Lineare Algebra I
Wintersemester 2021/22
Universität Regensburg

Marc Hoyois

30. Januar 2022

Inhaltsverzeichnis

Einführung

Diese Vorlesung hat zwei allgemeine Ziele:

- Das erste Ziel ist natürlich die lineare Algebra kennezulernen. Die lineare Algebra ist ein besonderer Bereich der Mathematik, indem sie in praktisch allen anderen mathematischen Bereichen verwendet wird, von der Analysis bis zu der Geometrie. Das steht zum Beispiel im Gegensatz zu der Analysis und der Geometrie, die nur für einen (wenn auch goßen) Teil der Mathematik relevant sind. Die lineare Algebra ist auch unerlässlich in der theoretischen Physik: Die beiden grundlegendsten Theorien der Physik, nämlich die Allgemeine Relativitätstheorie und die Quantenfeldtheorie, benutzen mehrere fortgeschrittene Begriffe aus der linearen Algebra. Weiter hat die lineare Algebra viele verschiedene Anwendungen in der heutigen Zeit, zum Beispiel für numerische Simulationen, Suchalgorithmen, maschinelles Lernen, usw.

- Das zweite Ziel, das sich auch mit der Vorlesung *Analysis* deckt, ist den Begriff des *mathematischen Beweises* kennenzulernern. Insbesondere werden Sie durch die Praxis lernen, was als mathematischer Beweis zählt, solche Beweise zu verstehen und selbst zu schreiben, sowie „einfache“ Beweise selbst herauszufinden.

Obwohl wir die Mathematik auf Deutsch besprechen, ist die mathematische Sprache ganz besonders, und man muss sich daran gewöhnen. Der wichtigste Unterschied zwischen der natürlichen Sprache und der mathematischen Sprache ist, dass in der mathematischen Sprache keine Mehrdeutigkeiten erlaubt sind: Jede Aussage muss eine ganz eindeutige Bedeutung haben, unabhängig von irgendeiner Auslegung.

Mathematische Texte wie dieses Skript bestehend aus folgenden Bausteinen:

- **Definitionen** führen neue Begriffe ein. Definitionen sind besonders wichtig in der Mathematik, denn sie notwendig sind, um den Rest des Textes überhaupt zu verstehen. Deswegen sollen Sie unbedingt alle Definitionen auswendig lernen.

- Es gibt verschiedene Arten von Aussagen:

  - **Sätze** sind Aussagen, die besonders wichtig oder schwierig sind.
  - **Propositionen** sind Aussagen, die nicht besonders schwierig sind.
  - **Lemmata** sind Hilfssätze, die in Beweisen weiterer Sätze verwendet werden.
  - **Korollare** sind Folgerungen vorheriger Sätze.

- Jede solche Aussage soll bewiesen werden, durch einen **Beweis**.

- Es gibt noch **Beispiele** und **Bemerkungen**, die auch wichtig sind.

In diesem Skript gibt es einige Sätze, die nicht bewiesen werden. Diesen Sätzen ist ein Stern vorangestellt: ***Satz**. Es gibt auch ein paar Sätze, die bewiesen werden, aber deren Beweise besonders kompliziert sind. Diesen Beweisen ist dann ein Stern vorangestellt: ***Beweis**. Solche Beweise müssen Sie nicht unbedingt lesen, aber Sie können es als Herausforderung nehmen, sie zu verstehen.

Überblick über die Vorlesung

In den Vorlesungen *Lineare Algebra I* und *II* werden wir folgende Themen studieren:

- Mengentheoretische Grundlagen: Mengen, Abbildungen und Relationen
- Grundlegende algebraische Strukturen: Gruppen und Körper
- Vektorräume
- Lineare Abbildungen
- Matrizen, Matrizenkalkül und die Determinante
- Lineare Gleichungssysteme
- Eigenwerte und Eigenvektoren
- Euklidische und unitäre Vektorräume
- Ringe und Moduln
- Endlich erzeugte Moduln über Hauptidealringen
- Normalformen linearer Abbildungen

Kapitel 1

Mengentheoretische Grundlagen

1.1 Logische Grundlagen

Eine vereinfachte Sichtweise der Mathematik ist, dass sie sich mit der Bestimmung der Wahrheit bzw. Falschheit von objektiven Aussagen beschäftigt. Beispiele von objektiven Aussagen sind „$1 + 1 = 3$“, „drei Punkte im Raum liegen gemeinsam auf mindestens einer Ebene“ und „jede gerade Zahl größer als 2 ist die Summe zweier Primzahlen“. Natürlich ist die erste Aussage falsch und die zweite wahr; die Wahrheit der dritten Aussage ist derzeit nicht bekannt.

Die mathematische Logik beschäftigt sich damit, grundlegende Begriffe wie „Aussage“ und „Wahrheit“ präzis zu definieren. In dieser Vorlesung versuchen wir nicht zu erläutern, was genau eine mathematische Aussage ist. Es gibt „atomare Aussagen“, die nicht aus kleineren Aussagen aufgebaut werden, z.B. „$1 + 1 = 3$“ oder „der Punkt P liegt auf der Ebene E“. Aus atomaren Aussagen können wir weitere Aussagen auf verschiedene Weise aufbauen:

- *Negation.* Jede Aussage φ besitzt eine gegenteilige Aussage „nicht φ“.
- *Logische Verknüpfungen.* Zu je zwei Aussagen φ und ψ kann man unter anderem die folgenden Aussagen bilden: „φ und ψ“, „φ oder ψ“, „φ impliziert ψ“ und „φ ist äquivalent zu ψ“.
- *Quantifizierung.* Wenn eine Aussage $\varphi(x)$ von einer Variablen x abhängt, dann kann man über diese Variable *quantifizieren*, um neue Aussagen zu erhalten: „für alle x gilt $\varphi(x)$“ (Allaussage) und „es existiert (mindestens) ein x, für das $\varphi(x)$ gilt“ (Existenzaussage).

In der folgenden Tabelle werden diese logischen Bausteine zusammengefasst:

symbolische Aussage	Bedeutung	Name	äquivalente Aussagen
$\neg\varphi$	nicht φ	Negation	
$\varphi \wedge \psi$	φ und ψ	Konjunktion	$\neg(\neg\varphi \vee \neg\psi)$
$\varphi \vee \psi$	φ oder ψ	Disjunktion	$\neg(\neg\varphi \wedge \neg\psi)$
$\varphi \Rightarrow \psi$	φ impliziert ψ, wenn φ, dann ψ, aus φ folgt ψ	Implikation	$\neg\varphi \vee \psi$
$\varphi \Leftrightarrow \psi$	φ ist äquivalent zu ψ, φ gilt genau dann, wenn ψ	Äquivalenz	$(\varphi \Rightarrow \psi) \wedge (\psi \Rightarrow \varphi)$
$\forall x\, \varphi(x)$	für alle/jedes x gilt $\varphi(x)$	Allaussage	$\neg\exists x\, \neg\varphi(x)$
$\exists x\, \varphi(x)$	es existiert/gibt ein x mit $\varphi(x)$	Existenzaussage	$\neg\forall x\, \neg\varphi(x)$

Die Symbole $\forall$ und $\exists$ heißen der *Allquantor* und der *Existenzquantor*. Man kann sich die Quantoren $\forall$ und $\exists$ als unendliche Verallgemeinerungen der logischen Verknüpfungen $\wedge$ und $\vee$ vorstellen. Manchmal verwendet man auch die Notation $\exists! x$ mit der Bedeutung „es existiert *genau ein* x“. Eigentlich ist $\exists! x\, \varphi(x)$ eine Abkürzung der Aussage

$$\exists x\, \varphi(x) \wedge \forall x \forall y((\varphi(x) \wedge \varphi(y)) \Rightarrow x = y).$$

Die Aussage $\forall x \forall y((\varphi(x) \wedge \varphi(y)) \Rightarrow x = y)$ heißt *Eindeutigkeitsaussage* für x in $\varphi(x)$: Sie bedeutet, dass *höchstens ein* x mit $\varphi(x)$ existiert.

Bemerkung 1.1.1. In mathematischen Texten (außer in der mathematischen Logik) schreibt man die logischen Symbole $\neg, \wedge, \vee, \forall, \exists$ sehr selten: sie werden eher auf Deutsch ausgeschrieben. Im weiteren Verlauf dieses Skriptes werden wir also diese Symbole nicht benutzen. Die Symbole $\forall$ und $\exists$ sind trotzdem nützlich, wenn man mit der Hand schreibt.

Bemerkung 1.1.2 (Reihenfolge der Quantoren)**.** Viele mathematische Aussagen fangen mit mehreren Quantoren an. Bei denen muss man beachten, dass die Reihenfolge verschiedener Quantoren wichtig ist. Sei zum Beispiel $\varphi(x, y)$ die Aussage „wenn x ein Punkt im Raum ist, dann ist y eine Gerade im Raum, auf der x liegt“. Die Aussage

$$\forall x \exists y\, \varphi(x, y)$$

bedeutet, dass jeder Punkt im Raum auf mindestens einer Gerade liegt (was wahr ist), und die Aussage

$$\exists y \forall x\, \varphi(x, y)$$

bedeutet, dass alle Punkte im Raum auf derselben Gerade liegen (was falsch ist). Für eine beliebige Aussage $\varphi(x, y)$ gilt die Implikation

$$\exists y \forall x\, \varphi(x, y) \Rightarrow \forall x \exists y\, \varphi(x, y),$$

aber nicht unbedingt die umgekehrte Implikation.

Wie kann man entscheiden, ob eine gegebene Aussage wahr oder falsch ist? Der Wahrheitswert der Aussagen $\neg\varphi$, $\varphi \wedge \psi$, $\varphi \vee \psi$, $\varphi \Rightarrow \psi$ und $\varphi \Leftrightarrow \psi$ kann man einfach bestimmen, wenn die Wahrheitswerte von φ und ψ bekannt sind. Dazu verwendet man die folgenden Wahrheitstabellen:

φ	ψ	$\neg\varphi$	$\varphi \wedge \psi$	$\varphi \vee \psi$	$\varphi \Rightarrow \psi$	$\varphi \Leftrightarrow \psi$
w	w	f	w	w	w	w
w	f	f	f	w	f	f
f	w	w	f	w	w	f
f	f	w	f	f	w	w

Bemerkung 1.1.3. Bei den Wahrheitstabellen der Disjunktion $\vee$ und der Implikation $\Rightarrow$ muss man folgendes beachten:

- Wenn beide φ und ψ wahr sind, dann ist „φ oder ψ“ auch wahr. In der Mathematik ist „oder“ nie exklusiv, d.h., es bedeutet nicht „entweder φ oder ψ“, sondern „φ oder ψ oder beide“. Mit den obigen Symbolen kann das exklusive Oder als $\neg(\varphi \Leftrightarrow \psi)$ geschrieben werden.
- Wenn φ falsch ist, dann ist „φ impliziert ψ“ immer *wahr*. In der Mathematik ist die Implikation immer so verstanden. Anders gesagt: Aus etwas Falschem folgt alles.

Bemerkung 1.1.4. Die obige Liste von logischen Verknüpfungen ist nicht erschöpfend. Andere Beispiele sind das exklusive Oder $\neg(\varphi \Leftrightarrow \psi)$ und die umgekehrte Implikation $\varphi \Leftarrow \psi$. Man kann jedoch zeigen, dass alle möglichen logischen Verknüpfungen (d.h., alle Wahrheitstabellen) als Kombinationen von $\neg$ und $\vee$ erhalten werden können.

Definition 1.1.5 (Tautologie)**.** Eine Aussage heißt *Tautologie*, wenn sie aus Aussagen $\varphi_1, \ldots, \varphi_n$ und den logischen Verknüpfungen $\neg, \wedge, \vee, \Rightarrow, \Leftrightarrow$ aufgebaut ist, so dass sich unter allen möglichen w/f-Belegungen der Aussagen φ_i der Wert w ergibt (gemäß den obigen Wahrheitstabellen).

Eine wichtige Eigenschaft des Begriffs der Tautologie ist seine *Berechenbarkeit*: Es ist immer möglich, automatisch und in endlich vielen Schritten nachzuprüfen, ob eine Aussage eine Tautologie ist oder nicht.

Beispiel 1.1.6 (Wichtige Tautologien)**.** Aussagen folgender Gestalt sind Tautologien:

(i) $\varphi \vee \neg\varphi$ (Satz vom ausgeschlossenen Dritten)

(ii) $\varphi \Leftrightarrow \neg\neg\varphi$ (Gesetz der doppelten Negation)

(iii) $\neg(\varphi \wedge \neg\varphi)$ (Satz vom ausgeschlossenen Widerspruch)

(iv) $(\varphi \Rightarrow \psi) \Leftrightarrow (\neg\psi \Rightarrow \neg\varphi)$ (Gesetz der Kontraposition)

(v) $(\varphi \wedge \neg\varphi) \Rightarrow \psi$ (Ex falso quodlibet/„aus Falschem [folgt] Beliebiges“)

(vi) $(\neg\varphi \Rightarrow (\psi \wedge \neg\psi)) \Rightarrow \varphi$ (Reductio ad absurdum/Widerspruchsbeweis)

(vii) $((\varphi \Rightarrow \psi) \wedge (\psi \Rightarrow \chi)) \Rightarrow (\varphi \Rightarrow \chi)$ (Syllogismus)

Als Beispiel überprüfen wir, dass das Gesetz der Kontraposition eine Tautologie ist:

φ	ψ	$\neg\varphi$	$\neg\psi$	$\varphi \Rightarrow \psi$	$\neg\psi \Rightarrow \neg\varphi$	$(\varphi \Rightarrow \psi) \Leftrightarrow (\neg\psi \Rightarrow \neg\varphi)$
w	w	f	f	w	w	w
w	f	f	w	f	f	w
f	w	w	f	w	w	w
f	f	w	w	w	w	w

Diese Tabelle zeigt, dass der Wahrheitswert von (iv) immer w ist, unabhängig von den Wahrheitswerten von φ und ψ. Das ist genau die Definition einer Tautologie.

Der Begriff der Wahrheit erweist sich als ziemlich subtil, wenn wir auch quantifizierte Aussagen betrachten wollen. Eigentlich sind wir in der Mathematik eher an dem konkreteren Begriff der *Beweisbarkeit* interessiert. Um den präzis zu definieren, benötigen wir *Axiome* und *Schlussregeln*. Axiome sind ausgewählte Aussagen, die als „wahr“ angenommen werden, und Schlussregeln schreiben vor, wie man Aussagen aus anderen Aussagen ableiten kann. Ein *Beweis* ist dann eine endliche Folge von Aussagen, in der jede Aussage entweder ein Axiom ist oder durch eine Schlussregel aus vorherigen Aussagen folgt. Eine Aussage φ heißt *beweisbar*, wenn ein Beweis existiert, dessen letzte Aussage φ ist.

Wir beschreiben jetzt ein solches System von Axiomen und Schlussregeln, das für die ganze Mathematik geeignet ist: Es ist die sogenannte *Prädikatenlogik erster Stufe mit Gleichheit*. Dies dient nur der Veranschaulichung, und es ist gar nicht wichtig, sich diese Axiome und Schlussregeln einzuprägen. Auf jeden Fall ist die folgende Beschreibung unvollständig, weil wir unter anderem nicht erläutert haben, was die atomaren Aussagen sind und was die „Variablen“ $x, y, \ldots$ sind.

- Die *aussagenlogischen Axiome* sind alle Tautologien.
- Die *quantorlogischen Axiome* sind:
  - $\forall x\, \varphi(x) \Rightarrow \varphi(a)$.
  - Falls x in ψ nicht vorkommt, $\forall x(\psi \Rightarrow \varphi(x)) \Rightarrow (\psi \Rightarrow \forall x\, \varphi(x))$.

- Die *Gleichheitsaxiome* sind:
  - $x = x$ (das *Identitätsaxiom*).
  - $x = y \Rightarrow (\varphi(x) \Leftrightarrow \varphi(y))$.
- Es gibt nur zwei Schlussregeln:
  - *Modus Ponens*: Aus φ und $\varphi \Rightarrow \psi$ kann man ψ herleiten.
  - *Allquantoreinführung*: Aus $\varphi(x)$ kann man $\forall x\, \varphi(x)$ herleiten.

In diesem System wird die Existenzaussage $\exists x\, \varphi(x)$ als Abkürzung von $\neg\forall x\, \neg\varphi(x)$ definiert, und sie erfordert keine weiteren Axiome.

Bemerkung 1.1.7. Die Schlussregel der Allquantoreinführung sieht vielleicht ein bisschen merkwürdig aus. Man soll sie wie folgt verstehen: Wenn man in diesem System eine Aussage $\varphi(x)$ mit einer freien Variablen x beweisen kann, bedeutet das, dass diese Aussage für ein *beliebiges* x gilt. Die selbständige unquantifizierte Aussage $\varphi(x)$ hat also dieselbe Bedeutung wie die Allaussage $\forall x\, \varphi(x)$.

Beispiel 1.1.8. Als Beispiel geben wir einen formalen Beweis der Aussage

$$\forall x \forall y (x \neq x \Rightarrow x = y),$$

wobei $x \neq x$ eine Abkürzung von $\neg(x = x)$ ist:

1. $x = x$	(Identitätsaxiom)
2. $x = x \Rightarrow (x \neq x \Rightarrow x = y)$	(Tautologie)
3. $x \neq x \Rightarrow x = y$	(Modus Ponens aus 1 und 2)
4. $\forall y (x \neq x \Rightarrow x = y)$	(Allquantoreinführung aus 3)
5. $\forall x \forall y (x \neq x \Rightarrow x = y)$	(Allquantoreinführung aus 4).

Beispiel 1.1.9. Ein berühmter Syllogismus lautet:

> Alle Menschen sind sterblich.
> Sokrates ist ein Mensch.
> Also ist Sokrates sterblich.

Sei $\mu(x)$ die Aussage „x ist ein Mensch“ und sei $\sigma(x)$ die Aussage „x ist sterblich“. In der Prädikatenlogik erster Stufe sieht dieser Syllogismus wie folgt aus:

1. $\forall x (\mu(x) \Rightarrow \sigma(x))$	(1. Annahme)
2. $\mu(\text{Sokrates})$	(2. Annahme)
3. $\forall x (\mu(x) \Rightarrow \sigma(x)) \Rightarrow (\mu(\text{Sokrates}) \Rightarrow \sigma(\text{Sokrates}))$	(quantorlogisches Axiom)
4. $\mu(\text{Sokrates}) \Rightarrow \sigma(\text{Sokrates})$	(Modus Ponens aus 1 und 3)
5. $\sigma(\text{Sokrates})$	(Modus Ponens aus 2 und 4).

Beispiel 1.1.10. Für den Exiztenzquantor $\exists$ gelten folgenden Aussagen:

- $\varphi(a) \Rightarrow \exists x\, \varphi(x)$
- Falls x in ψ nicht vorkommt, $\forall x (\varphi(x) \Rightarrow \psi) \Rightarrow (\exists x\, \varphi(x) \Rightarrow \psi)$

Diese Aussagen folgen aus den entsprechenden quantorlogischen Axiomen. Hier ist zum Beispiel ein Beweis der ersten Aussage:

1. $\forall x\, \neg\varphi(x) \Rightarrow \neg\varphi(a)$	(quantorlogisches Axiom)
2. $(\forall x\, \neg\varphi(x) \Rightarrow \neg\varphi(a)) \Rightarrow (\varphi(a) \Rightarrow \neg\forall x\, \neg\varphi(x))$	(Tautologie)
3. $\varphi(a) \Rightarrow \neg\forall x\, \neg\varphi(x)$	(Modus Ponens aus 1 und 2).

1.1.1 Beispiele von Beweisen

Zum Aufwärmen besprechen wir ein paar Beweise von Elementarsätzen in der Algebra: dem Satz von Euklid, dass unendlich viele Primzahlen existieren, und dem „Satz der Pythagoreer“, dass $\sqrt{2}$ (die Länge der Diagonale in einem Einheitsquadrat) keine rationale Zahl ist. Dazu werden wir einige Begriffe benutzen, die später in der Vorlesung präziser eingeführt werden.

Die *natürlichen Zahlen* sind die Zahlen 0, 1, 2, 3, und so weiter (siehe Abschnitt 1.2.1 für eine mengentheoretische Definition). Sind m und n natürliche Zahlen, so sagt man „m teilt n“ oder „n ist durch m teilbar“, wenn eine natürliche Zahl r mit $n = m \cdot r$ existiert. Eine *Primzahl* ist eine natürliche Zahl, die nicht gleich 1 ist und die nur durch 1 und sich selbst teilbar ist. Die kleinste Primzahl ist 2, dann kommen 3, 5, 7, 11, usw. (Die natürliche Zahl 0 ist durch jede andere natürliche Zahl teilbar, und damit keine Primzahl.)

Satz 1.1.11 (Euklid)**.** *Es gibt unendlich viele Primzahlen.*

Um diesen Satz zu beweisen brauchen wir das folgende Lemma:

Lemma 1.1.12. *Sei $n \geq 2$ eine natürliche Zahl. Dann existiert eine Primzahl p, die n teilt.*

Beweis. Wir verwenden das Prinzip der vollständigen Induktion (Korollar 1.2.23). Damit dürfen wir annehmen, dass jede natürliche Zahl m mit $2 \leq m < n$ durch eine Primzahl teilbar ist (diese Aussage heißt die *Induktionsvoraussetzung*). Wenn n schon eine Primzahl ist, können wir $p = n$ nehmen (da n durch sich selbst teilbar ist). Wir nehmen jetzt an, dass n keine Primzahl ist. Nach Definition von Primzahl existiert dann eine natürliche Zahl $m \neq 1, n$, die n teilt. Das heißt, es gilt $n = m \cdot r$ mit einer natürichen Zahl r. Da $n \neq 0, m$, ist $r \geq 2$ und damit ist $m < n$. Nach der Induktionsvoraussetzung existiert eine Primzahl p, die m teilt. Da n durch m teilbar ist und m durch p teilbar ist, ist auch n durch p teilbar, wie gewünscht. □

Bemerkung 1.1.13. Lemma 1.1.12 hat die folgende logische Gestalt:

$$\forall n(\varphi(n) \Rightarrow \exists p(\psi(p) \wedge \chi(p, n))),$$

wobei $\varphi(n)$, $\psi(p)$ und $\chi(p, n)$ sind die Aussagen „n ist eine natürliche Zahl und $n \geq 2$“, „p ist eine Primzahl“ und „p teilt n“.

Beweis vom Satz 1.1.11. Wir verwenden einen Widerspruchsbeweis. Angenommen, es gibt nur endlich viele Primzahlen $p_1, p_2, \ldots, p_k$. Sei

$$q = p_1 \cdot p_2 \cdot \ldots \cdot p_k$$

das Produkt aller Primzahlen und sei $n = q + 1$. Nach dem Lemma 1.1.12 existiert eine Primzahl p, die n teilt. Es existiert also eine natürliche Zahl m mit $n = p \cdot m$. Da q das Produkt aller Primzahlen ist, ist q durch p teilbar: Es existiert eine natürliche Zahl r mit $q = p \cdot r$. Dann

$$1 = n - q = p \cdot m - p \cdot r = p \cdot (m - r).$$

Insbesondere ist 1 durch p teilbar, und damit ist $p = 1$. Aber $p \neq 1$ nach Definition einer Primzahl, was ein Widerspruch ist. □

Der Beweis vom Lemma 1.1.12 liefert auch die folgende stärkere Aussage: Jede natürliche Zahl $n \geq 2$ ist ein Produkt von Primzahlen (das gilt auch für $n = 1$, wenn man 1 als das leere Produkt betrachtet). Denn im Beweis sind beide m und r Produkte von Primzahlen nach der Induktionsvoraussetzung, und deshalb ist auch $n = m \cdot r$ ein Produkt von Primzahlen. Diese stärkere Aussage ist die Existenzaussage im folgenden wichtigen Satz; die Eindeutigkeitsaussage ist schwieriger und wird in der Vorlesung *Lineare Algebra II* besprochen:

***Satz 1.1.14** (Fundamentalsatz der Arithmetik)**.** *Jede natürliche Zahl $n \geq 1$ läßt sich eindeutig als Produkt von Primzahlen darstellen. Genauer existiert zu jeder Primzahl p genau eine natürliche Zahl $v_p(n)$, so dass $v_p(n) \neq 0$ nur für endlich viele Primzahlen p und*

$$n = \prod_{p \text{ Primzahl}} p^{v_p(n)}.$$

Die rechte Seite der obigen Gleichung ist eine Abkürzung des Produkts

$$p_1^{v_{p_1}(n)} \cdot p_2^{v_{p_2}(n)} \cdot p_3^{v_{p_3}(n)} \cdot \ldots,$$

wobei $p_1, p_2, p_3, \ldots$ alle Primzahlen in aufsteigender Reihenfolge sind. Das ist also ein Produkt unendlich vieler Zahlen (nach dem Satz von Euklid), aber es ist trotzdem sinnvoll, da nur endlich viele dieser Zahlen nicht gleich 1 sind. Zum Beispiel:

$$\begin{aligned} 10 &= 2 \cdot 5, \\ 56 &= 2^3 \cdot 7, \\ 60 &= 2^2 \cdot 3 \cdot 5. \end{aligned}$$

Die im obigen Sinne eindeutige Darstellung von n als Produkt von Primzahlen heißt die *Primfaktorzerlegung* von n. Der Exponent $v_p(n)$ ist die *Vielfachheit* der Primzahl p in n.

Wir wenden uns nun dem Thema der Irrationalität von $\sqrt{2}$ zu. Die *rationalen Zahlen* sind reellen Zahlen, die als Bruch a/b zweier ganzen Zahlen a, b mit $b \neq 0$ dargestellt werden können. In einer solchen Bruchdarstellung kann man immer annehmen, dass a und b *teilerfremd* sind, d.h., dass 1 die einzige natürliche Zahl ist, die beide a und b teilt (sonst kann man a und b durch einen gemeinsamen Teiler dividieren, ohne die Zahl a/b zu verändern).

In den reellen Zahlen besitzt jede Zahl $r \geq 0$ eine Quadratwurzel $\sqrt{r}$, die die einzige ≥ 0 reelle Zahl ist, deren Quadrat gleich r ist. Der folgende Satz bedeutet, dass die reelle Zahl $\sqrt{2}$ keine rationale Zahl ist.

Satz 1.1.15. *Es gibt keine rationale Zahl x mit $x^2 = 2$.*

Um diesen Satz zu beweisen brauchen wir wieder ein Lemma. Eine natürliche Zahl heißt bekanntlich *gerade*, wenn sie durch 2 teilbar ist.

Lemma 1.1.16. *Eine natürliche Zahle n ist genau dann gerade, wenn ihr Quadrat n^2 gerade ist.*

Beweis. Das ist eine Aussage der Gestalt $\varphi \Leftrightarrow \psi$. Um sie zu beweisen, brauchen wir beide Implikationen $\varphi \Rightarrow \psi$ und $\varphi \Leftarrow \psi$ zu beweisen.

Zu $\Rightarrow$. Sei n gerade. Das heißt, es existiert eine natürliche Zahl m, so dass $n = 2m$. Man berechnet:

$$n^2 = (2m)^2 = 2^2 m^2 = 2(2m^2).$$

Also ist n^2 auch gerade.

Zu $\Leftarrow$. Wir verwenden das Gesetz der Kontraposition: Um eine Aussage der Form $\varphi \Rightarrow \psi$ zu beweisen, genügt es die Kontraposition $\neg\psi \Rightarrow \neg\varphi$ zu beweisen. Es genügt also zu zeigen, dass n^2 *un*gerade ist, wenn n *un*gerade ist. Die Zahl $n - 1$ ist dann gerade, und hat somit die Form $2m$. Es gilt dann $n = 2m + 1$ und man berechnet:

$$\begin{aligned} n^2 = (2m+1)^2 &= (2m+1)(2m+1) = 2m(2m+1) + 1(2m+1) \\ &= 4m^2 + 2m + 2m + 1 = 4m^2 + 4m + 1 = 2(2m^2 + 2m) + 1. \end{aligned}$$

Da $2(2m^2 + 2m)$ gerade ist, ist n^2 ungerade. □

Beweis vom Satz 1.1.15. (Widerspruchsbeweis.) Angenommen, es existiert eine rationale Zahl x mit $x^2 = 2$. Nach Definition der rationalen Zahlen läßt sich x als Bruch $\pm a/b$ mit natürlichen Zahlen a, b darstellen, wobei a und b teilerfremd sind. Es gilt dann $(a/b)^2 = 2$, also $a^2 = 2b^2$. Insbesondere sind a^2 und daher a gerade, nach Lemma 1.1.16. Es gilt also $a = 2n$ mit einer natürlichen Zahl n. Dann ist $2b^2 = 4n^2$, also $b^2 = 2n^2$, und damit ist b^2 gerade. Nach Lemma 1.1.16 ist auch b gerade. Dass a und b beide gerade sind, steht im Widerspruch zur Annahme, dass a und b teilerfremd sind. □

1.2 Mengen

Ohne Axiome kann man nichts beweisen. Ebenfalls kann mann nichts aus nichts *definieren*. Deswegen werden grundlegende mathematische Objekte nicht direkt definiert; stattdessen muss man grundlegende mathematische Objekte *indirekt* durch ein Axiomsystem definieren. Die Axiome sagen uns nicht, *was* genau diese Objekte sind, sondern *wie* man mit diesen Objekten umgehen kann.

Es hat sich herausgestellt, dass *Mengen* gute primitive Objekte sind, auf denen fast alle die Mathematik beruhen kann. Der moderne mathematische Begriff von „Menge“ wurde von Georg Cantor am Ende des 19. Jahrhunderts eingeführt. Sein 1895 Artikel *Beiträge zur Begründung der transfiniten Mengenlehre* beginnt mit folgendem Absatz:

> Unter einer „Menge“ verstehen wir jede Zusammenfassung M von bestimmten wohlunterschiedenen Objekten m unserer Anschauung oder unseres Denkens (welche Elemente von M genannt werden) zu einem Ganzen.

Wir werden mit so einem Begriff von „Menge“ arbeiten. Zum Beispiel, wenn man eine bestimmte Liste $a, b, c, \dots$ von Objekten betrachtet, darf man diese Objekte in einer Menge zusammenfassen. Diese Menge wird dann mit

$$\{a, b, c, \dots\}$$

bezeichnet. Wenn $\varphi(x)$ eine logische Aussage ist, darf man auch nach Cantors obiger Definition die Menge aller Objekte x betrachten, für die die Aussage $\varphi(x)$ gilt. Diese Menge wird mit

$$\{x \mid \varphi(x)\} \quad \text{oder} \quad \{x : \varphi(x)\}$$

bezeichnet.

Notation 1.2.1. Die Aussage „m ist ein Element von M“ wird mit der Notation $m \in M$ abgekürzt. Man sagt auch „m liegt in M“ oder „M enthält m“. Man schreibt $m \notin M$ für die gegenteilige Aussage, d.h., falls m kein Element von M ist.

Leider ist Cantors Definition einer Menge keine präzise mathematische Definition, und sie führt sehr schnell zu einem berühmten logischen Paradoxon:

Paradoxon 1.2.2 (Die Russelsche Antinomie)**.** Wir betrachten die Menge

$$R := \{x \mid x \text{ ist eine Menge und } x \notin x\}.$$

In Worten ist R die Menge aller Mengen, die kein Element von sich selbst sind. Nach Definition von R gilt also

$$x \in R \iff x \notin x$$

für alle Mengen x. Frage: Stimmt $R \in R$ oder nicht? Wenn wir x durch R in der obigen Äquivalenz ersetzen, erhalten wir

$$R \in R \iff R \notin R.$$

Das ist eine Aussage der Gestalt $\varphi \Leftrightarrow \neg\varphi$, also ein Widerspruch!

Die Russelsche Antinomie zeigt folgendes: Um einen widerspruchfreien Begriff von Menge zu erhalten, darf $\{x | x \notin x\}$ *nicht* eine Menge sein. Eine natürliche Frage ist dann: Für welche logischen Aussagen $\varphi(x)$ darf man die Menge $\{x \mid \varphi(x)\}$ bilden? Diese Frage ist schwierig, aber sie kann durch die *axiomatische Mengenlehre* ausführlich beantwortet werden.

