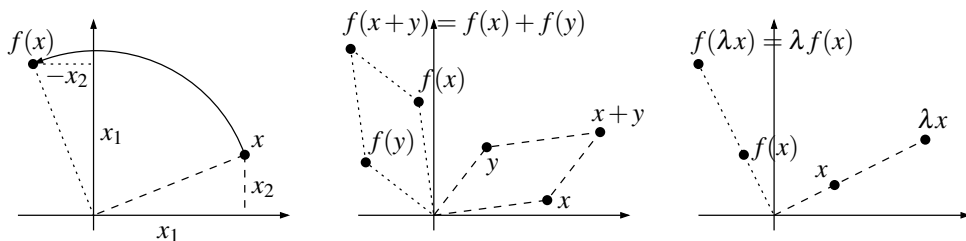
$$f_A(x+y) = A(x+y) = Ax + Ay = f_A(x) + f_A(y) \quad \text{und} \quad f_A(\lambda x) = A(\lambda x) = \lambda(Ax) = \lambda f_A(x).$$

Dies ist sicher das mit Abstand wichtigste Beispiel für eine lineare Abbildung – in der Tat werden wir in Abschnitt 16.C noch sehen, dass sogar *jede* lineare Abbildung von  $K^n$  nach  $K^m$  so geschrieben werden kann.

Konkret erhalten wir z. B. für die Matrix

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathbb{R}^{2 \times 2} \quad \text{die lineare Abbildung} \quad f_A: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}.$$

Geometrisch beschreibt sie wie im Bild unten links eine Vierteldrehung um den Ursprung. An den anderen beiden Bildern kann man die Morphismuseigenschaft auch gut anschaulich ablesen: Im mittleren Bild ist z. B. das gepunktete Parallelogramm aus der Drehung des gestrichelten entstanden, und der äußerste Punkt ergibt sich damit sowohl durch Addition der Punkte  $f(x)$  und  $f(y)$  als auch durch Drehung von  $x+y$ , d. h. es ist  $f(x) + f(y) = f(x+y)$ . Entsprechendes gilt für die Skalarmultiplikation im rechten Bild.



Anschaulich ist damit auch schon erkennbar, dass Drehungen um andere Winkel um den Ursprung ebenfalls Morphismen sein sollten. Wir werden solche allgemeinen Drehungen später in Beispiel ?? und Abschnitt ?? noch untersuchen.

- (c) Die Abbildung

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto x_1^2 + x_2$$

ist nicht linear, denn es ist z. B.

$$f \begin{pmatrix} -1 \\ 0 \end{pmatrix} + f \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1 + 1 = 2 \neq 0 = f \begin{pmatrix} 0 \\ 0 \end{pmatrix} = f \left( \begin{pmatrix} -1 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right).$$

- (d) Für alle  $m, n \in \mathbb{N}$  ist die Transposition von Matrizen

$$f: K^{m \times n} \rightarrow K^{n \times m}, A \mapsto A^\top$$

eine lineare Abbildung, denn für alle  $A = (a_{i,j})_{i,j}, B = (b_{i,j})_{i,j} \in K^{m \times n}$  und  $\lambda \in K$  gilt

$$f(A+B) = (A+B)^\top = (a_{j,i} + b_{j,i})_{i,j} = (a_{j,i})_{i,j} + (b_{j,i})_{i,j} = A^\top + B^\top = f(A) + f(B)$$

$$\text{und} \quad f(\lambda A) = (\lambda a_{j,i})_{i,j} = \lambda (a_{j,i})_{i,j} = \lambda A^\top = \lambda f(A).$$

- (e) Zum Vektorraum  $\text{Pol}(\mathbb{R}, \mathbb{R})$  aller reellen Polynome aus Beispiel 13.12 (c) betrachten wir die Abbildung  $f: \text{Pol}(\mathbb{R}, \mathbb{R}) \rightarrow \text{Pol}(\mathbb{R}, \mathbb{R})$ ,  $\varphi \mapsto \varphi'$ , die jedem Polynom  $\varphi: x \mapsto \sum_{k=0}^n a_k x^k$  seine Ableitung

$$\varphi': x \mapsto \sum_{k=1}^n k a_k x^{k-1} \quad (*)$$

zuordnet.

Wenn ihr die Ableitung bereits aus der Analysis kennt, wisst ihr aus Beispiel ?? ?? schon, dass die Ableitung eines Polynoms durch (\*) gegeben ist, und dass  $f$  eine lineare Abbildung ist, da nach Satz ?? ?? und Beispiel ?? ?? für alle differenzierbaren Funktionen  $\varphi$  und  $\psi$  sowie  $\lambda \in \mathbb{R}$  sowohl  $(\varphi + \psi)' = \varphi' + \psi'$  als auch  $(\lambda \varphi)' = \lambda \varphi'$  gilt.

Wenn ihr die Ableitung aus der Analysis noch nicht kennt, könnt ihr (\*) für die Zwecke der linearen Algebra einfach als *Definition* der Ableitung eines Polynoms ansehen. Man rechnet dann schnell nach, dass  $f$  wirklich eine lineare Abbildung ist: Für zwei Polynome  $\varphi(x) = \sum_{k=0}^n a_k x^k$  und  $\psi(x) = \sum_{k=0}^n b_k x^k$  (wobei  $n$  der größere der beiden Grade ist, so dass sowohl  $\varphi$  als auch  $\psi$  so geschrieben werden können) sowie  $\lambda \in \mathbb{R}$  ist

$$\begin{aligned} (\varphi + \psi)'(x) &= \sum_{k=1}^n k(a_k + b_k)x^{k-1} = \sum_{k=1}^n k a_k x^{k-1} + \sum_{k=1}^n k b_k x^{k-1} = \varphi'(x) + \psi'(x) \\ \text{und } (\lambda \varphi)'(x) &= \sum_{k=1}^n k(\lambda a_k)x^{k-1} = \lambda \sum_{k=1}^n k a_k x^{k-1} = \lambda \varphi'(x). \end{aligned}$$

**Aufgabe 16.4.** Untersuche, ob die folgenden Abbildungen  $f$  zwischen  $\mathbb{R}$ -Vektorräumen linear sind:

- (a)  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} ax_2^2 + bx_2 \\ cx_1x_2 \end{pmatrix}$  in Abhängigkeit von fest gegebenen  $a, b, c \in \mathbb{R}$ .  
 (b)  $f: V \rightarrow V$ ,  $\varphi \mapsto f(\varphi)$  mit  $f(\varphi)(x) = x^2 \varphi(x)$  für alle  $x \in \mathbb{R}$ .

**Aufgabe 16.5.** Wir betrachten die Vektoren

$$x_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, x_2 = \begin{pmatrix} 1 \\ 3 \\ 4 \end{pmatrix}, x_3 = \begin{pmatrix} 3 \\ 1 \\ 5 \end{pmatrix} \in \mathbb{R}^3 \quad \text{sowie} \quad y_1 = \begin{pmatrix} 4 \\ 3 \end{pmatrix}, y_2 = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \in \mathbb{R}^2.$$

Man zeige:

- (a) Es gibt einen Morphismus  $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  mit  $f(x_i) = e_i$  für alle  $i \in \{1, 2, 3\}$ .  
 (b) Es gibt einen Morphismus  $g: \mathbb{R}^3 \rightarrow \mathbb{R}^2$  mit  $g(x_i) = y_i$  für alle  $i \in \{1, 2, 3\}$ .

(Die Morphismen müssen nicht explizit angegeben werden.)

Wir wollen nun einige elementare Eigenschaften von Morphismen zeigen und beginnen damit, dass Bilder und Urbilder (im Sinne von Definition 2.11) von Unterräumen unter Morphismen immer wieder Unterräume sind.

**Lemma 16.6** (Bilder und Urbilder von Unterräumen). *Es sei  $f: V \rightarrow W$  eine lineare Abbildung. Dann gilt:*

- (a) Für jeden Unterraum  $U$  von  $V$  ist  $f(U)$  ein Unterraum von  $W$ .  
 (b) Für jeden Unterraum  $U$  von  $W$  ist  $f^{-1}(U)$  ein Unterraum von  $V$ .

*Beweis.* Die Aussagen ergeben sich durch einfaches Nachprüfen der Definition 13.8:

- (a) Wegen  $0 \in U$  ist nach Bemerkung 16.2 zunächst auch  $0 = f(0) \in f(U)$ . Wir zeigen nun die Abgeschlossenheit von  $f(U)$  bezüglich der Vektoraddition. Es seien dazu  $x, y \in f(U)$ , d. h.  $x = f(u)$  und  $y = f(v)$  für gewisse  $u, v \in U$ . Dann ist auch  $u + v \in U$ , und damit folgt  $x + y = f(u) + f(v) = f(u + v) \in f(U)$ .

Genauso zeigt man die Abgeschlossenheit unter der Skalarmultiplikation.

- (b) Wegen  $f(0) = 0 \in U$  ist zunächst einmal  $0 \in f^{-1}(U)$ . Wir zeigen jetzt die Abgeschlossenheit von  $f^{-1}(U)$  unter der Vektoraddition. Dazu seien  $x, y \in f^{-1}(U)$ , d. h.  $x, y \in V$  mit  $f(x), f(y) \in U$ . Dann ist auch  $f(x+y) = f(x) + f(y) \in U$ , also  $x+y \in f^{-1}(U)$ .

Analog ergibt sich die Abgeschlossenheit unter der Skalarmultiplikation.  $\square$

Die wichtigsten Spezialfälle dieses Lemmas sind die folgenden, die wir in Definition 15.10 und Lemma 15.11 schon im Zusammenhang mit Matrizen kennengelernt hatten:

**Definition 16.7** (Bild, Rang und Kern eines Morphismus). Es sei  $f: V \rightarrow W$  eine lineare Abbildung von  $K$ -Vektorräumen.

- (a) Die Menge  $\text{Im } f := f(V) = \{f(x) : x \in V\}$  heißt das **Bild** von  $f$ ; nach Lemma 16.6 (a) ist sie ein Unterraum von  $W$ . Die Dimension dieses Unterraums heißt der **Rang** von  $f$ , geschrieben  $\text{rk } f := \dim \text{Im } f \in \mathbb{N} \cup \{\infty\}$ .
- (b) Die Menge  $\text{Ker } f := f^{-1}(\{0\}) = \{x \in V : f(x) = 0\}$  heißt der **Kern** von  $f$ ; nach Lemma 16.6 (b) ist sie ein Unterraum von  $V$ .

**Bemerkung 16.8.** Für eine Matrix  $A \in K^{m \times n}$  mit  $m, n \in \mathbb{N}$  hatten wir in Beispiel 16.3 (b) die lineare Abbildung  $f_A: K^n \rightarrow K^m$ ,  $x \mapsto Ax$  konstruiert. Vergleichen wir nun Definition 16.7 mit den Definitionen 15.10 und 15.11, so sehen wir also unmittelbar, dass

$$\text{Im } f_A = \text{Im } A, \quad \text{rk } f_A = \text{rk } A \quad \text{und} \quad \text{Ker } f_A = \text{Ker } A$$

gilt. Insbesondere können wir damit Bild, Rang und Kern solcher linearen Abbildungen mit den Methoden aus Abschnitt 15.C auch explizit berechnen.

Offensichtlich ist eine lineare Abbildung  $f: V \rightarrow W$  nach Definition genau dann surjektiv, wenn  $\text{Im } f = W$  gilt. Wir wollen jetzt ein analoges Kriterium auch für die Injektivität zeigen.

**Lemma 16.9.** Eine lineare Abbildung  $f: V \rightarrow W$  ist genau dann injektiv, wenn  $\text{Ker } f = \{0\}$  gilt.

*Beweis.*

„ $\Rightarrow$ “: Ist  $f$  injektiv, so hat der Nullvektor höchstens ein Urbild unter  $f$ . Wegen  $f(0) = 0$  ist das Urbild des Nullvektors also genau der Nullvektor, d. h. es ist  $\text{Ker } f = \{0\}$ .

„ $\Leftarrow$ “: Es sei  $\text{Ker } f = \{0\}$ . Weiterhin seien  $x, y \in V$  mit  $f(x) = f(y)$ . Wegen der Linearität von  $f$  gilt dann  $f(x-y) = f(x) - f(y) = 0$ , mit  $\text{Ker } f = \{0\}$  also  $x-y=0$ . Damit folgt  $x=y$ , d. h.  $f$  ist injektiv.  $\square$

**Lemma 16.10** (Umkehrabbildungen und Verkettungen). Es sei  $f: V \rightarrow W$  ein Morphismus von  $K$ -Vektorräumen. Dann gilt:

- (a) Ist  $f$  bijektiv, so ist auch die Umkehrabbildung  $f^{-1}$  ein Morphismus.
- (b) Ist  $g: W \rightarrow Z$  ein weiterer Morphismus von  $K$ -Vektorräumen, so ist auch  $g \circ f: V \rightarrow Z$  ein Morphismus.

*Beweis.*

- (a) Es seien  $x, y \in W$ ; wir setzen  $u = f^{-1}(x)$  und  $v = f^{-1}(y)$ , also  $x = f(u)$  und  $y = f(v)$ . Dann gilt

$$\begin{aligned} f^{-1}(x+y) &= f^{-1}(f(u) + f(v)) \\ &= f^{-1}(f(u+v)) && (f \text{ ist ein Morphismus}) \\ &= u+v && (f^{-1} \text{ ist die Umkehrabbildung von } f) \\ &= f^{-1}(x) + f^{-1}(y). \end{aligned}$$

Analog zeigt man die Verträglichkeit mit der Skalarmultiplikation.

(b) Für alle  $x, y \in V$  gilt

$$\begin{aligned} g(f(x+y)) &= g(f(x) + f(y)) && (f \text{ ist ein Morphismus}) \\ &= g(f(x)) + g(f(y)) && (g \text{ ist ein Morphismus}). \end{aligned}$$

Genauso ergibt sich die Verträglichkeit mit der Skalarmultiplikation.  $\square$

**Aufgabe 16.11.** Man zeige: Sind  $f, g: V \rightarrow W$  zwei lineare Abbildungen, so gilt

- (a)  $\text{Im}(f + g) \leq \text{Im} f + \text{Im} g$ ;  
 (b)  $\text{Ker} f \cap \text{Ker} g \leq \text{Ker}(f + g)$ .

Weiterhin gebe man in beiden Fällen ein Beispiel an, das zeigt, dass man im Allgemeinen nicht „ $\leq$ “ durch „ $=$ “ ersetzen kann.

**Aufgabe 16.12.** Für ein gegebenes  $n \in \mathbb{N}_{>0}$  betrachten wir die Abbildung

$$f: \text{Pol}_{n-1}(\mathbb{R}, \mathbb{R}) \rightarrow \mathbb{R}^n, \varphi \mapsto \begin{pmatrix} \varphi(1) \\ \vdots \\ \varphi(n) \end{pmatrix}.$$

- (a) Zeige, dass  $f$  linear ist.  
 (b) Bestimme Kern und Bild von  $f$ , und gib dabei von jedem Element von  $\text{Im} f$  explizit ein Urbild an.

**Aufgabe 16.13.** Es sei  $f: V \rightarrow W$  eine lineare Abbildung zwischen  $K$ -Vektorräumen. Ferner sei  $U$  ein Unterraum von  $V$  mit  $U \cap \text{Ker} f = \{0\}$  und  $U + \text{Ker} f = V$ .

Zeige, dass die Abbildung  $f|_U: U \rightarrow \text{Im} f$  bijektiv ist.

**Aufgabe 16.14.** Beweise für jede lineare Abbildung  $f: V \rightarrow V$  mit gleichem Start- und Zielraum, dass

$$f \circ f = f \iff \text{Ker}(\text{id}_V - f) = \text{Im} f.$$

Wie wir nun zum Abschluss dieses Abschnitts noch sehen wollen, haben bijektive Morphismen wie in Lemma 16.10 (a) in der Praxis eine besondere Bedeutung. Sie haben daher auch einen besonderen Namen:

**Definition 16.15** (Isomorphismen). Es seien  $V$  und  $W$  zwei  $K$ -Vektorräume.

- (a) Einen bijektiven Morphismus  $f: V \rightarrow W$  (der nach Lemma 16.10 (a) also einen Umkehrmorphismus  $f^{-1}: W \rightarrow V$  besitzt) bezeichnet man als **(Vektorraum-)Isomorphismus**.  
 (b)  $V$  und  $W$  heißen **isomorph** (in Zeichen:  $V \cong W$ ), wenn es einen Isomorphismus  $f: V \rightarrow W$  zwischen ihnen gibt.

**Beispiel 16.16.** Anschaulich bedeutet ein Isomorphismus  $f$  zwischen zwei Vektorräumen  $V$  und  $W$ , dass diese beiden Räume „als Vektorräume ununterscheidbar“ sind: Die Objekte in  $V$  und  $W$  sind zwar unterschiedlich benannt, aber in allen Rechnungen können wir jederzeit mit der bijektiven Abbildung  $f$  bzw. der inversen Abbildung  $f^{-1}$  zwischen den beiden Darstellungen in  $V$  und  $W$  hin- und herwechseln, ohne das Endergebnis zu ändern. Die folgenden beiden Beispiele verdeutlichen dies; viele weitere werden wir im nächsten Abschnitt noch kennenlernen.

(a) Der Unterraum

$$V = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} \in \mathbb{R}^3 : x_1, x_2 \in \mathbb{R} \right\} \leq \mathbb{R}^3 \quad \text{ist mit} \quad f: V \rightarrow \mathbb{R}^2, \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

isomorph zu  $\mathbb{R}^2$ . In der Tat ist in diesem Beispiel offensichtlich, dass  $f$  linear und bijektiv ist. Auch anschaulich ist in diesem Fall klar, dass  $V$  und  $\mathbb{R}^2$  „im Prinzip ununterscheidbar“ sind, denn beide Räume sind einfach die reelle Ebene – die im Fall von  $V$  lediglich als Koordinatenebene in den  $\mathbb{R}^3$  eingebettet ist.

- (b) Für alle  $m, n \in \mathbb{N}$  ist der Raum  $K^{m \times n}$  aller  $m \times n$ -Matrizen über  $K$  isomorph zu  $K^{mn}$ : Ein Isomorphismus  $f: K^{m \times n} \rightarrow K^{mn}$  ist einfach dadurch gegeben, dass man die Einträge einer Matrix von links oben nach rechts unten nun untereinander in einen Vektor in  $K^{mn}$  schreibt. Auch hier ist klar, dass  $f$  linear und bijektiv ist, also dass die Anordnung der Zahlen – einmal als rechteckiges Schema und einmal untereinander geschrieben – nichts an der Vektorraumstruktur ändert, da Addition und Skalarmultiplikation in beiden Fällen einfach komponentenweise ausgeführt werden. (Dass es in  $K^{m \times n}$  eine Matrixmultiplikation gibt, in  $K^{mn}$  jedoch nicht, spielt hierbei keine Rolle, da dies nicht Teil der Vektorraumaxiome ist.)

35

Wie erwartet haben isomorphe Vektorräume als „ununterscheidbare Vektorräume“ auch immer die gleiche Dimension. Dies wollen wir zum Abschluss dieses Abschnitts noch formal beweisen.

**Lemma 16.17.** *Es sei  $f: V \rightarrow W$  ein Morphismus zwischen endlich-dimensionalen  $K$ -Vektorräumen.*

- (a) *Ist  $f$  surjektiv, so bildet  $f$  jedes Erzeugendensystem von  $V$  auf ein Erzeugendensystem von  $W$  ab. Insbesondere gilt dann also  $\dim V \geq \dim W$ .*
- (b) *Ist  $f$  injektiv, so bildet  $f$  jede linear unabhängige Familie in  $V$  auf eine linear unabhängige Familie in  $W$  ab. Insbesondere gilt dann also  $\dim V \leq \dim W$ .*
- (c) *Ist  $f$  ein Isomorphismus, so bildet  $f$  jede Basis von  $V$  auf eine Basis von  $W$  ab. Insbesondere gilt dann also  $\dim V = \dim W$ .*

*Beweis.* Es sei  $B = (x_1, \dots, x_n)$  eine Familie von Vektoren in  $V$ . Wir betrachten die abgebildete Familie  $f(B) := (f(x_1), \dots, f(x_n))$  von Vektoren in  $W$ .

- (a) Es seien  $B$  ein Erzeugendensystem von  $V$  und  $y \in W$  beliebig. Da  $f$  surjektiv ist, gibt es zunächst ein  $x \in V$  mit  $f(x) = y$ . Weil  $B$  ein Erzeugendensystem von  $V$  ist, können wir dann  $x = \lambda_1 x_1 + \dots + \lambda_n x_n$  für gewisse  $\lambda_1, \dots, \lambda_n \in K$  schreiben. Damit ist aber auch

$$y = f(x) = f(\lambda_1 x_1 + \dots + \lambda_n x_n) = \lambda_1 f(x_1) + \dots + \lambda_n f(x_n) \in \text{Lin } f(B).$$

Also ist  $f(B)$  ein Erzeugendensystem von  $W$ .

Wenden wir dies auf eine Basis  $B$  von  $V$  an, so erhalten wir aus Folgerung 14.21 (a) sofort  $\dim V = n \geq \dim W$ .

- (b) Es seien  $B$  linear unabhängig und  $\lambda_1, \dots, \lambda_n \in K$  mit  $\lambda_1 f(x_1) + \dots + \lambda_n f(x_n) = 0$ , also (weil  $f$  linear ist) mit

$$f(\lambda_1 x_1 + \dots + \lambda_n x_n) = 0.$$

Nach Voraussetzung ist  $f$  injektiv, d. h. es ist  $\text{Ker } f = \{0\}$  aufgrund von Lemma 16.9. Also folgt bereits  $\lambda_1 x_1 + \dots + \lambda_n x_n = 0$ , und damit auch  $\lambda_1 = \dots = \lambda_n = 0$ , da  $B$  linear unabhängig ist. Die Familie  $f(B)$  ist somit linear unabhängig.

Wenden wir dies wieder auf eine Basis  $B$  von  $V$  an, so erhalten wir aus Folgerung 14.21 (b) diesmal  $\dim V = n \leq \dim W$ .

- (c) Dies ist einfach (a) kombiniert mit (b). □

## 16.B Die Klassifikation endlich-dimensionaler Vektorräume

Immer wenn man eine mathematische Struktur eingeführt und dazu ein paar Beispiele untersucht hat, fragt man sich in der Regel, ob man vielleicht sogar eine *vollständige* Liste aller Beispiele angeben kann – in der linearen Algebra also, ob wir eine vollständige Liste aller (endlich-dimensionalen) Vektorräume hinschreiben können. Dabei soll „vollständig“ immer „vollständig bis auf Isomorphie“ bedeuten, da wir ja gerade in Beispiel 16.16 gesehen haben, dass isomorphe Vektorräume von ihrer Struktur her ohnehin ununterscheidbar sind, so dass es uns natürlich reichen sollte, wenn dann einer von ihnen in unserer Liste steht.