In der Praxis der Mathematik ist es gar nicht wichtig, die genauen Axiome der Mengenlehre zu kennen; ein gutes intuitives Verständnis von Mengen und Mengenoperationen ist hinreichend. In diesem Abschnitt erklären wir einige Konstruktionen mit Mengen, die durch die axiomatische Mengenlehre begründet werden können, und die für den größten Teil der Mathematik ausreichen. Das unterliegende System von Axiomen, das wir implizit benutzen werden, heißt die *Zermelo–Fraenkel-Mengenlehre mit Auswahlaxiom* oder *ZFC* (das C steht für „Choice“, das englische Wort für Auswahl).

Definition 1.2.3 (Gleichheit von Mengen)**.** Zwei Mengen A und B sind genau dann *gleich*, in Zeichen $X = Y$, wenn sie dieselben Elemente enthalten.

Zum Beispiel, $\{0, 0, 1\} = \{0, 1\} = \{1, 0\}$.

Definition 1.2.4 (Die leere Menge)**.** Die *leere Menge*, $\varnothing$ oder $\{\}$, ist die Menge, die keine Elemente enthält.

Die letzten zwei Definitionen entsprechen Axiomen der Mengenlehre, nämlich das *Extensionalitätsaxiom* und das *Leermengenaxiom*. Nach Definition 1.2.3 ist die leere Menge eindeutig, d.h.: Enthalten A und B keine Elemente, so gilt $A = B$. Deswegen darf man wirklich „*die* leere Menge“ sagen.

Definition 1.2.5 (Teilmenge)**.** Seien A, B Mengen. Die Menge A heißt *Teilmenge* von B, in Zeichen $A \subset B$, falls jedes Element von A ein Element von B ist. Man sagt auch „A ist in B enthalten“.

Bemerkung 1.2.6. Manche Quellen verwenden die Notation $A \subseteq B$, wenn A eine Teilmenge von B ist, und schreiben $A \subset B$, nur wenn $A \neq B$. In diesem Skript werden wir $\subseteq$ nicht verwenden.

Wie die Russelsche Antinomie zeigt, nicht alle logischen Aussagen $\varphi(x)$ bilden gültige Mengen $\{x \mid \varphi(x)\}$. Aber wenn man die Objekte x nur innerhalb einer bestimmten Menge A nehmen, dann gibt es keine Probleme, und man darf folgende Teilmenge von A immer betrachten:[1]

$$\{x \mid x \in A \text{ und } \varphi(x)\}.$$

Diese Menge wird auch mit

$$\{x \in A \mid \varphi(x)\}$$

bezeichnet.

Definition 1.2.7 (Vereinigung, Durchschnitt, Komplement)**.** Seien A, B Mengen.

- Die *Vereinigung* von A und B ist die Menge

$$A \cup B := \{x \mid x \in A \text{ oder } x \in B\}.$$

- Der *Durchschnitt* oder die *Schnittmenge* von A und B ist die Menge

$$A \cap B := \{x \mid x \in A \text{ und } x \in B\}.$$

  Die Mengen A und B sind *disjunkt*, falls $A \cap B = \varnothing$.

- Das *Komplement* von B in A, oder die *Differenz* von A und B, ist die Menge

$$A \setminus B := \{x \mid x \in A \text{ und } x \notin B\}.$$

[1] Das folgt aus dem *Aussonderungsaxiom* der Mengenlehre.

Bemerkung 1.2.8. In der obigen Definition haben wir das Symbol := geschrieben. Der Doppelpunkt betont, dass die linke Seite durch die rechte Seite definiert wird.

Proposition 1.2.9. *Seien A, B, C Mengen.*

(i) *$A = B$ genau dann, wenn $A \subset B$ und $B \subset A$.*

(ii) *Ist $A \subset B$ und $B \subset C$, so ist $A \subset C$.*

(iii) *Die Operationen $\cup$ und $\cap$ sind* kommutativ*, d.h.,*

$$A \cup B = B \cup A, \qquad A \cap B = B \cap A.$$

(iv) *Die Operationen $\cup$ und $\cap$ sind* assoziativ*, d.h.,*

$$A \cup (B \cup C) = (A \cup B) \cup C, \qquad A \cap (B \cap C) = (A \cap B) \cap C.$$

(v) *Es gilt die* de Morganschen Gesetze

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$
$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$

(vi) *Es gilt*

$$A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$$
$$A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C).$$

Beweis. Alle Aussagen folgen unmittelbar aus den Definitionen. Wir beweisen das erste de Morgansche Gesetz

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

als Beispiel. Nach (i) braucht man zu zeigen, dass beide Inklusionen $\subset$ und $\supset$ gelten.

Zu $\subset$. Sei $x \in A \cup (B \cap C)$. Das heißt, $x \in A$ oder $x \in B \cap C$. Falls $x \in A$, dann $x \in A \cup D$ für irgendeine Menge D; insbesondere gilt $x \in A \cup B$ und $x \in A \cup C$, d.h., $x \in (A \cup B) \cap (A \cup C)$. Falls $x \in B \cap C$, dann $x \in B$ und daher $x \in A \cup B$, und auch $x \in C$ und daher $x \in A \cup C$, also $x \in (A \cup B) \cap (A \cup C)$.

Zu $\supset$. Sei $x \in (A \cup B) \cap (A \cup C)$. Das heißt, $x \in A \cup B$ und $x \in A \cup C$. Ist $x \in A$, so folgt $x \in A \cup (B \cap C)$. Sonst muss es sein, dass $x \in B$ und $x \in C$, d.h., $x \in B \cap C$. In diesem Fall gilt dann auch $x \in A \cup (B \cap C)$. □

Bemerkung 1.2.10. Die erste Aussage in Proposition 1.2.9 ist trivial aber ganz wichtig in der Praxis: Um zu beweisen, dass zwei Mengen A und B gleich sind, muss man eigentlich zwei verschiedene Aussagen beweisen: dass jedes Element von A in B liegt, und umgekehrt dass jedes Element von B in A liegt.

Definition 1.2.11 (Potenzmenge)**.** Sei X eine Menge. Die *Potenzmenge* von X, $\mathcal{P}(X)$, ist die Menge aller Teilmengen von X:

$$\mathcal{P}(X) := \{A \mid A \subset X\}.$$

Die Existenz der Potenzmenge irgendeiner Menge ist wieder ein Axiom der Mengenlehre, das *Potenzmengenaxiom*.

Bemerkung 1.2.12. Die leere Menge ist eine Teilmenge jeder Menge, weil jedes Element der leeren Menge ein Element aller anderen Mengen ist. Also für alle Mengen X gilt $\varnothing \in \mathcal{P}(X)$. Für alle Mengen X gilt auch $X \in \mathcal{P}(X)$, da $X \subset X$.

Seien a und b mathematische Objekte. Das *Paar* (a, b) ist ein neues Objekt mit folgender Eigenschaft: Zwei Paare (a, b) und (a', b') sind genau dann gleich, wenn $a = a'$ und $b = b'$. In der Mengenlehre kann man das Paar (a, b) als die Menge $\{\{a\}, \{a, b\}\}$ definieren. Allgemeiner, aus n Objekten $a_1, \ldots, a_n$ kann mann das *n-Tupel* $(a_1, \ldots, a_n)$ bilden.[2]

Bemerkung 1.2.13. Das Paar (a, b) ist nicht mit der Menge $\{a, b\}$ zu verwechseln. Zum Beispiel, es gilt immer $\{a, b\} = \{b, a\}$, aber $(a, b) = (b, a)$ gilt nur, wenn $a = b$. Das heißt, in einem Paar (a, b) ist die Reihenfolge von a und b relevant. Eine ähnliche Bemerkung gilt für n-Tupel.

Definition 1.2.14 (Produkt, Summe)**.** Seien A, B Mengen. Das (kartesische) *Produkt* von A und B ist die Menge aller Paare (a, b) mit $a \in A$ und $b \in B$:

$$A \times B := \{(a, b) \mid a \in A \text{ und } b \in B\}.$$

Die *Summe* oder *disjunkte Vereinigung* von A und B ist die Menge

$$A \sqcup B := (\{1\} \times A) \cup (\{2\} \times B).$$

In der folgenden Tabelle sind die bisher eingeführten Notationen zusammengefasst:

Notation	Bedeutung
$a \in A$	a ist ein Element von A, a liegt in A
$A \subset B$	A ist eine Teilmenge von B
$\{x \in A \mid \varphi(x)\}$	die Menge aller $x \in A$, für die $\varphi(x)$ gilt
$\varnothing$	die leere Menge
$A \cup B$	die Vereinigung von A und B
$A \cap B$	der Durchschnitt von A und B
$A \setminus B$	das Komplement von B in A
$A \times B$	das Produkt von A und B
$A \sqcup B$	die Summe von A und B
$\mathcal{P}(A)$	die Potenzmenge von A

Definitionen 1.2.7 und 1.2.14 können auf mehr als zwei Mengen verallgemeinert werden. Zum Beispiel ist das n-fache Produkt

$$A_1 \times A_2 \times \cdots \times A_n$$

die Menge aller n-Tupel $(a_1, a_2, \ldots, a_n)$ mit $a_i \in A_i$ für jedes $i \in \{1, \ldots, n\}$. Um diese Konstruktionen auf sogar unendlich viele Mengen zu verallgemeinern, brauchen wir den Begriff der *Mengenfamilie*. Eine Mengenfamilie $(A_i)_{i \in I}$ mit Indexmenge I ordnet jedem Element $i \in I$ eine Menge A_i zu. Dieses „Zuordnen“ kann präziser als eine Abbildung $i \mapsto A_i$ mit Definitionsbereich I definiert werden (siehe Abschnitt 1.3).

Definition 1.2.15 (Vereinigung, Durchschnitt, Summe und Produkt von Familien)**.** Sei I eine Menge und $(A_i)_{i \in I}$ eine Mengenfamilie mit Indexmenge I.

- Die *Vereinigung* der Familie $(A_i)_{i \in I}$ ist die Menge

$$\bigcup_{i \in I} A_i := \{x \mid \text{es existiert } i \in I \text{ mit } x \in A_i\}.$$

- Falls $I \neq \varnothing$, ist der *Durchschnitt* oder die *Schnittmenge* der Familie $(A_i)_{i \in I}$ die Menge

$$\bigcap_{i \in I} A_i := \{x \mid \text{für alle } i \in I \text{ gilt } x \in A_i\}.$$

[2] Man kann zum Beispiel dieses n-Tupel als iteriertes Paar $(a_1, (a_2, (\ldots, a_n) \ldots))$ definieren.

- Die *Summe* oder *disjunkte Vereinigung* der Familie $(A_i)_{i\in I}$ ist die Menge

$$\coprod_{i\in I} A_i := \{(i,x) \mid i \in I \text{ und } x \in A_i\}.$$

- Das *Produkt* der Familie $(A_i)_{i\in I}$ ist die Menge

$$\prod_{i\in I} A_i := \{(a_i)_{i\in I} \mid \text{für alle } i \in I \text{ gilt } a_i \in A_i\}.$$

Die Definition der Vereinigung einer Familie wird durch das *Vereinigungsaxiom* der Mengenlehre begründet. Die anderen Konstruktionen in der obigen Definition können wie folgt begründet werden: $\bigcap_{i\in I} A_i$ ist eine Teilmenge von irgendeinem A_i, $\coprod_{i\in I} A_i$ ist gleich der Vereinigung $\bigcup_{i\in I} (\{i\} \times A_i)$, und $\prod_{i\in I} A_i$ kann präziser als Teilmenge von $\mathcal{P}\left(I \times \bigcup_{i\in I} A_i\right)$ definiert werden.

Bemerkung 1.2.16. In der Definition von $\bigcap_{i\in I} A_i$ ist es notwendig, $I \neq \varnothing$ vorauszusetzen. Sonst würde der Durchschnitt $\bigcap_{i\in I} A_i$ die „universelle Menge“ sein, d.h., die Menge *aller* Objekte. Aber die universelle Menge existiert nicht, sonst würde die Russelsche Menge als Teilmenge davon auch existieren, and das ist bekanntlich nicht möglich.

Wenn man nur Familien von Teilmengen einer festen Menge X betrachtet (was fast immer der Fall ist), ist es sinnvoll ihre Vereinigungen und Durchschnitte als $\{x \in X \mid \dots\}$ zu definieren. Mit dieser Veränderung ist der Durchschnitt der Familie mit Indexmenge $\varnothing$ gleich X.

1.2.1 Die natürlichen Zahlen

Die *natürlichen Zahlen* sind die Zahlen 0, 1, 2, 3, usw. Die Menge aller natürlichen Zahlen wird mit $\mathbb{N}$ bezeichnet:

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

Die genaue Beschaffenheit der natürlichen Zahlen ist nicht wichtig, und es gibt mehrere mögliche mengentheoretische Definitionen. Eine Standarddefinition ist:

$$0 := \varnothing,\ 1 := \{0\},\ 2 := \{0,1\},\ 3 := \{0,1,2\},\ \text{usw.}$$

Mit dieser Definition ist die gewöhnliche Ordnungsrelation $\leq$ zwischen natürlichen Zahlen einfach die Teilmengenrelation $\subset$.

Übrigens würde die obige Definition von $\mathbb{N}$ in der formalen Mengenlehre nicht sinnvoll sein (was bedeutet denn „...“?). Dass eine solche Menge $\mathbb{N}$ trotzdem existiert folgt aus der *Unendlichkeitsaxiom* der Mengenlehre.

Bemerkung 1.2.17 (Ist 0 eine natürliche Zahl?)**.** In manche Quellen werden die natürlichen Zahlen als $\mathbb{N} = \{1, 2, 3, \dots\}$ definiert, d.h., die Null wird nicht als natürliche Zahl betrachtet. Dann wird $\mathbb{N} \cup \{0\}$ mit $\mathbb{N}_0$ bezeichnet. Diese alternative Konvention ist populär in der Analysis, weil die Folge $(1/n)_{n\in\mathbb{N}}$ oft verwendet wird, in der 0 kein Element von $\mathbb{N}$ sein darf. In der Algebra benutzt man eher unsere Konvention, so dass z.B. $(\mathbb{N}, +)$ ein Monoid ist (siehe Bemerkung 2.1.13).

Bemerkung 1.2.18 (ganze, rationale, reelle, komplexe Zahlen)**.** Es gibt bekanntlich mehrere Erweiterungen der natürlichen Zahlen $\mathbb{N}$:

- die ganzen Zahlen $\mathbb{Z}$;
- die rationalen Zahlen $\mathbb{Q}$;
- die reellen Zahlen $\mathbb{R}$;

- die komplexen Zahlen $\mathbb{C}$.

Im Kapitel 2 werden wir erklären, wie die Mengen $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ und $\mathbb{C}$ konstruiert werden können.

Folgender Satz ist eine fundamentale Eigenschaft der natürlichen Zahlen. Wir nehmen diesen Satz ohne Beweis an, da ein Beweis eine detaillierte Beschreibung der unterliegenden Axiome erfordern würde.

***Satz 1.2.19** (Wohlordnungsprinzip)**.** *Jede nichtleere Teilmenge $A \subset \mathbb{N}$ besitzt ein* kleinstes *Element, d.h., ein Element $a \in A$, so dass $a \leq b$ für alle $b \in A$.*

Korollar 1.2.20 (Induktionsprinzip)**.** *Sei $A \subset \mathbb{N}$ eine Teilmenge mit folgenden Eigenschaften:*

(i) (Induktionsanfang) $0 \in A$.

(ii) (Induktionsschritt) *Für alle $n \in \mathbb{N}$ gilt:*

$$n \in A \implies n+1 \in A.$$

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

Beweis. (Widerspruchsbeweis.) Angenommen, $A \neq \mathbb{N}$. Dann ist das Komplement $B := \mathbb{N} \setminus A$ nicht leer. Nach dem Wohlordnungsprinzip (Satz 1.2.19) hat B ein kleinstes Element $b \in B$. Da $0 \in A$ nach (i) ist $b \neq 0$. Also existiert $n \in \mathbb{N}$ mit $b = n+1$. Da b das kleinste Element von B ist, ist $n \notin B$, d.h., $n \in A$. Nach (ii) ist dann $b = n+1 \in A$. Also liegt b in $A \cap B = \varnothing$, im Widerspruch zur Definition von $\varnothing$. □

Das Induktionsprinzip wird häufig verwendet, um eine gegebene Aussage $\varphi(n)$ für alle natürlichen Zahlen $n \in \mathbb{N}$ zu beweisen. Dazu wenden wir das Induktionsprinzip auf die folgende Menge an:

$$A = \{n \in \mathbb{N} \mid \varphi(n)\}.$$

Es genügt also zu zeigen, dass $\varphi(0)$ gilt, und dass für alle $n \in \mathbb{N}$ die Implikation $\varphi(n) \Rightarrow \varphi(n+1)$ gilt. Dieses Beweisverfahren heißt *Beweis durch Induktion*. Während die Implikation $\varphi(n) \Rightarrow \varphi(n+1)$ im Induktionsschritt bewiesen wird, heißt die Aussage $\varphi(n)$ die *Induktionsvoraussetzung*.

Beispiel 1.2.21. Als Beispiel zum Induktionsprinzip beweisen wir folgende Aussage: Für alle $n \in \mathbb{N}$ ist die Summe aller natürlichen Zahlen $\leq n$ gleich $\frac{n(n+1)}{2}$. In Zeichen:

$$\sum_{k=0}^{n} k = \frac{n(n+1)}{2}. \tag{1.2.22}$$

Hierzu betrachten wir die Menge

$$A = \{n \in \mathbb{N} \mid (1.2.22) \text{ gilt}\} \subset \mathbb{N}.$$

- *Induktionsanfang.* Es gilt $0 \in A$, da $\sum_{k=0}^{0} k = 0 = \frac{0(0+1)}{2}$.
- *Induktionsschritt.* Sei $n \in A$. Dann

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

  Dabei haben wir die Induktionsvoraussetzung in der zweiten Gleichung verwendet. Diese Berechnung zeigt, dass $n+1 \in A$.

Nach dem Induktionsprinzip gilt dann $A = \mathbb{N}$. Das heißt, (1.2.22) gilt für alle $n \in \mathbb{N}$, wie behauptet.

Folgendes Korollar ist eine stärkere Variante des Induktionsprinzips: Um zu beweisen, dass eine Aussage $\varphi(n)$ für alle $n \in \mathbb{N}$ gilt, darf man voraussetzen, dass $\varphi(m)$ für alle $m < n$ gilt.

Korollar 1.2.23 (Prinzip der vollständigen Induktion)**.** *Sei $A \subset \mathbb{N}$ eine Teilmenge, so dass folgendes gilt für alle $n \in \mathbb{N}$:*

$$\{m \in \mathbb{N} \mid m < n\} \subset A \implies n \in A.$$

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

Beweis. Sei $\mathbb{N}_{<n} := \{m \in \mathbb{N} | m < n\}$. Wir wenden das Induktionsprinzip 1.2.20 mit folgender Menge an:

$$A' := \{n \in \mathbb{N} \mid \mathbb{N}_{<n} \subset A\} \subset \mathbb{N}.$$

- *Induktionsanfang.* Es gilt $0 \in A'$, da $\mathbb{N}_{<0} = \varnothing \subset A$.
- *Induktionsschritt.* Es sei $n \in A'$, d.h., $\mathbb{N}_{<n} \subset A$. Dann folgt aus der Voraussetzung, dass $n \in A$. Also $\mathbb{N}_{<n+1} = \mathbb{N}_{<n} \cup \{n\} \subset A$, d.h., $n + 1 \in A'$.

Aus dem Induktionsprinzip folgt, dass $A' = \mathbb{N}$. Insbesondere, für jedes $n \in \mathbb{N}$, liegt $n + 1$ in A', d.h., $\mathbb{N}_{<n+1} \subset A$, und daher $n \in A$. □

Der Beweis vom Lemma 1.1.12 war eine Anwendung des Prinzips der vollständigen Induktion.

1.3 Abbildungen

Seien X, Y Mengen. Eine *Abbildung* f von X nach Y soll etwas sein, das jedem Element x von X ein Element $f(x)$ von Y zuordnet. Folgende Definition ist der präzise mengentheoretische Ausdruck dieser Idee:

Definition 1.3.1 (Abbildung, Wert, Urbild, Definitionsmenge, Zielmenge, Graph)**.** Eine *Abbildung* ist ein Tripel $f = (X, Y, \Gamma)$, wobei X und Y Mengen sind und $\Gamma \subset X \times Y$ eine Teilmenge ihres kartesischen Produkts ist, mit folgender Eigenschaft:

$$\text{Zu jedem } x \in X \text{ gibt es } \textit{genau ein } y \in Y \text{ mit } (x, y) \in \Gamma.$$

Man sagt „f ist eine Abbildung *von X nach Y*“. Das einzige Element $y \in Y$ mit $(x, y) \in \Gamma$ heißt der *Wert* von f in x und wird mit $f(x)$ bezeichnet. Man sagt auch, dass x ein *Urbild* von y unter f ist.

Die Menge X heißt *Definitionsmenge* oder *Definitionsbereich* von f.
Die Menge Y heißt *Zielmenge* oder *Zielbereich* von f.
Die Menge Γ heißt *Graph* von f und wird auch mit Γ_f bezeichnet.

Notation 1.3.2. Die Notation $f\colon X \to Y$ bedeutet, dass f eine Abbildung von X nach Y ist. In diesem Zusammenhang, die Notation $x \mapsto y$ bedeutet, dass y der Wert von f in x ist, d.h., $y = f(x)$. Man sagt auch „f bildet x auf y ab“.

Notation 1.3.3. Wir bezeichnen die Menge aller Abbildungen von X nach Y mit $\mathrm{Abb}(X, Y)$ oder Y^X. Sie ist eine Teilmenge von $\{X\} \times \{Y\} \times \mathcal{P}(X \times Y)$.

Bemerkung 1.3.4. Abbildungen heißen auch *Funktionen*, aber das Wort „Funktion“ wird oft für Abbildungen nach $\mathbb{R}$ oder $\mathbb{C}$ vorbehalten.

Beispiel 1.3.5. Abbildungen können auf verschiedene Weise definiert werden.

(i) Wenn die Definitionsmenge endlich ist, kann man einfach alle Werte ausdrücklich geben. Zum Beispiel, folgende Liste definiert eine Abbildung f von $\{0,1,2\}$ nach $\mathbb{N}$:

$$\begin{aligned} f\colon \{0,1,2\} &\to \mathbb{N}, \\ 0 &\mapsto 5, \\ 1 &\mapsto 1, \\ 2 &\mapsto 5. \end{aligned}$$

(ii) Man kann auch eine Abbildung durch eine „Formel“ definieren. Zum Beispiel:

$$\begin{aligned} f\colon \mathbb{R} &\to \mathbb{R}, \\ x &\mapsto 2x^2 + 1. \end{aligned}$$

(iii) Man kann Abbildungen durch *Fallunterscheidung* definieren. Zum Beispiel:

$$\begin{aligned} f\colon \mathbb{R} &\to \mathbb{R}, \\ x &\mapsto \begin{cases} 2x^2 + 1, & \text{falls } x \geq 0, \\ x + 1, & \text{falls } x < 0. \end{cases} \end{aligned}$$

Hier ist es wichtig, dass die beide Fälle „$x \geq 0$“ und „$x < 0$“ miteinander ausschließlich sind und den ganzen Definitionsbereich überdecken.

(iv) Ein weiteres Beispiel ist:

$$\begin{aligned} f\colon \mathcal{P}(\mathbb{N}) \setminus \{\varnothing\} &\to \mathbb{N}, \\ A &\mapsto \text{das kleinste Element von } A. \end{aligned}$$

Diese Definition ist sinnvoll, da jede nichtleere Teilmenge von $\mathbb{N}$ ein kleinstes Element enthält (nach dem Wohlordnungsprinzip 1.2.19), und dieses Element eindeutig ist.

(v) Die folgende Variante von (iv) ist *keine* wohldefinierte Abbildung, weil die leere Teilmenge kein kleinstes Element besitzt:

$$\begin{aligned} f\colon \mathcal{P}(\mathbb{N}) &\to \mathbb{N}, \\ A &\mapsto \text{das kleinste Element von } A. \end{aligned}$$

Auch folgende Variante ist nicht sinnvoll, da die meisten Teilmengen von $\mathbb{N}$ mehr als ein Element enthalten:

$$\begin{aligned} f\colon \mathcal{P}(\mathbb{N}) \setminus \{\varnothing\} &\to \mathbb{N}, \\ A &\mapsto \text{ein Element von } A. \end{aligned}$$

Beispiel 1.3.6. Algebraische Operationen zwischen Zahlen, wie $+$, $-$ oder $\cdot$, können mithilfe des kartesischen Produkts als Abbildungen aufgefasst werden. Zum Beispiel, die Addition und Multiplikation von natürlichen Zahlen sind Abbildungen

$$\begin{aligned} +\colon \mathbb{N} \times \mathbb{N} \to \mathbb{N}, &\quad (n,m) \mapsto n + m, \\ \cdot\colon \mathbb{N} \times \mathbb{N} \to \mathbb{N}, &\quad (n,m) \mapsto n \cdot m. \end{aligned}$$

Bemerkung 1.3.7. Zu jeder Menge X gibt es genau eine Abbildung von $\varnothing$ nach X. Denn das kartesische Product $\varnothing \times X = \varnothing$ hat genau eine Teilmenge, nämlich $\varnothing$, und das Tripel $(\varnothing, X, \varnothing)$ ist eine Abbildung, weil eine Aussage der Gestalt „für alle $x \in \varnothing \ldots$“ immer wahr ist. Auf der anderen Seite gibt es eine Abbildung von X nach $\varnothing$, nur wenn X leer ist.

Bemerkung 1.3.8 (Gleichheit von Abbildungen)**.** Zwei Abbildungen f und g sind genau dann gleich, wenn sie dieselbe Definitionsmenge und dieselbe Zielmenge haben, und außerdem gilt $f(x) = g(x)$ für alle x aus der gemeinsamen Definitionsmenge.

Definition 1.3.9 (Bild, Urbild)**.** Sei $f\colon X \to Y$ eine Abbildung von X nach Y.

- Sei $A \subset X$ eine Teilmenge. Das *Bild* von A unter f ist die Menge

$$f(A) := \{f(x) \mid x \in A\} \subset Y.$$

  Das Bild von X selbst, $f(X)$, heißt die *Bildmenge* oder das *Bild* von f.

- Sei $B \subset Y$ eine Teilmenge. Das *Urbild* von B unter f ist die Menge

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

Bemerkung 1.3.10. Sei $f\colon X \to Y$ eine Abbildung. Die Notation $f(x)$ hat jetzt zwei verschiedene Bedeutungen: Wenn x ein Element von X ist, dann ist $f(x)$ ein Element von Y, nämlich der Wert von f in x. Aber wenn x eine Teilmenge von X ist, dann ist $f(x)$ eine Teilmenge von Y, nämlich das Bild von x unter f. Es könnte leider sein, dass x ein Element *sowie* eine Teilmenge von X ist (z.B. $X = \{\varnothing\}$ und $x = \varnothing$). In diesem Fall haben wir einen Problem, da $f(x)$ nicht wohldefiniert ist. Zum Glück ist das kein ernstes Problem: in der Praxis wird es immer klar sein, ob wir x als Element oder als Teilmenge von X auffassen. In der Mathematik sind solche harmlosen Zweideutigkeiten ziemlich häufig, weil es viel mehr mathematische Begriffe als verfügbare Symbole/Namen gibt. Ein anderes Beispiel: Das Wort „Urbild“ hat schon zwei verschiedene Bedeutungen (Definitionen 1.3.1 und 1.3.9)!

Proposition 1.3.11 (Eigenschaften des Bilds und des Urbilds)**.** *Sei $f\colon X \to Y$ eine Abbildung.*

(i) *Für jede Teilmenge $A \subset X$ gilt $A \subset f^{-1}(f(A))$.*

(ii) *Für jede Teilmenge $B \subset Y$ gilt $f(f^{-1}(B)) \subset B$.*

(iii) *Für jede Teilmengen $A, A' \subset X$ gelten:*

$$\begin{aligned} f(A \cup A') &= f(A) \cup f(A'), \\ f(A \cap A') &\subset f(A) \cap f(A'), \\ f(A \setminus A') &\supset f(A) \setminus f(A'). \end{aligned}$$

(iv) *Für jede Teilmengen $B, B' \subset Y$ gelten:*

$$\begin{aligned} f^{-1}(B \cup B') &= f^{-1}(B) \cup f^{-1}(B'), \\ f^{-1}(B \cap B') &= f^{-1}(B) \cap f^{-1}(B'), \\ f^{-1}(B \setminus B') &= f^{-1}(B) \setminus f^{-1}(B'). \end{aligned}$$

Beweis. Jede Aussage folgt unmittelbar aus den Definitionen. □

Definition 1.3.12 (Identität, Komposition)**.**

- Sei X eine Menge. Die *Identität* auf X ist die Abbildung

$$\begin{aligned} \mathrm{id}_X\colon X &\to X, \\ x &\mapsto x. \end{aligned}$$

  Das heißt: $\mathrm{id}_X = (X, X, \Delta_X)$, wobei $\Delta_X = \{(x, x) \mid x \in X\}$ die diagonale Teilmenge ist.

- Seien $f\colon X \to Y$ und $g\colon Y \to Z$ Abbildungen. Die *Komposition* oder *Verkettung* von f und g ist die Abbildung

$$\begin{aligned} g \circ f\colon X &\to Z, \\ x &\mapsto g(f(x)). \end{aligned}$$

  Die Notation $g \circ f$ wird als „g nach f“ gelesen.

Proposition 1.3.13 (Eigenschaften von Komposition)**.**

(i) *Sei $f\colon X \to Y$ eine Abbildung. Dann $f \circ \mathrm{id}_X = f$ und $\mathrm{id}_Y \circ f = f$.*

(ii) (Assoziativität der Komposition) *Seien $f\colon X \to Y$, $g\colon Y \to Z$ und $h\colon Z \to W$ drei Abbildungen. Dann gilt*

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

Beweis. Um zu beweisen, dass zwei Abbildungen mit derselben Definitions- und Zielmenge gleich sind, ist zu zeigen, dass sie denselben Wert in jedem Element ihrer Definitionsmenge nehmen (Bemerkung 1.3.8). Dies folgt unmittelbar aus den Definitionen von id und $\circ$. Zum Beispiel, für jedes $x \in X$ gilt:

$$\begin{aligned} (f \circ \mathrm{id}_X)(x) &= f(\mathrm{id}_X(x)) && \text{(Definition von } \circ) \\ &= f(x) && \text{(Definition von } \mathrm{id}_X). \end{aligned}$$ □

Definition 1.3.14 (Einschränkung)**.** Sei $f\colon X \to Y$ eine Abbildung und $A \subset X$ eine Teilmenge. Die *Einschränkung* von f auf A ist die wie folgt definierte Abbildung:

$$\begin{aligned} f|_A\colon A &\to Y, \\ x &\mapsto f(x). \end{aligned}$$

Das heißt: $f|_A = (A, Y, \Gamma_f \cap (A \times Y))$.

Definition 1.3.15 (Inklusionsabbildung)**.** Sei X eine Menge und $A \subset X$ eine Teilmenge. Die Abbildung

$$\begin{aligned} i_A\colon A &\to X, \\ x &\mapsto x, \end{aligned}$$

heißt die *Inklusionsabbildung* oder *Inklusion* von A in X.

Bemerkung 1.3.16. Es gilt $i_A = \mathrm{id}_X|_A$ und $f|_A = f \circ i_A$.

Definition 1.3.17 (kanonische Projektionen)**.** Seien A, B Mengen. Die Abbildungen

$$\begin{aligned} \pi_1\colon A \times B &\to A, \\ (a,b) &\mapsto a, \end{aligned} \qquad \begin{aligned} \pi_2\colon A \times B &\to B, \\ (a,b) &\mapsto b, \end{aligned}$$

heißen die *kanonischen Projektionen* aus $A \times B$ auf die Faktoren.