Bei vielen mathematischen Strukturen ist eine derartige Klassifikation schlichtweg aussichtslos, weil es viel zu viele Beispiele gibt, die auch keinem ersichtlichen Schema folgen. Dies ist z. B. bei Gruppen (siehe Definition 3.1) der Fall – niemand kann eine vollständige Liste aller Gruppen (bis auf Isomorphie) angeben. Falls ihr die Parallelvorlesung „Algebraische Strukturen“ hört, in der Gruppen ausführlich untersucht werden, werdet ihr dort also nahezu keine Resultate in diese Richtung finden.

Es ist eine Besonderheit der linearen Algebra, dass dies bei endlich-dimensionalen Vektorräumen anders ist: Wie wir jetzt sehen wollen, sind diese Vektorräume genau durch ihre Dimension klassifiziert, d. h. zu jedem Körper  $K$  und jeder natürlichen Zahl  $n \in \mathbb{N}$  gibt es bis auf Isomorphie *genau einen*  $K$ -Vektorraum dieser Dimension, nämlich  $K^n$ . Dies wird unsere weitere Arbeit ganz wesentlich vereinfachen: Wenn wir dann Aussagen über beliebige endlich-dimensionale Vektorräume beweisen wollen, genügt es deswegen nämlich, nur die Vektorräume  $K^n$  zu betrachten – und mit denen lässt es sich natürlich deutlich leichter arbeiten als mit abstrakten Vektorräumen.

In der Tat haben wir alle Ideen zum Beweis dieser Klassifikation schon gesehen: Zu einem  $n$ -dimensionalen Vektorraum  $V$  können wir nach Satz 14.11 eine Basis  $B$  wählen, und dann jeden Vektor in  $V$  nach Lemma 14.8 eindeutig durch seine Koordinaten bezüglich  $B$  beschreiben, die einen Vektor in  $K^n$  bilden. Diese Zuordnung liefert wie folgt den gewünschten Isomorphismus zwischen  $V$  und  $K^n$ .

**Konstruktion 16.18** (Koordinatenabbildungen). Es sei  $V$  ein endlich erzeugter Vektorraum der Dimension  $n = \dim V$  mit gegebener Basis  $B = (x_1, \dots, x_n)$ . Da jeder Vektor  $x \in V$  nach Lemma 14.8 eine eindeutige Darstellung  $x = \lambda_1 x_1 + \dots + \lambda_n x_n$  mit  $\lambda_1, \dots, \lambda_n \in K$  besitzt, gibt es dann eine bijektive Abbildung

$$\Phi_B: V \rightarrow K^n, \lambda_1 x_1 + \dots + \lambda_n x_n \mapsto \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

Man nennt  $\Phi_B$  die **Koordinatenabbildung** zur Basis  $B$ , und  $\Phi_B(x) \in K^n$  für ein  $x \in V$  den **Koordinatenvektor** von  $x$  bezüglich  $B$ . Offensichtlich ist  $\Phi_B$  linear, und damit ein Isomorphismus zwischen  $V$  und  $K^n$ .

**Beispiel 16.19.**

- (a) Für die Standardbasis  $B = (e_1, \dots, e_n)$  von  $K^n$  mit  $n \in \mathbb{N}$  ist die Koordinatenabbildung  $\Phi_B: K^n \rightarrow K^n$  natürlich einfach die identische Abbildung.
- (b) Es seien  $n \in \mathbb{N}$  und  $B = (1, x, \dots, x^n)$  die Basis von  $\text{Pol}_n(\mathbb{R}, \mathbb{R})$  aus Beispiel 14.5 (d). Die zugehörige Koordinatenabbildung

$$\Phi_B: \text{Pol}_n(\mathbb{R}, \mathbb{R}) \rightarrow \mathbb{R}^{n+1}, a_0 + a_1 x + \dots + a_n x^n \mapsto \begin{pmatrix} a_0 \\ \vdots \\ a_n \end{pmatrix}$$

ordnet dann einem Polynom genau seine Koeffizienten zu (in der Reihenfolge, die wir für die Basis  $B$  gewählt haben). Also ist  $\text{Pol}_n(\mathbb{R}, \mathbb{R}) \cong \mathbb{R}^{n+1}$ . Im Sinne von Beispiel 16.16 ist dieser Isomorphismus auch hier anschaulich klar, da der Vektor der Koeffizienten  $a_0, \dots, a_n$  dieselben Informationen enthält wie das daraus gebildete Polynom  $a_0 + a_1 x + \dots + a_n x^n$ .

**Folgerung 16.20** (Klassifikation endlich-dimensionaler Vektorräume). *Jeder endlich-dimensionale  $K$ -Vektorraum  $V$  ist (über die Koordinatenabbildung zu einer beliebigen Basis) isomorph zu  $K^n$  für ein eindeutig bestimmtes  $n \in \mathbb{N}$ , nämlich für  $n = \dim V$ .*

*Beweis.* Die Existenz eines Isomorphismus haben wir in Konstruktion 16.18 gesehen, die Eindeutigkeit von  $n$  in Lemma 16.17 (c). □

Mit diesem wichtigen Ergebnis können wir nun wie angekündigt viele Aussagen und Rechenmethoden von  $K^n$  auf beliebige endlich-dimensionale Vektorräume übertragen. Als Erstes betrachten wir dazu den offensichtlichen Isomorphismus  $K^n \times K^m \cong K^{n+m}$  für alle  $m, n \in \mathbb{N}$ , der sich ergibt,

da die Elemente beider Seiten einfach durch  $n + m$  skalare Komponenten gegeben sind. Für endlich erzeugte Vektorräume ergibt sich daraus die folgende Dimensionsaussage.

**Folgerung 16.21** (Dimensionsformel für Produkte). *Sind  $V$  und  $W$  zwei endlich-dimensionale Vektorräume, so gilt*

$$\dim(V \times W) = \dim V + \dim W.$$

*Eine analoge Aussage gilt natürlich auch für Produkte von mehr als zwei Vektorräumen.*

*Beweis.* Nach Folgerung 16.20 gibt es Isomorphismen  $f: V \rightarrow K^n$  und  $g: W \rightarrow K^m$  mit  $n = \dim V$  und  $m = \dim W$ . Dann ist aber auch

$$V \times W \rightarrow K^n \times K^m \cong K^{n+m}, (x, x') \mapsto (f(x), g(x'))$$

ein Isomorphismus (mit Umkehrabbildung  $(y, y') \mapsto (f^{-1}(y), g^{-1}(y'))$ ), und damit ergibt sich aus Folgerung 16.20

$$\dim(V \times W) = n + m = \dim V + \dim W. \quad \square$$

Auch die Algorithmen zur Bestimmung von Erzeugendensystemen und linearer Unabhängigkeit, zur Basisauswahl und -ergänzung, oder zur Berechnung von Durchschnitten und Summen von Unterräumen können wir über Koordinatenabbildungen von  $K^n$  auf beliebige endlich-dimensionale Vektorräume übertragen. Am besten sieht man dies vermutlich an einem Beispiel:

**Beispiel 16.22.** Im Vektorraum  $\text{Pol}_3(\mathbb{R}, \mathbb{R})$  aller reellen Polynome vom Grad höchstens 3 seien

$$U_1 = \text{Lin}(1 + 2x^3, x + x^2 + x^3) \quad \text{und} \quad U_2 = \text{Lin}(1 + x + 2x^3, x^2 + x^3).$$

Um mit diesen Unterräumen konkrete Berechnungen durchzuführen, können wir gemäß Beispiel 14.5 (d) die Basis  $B = (1, x, x^2, x^3)$  von  $\text{Pol}_3(\mathbb{R}, \mathbb{R})$  wählen und die Koordinatenvektoren bezüglich  $B$  benutzen. Konkret sind diese Koordinatenvektoren der oben gegebenen Polynome die Vektoren

$$\Phi_B(1 + 2x^3) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 2 \end{pmatrix}, \Phi_B(x + x^2 + x^3) = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \Phi_B(1 + x + 2x^3) = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 2 \end{pmatrix}, \Phi_B(x^2 + x^3) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \quad (*)$$

in  $\mathbb{R}^4$ . Damit können wir z. B. berechnen:

- Da die ersten beiden Vektoren in (\*) linear unabhängig in  $\mathbb{R}^4$  sind, sind die Polynome  $1 + 2x^3$  und  $x + x^2 + x^3$  nach Lemma 16.17 (b) linear unabhängig in  $U_1$ , und damit eine Basis von  $U_1$ . Es ist also  $\dim U_1 = 2$ . Analog sieht man auch  $\dim U_2 = 2$ .
- Die ersten beiden Vektoren in (\*) bilden eine  $4 \times 2$ -Matrix in Spaltenstufenform mit 2 Stufen und den Stufenzeilen 1 und 2. Nach Algorithmus 15.41 (d) können sie also mit den Einheitsvektoren  $e_3$  und  $e_4$  zu einer Basis von  $\mathbb{R}^4$  ergänzt werden. Anwenden des Isomorphismus  $\Phi_B^{-1}$  liefert daher mit Lemma 16.17 (c), dass  $\Phi_B^{-1}(e_3) = x^2$  und  $\Phi_B^{-1}(e_4) = x^3$  die Polynome  $1 + 2x^3$  und  $x + x^2 + x^3$  zu einer Basis von  $\text{Pol}_3(\mathbb{R}, \mathbb{R})$  ergänzen.
- Um eine Basis des Durchschnitts  $U_1 \cap U_2$  zu bestimmen, berechnen wir zunächst das Bild dieses Durchschnitts unter  $\Phi_B$ . Diese Rechnung haben wir schon in Algorithmus 14.27 (a) durchgeführt: Es ist

$$\Phi_B(U_1) \cap \Phi_B(U_2) = \text{Lin} \left( \begin{pmatrix} 1 \\ 0 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} \right) \cap \text{Lin} \left( \begin{pmatrix} 1 \\ 1 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right) = \text{Lin} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 3 \end{pmatrix} \subset \mathbb{R}^4.$$

Anwenden von  $\Phi_B^{-1}$  liefert also  $U_1 \cap U_2 = \text{Lin}(1 + x + x^2 + 3x^3)$ ; insbesondere hat  $U_1 \cap U_2$  damit die Dimension 1.

## 16.C Abbildungsmatrizen

Im letzten Abschnitt haben wir endlich-dimensionale Vektorräume klassifiziert: Jeder solche Vektorraum  $V$  ist (über die Koordinatenabbildung zu einer beliebigen Basis) isomorph zu  $K^n$  mit  $n = \dim V$ . Mit diesem sehr nützlichen Ergebnis der linearen Algebra kann man sowohl Beweise (wie in Folgerung 16.21) als auch Rechnungen (wie in Beispiel 16.22) in  $V$  auf solche in  $K^n$  zurückzuführen, wo sie z. B. mit Hilfe von Matrizen und dem Gauß-Algorithmus oft einfacher durchführbar sind.

Ein ähnliches Klassifikationsproblem wollen wir nun auch für Morphismen untersuchen: Ist  $W$  ein weiterer endlich-dimensionaler Vektorraum der Dimension  $m$ , können wir dann alle linearen Abbildungen von  $V$  nach  $W$  klassifizieren, also die Menge  $\text{Hom}(V, W)$  beschreiben? Wenn ihr auch die „Algebraischen Strukturen“ hört, habt ihr dort vielleicht schon das eine oder andere Mal das analoge Problem für Gruppen gesehen – also alle Morphismen zwischen zwei gegebenen Gruppen zu bestimmen – und gesehen, dass dies in der Regel gar nicht so einfach ist und sehr von den konkret betrachteten Gruppen abhängt. Aber auch hier wollen wir jetzt sehen, dass sich die lineare Algebra wieder viel schöner verhält und eine ganz explizite Beschreibung von  $\text{Hom}(V, W)$  ermöglicht.

Im Sinne unserer Klassifikationsidee sollten wir dazu natürlich  $V$  und  $W$  durch die dazu isomorphen Vektorräume  $K^n$  und  $K^m$  ersetzen können. Wir betrachten daher jetzt zunächst einmal die Menge  $\text{Hom}(K^n, K^m)$  aller linearen Abbildungen von  $K^n$  nach  $K^m$  und untersuchen anschließend, wie der allgemeine Fall darauf zurückgeführt werden kann.

Im Fall des Start- und Zielraums  $K^n$  bzw.  $K^m$  haben wir in Beispiel 16.3 (b) schon viele Morphismen gesehen, nämlich zu jeder Matrix  $A \in K^{m \times n}$  die lineare Abbildung

$$f_A: K^n \rightarrow K^m, x \mapsto Ax.$$

Wir werden nun zeigen, dass in der Tat jede lineare Abbildung von  $K^n$  nach  $K^m$  von dieser Form ist – und zwar sogar mit einer eindeutig bestimmten Matrix  $A$ . In diesem Sinne sind lineare Abbildungen in  $\text{Hom}(K^n, K^m)$  und Matrizen in  $K^{m \times n}$  also „im Prinzip dasselbe“. Wie ihr vielleicht schon vermutet, wird dies dann exakt formuliert in der Sprechweise der linearen Algebra wieder heißen, dass  $\text{Hom}(K^n, K^m)$  und  $K^{m \times n}$  isomorphe Vektorräume sind.

**Satz und Definition 16.23** (Lineare Abbildungen  $K^n \rightarrow K^m$  und Abbildungsmatrizen). *Zu jeder linearen Abbildung  $f: K^n \rightarrow K^m$  mit  $m, n \in \mathbb{N}$  gibt es genau eine Matrix  $A \in K^{m \times n}$  mit  $f = f_A$  wie in Beispiel 16.3 (b), also mit*

$$f(x) = Ax \quad \text{für alle } x \in K^n,$$

nämlich  $A = (f(e_1) | \cdots | f(e_n))$ . Wir nennen sie die **Abbildungsmatrix** von  $f$  und bezeichnen sie mit  $A_f$ .

*Beweis.* Die geforderte Bedingung legt nach Beispiel 15.6 (c) für alle  $j = 1, \dots, n$  die  $j$ -te Spalte von  $A$  fest zu  $Ae_j = f(e_j)$ . Damit ist  $A$  eindeutig bestimmt als  $(f(e_1) | \cdots | f(e_n))$ . Da  $f$  linear ist, folgt aus dieser Beziehung für die Einheitsvektoren aber auch für alle  $x \in K^n$  mit Koordinaten  $x_1, \dots, x_n$

$$f(x) = f(x_1e_1 + \cdots + x_n e_n) = x_1 f(e_1) + \cdots + x_n f(e_n) = x_1 Ae_1 + \cdots + x_n Ae_n = Ax. \quad \square$$

**Folgerung 16.24.** Für alle  $m, n \in \mathbb{N}$  gilt:

- (a)  $\text{Hom}(K^n, K^m)$  ist ein Unterraum von  $\text{Abb}(K^n, K^m)$ , und die Abbildung

$$K^{m \times n} \rightarrow \text{Hom}(K^n, K^m), A \mapsto f_A$$

ist ein Isomorphismus mit Umkehrabbildung  $\text{Hom}(K^n, K^m) \rightarrow K^{m \times n}, f \mapsto A_f$ .

Insbesondere ist also  $\dim \text{Hom}(K^n, K^m) = m \cdot n$ .

- (b) Unter dem Isomorphismus aus (a) entspricht die Matrixmultiplikation der Verkettung von Morphismen, d. h. für alle linearen Abbildungen  $f: K^n \rightarrow K^m$  und  $g: K^m \rightarrow K^p$  gilt

$$A_{g \circ f} = A_g \cdot A_f.$$

*Beweis.*

- (a) Die Abbildung  $K^{m \times n} \rightarrow \text{Abb}(K^n, K^m)$ ,  $A \mapsto f_A$  ist linear, denn für alle  $A, B \in K^{m \times n}$  sowie  $x \in K^n$  und  $\lambda \in K$  gilt

$$f_{A+B}(x) = (A+B)x = Ax + Bx = f_A(x) + f_B(x) \quad \text{und} \quad f_{\lambda A}(x) = (\lambda A)x = \lambda(Ax) = \lambda f_A(x),$$

und damit  $f_{A+B} = f_A + f_B$  und  $f_{\lambda A} = \lambda f_A$ . Nach Satz 16.23 ist diese Abbildung injektiv mit Bild  $\text{Hom}(K^n, K^m)$ . Also ist  $\text{Hom}(K^n, K^m)$  nach Lemma 16.6 (a) ein Unterraum von  $\text{Abb}(K^n, K^m)$ , und die eingeschränkte Abbildung  $K^{m \times n} \rightarrow \text{Hom}(K^n, K^m)$ ,  $A \mapsto f_A$  ist ein Isomorphismus. Nach Definition der Abbildungsmatrix ist ihre Umkehrabbildung genau  $\text{Hom}(K^n, K^m) \rightarrow K^{m \times n}$ ,  $f \mapsto A_f$ .

- (b) Für alle  $x \in K^n$  gilt

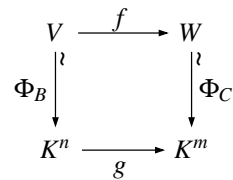
$$(g \circ f)(x) = g(f(x)) = g(A_f \cdot x) = A_g \cdot A_f \cdot x,$$

d. h.  $A_g \cdot A_f$  ist die Abbildungsmatrix von  $g \circ f$ . □

36

**Konstruktion 16.25** (Lineare Abbildungen zwischen endlich-dimensionalen Vektorräumen). Wie angekündigt wollen wir die Ergebnisse von Satz 16.23 und Folgerung 16.24 jetzt auf Morphismen zwischen beliebigen endlich-dimensionalen Vektorräumen  $V$  und  $W$  übertragen.

Die Idee dazu ist im Diagramm rechts dargestellt: Wir wählen Basen  $B$  und  $C$  von  $V$  bzw.  $W$ , die zu den vertikal eingezeichneten Koordinatenabbildungen  $\Phi_B: V \rightarrow K^n$  und  $\Phi_C: W \rightarrow K^m$  mit  $n = \dim V$  und  $m = \dim W$  führen. Die Schlangen am Beginn dieser Pfeile deuten an, dass es sich dabei um Isomorphismen handelt. Mit  $\Phi_B$  und  $\Phi_C$  wollen wir nun Abbildungen  $f: V \rightarrow W$  mit Abbildungen  $g: K^n \rightarrow K^m$  in Beziehung setzen.



Ist z. B.  $f: V \rightarrow W$  gegeben, so können wir daraus eine Abbildung  $g: K^n \rightarrow K^m$  konstruieren, indem wir im Diagramm „den Umweg über  $f$  nehmen“, also durch

$$g = \Phi_C \circ f \circ \Phi_B^{-1} \quad \text{bzw.} \quad f = \Phi_C^{-1} \circ g \circ \Phi_B.$$

Genauso können wir mit diesen Gleichungen natürlich auch umgekehrt  $f$  aus  $g$  konstruieren. Nach Lemma 16.10 (b) ist  $f$  dabei genau dann ein Morphismus, wenn  $g$  einer ist. Da ein Morphismus  $g: K^n \rightarrow K^m$  nun nach Satz 16.23 immer die Form  $x \mapsto Ax$  mit  $A \in K^{m \times n}$  hat, erhalten wir daraus den Morphismus

$$f: V \rightarrow W, \quad x \mapsto \Phi_C^{-1}(A \cdot \Phi_B(x)), \quad \text{bzw. mit} \quad \Phi_C(f(x)) = A \cdot \Phi_B(x) \quad \text{für alle } x \in V.$$

Wir bezeichnen diesen Morphismus mit  $f_A^{B,C}: V \rightarrow W$ . Beachte, dass er sich von der bisherigen Formel aus Satz 16.23 nur um die Koordinatenabbildungen unterscheidet: Während wir die Matrix  $A$  bisher direkt mit dem Startvektor multipliziert haben, um den Zielvektor zu erhalten, müssen wir sie jetzt mit dem *Koordinatenvektor* des Startvektors multiplizieren, um den *Koordinatenvektor* des Zielvektors zu erhalten. Im Fall  $V = K^n$  und  $W = K^m$ , und wenn  $B$  und  $C$  die Standardbasen sind, sind die Koordinatenabbildungen natürlich die Identität, und wir erhalten genau die bisherige Formel zurück.

Die obigen Formeln erscheinen auf den ersten Blick vielleicht etwas unübersichtlich. In der Praxis ergibt sich ihre Anwendung aber oft ganz von selbst, da wir eine Matrix ja gar nicht mit einem Vektor in einem allgemeinen Vektorraum  $V$  multiplizieren können und beim Matrixprodukt auch kein Vektor eines allgemeinen Vektorraums  $W$  herauskommt – so dass klar ist, dass wir hier mit den jeweiligen Koordinatenvektoren arbeiten müssen.

Mit dieser Konstruktion können wir nun wie gewünscht Satz 16.23 und Folgerung 16.24 auf beliebige endlich-dimensionale Vektorräume übertragen:

**Satz und Definition 16.26** (Lineare Abbildungen  $V \rightarrow W$  und Abbildungsmatrizen). *Es seien  $V$  und  $W$  zwei endlich-dimensionale Vektorräume mit  $n := \dim V$  und  $m := \dim W$ . Ferner wählen wir Basen  $B = (x_1, \dots, x_n)$  und  $C = (y_1, \dots, y_m)$  von  $V$  bzw.  $W$ .*

Dann gibt es zu jeder linearen Abbildung  $f: V \rightarrow W$  genau eine Matrix  $A \in K^{m \times n}$  mit  $f = f_A^{B,C}$  wie in Konstruktion 16.25, also mit

$$\Phi_C(f(x)) = A \cdot \Phi_B(x) \quad \text{für alle } x \in V,$$

nämlich  $A = (\Phi_C(f(x_1)) | \cdots | \Phi_C(f(x_n)))$ . Wir nennen sie die **Abbildungsmatrix** von  $f$  bezüglich  $B$  und  $C$  und bezeichnen sie mit  $A_f^{B,C}$ .

*Beweis.* Wir verwenden die Bezeichnungen aus Konstruktion 16.25. Zu  $f$  gibt es dann zunächst genau eine lineare Abbildung  $g: K^n \rightarrow K^m$  mit  $f = \Phi_C^{-1} \circ g \circ \Phi_B$ , und dazu dann genau eine Abbildungsmatrix  $A$  wie in Satz 16.23, also mit

$$f(x) = \Phi_C^{-1}(A \cdot \Phi_B(x)) \quad \Leftrightarrow \quad \Phi_C(f(x)) = A \cdot \Phi_B(x) \quad \text{für alle } x \in V.$$

Wegen  $\Phi_B(x_j) = e_j$  für  $j = 1, \dots, n$  ist dabei die  $j$ -te Spalte von  $A$  wie behauptet gegeben durch

$$Ae_j = A \cdot \Phi_B(x_j) = \Phi_C(f(x_j)). \quad \square$$

**Folgerung 16.27.** Es seien wieder  $V$  und  $W$  zwei endlich-dimensionale Vektorräume mit  $n := \dim V$  und  $m := \dim W$  sowie Basen  $B = (x_1, \dots, x_n)$  bzw.  $C = (y_1, \dots, y_m)$ . Dann gilt:

- (a)  $\text{Hom}(V, W)$  ist ein Unterraum von  $\text{Abb}(V, W)$ , und die Abbildung

$$K^{m \times n} \rightarrow \text{Hom}(V, W), \quad A \mapsto f_A^{B,C}$$

ist ein Isomorphismus mit Umkehrabbildung  $\text{Hom}(V, W) \rightarrow K^{m \times n}$ ,  $f \mapsto A_f^{B,C}$ .

Insbesondere ist also  $\dim \text{Hom}(V, W) = \dim V \cdot \dim W$ .

- (b) Unter dem Isomorphismus aus (a) entspricht die Matrixmultiplikation der Verkettung von Morphismen, d. h. ist  $f: V \rightarrow W$  linear, so gilt für alle linearen Abbildungen  $g: W \rightarrow Z$  in einen weiteren Vektorraum  $Z$  mit gegebener Basis  $D$

$$A_{g \circ f}^{B,D} = A_g^{C,D} \cdot A_f^{B,C}.$$

*Beweis.*

- (a) Wie im Beweis von Folgerung 16.24 (a) rechnet man auch hier wieder leicht nach, dass die Abbildung  $K^{m \times n} \rightarrow \text{Abb}(V, W)$ ,  $A \mapsto f_A^{B,C}$  linear ist. Da sie nach Satz 16.26 injektiv mit Bild  $\text{Hom}(V, W)$  ist, folgt die Behauptung.
- (b) Nach Definition 16.26 der Abbildungsmatrizen von  $f$  und  $g$  gilt

$$\Phi_C(f(x)) = A_f^{B,C} \cdot \Phi_B(x) \quad \text{für alle } x \in V$$

$$\text{sowie} \quad \Phi_D(g(y)) = A_g^{C,D} \cdot \Phi_C(y) \quad \text{für alle } y \in W.$$

Setzen wir dies mit  $y = f(x)$  ineinander ein, so erhalten wir für alle  $x \in V$

$$\Phi_D(g(f(x))) = A_g^{C,D} \cdot A_f^{B,C} \cdot \Phi_B(x).$$

Dies bedeutet genau, dass  $A_g^{C,D} \cdot A_f^{B,C}$  die Abbildungsmatrix von  $g \circ f$  bezüglich der Startbasis  $B$  und der Zielbasis  $D$  ist.  $\square$

**Beispiel 16.28.** Es seien  $V = \text{Pol}_2(\mathbb{R}, \mathbb{R})$  mit der Basis  $B = (1, x, x^2)$  und  $W = \text{Pol}_1(\mathbb{R}, \mathbb{R})$  mit der Basis  $C = (1, x)$  (siehe Beispiel 14.5 (d)). Wir betrachten wie in Beispiel 16.3 (e) die lineare Abbildung  $f: V \rightarrow W$ ,  $\varphi \mapsto \varphi'$ , die einem Polynom seine Ableitung zuordnet.

- (a) Um die Abbildungsmatrix  $A_f^{B,C}$  zu bestimmen, müssen wir nach Definition 16.26 die Basisvektoren von  $B$  abbilden und das Ergebnis als Linearkombinationen der Basisvektoren von  $C$  schreiben: Es ist

$$f(1) = 1' = 0 = 0 \cdot 1 + 0 \cdot x,$$

$$f(x) = x' = 1 = 1 \cdot 1 + 0 \cdot x,$$

$$f(x^2) = (x^2)' = 2x = 0 \cdot 1 + 2 \cdot x.$$

Die Koeffizienten dieser Linearkombinationen bilden wie üblich die zugehörigen Koordinatenvektoren; wir schreiben sie in die Spalten der gesuchten Abbildungsmatrix und erhalten so

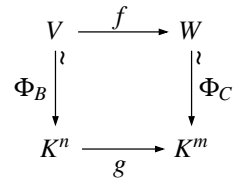
$$\Phi_C(f(1)) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Phi_C(f(x)) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \Phi_C(f(x^2)) = \begin{pmatrix} 0 \\ 2 \end{pmatrix} \Rightarrow A_f^{B,C} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

- (b) Auch die umgekehrte Richtung, aus der Abbildungsmatrix die Abbildung zu rekonstruieren, ist nicht weiter schwierig, wenn man sich daran erinnert, dass die Matrix immer nur Koordinatenvektoren sieht. Angenommen, wir wollen  $f(\varphi)$  für  $\varphi = 2x^2 + 3x + 4$ , also letztlich die Ableitung  $\varphi'$ , nur aus der Kenntnis der Abbildungsmatrix  $A_f^{B,C}$  bestimmen. Dann brauchen wir zunächst den Koordinatenvektor  $\Phi_B(\varphi)$  und können diesen dann an die Abbildungsmatrix multiplizieren: Wegen  $\varphi = 4 \cdot 1 + 3 \cdot x + 2 \cdot x^2$  ist

$$\Phi_B(\varphi) = \begin{pmatrix} 4 \\ 3 \\ 2 \end{pmatrix}, \text{ und damit } A_f^{B,C} \cdot \Phi_B(\varphi) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 4 \\ 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}.$$

Nach Satz 16.26 ist dies nun der Koordinatenvektor  $\Phi_C(f(\varphi))$  des Bildes  $f(\varphi)$ . Damit ist  $f(\varphi) = 3 \cdot 1 + 4 \cdot x = 4x + 3$  (was in der Tat die Ableitung von  $\varphi$  ist).

**Bemerkung 16.29** (Bild, Rang und Kern von Abbildungsmatrizen). In Verallgemeinerung von Bemerkung 16.8 lassen sich Bild, Rang und Kern einer linearen Abbildung zwischen endlich-dimensionalen Vektorräumen wie erwartet auf die zugehörigen Abbildungsmatrizen zurückführen und so mit den Methoden aus Abschnitt 15.C auch leicht berechnen. Dazu betrachten wir noch einmal das Diagramm rechts aus Konstruktion 16.25, in dem die Abbildung  $g$  gegeben ist durch  $x \mapsto A_f^{B,C} \cdot x$ . Damit ist



$$\text{Im } f = f(V) = \Phi_C^{-1}(g(\Phi_B(V))) = \Phi_C^{-1}(g(K^n)) = \Phi_C^{-1}(\text{Im } g) = \Phi_C^{-1}(\text{Im } A_f^{B,C})$$

$$\text{und } \text{Ker } f = f^{-1}(\{0\}) = \Phi_B^{-1}(g^{-1}(\Phi_C(\{0\}))) = \Phi_B^{-1}(g^{-1}(\{0\})) = \Phi_B^{-1}(\text{Ker } g) = \Phi_B^{-1}(\text{Ker } A_f^{B,C}).$$

Da Isomorphismen nach Lemma 16.17 (c) die Dimension erhalten, folgt damit auch  $\text{rk } f = \text{rk } A_f^{B,C}$ .

Für die konkrete Abbildung  $f$  und die Basen  $B$  und  $C$  aus Beispiel 16.28 ist z. B. mit der dort berechneten Abbildungsmatrix

$$\text{Im } A_f^{B,C} = \mathbb{R}^2 \Rightarrow \text{Im } f = \Phi_C^{-1}(\mathbb{R}^2) = \text{Pol}_1(\mathbb{R}, \mathbb{R}),$$

insbesondere ist also  $\text{rk } f = 2$  und  $f$  ist surjektiv. Für den Kern der betrachteten Ableitungsabbildung gilt

$$\text{Ker } A_f^{B,C} = \text{Lin}(e_1) \Rightarrow \text{Ker } f = \Phi_B^{-1}(\text{Lin}(e_1)) = \text{Lin}(1),$$

hier erhalten wir also erwartungsgemäß den Unterraum aller konstanten Polynome.

Mit diesen Ergebnissen können wir nun alle unsere Ergebnisse zu Matrizen auf Morphismen zwischen endlich-dimensionalen Vektorräumen übertragen. Hier sind ein paar sehr wichtige Beispiele dafür:

**Folgerung 16.30.** Für jede lineare Abbildung  $f: V \rightarrow W$  zwischen endlich-dimensionalen Vektorräumen  $V$  und  $W$  gilt:

- (a)  $\text{rk } f \leq \dim V$  und  $\text{rk } f \leq \dim W$ .
- (b) Für jede weitere lineare Abbildung  $g: W \rightarrow Z$  in einen endlich-dimensionalen Vektorraum  $Z$  gilt  $\text{rk}(g \circ f) \leq \text{rk } f$  und  $\text{rk}(g \circ f) \leq \text{rk } g$ .
- (c) (**Dimensionsformel für Morphismen**)  $\dim \text{Im } f + \dim \text{Ker } f = \dim V$ .
- (d) Ist  $V = W$ , so ist  $f$  genau dann surjektiv, wenn  $f$  injektiv ist.

*Beweis.* Wir wählen beliebige Basen der Vektorräume  $V$ ,  $W$  und  $Z$  und betrachten die Abbildungsmatrizen  $A \in K^{m \times n}$  und  $B \in K^{n \times p}$  bezüglich dieser Basen (mit  $n = \dim V$ ,  $m = \dim W$  und  $p = \dim Z$ ). Nach Bemerkung 16.29 ergeben sich die Teile (a), (b) und (c) der Folgerung dann unmittelbar aus den entsprechenden Resultaten für die Matrizen  $A$  und  $B$  aus Bemerkung 15.12 (b), Lemma 15.14 bzw. Satz 15.33 (a). So gilt z. B. für Teil (c)

$$\dim \operatorname{Im} f + \dim \operatorname{Ker} f \stackrel{16.29}{=} \dim \operatorname{Im} A + \dim \operatorname{Ker} A \stackrel{15.33(a)}{=} n = \dim V.$$

Teil (d) folgt nun unmittelbar aus dieser Dimensionsformel (c), denn die Surjektivität von  $f$  ist äquivalent zu  $\dim \operatorname{Im} f = \dim W$  (also  $\dim \operatorname{Im} f = \dim V$ ), und die Injektivität von  $f$  nach Lemma 16.9 zu  $\dim \operatorname{Ker} f = 0$ .  $\square$

**Bemerkung 16.31.** Folgerung 16.30 (d) sieht zwar vielleicht so aus, als ob die Dimension von  $V$  hierbei keine Rolle spielt, ist aber für einen unendlich-dimensionalen Raum  $V$  falsch: Betrachten wir z. B. mit dem Vektorraum  $V = \operatorname{Abb}(\mathbb{N}, \mathbb{R})$  aller reellen Zahlenfolgen die „Verschiebomorphismen“

$$\begin{aligned} f: V &\rightarrow V, (x_0, x_1, x_2, \dots) \mapsto (0, x_0, x_1, x_2, \dots) \\ \text{und } g: V &\rightarrow V, (x_0, x_1, x_2, \dots) \mapsto (x_1, x_2, x_3, \dots), \end{aligned}$$

so ist  $f$  injektiv aber nicht surjektiv, und  $g$  surjektiv aber nicht injektiv.

Als weitere einfache Konsequenz aus Folgerung 16.27 erhalten wir außerdem, dass sich lineare Abbildungen stets auf einer Basis des Startraums beliebig vorgeben lassen und dann eindeutig bestimmt sind (siehe auch Aufgabe 16.5):

**Folgerung 16.32.** Es seien  $V$  und  $W$  endlich-dimensionale Vektorräume,  $B = (x_1, \dots, x_n)$  eine Basis von  $V$  und  $(y_1, \dots, y_n)$  eine beliebige Familie (mit gleich vielen Elementen) in  $W$ .

Dann gibt es genau eine lineare Abbildung  $f: V \rightarrow W$  mit  $f(x_i) = y_i$  für alle  $i = 1, \dots, n$ .

*Beweis.* Wähle eine Basis  $C$  von  $W$  und betrachte die Abbildungsmatrix  $A_f^{B,C}$  der gesuchten Morphismen, die nach Satz 16.26 die Form  $A_f^{B,C} = (\Phi_C(f(x_1)) \mid \dots \mid \Phi_C(f(x_n)))$  hat. Die Bedingungen  $f(x_i) = y_i$  für alle  $i$  sind offensichtlich äquivalent dazu, dass diese Abbildungsmatrix gleich  $A_f^{B,C} = (\Phi_C(y_1) \mid \dots \mid \Phi_C(y_n))$  ist, und liefern damit genau einen solchen Morphismus.  $\square$

**Aufgabe 16.33.** Es sei  $V = \operatorname{Pol}_2(\mathbb{R}, \mathbb{R})$  mit Basis  $B = (1, x, x^2)$ . Wir betrachten die Abbildung  $f: V \rightarrow V$  mit  $f(\varphi)(x) = \varphi'(x+1) + x\varphi(1)$ , wobei  $\varphi'$  die Ableitung von  $\varphi$  bezeichnet.

- Zeige, dass  $f$  eine lineare Abbildung ist.
- Berechne die Abbildungsmatrix  $A_f^{B,B}$ .
- Berechne  $\operatorname{Ker} f$ .

**Aufgabe 16.34.** Es seien

$$U = \{x \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0\} \quad \text{und} \quad f: \mathbb{R}^3 \rightarrow \mathbb{R}^3, x \mapsto Ax \quad \text{mit} \quad A = \begin{pmatrix} 0 & 2 & 1 \\ 2 & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix}.$$

Zeige, dass  $f(U) \subset U$  gilt, bestimme eine Basis  $B$  von  $U$ , und berechne die Abbildungsmatrix der eingeschränkten Abbildung  $f|_U: U \rightarrow U$  bezüglich der Basis  $B$  im Start- und Zielraum.

**Aufgabe 16.35.** Bestimme die Dimension von  $\operatorname{Ker}(g \circ f)$ , wenn  $f: K^8 \rightarrow K^5$  eine surjektive und  $g: K^5 \rightarrow K^7$  eine injektive lineare Abbildung ist.

**Aufgabe 16.36.** Es seien  $V$  ein endlich-dimensionaler Vektorraum und  $f: V \rightarrow V$  eine lineare Abbildung mit  $f \circ f = f$ . Man zeige:

- $\operatorname{Im} f \cap \operatorname{Ker} f = \{0\}$ .

- (b) Es gibt eine Basis  $B$  von  $V$ , so dass die Abbildungsmatrix von  $f$  mit gleicher Start- und Zielbasis  $B$  die Form

$$A_f^{B,B} = \left( \begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array} \right)$$

mit  $r = \text{rk } f$  hat.

**Aufgabe 16.37.** Es sei  $A \in K^{m \times n}$  eine Matrix vom Rang  $r := \text{rk } A$ . Zeige, dass  $A$  dann als Matrixprodukt  $A = BC$  mit  $B \in K^{m \times r}$  und  $C \in K^{r \times n}$  geschrieben werden kann.

(Hinweis: Betrachte zugehörige lineare Abbildungen und verwende Folgerung 16.27 (b).)

Nach Konstruktion hängen unsere gerade eingeführten Abbildungsmatrizen  $A_f^{B,C}$  natürlich von der (letztlich willkürlichen) Wahl der Basen  $B$  und  $C$  im Start- bzw. Zielraum der Abbildung  $f$  ab. Wir wollen daher nun untersuchen, wie sich diese Abbildungsmatrizen ändern, wenn man zu anderen Basen übergeht. Dazu benötigen wir die sogenannten Basiswechselmatrizen, die einfach nur ein Spezialfall von Abbildungsmatrizen sind.

**Definition 16.38** (Basiswechselmatrizen). Es seien  $B = (x_1, \dots, x_n)$  und  $C = (y_1, \dots, y_n)$  zwei Basen eines endlich-dimensionalen  $K$ -Vektorraums  $V$ . Dann heißt die Abbildungsmatrix der Identität  $\text{id}_V$  bezüglich der Startbasis  $B$  und Zielbasis  $C$ , nach Definition 16.26 also

$$A^{B,C} := A_{\text{id}}^{B,C} = (\Phi_C(x_1) \mid \dots \mid \Phi_C(x_n)) \in K^{n \times n},$$

die **Basiswechselmatrix** von  $B$  nach  $C$ .

**Beispiel 16.39.** Im Vektorraum  $\text{Pol}_1(\mathbb{R}, \mathbb{R})$  betrachten wir die beiden Basen

$$B = (1, x) \quad \text{und} \quad C = (x+2, -1).$$

Analog zu Beispiel 16.28 (a) schreiben wir die beiden Basisvektoren von  $B$  als Linearkombinationen der Basisvektoren von  $C$ :

$$1 = 0 \cdot (x+2) + (-1) \cdot (-1) \quad \text{und} \quad x = 1 \cdot (x+2) + 2 \cdot (-1).$$

Damit ist

$$\Phi_C(1) = \begin{pmatrix} 0 \\ -1 \end{pmatrix} \quad \text{und} \quad \Phi_C(x) = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad \Rightarrow \quad A^{B,C} = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}.$$

**Bemerkung 16.40.** Es sei  $V$  ein endlich-dimensionaler Vektorraum mit Basis  $B$ .

- (a) Offensichtlich ist stets  $A^{B,B} = E$ .  
 (b) Ist  $f: V \rightarrow W$  ein Isomorphismus zwischen  $V$  und einem weiteren Vektorraum  $W$  mit Basis  $C$ , so gilt nach Folgerung 16.27 (b)

$$A_{f^{-1}}^{C,B} \cdot A_f^{B,C} = A_{f^{-1} \circ f}^{B,B} = A^{B,B} \stackrel{(a)}{=} E.$$

Die Abbildungsmatrix  $A_f^{B,C}$  zu einem Isomorphismus  $f$  ist also invertierbar mit inverser Matrix  $A_{f^{-1}}^{C,B}$ . Insbesondere erhalten wir daraus für  $V = W$  und  $f = \text{id}_V$ , dass Basiswechselmatrizen immer invertierbar sind mit  $(A^{B,C})^{-1} = A^{C,B}$ .

Umgekehrt sind invertierbare Matrizen in folgendem Sinne auch immer Basiswechselmatrizen:

**Lemma 16.41.** Es sei  $B$  eine Basis eines endlich erzeugten Vektorraums  $V$ . Ferner sei  $T \in \text{GL}(n, K)$  eine invertierbare Matrix mit  $n = \dim V$ . Dann gilt:

- (a) Es gibt eine Basis  $C$  von  $V$  mit  $A^{C,B} = T$ .  
 (b) Es gibt eine Basis  $C$  von  $V$  mit  $A^{B,C} = T$ .

*Beweis.*

- (a) Wir setzen  $y_i = \Phi_B^{-1}(Te_i)$  für  $i = 1, \dots, n$ . Da  $T$  invertierbar und  $\Phi_B^{-1}$  ein Isomorphismus ist, ist auch  $K^n \rightarrow V$ ,  $x \mapsto \Phi_B^{-1}(Tx)$  ein Isomorphismus, und bildet nach Lemma 16.17 (c) damit die Standardbasis  $(e_1, \dots, e_n)$  auf eine Basis  $C := (y_1, \dots, y_n)$  von  $V$  ab.

Für  $i = 1, \dots, n$  ist die  $i$ -te Spalte der Basiswechsellmatrix  $A^{C,B}$  nun nach Definition 16.38

$$\Phi_B(y_i) = \Phi_B(\Phi_B^{-1}(Te_i)) = Te_i,$$

also die  $i$ -te Spalte von  $T$ . Damit ist wie gewünscht  $A^{C,B} = T$ .

- (b) Nach (a) gibt es eine Basis  $C$  von  $V$  mit  $A^{C,B} = T^{-1}$ , nach Bemerkung 16.40 (b) also mit  $A^{B,C} = T$ .  $\square$

37

Mit diesen Basiswechsellmatrizen können wir nun konkret angeben, wie sich Abbildungsmatrizen unter einem Basiswechsel transformieren.

**Satz 16.42** (Verhalten von Abbildungsmatrizen unter Basiswechsel). *Es sei  $f: V \rightarrow W$  ein Morphismus zwischen endlich erzeugten Vektorräumen mit  $n := \dim V$  und  $m := \dim W$  sowie gegebenen Basen  $B$  und  $C$  von  $V$  bzw.  $W$ .*

- (a) *Sind  $B'$  und  $C'$  zwei weitere Basen von  $V$  bzw.  $W$ , so gilt*

$$A_f^{B',C'} = A^{C,C'} \cdot A_f^{B,C} \cdot A^{B',B}.$$

- (b) *Sind umgekehrt  $S \in \text{GL}(m, K)$  und  $T \in \text{GL}(n, K)$  zwei invertierbare Matrizen, so gibt es Basen  $B'$  und  $C'$  von  $V$  bzw.  $W$ , so dass  $A^{C,C'} = S$  und  $A^{B',B} = T$  gilt, und damit*

$$A_f^{B',C'} = S \cdot A_f^{B,C} \cdot T.$$

*Beweis.*

- (a) Dies folgt sofort durch doppelte Anwendung von Folgerung 16.27 (b) auf den Morphismus

$$f = \text{id}_W \circ f \circ \text{id}_V: V \rightarrow V \rightarrow W \rightarrow W$$

mit den Basen  $B', B, C, C'$  in den vier Räumen dieser Abbildungskette.

- (b) Nach Lemma 16.41 existieren Basen  $B'$  und  $C'$  mit  $A^{C,C'} = S$  und  $A^{B',B} = T$ ; die behauptete Formel für die Abbildungsmatrix ergibt sich dann aus (a).  $\square$

**Beispiel 16.43.** Es sei  $f: \text{Pol}_2(\mathbb{R}, \mathbb{R}) \rightarrow \text{Pol}_1(\mathbb{R}, \mathbb{R})$ ,  $\varphi \mapsto \varphi'$  die Ableitungsabbildung wie in Beispiel 16.28. Dort hatten wir die Abbildungsmatrix  $A_f^{B,C}$  für die Basen  $B = (1, x, x^2)$  und  $C = (1, x)$  im Start- bzw. Zielraum berechnet. Wollen wir nun stattdessen die Abbildungsmatrix  $A_f^{B',C'}$  für die Basis  $C' = (x+2, -1)$  von  $\text{Pol}_1(\mathbb{R}, \mathbb{R})$  bestimmen, so können wir entweder wieder das gleiche Verfahren wie in Beispiel 16.28 anwenden, oder die Formel aus Satz 16.42 (a) benutzen: Mit der in Beispiel 16.39 berechneten Basiswechsellmatrix  $A^{C,C'}$  ist

$$A_f^{B',C'} = A^{C,C'} \cdot A_f^{B,C} = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 2 \\ 0 & -1 & 4 \end{pmatrix}.$$

Nach Satz 16.42 beschreiben zwei Matrizen derselben Größe also genau dann die gleiche lineare Abbildung – nur bezüglich evtl. verschiedener Basen – wenn sie durch Multiplikation mit invertierbaren Matrizen von links und rechts auseinander hervorgehen. Es ist daher nützlich, die folgende Notation einzuführen.