Allgemeiner, sei $(A_i)_{i \in I}$ eine Mengenfamilie mit Indexmenge I und sei $e \in I$. Die e-te kanonische Projektion ist die Abbildung

$$\begin{aligned} \pi_e\colon \prod_{i \in I} A_i &\to A_e, \\ (a_i)_{i \in I} &\mapsto a_e. \end{aligned}$$

Bemerkung 1.3.18. Das Wort „kanonisch“ hat keine präzise Bedeutung in der Mathematik, aber es ist trotzdem häufig verwendet. Es könnte viele verschiedene Abbildungen $A \times B \to A$ sein, aber ohne weitere Informationen gibt es nur eine, die „besonders“ ist, nämlich π_1. Deswegen ist π_1 die „kanonische Abbildung“ $A \times B \to A$. Ein anderes Beispiel: Wenn A eine Teilmenge von X ist, dann würde die kanonische Abbildung $A \to X$ die Inklusionsabbildung sein.

Definition 1.3.19 (injektiv, surjektiv, bijektiv)**.** Sei $f\colon X \to Y$ eine Abbildung.

- f heißt *injektiv*, wenn jedes Element von Y *höchstens ein* Urbild unter f besitzt.
- f heißt *surjektiv*, wenn jedes Element von Y *mindestens ein* Urbild unter f besitzt.
- f heißt *bijektiv*, wenn jedes Element von Y *genau ein* Urbild unter f besitzt.

Nach Definition gilt also: bijektiv $\iff$ injektiv und surjektiv. Eine injektive/surjektive/bijektive Abbildung wird auch als Injektion/Surjektion/Bijektion bezeichnet.

Beispiel 1.3.20.

(i) Die Abbildung $f\colon \{0,1,2\} \to \mathbb{N}$ aus Beispiel 1.3.5(i) ist nicht injektiv, da sie denselben Wert 5 in zwei verschiedenen Elementen nimmt (d.h., 5 hat zwei Urbilder unter f, nämlich 0 und 2). Sie ist nicht surjektiv, da $57 \notin f(\{0,1,2\})$ (d.h., 57 hat kein Urbild unter f).

(ii) Folgende Abbildung ist injektiv (und nicht surjektiv):

$$\begin{aligned} f\colon \{0,1,2\} &\to \mathbb{N}, \\ 0 &\mapsto 5, \\ 1 &\mapsto 1, \\ 2 &\mapsto 0. \end{aligned}$$

(iii) Die Abbildung $f\colon \mathbb{R} \to \mathbb{R}$ aus Beispiel 1.3.5(ii) ist weder injektiv noch surjektiv. Die aus Beispiel 1.3.5(iii) ist bijektiv.

(iv) Die Abbildung $f\colon \mathcal{P}(\mathbb{N}) \setminus \{\varnothing\} \to \mathbb{N}$ aus Beispiel 1.3.5(iv) ist surjektiv aber nicht injektiv.

(v) Die Abbildung

$$\begin{aligned} &\mathbb{Z} \to \mathbb{N}, \\ &n \mapsto \begin{cases} 2n, & \text{falls } n \geq 0, \\ -(2n+1), & \text{falls } n < 0, \end{cases} \end{aligned}$$

ist bijektiv.

(vi) Die einzige Abbildung $\varnothing \to X$ ist injektiv (es gibt nichts zu zeigen!). Sie ist genau dann surjektiv, wenn X leer ist.

Beispiel 1.3.21. Seien A und B zwei Mengen. Es gibt eine kanonische Abbildung

$$\begin{aligned} A \sqcup B &\to A \cup B, \\ (1,a) &\mapsto a, \\ (2,b) &\mapsto b. \end{aligned}$$

(Nach Definition 1.2.14 ist $A \sqcup B$ eine Teilmenge von $\{1,2\} \times (A \cup B)$, und diese Abbildung ist die Einschränkung der zweiten kanonischen Projektion π_2.) Diese Abbildung ist immer surjektiv. Sie ist genau dann injektiv (und damit bijektiv), wenn A und B disjunkt sind.

Definition 1.3.22 (Umkehrabbildung)**.** Sei $f\colon X \to Y$ eine Abbildung. Eine Abbildung $g\colon Y \to X$ heißt *Umkehrabbildung* oder *inverse Abbildung* von f, falls

$$g \circ f = \mathrm{id}_X \quad \text{und} \quad f \circ g = \mathrm{id}_Y .$$

Wenn sie existiert, eine Umkehrabbildung von $f\colon X \to Y$ ist *eindeutig bestimmt*, denn: Sind g und g' zwei Umkehrabbildungen von f, so gilt

$$\begin{aligned}
g &= g \circ \mathrm{id}_Y && \text{(Proposition 1.3.13(i))}\\
&= g \circ (f \circ g') && (g' \text{ invers zu } f)\\
&= (g \circ f) \circ g' && \text{(Proposition 1.3.13(ii))}\\
&= \mathrm{id}_X \circ g' && (g \text{ invers zu } f)\\
&= g'. && \text{(Proposition 1.3.13(i)).}
\end{aligned}$$

Deswegen kann man von *der* Umkehrabbildung von f sprechen, und sie mit f^{-1} bezeichnen.

Satz 1.3.23. *Sei $f\colon X \to Y$ eine Abbildung. Die folgenden Aussagen sind äquivalent:*

(i) *f ist bijektiv.*

(ii) *f besitzt eine Umkehrabbildung.*

Beweis. Zu (i) $\Rightarrow$ (ii). Sei $f = (X, Y, \Gamma)$ bijektiv. Wir betrachten die Menge

$$\Gamma' := \{(y, x) \in Y \times X \mid (x, y) \in \Gamma\}.$$

Dann ist $g = (Y, X, \Gamma')$ eine Abbildung, denn: Sei $y \in Y$. Da f injektiv ist, gibt es höchstens ein $x \in X$ mit $(y, x) \in \Gamma'$. Da f surjetiv ist, gibt es mindestens ein $x \in X$ mit $(y, x) \in \Gamma'$. Also gibt es genau ein $x \in X$ mit $(y, x) \in \Gamma'$, d.h., g ist eine Abbildung. Nach Konstruktion gilt $g \circ f = \mathrm{id}_X$ und $f \circ g = \mathrm{id}_Y$, d.h., g ist die Umkehrabbildung von f.

Zu (ii) $\Rightarrow$ (i). Sei g eine Umkehrabbildung von f, und sei $y \in Y$. Zu zeigen ist, dass y genau ein Urbild unter f besitzt. Da $f \circ g = \mathrm{id}_Y$ ist $g(y)$ ein Urbild von y unter f, also gibt es mindestens ein Urbild. Seien x, x' zwei Urbilder von y unter f, d.h., $f(x) = y$ und $f(x') = y$. Da $g \circ f = \mathrm{id}_X$ gilt

$$x = g(f(x)) = g(y) = g(f(x')) = x'.$$

Also ist das Urbild von y eindeutig, wie gewünscht. □

Beispiel 1.3.24. Zu jeder Menge X gibt es eine bijektive Abbildung

$$\begin{aligned}
\mathrm{Abb}(X, \{0, 1\}) &\to \mathcal{P}(X),\\
f &\mapsto f^{-1}(\{1\}).
\end{aligned}$$

Die Umkehrabbildung bildet eine Teilmenge $A \in \mathcal{P}(X)$ auf die Abbildung $\chi_A\colon X \to \{0, 1\}$ ab, die durch

$$\chi_A(x) = \begin{cases} 1, & \text{falls } x \in A,\\ 0, & \text{andernfalls} \end{cases}$$

definiert wird. Die Abbildung χ_A heißt die *charakteristische Funktion* der Teilmenge A.

Proposition 1.3.25. *Seien X, Y, Z Mengen und seien $f\colon X \to Y$ und $g\colon Y \to Z$ Abbildungen.*

(i) *Sind f und g injektiv, so ist $g \circ f$ injektiv.*

(ii) *Sind f und g surjektiv, so ist $g \circ f$ surjektiv.*

(iii) *Ist $g \circ f$ injektiv, so ist f injektiv.*

(iv) *Ist $g \circ f$ surjektiv, so ist g surjektiv.*

Beweis. Wir beweisen stellvertretend (i) und (iv).

Zu (i). Zu zeigen ist, dass jedes $z \in Z$ höchstens ein Urbild unter $g \circ f$ besitzt. Seien $x, x' \in X$ zwei Urbilder von z, d.h., $g(f(x)) = g(f(x')) = z$. Aus der Injektivität von g folgt $f(x) = f(x')$, und aus der Injektivität von f folgt wiederum $x = x'$.

Zu (iv). Zu zeigen ist, dass jedes $z \in Z$ mindestens ein Urbild unter g besitzt. Da $g \circ f$ surketiv ist, gibt es ein $x \in X$ mit $g(f(x)) = z$. Insbesondere ist $f(x)$ ein Urbild von z unter g. □

Notation 1.3.26. Wenn wir die Injektivität bzw. Surjektivität einer Abbildung $f\colon X \to Y$ betonen wollen, schreiben wir manchmal $f\colon X \hookrightarrow Y$ bzw. $f\colon X \twoheadrightarrow Y$. Eine alternative Notation für injektive Abbildungen ist $f\colon X \rightarrowtail Y$.

Manchmal ist es nützlich, eine Abbildung $I \to X$ als eine „Familie von Elementen von X“ aufzufassen:

Definition 1.3.27 (Familie, Folge)**.** Seien I und X Mengen. Eine *Familie* $(x_i)_{i \in I}$ in X mit Indexmenge I ist einfach eine Abbildung

$$\begin{aligned} I &\to X, \\ i &\mapsto x_i. \end{aligned}$$

Eine *Folge* $(x_n)_{n \in \mathbb{N}}$ in X ist eine Familie in X mit Indexmenge $\mathbb{N}$.

1.3.1 Mächtigkeit

Definition 1.3.28 (Gleichmächtigkeit)**.** Zwei Mengen X and Y sind *gleichmächtig*, in Zeichen $|X| = |Y|$, wenn eine bijektive Abbildung $X \to Y$ existiert.

Definition 1.3.29 (Endlichkeit, Abzählbarkeit)**.** Sei X eine Menge.

- X heißt *endlich*, wenn eine natürliche Zahl n existiert, so dass X und $\{1, 2, \ldots, n\}$ gleichmächtig sind. Die natürliche Zahl n ist dann eindeutig bestimmt; sie heißt die *Mächtigkeit* oder *Kardinalität* von X, und wird mit $|X|$ bezeichnet.
- X heißt *unendlich*, wenn sie nicht endlich ist.
- X heißt *abzählbar*, wenn entweder X endlich ist oder X und $\mathbb{N}$ gleichmächtig sind.
- X heißt *überabzählbar*, wenn sie nicht abzählbar ist.

Bemerkung 1.3.30. In der Definition 1.3.29 betrachten wir die Menge $\{1, 2, \ldots, n\}$ mit n einer beliebigen natürlichen Zahl. Wenn $n = 0$ verstehen wir $\{1, 2, \ldots, n\}$ als die leere Menge. Insbesondere ist eine Menge genau dann leer, wenn ihre Mächtigkeit gleich null ist.

Beispiel 1.3.31.

(i) $\mathbb{N}$, $\mathbb{Z}$, und $\mathbb{Q}$ sind paarweise gleichmächtig, und damit abzählbar. Eine bijektive Abbildung zwischen $\mathbb{N}$ und $\mathbb{Z}$ wurde im Beispiel 1.3.20(v) gegeben. Die Abzählbarkeit von $\mathbb{Q}$ folgt aus Cantors erstem Diagonalargument.

(ii) $\mathbb{R}$ und $\mathbb{C}$ sind gleichmächtig und überabzählbar. Die Unabzählbarkeit von $\mathbb{R}$ folgt aus Cantors zweitem Diagonalargument.

(iii) Eine Menge X und ihre Potenzmenge $\mathcal{P}(X)$ sind nie gleichmächtig, d.h., $\mathcal{P}(X)$ ist „wirklich größer“ als X. Denn es wäre eine surjektive Abbildung $f\colon X \to \mathcal{P}(X)$. Sei $M = \{x \in X \mid x \notin f(x)\}$. Da f surjektiv ist, existiert ein $m \in X$ mit $f(m) = M$. Nach Definition von M gilt dann: $m \in M \iff m \notin M$, was ein Widerspruch ist.

(iv) Man kann zeigen, dass die Menge $\mathbb{N}^{\mathbb{N}}$ aller Abbildungen von $\mathbb{N}$ nach $\mathbb{N}$ (alias Folgen in $\mathbb{N}$) zu $\mathbb{R}$ gleichmächtig ist. Die Teilmenge $\mathbb{N}^{(\mathbb{N})} \subset \mathbb{N}^{\mathbb{N}}$ aller Folgen, die schließlich null sind, ist aber abzählbar. Man kann eine bijektive Abbildung von $\mathbb{N}^{(\mathbb{N})}$ nach $\mathbb{N}$ explizit definieren:

$$\begin{aligned} \mathbb{N}^{(\mathbb{N})} &\to \mathbb{N}, \\ (a_n)_{n\in\mathbb{N}} &\mapsto (p_0^{a_0} p_1^{a_1} p_2^{a_2} \dots) - 1, \end{aligned}$$

wobei $p_0 = 2$, $p_1 = 3$, $p_2 = 5$, usw. alle Primzahlen in aufsteigender Reihenfolge sind. Hierbei werden der *Satz von Euklid* und der *Fundamentalsatz der Arithmetik* verwendet (Sätze 1.1.11 und 1.1.14): Es gibt unendlich viele Primzahlen, und jede natürliche Zahl ≥ 1 läßt sich eindeutig als Produkt von Primzahlen darstellen.

Bemerkung 1.3.32. Seien A, B endliche Mengen und seien $a, b \in \mathbb{N}$ ihre jeweiligen Mächtigkeiten. Dann:

- Die Mächtigkeit der Summe $A \sqcup B$ ist $a + b$.
- Die Mächtigkeit des Produkts $A \times B$ ist ab.
- Die Mächtigkeit der Menge $\mathrm{Abb}(A, B)$ ist b^a. (Deswegen wird diese Menge auch mit B^A bezeichnet.)

Seien X, Y endliche Mengen. Es gilt $|X| \leq |Y|$ genau dann, wenn eine injektive Abbildung $X \hookrightarrow Y$ existiert. Also wenn injektive Abbildungen von X nach Y sowie von Y nach X existieren, dann sind X und Y gleichmächtig. Folgender Satz verallgemeinert diese Beobachtung auf unendliche Mengen:

Satz 1.3.33 (Satz von Cantor–Bernstein–Schröder)**.** *Seien X, Y Mengen. Wenn injektive Abbildungen $f\colon X \hookrightarrow Y$ und $g\colon Y \hookrightarrow X$ existieren, dann sind X und Y gleichmächtig.*

**Beweis.* Wir definieren eine Folge von Teilmengen $A_0, A_1, A_2, \dots \subset X$ wie folgt:

$$\begin{aligned} A_0 &= X \setminus g(Y), \\ A_{n+1} &= g(f(A_n)), \end{aligned}$$

und wir setzen

$$A := \bigcup_{n\in\mathbb{N}} A_n \subset X.$$

Nach Definition gilt

$$g(f(A)) = g\left(f\left(\bigcup_{n\in\mathbb{N}} A_n\right)\right) = \bigcup_{n\in\mathbb{N}} g(f(A_n)) = \bigcup_{n\in\mathbb{N}} A_{n+1},$$

und daher

$$A = A_0 \cup g(f(A)). \tag{1.3.34}$$

Es gilt $X \setminus g(Y) = A_0 \subset A$, und daher $X \setminus A \subset g(Y)$. Das heißt, jedes $x \in X \setminus A$ liegt im Bild von g. Da g injektiv ist, gibt es eigentlich *genau ein* Urbild von x unter g, das wir mit $g^{-1}(x)$ bezeichnen. Ebenso hat jedes $y \in f(A)$ genau ein Urbild $f^{-1}(y)$ unter f. Deswegen darf man definieren:

$$\begin{aligned} h\colon X &\to Y, & k\colon Y &\to X, \\ x &\mapsto \begin{cases} f(x), & \text{falls } x \in A, \\ g^{-1}(x), & \text{falls } x \in X \setminus A. \end{cases} & y &\mapsto \begin{cases} g(y), & \text{falls } y \in Y \setminus f(A), \\ f^{-1}(y), & \text{falls } y \in f(A). \end{cases} \end{aligned}$$

Wir behaupten, dass h bijektiv ist, mit Umkehrabbildung k.

- $k \circ h = \mathrm{id}_X$: Sei $x \in X$. Ist $x \in A$, so ist $f(x) \in f(A)$, und daher $k(h(x)) = k(f(x)) = f^{-1}(f(x)) = x$. Ist $x \notin A$, so ist $g^{-1}(x) \notin f(A)$, sonst wäre $x = g(g^{-1}(x)) \in g(f(A)) \subset A$ nach (1.3.34). Also gilt $k(h(x)) = k(g^{-1}(x)) = g(g^{-1}(x)) = x$.
- $h \circ k = \mathrm{id}_Y$: Sei $y \in Y$. Ist $y \in f(A)$, so ist $f^{-1}(y) \in A$, und daher $h(k(y)) = h(f^{-1}(y)) = f(f^{-1}(y)) = y$. Ist $y \notin f(A)$, so ist $g(y) \notin g(f(A))$ nach der Injektivität von g und $g(y) \notin A_0$ nach Definition von A_0, also $g(y) \notin A$ nach (1.3.34). Also gilt $h(k(y)) = h(g(y)) = g^{-1}(g(y)) = y$. □

Bemerkung 1.3.35. Es ist auch der Fall, dass X und Y gleichmächtig sind, wenn surjektive Abbildungen $f\colon X \twoheadrightarrow Y$ und $g\colon Y \twoheadrightarrow X$ existieren, oder wenn eine injektive Abbildung $f\colon X \hookrightarrow Y$ sowie eine surjektive Abbildung $g\colon X \twoheadrightarrow Y$ existieren. Diese Aussagen folgen aus dem Satz 1.3.33, weil jede surjektive Abbildung $f\colon X \twoheadrightarrow Y$ einen Schnitt $s\colon Y \to X$ besitzt, der eine injektive Abbildung ist (siehe Proposition 1.4.13(i)).

Nach dem Satz von Cantor–Bernstein–Schröder kann man von der Mächtigkeit unendlicher Mengen vernünftig sprechen. Man sagt zum Beispiel, dass die Mächtigkeit von X kleiner oder gleich der von Y ist, in Zeichen $|X| \leq |Y|$, wenn eine injektive Abbildung von X nach Y existiert. Dann sind zwei Mengen X und Y genau dann gleichmächtig, wenn $|X| \leq |Y|$ und $|Y| \leq |X|$. Eine Menge X ist genau dann abzählbar, wenn $|X| \leq |\mathbb{N}|$, und sie ist genau dann unendlich, wenn $|\mathbb{N}| \leq |X|$. Man kann auch zeigen, dass je zwei Mengen X, Y vergleichbare Mächtigkeiten haben, d.h., es gilt $|X| \leq |Y|$ oder $|Y| \leq |X|$ (dazu braucht man das Auswahlaxiom, siehe Abschnitt 1.4.2).

Man soll beachten, dass sich der Begriff der Mächtigkeit bei unendlichen Mengen manchmal anders als bei endlichen Mengen verhält. Zum Beispiel:

***Satz 1.3.36** (Mächtigkeit unendlicher Vereinigungen)**.** *Sei $(A_i)_{i\in I}$ eine Mengenfamilie mit Vereinigung $A = \bigcup_{i\in I} A_i$. Ist I unendlich und gilt $|A_i| \leq |I|$ für alle $i \in I$, so gilt auch $|A| \leq |I|$.*

Insbesondere: Eine abzählbare Vereinigung abzählbarer Mengen ist wieder abzählbar.

1.4 Relationen

Definition 1.4.1 (Relation, reflexiv, transitiv, symmetrisch, antisymmetrisch, total)**.** Sei X eine Menge. Eine *Relation* R auf X ist eine Teilmenge $R \subset X \times X$. Man sagt „x steht in Relation zu y bzgl. R“ und schreibt xRy, falls $(x, y) \in R$.

Eine Relation R auf X heißt:

- *reflexiv*, wenn xRx für alle $x \in X$;
- *transitiv*, wenn für alle $x, y, z \in X$,
$$xRy \text{ und } yRz \Longrightarrow xRz;$$
- *symmetrisch*, wenn für alle $x, y \in X$,
$$xRy \Longrightarrow yRx;$$
- *antisymmetrisch*, wenn für alle $x, y \in X$,
$$xRy \text{ und } yRx \Longrightarrow x = y;$$
- *total*, wenn für alle $x, y \in X$, xRy oder yRx.

Definition 1.4.2 (Äquivalenzrelation, partielle Ordnung, totale Ordnung)**.** Sei R eine Relation auf einer Menge X.

- R heißt *Äquivalenzrelation*, falls R reflexiv, transitiv und symmetrisch ist.
- R heißt *partielle Ordnung*, falls R reflexiv, transitiv und antisymmetrisch ist. Man sagt dann auch, dass das Paar (X, R) eine partiell geordnete Menge ist.
- R heißt *totale Ordnung*, falls R eine partielle Ordnung ist und außerdem total ist. Man sagt dann auch, dass das Paar (X, R) eine total geordnete Menge ist.

Beispiel 1.4.3.

(i) Die Identitäts- oder Gleichheitsrelation $=$ zwischen Elementen von X ist eine Äquivalenzrelation auf X. Sie entspricht der diagonalen Teilmenge

$$\Delta_X = \{(x, x) \mid x \in X\} \subset X \times X.$$

(ii) Die gewöhnliche Relation $\leq$ auf $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ bzw. $\mathbb{R}$ ist eine totale Ordnung.

(iii) Die Inklusionsrelation $\subset$ auf der Potenzmenge $\mathcal{P}(X)$ ist eine partielle Ordnung. Hat X mindestens zwei Elemente, so ist diese Ordnung *nicht* total.

(iv) Gleichmächtigkeit ist eine Äquivalenzrelation auf $\mathcal{P}(X)$. Denn seien A, B, C Teilmengen von X:

- *Zur Reflexivität:* Die Identität $\mathrm{id}_A \colon A \to A$ ist bijektiv, also ist A zu sich selbst gleichmächtig.
- *Zur Symmetrie:* Nach Satz 1.3.23 besitzt jede bijektive Abbildung eine Umkehrabbildung, die wieder bijektiv ist.
- *Zur Transitivität:* Nach Proposition 1.3.25(i,ii) ist die Komposition zweier bijektiven Abbildungen wieder bijektiv.

(v) Die Teilbarkeitsrelation $|$ auf $\mathbb{N}$ ist so definiert:

$$n|m \iff \text{es existiert } k \in \mathbb{N} \text{ mit } m = kn.$$

Sie ist eine partielle Ordnung auf $\mathbb{N}$, die nicht total ist.

Man kann ebenso die Teilbarkeitsrelation $|$ auf $\mathbb{Z}$ definieren. Auf $\mathbb{Z}$ ist diese Relation immer noch reflexiv und transitiv, aber sie ist nicht mehr antisymmetrisch, weil z.B. $2|-2$ und $-2|2$.

1.4.1 Quotient einer Menge modulo einer Äquivalenzrelation

Definition 1.4.4 (Äquivalenzklasse, Quotientenmenge, Quotientenabbildung)**.** Sei $\sim$ eine Äquivalenzrelation auf einer Menge X.

- Sei $x \in X$. Die Menge
$$[x] := \{y \in X \mid y \sim x\} \subset X$$
heißt die *Äquivalenzklasse* von x bzgl. $\sim$.
- Die *Quotientenmenge* von X bzgl. $\sim$ ist die Menge aller Äquivalenzklassen
$$X/\!\sim := \{[x] \mid x \in X\} \subset \mathcal{P}(X)$$
(gelesen „X modulo $\sim$“ oder „X durch $\sim$“).
- Die surjektive Abbildung
$$\begin{aligned} X &\twoheadrightarrow X/\!\sim, \\ x &\mapsto [x], \end{aligned}$$
heißt die *Quotientenabbildung* oder die *kanonische Abbildung*.

Definition 1.4.5 (Partition)**.** Sei X eine Menge. Eine *Partition* von X ist eine Menge $\mathcal{A}$ von nichtleeren Teilmengen von X, d.h., $\mathcal{A} \subset \mathcal{P}(X) \setminus \{\varnothing\}$, so dass jedes Element von X in *genau einem* Element von $\mathcal{A}$ liegt.

Proposition 1.4.6. *Sei $\sim$ eine Äquivalenzrelation auf einer Menge X. Dann ist $X/\sim$ eine Partition von X.*

Beweis. Da $\sim$ reflexiv ist, liegt jedes x in seiner Äquivalenzklasse $[x]$. Insbesondere besteht $X/\sim$ aus nichtleeren Teilmengen von X. Seien $[x_1], [x_2] \in X/\sim$ zwei Äquivalenzklassen mit $x \in [x_1]$ und $x \in [x_2]$. Ziel ist zu zeigen, dass $[x_1] = [x_2]$. Da die Situation symmetrisch ist, genügt es zu zeigen, dass $[x_1] \subset [x_2]$. Sei $y \in [x_1]$. Nach Definition von $[-]$ gilt $x \sim x_1$, $x \sim x_2$ und $y \sim x_1$. Da $\sim$ symmetrisch ist, gilt $x_1 \sim x$. Da $\sim$ transitiv ist und $y \sim x_1 \sim x \sim x_2$, gilt $y \sim x_2$, d.h., $y \in [x_2]$. □

Bemerkung 1.4.7. Umgekehrt, wenn $\mathcal{A} \subset \mathcal{P}(X)$ eine Partition von X ist, ist dann die Relation

$$x \sim_{\mathcal{A}} y \iff \text{es existiert } A \in \mathcal{A} \text{ mit } x, y \in A$$

eine Äquivalenzrelation auf X, so dass $X/\sim_{\mathcal{A}} = \mathcal{A}$.

Die Abbildungen $\sim \mapsto X/\sim$ und $\mathcal{A} \mapsto \sim_{\mathcal{A}}$ bilden eigentlich zueinander inverse Bijektionen zwischen der Menge aller Äquivalenzrelationen auf X und der Menge aller Partitionen von X.

Beispiel 1.4.8 (Kongruenzrelation)**.** Sei n eine natürliche Zahl und sei

$$n\mathbb{Z} := \{nx \mid x \in \mathbb{Z}\}.$$

Man definiert eine Relation $\equiv_n$ auf $\mathbb{Z}$ wie folgt:

$$x \equiv_n y \iff x - y \in n\mathbb{Z}.$$

Man schreibt üblicherweise

$$x \equiv y \mod n$$

(gelesen „x ist kongruent zu y modulo n“) statt $x \equiv_n y$. Man kann leicht nachprüfen, dass $\equiv_n$ eine Äquivalenzrelation auf $\mathbb{Z}$ ist. Die Quotientenmenge $\mathbb{Z}/\equiv_n$ wird mit $\mathbb{Z}/n\mathbb{Z}$ bezeichnet, und ihre Elemente heißen *Restklassen modulo n*. Falls $n \neq 0$, ist die Menge $\mathbb{Z}/\equiv_n$ endlich der Mächtigkeit n: Die Komposition der Inklusionsabbildung und der Quotientenabbildung

$$\{1, \dots, n\} \hookrightarrow \mathbb{Z} \twoheadrightarrow \mathbb{Z}/n\mathbb{Z}$$

ist bijektiv.

Satz 1.4.9 (universelle Eigenschaft der Quotientenmenge)**.** *Sei $\sim$ eine Äquivalenzrelation auf einer Menge X, $q\colon X \to X/\sim$ die Quotientenabbildung, und $f\colon X \to Y$ eine beliebige Abbildung. Folgende Aussagen sind äquivalent:*

(i) *Für alle $x, x' \in X$ gilt*

$$x \sim x' \implies f(x) = f(x'). \tag{1.4.10}$$

(ii) *Es existiert eine Abbildung $\bar{f}\colon X/\sim \to Y$ mit $f = \bar{f} \circ q$.*

Außerdem ist die Abbildung $\bar{f}$ eindeutig bestimmt (wenn sie existiert).

Beweis. Wir müssen die beiden Implikationen (i) ⇒ (ii) und (ii) ⇒ (i) beweisen, und dann auch die letzte Aussage.

Zu (i) ⇒ (ii). Wir betrachten die Teilmenge

$$\bar{\Gamma} = \{(A, y) \in X/\sim \times Y \mid y \in f(A)\} \subset X/\sim \times Y,$$

und wir setzen $\bar{f} := (X/\!\sim, Y, \bar{\Gamma})$. Nach Definition liegt ein Paar (A, y) in $\bar{\Gamma}$ genau dann, wenn es ein Element $x \in A$ gibt, so dass $f(x) = y$. Sind $x, x' \in A$ zwei solche Elemente, so gilt $x \sim x'$ und daher $f(x) = f(x')$ nach (1.4.10). Dies zeigt, dass es zu jeder Äquivalenzklasse $A \in X/\!\sim$ genau einem $y \in Y$ mit $(A, y) \in \bar{\Gamma}$ gibt. Also ist $\bar{f}$ eine Abbildung, und es ist jetzt klar, dass $\bar{f} \circ q = f$.

Zu (i) $\Rightarrow$ (ii). Angenommen, $\bar{f}$ existiert. Seien $x, x' \in X$ mit $x \sim x'$. Dann gilt:

$$f(x) = \bar{f}(q(x)) = \bar{f}(q(x')) = f(x').$$

Also ist die Bedingung (1.4.10) erfüllt.

Zur Eindeutigkeit. Seien $\bar{f}$ und $\tilde{f}$ zwei Abbildungen $X/\!\sim \to Y$, so dass $f = \bar{f} \circ q$ und $f = \tilde{f} \circ q$. Sei $[x] \in X/\!\sim$ eine beliebige Äquivalenzklasse. Dann gilt:

$$\bar{f}([x]) = \bar{f}(q(x)) = f(x) = \tilde{f}(q(x)) = \tilde{f}([x]).$$

Also gilt $\bar{f} = \tilde{f}$. □

Die folgende Situation kommt sehr häufig vor: man möchte eine Abbildung $f\colon X/\!\sim \to Y$ durch

$$f([x]) = g(x)$$

definieren, wobei g eine gewisse Abbildung von X nach Y ist. Bei einer solchen Definition muss man immer nachprüfen, dass g der Bedingung (1.4.10) genügt, d.h., dass $x \sim x' \Rightarrow g(x) = g(x')$. In diesem Fall sagt man, dass f durch die obige Gleichung *wohldefiniert* ist. Sonst wäre eine solche Definition *nicht sinnvoll*.

Bemerkung 1.4.11. Die Aussage vom Satz 1.4.9 kann man auch auf folgende Weise formulieren. Die Abbildung

$$\begin{aligned} \mathrm{Abb}(X/\!\sim, Y) &\to \mathrm{Abb}(X, Y), \\ g &\mapsto g \circ q, \end{aligned}$$

ist injektiv, und ihr Bild besteht genau aus diesen Abbildungen $f\colon X \to Y$, die die Bedingung (1.4.10) erfüllen.

Bemerkung 1.4.12 (kommutative Diagramme)**.** Eine Situation mit mehreren Mengen und Abbildungen zwischen denen läßt sich oft gut durch ein Diagramm veranschaulichen. Zum Beispiel: Das Dreieck

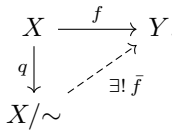
$$\begin{array}{ccc} X & \xrightarrow{h} & Z \\ {\scriptstyle f}\downarrow & \nearrow_{g} & \\ Y & & \end{array}$$

stellt drei Mengen X, Y, Z und drei Abbildungen f, g, h mit gegebenen Definitions- und Zielmengen dar. Ein solches Dreieck heißt *kommutativ*, wenn $h = g \circ f$. Im Allgemeinen, ein Diagramm von Mengen und Abbildungen heißt *kommutativ*, wenn folgendes gilt: Für alle zwei Mengen X, Y im Diagramm stimmen alle möglichen Kompositionen von Abbildungen von X nach Y überein. Beispielsweise ist das Quadrat

$$\begin{array}{ccc} X & \xrightarrow{h} & Z \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle k} \\ Y & \xrightarrow[g]{} & W \end{array}$$

genau dann kommutativ, wenn $g \circ f = k \circ h$.