**Definition 16.44** (Äquivalente Matrizen). Zwei Matrizen  $A, A' \in K^{m \times n}$  mit  $m, n \in \mathbb{N}$  heißen **äquivalent** zueinander, wenn es invertierbare Matrizen  $S \in \text{GL}(m, K)$  und  $T \in \text{GL}(n, K)$  gibt mit  $A' = SAT$ .

**Bemerkung 16.45.**

- (a) Wie man leicht nachprüfen kann, ist die Äquivalenz von Matrizen in der Tat eine Äquivalenzrelation im Sinne von Definition 2.31.

- (b) Nach Satz 16.42 sind zwei Matrizen genau dann äquivalent zueinander, wenn sie die gleiche lineare Abbildung, nur bezüglich evtl. anderer Basen im Start- und Zielraum beschreiben.

Wir werden in Bemerkung 17.30 noch sehen, dass diese Bedingung sehr leicht überprüft werden kann: Sie ist einfach äquivalent dazu, dass die beiden Matrizen denselben Rang haben. Eine Richtung dieser Aussage können wir aber jetzt schon zeigen.

**Folgerung 16.46** (Äquivalente Matrizen haben denselben Rang). *Es seien  $m, n \in \mathbb{N}$  und  $A \in K^{m \times n}$ . Für alle  $S \in GL(m, K)$  und  $T \in GL(n, K)$  gilt dann  $\text{rk}(SAT) = \text{rk}A$ .*

*Beweis.* Es sei  $f = f_A: K^n \rightarrow K^m$ . Für die Standardbasen  $B$  und  $C$  von  $K^n$  bzw.  $K^m$  ist dann also  $A_f^{B,C} = A$ . Nach Satz 16.42 (b) gibt es nun Basen  $B'$  und  $C'$  von  $K^n$  bzw.  $K^m$  mit  $A_f^{B',C'} = SAT$ . Mit Bemerkung 16.29 erhalten wir also wie behauptet  $\text{rk}(SAT) = \text{rk}A_f^{B',C'} = \text{rk}f = \text{rk}A$ .  $\square$

**Aufgabe 16.47.** Wir betrachten die Basen

$$B = \left( \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right) \quad \text{und} \quad C = \left( \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right)$$

von  $\mathbb{R}^3$  bzw.  $\mathbb{R}^2$  sowie die lineare Abbildung  $f: \mathbb{R}^3 \rightarrow \mathbb{R}^2$  mit

$$A_f^{B,C} = \begin{pmatrix} 2 & -1 & 0 \\ 4 & 0 & 2 \end{pmatrix}.$$

Berechne die Abbildungsmatrix  $A_f$  von  $f$  bezüglich der Standardbasen von  $\mathbb{R}^3$  und  $\mathbb{R}^2$ .

**Aufgabe 16.48.** Zum Vektorraum  $V = \text{Pol}_2(\mathbb{R}, \mathbb{R})$  mit der Basis  $B = (1, x, x^2)$  betrachten wir die lineare Abbildung  $f: V \rightarrow V$  mit  $f(\varphi)(x) = \varphi(x+1)$  für alle  $\varphi \in V$ .

- Berechne die Abbildungsmatrix  $A_f^{B,B}$ .
- Zeige, dass  $A_f^{B,B}$  invertierbar ist, und berechne die inverse Matrix  $(A_f^{B,B})^{-1}$ .
- Zeige, dass es zu jeder Basis  $C'$  von  $V$  eine Basis  $C$  von  $V$  gibt mit  $A_f^{C,C'} = E$ .
- Zeige, dass es zu keiner Basis  $C'$  von  $V$  eine Basis  $C$  von  $V$  gibt mit  $A_f^{C,C'} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ .

**Aufgabe 16.49** (Berechnung von Basiswechselmatrizen). Für ein  $n \in \mathbb{N}$  seien  $B = (x_1, \dots, x_n)$  und  $C = (y_1, \dots, y_n)$  zwei Basen von  $K^n$ . Wir setzen  $A := (y_1 | \dots | y_n | x_1 | \dots | x_n) \in K^{2 \times 2n}$  und bringen die linke Hälfte dieser Matrix mit elementaren Zeilenumformungen in reduzierte Zeilenstufenform, führen dabei aber alle Umformungen mit der gesamten Matrix durch.

Zeige, dass dann in der rechten Hälfte der Matrix genau die Basiswechselmatrix  $A^{B,C}$  steht.

## 17. Komplemente und Quotientenräume

In diesem Kapitel wollen wir uns mit der folgenden Frage beschäftigen, die in der linearen Algebra oft auftritt: Gegeben sei ein Unterraum  $U$  eines Vektorraums  $V$ . Können wir dann jeden Vektor  $x \in V$  in einem gewissen Sinne eindeutig in einen „Anteil in  $U$ “ und einen „Restteil“ zerlegen?

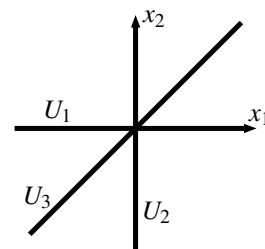
Diese Frage ist so sicher erst einmal nicht mathematisch exakt formuliert, und in der Tat werden wir in den beiden Abschnitten dieses Kapitels zwei ganz verschiedene Arten sehen, wie man sie interpretieren kann. Das folgende Beispiel zeigt aber schon einmal in einem sehr einfachen Fall, was damit gemeint sein kann: Sind  $V = \mathbb{R}^2$  und  $U = \text{Lin}(e_1)$  die horizontale Koordinatenachse, so können wir natürlich jeden Vektor  $x \in V$  eindeutig als  $x = x_1 e_1 + x_2 e_2$  mit seinen Koordinaten  $x_1$  und  $x_2$  bezüglich der Standardbasis schreiben. Wir haben  $x$  damit also geschrieben als Summe von einem Anteil  $x_1 e_1$  in  $U$  und einem Rest  $x_2 e_2$ , der in  $\text{Lin}(e_2)$  liegt. Diese recht naheliegende Idee, einen Vektor eindeutig in eine Summe zu zerlegen, bei der jeder Summand in einem gegebenen Unterraum liegt, wollen wir jetzt im ersten Abschnitt dieses Kapitels untersuchen.

### 17.A Direkte Summen und Komplemente

Wir betrachten noch einmal die Konstruktion der Summe  $U_1 + \dots + U_n$  von Unterräumen  $U_1, \dots, U_n$  eines Vektorraums  $V$  wie in Lemma 13.13:

**Bemerkung 17.1** (Eindeutigkeit der Summendarstellung). Jeder Vektor in einer Summe  $U_1 + \dots + U_n$  von Unterräumen eines Vektorraums  $V$  lässt sich nach Definition als  $x_1 + \dots + x_n$  mit  $x_i \in U_i$  für alle  $i = 1, \dots, n$  schreiben. Allerdings ist diese Darstellung im Allgemeinen natürlich nicht eindeutig: Betrachten wir z. B. wie im Bild rechts die drei Ursprungsgeraden

$$U_1 = \text{Lin} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad U_2 = \text{Lin} \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad U_3 = \text{Lin} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$



in  $\mathbb{R}^2$ , so hat der Vektor

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 \\ 0 \end{pmatrix}}_{\in U_1} + \underbrace{\begin{pmatrix} 0 \\ 1 \end{pmatrix}}_{\in U_2} + \underbrace{\begin{pmatrix} 0 \\ 0 \end{pmatrix}}_{\in U_3} = \underbrace{\begin{pmatrix} 0 \\ 0 \end{pmatrix}}_{\in U_1} + \underbrace{\begin{pmatrix} 0 \\ 0 \end{pmatrix}}_{\in U_2} + \underbrace{\begin{pmatrix} 1 \\ 1 \end{pmatrix}}_{\in U_3} \in U_1 + U_2 + U_3 = \mathbb{R}^2$$

zwei verschiedene Darstellungen dieser Art. Ist die Darstellung jedoch immer eindeutig, so geben wir dieser Situation einen besonderen Namen:

**Definition 17.2** (Direkte Summe von Unterräumen). Es seien  $U_1, \dots, U_n$  Untervektorräume eines  $K$ -Vektorraums  $V$  und  $U = U_1 + \dots + U_n$ . Hat jedes  $x \in U$  eine *eindeutige* Darstellung der Form  $x = x_1 + \dots + x_n$  mit  $x_i \in U_i$  für alle  $i = 1, \dots, n$ , so nennt man die Summe **direkt**. Möchte man dies auch in der Notation andeuten, so schreibt man dafür  $U = U_1 \oplus \dots \oplus U_n$ .

Die Summe in Bemerkung 17.1 ist also nicht direkt – was natürlich einfach daran liegt, dass die drei aufspannenden Vektoren von  $U_1$ ,  $U_2$  und  $U_3$  linear abhängig sind. In der Tat kann man sich direkte Summen als eine Verallgemeinerung des Konzepts der linearen Unabhängigkeit auf Unterräume vorstellen.

**Lemma 17.3** (Alternatives Kriterium für direkte Summen). Die Summe  $U_1 + \dots + U_n$  von Unterräumen  $U_1, \dots, U_n$  eines  $K$ -Vektorraums  $V$  ist genau dann direkt, wenn die Abbildung

$$f: U_1 \times \dots \times U_n \rightarrow U_1 + \dots + U_n, (x_1, \dots, x_n) \mapsto x_1 + \dots + x_n$$

ein Isomorphismus ist. Ist  $V$  endlich-dimensional, so gilt in diesem Fall also die **Dimensionsformel für direkte Summen**

$$\dim U_1 + \cdots + \dim U_n = \dim(U_1 \oplus \cdots \oplus U_n).$$

*Beweis.* Es ist klar, dass die Abbildung  $f$  in jedem Fall linear ist; außerdem ist  $f$  nach Definition der Summe  $U_1 + \cdots + U_n$  stets surjektiv. Injektiv ist  $f$  genau dann, wenn für alle  $x_i, y_i \in U_i$  aus  $x_1 + \cdots + x_n = y_1 + \cdots + y_n$  bereits  $(x_1, \dots, x_n) = (y_1, \dots, y_n)$ , also  $x_i = y_i$  für alle  $i$  folgt. Dies bedeutet nach Definition 17.2 aber genau, dass die Summe direkt ist.

Ist  $V$  darüber hinaus endlich-dimensional, so gilt dies nach Lemma 14.23 (a) auch für die Unterräume  $U_1, \dots, U_n$ . Da endlich erzeugte isomorphe Vektorräume nach Lemma 16.17 (c) die gleiche Dimension haben, ist also

$$\dim U_1 + \cdots + \dim U_n \stackrel{16.21}{=} \dim(U_1 \times \cdots \times U_n) = \dim(U_1 \oplus \cdots \oplus U_n). \quad \square$$

Im Fall von nur zwei Unterräumen kann man besonders einfach feststellen, ob ihre Summe direkt ist:

**Lemma 17.4.** Die Summe  $U_1 + U_2$  von zwei Unterräumen  $U_1$  und  $U_2$  eines  $K$ -Vektorraums  $V$  ist genau dann direkt, wenn  $U_1 \cap U_2 = \{0\}$ .

*Beweis.* Nach Lemma 17.3 ist die Summe  $U_1 + U_2$  genau dann direkt, wenn die Abbildung

$$f: U_1 \times U_2 \rightarrow U_1 + U_2, (x_1, x_2) \mapsto x_1 + x_2$$

ein Isomorphismus ist. Wir hatten im Beweis dieses Lemmas aber auch schon gesehen, dass  $f$  stets linear und surjektiv ist. Also ist die Summe  $U_1 + U_2$  genau dann direkt, wenn  $f$  injektiv ist, d. h. nach Lemma 16.9 genau dann, wenn  $\text{Ker } f = \{(0, 0)\}$  gilt. Nun ist aber

$$\begin{aligned} \text{Ker } f &= \{(x_1, x_2) : x_1 \in U_1, x_2 \in U_2, x_1 + x_2 = 0\} \\ &= \{(x_1, -x_1) : x_1 \in U_1, -x_1 \in U_2\} \\ &= \{(x_1, -x_1) : x_1 \in U_1 \cap U_2\}, \end{aligned}$$

und damit ist wie behauptet genau dann  $\text{Ker } f = \{(0, 0)\}$ , wenn  $U_1 \cap U_2 = \{0\}$ . □

**Beispiel 17.5.**

- Die Summe  $U_1 + U_2$  der  $x_1$ -Achse und der  $x_2$ -Achse in  $\mathbb{R}^3$  in Beispiel 13.14 ist direkt, denn in diesem Fall ist natürlich  $U_1 \cap U_2 = \{0\}$ . In der Tat sieht man in diesem Beispiel auch sofort, dass sich jeder Vektor in der  $x_1$ - $x_2$ -Ebene  $U_1 + U_2$  wie in der Einleitung zu diesem Kapitel eindeutig als Summe von einem Vektor in  $U_1$  und einem in  $U_2$  schreiben lässt.
- Die Summe  $U_1 + U_2 + U_3$  in Bemerkung 17.1 ist hingegen nicht direkt, wie wir dort bereits gesehen hatten. Allerdings ist in diesem Fall trotzdem  $U_1 \cap U_2 \cap U_3 = \{0\}$  – was zeigt, dass sich die Aussage von Lemma 17.4 nicht genauso auf mehr als zwei Summanden übertragen lässt. Zu Lemma 17.4 analoge Aussagen für allgemeine Summen sind stattdessen die folgenden.

**Aufgabe 17.6.** Es seien  $n \in \mathbb{N}$  und  $U_1, \dots, U_n$  Unterräume eines endlich-dimensionalen Vektorraums. Zeige, dass die folgenden Aussagen äquivalent sind:

- Die Summe  $U_1 + \cdots + U_n$  ist direkt.
- Sind  $x_i \in U_i$  für  $i = 1, \dots, n$  so dass  $x_1 + \cdots + x_n = 0$  ist, so gilt bereits  $x_1 = \cdots = x_n = 0$ .
- $U_i \cap (U_1 + \cdots + U_{i-1} + U_{i+1} + \cdots + U_n) = \{0\}$  für alle  $i = 1, \dots, n$ .
- $U_i \cap (U_{i+1} + \cdots + U_n) = \{0\}$  für alle  $i = 1, \dots, n-1$ .

**Aufgabe 17.7.** Es seien  $U_1, \dots, U_n$  Unterräume eines endlich erzeugten Vektorraums  $V$ . Zeige in Ergänzung zu Lemma 17.3, dass die Summe  $U_1 + \cdots + U_n$  genau dann direkt ist, wenn

$$\dim(U_1 + \cdots + U_n) = \dim U_1 + \cdots + \dim U_n,$$

und dass man in diesem Fall eine Basis von  $U_1 + \cdots + U_n$  erhält, indem man Basen von  $U_1, \dots, U_n$  vereinigt.

Sind nun  $U$  ein Unterraum eines Vektorraums  $V$  und  $U'$  ein weiterer Unterraum mit  $U \oplus U' = V$ , so haben wir genau die am Anfang dieses Kapitels beschriebene Situation, dass sich jeder Vektor  $x \in V$  eindeutig als  $x = y + z$  schreiben lässt, wobei  $y$  in  $U$  liegt und  $z$  als „Restteil“ des Vektors in  $U'$  aufgefasst werden kann. Diese Situation hat einen besonderen Namen:

**Definition 17.8** (Komplemente). Es sei  $U$  ein Unterraum eines  $K$ -Vektorraums  $V$ . Ein Unterraum  $U' \leq V$  heißt **Komplement** oder **komplementärer Unterraum** von  $U$  in  $V$ , wenn  $U \oplus U' = V$  (nach Lemma 17.4 also  $U + U' = V$  und  $U \cap U' = \{0\}$ ) gilt.

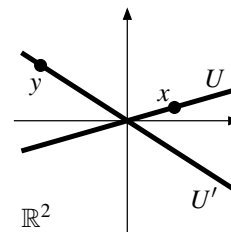
**Bemerkung 17.9** (Dimensionsformel für Komplemente). Nach Lemma 17.3 gilt für jedes Komplement  $U'$  eines Untervektorraums  $U$  in einem endlich-dimensionalen Vektorraum  $V$  die Dimensionsformel  $\dim U + \dim U' = \dim V$ , also

$$\dim U' = \dim V - \dim U.$$

**Beispiel 17.10** (Nichteindeutigkeit von Komplementen). Wie im Bild unten rechts seien  $U = \text{Lin}(x)$  und  $U' = \text{Lin}(y)$  zwei verschiedene Ursprungsgeraden in  $\mathbb{R}^2$ . Da  $x$  und  $y$  dann keine Vielfachen voneinander sind, sind diese beiden Vektoren also linear unabhängig und bilden damit eine Basis des zweidimensionalen Vektorraums  $\mathbb{R}^2$ .

Es ist also  $U + U' = \text{Lin}(x, y) = \mathbb{R}^2$ , außerdem gilt natürlich  $U \cap U' = \{0\}$ . Also ist  $U'$  ein Komplement von  $U$ .

Da es zu einer gegebenen Ursprungsgeraden  $U$  in  $\mathbb{R}^2$  aber natürlich unendlich viele Geraden  $U' \neq U$  gibt, folgt daraus insbesondere, dass Komplemente von Unterräumen in der Regel nicht eindeutig sind. Wir wollen nun aber sehen, dass Komplemente zumindest im endlich-dimensionalen Fall stets existieren:



**Satz 17.11** (Existenz von Komplementen). *Jeder Unterraum  $U$  eines endlich-dimensionalen Vektorraums  $V$  besitzt ein Komplement.*

*Beweis.* Nach Lemma 14.23 (a) ist  $U$  endlich erzeugt, besitzt damit also nach Satz 14.11 eine Basis  $(x_1, \dots, x_n)$ . Wir ergänzen sie gemäß Folgerung 14.16 zu einer Basis  $(x_1, \dots, x_n, y_1, \dots, y_m)$  von  $V$  und zeigen, dass  $U' := \text{Lin}(y_1, \dots, y_m)$  dann ein Komplement von  $U$  ist:

- Es ist offensichtlich  $U + U' = V$ , denn nach Beispiel 13.14 ist

$$U + U' = \text{Lin}(x_1, \dots, x_n) + \text{Lin}(y_1, \dots, y_m) = \text{Lin}(x_1, \dots, x_n, y_1, \dots, y_m) = V.$$

- Nach der Dimensionsformel aus Satz 14.25 ist damit

$$\dim(U \cap U') = \dim U + \dim U' - \underbrace{\dim(U + U')}_{=V} = n + m - (n + m) = 0,$$

also auch  $U \cap U' = \{0\}$ . □

**Bemerkung 17.12** (Berechnung von Komplementen). Beachte, dass der Beweis von Satz 17.11 konstruktiv ist, d. h. auch die konkrete Berechnung eines Komplements ermöglicht: Möchte man ein Komplement  $U'$  zu einem Unterraum  $U$  eines endlich-dimensionalen Vektorraums  $V$  berechnen, muss man nur eine Basis von  $U$  zu einer Basis von  $V$  ergänzen; die dafür hinzugenommenen Vektoren bilden dann eine Basis eines Komplements  $U'$ .

So haben wir z. B. in Beispiel 16.22 (b) im Raum  $\text{Pol}_3(\mathbb{R}, \mathbb{R})$  aller Polynome vom Grad höchstens 3 die Basis  $(1 + 2x^3, x + x^2 + x^3)$  von  $U := \text{Lin}(1 + 2x^3, x + x^2 + x^3)$  mit den Polynomen  $x^2$  und  $x^3$  zu einer Basis von  $\text{Pol}_3(\mathbb{R}, \mathbb{R})$  ergänzt; dementsprechend ist  $U' := \text{Lin}(x^2, x^3)$  also ein Komplement von  $U$  in  $V$ .

**Aufgabe 17.13.** Es seien  $U_1, U_2, U_3$  und  $U$  Unterräume eines Vektorraums  $V$  mit  $U_1 \oplus U_2 = U$  und  $U \oplus U_3 = V$ .

Zeige, dass dann  $U_1 \oplus U_2 \oplus U_3 = V$  gilt.

### 17.B Quotientenräume

Komplemente von Unterräumen sind in der Praxis sehr nützlich: Wir haben gerade gesehen, dass ein Komplement  $U'$  eines Unterraums  $U$  in einem Vektorraum  $V$  die „Restteile“ von Vektoren in  $V$  misst, wenn man ihren Anteil in  $U$  heraus nimmt. Unschön ist an dieser Konstruktion allerdings, dass ein Komplement nach Beispiel 17.10 nicht eindeutig bestimmt und damit ein recht unnatürliches Objekt ist. Was im obigen Sinne der Restteil eines Vektors in  $V$  nach Herausnehmen des Anteils in  $U$  ist, lässt sich also nicht beantworten, solange man nicht eine (letztlich willkürliche) Wahl eines Komplements von  $U$  in  $V$  getroffen hat.

Wir wollen nun eine deutlich schönere Konstruktion einführen, die solche Restteile auch ohne willkürliche Wahlen messen kann. Der Preis dafür ist, dass der Vektorraum, der diese Restteile auf ganz natürliche Art beschreibt, kein *Unterraum* von  $V$  mehr ist, sondern ein sogenannter *Quotientenraum*: ein Raum von Äquivalenzklassen von Vektoren in  $V$  wie in Abschnitt 2.B, wobei wir zwei Vektoren in  $V$  miteinander identifizieren wollen, wenn sie sich um ein Element von  $U$  voneinander unterscheiden. Diejenigen von euch, die auch die Vorlesung „Algebraische Strukturen“ besuchen, kennen diese Idee vermutlich bereits von den Faktorgruppen [G, Kapitel 6].

**Lemma und Definition 17.14.** *Es seien  $V$  ein  $K$ -Vektorraum und  $U \leq V$  ein fest gewählter Unterraum. Dann ist durch*

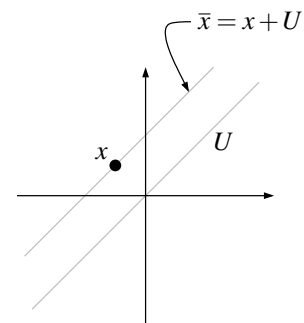
$$x \sim y \iff x - y \in U \quad \text{für alle } x, y \in V$$

*eine Äquivalenzrelation auf  $V$  definiert. Für die Äquivalenzklasse eines Vektors  $x \in V$  bezüglich dieser Relation gilt*

$$\bar{x} = x + U := \{x + u : u \in U\},$$

*d. h.  $\bar{x}$  ist (wie in Beispiel 13.12 (c) und im Bild rechts) ein verschobener Unterraum mit Aufpunkt  $x$ . Man nennt dies auch einen affinen Unterraum mit Aufpunkt  $x$ .*

*Die Menge  $V / \sim$  aller Äquivalenzklassen bezüglich dieser Relation bezeichnet man mit  $V/U$ .*



*Beweis.* Wir zeigen zunächst, dass  $\sim$  eine Äquivalenzrelation wie in Definition 2.31 ist.