Die universelle Eigenschaft der Quotientenmenge läßt sich dann wie folgt formulieren: Wenn $f\colon X \to Y$ die Implikation $x \sim x' \Rightarrow f(x) = f(x')$ erfüllt, dann existiert genau eine Abbildung $\bar{f}\colon X/\!\sim \to Y$, so dass folgendes Dreieck kommutiert:



1.4.2 Das Auswahlaxiom und das Zornsche Lemma

Ein besonderes Axiom der Mengenlehre ist das *Auswahlaxiom*:

> Sei $(A_i)_{i\in I}$ eine Mengenfamilie nichtleerer Mengen. Dann existiert eine Abbildung $w\colon I \to \bigcup_{i\in I} A_i$ mit $w(i) \in A_i$ für alle $i \in I$.

Eine solche Abbildung w heißt *Auswahlfunktion* für die Mengenfamilie $(A_i)_{i\in I}$. Die Auswahlfunktion w wählt ein Element $w(i)$ aus jedem A_i aus. Dieses Axiom scheint sehr intuitiv zu sein, und eigentlich kann man es aus den anderen Axiomen beweisen, wenn die Indexmenge I endlich ist. Es gibt viele hilfreiche Folgerungen des Auswahlaxioms:

Proposition 1.4.13 (Folgerungen des Auswahlaxioms)**.**

(i) *Sei $f\colon X \to Y$ eine surjektive Abbildung. Dann existiert ein* Schnitt *von f, d.h., eine Abbildung $s\colon Y \to X$ mit $f \circ s = \mathrm{id}_Y$.*

(ii) *Sei $(A_i)_{i\in I}$ eine Mengenfamilie nichtleerer Mengen. Dann ist das Produkt $\prod_{i\in I} A_i$ nicht leer.*

Beweis. Zu (i). Wir betrachten die Mengenfamilie $(f^{-1}(\{y\}))_{y\in Y}$, deren Vereinigung gleich X ist. Da f surjektiv ist, ist kein Urbild $f^{-1}(\{y\})$ leer. Nach dem Auswahlaxiom existiert also eine Abbildung $s\colon Y \to X$ mit $s(y) \in f^{-1}(\{y\})$ für alle $y \in Y$, d.h., $f \circ s = \mathrm{id}_Y$.

Zu (ii). Sei w eine Auswahlfunktion für die Familie $(A_i)_{i\in I}$. Dann $(w(i))_{i\in I}$ ist ein Element des Produkts $\prod_{i\in I} A_i$. □

In diesem Abschnitt erklären wir eine weitere Folgerung des Auswahlaxioms, die viele Anwendungen in der gesamten Mathematik hat: das sogenannte *Zornsche Lemma*. Dazu benötigen wir ein paar Definitionen zu geordneten Mengen.

Definition 1.4.14 (kleinstes/größtes Element, minimales/maximales Element, untere/obere Schranke)**.** Sei $(X, \leq)$ eine partiell geordnete Menge und $A \subset X$ eine Teilmenge.

- Ein *kleinstes Element* von A ist ein Element $a \in A$ mit $a \leq b$ für alle $b \in A$.
- Ein *größtes Element* von A ist ein Element $a \in A$ mit $b \leq a$ für alle $b \in A$.
- Ein *minimales Element* von A ist ein Element $a \in A$ mit folgender Eigenschaft: Für alle $b \in A$ mit $b \leq a$ gilt $b = a$.
- Ein *maximales Element* von A ist ein Element $a \in A$ mit folgender Eigenschaft: Für alle $b \in A$ mit $a \leq b$ gilt $b = a$.
- Eine *untere Schranke* von A ist ein Element $x \in X$ mit $x \leq a$ für alle $a \in A$.
- Eine *obere Schranke* von A ist ein Element $x \in X$ mit $a \leq x$ für alle $a \in A$.

Bemerkung 1.4.15. Wenn A ein kleinstes Element a besitzt, dann ist auch a das eindeutige minimale Element von A sowie eine untere Schranke von A. Aber im Allgemeinen ist ein minimales Element kein kleinstes Element, und A kann mehrere minimale Elemente enthalten. Falls $(X, \leq)$ eine *total* geordnete Menge ist, gibt es aber keinen Unterschied zwischen „kleistes Element“ und „minimales Element“.

Definition 1.4.16 (Kette)**.** Sei $(X, \leq)$ eine partiell geordnete Menge. Eine *Kette* in X ist eine Teilmenge $K \subset X$, so dass die Einschränkung von $\leq$ auf K eine totale Ordnung ist.

Definition 1.4.17 (Wohlordnung)**.** Eine *Wohlordnung* auf einer Menge X ist eine partielle Ordnung $\leq$ mit folgender Eigenschaft: Jede nichtleere Teilmenge $A \subset X$ besitzt ein kleinstes Element. Man sagt dann auch, dass $(X, \leq)$ eine wohlgeordnete Menge ist.

Beispiel 1.4.18. Das Wohlordnungsprinzip 1.2.19 ist genau die Aussage, dass $(\mathbb{N}, \leq)$ eine wohlgeordnete Menge ist. Wenn man $\mathbb{N}$ ein neues Element ∞ hinzufügt, so dass $n \leq \infty$ für alle $n \in \mathbb{N}$, dann ist $(\mathbb{N} \cup \{\infty\}, \leq)$ auch eine wohlgeordnete Menge.

Bemerkung 1.4.19. Eine Wohlordnung ist insbesondere eine totale Ordnung, da jede Teilmenge der Gestalt $\{x, y\}$ ein kleinstes Element besitzt. Jede Teilmenge einer wohlgeordneten Menge ist auch wohlgeordnet.

Beispiel 1.4.20. Die total geordneten Mengen $(\mathbb{Z}, \leq)$, $(\mathbb{Q}, \leq)$ und $(\mathbb{R}, \leq)$ sind *nicht* wohlgeordnet, weil sie selbst kein kleinstes Element besitzen. Die total geordneten Mengen $(\mathbb{Q}_{\geq 0}, \leq)$ und $(\mathbb{R}_{\geq 0}, \leq)$ sind auch nicht wohlgeordnet. Zum Beispiel besitzen die Teilmengen $\mathbb{Q}_{>0}$ und $\mathbb{R}_{>0}$ kein kleinstes Element.

Satz 1.4.21 (Das Zornsche Lemma)**.** *Sei $(X, \leq)$ eine partiell geordnete Menge, in der jede Kette eine obere Schranke besitzt. Dann besitzt $(X, \leq)$ ein maximales Element.*

Der Beweis ist ziemlich kompliziert, und es ist jetzt nicht wichtig, ihn zu verstehen.

**Beweis.* (Widerspruchsbeweis.) Wir nehmen an, dass X kein maximales Element besitzt. Sei $\mathcal{W} \subset \mathcal{P}(X)$ die Menge aller Teilmengen von X, die bzgl. $\leq$ wohlgeordnet sind. Insbesondere ist jedes $K \in \mathcal{W}$ eine Kette. Zu jedem $K \in \mathcal{W}$ sei $X_{\geq K}$ die Menge aller oberen Schranken von K, d.h., aller $x \in X$ mit $k \leq x$ für alle $k \in K$.

Nach Voraussetzung ist $X_{\geq K}$ nicht leer. Behauptung: $X_{\geq K} \setminus K$ ist nicht leer. Sonst wäre jede obere Schranke von K in K liegen. Da K total geordnet ist, würde es daraus folgen, dass jedes $x \in X_{\geq K}$ ein maximales Element von X ist, im Widerspruch zur Annahme.

Also besteht die Mengenfamilie $(X_{\geq K} \setminus K)_{K \in \mathcal{W}}$ aus nichtleeren Mengen. Nach dem Auswahlaxiom existiert eine Abbildung $s \colon \mathcal{W} \to X$ mit $s(K) \in X_{\geq K} \setminus K$ für alle $K \in \mathcal{W}$. Anders gesagt, s wählt zu jedem $K \in \mathcal{W}$ eine obere Schranke $s(K)$, die außerhalb von K liegt.

Eine Teilmenge $K \subset X$ nennen wir *s-induktiv*, falls $K \in \mathcal{W}$ und $x = s(K_{<x})$ für alle $x \in K$, wobei

$$K_{<x} := \{y \in K \mid y \leq x \text{ und } y \neq x\}.$$

Zum Beispiel, die leere Teilmenge ist s-induktiv, $\{s(\varnothing)\}$ ist s-induktiv, und ist K s-induktiv, so ist $K \cup \{s(K)\}$ s-induktiv. Sei $V \subset X$ die Vereinigung aller s-induktiven Teilmengen von X. Wir zeigen im Folgenden, dass V s-induktiv ist. Dann ist $V \cup \{s(V)\}$ eine s-induktive Menge, die keine Teilmenge von V ist, im Widerspruch zur Definition von V. Damit wird der Satz bewiesen.

Ist A eine Teilmenge von $K \in \mathcal{W}$, so schreibt man $A \prec K$, falls A nach unten abgeschlossen ist, das heißt: Für alle $a \in A$ und $k \in K$, ist $k \leq a$, so ist $k \in A$.

Behauptung. Seien $K, L \subset X$ s-induktive Teilmengen. Dann gilt $K \prec L$ oder $L \prec K$.

Mithilfe dieser Behauptung können wir beweisen, dass V s-induktiv ist:

- *V ist wohlgeordnet.* Sei $A \subset V$ eine nichtleere Teilmenge. Nach Definition von V existiert eine s-induktive Menge K, so dass $A \cap K \neq \varnothing$. Da K wohlgeordnet ist, existiert ein kleinstes Element $a \in A \cap K$. Wir beweisen, dass a ein kleinstes Element von A ist. Sei $b \in A$ ein beliebiges Element. Liegt b auch in K, so ist $a \leq b$. Andernfalls, sei L eine s-induktive Menge mit $b \in L \setminus K$. Nach der Behauptung gilt dann $K \prec L$, und daher $a \leq b$.

- *V ist s-induktiv.* Sei $x \in V$ und sei K eine s-induktive Teilmenge mit $x \in K$. Aus der Behauptung folgt $V_{<x} = K_{<x}$. Also $s(V_{<x}) = s(K_{<x}) = x$.

Beweis der Behauptung. Nach Definition von $\prec$, eine beliebiege Vereinigung von $\prec K$ Teilmengen ist wieder $\prec K$. Sei A die Vereinigung aller Mengen, die $\prec K$ *und* $\prec L$ sind. Dann ist A wieder $\prec K$ und $\prec L$. Zu zeigen ist, dass $A = K$ oder $A = L$. Angenommen, $A \neq K$ und $A \neq L$. Da K wohlgeordnet ist, existiert ein kleinstes Element $k \in K \setminus A$; insbesondere gilt $K_{<k} \subset A$. Da A in K nach unten abgeschlossen ist, gelten auch $A \cup \{k\} \prec K$ und $A \subset K_{<k}$, also $A = K_{<k}$. Insbesondere ist $s(A) = s(K_{<k}) = k$, und daher $A \cup \{s(A)\} \prec K$. Wenn wir in diesem Argument K durch L ersetzen, erhalten wir $A \cup \{s(A)\} \prec L$. Aber das steht im Widerspruch zur Definition von A, da $A \cup \{s(A)\}$ keine Teilmenge von A ist. □

Korollar 1.4.22 (Wohlordnungssatz)**.** *Sei X eine beliebige Menge. Dann existiert auf X eine Wohlordnung.*

**Beweis.* Dieser Beweis ist eine typische Anwendung des Zornschen Lemmas. Wir betrachten folgende Menge:

$$W = \{(Y, R) \mid Y \subset X \text{ und } R \subset Y \times Y \text{ ist eine Wohlordnung auf } Y\} \subset \mathcal{P}(X) \times \mathcal{P}(X \times X),$$

und definieren auf W die folgende Relation $\prec$:

$$(Y, R) \prec (Y', R') \iff Y \subset Y',\ R = R' \cap (Y \times Y) \\ \text{und } Y \text{ ist in } Y' \text{ bzgl. } R' \text{ nach unten abgeschlossen.}$$

Man kann leicht nachprüfen, dass $\prec$ eine partielle Ordnung auf W ist. Als nächstes überprüfen wir, dass $(W, \prec)$ der Bedingung des Zornschen Lemmas genügt, d.h., dass jede Kette in W eine obere Schranke besitzt. Sei $K \subset W$ eine Kette. Wir setzen

$$Y_\infty = \bigcup_{(Y,R)\in K} Y \quad \text{und} \quad R_\infty = \bigcup_{(Y,R)\in K} R.$$

Dann ist das Paar (Y_∞, R_∞) ein Element von W, und es ist eine obere Schranke von K:

- *R_∞ ist eine Wohlordnung auf Y_∞.* Sei $A \subset Y_\infty$ eine nichtleere Teilmenge. Es gibt $(Y, R) \in K$, so dass $A \cap Y \neq \varnothing$. Sei a das kleinste Element von $A \cap Y$ bzgl. R. Dann ist a eigentlich das kleinste Element von A bzgl. R_∞. Denn sei $b \in A$, und sei $(Y', R') \in K$ mit $b \in Y'$. Falls b auch in Y liegt, gilt $aR_\infty b$. Andernfalls, da K eine Kette ist, gilt $(Y, R) \prec (Y', R')$ und $b \in Y' \setminus Y$. Da Y in Y' nach unten abgeschlossen ist, gilt dann auch $aR_\infty b$.
- *Für jedes $(Y, R) \in K$ gilt $(Y, R) \prec (Y_\infty, R_\infty)$.* Die Inklusion $R \subset R_\infty \cap (Y \times Y)$ ist klar. Zur anderen Inklusion, sei $(y, z) \in R_\infty \cap (Y \times Y)$. Es existiert dann $(Y', R') \in K$ mit $yR'z$. Da K eine Kette ist, gilt $R' \subset R$ oder $R = R' \cap (Y \times Y)$. In beiden Fällen folgt yRz. Es bleibt zu zeigen, dass Y in Y_∞ nach unten abgeschlossen ist. Sei $y \in Y$ und $z \in Y_\infty$ mit $zR_\infty y$, und sei $(Y', R') \in K$ mit $z \in Y'$. Falls $(Y', R') \prec (Y, R)$ ist $z \in Y$. Andernfalls, Y ist in Y' enthalten und nach unten abgeschlossen, und somit ist auch $z \in Y$.

Nach dem Zornschen Lemma besitzt W ein maximales Element (Y, R). Es bleibt zu zeigen, dass $Y = X$; dann ist R eine Wohlordnung auf ganz X, wie gewünscht. Sei also $x \in X$. Auf $Y \cup \{x\}$ kann man folgende Relation R' definieren:

$$(y, z) \in R' \iff (y, z) \in R \text{ oder } (z \notin Y \text{ und } z = x).$$

Dann ist R' eine Wohlordnung auf $Y \cup \{x\}$, und $(Y, R) \prec (Y \cup \{x\}, R')$. Da (Y, R) maximal bzgl. $\prec$ ist, erhalten wir $Y = Y \cup \{x\}$, d.h., $x \in Y$. □

Kapitel 2

Gruppen und Körper

2.1 Gruppen

Definition 2.1.1 (Gruppe)**.** Eine *Gruppe* ist ein Paar $(G, \cdot)$, bestehend aus einer Menge G und einer Abbildung

$$\cdot : G \times G \to G, \quad (g, h) \mapsto g \cdot h$$

(die *Verknüpfung* der Gruppe), mit folgenden Eigenschaften:

(i) Die Verknüpfung ist *assoziativ*, d.h., für alle $g, h, k \in G$ gilt

$$g \cdot (h \cdot k) = (g \cdot h) \cdot k.$$

(ii) Es existiert ein *neutrales Element* bzgl. $\cdot$, d.h., ein Element $e \in G$ so dass

$$e \cdot g = g \quad \text{und} \quad g \cdot e = g$$

für alle $g \in G$.

(iii) Jedes $g \in G$ besitzt ein *inverses Element*, d.h., ein Element $h \in G$ so dass

$$g \cdot h \quad \text{und} \quad h \cdot g$$

neutrale Elemente sind.

Beispiel 2.1.2. $(\mathbb{Z}, +)$ ist eine Gruppe: Die Addition von ganzen Zahlen ist assoziativ, $0 \in \mathbb{Z}$ ist ein neutrales Element bzgl. $+$, und $-n$ ist ein inverses Element von n. In ähnlicher Weise, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$ und $(\mathbb{C}, +)$ sind Gruppen.

Andererseits ist $(\mathbb{N}, +)$ keine Gruppe. Die Bedingungen (i) und (ii) der Definition 2.1.1 sind erfüllt, aber (iii) nicht: Eine positive natürliche Zahl hat kein inverses Element bzgl. $+$ in $\mathbb{N}$.

Proposition 2.1.3 (Eindeutigkeit von neutralen und inversen Elementen)**.** *Sei $(G, \cdot)$ eine Gruppe.*

(i) *Sind $e, e' \in G$ neutrale Elemente bzgl. $\cdot$, so gilt $e = e'$.*

(ii) *Sei $g \in G$. Sind $h, h' \in G$ inverse Elemente von g, so gilt $h = h'$.*

Beweis. Zu (i). Seien $e, e' \in G$ neutrale Elemente. Dann

$$e = e \cdot e' = e',$$

wobei die erste Gleichung gilt, weil e' ein neutrales Element ist, und die zweite Gleichung gilt, weil e ein neutrales Element ist.

Zu (ii). Seien $h, h' \in G$ inverse Elemente von g, und sei $e \in G$ das eindeutige neutrale Element (nach (i)). Dann

$$\begin{aligned} h &= h \cdot e && (e \text{ neutral}) \\ &= h \cdot (g \cdot h') && (h' \text{ invers zu } g) \\ &= (h \cdot g) \cdot h' && (\cdot \text{ assoziativ}) \\ &= e \cdot h' && (h \text{ invers zu } g) \\ &= h' && (e \text{ neutral}). \end{aligned}$$ □

Notation 2.1.4. Wegen Proposition 2.1.3 darf man „*das* neutrale Element“ und „*das* inverse Element von g“ sagen, da die entsprechenden mathematischen Objekte eindeutig sind. Das neutrale Element einer Gruppe wird mit e oder 1 bezeichnet, und das inverse Element von einem Element g wird mit g^{-1} bezeichnet. Das Symbol $\cdot$ schreibt man normalerweise gar nicht, d.h., man schreibt eher gh statt $g \cdot h$.

Bemerkung 2.1.5. In einer Gruppe gilt $(gh)^{-1} = h^{-1}g^{-1}$ (Reihenfolge beachten!). Nach der Eindeutigkeit des inversen Element genügt es zu zeigen, dass $h^{-1}g^{-1}$ ein inverses Element von gh ist. Tatsächlich gilt:

$$(h^{-1}g^{-1})(gh) = h^{-1}((g^{-1}g)h) = h^{-1}(eh) = h^{-1}h = e,$$

und ebenso $(gh)(h^{-1}g^{-1}) = e$.

Notation 2.1.6. Wegen der Assoziativität der Verknupfung in einer Gruppe $(G, \cdot)$, darf man unmissverständlich $g \cdot h \cdot k$ schreiben: Es macht keinen Unterschied, ob wir die Klammern um $g \cdot h$ oder um $h \cdot k$ setzen. Allgemeiner, für jede endliche Liste von Elementen $g_1, g_2, \ldots, g_n \in G$, darf man das Produkt $g_1 \cdot g_2 \cdot \ldots \cdot g_n \in G$ ohne Klammern schreiben. Man schreibt auch

$$\prod_{i=1}^{n} g_i := g_1 \cdot g_2 \cdot \ldots \cdot g_n.$$

Wenn $n = 0$ verstehen wir das „leere Produkt“ als das neutrale Element $e \in G$. Wenn alle Elemente g_i dasselbe Element g sind, schreibt man einfach g^n für das n-fache Produkt von g mit sich selbst. Man setzt auch $g^0 := e$ und $g^{-n} := (g^n)^{-1}$. So wird die Potenz g^n für alle ganzen Zahlen $n \in \mathbb{Z}$ definiert.

Proposition 2.1.7 (Eigenschaften der Potenzen)**.** *Sei $(G, \cdot)$ eine Gruppe und sei $g \in G$.*

(i) *Für alle $m, n \in \mathbb{Z}$ gilt $g^m \cdot g^n = g^{m+n}$.*

(ii) *Für alle $m, n \in \mathbb{Z}$ gilt $(g^n)^m = g^{mn}$.*

Beweis. Zu (i). Wir betrachten zunächst den Fall $m \in \mathbb{N}$, in dem wir Induktion über m verwenden.

- *Induktionsanfang.* Wenn $m = 0$, gilt $g^0 \cdot g^n = e \cdot g^n = g^n = g^{0+n}$.
- *Induktionsschritt.* Es gilt

  $$g^{m+1} \cdot g^n = g \cdot g^m \cdot g^n = g \cdot g^{m+n} = g^{m+1+n},$$

  wobei die zweite Gleichung aus der Induktionsvoraussetzung folgt.

Damit ist (i) bewiesen für alle $m \geq 0$. Der Fall $n \geq 0$ wird mit einem ähnlichen Argument erledigt. Wenn $m < 0$, dann gilt

$$g^m \cdot g^n = (g^{-n} \cdot g^{-m})^{-1} = (g^{-(m+n)})^{-1} = g^{m+n}.$$

Dabei haben wir die Formel $(gh)^{-1} = h^{-1}g^{-1}$ in der ersten Gleichung und den Fall $n \geq 0$ (mit $-m$ anstelle von n) in der zweiten Gleichung verwendet.

Zu (ii). Wir beweisen zunächst den Fall $m \in \mathbb{N}$ durch Induktion über m.

- *Induktionsanfang.* Wenn $m = 0$, gilt $(g^n)^0 = e = g^0 = g^{0n}$.
- *Induktionsschritt.* Es gilt

$$(g^n)^{m+1} = g^n \cdot (g^n)^m = g^n \cdot g^{mn} = g^{n+mn} = g^{(m+1)n},$$

wobei die zweite Gleichung aus der Induktionsvoraussetzung folgt, und die dritte aus (i).

Damit ist (ii) bewiesen für alle $m \geq 0$. Wenn $m < 0$ ist der Beweis mit der folgenden Berechnung abgeschlossen:

$$(g^n)^m = ((g^n)^{-m})^{-1} = (g^{-mn})^{-1} = g^{mn}. \qquad \square$$

Bemerkung 2.1.8. Sei $(G, \cdot)$ eine Gruppe, $g, h \in G$ und $n \in \mathbb{Z}$. Im Allgemeinen gilt die Gleichung $(gh)^n = g^n h^n$ nur für $n = 0$ und $n = 1$. Zum Beispiel, $(gh)^2 = ghgh$ muss nicht gleich $g^2h^2 = gghh$ sein. Diese Gleichung gilt aber für alle $n \in \mathbb{Z}$, wenn die Gruppe abelsch ist (siehe Definition 2.1.9).

Definition 2.1.9 (abelsche Gruppe)**.** Eine Gruppe $(G, \cdot)$ heißt *abelsch*, falls ihre Verknüpfung *kommutativ* ist, d.h., für alle $g, h \in G$ gilt

$$g \cdot h = h \cdot g.$$

Beispiel 2.1.10. Die Gruppe $(\mathbb{Z}, +)$ ist abelsch, da $n + m = m + n$ für alle ganzen Zahlen n, m. In ähnlicher Weise, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$ und $(\mathbb{C}, +)$ sind abelsche Gruppen.

Notation 2.1.11. Bei abelschen Gruppen verwendet man häufig die *additive Notation* statt der *multiplikativen* Notation (Notation 2.1.4): d.h., man schreibt $+$ für die Verknüpfung, 0 für das neutrale Element und $-a$ für das Inverse von a. Man schreibt auch $a-b$ als Abkürzung von $a + (-b)$. Ist $a_1, a_2, \ldots, a_n$ eine Liste von Elementen, so schreibt man

$$\sum_{i=1}^{n} a_i := a_1 + a_2 + \cdots + a_n.$$

Wenn $n = 0$ verstehen wir die „leere Summe“ als das neutrale Element 0. Wenn alle a_i gleich a sind, schreibt man einfach $n \cdot a$ oder na für diese Summe. Man setzt auch $0 \cdot a := 0$ und $(-n) \cdot a := -(n \cdot a)$. So wird $n \cdot a$ für alle ganzen Zahlen $n \in \mathbb{Z}$ definiert.

Notation 2.1.12. Oft unterdrückt man die Verknüpfung in der Notation für eine Gruppe $(G, \cdot)$. Das heißt, man sagt üblicherweise „Sei G eine Gruppe“ und nicht „Sei $(G, \cdot)$ eine Gruppe“. In diesem Fall wird die Verknüpfung standardmäßig mit $\cdot$ bezeichnet (oder mit $+$ im Fall einer abelschen Gruppe).

Bemerkung 2.1.13. Wenn wir in der Definition einer Gruppe (Definition 2.1.1) auf das dritte Axiom verzichten, erhalten wir den Begriff des *Monoids*. In einem Monoid ist das neutrale Element immer noch eindeutig bestimmt: Der Beweis von Proposition 2.1.3(i) benötigt keine inverse Elemente. Zum Beispiel ist $(\mathbb{N}, +)$ ein (abelsches) Monoid aber keine Gruppe.

2.2 Beispiele von Gruppen

2.2.1 Die ganzen Zahlen

Die natürlichen Zahlen $\mathbb{N}$ bilden *keine* Gruppe bezüglich Addition, denn die positiven natürlichen Zahlen besitzen keine inverse Elemente. Um $(\mathbb{N}, +)$ zu einer Gruppe zu erweitern, brauchen wir die negativen Zahlen hinzuzufügen. Dann erhalten wir die ganzen Zahlen $\mathbb{Z}$,

und $(\mathbb{Z}, +)$ ist eine abelsche Gruppe. In diesem Abschnitt erklären wir, wie $\mathbb{Z}$ aus $\mathbb{N}$ eigentlich konstruiert werden kann.

Die Idee ist, dass jede ganze Zahl als Differenz zweier natürlichen Zahlen n, m dargestellt werden kann. Auf diese Weise bestimmt jedes Paar $(n, m) \in \mathbb{N} \times \mathbb{N}$ eine ganze Zahl, nämlich $n - m$. Aber diese Darstellung ist nicht eindeutig: Es gilt $n - m = n' - m'$ genau dann, wenn $n + m' = n' + m$. Deswegen führen wir folgende Relation $\sim$ auf $\mathbb{N} \times \mathbb{N}$ ein:

$$(n, m) \sim (n', m') \iff n + m' = n' + m.$$

Man kann leicht nachprüfen, dass $\sim$ eine Äquivalenzrelation auf $\mathbb{N} \times \mathbb{N}$ ist. Als Beispiel überprüfen wir die Transitivität: Falls $(n, m) \sim (n', m')$ und $(n', m') \sim (n'', m'')$, dann gilt

$$\begin{aligned}
(n + m'') + m' &= (n + m') + m'' && \text{(Assoziativität und Kommutativität von } +) \\
&= (n' + m) + m'' && \text{(da } (n, m) \sim (n', m')) \\
&= (n' + m'') + m && \text{(Assoziativität und Kommutativität von } +) \\
&= (n'' + m') + m && \text{(da } (n', m') \sim (n'', m'')) \\
&= (n'' + m) + m', && \text{(Assoziativität und Kommutativität von } +)
\end{aligned}$$

und daher $n + m'' = n'' + m$, d.h., $(n, m) \sim (n'', m'')$ (hierbei wurde benutzt, dass für naturliche Zahlen n, m, p gilt: $n + p = m + p \Rightarrow n = m$).

Wir setzen nun

$$\mathbb{Z} := (\mathbb{N} \times \mathbb{N})/\sim.$$

Mit dieser Definition ist $\mathbb{N}$ keine Teilemenge von $\mathbb{Z}$, was eher unangenehm ist. In der Praxis möchten wir auf jeden Fall $\mathbb{N}$ als Teilmenge von $\mathbb{Z}$ auffassen. Hierzu betrachten wir die Abbildung

$$i \colon \mathbb{N} \to \mathbb{Z}, \quad n \mapsto [(n, 0)].$$

Die Abbildung i ist *injektiv*. Denn seien $n, n' \in \mathbb{N}$ mit $[(n, 0)] = [(n', 0)]$; dann gilt $(n, 0) \sim (n', 0)$, d.h. $n + 0 = n' + 0$, also $n = n'$. Insbesondere induziert i eine Bijektion zwischen $\mathbb{N}$ und seinem Bild $i(\mathbb{N}) \subset \mathbb{Z}$, und so können wir $\mathbb{N}$ mit einer Teilmenge von $\mathbb{Z}$ identifizieren. Dementsprechend führen wir folgende Notation ein:

Notation 2.2.1. Sei $n \in \mathbb{N}$. Die Äquivalenzklasse $[(n, 0)] \in \mathbb{Z}$ bezeichnen wir auch mit n. Die Äquivalenzklasse $[(0, n)] \in \mathbb{Z}$ bezeichnen wir mit $-n$.

Wir möchten jetzt die arithmetischen Operationen $+$ und $\cdot$ von $\mathbb{N}$ auf $\mathbb{Z}$ fortsetzen, sowie die totale Ordnungsrelation $\leq$. Die Definitionen sind klar, wenn man sich die Äquivalenzklasse $[(n, m)]$ als die Differenz $n - m$ vorstellt:

$$\begin{aligned}
[(n_1, m_1)] + [(n_2, m_2)] &= [(n_1 + n_2, m_1 + m_2)], \\
[(n_1, m_1)] \cdot [(n_2, m_2)] &= [(n_1 n_2 + m_1 m_2, n_1 m_2 + n_2 m_1)], \\
[(n_1, m_1)] \leq [(n_2, m_2)] &\iff n_1 + m_2 \leq n_2 + m_1.
\end{aligned}$$

Da es Äquivalenzklassen auf der linken Seite dieser Definitionen gibt, muss man hier nachprüfen, dass alles wohldefiniert ist. Zum Beispiel, um zu zeigen, dass $+$ auf $\mathbb{Z}$ wohldefiniert ist, ist folgendes zu beweisen:

$$(n_1, m_1) \sim (n_1', m_1') \text{ und } (n_2, m_2) \sim (n_2', m_2') \implies (n_1 + n_2, m_1 + m_2) \sim (n_1' + n_2', m_1' + m_2').$$

Dies folgt unmittelbar aus der Definition von $\sim$.

Bemerkung 2.2.2. Diese Konstruktion von $\mathbb{Z}$ aus $\mathbb{N}$ ist ein Sonderfall einer allgemeineren Konstruktion, die eine Gruppe aus irgendeinem Monoid liefert.

2.2.2 Symmetrische Gruppen

Definition 2.2.3 (Permutation)**.** Sei X eine Menge. Eine *Permutation* von X ist eine bijektive Abbildung von X nach X. Die Menge aller Permutationen von X wird mit S_X bezeichnet.

Proposition 2.2.4. *Sei X eine Menge. Dann ist $(S_X, \circ)$ eine Gruppe. Falls X mindestens drei verschiedene Elemente besitzt, ist diese Gruppe nicht abelsch.*

Beweis. Komposition von Abbildungen ist assoziativ und die Identität id_X ist ein neutrales Element (Proposition 1.3.13). Jedes Element von S_X hat ein inverses Element nach Satz 1.3.23. Also ist $(S_X, \circ)$ eine Gruppe.

Zu je zwei Elementen $a, b \in X$ gibt es eine bijektive Abbildung $\tau_{a,b} \in S_X$, die wie folgt definiert wird:

$$\tau_{a,b}(x) = \begin{cases} b, & \text{falls } x = a, \\ a, & \text{falls } x = b, \\ x, & \text{andernfalls.} \end{cases}$$

Sind $a, b, c \in X$ drei verschiedene Elemente, so gilt $\tau_{b,c} \circ \tau_{a,b} \neq \tau_{a,b} \circ \tau_{b,c}$, denn:

$$\tau_{b,c}(\tau_{a,b}(a)) = c \quad \text{aber} \quad \tau_{a,b}(\tau_{b,c}(a)) = b.$$

Also ist $(S_X, \circ)$ nicht abelsch. □

Definition 2.2.5 (symmetrische Gruppe)**.** Die Gruppe $(S_X, \circ)$ heißt die *symmetrische Gruppe* von X. Falls $X = \{1, \ldots, n\}$ mit $n \in \mathbb{N}$ schreibt man S_n (auch Σ_n, $\mathfrak{S}_n$) statt S_X. Die Gruppe $(S_n, \circ)$ heißt die *symmetrische Gruppe vom Grad n*.