**Reflexivität:** Für alle  $x \in V$  gilt  $x - x = 0 \in U$  nach Definition 13.8 (a), und damit  $x \sim x$ .

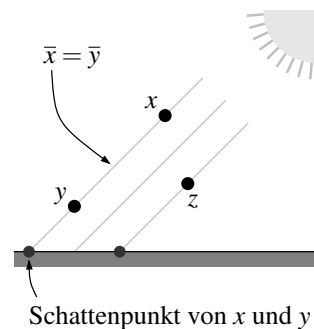
**Symmetrie:** Es seien  $x, y \in V$  mit  $x \sim y$ , also  $x - y \in U$ . Dann ist nach Definition 13.8 (c) auch  $(-1)(x - y) = y - x \in U$ , und damit  $y \sim x$ .

**Transitivität:** Sind  $x, y, z \in V$  mit  $x \sim y$  und  $y \sim z$ , also  $x - y \in U$  und  $y - z \in U$ , so ist nach Definition 13.8 (b) auch  $(x - y) + (y - z) = x - z \in U$ , und damit  $x \sim z$ .

Also ist  $\sim$  eine Äquivalenzrelation. Für die Klasse  $\bar{x}$  eines Vektors  $x$  gilt nun nach Definition 2.31

$$\bar{x} = \{y \in V : y - x \in U\} = \{y \in V : y - x = u \text{ für ein } u \in U\} = \{x + u : u \in U\}. \quad \square$$

**Bemerkung 17.15** (Anschauliche Deutung von  $V/U$ ). Die geometrische Bedeutung des Raumes  $V/U$  lässt sich am besten wie im Bild rechts erläutern, in dem  $V = \mathbb{R}^2$  und  $U = \text{Lin} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$  ist. Dort scheint die Sonne mit parallelen (hell eingezeichneten) Strahlen in Richtung von  $U$  und wirft dabei von jedem Punkt in  $V$  einen Schatten auf den Boden. In diesem Bild ist die Klasse  $\bar{x} \in V/U$  eines Punktes  $x \in V$  gerade der Sonnenstrahl durch  $x$ . Zwei Punkte in  $V$  bestimmen also genau dann den gleichen Punkt in  $V/U$ , wenn sie auf dem gleichen Sonnenstrahl liegen, d. h. denselben Schattenpunkt auf dem Boden werfen. Im Bild rechts ist also  $\bar{x} = \bar{y} \neq \bar{z}$ .



In diesem Sinne kann man sich  $V/U$  damit als eine „Schattenwelt“ von  $V$  vorstellen, die zwar jeden Punkt von  $V$  sieht, aber nur mit einem Teil seiner Informationen: Der „Abstand zur Sonne“ eines Punktes in  $V$  ist anhand des Schattenbildes nicht mehr zu rekonstruieren. Für einen Vektor  $x \in V$  nimmt die Klasse  $\bar{x}$  also wie beabsichtigt „den Anteil in  $U$  heraus“.

**Bemerkung 17.16.** Für zwei Vektoren  $x, y \in V$  gilt nach Satz 2.33 (a) genau dann  $\bar{x} = \bar{y}$  in  $V/U$ , wenn  $x \sim y$  ist. Wir sehen mit Definition 17.14 also für alle  $x, y \in V$  in  $V/U$ :

$$\begin{array}{l} \bar{x} = \bar{y} \Leftrightarrow x - y \in U, \\ \text{insbesondere also } \bar{x} = \bar{0} \Leftrightarrow x \in U. \end{array}$$

Mit diesen Rechenregeln kann man Gleichungen zwischen Äquivalenzklassen in  $V/U$  immer auf Aussagen über die Repräsentanten in  $V$  zurückführen. So ist in der Situation von Bemerkung 17.15 beispielsweise

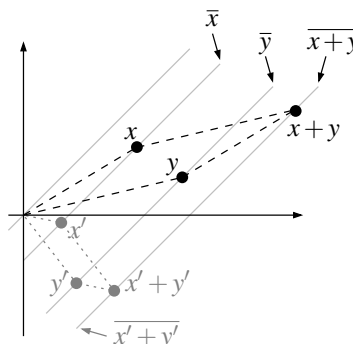
$$\overline{\begin{pmatrix} 3 \\ 5 \end{pmatrix}} = \overline{\begin{pmatrix} 6 \\ 8 \end{pmatrix}}, \quad \text{weil} \quad \begin{pmatrix} 3 \\ 5 \end{pmatrix} - \begin{pmatrix} 6 \\ 8 \end{pmatrix} = \begin{pmatrix} -3 \\ -3 \end{pmatrix} \in U.$$

Allerdings fehlt uns noch ein letzter Schritt: Bisher ist der Raum  $V/U$  nur eine Menge ohne weitere Struktur. Um ihn im Rahmen der linearen Algebra untersuchen zu können, müssen wir ihn selbst wieder zu einem Vektorraum machen, also auf ihm eine Vektoraddition und Skalarmultiplikation definieren und zeigen, dass damit dann die Vektorraumeigenschaften für  $V/U$  gelten.

Die Idee hierfür ist sehr einfach und im Bild rechts dargestellt: Wollen wir die verschobenen Unterräume  $\bar{x}$  und  $\bar{y}$  in  $V/U$  addieren, so addieren wir einfach wie im oberen Teil des Bildes die Aufpunkte  $x$  und  $y$  und verwenden den so erhaltenen Punkt  $x + y$  als Aufpunkt für die Summe, d. h. wir setzen

$$\bar{x} + \bar{y} := \overline{x + y}. \quad (*)$$

Allerdings müssen wir dabei etwas aufpassen: Wir hätten für dieselben verschobenen Unterräume statt  $x$  und  $y$  ja auch wie im unteren Teil des Bildes genauso gut andere Aufpunkte  $x'$  bzw.  $y'$  wählen können und hätten dann als Ergebnis den verschobenen Unterraum  $\overline{x' + y'}$  erhalten!



Damit die Vorschrift (\*) wirklich widerspruchsfrei eine Verknüpfung auf  $V/U$  definiert, müssen wir also überprüfen, dass der verschobene Unterraum  $\overline{x' + y'}$  derselbe ist wie  $\overline{x + y}$ , d. h. dass das *Endergebnis nicht von der Wahl der Aufpunkte abhängt*. Man sagt dazu auch, dass wir die *Wohldefiniertheit* von (\*) überprüfen müssen. Eine solche Überprüfung ist immer dann nötig, wenn wir eine Funktion auf einer Menge von Äquivalenzklassen (hier:  $V/U$ ) definieren wollen und bei der Konstruktion die Wahl eines Repräsentanten einer Äquivalenzklasse (hier: eines Aufpunktes eines verschobenen Unterrumes) verwenden. Die allgemeine Situation ist die folgende:

**Notation 17.17** (Wohldefiniertheit). Es sei  $\sim$  eine Äquivalenzrelation auf einer Menge  $M$ . Möchte man auf der Menge  $M/\sim$  der Äquivalenzklassen eine Abbildung in eine andere Menge  $N$  definieren, so ist die Idee hierfür in der Regel, dass man eine Abbildung  $g: M \rightarrow N$  wählt und dann

$$f: M/\sim \rightarrow N, \quad f(\bar{x}) := g(x) \quad (*)$$

setzt. Man möchte das Bild einer Äquivalenzklasse unter  $f$  also dadurch definieren, dass man einen Repräsentanten dieser Klasse wählt und diesen dann mit  $g$  abbildet. Damit dies nun  $f$  widerspruchsfrei definiert, brauchen wir offensichtlich, dass das Ergebnis dieser Vorschrift nicht von der Wahl des Repräsentanten abhängt: Sind  $x, y \in M$  äquivalent zueinander, sind sie also Repräsentanten derselben Äquivalenzklasse, so muss  $g(x) = g(y)$  gelten. Mit anderen Worten benötigen wir

$$g(x) = g(y) \quad \text{für alle } x, y \in M \text{ mit } \bar{x} = \bar{y},$$

damit die Definition (\*) widerspruchsfrei ist. Statt „widerspruchsfrei“ sagt man in diesem Fall wie oben schon erwähnt in der Regel, dass  $f$  durch die Vorschrift (\*) **wohldefiniert** ist. Die Wohldefiniertheit einer Funktion muss man also immer dann nachprüfen, wenn der Startraum der Funktion eine Menge von Äquivalenzklassen ist und die Funktionsvorschrift Repräsentanten dieser Klassen benutzt. Oder noch etwas allgemeiner: Wenn eine Funktionsvorschrift an irgendeiner Stelle eine Wahl beinhaltet, muss man sich vergewissern, dass der letztliche Funktionswert von dieser Wahl unabhängig ist.

Nach diesen Vorbemerkungen können wir die Menge  $V/U$  nun wie angekündigt zu einem Vektorraum machen:

**Satz und Definition 17.18** (Quotientenräume). *Es sei  $U$  ein Unterraum eines  $K$ -Vektorraums  $V$ . Dann sind die Verknüpfungen*

$$\bar{x} + \bar{y} := \overline{x+y} \quad \text{und} \quad \lambda \cdot \bar{x} := \overline{\lambda x} \quad \text{für } x, y \in V \text{ und } \lambda \in K$$

auf  $V/U$  wohldefiniert und machen  $V/U$  zu einem  $K$ -Vektorraum. Man nennt ihn den **Quotientenraum** bzw. **Faktorraum** von  $V$  nach  $U$ .

Man spricht  $V/U$  auch als „ $V$  modulo  $U$ “ und nennt  $\bar{x} \in V/U$  für ein  $x \in V$  die **Restklasse** von  $x$  modulo  $U$ .

*Beweis.* Wir zeigen zunächst die Wohldefiniertheit der Addition: Sind  $x, x', y, y' \in V$  mit  $\bar{x} = \bar{x}'$  und  $\bar{y} = \bar{y}'$ , so bedeutet dies nach Bemerkung 17.16 genau  $x - x' \in U$  und  $y - y' \in U$ . Nach Definition 13.8 (a) ist dann aber auch  $(x + y) - (x' + y') = (x - x') + (y - y') \in U$  – was wiederum nach Bemerkung 17.16 genau  $\overline{x+y} = \overline{x'+y'}$  bedeutet. Also ist die Addition auf  $V/U$  wohldefiniert.

Genauso zeigt man die Wohldefiniertheit der Skalarmultiplikation: Sind  $\lambda \in K$  und  $x, x' \in V$  mit  $\bar{x} = \bar{x}'$ , also  $x - x' \in U$ , so ist nach Definition 13.8 (a) auch  $\lambda x - \lambda x' = \lambda(x - x') \in U$  und damit  $\overline{\lambda x} = \overline{\lambda x'}$ .

Die Vektorraumaxiome für  $V/U$  ergeben sich nun unmittelbar aus denen von  $V$ . So erhält man z. B. die Assoziativität der Vektoraddition durch die einfache Rechnung

$$(\bar{x} + \bar{y}) + \bar{z} = \overline{x+y} + \bar{z} = \overline{(x+y) + z} = \overline{x + (y+z)} = \bar{x} + \overline{y+z} = \bar{x} + (\bar{y} + \bar{z})$$

für alle  $x, y, z \in V$ , wobei die mittlere Gleichheit die Assoziativität in  $V$  ist und sich die anderen Gleichungen aus der Definition der Addition in  $V/U$  ergeben. Die übrigen Eigenschaften überprüft man genauso; der Nullvektor in  $V/U$  ist die Klasse  $\bar{0}$  des Nullvektors in  $V$  bzw. der unverschobene Unterraum  $U$ , das additive Inverse eines Elements  $\bar{x} \in V/U$  ist  $\overline{-x}$ .  $\square$

Aufgrund der anschaulichen Deutung von Komplementen und Quotientenräumen sollte es nicht verwundern, dass wir nun zeigen können, dass diese beiden Konzepte letztlich das gleiche beschreiben, also als Vektorräume isomorph zueinander sind. Wie oben schon erwähnt ist der Vorteil des Komplements lediglich, dass es als Unterraum des ursprünglichen Vektorraums anschaulich leichter zu verstehen ist; der Vorteil des Quotientenraums ist dagegen, dass er ohne willkürliche Wahlen konstruiert werden kann und damit aus mathematischer Sicht das natürlichere Objekt ist.

**Satz 17.19** (Quotientenräume und Komplemente). *Es seien  $U$  ein Untervektorraum eines  $K$ -Vektorraums  $V$  und  $U'$  ein Komplement von  $U$ . Dann ist die Abbildung*

$$f: U' \rightarrow V/U, x \mapsto \bar{x}$$

ein Isomorphismus.

Ist  $V$  endlich-dimensional, so gilt also insbesondere die **Dimensionsformel für Quotientenräume**

$$\dim V/U = \dim V - \dim U.$$

*Beweis.* Um zu zeigen, dass  $f$  ein Isomorphismus ist, müssen wir die folgenden Dinge überprüfen:

- $f$  ist eine lineare Abbildung, denn für alle  $x, y \in U'$  ist

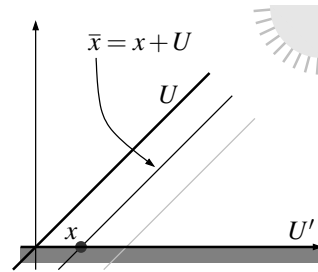
$$f(x+y) = \overline{x+y} = \overline{x} + \overline{y} = f(x) + f(y),$$

und eine analoge Aussage gilt natürlich für die Skalarmultiplikation.

- $f$  ist surjektiv: Es sei  $\bar{x} \in V/U$  beliebig, also  $x \in V$ . Wegen  $V = U + U'$  können wir  $x = x_1 + x_2$  mit  $x_1 \in U$  und  $x_2 \in U'$  schreiben. Dann liegt  $x_2$  in der Definitionsmenge  $U'$  von  $f$ , und es gilt  $f(x_2) = \bar{x}_2 = \bar{x}$  nach Bemerkung 17.16, da  $x - x_2 = x_1 \in U$ . Also ist  $f$  surjektiv.
- $f$  ist injektiv: Es sei  $x \in U'$  mit  $f(x) = \bar{x} = \bar{0}$ , also  $x \in U$  nach Bemerkung 17.16. Dann ist aber  $x \in U \cap U' = \{0\}$ . Damit ist  $f$  nach Lemma 16.9 injektiv.

Die Zusatzaussage folgt nun sofort daraus, dass das Komplement  $U'$  nach Bemerkung 17.9 die Dimension  $\dim V - \dim U$  hat. □

**Bemerkung 17.20.** Das Bild rechts illustriert in der Situation von Bemerkung 17.15 noch einmal, dass der Morphismus  $f$  aus Satz 17.19 bijektiv ist: Die Bodenlinie  $U'$  ist nach Beispiel 17.10 ein Komplement der Richtung  $U$  der Sonnenstrahlen. Die Abbildung  $f$  ordnet nun jedem Punkt  $x \in U'$  auf dem Boden den Sonnenstrahl  $\bar{x} \in V/U$  durch diesen Punkt zu, und liefert offensichtlich eine Bijektion zwischen den Bodenpunkten und der Menge der Sonnenstrahlen. Wenn wir in Bemerkung 17.15 gesagt haben, dass  $V/U$  die „Schattenwelt“ auf dem Boden ist, haben wir dabei also schon den Isomorphismus zwischen dem eigentlichen Quotientenraum  $V/U$  und dem Boden  $U'$  verwendet.



**Bemerkung 17.21** (Basen von Quotientenräumen). Nach Satz 17.19 (und Lemma 16.17 (c)) erhalten wir im endlich-dimensionalen Fall eine Basis des Quotientenraums  $V/U$ , indem wir eine Basis eines Komplements von  $U$  wählen und die Restklassen dieser Vektoren in  $V/U$  nehmen. Kombinieren wir dies mit dem Verfahren aus Bemerkung 17.12, so bedeutet dies: Wir können eine Basis von  $V/U$  konstruieren, indem wir eine Basis von  $U$  zu einer Basis von  $V$  ergänzen und dann die Restklassen der hinzugenommenen Vektoren wählen.

Im Beispiel  $V = \mathbb{R}^2$  und  $U = \text{Lin}(v)$  mit  $v = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$  aus Bemerkung 17.15 ergänzt z. B.  $e_1$  den Vektor  $v$  zu einer Basis von  $\mathbb{R}^2$ , und damit ist  $(\bar{e}_1)$  eine Basis von  $V/U$ .

Damit müssen wir z. B. den Vektor  $\bar{e}_2 \in V/U$  als Linearkombination dieser Basis, also als Vielfaches von  $\bar{e}_1$  schreiben können. Dies ist hier auch einfach zu sehen: Wegen  $e_1 + e_2 = v \in U$  ist  $\overline{e_1 + e_2} = \bar{0}$  in  $V/U$ , also  $\bar{e}_2 = -\bar{e}_1$ . Im Bild von Bemerkung 17.20 bedeutet dies einfach, dass die Vektoren  $e_2$  und  $-e_1$  auf dem gleichen Sonnenstrahl liegen.

**Aufgabe 17.22.** Es sei  $U = \{x \in \mathbb{R}^3 : -2x_1 + x_2 + x_3 = x_1 - 2x_3 = 0\} \leq \mathbb{R}^3$ .

Sind die Vektoren  $\overline{\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}}$  und  $\overline{\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}}$  linear unabhängig in  $\mathbb{R}^3/U$ ?

**Aufgabe 17.23.**

- (a) Es sei  $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3/U, x \mapsto \bar{x}$  mit  $U = \text{Lin}(e_1 - 2e_2 + e_3)$ .

Bestimme eine Basis  $B$  von  $\mathbb{R}^3/U$  sowie die zugehörige Abbildungsmatrix  $A_f^{E,B}$  für die Standardbasis  $E$  von  $\mathbb{R}^3$ .

- (b) Es sei  $U$  ein Unterraum eines endlich-dimensionalen  $K$ -Vektorraums  $V$ .

Man zeige: Ist  $(x_1, \dots, x_n)$  eine Basis von  $U$ , und sind  $y_1, \dots, y_m \in V$  so dass  $(\bar{y}_1, \dots, \bar{y}_m)$  eine Basis von  $V/U$  ist, dann ist  $(x_1, \dots, x_n, y_1, \dots, y_m)$  eine Basis von  $V$ .

**Aufgabe 17.24.**

- (a) Zeige, dass  $f: \text{Pol}_n(\mathbb{R}, \mathbb{R})/\text{Pol}_1(\mathbb{R}, \mathbb{R}) \rightarrow \mathbb{R}$ ,  $\bar{\varphi} \mapsto \varphi''(1)$  für alle  $n \in \mathbb{N}_{>0}$  wohldefiniert ist, und bestimme die Dimension von  $\text{Ker } f$ .
- (b) Es seien  $U$  ein Unterraum eines  $K$ -Vektorraums  $V$  und  $f: V \rightarrow V$  eine lineare Abbildung. Zeige, dass

$$g: V/U \rightarrow V/U, \bar{x} \mapsto \overline{f(x)}$$

genau dann eine wohldefinierte lineare Abbildung ist, wenn  $f(U) \subset U$  gilt.

Mit Hilfe von Quotientenräumen wollen wir jetzt sehen, wie man aus jeder linearen Abbildung  $f: V \rightarrow W$  „einen Isomorphismus machen kann“. Die Idee hierfür ist sehr einfach: Natürlich kann man  $f$  zunächst einmal surjektiv machen, indem man den Zielraum  $W$  durch den Bildraum  $\text{Im } f$  ersetzt. Um  $f$  auch noch injektiv zu machen, also gemäß Lemma 16.9 den Kern zu  $\{0\}$  zu machen, können wir einfach den Startraum  $V$  durch den Quotientenraum  $V/\text{Ker } f$  ersetzen: Auf diese Art werden alle Elemente des Kerns von  $f$  miteinander identifiziert, so dass der Kern der neuen Abbildung auf dem Quotientenraum nur noch aus dem einen Element  $\bar{0} = \text{Ker } f$  besteht.

**Satz 17.25 (Homomorphiesatz).** Für jede lineare Abbildung  $f: V \rightarrow W$  ist die Abbildung

$$g: V/\text{Ker } f \rightarrow \text{Im } f, \bar{x} \mapsto f(x)$$

(wohldefiniert und) ein Isomorphismus.

*Beweis.* Wir müssen einige Dinge überprüfen:

- Die Abbildung  $g$  ist wohldefiniert: Sind  $x, y \in V$  mit  $\bar{x} = \bar{y}$ , also  $x - y \in \text{Ker } f$  nach Bemerkung 17.16, so ist  $f(x - y) = f(x) - f(y) = 0$  und damit  $f(x) = f(y)$ .
- Die Abbildung  $g$  ist linear: Für  $x, y \in V$  gilt

$$g(\bar{x} + \bar{y}) = g(\overline{x + y}) = f(x + y) = f(x) + f(y) = g(\bar{x}) + g(\bar{y});$$

analog folgt auch die Verträglichkeit mit der Skalarmultiplikation.

- Die Abbildung  $g$  ist surjektiv: Dies ist klar nach Definition von  $\text{Im } f$ , denn jedes Element in  $\text{Im } f$  ist ja von der Form  $f(x) = g(\bar{x})$  für ein  $x \in V$ .
- Die Abbildung  $g$  ist injektiv: Nach Lemma 16.9 genügt es dafür zu zeigen, dass  $\text{Ker } g = \{\bar{0}\}$ . Es sei also  $x \in V$  mit  $g(\bar{x}) = 0$ . Dann ist  $f(x) = 0$ , also  $x \in \text{Ker } f$  und damit  $\bar{x} = \bar{0} \in V/\text{Ker } f$  nach Bemerkung 17.16.  $\square$

**Bemerkung 17.26** (Dimensionsformel aus dem Homomorphiesatz). Insbesondere liefert der Homomorphiesatz für eine lineare Abbildung  $f: V \rightarrow W$  zwischen endlich-dimensionalen Vektorräumen einen alternativen Beweis der Dimensionsformel für Morphismen aus Folgerung 16.30 (c), der nicht den Umweg über Matrizen und die algorithmischen Methoden aus Abschnitt 15.C nimmt: Da Isomorphismen nach Lemma 16.17 (c) die Dimension erhalten, folgt aus dem Homomorphiesatz 17.25 mit der Dimensionsformel für Quotientenräume aus Satz 17.19 sofort

$$\dim V - \dim \text{Ker } f = \dim \text{Im } f, \quad \text{also} \quad \dim \text{Im } f + \dim \text{Ker } f = \dim V.$$

**Beispiel 17.27** (Anschauliche Deutung des Homomorphiesatzes). Als anschauliches Beispiel für den Homomorphiesatz können wir noch einmal die „Schattenwelt“ aus Bemerkung 17.15 und Bemerkung 17.20 betrachten. Ist  $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  die lineare Abbildung, die einen Punkt auf seinen Schattenpunkt auf den Boden abbildet, so ist  $\text{Ker } f = U$  der Sonnenstrahl durch 0 und  $\text{Im } f = U'$  der Boden. Satz 17.25 gibt uns dann den Isomorphismus  $g: \mathbb{R}^2/U \rightarrow U'$ , der jeden Sonnenstrahl auf seinen Bodenpunkt abbildet und genau die Umkehrung des Isomorphismus aus Satz 17.19 ist.

**Aufgabe 17.28.** Die lineare Abbildung, die der Situation in Beispiel 17.27 entspricht, ist

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} x_1 - x_2 \\ 0 \end{pmatrix}.$$

Überprüfe den Homomorphiesatz in diesem Fall explizit, d. h. zeige durch eine direkte Rechnung, dass die Abbildung

$$g: \mathbb{R}^2 / \text{Lin} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rightarrow \text{Lin} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \overline{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}} \mapsto \begin{pmatrix} x_1 - x_2 \\ 0 \end{pmatrix}$$

wohldefiniert, linear, surjektiv und injektiv ist.

Als Anwendung des Homomorphiesatzes wollen wir zum Abschluss dieses Kapitels noch einmal Abbildungsmatrizen wie in Satz 16.26 und Folgerung 16.27 betrachten. Dort hatten wir gesehen, dass jeder Morphismus  $f: V \rightarrow W$  zwischen endlich-dimensionalen Vektorräumen nach Wahl von Basen  $B$  und  $C$  von  $V$  bzw.  $W$  eindeutig durch seine Abbildungsmatrix  $A_f^{B,C}$  beschrieben werden kann. Allerdings ist die Wahl solcher Basen natürlich oft willkürlich, und eine andere Wahl führt auch zu einer anderen Matrix – nämlich nach Satz 16.42 zu einer Matrix der Form  $S \cdot A_f^{B,C} \cdot T$ , wobei  $S$  und  $T$  die zugehörigen Basiswechselformen sind.

Es ist daher eine naheliegende Frage, in wie weit wir zumindest durch eine geschickte Wahl der Basen  $B$  und  $C$  erreichen können, dass die Abbildungsmatrix  $A_f^{B,C}$  möglichst einfach wird – also z. B. viele Nullen enthält. Dabei ist nach Bemerkung 16.29 natürlich klar, dass  $A_f^{B,C}$  zumindest denselben Rang wie  $f$  haben muss. Ansonsten besagt der folgende Satz aber, dass wir immer die einfachste mögliche Matrix von diesem Rang erhalten können. Der Beweis ist dabei auch konstruktiv, gibt also ein Verfahren an, wie  $B$  und  $C$  gefunden werden können.

**Satz und Definition 17.29** (Normalform von Abbildungsmatrizen). *Es sei  $f: V \rightarrow W$  ein Morphismus vom Rang  $r$  zwischen endlich-dimensionalen Vektorräumen mit  $n = \dim V$  und  $m = \dim W$ . Dann gibt es Basen  $B$  und  $C$  von  $V$  bzw.  $W$ , so dass die Abbildungsmatrix von  $f$  bezüglich dieser Basen die Form*

$$A_f^{B,C} = \left( \begin{array}{c|c} \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array} \right)$$

hat, d. h. so dass Einsen genau auf den ersten  $r$  Diagonalpositionen stehen, und sonst überall Nullen. Man sagt, dass eine solche Abbildungsmatrix in **Normalform** ist. (Beachte dabei, dass unter der Einheitsmatrix  $E_r$  genau  $m - r$  Nullzeilen, rechts von der Einheitsmatrix hingegen  $n - r$  Nullspalten stehen – die Matrix  $A_f^{B,C}$  ist also nicht notwendig quadratisch.)

*Beweis.* Nach der Dimensionsformel aus Folgerung 16.30 (c) ist  $\dim \text{Ker } f = \dim V - \text{rk } f = n - r$ . Wir können also eine Basis  $(x_{r+1}, \dots, x_n)$  von  $\text{Ker } f$  wählen und zu einer Basis  $B = (x_1, \dots, x_n)$  von  $V$  ergänzen. Damit ist  $(\bar{x}_1, \dots, \bar{x}_r)$  nach Bemerkung 17.21 eine Basis des Quotientenraums  $V / \text{Ker } f$ .

Nach dem Homomorphiesatz 17.25 ist nun aber die Abbildung  $V / \text{Ker } f \rightarrow \text{Im } f$ ,  $\bar{x} \mapsto f(x)$  ein Isomorphismus, und damit ist  $(y_1, \dots, y_r)$  mit  $y_i := f(x_i)$  für  $i = 1, \dots, r$  nach Lemma 16.17 (c) eine Basis von  $\text{Im } f$ . Wir ergänzen diese schließlich noch zu einer Basis  $C = (y_1, \dots, y_m)$  von  $W$ . Wegen

$$f(x_i) = \begin{cases} y_i & \text{für } i \leq r, \\ 0 & \text{für } i > r, \end{cases} \quad \text{also} \quad \Phi_C(f(x_i)) = \begin{cases} e_i & \text{für } i \leq r, \\ 0 & \text{für } i > r, \end{cases}$$

hat die Matrix  $A_f^{B,C}$  nach Definition 16.26 dann die gewünschte Form  $(e_1 | \dots | e_r | 0 | \dots | 0)$ .  $\square$

**Bemerkung 17.30** (Normalform von Matrizen bezüglich Äquivalenz). Nach Bemerkung 16.45 (b) können wir Satz 17.29 auch so formulieren: Ist  $A \in K^{m \times n}$  eine beliebige Matrix vom Rang  $r$ , so ist  $A$  äquivalent zur Matrix

$$\left( \begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array} \right), \quad (*)$$

d. h. es gibt invertierbare Matrizen  $S \in GL(m, K)$  und  $T \in GL(n, K)$ , so dass  $SAT$  diese eine spezielle Form hat. Insbesondere sind also alle Matrizen vom gleichen Rang zueinander äquivalent. Zusammen mit Folgerung 16.46 bedeutet dies, dass zwei Matrizen (derselben Größe) genau dann zueinander äquivalent sind, wenn sie denselben Rang haben.

Analog zu Satz 17.29 nennt man (\*) auch die **Normalform** von  $A$  (bezüglich der Äquivalenz von Matrizen); sie ist die „einfachste“ Matrix in der Äquivalenzklasse von  $A$ . Normalformen bezüglich anderer Matrixtransformationen werden wir z. B. in Kapitel ?? und Satz ?? noch kennenlernen.

**Bemerkung 17.31** (Alternativer Beweis von  $\text{rk} A = \text{rk}(A^T)$  mit Normalformen). Als Beispiel für die Nützlichkeit von Normalformen wollen wir damit nun einen alternativen Beweis der Gleichung  $\text{rk} A = \text{rk}(A^T)$  für jede Matrix  $A \in K^{m \times n}$  geben.

Bisher hatten wir diese sehr wichtige Aussage in Satz 15.38 durch den Vergleich von zwei verschiedenen Verfahren zur Berechnung von  $\text{Im} A$  gezeigt. Auch wenn dies natürlich ein vollständiger Beweis war, gibt er dennoch durch seine algorithmische Struktur wenig Erkenntnisse darüber, welche Idee dahinter steckt – das Ergebnis ist einfach am Ende aus einer nicht besonders gut überschaubaren Rechnung so herausgekommen.

Im Gegensatz dazu geben Normalformen einen sehr einfachen und sofort verständlichen Beweis dieser Aussage: Da sich der Rang einer Matrix beim Übergang zu einer äquivalenten Matrix nach Folgerung 16.46 nicht ändert, können wir ohne Einschränkung annehmen, dass die Matrix in Normalform bezüglich Äquivalenz ist. Für solche Matrizen ist die Aussage aber offensichtlich.

Die ausführliche Formulierung dieses Beweises lautet so: Nach Bemerkung 17.30 gibt es Matrizen  $S \in GL(m, K)$  und  $T \in GL(n, K)$ , so dass

$$SAT = \left( \begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array} \right) \quad \text{mit } r = \text{rk} A$$

gilt. Transponieren dieser Gleichung liefert dann mit Lemma 15.7 (d) die Matrix

$$T^T A^T S^T = \left( \begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array} \right)^T = \left( \begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array} \right) \in K^{n \times m},$$

nach Folgerung 16.46 ist also

$$\text{rk}(A^T) = \text{rk}(T^T A^T S^T) = \text{rk} \left( \begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array} \right) = r = \text{rk} A.$$

**Aufgabe 17.32.** Es sei  $A = \begin{pmatrix} 0 & 2 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 \end{pmatrix} \in \mathbb{R}^{4 \times 5}$ .

- Für die Abbildung  $f: \mathbb{R}^5 \rightarrow \mathbb{R}^4$ ,  $x \mapsto Ax$  bestimme man Basen  $B$  von  $\mathbb{R}^5$  und  $C$  von  $\mathbb{R}^4$ , so dass die Abbildungsmatrix  $A_f^{B,C}$  in Normalform ist.
- Bestimme  $S \in GL(4, \mathbb{R})$  und  $T \in GL(5, \mathbb{R})$  so, dass  $SAT$  in Normalform ist.
- Es sei  $f: \text{Pol}_4(\mathbb{R}, \mathbb{R}) \rightarrow \text{Pol}_3(\mathbb{R}, \mathbb{R})$  mit  $f(\varphi)(x) = x\varphi''(x)$  für alle  $x \in \mathbb{R}$ .  
Gibt es Basen  $B$  und  $C$  von  $\text{Pol}_4(\mathbb{R}, \mathbb{R})$  bzw.  $\text{Pol}_3(\mathbb{R}, \mathbb{R})$ , so dass  $A_f^{B,C} = A$  gilt?

**Aufgabe 17.33.**

- Bestimme Basen  $B$  und  $C$  von  $\mathbb{R}^{2 \times 2}$ , so dass die Abbildungsmatrix  $A_f^{B,C}$  der Transpositionsabbildung  $f: \mathbb{R}^{2 \times 2} \rightarrow \mathbb{R}^{2 \times 2}$ ,  $M \mapsto M^T$  in Normalform ist.
- Kann diese Normalform auch mit derselben Basis im Start- und Zielraum erreicht werden?

## 18. Determinanten

Zum Ende des ersten Teils der linearen Algebra wollen wir jetzt noch die sogenannten Determinanten einführen, die beim Rechnen mit Matrizen ein unverzichtbares Hilfsmittel sind. Determinanten haben sehr viele schöne Eigenschaften und können demzufolge auch auf viele verschiedene Arten motiviert werden. Eine mögliche Herangehensweise ist, dass man nach einem einfachen Kriterium für die Invertierbarkeit quadratischer Matrizen sucht, so wie in dem folgenden einfachen Lemma für  $2 \times 2$ -Matrizen:

**Lemma 18.1.** *Eine  $2 \times 2$ -Matrix*

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}$$

über einem Körper  $K$  ist genau dann invertierbar, wenn  $a_{1,1}a_{2,2} - a_{1,2}a_{2,1} \neq 0$  gilt.

*Beweis.* Wir unterscheiden zwei Fälle:

Fall 1: Ist  $a_{1,1} = 0$ , so ist  $A$  genau dann invertierbar, wenn  $a_{1,2} \neq 0$  und  $a_{2,1} \neq 0$  gilt – denn wenn diese beiden Einträge ungleich Null sind, sind die beiden Spalten von  $A$  offensichtlich linear unabhängig (so dass dann  $\text{rk}A = 2$  ist), während  $A$  andernfalls eine Nullzeile oder Nullspalte enthält und somit höchstens Rang 1 haben kann. Im Fall  $a_{1,1} = 0$  sind die Bedingungen  $a_{1,2} \neq 0$  und  $a_{2,1} \neq 0$  aber äquivalent zu  $a_{1,1}a_{2,2} - a_{1,2}a_{2,1} \neq 0$ .

Fall 2: Ist  $a_{1,1} \neq 0$ , so wenden wir den Gauß-Algorithmus aus Satz 15.27 an, um  $\text{rk}A$  zu berechnen:

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} \xrightarrow{\frac{1}{a_{1,1}} Z_1 \rightarrow Z_1} \begin{pmatrix} 1 & \frac{a_{1,2}}{a_{1,1}} \\ a_{2,1} & a_{2,2} \end{pmatrix} \xrightarrow{Z_2 - a_{2,1} Z_1 \rightarrow Z_2} \begin{pmatrix} 1 & \frac{a_{1,2}}{a_{1,1}} \\ 0 & a_{2,2} - \frac{a_{1,2}a_{2,1}}{a_{1,1}} \end{pmatrix}.$$

Diese Matrix hat genau dann Rang 2, ist also genau dann invertierbar, wenn  $a_{2,2} - \frac{a_{1,2}a_{2,1}}{a_{1,1}} \neq 0$ , d. h. wenn  $a_{1,1}a_{2,2} - a_{1,2}a_{2,1} \neq 0$  gilt.  $\square$

Die Zahl  $a_{1,1}a_{2,2} - a_{1,2}a_{2,1}$  werden wir später die *Determinante*  $\det A$  von  $A$  nennen (weil sie „determiniert“, ob  $A$  invertierbar ist oder nicht).

Unser Ziel in diesem Kapitel ist es, Lemma 18.1 auf größere quadratische Matrizen zu verallgemeinern, also zu jeder Matrix  $A \in K^{n \times n}$  eine Zahl  $\det A \in K$  zu definieren, die ein Polynom in den Einträgen von  $A$  ist und (neben vielen anderen schönen Eigenschaften) genau dann ungleich Null ist, wenn  $A$  invertierbar ist.

### 18.A Die Konstruktion der Determinante

Leider ist eine direkte Angabe der Determinante einer quadratischen Matrix  $A \in K^{n \times n}$  als polynomieller Ausdruck in den Einträgen von  $A$  so wie in Lemma 18.1 für allgemeines  $n$  zwar möglich (siehe Bemerkung 18.14), aber auch recht kompliziert. Wir wollen daher hier den für euch wahrscheinlich etwas ungewohnten Zugang wählen, die Determinante über ihre Eigenschaften zu definieren, d. h. als eine Funktion  $A \mapsto \det A$  auf den  $n \times n$ -Matrizen, die eine gewisse „Wunschliste“ elementarer Eigenschaften erfüllt. Im Anschluss werden wir dann zeigen, dass unsere Wunschliste wirklich erfüllbar ist und die Determinante in der Tat auch eindeutig bestimmt.

Hier ist nun unsere Wunschliste:

**Definition 18.2** (Determinante). Es seien  $K$  ein Körper und  $n \in \mathbb{N}_{>0}$  gegeben. Eine Abbildung  $\det: K^{n \times n} \rightarrow K$  heißt **Determinante** (von  $n \times n$ -Matrizen), wenn gilt:

- (a) („det ist multilinear“) Die Funktion  $\det$  ist *linear in jeder Zeile*, d. h. für alle  $k \in \{1, \dots, n\}$  und  $\lambda \in K$  gilt

$$\det \begin{pmatrix} a_1 \\ \vdots \\ a_k + a'_k \\ \vdots \\ a_n \end{pmatrix} = \det \begin{pmatrix} a_1 \\ \vdots \\ a_k \\ \vdots \\ a_n \end{pmatrix} + \det \begin{pmatrix} a_1 \\ \vdots \\ a'_k \\ \vdots \\ a_n \end{pmatrix} \quad \text{und} \quad \det \begin{pmatrix} a_1 \\ \vdots \\ \lambda a_k \\ \vdots \\ a_n \end{pmatrix} = \lambda \cdot \det \begin{pmatrix} a_1 \\ \vdots \\ a_k \\ \vdots \\ a_n \end{pmatrix},$$

wobei  $a_1, \dots, a_k, a'_k, \dots, a_n \in K^{1 \times n}$  die Zeilen der jeweiligen (quadratischen) Matrizen bezeichnen. (Halten wir also alle Zeilen bis auf die  $k$ -te fest, so haben wir genau eine lineare Abbildung in der  $k$ -ten Zeile im Sinne von Definition 16.1.)

- (b) („det ist alternierend“) Stimmen zwei Zeilen von  $A \in K^{n \times n}$  überein, so ist  $\det A = 0$ .  
 (c) („det ist normiert“) Es gilt  $\det(E_n) = 1$ .

**Beispiel 18.3.** Die Funktion

$$\det: K^{2 \times 2} \rightarrow K, \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} \mapsto a_{1,1}a_{2,2} - a_{1,2}a_{2,1}$$

aus Lemma 18.1 ist eine Determinante:

- (a)  $\det$  ist multilinear: Die Additivität in der ersten Zeile ergibt sich z. B. aus der Rechnung

$$\begin{aligned} \det \begin{pmatrix} a_{1,1} + a'_{1,1} & a_{1,2} + a'_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} &= (a_{1,1} + a'_{1,1})a_{2,2} - (a_{1,2} + a'_{1,2})a_{2,1} \\ &= a_{1,1}a_{2,2} - a_{1,2}a_{2,1} + a'_{1,1}a_{2,2} - a'_{1,2}a_{2,1} \\ &= \det \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} + \det \begin{pmatrix} a'_{1,1} & a'_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}; \end{aligned}$$

die anderen Linearitätseigenschaften folgen natürlich genauso.

- (b)  $\det$  ist alternierend: Sind die beiden Zeilen der Matrix gleich, so ist

$$\det \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{1,1} & a_{1,2} \end{pmatrix} = a_{1,1}a_{1,2} - a_{1,2}a_{1,1} = 0.$$

- (c)  $\det$  ist normiert, denn natürlich ist  $\det(E_2) = 1$ .

**Bemerkung 18.4.**

- (a) Wir werden in Folgerung 18.8 und Satz 18.12 noch sehen, dass es zu jedem Körper  $K$  und jedem  $n \in \mathbb{N}_{>0}$  in der Tat genau eine Determinante  $\det: K^{n \times n} \rightarrow K$  gibt, dass Definition 18.2 die Determinante also widerspruchsfrei und eindeutig festlegt. Solange wir dies noch nicht gezeigt haben, sollten wir aber korrekterweise immer von *einer* Determinante (und nicht von *der* Determinante) sprechen.
- (b) Enthält  $A$  eine Nullzeile, so können wir aus dieser Zeile den Faktor 0 herausziehen und erhalten aus der Linearitätseigenschaft in dieser Zeile sofort, dass dann  $\det A = 0$  sein muss.
- (c) Aus Eigenschaft (b) der Definition 18.2 einer Determinante folgt, dass sich  $\det A$  beim Vertauschen zweier Zeilen mit  $-1$  multipliziert, also genau das Vorzeichen ändert (daher kommt auch der Name „alternierend“ für diese Eigenschaft): Für alle  $k, l \in \{1, \dots, n\}$  mit  $k \neq l$  ergibt sich zusammen mit der Multilinearität nämlich (wobei die angegebenen Einträge in den

Zeilen  $k$  bzw.  $l$  stehen)

$$\begin{aligned} \det \underbrace{\begin{pmatrix} \vdots \\ a_k + a_l \\ \vdots \\ a_k + a_l \\ \vdots \end{pmatrix}}_{=0} &= \det \begin{pmatrix} \vdots \\ a_k \\ \vdots \\ a_k + a_l \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ a_l \\ \vdots \\ a_k + a_l \\ \vdots \end{pmatrix} \\ &= \det \underbrace{\begin{pmatrix} \vdots \\ a_k \\ \vdots \\ a_k \\ \vdots \end{pmatrix}}_{=0} + \det \begin{pmatrix} \vdots \\ a_k \\ \vdots \\ a_l \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ a_l \\ \vdots \\ a_k \\ \vdots \end{pmatrix} + \det \underbrace{\begin{pmatrix} \vdots \\ a_l \\ \vdots \\ a_l \\ \vdots \end{pmatrix}}_{=0}, \end{aligned}$$

und damit

$$\det \begin{pmatrix} \vdots \\ a_k \\ \vdots \\ a_l \\ \vdots \end{pmatrix} = -\det \begin{pmatrix} \vdots \\ a_l \\ \vdots \\ a_k \\ \vdots \end{pmatrix}.$$

- (d) Analog zu (c) wollen wir jetzt untersuchen, was mit einer Determinante passiert, wenn wir in einer Matrix  $A$  für gegebenes  $k \in \{1, \dots, n\}$  die  $k$ -te Zeile unter Beibehaltung der Reihenfolge der anderen Zeilen ganz nach oben schieben. Wir können dies wie folgt durch  $k-1$  Vertauschungen zweier benachbarter Zeilen erreichen:

$$A = \begin{pmatrix} \vdots \\ a_{k-2} \\ a_{k-1} \\ a_k \\ a_{k+1} \\ \vdots \end{pmatrix} \begin{array}{c} \curvearrowright \\ \curvearrowright \\ \curvearrowright \end{array} \longrightarrow \begin{pmatrix} \vdots \\ a_{k-2} \\ a_k \\ a_{k-1} \\ a_{k+1} \\ \vdots \end{pmatrix} \longrightarrow \dots \longrightarrow \begin{pmatrix} a_k \\ a_1 \\ \vdots \\ a_{k-1} \\ a_{k+1} \\ \vdots \end{pmatrix}.$$

Da sich bei jeder dieser Vertauschungen nach (c) das Vorzeichen der Determinante ändert, ändert das gesamte Verschieben der  $k$ -ten Zeile ganz nach oben die Determinante von  $A$  also um einen Faktor  $(-1)^{k-1}$ .

Um die weiteren Eigenschaften von Determinanten zu untersuchen, beginnen wir zunächst mit den Elementarmatrizen.

**Lemma 18.5** (Determinanten von Elementarmatrizen). *Es sei  $\det: K^{n \times n} \rightarrow K$  eine Determinante. Dann gilt für alle  $A \in K^{n \times n}$  sowie für alle  $n \times n$ -Elementarmatrizen  $F_k(\lambda)$  und  $F_{k,l}(\lambda)$  aus Konstruktion 15.24:*

- (a)  $\det(F_k(\lambda) \cdot A) = \lambda \det A$ .
- (b)  $\det(F_{k,l}(\lambda) \cdot A) = \det A$ .