Beispiel 2.2.6. Die Gruppe S_2 hat genau zwei Elemente: die Identität id und die Abbildung $\tau\colon \{1, 2\} \to \{1, 2\}$, die 1 und 2 austauscht. Es gilt $\tau \circ \tau = \mathrm{id}$.

Bemerkung 2.2.7. Durch Induktion kann man leicht nachprüfen, dass die Mächtigkeit von S_n gleich $n!$ ist (wobei $0! = 1$).

Bemerkung 2.2.8 (indizierte Summen in einer abelschen Gruppe)**.** Sei A eine abelsche Gruppe (allgemeiner, ein abelsches Monoid) und $a_1, \ldots, a_n \in A$. Nach der Kommutativität von $+$ ist die Summe $\sum_{k=1}^{n} a_i$ *unabhängig* von der Reihenfolge dieser Elemente. Genauer heißt das folgendes: Für jede Permutation $\sigma \in S_n$ der Indexmenge $\{1, \ldots, n\}$ gilt:

$$\sum_{k=1}^{n} a_k = \sum_{k=1}^{n} a_{\sigma(k)}.$$

Sei nun I eine endliche Menge und $(a_i)_{i \in I}$ eine Familie von Elementen von A mit Indexmenge I. Dann kann man die Summe $\sum_{i \in I} a_i \in A$ wie folgt definieren. Man wählt eine Bijektion $f\colon \{1, \ldots, n\} \to I$ und setzt

$$\sum_{i \in I} a_i := \sum_{k=1}^{n} a_{f(k)}.$$

Die rechte Seite ist unabhängig von der Wahl von f, denn: Seien $f, g\colon \{1, \ldots, n\} \to I$ zwei Bijektionen. Dann ist $\sigma = f^{-1} \circ g$ eine Permutation von $\{1, \ldots, n\}$, und es gilt

$$\sum_{k=1}^{n} a_{f(k)} = \sum_{k=1}^{n} a_{f(\sigma(k))} = \sum_{k=1}^{n} a_{g(k)}.$$

2.3 Körper

Der Begriff des Körpers ist eine Abstraktion der arithmetischen Struktur der rationalen, reellen und komplexen Zahlen.

Definition 2.3.1 (Körper)**.** Ein *Körper* ist ein Tripel $(K, +, \cdot)$, bestehend aus einer Menge K und Abbildungen

$$\begin{aligned} +\colon K \times K \to K, &\quad (x, y) \mapsto x + y, \\ \cdot\colon K \times K \to K, &\quad (x, y) \mapsto x \cdot y, \end{aligned}$$

die als *Addition* und *Multiplikation* bezeichnet werden, mit folgenden Eigenschaften:

(i) $(K, +)$ ist eine abelsche Gruppe. Das heißt, die Verknüpfung $+$ ist assoziativ und kommutativ, sie besitzt ein neutrales Element 0, und jedes Element von K besitzt ein inverses Element bzgl. $+$.

(ii) Die Verknüpfung $\cdot$ ist assoziativ, kommutativ, und besitzt ein neutrales Element 1.

(iii) Jedes Element von $K \setminus \{0\}$ besitzt ein inverses Element bzgl. $\cdot$.

(iv) Es gilt $0 \neq 1$.

(v) Es gilt das *Distributivgesetz*: Für alle $x, y, z \in K$,

$$x \cdot (y + z) = x \cdot y + x \cdot z.$$

Beispiel 2.3.2. $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$ und $(\mathbb{C}, +, \cdot)$ sind Körper. Siehe Abschnitt 2.4 für die formalen Konstruktionen von $\mathbb{Q}$, $\mathbb{R}$ und $\mathbb{C}$.

Bemerkung 2.3.3. Da die Verknüpfung $\cdot$ kommutativ ist, gilt auch in einem Körper das umgekehrte Distributivgesetz

$$(x + y) \cdot z = x \cdot z + y \cdot z.$$

Durch Induktion kann man außerdem folgendes verallgemeinertes Distributivgesetz beweisen:

$$x \cdot \left(\sum_{i=1}^{n} y_i \right) = \sum_{i=1}^{n} x \cdot y_i.$$

Dies gilt auch für $n = 0$ nach Proposition 2.3.8(i).

Bemerkung 2.3.4. Nach Definition gilt in einem Körper $0 \neq 1$. Ein Körper enthält deshalb mindestens zwei Elemente. Eigentlich existiert ein Körper mit genau zwei Elementen, siehe Abschnitt 2.4.4.

Bemerkung 2.3.5. Wenn wir in der Definition 2.3.1 auf Axiome (iii) und (iv) verzichten, erhalten wir den Begriff des *kommutativen Ringes*. Zum Beispiel ist $(\mathbb{Z}, +, \cdot)$ ein kommutativer Ring, aber kein Körper. Kommutative Ringe spielen eine wichtige Rolle in mehreren Bereichen der Mathematik, und sie werden in der späteren Vorlesung *Kommutative Algebra* untersucht.

Notation 2.3.6. In einem Körper verwenden wir ausschließlich die additive Notation 2.1.11 für die erste Verknüpfung und die multiplikative Notation 2.1.4 für die zweite Verknüpfung. Insbesondere schreiben wir $-x$ für das inverse Element von x bzgl. $+$ und (falls $x \neq 0$) x^{-1} für das inverse Element bzgl. $\cdot$. Außerdem verwenden wir die Bruchnotation

$$\frac{x}{y} := x \cdot y^{-1},$$

falls $y \neq 0$.

Notation 2.3.7. Wie bei Gruppen unterdrückt man oft die Verknüpfungen in der Notation für einen Körper $(K, +, \cdot)$. Das heißt, man sagt üblicherweise „Sei K ein Körper“ und nicht „Sei $(K, +, \cdot)$ ein Körper“.

Proposition 2.3.8 (Rechnen in Körpern)**.** *Sei $(K, +, \cdot)$ ein Körper.*

(i) *Für alle $x \in K$ gilt $0 \cdot x = 0$.*

(ii) *Für alle $x \in K$ gilt $(-1) \cdot x = -x$. Insbesondere ist $(-1) \cdot (-1) = 1$.*

(iii) (Nullteilerfreiheit) *Sind $x, y \in K \setminus \{0\}$, so ist $x \cdot y \in K \setminus \{0\}$.*

Beweis. Zu (i). Nach dem Distributivgesetz gilt $(0 + 0) \cdot x = 0 \cdot x + 0 \cdot x$. Da $0 + 0 = 0$, erhalten wir

$$0 \cdot x = 0 \cdot x + 0 \cdot x.$$

Da $(K, +)$ eine Gruppe ist, dürfen wir $0 \cdot x$ von beiden Seiten subtrahieren. Also $0 = 0 \cdot x$.

Zu (ii). Nach der Eindeutigkeit des inversen Element (Proposition 2.1.3) genügt es zu zeigen, dass $(-1) \cdot x$ ein inverses Element von x bezüglich $+$ ist, d.h., dass $(-1) \cdot x + x = 0$. Wir berechnen:

$$\begin{aligned}
(-1) \cdot x + x &= (-1) \cdot x + 1 \cdot x && \text{(1 neutral)} \\
&= ((-1) + 1) \cdot x && \text{(Distributivität)} \\
&= 0 \cdot x \\
&= 0. && \text{(nach (i))}
\end{aligned}$$

Zu (iii). Wir beweisen die Kontraposition. Seien $x, y \in K$ mit $x \cdot y = 0$. Zu zeigen ist, dass $x = 0$ oder $y = 0$. Falls $y \neq 0$, existiert ein inverses Element y^{-1}, so dass $y \cdot y^{-1} = 1$. Also gilt

$$\begin{aligned}
x &= x \cdot 1 && \text{(1 neutral)} \\
&= x \cdot (y \cdot y^{-1}) \\
&= (x \cdot y) \cdot y^{-1} && \text{(Assoziativität)} \\
&= 0 \cdot y^{-1} && \text{(da } x \cdot y = 0\text{)} \\
&= 0, && \text{(nach (i))}
\end{aligned}$$

wie gewünscht. □

Korollar 2.3.9. *Ist $(K, +, \cdot)$ ein Körper, so ist $(K \setminus \{0\}, \cdot)$ eine abelsche Gruppe.*

Beweis. Zuerst muss man beobachten, dass die Multiplikation $(K \setminus \{0\}) \times (K \setminus \{0\})$ auf $K \setminus \{0\}$ abbildet. Dies folgt aus Proposition 2.3.8(iii). Nach Definition eines Körpers ist die Verknüpfung $\cdot$ assoziativ und kommutativ, und sie besitzt ein neutrales Element $1 \in K \setminus \{0\}$. Es bleibt zu zeigen, dass jedes $x \in K \setminus \{0\}$ ein inverses Element in $K \setminus \{0\}$ besitzt. Nach Definition eines Körpers gibt es ein inverses Element $x^{-1} \in K$. Da $x^{-1} \cdot x = 1 \neq 0$, folgt aus Proposition 2.3.8(i), dass $x^{-1} \neq 0$. □

Notation 2.3.10. Sei $(K, +, \cdot)$ ein Körper. Die Menge $K \setminus \{0\}$ wird oft mit K^* oder $K^\times$ bezeichnet, besonders wenn man diese Menge als abelsche Gruppe bzgl. $\cdot$ betrachtet.

Bemerkung 2.3.11. Wenn wir in der Definition 2.3.1 die Axiome (ii) und (iii) durch das einfachere Axiom „$(K, \cdot)$ ist eine abelsche Gruppe“ ersetzen, dann erhalten wir einen Begriff, der keine Beispiele besitzt. Denn in einem solchen K würde 0 ein Inverses 0^{-1} haben, so dass $0 \cdot 0^{-1} = 1$. Auf der anderen Seite ist $0 \cdot 0^{-1} = 0$ nach Proposition 2.3.8(i), und somit $0 = 1$, im Widerspruch zum Axiom (iv).

Zur Erinnerung (Notation 2.1.11), in der additiven Gruppe $(K, +)$ eines Körpers haben wir für alle $n \in \mathbb{Z}$ und $x \in K$ das ganzzahlige Vielfache $n \cdot x$ definiert. Insbesondere gibt es eine Abbildung

$$\begin{aligned} &\mathbb{Z} \to K, \\ &n \mapsto n \cdot 1, \end{aligned}$$

wobei $1 \in K$ das neutrale Element bzgl. $\cdot$ ist. Man schreibt oft einfach n für das Element $n \cdot 1$ von K. Wenn $K = \mathbb{Q}$, $\mathbb{R}$ oder $\mathbb{C}$, ist diese Abbildung injektiv, und zwar die gewöhnliche Inklusion von $\mathbb{Z}$ in diesen Körpern. Im Abschnitt 2.4.4 geben wir Beispiele von Körpern, in denen $n \cdot 1 = 0$ mit einem $n \neq 0$. Im Vorgriff auf diese Beispiele führen wir folgende Definition ein:

Definition 2.3.12 (Charakteristik eines Körpers)**.** Sei K ein Körper. Die *Charakteristik* von K, $\mathrm{char}(K)$, ist die wie folgt definierte natürliche Zahl:

- Falls $n \cdot 1 \neq 0$ für alle $n \in \mathbb{N} \setminus \{0\}$, ist $\mathrm{char}(K) = 0$.
- Andernfalls ist $\mathrm{char}(K)$ das kleinste $n \in \mathbb{N} \setminus \{0\}$ mit $n \cdot 1 = 0$.

In einem Körper K der Charakteristik $n \neq 0$ gilt $n \cdot x = 0$ für alle $x \in K$, denn:

$$n \cdot x = x + \cdots + x = (1 + \cdots + 1) \cdot x = (n \cdot 1) \cdot x = 0 \cdot x = 0.$$

Proposition 2.3.13. *Sei K ein Körper. Dann ist die Charakteristik von K entweder 0 oder eine Primzahl.*

Beweis. Sei $n = \mathrm{char}(K)$. Wir nehmen an, dass $n \neq 0$. Da $1 \neq 0$ in K, ist auch $n \neq 1$. Es sei $n = rs$ mit natürlichen Zahlen $r, s \in \mathbb{N}$. Zu zeigen ist, dass $r = n$ oder $s = n$. Es gilt

$$0 = n \cdot 1 = r \cdot (s \cdot 1) = (r \cdot 1) \cdot (s \cdot 1),$$

wobei die letzte Gleichung aus dem Distributivgesetz folgt. Aus der Nullteilerfreiheit von K (Proposition 2.3.8(iii)) folgt $r \cdot 1 = 0$ oder $s \cdot 1 = 0$. Aber n ist nach Definition die kleinste natürliche Zahl mit $n \cdot 1 = 0$. Es muss also $r = n$ oder $s = n$ sein, wie gewünscht. □

Beispiel 2.3.14. In einem Körper K der Charakteristik 2 gilt $(a + b)^2 = a^2 + b^2$ für alle $a, b \in K$. Denn in einem beliebigen Körper gilt $(a+b)^2 = a^2 + 2ab + b^2$, wobei $2ab = ab + ab$ (dies folgt aus der Distributivität von $\cdot$ über $+$ und der Kommutativität von $\cdot$), und $2ab = 0$ in K.

Allgemeiner, wenn die Charakteristik von K eine Primzahl p ist, folgt aus dem binomischen Lehrsatz, dass $(a + b)^p = a^p + b^p$ für alle $a, b \in K$.

2.4 Beispiele von Körpern

2.4.1 Die rationalen Zahlen

Die ganzen Zahlen $\mathbb{Z}$ bilden *keinen* Körper bezüglich Addition und Multiplikation, denn die ganzen Zahlen außer ± 1 besitzen keine inverse Elemente bezüglich Multiplikation. Um $(\mathbb{Z}, +, \cdot)$ zu einem Körper zu erweitern, brauchen wir Brüche ganzer Zahlen hinzuzufügen. Dann erhalten wir die rationalen Zahlen $\mathbb{Q}$, und $(\mathbb{Q}, +, \cdot)$ ist ein Körper. In diesem Abschnitt erklären wir, wie $\mathbb{Q}$ aus $\mathbb{Z}$ eigentlich konstruiert werden kann.

Die Konstruktion von $\mathbb{Q}$ aus $\mathbb{Z}$ ist ganz analog zu der Konstruktion von $\mathbb{Z}$ aus $\mathbb{N}$, wobei die Multiplikation die Rolle der Addition übernimmt. Jede rationale Zahl kann als Bruch zweier ganzen Zahlen a, b mit $b \neq 0$ dargestellt werden, aber diese Darstellung ist nicht

eindeutig: Es gilt $a/b = a'/b'$ genau dann, wenn $ab' = a'b$. Deswegen führen wir folgende Äquivalenzrelation $\sim$ auf $\mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$ ein:

$$(a, b) \sim (a', b') \iff ab' = a'b,$$

und setzen wir:

$$\mathbb{Q} := (\mathbb{Z} \times (\mathbb{Z} \setminus \{0\}))/\sim.$$

Durch die injektive Abbildung

$$\begin{aligned} &\mathbb{Z} \to \mathbb{Q}, \\ &a \mapsto [(a, 1)], \end{aligned}$$

können wir $\mathbb{Z}$ mit einer Teilmenge von $\mathbb{Q}$ identifizieren.

Weiter definieren wir die arithmetischen Operationen $+$ und $\cdot$ und die totale Ordnungsrelation $\leq$ wie folgt (nach den üblichen Bruchregeln):

$$\begin{aligned} [(a_1, b_1)] + [(a_2, b_2)] &= [(a_1 b_2 + a_2 b_1, b_1 b_2)], \\ [(a_1, b_1)] \cdot [(a_2, b_2)] &= [(a_1 a_2, b_1 b_2)], \\ [(a_1, b_1)] \leq [(a_2, b_2)] &\iff \begin{cases} a_1 b_2 \leq a_2 b_1, & \text{falls } b_1 b_2 > 0, \\ a_1 b_2 \geq a_2 b_1, & \text{falls } b_1 b_2 < 0. \end{cases} \end{aligned}$$

Man sollte natürlich nachprüfen, dass $+$, $\cdot$ und $\leq$ durch diese Formeln wohldefiniert sind, was nicht schwierig ist. Außerdem setzen $+$, $\cdot$ und $\leq$ auf $\mathbb{Q}$ die entsprechenden Verknüpfungen bzw. Relation von $\mathbb{Z}$ fort.

2.4.2 Die reellen Zahlen

Im Vergleich zu den Konstruktionen von $\mathbb{Z}$ und $\mathbb{Q}$ ist die Konstruktion der reellen Zahlen $\mathbb{R}$ deutlich komplizierter, und sie benötigt etwas *analytisch*.

Die Idee ist, dass es in der Menge $\mathbb{Q}$ der rationalen Zahlen „Löcher“ bezüglich der gewöhnlichen Ordnung $\leq$ gibt, und dass wir die reellen Zahlen erhalten, indem wir diese Löcher ausfüllen. Zum Beispiel gibt es keine rationale Zahl x mit $x^2 = 2$ (Satz 1.1.15). Trotzdem kann man die rationalen Zahlen in zwei Teilmengen unterteilen, je nachdem x^2 kleiner als oder größer als 2 ist:

$$\begin{aligned} \mathbb{Q}_{<\sqrt{2}} &:= \{x \in \mathbb{Q} \mid x < 0 \text{ oder } x^2 < 2\}, \\ \mathbb{Q}_{>\sqrt{2}} &:= \{x \in \mathbb{Q} \mid x \geq 0 \text{ und } x^2 > 2\}. \end{aligned}$$

Dann ist jede Zahl in $\mathbb{Q}_{<\sqrt{2}}$ kleiner als jede Zahl in $\mathbb{Q}_{>\sqrt{2}}$, und es gilt:

$$\mathbb{Q}_{<\sqrt{2}} \cup \mathbb{Q}_{>\sqrt{2}} = \mathbb{Q}, \quad \mathbb{Q}_{<\sqrt{2}} \cap \mathbb{Q}_{>\sqrt{2}} = \varnothing.$$

Aber $\mathbb{Q}_{<\sqrt{2}}$ hat kein größtes Element, und $\mathbb{Q}_{>\sqrt{2}}$ hat kein kleinstes Element. In diesem Sinne gibt es ein „Loch“ zwischen $\mathbb{Q}_{<\sqrt{2}}$ und $\mathbb{Q}_{>\sqrt{2}}$, und dieses Loch wird durch die reelle Zahl $\sqrt{2}$ ausgefüllt.

Wir erklären jetzt die zwei häufigsten Konstruktionen von $\mathbb{R}$: die „Dedekindschen“ reellen Zahlen $\mathbb{R}_{\text{Dedekind}}$ und die „Cauchyschen“ reellen Zahlen $\mathbb{R}_{\text{Cauchy}}$. Die zweite Konstruktion wird in der Volesung *Analysis I* ausführlicher behandelt. Es gibt noch weitere Konstruktionen von $\mathbb{R}$, aber die gewählte Konstruktion ist nicht wichtig, solange die Ausgabe ein *vollständiger angeordneter Körper* ist.

Definition 2.4.1 (Dedekindscher Schnitt)**.** Ein *Dedekindscher Schnitt* auf $\mathbb{Q}$ ist eine Teilmenge $A \subset \mathbb{Q}$ mit folgenden Eigenschaften:

- $A \neq \varnothing$ und $A \neq \mathbb{Q}$.

- A ist in $\mathbb{Q}$ nach unten abgeschlossen, d.h., ist $a \in A$ und ist $b \leq a$, so ist $b \in A$.
- A enthält kein größtes Element.

Ein Dedekindscher Schnitt heißt *rational*, falls das Komplement von A ein kleinstes Element besitzt.

Man kann die reellen Zahlen als die Menge aller Dedekindschen Schnitte auf $\mathbb{Q}$ definieren:

$$\mathbb{R}_{\text{Dedekind}} := \{A \mid A \text{ ist ein Dedekindscher Schnitt auf } \mathbb{Q}\} \subset \mathcal{P}(\mathbb{Q}).$$

Mit dieser Definition gibt es eine injektive Abbildung

$$\mathbb{Q} \to \mathbb{R}_{\text{Dedekind}}, \quad q \mapsto \mathbb{Q}_{<q} = \{x \in \mathbb{Q} \mid x < q\},$$

deren Bild genau aus den rationalen Schnitten besteht. Die Addition auf $\mathbb{R}_{\text{Dedekind}}$, sowie die Ordnungsrelation $\leq$, kann man auch leicht definieren:

$$\begin{aligned} A + B &= \{a + b \mid a \in A \text{ und } b \in B\}, \\ A \leq B &\iff A \subset B, \end{aligned}$$

wobei $+$ auf der rechten Seite die gewöhnliche Verknüpfung auf $\mathbb{Q}$ ist. Die Multiplikation ist etwas mühsamer zu definieren. Falls $\mathbb{Q}_{<0} \subset A$ und $\mathbb{Q}_{<0} \subset B$ (d.h., A und B entsprechen ≥ 0 reellen Zahlen), setzen wir

$$A \cdot B = \mathbb{Q}_{<0} \cup \{a \cdot b \mid a \in A \cap \mathbb{Q}_{\geq 0} \text{ und } b \in B \cap \mathbb{Q}_{\geq 0}\}.$$

Mithilfe der üblichen Vorzeichenregeln kann man diese Multiplikation auf ganz $\mathbb{R}_{\text{Dedekind}}$ fortsetzen. Man kann dann beweisen, dass $(\mathbb{R}_{\text{Dedekind}}, +, \cdot)$ ein Körper ist, und dass die obige injektive Abbildung $\mathbb{Q} \to \mathbb{R}_{\text{Dedekind}}$ identifiziert $\mathbb{Q}$ mit einem Teilkörper von $\mathbb{R}_{\text{Dedekind}}$.

Definition 2.4.2 (Cauchyfolge in $\mathbb{Q}$)**.** Eine Folge $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{Q}$ heißt *Cauchyfolge* wenn folgendes gilt: Zu jeder rationalen Zahl $\varepsilon > 0$ gibt es eine natürliche Zahl $N \in \mathbb{N}$, so dass für alle $n, m \geq N$ gilt $|x_n - x_m| < \varepsilon$. Sei Cauchy($\mathbb{Q}$) die Menge aller Cauchyfolgen in $\mathbb{Q}$.

Zum Beispiel ist die konstante Folge $(q)_{n\in\mathbb{N}}$ mit $q \in \mathbb{Q}$ eine Cauchyfolge. Man definiert eine Äquivalenzrelation $\sim$ auf Cauchy($\mathbb{Q}$) wie folgt: $(x_n)_{n\in\mathbb{N}} \sim (y_n)_{n\in\mathbb{N}}$ genau dann, wenn folgendes gilt: Zu jeder rationalen Zahl $\varepsilon > 0$ gibt es eine natürliche Zahl $N \in \mathbb{N}$, so dass für alle $n \geq N$ gilt $|x_n - y_n| < \varepsilon$. Man kann dann die reellen Zahlen als die Quotientenmenge

$$\mathbb{R}_{\text{Cauchy}} := \text{Cauchy}(\mathbb{Q})/\sim$$

definieren. Mit dieser Definition gibt es eine injektive Abbildung

$$\mathbb{Q} \to \mathbb{R}_{\text{Cauchy}}, \quad q \mapsto [(q)_{n\in\mathbb{N}}],$$

(Zur Injektivität: Falls $q \neq q'$, sei $\varepsilon = |q - q'|/2 \in \mathbb{Q}$. Dann $|q - q'| \not< \varepsilon$. Aus der Definition von $\sim$ folgt, dass $(q)_{n\in\mathbb{N}} \not\sim (q')_{n\in N}$.) Die Addition und Multiplikation auf $\mathbb{R}_{\text{Cauchy}}$ werden so definiert:

$$\begin{aligned} [(x_n)_{n\in\mathbb{N}}] + [(y_n)_{n\in\mathbb{N}}] &= [(x_n + y_n)_{n\in\mathbb{N}}], \\ [(x_n)_{n\in\mathbb{N}}] \cdot [(y_n)_{n\in\mathbb{N}}] &= [(x_n \cdot y_n)_{n\in\mathbb{N}}]. \end{aligned}$$

Hier ist es notwendig, ein paar Tatsachen nachzuprüfen: erstens, dass $(x_n + y_n)_{n\in\mathbb{N}}$ bzw. $(x_n \cdot y_n)_{n\in\mathbb{N}}$ eine Cauchyfolge ist, wenn $(x_n)_{n\in\mathbb{N}}$ und $(y_n)_{n\in\mathbb{N}}$ Cauchyfolgen sind, und zweitens, dass $+$ und $\cdot$ auf $\sim$-Äquivalenzklassen wohldefiniert sind. Man kann auch die strenge Ordnungrelation $<$ wie folgt definieren: $[(x_n)_{n\in\mathbb{N}}] < [(y_n)_{n\in\mathbb{N}}]$ genau dann, wenn es ein $N \in \mathbb{N}$ und ein $\varepsilon \in \mathbb{Q}_{>0}$ gibt, so dass $x_n + \varepsilon < y_n$ für alle $n \geq N$.

Die Mengen $\mathbb{R}_{\text{Dedekind}}$ und $\mathbb{R}_{\text{Cauchy}}$ sehen ganz verschieden aus. Um sie zu vergleichen, definieren wir eine Abbildung

$$\begin{aligned} v\colon \text{Cauchy}(\mathbb{Q}) &\to \mathbb{R}_{\text{Dedekind}}, \\ (x_n)_{n\in N} &\mapsto \{x \in \mathbb{Q} \mid \text{es existiert } N \in \mathbb{N}, \text{ so dass } x < x_n \text{ für alle } n \geq N\}. \end{aligned}$$

Man kann leicht nachprüfen, dass

$$(x_n)_{n\in N} \sim (y_n)_{n\in N} \Longrightarrow v((x_n)_{n\in N}) = v((y_n)_{n\in N}).$$

Nach der universellen Eigenschaft der Quotientenmenge (Satz 1.4.9) erhalten wir eine induzierte Abbildung

$$\bar{v}\colon \text{Cauchy}(\mathbb{Q})/\!\sim\; = \mathbb{R}_{\text{Cauchy}} \to \mathbb{R}_{\text{Dedekind}}.$$

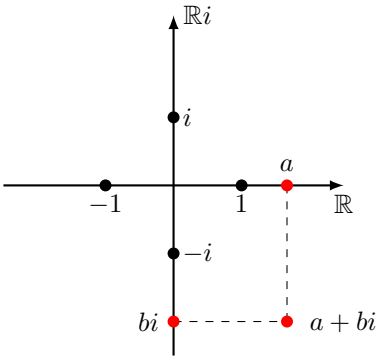
Man kann dann zeigen, dass $\bar{v}$ bijektiv ist, und zwar ein *Isomorphismus* von angeordneten Körpern, d.h., $\bar{v}$ ist kompatibel mit den arithmetischen Operationen und den Ordnungsrelationen, die wir auf beiden Seiten definiert haben. Es gilt zum Beispiel $\bar{v}(x+y) = \bar{v}(x)+\bar{v}(y)$, wobei $+$ auf der linken Seite die Addition von Äquivalenzklassen von Cauchyfolgen ist, und $+$ auf der rechten Seite die Addition von Dedekindschen Schnitten ist.

2.4.3 Die komplexen Zahlen

Eine komplexe Zahl ist ein Ausdruck der Gestalt $a + bi$ mit $a, b \in \mathbb{R}$:

$$\mathbb{C} := \{a + bi \mid a, b \in \mathbb{R}\}.$$

Die reelle Zahl a bzw. b heißt der *Realteil* bzw. der *Imaginärteil* der komplexen Zahl $a + bi$. Mengentheoretisch kann man einfach $\mathbb{C}$ als $\mathbb{R} \times \mathbb{R}$ definieren, wobei ein Paar (a, b) als die komplexe Zahl $a + bi$ aufgefasst wird. Geometrisch kann man sich also die komplexe Zahl $a + bi$ als einen Punkt auf der Ebene vorstellen:



Mit folgenden Definitionen ist $(\mathbb{C}, +, \cdot)$ ein Körper:

$$\begin{aligned} (a + bi) + (c + di) &= (a + c) + (b + d)i, \\ (a + bi) \cdot (c + di) &= (ac - bd) + (ad + bc)i. \end{aligned}$$

Insbesondere ist $i^2 = -1$, wobei $i := 0 + 1i$. Umgekehrt folgen die obigen Formeln für $+$ und $\cdot$ aus $i^2 = -1$ und den Körperaxiomen. Wir identifizieren $\mathbb{R}$ mit einem Teilkörper von $\mathbb{C}$ mit Hilfe der injektiven Abbildung

$$\mathbb{R} \to \mathbb{C}, \quad a \mapsto a + 0i.$$

Bemerkung 2.4.3. Im Gegensatz zu $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ und $\mathbb{R}$, gibt es auf $\mathbb{C}$ keine vernünftige totale Ordnung $\leq$. Insbesondere gibt es keine „positive“ und „negative“ komplexe Zahlen.

Beispiel 2.4.4. Eine komplexe Zahl $a + bi$ heißt *rational*, falls $a, b \in \mathbb{Q}$. Die rationalen komplexen Zahlen bilden einen Teilkörper $\mathbb{Q}(i) \subset \mathbb{C}$.

Man kann die Konstruktion der Zahlenmengen $\mathbb{Z}$ und $\mathbb{Q}$ damit motivieren, dass man bestimmte algebraische Gleichungen lösen will. Gleichungen der Gestalt

$$x + a = b, \quad a, b \in \mathbb{N},$$

since nicht immer mit $x \in \mathbb{N}$ lösbar. Deswegen führen wir die ganzen Zahlen $\mathbb{Z}$ ein, und dann haben *alle* Gleichungen

$$x + a = b, \quad a, b \in \mathbb{Z},$$

eine Lösung $x \in \mathbb{Z}$. In ähnlicher Weise, Gleichungen der Gestalt

$$a \cdot x = b, \quad a, b \in \mathbb{Z}, \quad a \neq 0,$$

sind nicht immer mit $x \in \mathbb{Z}$ lösbar. Deswegen führen wir die rationalen Zahlen $\mathbb{Q}$ ein, mit denen alle solchen Gleichungen lösbar sind. Der Übergang von $\mathbb{Q}$ nach $\mathbb{R}$ ist von anderer Art: Es handelt sich um eine analytische und nicht algebraische Konstruktion (obwohl es auch algebraische Gleichungen gibt, die in $\mathbb{R}$ aber nicht in $\mathbb{Q}$ lösbar sind, z.B. $x^2 = 2$). Nun ist die Gleichung

$$x^2 = a, \quad a \in \mathbb{R},$$

nur mit $x \in \mathbb{R}$ lösbar, wenn $a \geq 0$. Dies motiviert die Einführung der komplexen Zahlen. Es stellt sich heraus, dass in $\mathbb{C}$ *alle* algebraische Gleichungen lösbar sind, und deswegen benötigen wir keine weitere Erweiterung von $\mathbb{C}$. Das ist der *Fundamentalsatz der Algebra*, den wir jetzt genauer formulieren.

Definition 2.4.5 (algebraisch abgeschlossener Körper)**.** Ein Körper K heißt *algebraisch abgeschlossen*, falls jede Gleichung der Gestalt

$$x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0$$

mit $n \geq 1$ und $a_i \in K$ eine Lösung $x \in K$ besitzt.

***Satz 2.4.6** (Fundamentalsatz der Algebra)**.** *Der Körper $\mathbb{C}$ ist algebraisch abgeschlossen.*

Im Gegensatz dazu sind $\mathbb{Q}$ und $\mathbb{R}$ nicht algebraisch abgeschlossen.

2.4.4 Endliche Körper

Sei $n \geq 1$ eine natürliche Zahl. Zur Erinnerung ist die Menge der Restklassen modulo n, $\mathbb{Z}/n\mathbb{Z}$, die Quotientenmenge von $\mathbb{Z}$ bezüglich der Kongruenzrelation $\equiv_n$ (Beispiel 1.4.8). Sie ist eine endliche Menge der Mächtigkeit n:

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

Die Äquivalenzrelation $\equiv_n$ ist kompatibel mit den arithmetischen Operationen $+$ und $\cdot$ im folgenden Sinne: Sind $x \equiv_n x'$ und $y \equiv_n y'$, so sind $x + y \equiv_n x' + y'$ und $x \cdot y \equiv_n x' \cdot y'$. Daraus folgt, dass die wie folgt definierten Operationen auf $\mathbb{Z}/n\mathbb{Z}$ wohldefiniert sind:

$$[x] + [y] := [x + y],$$
$$[x] \cdot [y] := [x \cdot y].$$

Mann kann sogar zeigen, dass $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ ein kommutativer Ring ist (siehe Bemerkung 2.3.5).