*Insbesondere gilt also  $\det F_k(\lambda) = \lambda$  und  $\det F_{k,l}(\lambda) = 1$ , und damit  $\det(F A) = \det F \cdot \det A$  für jede Elementarmatrix  $F$  und jede beliebige quadratische Matrix  $A$ .*

*Beweis.* Es seien  $a_1, \dots, a_n \in K^{1 \times n}$  die Zeilen von  $A$ . Nach Konstruktion 15.24 entspricht eine Multiplikation von  $A$  mit einer Elementarmatrix von links genau einer elementaren Zeilenumformung. Damit erhalten wir mit den Eigenschaften (a) und (b) aus Definition 18.2

$$\det(F_k(\lambda) \cdot A) = \det \begin{pmatrix} \vdots \\ \lambda a_k \\ \vdots \end{pmatrix} = \lambda \det \begin{pmatrix} \vdots \\ a_k \\ \vdots \end{pmatrix} = \lambda \det A$$

und

$$\det(F_{k,l}(\lambda) \cdot A) = \det \begin{pmatrix} \vdots \\ a_k + \lambda a_l \\ \vdots \\ a_l \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ a_k \\ \vdots \\ a_l \\ \vdots \end{pmatrix} + \underbrace{\lambda \det \begin{pmatrix} \vdots \\ a_l \\ \vdots \\ a_l \\ \vdots \end{pmatrix}}_{=0} = \det \begin{pmatrix} \vdots \\ a_k \\ \vdots \\ a_l \\ \vdots \end{pmatrix} = \det A,$$

was die beiden Teile des Lemmas zeigt. Die Determinanten der Elementarmatrizen erhält man daraus für  $A = E_n$ .  $\square$

Aus diesem einfachen Lemma folgt nun bereits die wahrscheinlich wichtigste Eigenschaft von Determinanten:

**Satz 18.6 (Produktsatz für Determinanten).** *Es sei  $\det: K^{n \times n} \rightarrow K$  eine Determinante. Dann gilt für alle  $A, B \in K^{n \times n}$ :*

- (a)  $\det(AB) = \det A \cdot \det B$ .
- (b)  $A$  ist genau dann invertierbar, wenn  $\det A \neq 0$ . In diesem Fall ist  $\det(A^{-1}) = \frac{1}{\det A}$ .

*Beweis.* Wir unterscheiden zwei Fälle:

Fall 1:  $A$  ist invertierbar. Dann ist  $A = F_1 \cdot \dots \cdot F_k$  nach Folgerung 15.36 ein Produkt von Elementarmatrizen. Durch  $k$ -fache Anwendung von Lemma 18.5 erhält man also

$$\det(AB) = \det(F_1 \cdot \dots \cdot F_k \cdot B) = \det F_1 \cdot \dots \cdot \det F_k \cdot \det B$$

sowie

$$\det A = \det(F_1 \cdot \dots \cdot F_k) = \det F_1 \cdot \dots \cdot \det F_k,$$

und damit wie behauptet  $\det(AB) = \det A \cdot \det B$ . Setzt man hier  $B = A^{-1}$  ein, so ergibt sich insbesondere  $\det A \cdot \det A^{-1} = \det(AA^{-1}) = \det E_n = 1$ , d. h. es gilt auch  $\det A \neq 0$  und  $\det(A^{-1}) = \frac{1}{\det A}$ .

Fall 2:  $A$  ist nicht invertierbar, also  $\text{rk} A < n$ . Bringen wir  $A$  dann mit einem Produkt  $F$  von Elementarmatrizen auf Zeilenstufenform  $FA$ , so hat  $FA$  weniger als  $n$  Stufen und damit am Ende (mindestens) eine Nullzeile. Also ist  $\det(FA) = 0$  nach Bemerkung 18.4 (b). Da  $F$  als Produkt von Elementarmatrizen invertierbar ist, bedeutet dies nach dem bereits gezeigten Fall 1 auch  $\det F \cdot \det A = 0$ ; wegen  $\det F \neq 0$  also  $\det A = 0$ .

Mit  $FA$  hat aber auch  $FAB$  eine Nullzeile. Damit folgt genauso wie oben auch  $\det(AB) = 0$ , also insbesondere  $\det(AB) = \det A \cdot \det B$ .  $\square$

**Bemerkung 18.7.** Im Gegensatz zu Produkten gibt es *keine* Formel für die Determinante  $\det(A+B)$  einer Summe von zwei Matrizen – insbesondere ist im Allgemeinen  $\det(A+B) \neq \det A + \det B$ !

Als Folgerung aus dem Produktsatz können wir nun bereits beweisen, dass die Eigenschaften aus Definition 18.2 eine Determinante eindeutig festlegen.

**Folgerung 18.8** (Eindeutigkeit der Determinante). *Zu jedem Körper  $K$  und  $n \in \mathbb{N}_{>0}$  gibt es höchstens eine Determinante  $\det: K^{n \times n} \rightarrow K$ .*

*Beweis.* Es sei  $A \in K^{n \times n}$ . Ist  $A$  nicht invertierbar, so ist nach Satz 18.6 notwendigerweise  $\det A = 0$ . Andernfalls ist  $A = F_1 \cdots F_k$  nach Folgerung 15.36 ein Produkt von Elementarmatrizen, und damit ist nach Satz 18.6 (a)

$$\det A = \det F_1 \cdots \det F_k.$$

Da die Determinante der Elementarmatrizen nach Lemma 18.5 aber durch Definition 18.2 eindeutig bestimmt ist, ist damit auch  $\det A$  durch diese Definition eindeutig festgelegt.  $\square$

Auf ganz ähnliche Art wollen wir nun zeigen, dass sich eine Determinante beim Transponieren der Matrizen nicht ändert.

**Folgerung 18.9.** *Ist  $A \in K^{n \times n}$  und  $\det: K^{n \times n} \rightarrow K$  eine Determinante, so gilt  $\det(A^T) = \det A$ .*

*Beweis.* Ist  $A$  nicht invertierbar, also  $\text{rk} A < n$ , so ist nach Bemerkung 15.40 auch  $A^T$  nicht invertierbar, und damit ist  $\det(A^T) = 0 = \det A$  nach Satz 18.6 (b).

Andernfalls ist  $A = F_1 \cdots F_k$  nach Folgerung 15.36 wieder ein Produkt von Elementarmatrizen. Da die zu zeigende Aussage für Elementarmatrizen aus Lemma 18.5 offensichtlich ist (es ist nämlich  $(F_k(\lambda))^T = F_k(\lambda)$  und  $(F_{k,l}(\lambda))^T = F_{l,k}(\lambda)$ ), folgt somit nach Lemma 15.7 (d) und Satz 18.6 (a)

$$\det(A^T) = \det((F_1 \cdots F_k)^T) = \det(F_k^T \cdots F_1^T) = \det(F_k^T) \cdots \det(F_1^T) = \det F_1 \cdots \det F_k = \det A. \quad \square$$

**Bemerkung 18.10.** Folgerung 18.9 besagt anschaulich, dass alle Eigenschaften, die für die Zeilen einer Determinante gelten, analog auch für die Spalten gelten. So ist eine Determinante z. B. auch linear in jeder Spalte (vgl. Definition 18.2 (a)) und ändert ihr Vorzeichen beim Vertauschen zweier Spalten (vgl. Bemerkung 18.4 (c)).

Um sicherzustellen, dass wir mit Definition 18.2 keine in sich widersprüchliche Wunschliste aufgeschrieben haben, kommen wir nun aber endlich zum bereits angekündigten Resultat, dass eine Determinante mit den geforderten Eigenschaften auch wirklich existiert. Wir werden die Funktionen  $\det: K^{n \times n} \rightarrow K$  rekursiv über  $n$  definieren und verwenden dazu die folgende Konstruktion, um Matrizen der Größe  $n$  auf solche der Größe  $n-1$  zurückzuführen.

**Definition 18.11** (Streichungsmatrix). Zu  $A = (a_{i,j})_{i,j} \in K^{n \times n}$  sowie  $k, l \in \{1, \dots, n\}$  sei

$$A'_{k,l} := \begin{pmatrix} a_{1,1} & \cdots & a_{1,l-1} & a_{1,l+1} & \cdots & a_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{k-1,1} & \cdots & a_{k-1,l-1} & a_{k-1,l+1} & \cdots & a_{k-1,n} \\ a_{k+1,1} & \cdots & a_{k+1,l-1} & a_{k+1,l+1} & \cdots & a_{k+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,l-1} & a_{n,l+1} & \cdots & a_{n,n} \end{pmatrix} \in K^{(n-1) \times (n-1)}$$

die Matrix, die man erhält, wenn man aus  $A$  die  $k$ -te Zeile und  $l$ -te Spalte herausstreicht. Wir bezeichnen diese Matrizen als **Streichungsmatrizen** zu  $A$ .

**Satz 18.12** (Existenz der Determinante). *Für alle  $n \in \mathbb{N}_{>0}$  definieren wir  $\det: K^{n \times n} \rightarrow K$  rekursiv über  $n$  durch die folgende Vorschrift:*

- Für  $n = 1$  setzen wir  $\det(a_{1,1}) := a_{1,1}$ .
- Für  $n > 1$  setzen wir

$$\det A := \sum_{k=1}^n (-1)^{k+1} a_{k,1} \det A'_{k,1},$$

wobei wie üblich  $a_{k,1}$  die Einträge der ersten Spalte von  $A$  und  $A'_{k,1}$  die zu diesen Einträgen gehörigen Streichungsmatrizen sind.

Dann ist  $\det$  eine (und damit nach Folgerung 18.8 „die“) Determinante für alle  $n$ .

40

Bevor wir diesen Satz beweisen, wollen wir uns ein paar Beispiele anschauen, um die angegebene rekursive Formel besser zu verstehen.

**Beispiel 18.13** (Determinante von  $2 \times 2$ - und  $3 \times 3$ -Matrizen).

(a) Für  $n = 2$  besagt die Formel aus Satz 18.12

$$\det \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} = (-1)^{1+1} a_{1,1} \det(a_{2,2}) + (-1)^{2+1} a_{2,1} \det(a_{1,2})$$

$$= a_{1,1} a_{2,2} - a_{2,1} a_{1,2}$$

und reproduziert damit die Formel aus Lemma 18.1.

(b) Für  $n = 3$  ergibt sich unter Benutzung des Ergebnisses aus (a)

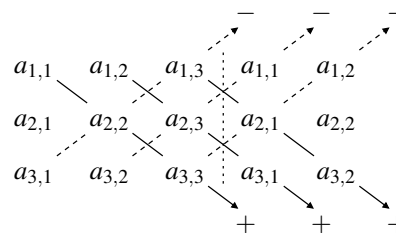
$$\det \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix} = (-1)^{1+1} a_{1,1} \det \begin{pmatrix} a_{2,2} & a_{2,3} \\ a_{3,2} & a_{3,3} \end{pmatrix} + (-1)^{2+1} a_{2,1} \det \begin{pmatrix} a_{1,2} & a_{1,3} \\ a_{3,2} & a_{3,3} \end{pmatrix}$$

$$+ (-1)^{3+1} a_{3,1} \det \begin{pmatrix} a_{1,2} & a_{1,3} \\ a_{2,2} & a_{2,3} \end{pmatrix}$$

$$= a_{1,1} a_{2,2} a_{3,3} - a_{1,1} a_{2,3} a_{3,2} - a_{2,1} a_{1,2} a_{3,3} + a_{2,1} a_{1,3} a_{3,2}$$

$$+ a_{3,1} a_{1,2} a_{2,3} - a_{3,1} a_{1,3} a_{2,2}.$$

Am einfachsten kann man sich diese Formel nach der sogenannten **Regel von Sarrus** merken: Bilden wir die  $3 \times 5$ -Matrix, in der wir neben der Matrix  $A$  die beiden ersten Spalten noch einmal wiederholen, so ergeben sich die 6 Terme der Determinante mit ihren Vorzeichen aus dem folgenden Schema:



Beachte aber, dass diese einfache Merkmethode *nur für  $n = 3$  gilt* – für größere  $n$  ist der komplett ausmultiplizierte Ausdruck für  $\det A$  deutlich komplizierter (und für konkrete numerische Berechnungen in der Tat auch nicht mehr geeignet).

**Bemerkung 18.14.** Diejenigen von euch, die aus der Parallelvorlesung „Algebraische Strukturen“ die symmetrische Gruppe  $S_n$  aller Permutationen von  $\{1, \dots, n\}$  kennen [G, Kapitel 2], können die Formel für die Determinante einer Matrix  $A = (a_{i,j})_{i,j} \in K^{n \times n}$  auch nicht-rekursiv als

$$\det A = \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot a_{1,\sigma(1)} \cdot \dots \cdot a_{n,\sigma(n)} \tag{*}$$

hinschreiben. Man sieht an dieser Darstellung also, dass die Determinante aus einer Summe von  $n!$  Termen besteht. Dabei ist jeder Term ein Produkt von genau  $n$  Einträgen von  $A$ , und zwar aus jeder Zeile und jeder Spalte genau einem. Aufsummiert wird über alle Möglichkeiten,  $n$  Einträge von  $A$  eben gerade so auszuwählen, dass man aus jeder Zeile und Spalte einen Eintrag genommen hat. Die Vorzeichen der einzelnen Terme sind immer genau das Vorzeichen der entsprechenden Permutation.

Wir werden die Formel (\*) in dieser Vorlesung aber nicht benötigen und daher auch nicht beweisen, dass sie wirklich mit der rekursiven Definition aus Satz 18.12 übereinstimmt bzw. die Eigenschaften von Definition 18.2 erfüllt.

Wir kommen nun aber endlich zum Beweis des Existenzsatzes 18.12.

*Beweis von Satz 18.12.* Wir überprüfen die drei Eigenschaften aus Definition 18.2 mit Induktion über  $n$ . Für  $n = 1$  sind alle Aussagen klar. Wir können also annehmen, dass  $n > 1$  ist und wir die Eigenschaften von Definition 18.2 für Matrizen der Größe  $n - 1$  bereits gezeigt haben; wir müssen sie nun für Matrizen der Größe  $n$  zeigen.

det ist multilinear: Der Ausdruck  $a_{1,1} \det A'_{i,1}$  ist linear in der ersten Zeile, da  $a_{1,1}$  natürlich linear in der ersten Zeile ist und  $A'_{i,1}$  nicht von der ersten Zeile abhängt. Die Ausdrücke  $a_{k,1} \det A'_{k,1}$  für  $k > 1$  sind ebenfalls linear in der ersten Zeile, da  $a_{k,1}$  nicht von der ersten Zeile abhängt und  $\det A'_{k,1}$  nach Induktionsvoraussetzung linear in der ersten Zeile ist. Damit ist auch  $\det A$  als Linearkombination dieser Ausdrücke linear in der ersten Zeile. Die Linearität in den anderen Zeilen folgt natürlich analog.

det ist alternierend: Wir bezeichnen die Zeilen von  $A$  mit  $a_1, \dots, a_n \in K^{1 \times n}$ . Weiterhin seien  $a'_1, \dots, a'_n \in K^{1 \times (n-1)}$  die Zeilen von  $A$ , bei denen man jeweils den ersten Eintrag herausgestrichen hat. Wir nehmen nun an, dass zwei Zeilen  $a_i$  und  $a_j$  von  $A$  übereinstimmen, und müssen zeigen, dass  $\det A = 0$  folgt. Ohne Beschränkung der Allgemeinheit sei dazu  $i > j$ .

Beachte, dass dann auch in den Streichungsmatrizen  $A'_{k,1}$  mit  $k \neq i$  und  $k \neq j$ , bei denen wir also weder die  $i$ -te noch die  $j$ -te Zeile herausgestrichen haben, jeweils zwei Zeilen übereinstimmen. Nach Induktionsvoraussetzung ist die Determinante aller dieser Streichungsmatrizen gleich 0, und damit bleibt in der rekursiven Formel für  $\det A$  nur der Ausdruck

$$\det A = (-1)^{i+1} a_{i,1} \det A'_{i,1} + (-1)^{j+1} a_{j,1} \det A'_{j,1} \quad (*)$$

übrig. Nun können wir wegen  $a'_i = a'_j$

$$\text{sowohl } A'_{i,1} = \begin{pmatrix} a'_1 \\ \vdots \\ a'_{j-1} \\ a'_j \\ a'_{j+1} \\ \vdots \\ a'_{i-1} \\ a'_{i+1} \\ \vdots \end{pmatrix} \quad \text{als auch } A'_{j,1} = \begin{pmatrix} a'_1 \\ \vdots \\ a'_{j-1} \\ a'_{j+1} \\ \vdots \\ a'_{i-1} \\ a'_i \\ a'_{i+1} \\ \vdots \end{pmatrix} \quad \text{auf die Form } A' := \begin{pmatrix} a'_i \\ a'_1 \\ \vdots \\ a'_{j-1} \\ a'_{j+1} \\ \vdots \\ a'_{i-1} \\ a'_{i+1} \\ \vdots \end{pmatrix}$$

bringen, indem wir die Zeile  $a'_j$  bzw.  $a'_i$  unter Beibehaltung der Reihenfolge der anderen Zeilen ganz nach oben schieben. Da det für Matrizen der Größe  $n-1$  nach Induktionsvoraussetzung eine Determinante ist, ändern sich dadurch die Vorzeichen von  $\det A'_{i,1}$  und  $\det A'_{j,1}$  wie in Bemerkung 18.4 (d): Da wir in  $A'_{i,1}$  die Zeile mit der Nummer  $j$ , in  $A'_{j,1}$  jedoch die Zeile mit der Nummer  $i-1$  nach oben schieben (im letzteren Fall fehlt ja die Zeile  $a'_j$  oberhalb von  $a'_i$ ), ist also

$$\det A'_{i,1} = (-1)^{j-1} \det A' \quad \text{und} \quad \det A'_{j,1} = (-1)^{i-2} \det A'$$

und damit nach (\*)

$$\det A = (-1)^{i+j} a_{i,1} \det A' + (-1)^{i+j-1} a_{j,1} \det A' = 0$$

wegen  $a_{i,1} = a_{j,1}$ .

det ist normiert: In der ersten Spalte der Einheitsmatrix sind natürlich der erste Eintrag gleich 1 und alle anderen gleich 0. Weiterhin ist die Streichungsmatrix des Eintrags links oben gerade  $E_{n-1}$ . Also folgt sofort

$$\det E_n = (-1)^{1+1} \cdot 1 \cdot \det E_{n-1} = 1.$$

Damit ist alles gezeigt. □

Insgesamt haben wir jetzt also gesehen, dass es für alle Körper  $K$  und  $n \in \mathbb{N}$  genau eine Determinante  $\det: K^{n \times n} \rightarrow K$  gibt. In Zukunft werden wir daher immer von *der* Determinante quadratischer Matrizen sprechen.

## 18.B Eigenschaften der Determinante

Im letzten Abschnitt haben wir die Determinante quadratischer Matrizen definiert und auch bereits ihre ersten wichtigen Eigenschaften gesehen. Wir wollen diese Untersuchung der Determinante jetzt fortsetzen und uns dabei als Erstes um ihre praktische Berechnung kümmern. In der Tat ist hierfür die rekursive Formel aus Satz 18.12 bereits sehr nützlich. Wir können sie allerdings noch etwas erweitern, denn dort ist ja momentan die erste Spalte der Matrix ausgezeichnet – obwohl aufgrund von Definition 18.2 natürlich klar sein sollte, dass die erste Spalte der Matrix keine besondere Rolle spielt. Wir sollten eine ähnliche Rekursionsformel also auch für die anderen Spalten (und aufgrund von Folgerung 18.9 in der Tat auch für die Zeilen) erwarten können. Dies besagt der folgende Satz.

**Satz 18.15 (Laplacescher Entwicklungssatz).** *Es sei  $A = (a_{i,j})_{i,j} \in K^{n \times n}$ .*

- (a) Für alle  $l \in \{1, \dots, n\}$  gilt  $\det A = \sum_{k=1}^n (-1)^{k+l} \cdot a_{k,l} \cdot \det A'_{k,l}$ .  
 (b) Für alle  $k \in \{1, \dots, n\}$  gilt  $\det A = \sum_{l=1}^n (-1)^{k+l} \cdot a_{k,l} \cdot \det A'_{k,l}$ .

Benutzt man diese Formeln, so sagt man auch, dass man die Determinante von  $A$  nach der  $l$ -ten Spalte bzw.  $k$ -ten Zeile entwickelt.

*Beweis.*

- (a) Es sei  $B = (b_{i,j})_{i,j}$  die Matrix, die man aus  $A$  erhält, indem man die Spalte  $l$  unter Beibehaltung der Reihenfolge der anderen Spalten ganz nach links schiebt. Nach den Bemerkungen 18.4 (d) und 18.10 ist dann  $\det A = (-1)^{l-1} \det B$ . Andererseits ist natürlich  $b_{k,1} = a_{k,l}$  und  $B'_{k,1} = A'_{k,l}$  für alle  $k \in \{1, \dots, n\}$ . Damit folgt wie behauptet nach Satz 18.12 angewendet auf  $B$

$$\det A = (-1)^{l-1} \det B = (-1)^{l-1} \sum_{k=1}^n (-1)^{k+1} b_{k,1} \det B'_{k,1} = \sum_{k=1}^n (-1)^{k+l} a_{k,l} \det A'_{k,l}.$$

- (b) Dies ergibt sich mit Bemerkung 18.10 sofort aus (a). □

**Beispiel 18.16** (Berechnung von Determinanten). Die Entwicklung nach Laplace ist oft die geschickteste Art, die Determinante einer Matrix  $A$  konkret zu berechnen – insbesondere wenn man nach einer Spalte oder Zeile entwickeln kann, in der bereits viele Einträge gleich 0 sind, so dass die entsprechenden Terme in der Summe wegfallen. In der Praxis empfiehlt es sich daher, zunächst mit elementaren Spalten- oder Zeilenumformungen eine Spalte oder Zeile zu erzeugen, in der nur ein Eintrag ungleich Null ist, und dann nach dieser Spalte bzw. Zeile zu entwickeln. Beachte, dass die Determinante dabei nach Lemma 18.5 ...

- mit  $\lambda$  multipliziert wird, wenn wir eine Spalte oder Zeile mit  $\lambda$  multiplizieren; und
- unverändert bleibt, wenn wir ein Vielfaches einer Spalte bzw. Zeile zu einer anderen addieren.

Hier ist ein Beispiel, bei dem wir der Reihe nach die erste von der dritten Spalte subtrahieren, nach der dritten Spalte entwickeln, und noch einmal nach der zweiten Zeile entwickeln: Es ist

$$\begin{aligned} \det \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 1 & 0 & 3 \end{pmatrix} &\stackrel{s_3-s_1 \rightarrow s_3}{=} \det \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 1 & 0 & 3 \end{pmatrix} = (-1)^{3+3} \cdot 2 \cdot \det \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 3 \end{pmatrix} \\ &= 2 \cdot (-1)^{2+1} \cdot 1 \cdot \det \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix} \\ &= -4. \end{aligned}$$

Ein besonders einfacher Fall – der aber dennoch häufig vorkommt – sind die sogenannten Dreiecksmatrizen, bei denen oberhalb oder unterhalb der Diagonale nur Nullen stehen.