Bemerkung 2.4.7. Wenn $n = 24$, ist uns die Addition auf $\mathbb{Z}/24\mathbb{Z}$ vom Rechnen mit Uhrzeiten sehr bekannt. Zum Beispiel können wir die Gleichung $[5] - [8] = [21]$ in $\mathbb{Z}/24\mathbb{Z}$ als „8 Stunden vor 5 Uhr ist 21 Uhr“ verstehen.

***Satz 2.4.8.** $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ *ist genau dann ein Körper, wenn n eine Primzahl ist.*

Der Körper $\mathbb{Z}/p\mathbb{Z}$ mit $p \in \mathbb{N}$ eine Primzahl wird auch mit $\mathbb{F}_p$ bezeichnet. Die Charakteristik von $\mathbb{F}_p$ ist gleich p.

Beispiel 2.4.9. Der Körper $\mathbb{F}_2$ hat genau zwei Elemente, nämlich 0 und 1. Die Addition und Multiplikation werden in folgenden Tabellen explizit dargelegt:

$\mathbb{F}_2 = \{0, 1\}$

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

Die Operationen + und · in diesem Körper können auch als logische Operationen aufgefasst werden: Wenn wir 0 als „falsch“ und 1 als „wahr“ interpretieren, dann entspricht + dem exklusiven Oder und · der Konjunktion $\wedge$.

Beispiel 2.4.10. Hier sind die Additions- und Multiplikationstabellen des Körpers $\mathbb{F}_3$:

$\mathbb{F}_3 = \{0, 1, -1\}$

$+$	0	1	-1
0	0	1	-1
1	1	-1	0
-1	-1	0	1

$\cdot$	0	1	-1
0	0	0	0
1	0	1	-1
-1	0	-1	1

Bemerkung 2.4.11. In der Vorlesung *Algebra* wird gezeigt, dass ein endlicher Körper mit q Elementen genau dann existiert, wenn q eine Primzahlpotenz ist, d.h., wenn $q = p^n$ mit einer Primzahl p und einer natürlichen Zahl $n \geq 1$. Außerdem ist ein solcher Körper eindeutig bis auf Isomorphie und wird mit $\mathbb{F}_q$ bezeichnet. Die Charakteristik von $\mathbb{F}_{p^n}$ ist gleich p. Zum Beispiel gibt es einen Körper $\mathbb{F}_4$ mit vier Elementen und folgender Addition bzw. Multiplikation:

$\mathbb{F}_4 = \{0, 1, \alpha, \beta\}$

$+$	0	1	α	β
0	0	1	α	β
1	1	0	β	α
α	α	β	0	1
β	β	α	1	0

$\cdot$	0	1	α	β
0	0	0	0	0
1	0	1	α	β
α	0	α	β	1
β	0	β	1	α

Bemerkung 2.4.12. Die endlichen Körper $\mathbb{F}_q$ sind nicht algebraisch abgeschlossen. Es existiert aber auch algebraisch abgeschlossene Körper der Primcharakteristik p.

Kapitel 3

Vektorraüme

In diesem Kapitel legen wir einen Körper K fest. Der Körper K heißt der *Grundkörper*, und die Elemente von K heißen *Skalare*. Wir verwenden üblicherweise griechische Buchstaben $\lambda, \mu, \ldots$ für Skalare.

3.1 Das prototypische Beispiel

Sei $n \in \mathbb{N}$. Das prototypische Beispiel eines Vektorraums über K ist die Menge K^n aller n-Tupel von Elementen von K:

$$K^n = \underbrace{K \times \cdots \times K}_{n \text{ mal}} = \{(x_1, \ldots, x_n) \mid x_i \in K\}.$$

Wir werden oft ein n-Tupel $(x_1, x_2, \ldots, x_n) \in K^n$ als *Spaltenvektor*

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

darstellen. Der Grund für eine solche Darstellung wird später im Zusammenhang mit der Multiplikation von Matrizen begründet werden (siehe Abschnitt 4.2.1)

Sei $i \in \{1, \ldots, n\}$. Die i-te kanonische Projektion von K^n auf K ist die Abbildung

$$\begin{aligned} \pi_i \colon K^n &\to K, \\ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} &\mapsto x_i. \end{aligned}$$

Ist $x \in K^n$, so schreibt man üblicherweise x_i für $\pi_i(x)$, so dass

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Das Element $x_i \in K$ heißt die i-te *Koordinate* des n-Tupels $x \in K^n$.

Bemerkung 3.1.1. Wenn $n = 0$ ist K^n das Produkt einer Mengenfamilie mit leerer Indexmenge. Nach Definition 1.2.15(iv) besteht also K^0 aus genau einem Element, dem „leeren Spaltenvektor“. Wenn $n = 1$ ist $K^n = K$.

Definition 3.1.2 (Addition und Skalarmultiplikation auf K^n)**.**

- Die *Addition* auf K^n ist die wie folgt definierte Abbildung:

$$+\colon K^n \times K^n \to K^n,$$
$$\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}\right) \mapsto \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}.$$

- Die *Skalarmultiplikation* auf K^n die wie folgt definierte Abbildung:

$$\cdot\colon K \times K^n \to K^n,$$
$$\left(\lambda, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) \mapsto \begin{pmatrix} \lambda \cdot x_1 \\ \vdots \\ \lambda \cdot x_n \end{pmatrix}.$$

Lemma 3.1.3. *Seien $G_1, \dots, G_n$ Gruppen. Dann ist das Produkt $G_1 \times \cdots \times G_n$ eine Gruppe mit der komponentenweisen Verknüpfung*

$$(g_1, \dots, g_n) \cdot (h_1, \dots, h_n) := (g_1 \cdot h_1, \dots, g_n \cdot h_n).$$

Die Gruppe $G_1 \times \cdots \times G_n$ ist abelsch, falls alle Gruppen $G_1, \dots, G_n$ abelsch sind.

Beweis. Alle Gruppenaxiome für $G_1 \times \cdots \times G_n$ folgen unmittelbar aus den Gruppenaxiomen für die einzelnen Faktoren G_i. Zur Assoziativität gilt nach Definition der Verknüpfung:

$$\big((g_1, \dots, g_n)(h_1, \dots, h_n)\big)(k_1, \dots, k_n) = ((g_1h_1)k_1, \dots, (g_nh_n)k_n),$$
$$(g_1, \dots, g_n)\big((h_1, \dots, h_n)(k_1, \dots, k_n)\big) = (g_1(h_1k_1), \dots, g_n(h_nk_n)),$$

und die rechten Seiten sind gleich nach der Assoziativität in jedem G_i. Im abelschen Fall wird die Kommutativität auf ähnliche Weise nachgeprüft. Das neutrale Element ist das n-Tupel der neutralen Elemente $(e, \dots, e)$. Das Inverse von $(g_1, \dots, g_n)$ ist $(g_1^{-1}, \dots, g_n^{-1})$. □

Bemerkung 3.1.4. Keine ähnliche Aussage gilt für Körper: Das Produkt $K_1 \times K_2$ zweier Körper ist *kein* Körper bzgl. der komponentweisen Verknüpfungen, da z.B. $(1, 0)$ kein multiplikatives Inverses besitzt (es ist jedoch ein kommutativer Ring).

Proposition 3.1.5 (Eigenschaften der Addition und der Skalarmultiplikation auf K^n)**.**

(i) *$(K^n, +)$ ist eine abelsche Gruppe.*

(ii) *Für alle $\lambda, \mu \in K$ und $x \in K^n$ gilt*

$$(\lambda \cdot \mu) \cdot x = \lambda \cdot (\mu \cdot x).$$

(iii) *Für alle $x \in K^n$ gilt*

$$1 \cdot x = x.$$

(iv) *Für alle $\lambda, \mu \in K$ und $x, y \in K^n$ gilt*

$$\lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y,$$
$$(\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot x.$$

Beweis. Aussage (i) ist der Sonderfall von Lemma 3.1.3 mit $G_1 = \cdots = G_n = K$. Die Beweise der Eigenschaften (ii)–(iv) sind ähnlich: Sie lassen sich komponentweise nachrechnen und folgen aus den entsprechenden Eigenschaften der Verknüpfungen $+$ und $\cdot$ des Körpers K. Wir beweisen stellvertretend Aussage (ii):

$$\begin{aligned}
(\lambda \cdot \mu) \cdot x &= \begin{pmatrix} (\lambda \cdot \mu) \cdot x_1 \\ \vdots \\ (\lambda \cdot \mu) \cdot x_n \end{pmatrix} && \text{(Definition der Skalarmultiplikation)} \\
&= \begin{pmatrix} \lambda \cdot (\mu \cdot x_1) \\ \vdots \\ \lambda \cdot (\mu \cdot x_n) \end{pmatrix} && \text{(Assoziativität von } \cdot \text{ in } K) \\
&= \lambda \cdot \begin{pmatrix} \mu \cdot x_1 \\ \vdots \\ \mu \cdot x_n \end{pmatrix} && \text{(Definition der Skalarmultiplikation)} \\
&= \lambda \cdot (\mu \cdot x). && \text{(Definition der Skalarmultiplikation)}
\end{aligned}$$

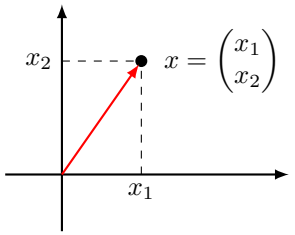
□

Definition 3.1.6 (Standardeinheitsvektoren)**.** Sei $i \in \{1, \ldots, n\}$. Der i-te *Standardeinheitsvektor* in K^n ist

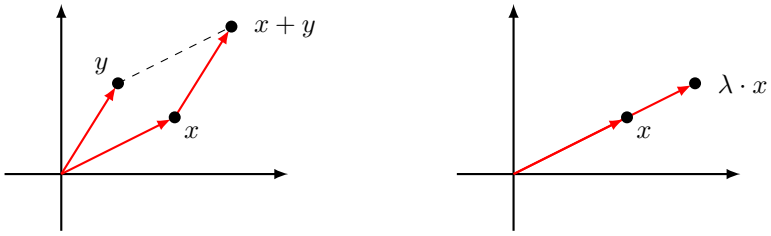
$$e_i := \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix},$$

wobei 1 in der i-te Zeile liegt und alle anderen Koordinaten null sind.

Wenn $K = \mathbb{R}$, können wir $\mathbb{R}^2$ und $\mathbb{R}^3$ als die 2-dimensionale Ebene und den 3-dimensionalen Raum der Euklidischen Geometrie auffassen. Das heißt, wir können Elemente von $\mathbb{R}^2$ und $\mathbb{R}^3$ mit Punkten der Ebene und des Raums identifizieren. Manchmal stellt man auch ein Element x aus $\mathbb{R}^2$ oder $\mathbb{R}^3$ mit einem Pfeil von dem Nullpunkt nach dem Punkt x dar:



Diese Darstellung ist hilfreich, um die Addition und Skalarmultiplikation auf geometrische Weise zu beschreiben: Man erhält die Summe $x + y$ zweier Elemente x und y, indem man der Pfeil von y längs des von x verschiebt, und man erhält das λ-Vielfache $\lambda \cdot x$ von x, indem man der Pfeil von x um den Faktor λ skaliert:



Diese geometrische Anschauung ist auch hilfreich in höherer Dimension oder bei anderen Körpern, selbst wenn man nicht zeichnen kann.

3.2 Vektorraüme

Wir abstrahieren jetzt die in Proposition 3.1.5 bewiesenen Eigenschaften von K^n zum Begriff des Vektorraums:

Definition 3.2.1 (Vektorraum)**.** Ein *Vektorraum* über K, oder K*-Vektorraum*, ist ein Tripel $(V, +, \cdot)$, bestehend aus einer Menge V und Abbildungen

$$\begin{aligned} +\colon V \times V \to V, &\quad (v, w) \mapsto v + w, \\ \cdot\colon K \times V \to V, &\quad (\lambda, v) \mapsto \lambda \cdot v, \end{aligned}$$

die als *Addition* und *Skalarmultiplikation* bezeichnet werden, mit folgenden Eigenschaften:

(i) $(V, +)$ ist eine abelsche Gruppe.

(ii) Für alle $\lambda, \mu \in K$ und $v \in V$ gilt

$$(\lambda \cdot \mu) \cdot v = \lambda \cdot (\mu \cdot v).$$

(iii) Für alle $v \in V$ gilt

$$1 \cdot v = v.$$

(iv) Für alle $\lambda, \mu \in K$ und $v, w \in V$ gilt

$$\begin{aligned} \lambda \cdot (v + w) &= \lambda \cdot v + \lambda \cdot w, \\ (\lambda + \mu) \cdot v &= \lambda \cdot v + \mu \cdot v. \end{aligned}$$

Elemente von V heißen *Vektoren*. Das neutrale Element 0 bzgl. $+$ heißt der *Nullvektor*.

Bemerkung 3.2.2. Im zweiten Axiom, $(\lambda \cdot \mu) \cdot v = \lambda \cdot (\mu \cdot v)$, das erste $\cdot$ ist die Multiplikation auf K, und die anderen drei $\cdot$ sind Skalarmultiplikation. Wegen diesem Axiom darf man einfach $\lambda \cdot \mu \cdot v$ schreiben. Allgemeiner, sind $\lambda_1, \ldots, \lambda_n$ Skalare, so darf man $\lambda_1 \cdot \ldots \cdot \lambda_n \cdot v$ ohne Klammern schreiben. Wie üblich wird das Symbol $\cdot$ oft ganz unterdrückt.

Beispiel 3.2.3 (Vektorraum der n-Tupel)**.** Für jedes $n \in \mathbb{N}$ ist $(K^n, +, \cdot)$ ein Vektorraum über K, wobei $+$ und $\cdot$ die Addition und Skalarmultiplication von n-Tupeln sind (Definition 3.1.2). Dies folgt aus Proposition 3.1.5.

Beispiel 3.2.4 (Körpererweiterungen als Vektorräume)**.** Sei L ein Körper. Ein Teilmenge $K \subset L$ heißt *Teilkörper*, wenn sich die Addition und Multiplikation auf L zu K einschränken und K mit diesen eingeschränkten Verknüpfungen einen Körper bildet. Man sagt dann auch, dass L eine *Körpererweiterung* von K ist. Zum Beispiel: $\mathbb{R}$ und $\mathbb{C}$ sind Körpererweiterungen von $\mathbb{Q}$, und $\mathbb{C}$ ist auch eine Körpererweiterung von $\mathbb{R}$.

Wenn L eine Körpererweiterung von K ist, dann bildet L mit seiner Addition und seiner auf $K \times L$ eingeschränkten Multiplikation einen K-Vektorraum: Die Axiome (i)–(iv) der Definition 3.2.1 sind Sonderfälle der Körperaxiome für L.

Beispiel 3.2.5 (Vektorräume von Abbildungen)**.** Sei V ein K-Vektorraum und X eine beliebige Menge. Dann ist die Menge $\mathrm{Abb}(X, V)$ aller Abbildungen von X nach V ein K-Vektorraum bezüglich der punktweisen Addition bzw. Skalarmultiplikation:

$$\begin{aligned} +\colon \mathrm{Abb}(X, V) \times \mathrm{Abb}(X, V) &\to \mathrm{Abb}(X, V), \\ (f, g) &\mapsto (x \mapsto f(x) + g(x)), \end{aligned}$$

$$\begin{aligned} \cdot\colon K \times \mathrm{Abb}(X, V) &\to \mathrm{Abb}(X, V), \\ (\lambda, f) &\mapsto (x \mapsto \lambda \cdot f(x)). \end{aligned}$$

Alle Axiome der Definition 3.2.1 können punktweise nachgeprüft werden und folgen aus den entsprechenden Axiomen für V.

Insbesondere haben wir den K-Vektorraum $\mathrm{Abb}(X, K)$ aller Abbildungen von X nach K, der auch mit K^X bezeichnet wird. Der K-Vektorraum K^n kann als Sonderfall dieser Konstruktion aufgefasst werden, nämlich mit $X = \{1, \ldots, n\}$: Effektiv ist ein n-Tupel in K nichts anderes als eine Abbildung $\{1, \ldots, n\} \to K$.

Proposition 3.2.6 (Rechnen in Vektorräumen)**.** *Sei V ein Vektorraum über K.*

(i) *Für alle $\lambda \in K$ gilt $\lambda \cdot 0 = 0$, wobei $0 \in V$ der Nullvektor ist.*

(ii) *Für alle $v \in V$ gilt $0 \cdot v = 0$. Dabei bezeichnet 0 auf der linken Seite den Nullskalar und auf der rechten Seite den Nullvektor.*

(iii) *Für alle $\lambda \in K$ und $v \in V$ gilt $(-\lambda) \cdot v = -(\lambda \cdot v)$ und $\lambda \cdot (-v) = -(\lambda \cdot v)$.*

(iv) *Sind $\lambda \in K \setminus \{0\}$ und $v \in V \setminus \{0\}$, so ist $\lambda \cdot v \in V \setminus \{0\}$.*

Beweis. Zu (i). Nach Axiom (iv) gilt

$$\lambda \cdot 0 = \lambda \cdot (0 + 0) = \lambda \cdot 0 + \lambda \cdot 0.$$

Da $(V, +)$ eine Gruppe ist, können wir $\lambda \cdot 0$ von beiden Seiten subtrahieren, und erhalten wir $0 = \lambda \cdot 0$.

Zu (ii). Ähnlicher Beweis: Nach Axiom (iv) gilt

$$0 \cdot v = (0 + 0) \cdot v = 0 \cdot v + 0 \cdot v,$$

und daher $0 = 0 \cdot v$.

Zu (iii). Nach der Eindeutigkeit von inversen Elementen genügt es zu zeigen, dass $(-\lambda) \cdot v$ und $\lambda \cdot (-v)$ inverse Elemente von $\lambda \cdot v$ bzgl. $+$ sind. Nach Axiom (iv) gilt

$$(-\lambda) \cdot v + \lambda \cdot v = ((-\lambda) + \lambda) \cdot v = 0 \cdot v,$$

und $0 \cdot v$ ist gleich null nach (ii). Nach Axiom (iv) gilt ebenfalls

$$\lambda \cdot (-v) + \lambda \cdot v = \lambda \cdot ((-v) + v) = \lambda \cdot 0,$$

und $\lambda \cdot 0$ ist gleich null nach (i).

Zu (iv). Wir beweisen die äquivalente Aussage: Ist $\lambda \cdot v = 0$ und $\lambda \neq 0$, so ist $v = 0$. Da K ein Körper ist hat λ ein Inverses λ^{-1} bzgl. $\cdot$. Dann gilt:

$$\begin{aligned}
v &= 1 \cdot v && \text{(Axiom (iii))} \\
&= (\lambda^{-1} \cdot \lambda) \cdot v && (\lambda^{-1} \text{ invers zu } \lambda) \\
&= \lambda^{-1} \cdot (\lambda \cdot v) && \text{(Axiom (ii))} \\
&= \lambda^{-1} \cdot 0 && \text{(Annahme)} \\
&= 0. && \text{(nach (i))}
\end{aligned}$$

□

3.2.1 Untervektorräume

Definition 3.2.7 (Untervektorraum)**.** Sei $(V, +, \cdot)$ ein Vektorraum über K. Eine Teilmenge $U \subset V$ heißt *Untervektorraum*, wenn sich die Abbildungen $+ \colon V \times V \to V$ und $\cdot \colon K \times V \to V$ zu Abbildungen $+ \colon U \times U \to U$ und $\cdot \colon K \times U \to U$ einschränken, und U mit diesen eingeschränkten Verknüpfungen ein K-Vektorraum ist.

Ist U ein Untervektorraum eines K-Vektorraums $(V, +, \cdot)$, so betrachten wir immer U als K-Vektorraum mit den eingeschränkten Verknüpfungen. Um Untervektorräume zu erkennen verwenden wir folgendes Kriterium:

Proposition 3.2.8 (Kriterium für Untervektorräume)**.** *Sei V ein Vektorraum über K. Eine Teilmenge $U \subset V$ ist genau dann ein Untervektorraum, wenn folgende drei Bedingungen erfüllt sind:*

(i) *U ist nicht leer.*

(ii) *Für alle $v, w \in U$ gilt $v + w \in U$.*

(iii) *Für alle $v \in U$ und $\lambda \in K$ gilt $\lambda \cdot v \in U$.*

Außerdem gilt in diesem Fall:

(iv) *Der Nullvektor $0 \in V$ liegt in U, und ist auch der Nullvektor von U.*

(v) *Für alle $v \in U$, der inverse Vektor $-v \in V$ liegt in U, und ist auch das Inverse von v in U.*

Beweis. Ist U ein Untervektorraum, so sind Bedingungen (ii) und (iii) nach Definition erfüllt. Außerdem ist U nicht leer, da es als Vektorraum einen Nullvektor enthält.

Umgekehrt, sei $U \subset V$ eine Teilmenge, die die Bedingungen (i)–(iii) erfüllt. Nach (ii) und (iii) schränken sich die Verknüpfungen $+$ und $\cdot$ auf U ein, und es ist zu zeigen, dass $(U, +, \cdot)$ ein K-Vektorraum ist. Die Axiome (ii)–(iv) in der Definition 3.2.1 eines Vektorraums, wie auch die Assoziativität und Kommutativität von $+$, gelten für alle $v, w \in V$ und $\lambda, \mu \in K$; insbesondere gelten sie für alle $v, w \in U$ und $\lambda, \mu \in K$. Es bleibt also zu zeigen, dass in U ein neutrales Element und inverse Elemente bzgl. $+$ existieren. Dazu genügt es (iv) und (v) zu beweisen. Nach Proposition 3.2.6(iii) und dem Vektorraumaxiom $1 \cdot v = v$ gilt

$$(-1) \cdot v = -(1 \cdot v) = -v$$

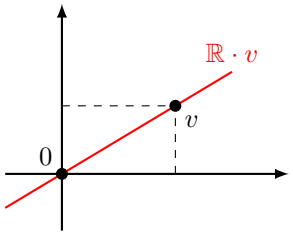
für alle $v \in V$. Aus (iii) folgt, dass $-v \in U$ für alle $v \in U$, d.h., es gilt (v). Nach (i) existiert ein $u \in U$. Da $u + (-u) = 0$, liegt 0 in U nach (v) und (ii), d.h., es gilt (iv). □

Beispiel 3.2.9. Sei V ein beliebiger K-Vektorraum. Dann sind $\{0\}$ und V Untervektorräume von V.

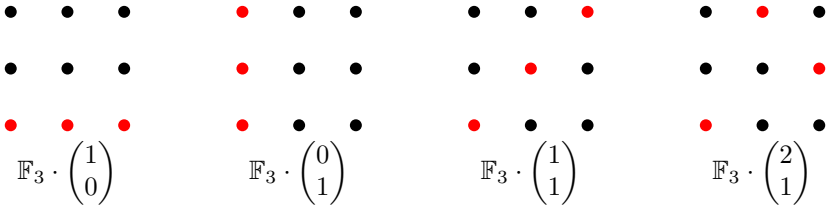
Beispiel 3.2.10 (Ursprungsgeraden)**.** Sei V ein K-Vektorraum und $v \in V \setminus \{0\}$. Dann ist die Teilmenge

$$K \cdot v := \{\lambda \cdot v \mid \lambda \in K\} \subset V$$

ein Untervektorraum. Dieser Untervektorraum heißt die *von v aufgespannte Gerade* in V. Wenn $K = \mathbb{R}$ und $V = \mathbb{R}^2$ oder $\mathbb{R}^3$, dann ist $\mathbb{R} \cdot v$ eine Gerade im üblichen Sinn: Sie ist nämlich die Gerade, die durch den Ursprung und v läuft:



Beispiel 3.2.11 (Ursprungsgeraden über endlichen Körpern)**.** Sei p eine Primzahl. Eine Ursprungsgerade in einem $\mathbb{F}_p$-Vektorraum besteht aus genau p Elementen. Folgendes Bild zeigt alle vier Ursprungsgeraden in $\mathbb{F}_3^2$:



Beispiel 3.2.12. Sei I eine beliebige Menge. Wir betrachten den K-Vektorraum $K^I = \mathrm{Abb}(I, K)$ aller Abbildungen von I nach K (siehe Beispiel 3.2.5). Sei $K^{(I)}$ seine Teilmenge bestehend aus aller Abbildungen, die außerhalb einer endlichen Teilmenge von I null sind:

$$K^{(I)} := \{f\colon I \to K \mid \text{es existiert } J \subset I \text{ endlich, so dass } f(I \setminus J) \subset \{0\}\}.$$

Dann ist $K^{(I)}$ ein Untervektorraum von K^I, denn: (i) Er enthält den Nullvektor; (ii) falls f_1 außerhalb J_1 und f_2 außerhalb J_2 null sind, dann ist $f_1 + f_2$ außerhalb $J_1 \cup J_2$ null; (iii) falls f außerhalb J null ist, dann ist $\lambda \cdot f$ ebenfalls außerhalb J null. Man beachte, dass $K^{(I)} = K^I$, wenn I endlich ist.

Beispiel 3.2.13 (konvergente Folgen)**.** Sei $\mathbb{R}^{\mathbb{N}}$ der $\mathbb{R}$-Vektorraum aller Folgen in $\mathbb{R}$. Dann ist die Teilmenge

$$\mathrm{Konv}(\mathbb{R}) := \{(x_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} \mid (x_n)_{n\in\mathbb{N}} \text{ konvergiert}\} \subset \mathbb{R}^{\mathbb{N}}$$

ein Untervektorraum. Dies folgt aus dem Kriterium 3.2.8, da die Summe zweier konvergenten Folgen konvergiert, wie auch die Skalarmultiplikation einer konvergenten Folge mit einer reellen Zahl.

Beispiel 3.2.14 (Funktionenräume)**.** Sei $K = \mathbb{R}$ und seien $a < b$ reelle Zahlen. Nach Beispiel 3.2.5 ist die Menge $\mathrm{Abb}([a, b], \mathbb{R})$ aller reellen Funktionen auf $[a, b]$, versehen mit der punktweisen Addition bzw. Skalarmultiplikation, ein $\mathbb{R}$-Vektorraum. In der Analysis betrachtet man viele verschiedene Sorten solcher Funktionen, und die entsprechenden Teilmengen von $\mathrm{Abb}([a, b], \mathbb{R})$ sind oft Untervektorräume. Das gilt zum Beispiel für stetige, differenzierbare, stetig differenzierbare, beliebig oft differenzierbare und Riemann-integrierbare Funktionen.

Bemerkung 3.2.15 (Abhängigkeit vom Grundkörper)**.** Sei $K \subset L$ eine Körpererweiterung (siehe Beispiel 3.2.4). Dann können wir jeden L-Vektorraum V als K-Vektorraum betrachten, indem wir die Skalarmultiplikation $\cdot\colon L \times V \to V$ auf $K \times V$ einschränken. Ob eine Teilmenge $U \subset V$ ein Untervektorraum ist, hängt davon ab, über welchem Körper wir V als Vektorraum betrachten. Zum Beispiel ist $\mathbb{R}$ ein Untervektorraum von $\mathbb{C}$, wenn wir $\mathbb{C}$ als $\mathbb{R}$-Vektorraum betrachten, aber nicht wenn wir $\mathbb{C}$ als $\mathbb{C}$-Vektorraum betrachten. Deswegen sollten wir eher von K-Untervektorräumen sprechen, wenn solche Mehrdeutigkeiten möglich sind.

Proposition 3.2.16 (Durchschnitt von Untervektorräumen)**.** *Sei V ein Vektorraum über K. Sind $U, W \subset V$ Untervektorräume, so ist $U \cap W \subset V$ ein Untervektorraum.*

Allgemeiner, ist $(U_i)_{i\in I}$ eine beliebige Familie von Untervektorräumen von V, so ist $\bigcap_{i\in I} U_i \subset V$ ein Untervektorraum.

Beweis. Wir beweisen die allgemeinere Aussage mithilfe der Proposition 3.2.8. Der Durchschnitt $\bigcap_{i\in I} U_i$ ist nicht leer, da jedes U_i den Nullvektor von V enthält. Die Bedingungen (ii) und (iii) für alle U_i implizieren dieselben Bedingungen für $\bigcap_{i\in I} U_i$. □

Definition 3.2.17 (erzeugter Untervektorraum, Erzeugendensystem)**.** Sei V ein Vektorraum über K und $E \subset V$ eine Teilmenge.

- Der *von E erzeugte Untervektorraum*, oder die *lineare Hülle* von E, ist

$$\mathrm{Span}_K(E) := \bigcap_{U\in\mathcal{U}(E)} U,$$

  wobei $\mathcal{U}(E)$ die Menge aller Untervektorräume $U \subset V$ mit $E \subset U$ ist. Nach Proposition 3.2.16 ist $\mathrm{Span}_K(E)$ ein Untervektorraum von V, und zwar der kleinste Untervektorraum, der E enthält.

- E heißt *Erzeugendensystem* von V, falls $\mathrm{Span}_K(E) = V$.

Ein Ausdruck der Gestalt

$$\sum_{i=1}^{n} \lambda_i \cdot v_i$$

mit $n \in \mathbb{N}$, $\lambda_i \in K$ und $v_i \in V$ heißt *Linearombination* der Vektoren v_i. Die folgende Proposition gibt eine explizite Beschreibung des Untervektorraums $\mathrm{Span}_K(E)$ als die Menge aller Linearkombinationen von Vektoren aus E. Insbesondere ist E genau dann ein Erzeugendensystem von V, wenn jeder Vektor aus V eine Linearkombination von Vektoren aus E ist.

Proposition 3.2.18. *Sei V ein Vektorraum über K und $E \subset V$ eine Teilmenge. Dann gilt*

$$\mathrm{Span}_K(E) = \left\{ \sum_{i=1}^{n} \lambda_i \cdot v_i \;\middle|\; n \in \mathbb{N},\ \lambda_i \in K,\ v_i \in E \right\}.$$

(Dabei ist die leere Summe $\sum_{i=1}^{0} \lambda_i \cdot v_i$ gleich 0, nach Notation 2.1.11.)

Beweis. Zu $\supset$. Nach Definition gilt $E \subset \mathrm{Span}_K(E)$. Da $\mathrm{Span}_K(E)$ ein Untervektorraum ist, jede Summe $\sum_{i=1}^{n} \lambda_i \cdot v_i$ auf der rechten Seite liegt in $\mathrm{Span}_K(E)$.

Zu $\subset$. Sei U die Menge auf der rechten Seite. Es gilt $E \subset U$. Nach Definition von $\mathrm{Span}_K(E)$ genügt es also zu zeigen, dass U ein Untervektorraum von V ist. Dazu verwenden wir das Kriterium 3.2.8. Es gilt $0 \in U$, insbesondere ist U nicht leer. Bedingung (ii) ist offensichtlich, und Bedingung (iii) folgt aus den Vektorraumaxiomen, da

$$\lambda \cdot \left(\sum_{i=1}^{n} \lambda_i \cdot v_i \right) = \sum_{i=1}^{n} (\lambda \cdot \lambda_i) \cdot v_i. \qquad \square$$

Beispiel 3.2.19.

(i) Es gilt $\mathrm{Span}_K(\varnothing) = \{0\}$, weil $\{0\}$ bereits ein Untervektorraum ist, der $\varnothing$ enthält.

(ii) Die Menge der Standardeinheitsvektoren $\{e_1, \ldots, e_n\} \subset K^n$ ist ein Erzeugendensystem von K^n, da jeder Vektor $x \in K^n$ als

$$x = x_1 \cdot e_1 + \cdots + x_n \cdot e_n$$

dargestellt werden kann.

(iii) Sei $E \subset \mathbb{Q}^n$ die Menge aller n-Tupel $(x_1, \ldots, x_n)$ mit $x_i > 7$ für alle i, d.h., $E = (\mathbb{Q}_{>7})^n$. Dann ist E ein Erzeugendensystem des $\mathbb{Q}$-Vektorraums $\mathbb{Q}^n$, denn: Sei $U \subset \mathbb{Q}^n$ ein Untervektorraum, der E enthält. Es gilt $(8, \ldots, 8) \in E$ und $e_i + (8, \ldots, 8) \in E$ für alle $i \in \{1, \ldots, n\}$, und damit $e_i \in U$. Aus (ii) folgt, dass $U = \mathbb{Q}^n$.

(iv) Sei I eine Menge und sei K^I der Vektorraum aller Abbildungen von I nach K (siehe Beispiel 3.2.5). Man kann jedem $i \in I$ einen „Standardeinheitsvektor“ $e_i \in K^I$ zuordnen, wobei

$$\begin{aligned} e_i \colon I &\to K, \\ j &\mapsto \begin{cases} 1, & \text{falls } j = i, \\ 0, & \text{falls } j \neq i. \end{cases} \end{aligned}$$

Im Gegensatz zu (ii), wenn I *unendlich* ist, ist die Menge $\{e_i \mid i \in I\}$ *kein* Erzeugendensystem von K^I. Denn jede Abbildung $f \colon I \to K$, die eine Linearkombination der Abbildungen e_i ist, ist gleich null außerhalb einer endlichen Teilmenge von I. Es gilt also

$$\mathrm{Span}_K(\{e_i \mid i \in I\}) = K^{(I)} \subset K^I,$$

wobei $K^{(I)}$ der im Beispiel 3.2.12 definierte Untervektorraum ist.

Beispiel 3.2.20. Sei $K = \mathbb{R}$ und seien $u, v, w \in \mathbb{R}^3$ drei Vektoren. Es gibt vier verschiedenen geometrischen Möglichkeiten für $\mathrm{Span}_{\mathbb{R}}(\{u, v, w\})$:

- Falls $u = v = w = 0$, dann ist $\mathrm{Span}_{\mathbb{R}}(\{u, v, w\}) = \{0\}$.
- Falls die drei Vektoren auf derselben Ursprungsgerade G liegen, und nicht alle null sind, dann ist $\mathrm{Span}_{\mathbb{R}}(\{u, v, w\}) = G$.
- Falls die drei Vektoren auf derselben Ursprungsebene E liegen, aber nicht auf irgendeiner Ursprungsgerade, dann ist $\mathrm{Span}_{\mathbb{R}}(\{u, v, w\}) = E$.
- Falls die drei Vektoren auf keiner gemeinsamen Ursprungsebene liegen, dann ist $\mathrm{Span}_{\mathbb{R}}(\{u, v, w\}) = \mathbb{R}^3$, d.h., $\{u, v, w\}$ ist ein Erzeugendensystem.

Definition 3.2.21 (endlich erzeugt)**.** Ein K-Vektorraum V heißt *endlich erzeugt*, falls er ein endliches Erzeugendensystem besitzt.

Beispiel 3.2.22.

(i) Für alle $n \in \mathbb{N}$ ist K^n endlich erzeugt, da $\{e_1, \ldots, e_n\}$ ein endliches Erzeugendensystem von K^n ist (siehe Beispiel 3.2.19(ii)).

(ii) Ist I eine unendliche Menge, so sind K^I und $K^{(I)}$ *nicht* endlich erzeugt. Das werden wir später beweisen: Siehe Bemerkung 3.3.28.

3.2.2 Quotientenvektorräume

Im Abschnitt 1.4.1 haben wir den wichtigen Begriff des *Quotienten* einer Menge modulo einer Äquivalenzrelation eingeführt. Wenn $\sim$ eine Äquivalenzrelation auf einem Vektorraum V ist, die mit der Vektorraumstruktur in geeigneter Weise verträglich ist, dann vererbt sich die Vektorraumstruktur auf die Quotientenmenge $V/\!\sim$. Dies führt zum Begriff des *Quotientenvektorraum.*

Proposition 3.2.23. *Sei V ein K-Vektorraum und $U \subset V$ ein Untervektorraum. Sei $\sim_U$ die wie folgt definierte Relation auf V:*

$$v \sim_U w \iff v - w \in U.$$

(i) *$\sim_U$ ist eine Äquivalenzrelation auf V.*

(ii) *Die Äquivalenzklasse eines Vektors v bzgl. $\sim_U$ ist*

$$v + U := \{v + u \mid u \in U\}.$$

(iii) *Die Quotientenmenge $V/\!\sim_U$ hat die Struktur eines K-Vektorraums mit folgender Addition und Skalarmultiplikation:*

$$(v + U) + (w + U) = (v + w) + U,$$
$$\lambda \cdot (v + U) = \lambda v + U.$$

Das neutrale Element ist die Äquivalenzklasse von $0 \in V$, d.h., $0 + U = U$, und das inverse Element von $v + U$ ist $(-v) + U$.

Beweis. Zu (i). Die Reflexivität von $\sim_U$ folgt aus $0 \in U$ und die Symmetrie folgt aus der Implikation $u \in U \Rightarrow -u \in U$. Zur Transitivität: Es seien $v \sim_U w$ und $w \sim_U x$, d.h., $v - w \in U$ und $w - x \in U$. Da die Addition zweier Vektoren aus U wieder in U liegt, gilt $v - x = (v - w) + (w - x) \in U$, d.h., $v \sim_U x$.

Zu (ii). Nach Definition 1.4.4 ist die Äquivalenzklasse von v gleich

$$\{w \in V \mid w - v \in U\} = \{w \in V \mid \text{es existiert } u \in U \text{ mit } w = v + u\} = v + U.$$

Zu (iii). Wir zeigen zunächst, dass die Verknüpfungen $+$ und $\cdot$ auf $V/\sim_U$ wohldefiniert sind. Danach folgen die Vektorraumaxiome für $V/\sim_U$ unmittelbar aus den entsprechenden Axiomen für V. Seien $v \sim_U v'$ und $w \sim_U w'$. Zu zeigen ist, dass

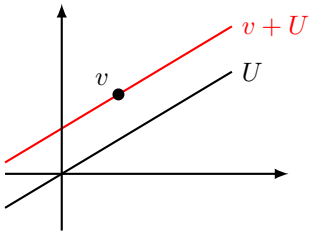
$$v + w \sim_U v' + w'$$
$$\text{und } \lambda v \sim_U \lambda v'.$$

Dies folgt aus den Gleichungen $(v + w) - (v' + w') = (v - v') + (w - w')$ und $\lambda v - \lambda v' = \lambda(v - v')$. □

Definition 3.2.24 (Quotientenvektorraum)**.** Sei V ein K-Vektorraum und $U \subset V$ ein Untervektorraum. Der in Proposition 3.2.23 definierte K-Vektorraum $V/\sim_U$ heißt der *Quotientenvektorraum* von V nach U und wird mit V/U (gelesen „V modulo U“ oder „V durch U“) bezeichnet.

Eine Teilmenge $A \subset V$ der Gestalt $A = v + U$ mit einem Untervektorraum U heißt *affiner Unterraum* von V. Affine Unterräume sind also verschobene Untervektorräume. Der Untervektorraum U ist durch A eindeutig bestimmt, da für jedes $v \in A$ gilt $U = A - v$. Der Quotientenvektorraum V/U ist also die Menge aller affinen Unterräume von V, die parallel zu U sind.

Folgendes Bild stellt einen Untervektorraum U von $\mathbb{R}^2$ und einen affinen Unterraum $v+U$ dar. Der Quotientenvektorraum $\mathbb{R}^2/U$ ist die Menge aller Geraden, die parallel zu U sind:



3.3 Basen und Dimension

Im K-Vektorraum K^n spielt die Familie der Standardeinheitsvektoren $(e_1, \ldots, e_n)$ aus Definition 3.1.6 eine besondere Rolle. Ein Grund dafür ist die folgende Eigenschaft: Jeder Vektor $x \in K^n$ kann als eine Summe

$$x = x_1 \cdot e_1 + \cdots + x_n \cdot e_n$$

dargestellt werden, mit *eindeutig bestimmten* Skalaren $x_1, \ldots, x_n \in K$. Wegen dieser Eigenschaft sagt man, dass die Familie $(e_1, \ldots, e_n)$ eine *Basis* von K^n ist. In diesem Abschnitt werden wir Basen in allgemeinen Vektorräumen definieren. Wir werden beweisen, dass jeder Vektorraum V eine Basis besitzt, und außerdem dass je zwei Basen von V dieselbe Länge haben. Die Länge einer Basis von V heißt dann die *Dimension* von V. Beispielsweise ist die Dimension von K^n gleich n.

3.3.1 Lineare Unabhängigkeit

Definition 3.3.1 (lineare Unabhängigkeit)**.** Sei V ein Vektorraum über K und sei $(v_i)_{i \in I}$ eine Familie von Elementen von V (d.h., eine Abbildung $I \to V$, $i \mapsto v_i$).

- Die Familie $(v_i)_{i\in I}$ heißt *linear unabhängig*, wenn folgendes gilt: Für jede endliche Teilmenge $J \subset I$ und Skalarfamilie $(\lambda_j)_{j\in J}$, ist $\sum_{j\in J} \lambda_j \cdot v_j = 0$, so folgt bereits $\lambda_j = 0$ für alle $j \in J$.
- Die Familie $(v_i)_{i\in I}$ heißt *linear abhängig*, wenn sie nicht linear unabhängig ist, d.h., wenn es eine endliche Teilmenge $J \subset I$ und eine Skalarfamilie $(\lambda_j)_{j\in J}$ existiert, so dass $\sum_{j\in J} \lambda_j \cdot v_j = 0$ und $\lambda_j \neq 0$ für mindestens ein $j \in J$.

Ein Ausdruck der Gestalt

$$\sum_{j\in J} \lambda_j \cdot v_j$$

mit $J \subset I$ einer endlichen Teilmenge heißt *Linearkombination* der Familie $(v_i)_{i\in I}$. Die Definition der linearen Unabhängigkeit wird oft folgendermaßen formuliert: Es gibt keine *nicht-triviale* Linearkombination der Familie $(v_i)_{i\in I}$, die gleich null ist. Dabei heißt eine Linearkombination $\sum_{j\in J} \lambda_j \cdot v_j$ nicht-trivial, falls $\lambda_j \neq 0$ für mindestens ein $j \in J$.

Beispiel 3.3.2. Die leere Familie $(v_i)_{i\in\varnothing}$ ist immer linear unabhängig.

Beispiel 3.3.3. Sei V ein K-Vektorraum und $(v_i)_{i\in I}$ eine Familie von Vektoren aus V.

(i) Gibt es $i \in I$ mit $v_i = 0$, so ist die Familie $(v_i)_{i\in I}$ linear abhängig: Die Linearkombination $1 \cdot v_i$ ist nicht-trivial aber gleich null.

(ii) Gibt es $i \neq j$ mit $v_i = v_j$, so ist die Familie $(v_i)_{i\in I}$ linear abhängig, da $v_i + (-1)\cdot v_j = 0$.

(iii) Folgendes Beispiel ist eine Verallgemeinerung von (i) und (ii). Es gebe einen Index $i \in I$ und eine endliche Teilmenge $J \subset I \setminus \{i\}$, so dass v_i eine Linearkombination von $(v_j)_{j\in J}$ ist, d.h., $v_i = \sum_{j\in J} \lambda_j \cdot v_j$ mit $\lambda_j \in K$. Dann ist die Familie $(v_i)_{i\in I}$ linear abhängig.

Beispiel 3.3.4. Sei $n \in \mathbb{N}$. Die Familie der Standardeinheitsvektoren $(e_i)_{i\in\{1,\dots,n\}}$ in K^n ist linear unabhängig. Denn seien $\lambda_1, \dots, \lambda_n \in K$ mit $\sum_{i=1}^n \lambda_i \cdot e_i = 0$. Nach Definition von e_i gilt

$$\sum_{i=1}^n \lambda_i \cdot e_i = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

Ein n-Tupel ist genau dann gleich null, wenn alle seinen Koordinaten gleich null sind. Es folgt also $\lambda_1 = \cdots = \lambda_n = 0$.

Allgemeiner, die Familie $(e_i)_{i\in I}$ im K-Vektorraum K^I ist linear unabhängig (siehe Beispiel 3.2.19(iv)).

Bemerkung 3.3.5. Ob eine Familie $(v_i)_{i\in I}$ linear unabhängig ist kann nicht „paarweise“ überprüft werden. Zum Beispiel sind alle drei Familien (e_1, e_2), $(e_1, e_1 + e_2)$ und $(e_2, e_1 + e_2)$ in K^2 linear unabhängig, aber die Familie $(e_1, e_2, e_1 + e_2)$ ist linear abhängig.

Beispiel 3.3.6. Sei $K = \mathbb{R}$ und $V = \mathbb{R}^3$.

- Ein einzelner Vektor $v \in \mathbb{R}^3$ ist genau dann linear unabhängig, wenn $v \neq 0$, d.h., wenn der von $\{v\}$ erzeugter Untervektorraum eine Gerade ist.
- Zwei Vektoren $v, w \in \mathbb{R}^3$ sind genau dann linear unabhängig, wenn der von $\{v, w\}$ erzeugter Untervektorraum eine Ebene ist.
- Drei Vektoren $u, v, w \in \mathbb{R}^3$ sind genau dann linear unabhängig, wenn der von $\{u, v, w\}$ erzeugter Untervektorraum der ganze Raum $\mathbb{R}^3$ ist.
- Ein Familie bestehend aus mehr als drei Vektoren in $\mathbb{R}^3$ kann nicht linear unabhängig sein.

Proposition 3.3.7 (Charakterisierung der linearen Abhängigkeit)**.** *Sei $(v_i)_{i\in I}$ eine Familie in einem K-Vektorraum V. Folgende Aussagen sind äquivalent:*

(i) *$(v_i)_{i\in I}$ ist linear abhängig.*

(ii) *Einer der Vektoren in $(v_i)_{i\in I}$ ist eine Linearkombination der anderen. Das heißt: Es existiert $k \in I$, eine endliche Teilmenge $J \subset I \setminus \{k\}$ und eine Familie $(\lambda_j)_{j\in J}$ von Skalaren, so dass $v_k = \sum_{j\in J} \lambda_j \cdot v_j$.*

(iii) *Es existiert eine Teilmenge $J \subsetneq I$, so dass $\mathrm{Span}_K(\{v_i \mid i \in I\}) = \mathrm{Span}_K(\{v_i \mid i \in J\})$.*

Beweis. Wir beweisen die Implikationen (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (i).

Zu (i) $\Rightarrow$ (ii). Nach Definition gibt es eine endliche Teilmenge $J \subset I$ und eine Familie von Skalaren $(\lambda_j)_{j\in J}$, die nicht alle null sind, so dass $\sum_{j\in J} \lambda_j \cdot v_j = 0$. Sei $k \in J$ ein Index mit $\lambda_k \neq 0$. Dann gilt

$$v_k = \sum_{j\in J\setminus\{k\}} \frac{-\lambda_j}{\lambda_k} \cdot v_j.$$

Zu (ii) $\Rightarrow$ (iii). Es seien k, J und $(\lambda_j)_{j\in J}$ wie in (ii). Wir zeigen, dass

$$\mathrm{Span}_K(\{v_i \mid i \in I\}) = \mathrm{Span}_K(\{v_i \mid i \in I \setminus \{k\}\}).$$

Zur Erinnerung ist $\mathrm{Span}_K(E)$ der Durchschnitt aller Untervektorräume von V, die E enthalten. Die Inklusion $\supset$ ist klar. Um die Inklusion $\subset$ zu überprüfen, ist also zu zeigen, dass folgendes für alle Untervektorräume U gilt:

$$\{v_i \mid i \in I \setminus \{k\}\} \subset U \implies v_k \in U.$$

Nach Voraussetzung ist $v_k = \sum_{j\in J} \lambda_j \cdot v_j$. Da $J \subset I \setminus \{k\}$, jedes v_j mit $j \in J$ liegt in U. Daraus folgt, dass $v_k \in U$.

Zu (iii) $\Rightarrow$ (i). Sei $k \in I\setminus J$. Da $v_k \in \mathrm{Span}_K(\{v_i | i \in J\})$ kann man nach Proposition 3.2.18 schreiben:

$$v_k = \sum_{s\in S} \lambda_s \cdot v_s,$$

wobei $S \subset J$ eine endliche Teilmenge ist und die λ_s Skalare sind. Dann gilt

$$(-1) \cdot v_k + \sum_{s\in S} \lambda_s \cdot v_s = 0.$$

Da $-1 \neq 0$ ist diese Linearkombination nicht trivial, also ist $(v_i)_{i\in I}$ linear abhängig. □

Um eine Charakterisierung der linearen *Un*abhängigkeit zu erhalten, brauchen wir eine Konstruktion.

Konstruktion 3.3.8. Sei $F = (v_i)_{i\in I}$ eine Familie von Vektoren in einem Vektorraum V über K. Diese Familie induziert eine Abbildung

$$\begin{aligned} \varphi_F \colon K^{(I)} &\to V, \\ (\lambda_i)_{i\in I} &\mapsto \sum_{i\in I} \lambda_i \cdot v_i. \end{aligned}$$

Dabei ist $K^{(I)}$ die Menge aller Abbildungen $I \to K$, die außerhalb einer endlichen Teilmenge von I null sind (siehe Beispiel 3.2.12). Obwohl I eine unendliche Menge sein könnte, ist die obige Summe $\sum_{i\in I} \lambda_i \cdot v_i$ sinnvoll, da nur endlich viele Summanden nicht null sind. Genauer könnte man schreiben:

$$\varphi_F((\lambda_i)_{i\in I}) = \sum_{i\in\{i\in I \mid \lambda_i\neq 0\}} \lambda_i \cdot v_i.$$

Man bemerkt, dass man die Familie F aus der Abbildung φ_F zurückbekommen kann, da $v_i = \varphi_F(e_i)$.

Proposition 3.3.9 (Charakterisierung der linearen Unabhängigkeit)**.** *Sei $F = (v_i)_{i\in I}$ eine Familie in einem K-Vektorraum V. Folgende Aussagen sind äquivalent:*

(i) *F ist linear unabhängig.*

(ii) *Die von F induzierte Abbildung $\varphi_F\colon K^{(I)} \to V$ ist injektiv.*

Beweis. Zu (i) $\Rightarrow$ (ii). Es seien $(\lambda_i)_{i\in I}$ und $(\mu_i)_{i\in I}$ zwei Elemente von $K^{(I)}$ mit

$$\sum_{i\in I} \lambda_i \cdot v_i = \sum_{i\in I} \mu_i \cdot v_i.$$

Diese Gleichung bedeutet, dass

$$\sum_{i\in J} \lambda_i \cdot v_i = \sum_{i\in J} \mu_i \cdot v_i,$$

wobei $J \subset I$ eine endliche Teilmenge ist, so dass $\lambda_i = \mu_i = 0$ für alle $i \in I \setminus J$. Daraus folgt:

$$\sum_{i\in J} (\lambda_i - \mu_i) \cdot v_i = 0.$$

Aus der linearen Unabhängigkeit von $(v_i)_{i\in I}$ folgt jetzt $\lambda_i - \mu_i = 0$ für alle $i \in J$. Also gilt $\lambda_i = \mu_i$ für alle $i \in I$, d.h., $(\lambda_i)_{i\in I} = (\mu_i)_{i\in I}$.

Zu (ii) $\Rightarrow$ (i). Sei $J \subset I$ eine endliche Teilmenge und $\sum_{j\in J} \lambda_j \cdot v_j$ eine Linearkombination, die gleich null ist. Zu zeigen ist, dass $\lambda_j = 0$ für alle $j \in J$. Man kann die Familie $(\lambda_j)_{j\in J}$ zu einer Familie $(\lambda_i)_{i\in I} \in K^{(I)}$ fortsetzen, indem man $\lambda_i = 0$ setzt, falls $i \in I \setminus J$. Die Abbildung $\varphi_F\colon K^{(I)} \to V$ bildet dann $(\lambda_i)_{i\in I}$ auf 0 ab. Sie bildet auch die Nullfamilie $(0)_{i\in I}$ auf 0 ab. Aus der Injektivität von φ_F folgt, dass $(\lambda_i)_{i\in I} = (0)_{i\in I}$. □

3.3.2 Basen

Definition 3.3.10 (erzeugende Familie)**.** Sei V ein Vektorraum über K. Eine Familie $(v_i)_{i\in I}$ in V heißt *erzeugend*, wenn $\{v_i \mid i \in I\}$ ein Erzeugendensystem ist. Man sagt auch, dass die Familie $(v_i)_{i\in I}$ selbst ein Erzeugendensystem ist.

Definition 3.3.11 (Basis)**.** Sei V ein Vektorraum über K. Eine *Basis* von V ist eine erzeugende linear unabhängige Familie in V.

Beispiel 3.3.12. Die Familie der Standardeinheitsvektoren $(e_1, \ldots, e_n)$ ist eine Basis von K^n: Sie ist erzeugend (Beispiel 3.2.19(ii)) und linear unabhängig (Beispiel 3.3.4). Diese Basis heißt *Standardbasis* von K^n.

Beispiel 3.3.13. Sei I eine beliebige Menge. Die Familie der Standardeinheitsvektoren $(e_i)_{i\in I}$ in K^I ist linear unabhängig (Beispiel 3.3.4) und erzeugt den Untervektorraum $K^{(I)}$ (Beispiel 3.2.19(iv)). Sie ist also eine Basis von $K^{(I)}$.

Bemerkung 3.3.14 (Unabhängigkeit der Reihenfolge)**.** Sei $(v_i)_{i\in I}$ eine Basis von V und $\sigma\colon I \to I$ eine Permutation von I (Definition 2.2.5). Dann ist $(v_{\sigma(i)})_{i\in I}$ wieder eine Basis von V. Insbesondere, für alle $\sigma \in S_n$, ist die Familie $(e_{\sigma(1)}, \ldots, e_{\sigma(n)})$ eine Basis von K^n.

Beispiel 3.3.15.

(i) Die folgenden Familien sind Basen von K^2: (e_2, e_1), $(-e_1, e_2)$, $(e_1, e_1 + e_2)$.

(ii) Ob die Familie $(e_1 + e_2, e_1 - e_2)$ eine Basis von K^2 ist, hängt von der Characteristik von K ab. Falls die Charakteristik von K gleich 2 ist, dann gilt $e_1 + e_2 = e_1 - e_2$, und

damit ist die Familie linear abhängig. Falls die Charakteristik von K nicht gleich 2 ist, also falls $2 \neq 0$ in K, dann existiert $1/2$ in K, und es gilt

$$e_1 = \frac{1}{2}((e_1 + e_2) + (e_1 - e_2)) \quad \text{und} \quad e_2 = \frac{1}{2}((e_1 + e_2) - (e_1 - e_2)).$$

Dies zeigt, dass die Familie $(e_1 + e_2, e_1 - e_2)$ erzeugend ist. Sie ist auch linear unabhängig, denn: Sei $\lambda, \mu \in K$ mit

$$\lambda(e_1 + e_2) + \mu(e_1 - e_2) = 0, \quad \text{d.h.,} \quad (\lambda + \mu)e_1 + (\lambda - \mu)e_2 = 0.$$

Nach der linearen Unabhängigkeit von (e_1, e_2) gilt $\lambda + \mu = 0$ und $\lambda - \mu = 0$. Aus der zweiten Gleichung folgt $\lambda = \mu$, und aus der ersten $2\lambda = 0$. Da $2 \neq 0$ in K, folgt $\lambda = 0$.

(iii) Die folgenden Familien sind Basen von K^3: (e_2, e_3, e_1), $(e_1, e_1 + e_2, e_1 + e_2 + e_3)$, $(e_1 + e_2, e_1 + e_3, e_1 + e_2 + e_3)$.

(iv) Die Familie $(e_1 + e_2, e_1 + e_3, e_2 + e_3)$ ist genau dann eine Basis von K^3, wenn die Charakteristik von K nicht gleich 2 ist.

Proposition 3.3.16. *Sei $(v_i)_{i\in I}$ eine Basis von einem K-Vektorraums V. Zu jedem Vektor $v \in V$ gibt es eine eindeutige Familie $(\lambda_i)_{i\in I}$ von Skalaren, die null außerhalb einer endlichen Teilmenge von I sind, so dass $v = \sum_{i\in I} \lambda_i \cdot v_i$.*

Beweis. Da $v \in \mathrm{Span}_K(\{v_i \mid i \in I\})$, eine solche Familie $(\lambda_i)_{i\in I}$ existiert nach Proposition 3.2.18. Die Eindeutigkeit der Familie folgt aus Proposition 3.3.9. □

Definition 3.3.17 (Koordinatenvektor)**.** Sei $B = (v_i)_{i\in I}$ eine Basis eines K-Vektorraums V und sei $v \in V$. Die Familie $(\lambda_i)_{i\in I} \in K^{(I)}$ aus Proposition 3.3.16 heißt der *Koordinatenvektor* von v bzgl. der Basis B und wird mit $[v]_B$ bezeichnet.

Bemerkung 3.3.18. Sei $F = (v_i)_{i\in I}$ eine Familie von Vektoren in V. Ob diese Familie erzeugend, linear unabhängig oder eine Basis ist, läßt sich durch die induzierte Abbildung

$$\begin{aligned} \varphi_F \colon K^{(I)} &\to V, \\ (\lambda_i)_{i\in I} &\mapsto \sum_{i\in I} \lambda_i \cdot v_i, \end{aligned}$$

aus Konstruktion 3.3.8 durchsichtig ausdrücken. Es gilt nämlich:

(i) F ist genau dann erzeugend, wenn φ_F surjektiv ist (nach Proposition 3.2.18).

(ii) F ist genau dann linear unabhängig, wenn φ_F injektiv ist (nach Proposition 3.3.9).

(iii) F ist genau dann eine Basis, wenn φ_F bijektiv ist (nach (i) und (ii)).

Falls F eine Basis ist, ist der Koordinatenvektor von einem $v \in V$ bzgl. F gleich $\varphi_F^{-1}(v) \in K^{(I)}$.

Proposition 3.3.19 (Charakterisierung von Basen)**.** *Sei V ein Vektorraum über K und $(v_i)_{i\in I}$ eine Familie in V. Dann sind die folgenden Aussagen äquivalent:*

(i) *$(v_i)_{i\in I}$ ist eine Basis von V.*

(ii) *$(v_i)_{i\in I}$ ist eine* maximale linear unabhängige Familie*, d.h., sie ist linear unabhängig, und für jede linear unabhängige Familie $(v_j)_{j\in J}$ mit $I \subset J$ gilt $I = J$.*

(iii) *$(v_i)_{i\in I}$ ist eine* minimale erzeugende Familie*, d.h., sie ist erzeugend, und für jede erzeugende Familie $(v_j)_{j\in J}$ mit $J \subset I$ gilt $I = J$.*

Beweis. Wir beweisen die beide Äquivalenzen (i) $\Leftrightarrow$ (ii) und (i) $\Leftrightarrow$ (iii).

Zu (i) $\Rightarrow$ (ii). Sei $(v_i)_{i\in I}$ eine Basis und $(v_j)_{j\in J}$ eine Familie mit $I \subset J$. Falls $I \neq J$, dann ist $(v_j)_{j\in J}$ linear abhängig nach Proposition 3.3.7 (iii) $\Rightarrow$ (i).

Zu (ii) $\Rightarrow$ (i). Zu zeigen ist, dass die Familie $(v_i)_{i\in I}$ erzeugend ist. Sei $v \in V$ ein beliebiger Vektor. Wenn man v zu der Familie $(v_i)_{i\in I}$ hinzufügt, erhält man aufgrund der Maximalität von $(v_i)_{i\in I}$ eine Familie, die linear abhängig ist. Es gibt also eine nicht-triviale Linearkombination der Vektoren v_i und v, die gleich null ist. Da $(v_i)_{i\in I}$ linear unabhängig ist, muss der Koeffizient von v in einer solchen Linearkombination nicht null sein. Daraus folgt, dass v eine Linearkombination der Vektoren v_i ist, wie gewünscht.

Zu (i) $\Rightarrow$ (iii). Sei $(v_i)_{i\in I}$ eine Basis und $J \subset I$ eine Teilmenge mit $J \neq I$. Zu zeigen ist, dass die Familie $(v_j)_{j\in J}$ nicht erzeugend ist. Aber wenn sie erzeugend wäre, dann wäre $(v_i)_{i\in I}$ linear abhängig sein, nach Proposition 3.3.7 (iii) $\Rightarrow$ (i).

Zu (iii) $\Rightarrow$ (i). Man hat zu zeigen, dass die Familie $(v_i)_{i\in I}$ linear unabhängig ist. Wenn nicht, dann existiert nach Proposition 3.3.7 (i) $\Rightarrow$ (iii) eine Teilmenge $J \subsetneq I$, so dass die Familie $(v_j)_{j\in J}$ erzeugend ist. Aber das steht im Widerspruch zur Minimalität von $(v_i)_{i\in I}$. □

Der folgende Satz ist einer der wichtigsten Struktursätze für Vektorräume. Wir werden ihn sehr häufig verwenden.

Satz 3.3.20 (Basisergänzungssatz)**.** *Sei V ein K-Vektorraum.*

(i) *Sei $(v_i)_{i\in I}$ eine erzeugende Familie in V und sei $J \subset I$ eine Teilmenge, so dass $(v_i)_{i\in J}$ linear unabhängig ist. Dann existiert eine Menge L mit $J \subset L \subset I$, so dass $(v_i)_{i\in L}$ eine Basis von V ist.*

Insbesondere:

(ii) *Jede erzeugende Familie $(v_i)_{i\in I}$ in V kann zu einer Basis eingeschränkt werden, d.h., es existiert eine Teilmenge $J \subset I$, so dass $(v_i)_{i\in J}$ eine Basis ist.*

(iii) *Jede linear unabhängige Familie $(v_i)_{i\in I}$ in V kann zu einer Basis ergänzt werden, d.h., es existiert eine Indexmenge $J \supset I$ und Vektoren v_j für $j \in J \setminus I$, so dass $(v_j)_{j\in J}$ eine Basis ist.*

Beweis. Aussage (ii) ist der Sonderfall von (i) mit $J = \varnothing$. Aussage (iii) folgt aus (i), indem wir zuerst die gegebene Familie $(v_i)_{i\in I}$ zu einer erzeugenden Familie ergänzen, z.B. zu einer Familie, die alle Vektoren aus V enthält.

Wir beweisen (i) zunächst im Spezialfall, dass I endlich ist, da der Beweis in diesem Fall einfacher ist. Wir beweisen die Existenz von L durch vollständige Induktion über die Mächtigkeit von $I \setminus J$ (Korollar 1.2.23). Falls $(v_i)_{i\in J}$ bereits erzeugend ist, kann man $L = J$ nehmen. Andernfalls gibt es einen Index $k \in I \setminus J$, so dass $v_k \notin \mathrm{Span}_K(\{v_i \mid i \in J\})$. Dann ist die Familie $(v_i)_{i\in J\cup\{k\}}$ linear unabhängig, denn: Sei

$$\sum_{i\in J\cup\{k\}} \lambda_i \cdot v_i = 0$$

mit $\lambda_i \in K$. Es gilt $\lambda_k = 0$, sonst wäre

$$v_k = -\sum_{i\in J} \frac{\lambda_i}{\lambda_k} \cdot v_i$$

und damit würde v_k in $\mathrm{Span}_K(\{v_i \mid i \in J\})$ liegen. Alle anderen λ_i sind dann auch null, da $(v_i)_{i\in J}$ linear unabhängig ist. Die Mächtigkeit von $I \setminus (J \cup \{k\})$ ist kleiner als die von $I \setminus J$. Nach der Induktionsvoraussetzung gibt es eine Menge L mit $J \cup \{k\} \subset L \subset I$, so dass $(v_i)_{i\in L}$ eine Basis von V ist. Damit ist (i) im Fall einer endlichen Menge I bewiesen.