**Definition 18.17** (Dreiecksmatrizen). Eine quadratische Matrix  $A = (a_{i,j})_{i,j} \in K^{n \times n}$  heißt **obere Dreiecksmatrix**, falls  $a_{i,j} = 0$  für alle  $i > j$  gilt, und **untere Dreiecksmatrix**, falls  $a_{i,j} = 0$  für alle  $i < j$  gilt. Obere bzw. untere Dreiecksmatrizen haben also die Form

$$\begin{pmatrix} a_{1,1} & & * \\ & \ddots & \\ 0 & & a_{n,n} \end{pmatrix} \quad \text{bzw.} \quad \begin{pmatrix} a_{1,1} & & 0 \\ & \ddots & \\ * & & a_{n,n} \end{pmatrix}.$$

Sind zusätzlich noch alle Einträge  $a_{i,i}$  auf der Diagonale gleich Null, so heißt  $A$  **echte (obere bzw. untere) Dreiecksmatrix**.

**Folgerung 18.18** (Determinante von Dreiecksmatrizen). Ist  $A = (a_{i,j})_{i,j} \in K^{n \times n}$  eine (obere oder untere) Dreiecksmatrix, so ist ihre Determinante gleich dem Produkt ihrer Einträge auf der Diagonale

$$\det A = a_{1,1} \cdot \cdots \cdot a_{n,n}.$$

*Beweis.* Da untere Dreiecksmatrizen beim Transponieren in obere übergehen, reicht es nach Folgerung 18.9, die Aussage für obere Dreiecksmatrizen zu zeigen. Wir beweisen die Aussage in diesem Fall mit Induktion über  $n$ ; der Fall  $n = 1$  ist dabei trivial. Für  $n > 1$  entwickeln wir  $\det A$  gemäß Satz 18.15 nach der 1. Spalte: Da hier nur der erste Eintrag ungleich Null ist, ergibt sich sofort nach Induktionsvoraussetzung

$$\det A = (-1)^{1+1} a_{1,1} \det A'_{1,1} = a_{1,1} \cdot (a_{2,2} \cdot \cdots \cdot a_{n,n}),$$

da auch  $A'_{1,1} \in K^{(n-1) \times (n-1)}$  eine obere Dreiecksmatrix (mit Diagonaleinträgen  $a_{2,2}, \dots, a_{n,n}$ ) ist.  $\square$

**Aufgabe 18.19.**

(a) Berechne  $\det(A^5)$  und  $\det(5A)$  für die Matrix

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 8 \end{pmatrix} \in \mathbb{R}^{3 \times 3}.$$

(b) Für  $a_1, \dots, a_n \in K \setminus \{0\}$  zeige man

$$\det \begin{pmatrix} 0 & 1 & 1 & \cdots & 1 \\ 1 & a_1 & 0 & \cdots & 0 \\ 1 & 0 & a_2 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 1 & 0 & \cdots & 0 & a_n \end{pmatrix} = - \left( \prod_{i=1}^n a_i \right) \cdot \left( \sum_{i=1}^n \frac{1}{a_i} \right).$$

**Aufgabe 18.20.** Es sei  $A \in K^{n \times n}$  eine quadratische Matrix, die eine Blockgestalt der Form

$$A = \left( \begin{array}{c|c} B & * \\ \hline 0 & C \end{array} \right)$$

hat, wobei  $B \in K^{m \times m}$  und  $C \in K^{(n-m) \times (n-m)}$  selbst quadratische Matrizen sind. Zeige, dass dann  $\det A = \det B \cdot \det C$  gilt.

(Hinweis: Es hilft, zunächst die Fälle zu betrachten, in denen eine der Matrizen  $B$  und  $C$  nicht invertierbar oder die Einheitsmatrix ist.)

Wir wollen nun noch zwei Ergebnisse zu Determinanten beweisen, die mehr aus theoretischer als aus rechnerischer Sicht interessant sind. Das erste betrifft inverse Matrizen: Ist  $A$  eine invertierbare Matrix, so haben wir in Satz 15.35 ja bereits gesehen, wie man  $A^{-1}$  konkret berechnen kann. Mit Hilfe von Determinanten können wir nun auch eine explizite Formel für  $A^{-1}$  angeben – die allerdings den Nachteil hat, dass sie bei konkreten Berechnungen relativ aufwendig ist, weil für jeden Eintrag von  $A^{-1}$  eine eigene Determinante berechnet werden muss.

**Satz 18.21** (Explizite Formel für die inverse Matrix). *Es sei  $A = (a_{i,j})_{i,j} \in K^{n \times n}$ .*

(a) *Ist  $C = (c_{i,j})_{i,j} \in K^{n \times n}$  die Matrix mit Einträgen*

$$c_{i,j} = (-1)^{i+j} \det A'_{j,i}$$

*(beachte die Vertauschung von Spalten- und Zeilenindizes bei der Streichungsmatrix!), so ist  $CA = AC = (\det A) \cdot E_n$ .*

(b) *Ist  $A$  invertierbar, so ist die inverse Matrix von  $A$  gegeben durch*

$$A^{-1} = \frac{1}{\det A} \cdot C = \left( (-1)^{i+j} \frac{\det A'_{j,i}}{\det A} \right)_{i,j}$$

*mit  $C$  wie in (a).*

*Beweis.* Für alle  $i, k = 1, \dots, n$  überprüfen wir den  $(i, k)$ -Eintrag des Matrixprodukts  $CA$ : Nach Definition 15.5 ist dies

$$\sum_{j=1}^n c_{i,j} a_{j,k} = \sum_{j=1}^n (-1)^{i+j} a_{j,k} \det A'_{j,i} \stackrel{18.15}{=} \det \begin{matrix} & & \text{Spalte } i \\ & & \downarrow \\ \begin{pmatrix} a_{1,1} & \cdots & a_{1,k} & \cdots & a_{1,n} \\ \vdots & & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,k} & \cdots & a_{n,n} \end{pmatrix} \end{matrix},$$

wobei die zweite Gleichung genau die Entwicklung nach Spalte  $i$  ist, und die Matrix auf der rechten Seite aus  $A$  entsteht, indem die Einträge aus Spalte  $k$  auch in Spalte  $i$  geschrieben werden. Die Determinante dieser Matrix ist aber 0 für  $i \neq k$  (da dann zwei gleiche Spalten existieren) und  $\det A$  für  $i = k$  (denn dann ist diese Matrix gleich  $A$ ). Damit ist  $CA = (\det A) E_n$ .

Analog zeigt man auch  $AC = (\det A) E_n$  und damit Teil (a). Die Formel in (b) folgt daraus natürlich sofort mit Division durch  $\det A$ . □

**Beispiel 18.22.** Für eine  $2 \times 2$ -Matrix

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}$$

hat die Matrix  $C$  aus Satz 18.21 die Einträge

$$c_{1,1} = (-1)^{1+1} \det(a_{2,2}) = a_{2,2}, \quad c_{1,2} = (-1)^{1+2} \det(a_{1,2}) = -a_{1,2},$$

und genauso  $c_{2,1} = -a_{2,1}$  und  $c_{2,2} = a_{1,1}$ . Damit ist nach Satz 18.21 (b) im Fall einer invertierbaren Matrix also

$$A^{-1} = \frac{1}{\det A} \cdot \begin{pmatrix} a_{2,2} & -a_{1,2} \\ -a_{2,1} & a_{1,1} \end{pmatrix}.$$

Eine konkrete Anwendung von Satz 18.21 ergibt sich bei der Lösung linearer Gleichungssysteme: Sind  $A \in GL(n, K)$  eine invertierbare Matrix und  $b \in K^n$ , so wissen wir bereits, dass das Gleichungssystem  $Ax = b$  für  $x$  die eindeutige Lösung  $x = A^{-1}b$  hat. Da wir gerade mit Hilfe von Determinanten eine explizite Formel für die inverse Matrix  $A^{-1}$  gefunden haben, überrascht es nicht, dass wir auch für die Koordinaten dieses Lösungsvektors  $x = A^{-1}b$  eine ähnliche explizite Formel herleiten können:

**Satz 18.23 (Cramersche Regel).** *Es seien  $A \in GL(n, K)$  und  $b \in K^n$ . Wir bezeichnen die Spalten von  $A$  mit  $a_1, \dots, a_n \in K^n$ . Dann ist die (nach Algorithmus 15.41 (a) eindeutige) Lösung des Gleichungssystems  $Ax = b$  der Vektor  $x \in K^n$  mit den Komponenten*

$$x_i = \frac{\det(a_1 | \cdots | a_{i-1} | b | a_{i+1} | \cdots | a_n)}{\det A}$$

*für  $i = 1, \dots, n$ .*

*Beweis.* Natürlich ist  $x = A^{-1}b$ . Nach Satz 18.21 (b) und Definition 15.5 der Matrixmultiplikation ist die  $i$ -te Komponente dieses Matrixprodukts gleich

$$x_i = \sum_{j=1}^n (-1)^{i+j} \frac{\det A'_{j,i}}{\det A} \cdot b_j \stackrel{18.15}{=} \frac{1}{\det A} \det(a_1 | \cdots | a_{i-1} | b | a_{i+1} | \cdots | a_n),$$

wobei  $b_j$  die  $j$ -te Komponente von  $b$  und die zweite Gleichheit die Entwicklung nach der  $i$ -ten Spalte ist.  $\square$

**Beispiel 18.24.** Wir wollen mit der Cramerschen Regel das lineare Gleichungssystem

$$\begin{array}{r} x_1 + x_2 = 2 \\ x_1 - 2x_2 = -1 \end{array} \quad \text{lösen, also} \quad \begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$

Dies ist sehr einfach: Es ist

$$x_1 = \frac{\det \begin{pmatrix} 2 & 1 \\ -1 & -2 \end{pmatrix}}{\det \begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix}} = \frac{-3}{-3} = 1 \quad \text{und} \quad x_2 = \frac{\det \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}}{\det \begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix}} = \frac{-3}{-3} = 1.$$

Beachte jedoch, dass es für konkrete Gleichungssysteme mit mehr als zwei Variablen sehr rechenaufwendig ist, die Cramersche Regel zu verwenden. Man wird diese Regel daher meistens nur für theoretische Überlegungen verwenden, in denen man eine konkrete Formel für die Lösung (und nicht nur ein Lösungsverfahren) braucht. Für numerische Berechnungen ist der Gauß-Algorithmus in Satz 15.33 wesentlich effizienter.

**Aufgabe 18.25.** Es sei  $A \in K^{n \times n}$  eine invertierbare Matrix mit ganzzahligen Einträgen.

Zeige, dass  $A^{-1}$  genau dann ebenfalls nur ganzzahlige Einträge hat, wenn  $\det A = \pm 1$  gilt.

## Literatur

- [B] A. Beutelspacher, *Lineare Algebra*, Vieweg-Verlag (2003)
- [BF] M. Barner, F. Flohr, *Analysis 1*, de Gruyter Lehrbuch (2000)
- [E] H.-D. Ebbinghaus et al., *Zahlen*, Springer-Verlag (1988)
- [Fi] G. Fischer, *Lineare Algebra*, Vieweg-Verlag (2002)
- [Fo1] O. Forster, *Analysis 1*, Vieweg-Verlag (2011)
- [Fo2] O. Forster, *Analysis 2*, Vieweg-Verlag (2010)
- [G] A. Gathmann, *Algebraische Strukturen*, Vorlesungsskript RPTU Kaiserslautern (2023),  
<https://agag-gathmann.math.rptu.de/ags>
- [GK] G.-M. Greuel, T. Keilen, *Lineare Algebra I und II*, Vorlesungsskript TU Kaiserslautern (1999),  
[www.math.uni-tuebingen.de/~keilen/download/LectureNotes/linearealgebra.pdf](http://www.math.uni-tuebingen.de/~keilen/download/LectureNotes/linearealgebra.pdf)
- [J] K. Jänich, *Lineare Algebra*, Springer-Verlag (2010)
- [K1] K. Königsberger, *Analysis 1*, Springer-Verlag (2003)
- [K2] K. Königsberger, *Analysis 2*, Springer-Verlag (2003)
- [M] T. Markwig, *Grundlagen der Mathematik*, Vorlesungsskript TU Kaiserslautern (2011),  
[www.math.uni-tuebingen.de/~keilen/download/LectureNotes/grundlagen11.pdf](http://www.math.uni-tuebingen.de/~keilen/download/LectureNotes/grundlagen11.pdf)

## Index

- $A^T$  56
- $A^{B,C}$  87
- $A_f$  82
- $A_f^{B,C}$  84
- Abb( $D, K$ ) 37
- Abb( $D, W$ ) 38
- Abbildung 15
  - alternierende 101
  - bijektive 17
  - identische 16
  - injektive 17
  - lineare 74
  - multilineare 101
  - surjektive 17
- Abbildungsmatrix 82, 84
- abelsche Gruppe 24
- Abgeschlossenheit
  - eines Unterraums 39
- Ableitung
  - eines Polynoms 76
- Äquivalenz
  - von Aussagen 7
  - von Matrizen 88
- Äquivalenzklasse 22
- Äquivalenzrelation 22
- affiner Unterraum 93
- alternierende Abbildung 101
- Assoziativität
  - der Matrixmultiplikation 57
  - der Skalarmultiplikation 35
  - der Verkettung 19
  - in Gruppen 24
- aufgespannter Unterraum 40
- Aussage 6
  - äquivalente 7
  - zusammengesetzte 7
- Aussageform 6
- Austauschlemma 47
- Austauschsatz von Steinitz 48
- Axiom 5
- Axiomensystem
  - von Zermelo und Fraenkel 12
- Basis 44, 53
- Basisauswahl 47, 71
- Basisergänzung 48, 71
- Basiswechselmatrix 87
- bijektive Abbildung 17
- Bild
  - einer Matrix 58, 66, 70
  - einer Menge 18
  - eines Elements 15
  - eines Morphismus 77
- Blockmatrixmultiplikation 57
- Cantor 11
- Cramersche Regel 109
- Definitionsmenge 15
- $\deg f$  33
- Determinante
  - Eindeutigkeit 103
  - einer Matrix 100
  - Existenz 104
- Determinantenproduktsatz 103
- Diagonale
  - einer Matrix 55
- Diagonaleintrag 55
- Differenzmenge 13
- $\dim V$  49
- Dimension
  - eines Vektorraums 49
- Dimensionsformel
  - für direkte Summen 91
  - für Durchschnitte 51
  - für Komplemente 92
  - für Matrizen 68
  - für Morphismen 85
  - für Produkte 81
  - für Quotientenräume 95
  - für Summen 51
- direkte Summe 90
- disjunkte Mengen 13
- disjunkte Vereinigung 13
- Distributivität 26
  - der Matrixmultiplikation 57
  - der Skalarmultiplikation 35
- Dreiecke
  - kongruente 21
- Dreiecksmatrix
  - echte 108
  - obere 108
  - untere 108
- E 60
- $e_i$  43
- $E_n$  60
- echte Dreiecksmatrix 108
- Einheitsmatrix 60
- Einheitsvektor 43
- Einschränkung 16
- Element
  - einer Familie 39
  - einer Menge 11
  - inverses 24
  - linksinverses 24
  - linksneutrales 24
  - neutrales 24
- elementare Spaltenumformung 63, 70
- elementare Zeilenumformung 63
- Elementarmatrix 63
- endlich erzeugter Vektorraum 43
- endlich-dimensionaler Vektorraum 49
- endliche Menge 12
- Entwicklungssatz von Laplace 107

- erweiterte Koeffizientenmatrix 67
- Erzeugendensystem 43, 53
- erzeugter Unterraum 40
- $f_A$  75
- $f_A^{B,C}$  83
- $F_k(\lambda)$  63
- $F_{k,l}(\lambda)$  63
- Faktorraum 95
- Familie
  - linear abhängige 44, 71
  - linear unabhängige 44, 53
  - von Vektoren 39
- Fibonacci-Folge 54
- Folge
  - von Fibonacci 54
- Fraenkel 12
- Funktion 15
  - alternierende 101
  - lineare 74
  - multilineare 101
- Funktionswert 15
- ganze Zahl 13
- Gauß
  - Algorithmus von 64
  - Summenformel von 30
- geordnetes Paar 13
- $GL(n, K)$  60
- Gleichungssystem
  - lineares 55, 68
- Grad
  - einer Polynomfunktion 31, 33
- Graph 16
- Gruppe 24
  - abelsche 24
  - kommutative 24
- Gruppenaxiome 24
- $\text{Hom}(V, W)$  74
- Homomorphiesatz 97
- Homomorphismus 74
- identische Abbildung 16
- $\text{Im } f$  77
- Indexverschiebung 29
- Induktion 30
- Induktionsanfang 30
- Induktionsannahme 30
- Induktionsschluss 30
- Induktionsschritt 30
- Induktionsvoraussetzung 30
- injektive Abbildung 17
- inverse Matrix 60, 69, 109
- inverses Element 24
- invertierbare Matrix 60
- isomorph 78
- Isomorphismus 78
- $K^{m \times n}$  55
- $\text{Ker } f$  77
- Kern
  - einer Matrix 58
  - eines Morphismus 77
- Klasse 22
- Koeffizient
  - einer Polynomfunktion 31
- Koeffizientenmatrix
  - erweiterte 67
- Koeffizientenvergleich
  - für Polynomfunktionen 33
- Körper 26
- Körperaxiome 26
- kommutative Gruppe 24
- Kommutativität 24
- Komplement
  - eines Unterraums 92
- komplementärer Unterraum 92
- Kongruenz 21
- Kongruenzklasse 21
- Kontraposition 9
- Koordinaten
  - eines Vektors 46
- Koordinatenabbildung 80
- Koordinatenvektor 80
- Laplacescher Entwicklungssatz 107
- leere Summe 29
- leeres Produkt 29
- Leitkoeffizient
  - einer Polynomfunktion 31
- Lemma 19
- $\text{Lin } B$  39, 53
- linear abhängig 44, 71
- linear unabhängig 44
- lineare Abbildung 74
- lineares Gleichungssystem 55, 68
- lineares Polynom 33
- Linearkombination 39, 53
  - nicht-triviale 44
- linksinverses Element 24
- linksneutrales Element 24
- Matrix 55
  - äquivalente 88
  - einer linearen Abbildung 82, 84
  - eines Basiswechsels 87
  - inverse 60, 69, 109
  - invertierbare 60
  - quadratische 55
  - transponierte 56
- Matrixmultiplikation 56
- Menge 11
  - endliche 12
  - leere 11
- modulo 95
- Morphismus 74
- multilineare Abbildung 101
- Multiplikation
  - von Matrizen 56
- Multiplizität
  - einer Nullstelle 34
- $\mathbb{N}$  12
- natürliche Zahl 12
- Negation 9
- neutrales Element 24
- nicht-triviale Linearkombination 44
- Normalform

- bezüglich Äquivalenz 99
  - einer Abbildungsmatrix 98
- normierte Polynomfunktion 31
- Nullmatrix 56
- Nullstelle 31
- Nullvektor 36
- Nullvektorraum 36
- obere Dreiecksmatrix 108
- Obermenge 11
- Ordnung
  - einer Nullstelle 34
- $\Phi_B$  80
- Paar
  - geordnetes 13
- Paradoxon
  - von Russell 12
- Partition
  - einer Menge 23
- $\text{Pol}(D, \mathbb{R})$  41
- $\text{Pol}_n(D, \mathbb{R})$  41
- Polynom 33
  - lineares 33
  - quadratisches 33
- Polynomdivision 32
- Polynomfunktion 21
  - normierte 31
- Potenz 28
- Potenzmenge 13
- Produkt
  - leeres 29
- Produktmenge 13
- Produktsatz
  - für Determinanten 103
- Produktzeichen 29
- $\mathbb{Q}$  13
- quadratische Matrix 55
- quadratisches Polynom 33
- Quantor 8
- Quotientenraum 95
- $\mathbb{R}$  12
- Rang
  - einer Matrix 58, 66
  - eines Morphismus 77
- rationale Zahl 13
- reduzierte Zeilenstufenform 64
- reelle Zahl 12
- Reflexivität
  - einer Relation 22
- Regel
  - von Cramer 109
  - von Sarrus 105
- Relation 15
- Repräsentant
  - einer Äquivalenzklasse 22
- Restklasse 95
- $\text{rk } f$  77
- Russell 12
- Russellsches Paradoxon 12
- Sarrus
  - Regel von 105
- Satz
  - von Laplace 107
  - von Steinitz 48
- Schnittmenge 13
- Skalar 36
- Skalarmultiplikation 35
- Spaltenstufenform 70
- Spaltenumformung 63, 70
- Standardbasis 45
- Startmenge 15
- Startraum 15
- Steinitzscher Austauschatz 48
- Streichungsmatrix 104
- Stufenspalte 64
- Summe
  - direkte 90
  - leere 29
  - von Unterräumen 41
- Summenformel
  - von Gauß 30
- Summenzeichen 28
- surjektive Abbildung 17
- Symmetrie
  - einer Relation 22
- Teilmenge 11
  - echte 12
- Transitivität
  - einer Relation 22
- transponierte Matrix 56
- trivialer Unterraum 41
- Umkehrabbildung 19
- Umkehrfunktion 19
- untere Dreiecksmatrix 108
- Unterraum 39
  - affiner 93
  - aufgespannter 40
  - erzeugter 40
  - komplementärer 92
  - trivialer 41
  - verschobener 93
- Untervektorraum 39
  - affiner 93
  - aufgespannter 40
  - erzeugter 40
  - komplementärer 92
  - trivialer 41
  - verschobener 93
- Urbild
  - einer Menge 18
  - eines Elements 17
- Variable 6
- Vektor 36
- Vektoraddition 35
- Vektorraum 35
  - endlich erzeugter 43
  - endlich-dimensionaler 49
  - isomorpher 78
- Vektorraumhomomorphismus 74
- Vektorraumisomorphismus 78
- Vereinigung 13
  - disjunkte 13

Vereinigungsmenge 13  
Verkettung 19  
Verknüpfung 24  
Verneinung 9  
verschobener Unterraum 93  
Vielfachheit  
  einer Nullstelle 34  
vollständige Induktion 30  
  
Wahrheitstafel 7  
Wert  
  einer Funktion 15  
Widerspruchsbeweis 9  
Wohldefiniertheit 95  
  
 $\mathbb{Z}$  13  
Zahl  
  ganze 13  
  natürliche 12  
  rationale 13  
  reelle 12  
Zeilenstufenform 64  
  reduzierte 64  
Zeilenumformung 63  
Zermelo 12  
Zielmenge 15  
Zielraum 15