Wir beweisen jetzt den allgemeinen Fall. Sei

$$\mathcal{A} := \{L \mid J \subset L \subset I \text{ und } (v_i)_{i\in L} \text{ ist linear unabhängig}\} \subset \mathcal{P}(I).$$

Wir betrachten $\mathcal{A}$ als partiell geordnete Menge bezüglich der Inklusionsrelation $\subset$. Wir behaupten, dass $\mathcal{A}$ ein maximales Element $L_{\max}$ besitzt, und dass die Familie $(v_i)_{i\in L_{\max}}$ eine Basis ist. Um zu zeigen, dass $\mathcal{A}$ ein maximales Element besitzt, verwenden wir das Zornsche Lemma 1.4.21: Es genügt zu zeigen, dass jede Kette $\mathcal{B} \subset \mathcal{A}$ eine obere Schranke besitzt. Falls $\mathcal{B} = \varnothing$, dann ist J eine obere Schranke von $\mathcal{B}$. Andernfalls betrachten wir die Vereinigung $B = \bigcup_{L\in\mathcal{B}} L$. Da $\mathcal{B}$ total geordnet und nicht leer ist, gibt es zu jeder endlichen Teilmenge $E \subset B$ ein $L \in \mathcal{B}$ mit $E \subset L$ (das kann mann leicht durch Induktion über $|E|$ nachprüfen). Für jede endliche Teilmenge $E \subset B$ ist also die Familie $(v_i)_{i\in E}$ linear unabhängig, und daher ist die ganze Familie $(v_i)_{i\in B}$ linear unabhängig. Das heißt: Es gilt $B \in \mathcal{A}$, und damit ist B eine obere Schranke von $\mathcal{B}$.

Nach dem Zornschen Lemma existiert also ein maximales Element $L_{\max} \in \mathcal{A}$. Es bleibt zu zeigen, dass $\{v_i \mid i \in L_{\max}\}$ ein Erzeugendensystem ist. Nach Voraussetzung ist $\{v_i \mid i \in I\}$ ein Erzeugendensystem. Deswegen genügt es zu zeigen, dass $v_k \in \mathrm{Span}_K(\{v_i \mid i \in L_{\max}\})$ für jedes $k \in I \setminus L_{\max}$. Da $L_{\max}$ maximal in $\mathcal{A}$ ist, ist die Familie $(v_i)_{i\in L_{\max}\cup\{k\}}$ linear abhängig: Es gibt Skalare λ_i, $i \in L_{\max} \cup \{k\}$, die nicht alle null sind, so dass

$$\sum_{i\in L_{\max}\cup\{k\}} \lambda_i \cdot v_i = 0.$$

Es muss eigentlich $\lambda_k \neq 0$ gelten, da $(v_i)_{i\in L_{\max}}$ linear unabhängig ist. Wir dürfen also durch λ_k dividieren, und erhalten

$$v_k = -\sum_{i\in L_{\max}} \frac{\lambda_i}{\lambda_k} \cdot v_i.$$

Also gilt $v_k \in \mathrm{Span}_K(\{v_i \mid i \in L_{\max}\})$, wie gewünscht. □

Bemerkung 3.3.21. Die Beweise des obigen Satzes im Fall einer endlichen Teilmenge I und im allgemeinen Fall waren sehr ähnlich. Der Unterschied war nur, dass wir im allgemeinen Fall das Induktionsprinzip durch das Zornsche Lemma ersetzen mussten. Das ist eigentlich ein typisches Anwendung des Zornschen Lemmas: Es ist oft der Fall, dass Aussagen, die bei endlichen Mengen durch Induktion bewiesen werden können, auch bei unendlichen Mengen mit dem Zornschen Lemma bewiesen werden können.

Beispiel 3.3.22. Die Familie $(e_1, e_2, e_3, e_1+e_2+e_3)$ in K^3 ist erzeugend, und die Teilfamilie $(e_1, e_1 + e_2 + e_3)$ ist linear unabhängig. Nach Satz 3.3.20(i) gibt es eine Basis zwischen den beiden Familien. In diesem Fall sind beide Familien $(e_1, e_2, e_1+e_2+e_3)$ und $(e_1, e_3, e_1+e_2+e_3)$ Basen von K^3.

Korollar 3.3.23 (Existenz von Basen)**.**

(i) *Jeder K-Vektorraum besitzt eine Basis.*

(ii) *Jeder endlich erzeugte K-Vektorraum besitzt eine endliche Basis.*

Beweis. Sei V ein K-Vektorraum und sei $(v_i)_{i\in I}$ eine erzeugende Familie, z.B. $(v)_{v\in V}$. Nach Satz 3.3.20(ii), kann diese Familie zu einer Basis eingeschränkt werden. Falls V endlich erzeugt ist, gibt es eine solche Familie mit einem endlichen I, and damit erhalten wir eine endliche Basis. □

Lemma 3.3.24 (Austauschlemma)**.** *Sei V ein K-Vektorraum, $(v_i)_{i\in I}$ eine Basis von V und $w \in V$. Sei $I' \subset I$ eine Teilmenge, so dass $w \notin \mathrm{Span}_K(\{v_i \mid i \in I'\})$; zum Beispiel, $I' = \varnothing$ und $w \neq 0$. Dann existiert ein Index $k \in I \setminus I'$, so dass die Familie, die sich aus $(v_i)_{i\in I}$ ergibt, wenn v_k gegen w ausgetauscht wird, wieder eine Basis von V ist.*

Beweis. Da $(v_i)_{i\in I}$ eine Basis ist, kann man schreiben

$$w = \sum_{i\in I} \lambda_i \cdot v_i, \tag{3.3.25}$$

wobei die Skalare $\lambda_i \in K$ alle null sind, außer endlich viele. Es gibt dann ein $k \in I \setminus I'$ mit $\lambda_k \neq 0$, sonst würde w in $\mathrm{Span}_K(\{v_i \mid i \in I'\})$ liegen. Sei $(\tilde{v}_i)_{i\in I}$ die Familie mit

$$\tilde{v}_i = \begin{cases} v_i, & \text{falls } i \neq k, \\ w, & \text{falls } i = k. \end{cases}$$

Die Famile $(\tilde{v}_i)_{i\in I}$ ist erzeugend, da

$$v_k = \frac{1}{\lambda_k}\left(w - \sum_{i\in I\setminus\{k\}} \lambda_i \cdot v_i\right) \in \mathrm{Span}_K(\{\tilde{v}_i \mid i \in I\}).$$

Die Familie $(\tilde{v}_i)_{i\in I}$ ist linear unabhängig, denn: Sei

$$\sum_{i\in I} \mu_i \cdot \tilde{v}_i = 0 \tag{3.3.26}$$

mit $(\mu_i)_{i\in I} \in K^{(I)}$. Aus (3.3.25) und (3.3.26) folgt:

$$0 = \mu_k \cdot w + \sum_{i\in I\setminus\{k\}} \mu_i \cdot v_i = (\mu_k \cdot \lambda_k) \cdot v_k + \sum_{i\in I\setminus\{k\}} (\mu_k \cdot \lambda_i + \mu_i) \cdot v_i.$$

Da $(v_i)_{i\in I}$ linear unabhängig ist, sind $\mu_k \cdot \lambda_k$ und $\mu_k \cdot \lambda_i + \mu_i$ mit $i \in I \setminus \{k\}$ alle null. Aus $\lambda_k \neq 0$ folgt jetzt, dass $\mu_k = 0$, und daher auch dass $\mu_i = 0$ für alle $i \in I \setminus \{k\}$. □

Satz 3.3.27 (alle Basen haben dieselbe Länge, endlicher Fall)**.** *Sei V ein endlich erzeugter K-Vektorraum.*

(i) *Ist $(v_i)_{i\in I}$ eine Basis von V, so ist die Menge I endlich.*

(ii) *Sind $(v_1, \ldots, v_n)$ und $(w_1, \ldots, w_m)$ zwei Basen von V, so gilt $n = m$.*

Beweis. Da V endlich erzeugt ist, existiert nach Korollar 3.3.23 eine Basis $(b_1, \ldots, b_n)$ von V mit $n \in \mathbb{N}$. Sei $(v_i)_{i\in I}$ eine beliebige Basis von V. Wir behaupten, dass es paarweise verschiedene Indizes $i_1, \ldots, i_n \in I$ gibt, so dass die Familie $(\tilde{v}_i)_{i\in I}$ mit

$$\tilde{v}_i = \begin{cases} b_k, & \text{falls } i = i_k \text{ mit } k \in \{1, \ldots, n\}, \\ v_i & \text{andernfalls,} \end{cases}$$

eine Basis von V ist. Dazu verwenden wir n-mal das Austauschlemma: Sind die Indizes $i_1, \ldots, i_{k-1}$ schon gefunden, wenden wir das Austauschlemma mit $I' = \{i_1, \ldots, i_{k-1}\}$ und $w = b_k$ an, um das Index i_k zu erhalten. Da $(b_1, \ldots, b_n)$ eine *maximale* linear unabhängige Familie ist (Proposition 3.3.19), folgt daraus, dass $I = \{i_1, \ldots, i_n\}$. Insbesondere ist I endlich der Mächtigkeit n, was beide (i) und (ii) beweist. □

Bemerkung 3.3.28. Ist I eine Menge, so hat $K^{(I)}$ die Basis $(e_i)_{i\in I}$ (Beispiel 3.3.13). Nach Satz 3.3.27(i), falls I unendlich ist, dann ist $K^{(I)}$ nicht endlich erzeugt. Da die Familie $(e_i)_{i\in I}$ zu einer Basis von K^I ergänzt werden kann (Satz 3.3.20(iii)), ist K^I ebenfalls nicht endlich erzeugt.

Man kann Satz 3.3.27(ii) auf beliebige Vektorräume verallgemeinern, mithilfe des Begriffs der Gleichmächtigkeit (Definition 1.3.28):

Satz 3.3.29 (alle Basen haben dieselbe Länge, allgemeiner Fall)**.** *Seien $(v_i)_{i\in I}$ und $(w_j)_{j\in J}$ zwei Basen eines K-Vektorraums V. Dann sind I und J gleichmächtig.*

Beweis. Falls V endlich erzeugt ist, folgt dies bereits aus dem Satz 3.3.27. Wir dürfen deshalb annehmen, dass I und J unendlich sind. Jedes v_i läßt sich eindeutig als Linearkombination der Vektoren w_j schreiben. Sei $J_i \subset J$ die endliche Teilmenge aller Indizes j, so dass der Koeffizient von w_j in dieser Linearkombination nicht null ist. Es gilt dann $v_i \in \mathrm{Span}_K(\{w_j \mid j \in J_i\})$ und daher

$$V = \mathrm{Span}_K(\{v_i \mid i \in I\}) \subset \mathrm{Span}_K\left(\left\{ w_j \;\middle|\; j \in \bigcup_{i\in I} J_i \right\}\right).$$

Da $(w_j)_{j\in J}$ eine Basis ist, und damit eine *minimale* erzeugende Familie (Proposition 3.3.19), gilt $\bigcup_{i\in I} J_i = J$. Aus dem Satz 1.3.36 folgt, dass $|J| \leq |I|$. Auf symmetrische Weise folgt $|I| \leq |J|$. Nach dem Satz von Cantor–Bernstein–Schröder (Satz 1.3.33) sind also I und J gleichmächtig. □

3.3.3 Dimension

Definition 3.3.30 (Dimension, endlich-dimensional, unendlich-dimensional)**.** Sei V ein K-Vektorraum. Die *Dimension* von V,

$$\dim_K(V) \in \mathbb{N} \cup \{\infty\},$$

wird wie folgt definiert:

- Falls V endlich erzeugt ist, ist $\dim_K(V) \in \mathbb{N}$ die Länge irgendeiner Basis von V, die nach Satz 3.3.27 eine wohldefinierte natürliche Zahl ist. In diesem Fall sagt man auch, dass V *endlich-dimensional* ist.
- Falls V nicht endlich erzeugt ist, setzt man $\dim_K(V) := \infty$. In diesem Fall heißt V *unendlich-dimensional*.

Bemerkung 3.3.31. Die Dimension eines unendlich-dimensionalen K-Vektorraums V kann genauer als die Mächtigkeit irgendeiner Basis von V definiert werden, die nach Satz 3.3.29 wohldefiniert ist.

Beispiel 3.3.32.

(i) Es gilt $\dim_K(K^n) = n$, da $(e_1, \ldots, e_n)$ eine Basis von K^n ist. Insbesondere gilt $\dim_K(K) = 1$.

(ii) Es gilt $\dim_K(V) = 0$ genau dann, wenn $V = \{0\}$ (die leere Familie ist eine Basis des trivialen Vektorraums $\{0\}$).

(iii) Ist $(v_1, \ldots, v_n)$ eine linear unabhängige Familie in einem K-Vektorraum V, so ist die Dimension des Untervektorraums $\mathrm{Span}_K(\{v_1, \ldots, v_n\})$ gleich n.

(iv) Sei I eine Menge. Der K-Vektorraum $K^{(I)}$ hat dann die Basis $(e_i)_{i\in I}$. Falls I unendlich ist, ist also $\dim_K(K^{(I)}) = \infty$. Im Sinne der Bemerkung 3.3.31 ist genauer die Dimension von $K^{(I)}$ gleich der Mächtigkeit von I.

Bemerkung 3.3.33 (Abhängigkeit vom Grundkörper)**.** Wenn $K \subset L$ eine Körpererweiterung ist, kann jeder L-Vektorraum V als K-Vektorraum betrachtet werden (siehe Bemerkung 3.2.15). Es gilt dann $\dim_L(V) \leq \dim_K(V)$, da jede L-Basis von V auch K-linear unabhängig ist, und daher zu einer K-Basis ergänzt werden kann (nach Satz 3.3.20(iii)). Im Allgemeinen ist aber $\dim_L(V) \neq \dim_V(K)$. Zum Beispiel, $\dim_{\mathbb{C}}(\mathbb{C}) = 1$ aber $\dim_{\mathbb{R}}(\mathbb{C}) = 2$: $(1, i)$ ist eine Basis von $\mathbb{C}$ als $\mathbb{R}$-Vektorraum.

Proposition 3.3.34 (Basen in endlich-dimensionalen Vektorräumen)**.** *Sei V ein Vektorraum über K mit $\dim_K(V) = n < \infty$, und sei $B = (v_1, \dots, v_n)$ eine Familie von n Vektoren aus V. Folgende Aussagen sind äquivalent:*

(i) *B ist eine Basis.*

(ii) *B ist erzeugend.*

(iii) *B ist linear unabhängig.*

Beweis. Nach Definition gelten (i) $\Rightarrow$ (ii) und (i) $\Rightarrow$ (iii). Falls B erzeugend bzw. linear unabhängig ist, dann kann B nach Satz 3.3.20 zu einer Basis B' eingeschränkt bzw. ergänzt werden. Die Basis B' muss nach Satz 3.3.27 aus n Vektoren bestehen, also gilt $B = B'$. Insbesondere war B bereits eine Basis. □

Proposition 3.3.35 (Dimension von Untervektorräumen)**.** *Sei V ein endlich-dimensionaler K-Vektorraum und $U \subset V$ ein Untervektorraum. Dann ist U auch endlich-dimensional und $\dim_K(U) \leq \dim_K(V)$. Falls $U \neq V$ gilt eigentlich $\dim_K(U) < \dim_K(V)$.*

Beweis. Nach Korollar 3.3.23 besitzt U eine Basis $(u_i)_{i\in I}$. Die Familie $(u_i)_{i\in I}$ ist insbesondere linear unabhängig, und kann nach Satz 3.3.20(iii) zu einer Basis von V ergänzt werden, die endlich ist nach Satz 3.3.27(i). Dies zeigt, dass I endlich ist und dass $\dim_K(U) \leq \dim_K(V)$.

Zur letzten Aussage beweisen wir die Kontraposition. Falls $\dim_K(V) = \dim_K(U)$, dann muss $(u_i)_{i\in I}$ bereits eine Basis von V sein, und daher muss $U = V$ sein. □

Definition 3.3.36 (Summe von Untervektorräumen)**.** Sei V ein K-Vektorraum und seien $U, W \subset V$ zwei Untervektorräume. Die *Summe* von U und W ist der Untervektorraum

$$U + W := \mathrm{Span}_K(U \cup W) \subset V.$$

Nach Proposition 3.2.18 gilt

$$U + W = \{u + w \mid u \in U \text{ und } w \in W\}.$$

Satz 3.3.37 (Dimensionsformel für Untervektorräume)**.** *Sei V ein endlich-dimensionaler K-Vektorraum und seien $U, W \subset V$ Untervektorräume. Dann gilt*

$$\dim_K(U + W) = \dim_K(U) + \dim_K(W) - \dim_K(U \cap W).$$

Es ist hilfreich, diese Dimensionsformel im Fall $K = \mathbb{R}$ und $V = \mathbb{R}^3$ explizit zu untersuchen. Untervektorraüme von $\mathbb{R}^3$ sind $\{0\}$, Ursprungsgeraden, Ursprungsebenen, und $\mathbb{R}^3$ selbst. Seien zum Beispiel $U, W \subset \mathbb{R}^3$ zwei Ursprungsebenen. Falls $U \neq W$, dann ist $U \cap W$ eine Gerade und ist $U + W = \mathbb{R}^3$, und die Dimensionsformel lautet $3 = 2 + 2 - 1$. Falls $U = W$, dann gilt $U = W = U \cap W = U + W$, und die Dimensionsformel lautet $2 = 2+2-2$. Es ist natürlich unmöglich, dass $U \cap W = \{0\}$; das läßt sich auch aus der Dimensionsformel ableiten, da $2 + 2 - 0 = 4$ aber $\dim_{\mathbb{R}}(U + W) \leq \dim_{\mathbb{R}}(\mathbb{R}^3) = 3$. In höherer Dimension gibt es jedoch Ebenen, die nur in einem Punkt treffen, z.B. die Ebenen $\mathrm{Span}_{\mathbb{R}}(\{e_1, e_2\})$ und $\mathrm{Span}_{\mathbb{R}}(\{e_3, e_4\})$ in $\mathbb{R}^4$.

Beweis. Es folgt aus Proposition 3.3.35, dass U, V, $U + V$, und $U \cap V$ endlich-dimensional sind. Sei $(v_1, \dots, v_n)$ eine Basis von $U \cap W$. Nach Satz 3.3.20(iii) können wir diese Basis zu einer Basis $(v_1, \dots, v_n, u_1, \dots, u_p)$ von U und zu einer Basis $(v_1, \dots, v_n, w_1, \dots, w_q)$ von W ergänzen. Dann gilt $\dim_K(U) = n + p$, $\dim_K(W) = n + q$, $\dim_K(U \cap W) = n$. Wir müssen also zeigen, dass

$$\dim_K(U + W) = (n + p) + (n + q) - n = n + p + q.$$

Dazu zeigen wir, dass die Familie

$$(v_1, \dots, v_n, u_1, \dots, u_p, w_1, \dots, w_q)$$

eine Basis von $U + W$ ist. Sie erzeugt $U + W$, weil sie Basen von U sowie W enthält. Es bleibt die lineare Unabhängigkeit nachzuprüfen. Sei also

$$\underbrace{\sum_{i=1}^{n} \lambda_i \cdot v_i}_{\in U \cap W} + \underbrace{\sum_{j=1}^{p} \mu_j \cdot u_j}_{\in U} + \underbrace{\sum_{k=1}^{q} \nu_k \cdot w_k}_{\in W} = 0 \tag{3.3.38}$$

mit $\lambda_i, \mu_j, \nu_k \in K$. Aus dieser Gleichung folgt, dass die Summe $\sum_{j=1}^{p} \mu_j \cdot u_j$ in $U \cap W$ liegt. Da $(v_1, \ldots, v_n)$ eine erzeugende Familie von $U \cap V$ ist, gibt es Skalare $\lambda'_i \in K$, so dass

$$\sum_{j=1}^{p} \mu_j \cdot u_j = \sum_{i=1}^{n} \lambda'_i \cdot v_i.$$

Aus der linearen Unabhängigkeit von $(v_1, \ldots, v_n, u_1, \ldots, u_p)$ folgt insbesondere, dass $\mu_1 = \cdots = \mu_p = 0$. Wenn wir die Rollen von U und W vertauschen, erhalten wir gleichfalls $\nu_1 = \cdots = \nu_q = 0$. Von der Gleichung (3.3.38) bleibt übrig

$$\sum_{i=1}^{n} \lambda_i \cdot v_i = 0.$$

Da $(v_1, \ldots, v_n)$ linear unabhängig ist, folgt schließlich $\lambda_1 = \cdots = \lambda_n = 0$. □

Definition 3.3.39 (komplementäre Untervektorräume)**.** Sei V ein K-Vektorraum. Zwei Untervektorräume $U, W \subset V$ heißen *komplementär*, wenn folgende Bedingungen gelten:

$$U + W = V \quad \text{und} \quad U \cap W = \{0\}.$$

Man sagt auch, dass W zu U in V komplementär ist, oder dass W ein *direktes Komplement* von U in V ist.

Beispiel 3.3.40. Sei $(v_i)_{i \in I}$ eine Basis eines K-Vektorraums V und seien $I', I'' \subset I$ disjunkte Teilmengen mit $I = I' \cup I''$. Dann sind die Untervektorräume $\mathrm{Span}_K(\{v_i \,|\, i \in I'\})$ und $\mathrm{Span}_K(\{v_i | i \in I''\})$ komplementär. Beispielsweise sind $\mathrm{Span}_K(\{e_1, e_3\})$ und $\mathrm{Span}_K(\{e_2, e_4\})$ komplementäre Untervektorräume von K^4.

Proposition 3.3.41 (Charakterisierung von komplementären Untervektorräumen)**.** *Sei V ein endlich-dimensionaler K-Vektorraum, und seien $U, W \subset V$ Untervektorräume. Die folgenden Aussagen sind äquivalent:*

(i) *U und W sind komplementär.*

(ii) *$U + W = V$ und $\dim_K(V) = \dim_K(U) + \dim_K(W)$.*

(iii) *$U \cap W = \{0\}$ und $\dim_K(V) = \dim_K(U) + \dim_K(W)$.*

Beweis. Die Implikationen (i) ⇒ (ii) und (i) ⇒ (iii) folgen aus dem Satz 3.3.37 und der Definition von komplementären Untervektorräumen. Wir nehmen jetzt an, dass $\dim_K(V) = \dim_K(U) + \dim_K(W)$. Nach dem Satz 3.3.37 gilt dann

$$\dim_K(V) = \dim_K(U + W) + \dim_K(U \cap W).$$

Falls $U + W = V$, dann gilt $\dim_K(U \cap W) = 0$, d.h., $U \cap W = \{0\}$. Falls $U \cap W = \{0\}$, dann gilt $\dim_K(U + W) = \dim_K(V)$, und daher $U + W = V$ nach Proposition 3.3.35. □

Beispiel 3.3.42. Eine Ursprungsgerade G und eine Ursprungsebene E in $\mathbb{R}^3$ sind genau dann komplementär, wenn $G \cap E = \{0\}$.

Proposition 3.3.43 (Existenz von komplementären Untervektorräumen)**.** *Sei V ein K-Vektorraum und $U \subset V$ ein Untervektorraum. Dann existiert ein Untervektorraum $W \subset V$, der komplementär zu U in V ist.*

Beweis. Sei $(v_i)_{i\in I}$ eine Basis von U. Nach Satz 3.3.20(iii) gibt es eine Menge $J \supset I$ und Vektoren $v_j \in V$ für alle $j \in J \setminus I$, so dass die ergänzte Familie $(v_j)_{j\in J}$ eine Basis von V ist. Sei dann W der von $\{v_j \mid j \in J \setminus I\}$ erzeugte Untervektorraum von V. Nach Konstruktion gilt $U + W = V$, und es bleibt zu zeigen, dass $U \cap W = \{0\}$. Jedes $v \in U \cap W$ kann als

$$v = \sum_{j\in I} \lambda_j \cdot v_j \quad \text{und} \quad v = \sum_{j\in J\setminus I} \lambda_j \cdot v_j$$

dargestellt werden. Aus der linearen Unabhängigkeit von $(v_j)_{j\in J}$ folgt, dass alle λ_j null sind, und daher dass $v = 0$. □

Bemerkung 3.3.44. Im Allgemeinen hat ein Untervektorraum $U \subset V$ viele verschiedene komplementäre Untervektorräume. Zum Beispiel, wenn $U \subset K^2$ die von e_1 aufgespannte Gerade ist, dann ist jede andere Ursprungsgerade komplementär zu U. Man darf also nicht von *dem* komplementären Untervektorraum sprechen.

Es ist möglich, je zwei Vektorräume U und W als komplementäre Untervektoräume eines größeren Vektorraum $U \oplus W$ zu betrachten:

Definition 3.3.45 (direkte Summe)**.** Seien U und W Vektorräume über K. Die *direkte Summe* $U \oplus W$ von U und W ist das kartesische Produkt $U \times W$, versehen mit den komponentenweisen Addition und Skalarmultiplikation, d.h.:

$$(u, w) + (u', w') = (u + u', w + w')$$
$$\lambda \cdot (u, w) = (\lambda u, \lambda w).$$

Man kann leicht nachprüfen, dass $U \oplus W$ mit diesen Verknüpfungen ein K-Vektorraum ist (siehe Lemma 3.1.3 für einen ähnlichen Beweis). Die direkte Summe $U \oplus W$ heißt auch das *Produkt* von U und W, und kann auch mit $U \times W$ bezeichnet werden. Nach Proposition 3.2.8 sind $U \times \{0\}$ und $\{0\} \times W$ Untervektorräume von $U \oplus W$, und es ist klar, dass sie komplementäre Untervektorräume sind. Außerdem können U und W durch die injektiven Abbildungen

$$\begin{aligned} U &\to U \oplus W, & W &\to U \oplus W, \\ u &\mapsto (u, 0) & w &\mapsto (0, w) \end{aligned}$$

mit diesen Untervektorräumen von $U \oplus W$ identifiziert werden. Das heißt, wenn wir den Vektor $u \in U$ mit dem Paar $(u, 0)$ identifizieren, dann ist die Addition bzw. die Skalarmultiplikation auf U dieselbe wie auf dem Untervektorraum $U \times \{0\} \subset U \oplus W$. Solche Identifizierungen werden wir später mit dem Begriff des Isomorphismus präziser machen (siehe Definition 4.1.14). Es gilt insbesondere

$$\dim_K(U \oplus W) = \dim_K(U) + \dim_K(W)$$

nach Satz 3.3.37.

Proposition 3.3.46 (Dimensionsformel für Quotientenvektorräume)**.** *Sei V ein endlichdimensionaler K-Vektorraum und $U \subset V$ ein Untervektorraum. Dann gilt*

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

Beweis. Nach Satz 3.3.20(iii) gibt es eine Basis $(v_i)_{i\in I}$ von V und eine Teilmenge $J \subset I$, so dass $(v_i)_{i\in J}$ eine Basis von U ist. Es genügt zu zeigen, dass $(v_i + U)_{i\in I\setminus J}$ eine Basis von V/U ist.

- $\{v_i + U \mid i \in I \setminus J\}$ *ist ein Erzeugendensystem.* Sei $v = \sum_{i \in I} \lambda_i v_i$ ein beliebiger Vektor aus V. Die Differenz $v - \sum_{i \in I \setminus J} \lambda_i v_i$ liegt in U, und daher gilt

$$v + U = \left(\sum_{i \in I \setminus J} \lambda_i \cdot v_i \right) + U = \sum_{i \in I \setminus J} \lambda_i \cdot (v_i + U)$$

in V/U.

- $(v_i + U)_{i \in I \setminus J}$ *ist linear unabhängig.* Es sei

$$\sum_{i \in I \setminus J} \lambda_i \cdot (v_i + U) = 0 + U$$

mit $\lambda_i \in K$. Diese Gleichung bedeutet, dass $\sum_{i \in I \setminus J} \lambda_i v_i \in U$. Es gibt also Skalare $\lambda_i \in K$ für alle $i \in J$, so dass

$$\sum_{i \in I \setminus J} \lambda_i \cdot v_i = \sum_{i \in J} \lambda_i \cdot v_i.$$

Aus der linearen Unabhängigkeit von $(v_i)_{i \in I}$ folgt, dass $\lambda_i = 0$ für alle $i \in I$, und insbesondere für alle $i \in I \setminus J$, wie gewünscht. □

Kapitel 4

Lineare Abbildungen

In diesem Kapitel wird ein Grundkörper K immer wieder festgelegt.

4.1 Lineare Abbildungen

Wir fangen mit der Definition an:

Definition 4.1.1 (lineare Abbildung)**.** Seien V, W zwei Vektorräume über K. Eine *lineare Abbildung*, oder genauer *K-lineare Abbildung*, von V nach W ist eine Abbildung $f\colon V \to W$ mit folgenden Eigenschaften:

(i) Für alle $v, v' \in V$ gilt
$$f(v + v') = f(v) + f(v').$$

(ii) Für alle $v \in V$ und $\lambda \in K$ gilt
$$f(\lambda \cdot v) = \lambda \cdot f(v).$$

Lineare Abbildungen heißen auch *Vektorraumhomorphismen*. Die Menge aller K-linearen Abbildungen von V nach W wird mit $\mathrm{Hom}_K(V, W)$ bezeichnet.

Bemerkung 4.1.2. Die linearen Abbildungen zwischen Vektorräumen sind die Abbildungen, die mit der Vektorraumstruktur verträglich sind. Jedes Mal, wenn wir eine Art von „Mengen mit Struktur“ einführen, wie z.B. Gruppen, Körper, Vektorräume, partiell geordnete Mengen usw., gibt es normalerweise eine entsprechende Art von Abbildungen zwischen denen, die mit dieser Struktur verträglich sind. Zum Beispiel:

- Seien G und H Gruppen. Ein *Gruppenhomomorphismus* von G nach H ist eine Abbildung $f\colon G \to H$, so dass für alle $g, g' \in G$ gilt: $f(g \cdot g') = f(g) \cdot f(g')$.
- Seien K und L Körper. Ein *Körperhomomorphismus* von K nach L ist eine Abbildung $f\colon K \to L$, die mit beiden Verknüpfungen $+$ und $\cdot$ verträglich ist und außerdem 1 auf 1 abbildet.
- Seien X und Y partiell geordnete Mengen. Eine *monotone Abbildung* (oder *Ordnungshomomorphismus*) von X nach Y ist eine Abbildung $f\colon X \to Y$, so dass für alle $x, x' \in X$ gilt: $x \leq x' \Rightarrow f(x) \leq f(y)$.

Dieses Phänomen ist der Ausgangspunkt der *Kategorientheorie*. Man beachte, dass eine K-lineare Abbildung von V nach W insbesondere ein Gruppenhomomorphismus von $(V, +)$ nach $(W, +)$ ist.

Beispiel 4.1.3 (Skalierung und Verschiebung)**.** Sei V ein K-Vektorraum.

(i) Für jeden Skalar $\lambda_0 \in K$, die Skalierungsabbildung

$$\begin{aligned} V &\to V, \\ v &\mapsto \lambda_0 \cdot v, \end{aligned}$$

ist linear. Bedingungen (i) und (ii) der Definition 4.1.1 folgen aus Axiomen (iv) und (ii) der Definition 3.2.1.

(ii) Sei $v_0 \in V$ ein Vektor. Ist $v_0 \neq 0$, so ist die Verschiebungsabbildung

$$\begin{aligned} V &\to V, \\ v &\mapsto v + v_0, \end{aligned}$$

nicht linear. Zum Beispiel, $(v + v') + v_0 \neq (v + v_0) + (v' + v_0) = (v + v') + 2v_0$.

Beispiel 4.1.4 (generische Beispiele)**.** Seien V, W Vektorräume über K und sei $U \subset V$ ein Untervektorraum. Die folgenden Abbildungen sind K-linear:

(i) Die Identität $\mathrm{id}_V \colon V \to V$.

(ii) Die *Nullabbildung* $0 \colon V \to W$, $v \mapsto 0$.

(iii) Die Inklusionsabbildung $U \to V$, $u \mapsto u$.

(iv) Die Quotientenabbildung $V \to V/U$, $v \mapsto v + U$.

(v) Die kanonischen Abbildungen

$$\begin{aligned} \iota_1 &\colon V \to V \oplus W, \quad v \mapsto (v, 0), \\ \iota_2 &\colon W \to V \oplus W, \quad w \mapsto (0, w). \end{aligned}$$

(vi) Die kanonischen Abbildungen

$$\begin{aligned} \pi_1 &\colon V \oplus W \to V, \quad (v, w) \mapsto v, \\ \pi_2 &\colon V \oplus W \to W, \quad (v, w) \mapsto w. \end{aligned}$$

Beispiel 4.1.5. Sei $F = (v_i)_{i \in I}$ eine Familie von Vektoren in einem K-Vektorraum V. Dann ist die Abbildung

$$\begin{aligned} \varphi_F &\colon K^{(I)} \to V, \\ (\lambda_i)_{i \in I} &\mapsto \sum_{i \in I} \lambda_i \cdot v_i, \end{aligned}$$

aus Konstruktion 3.3.8 K-linear.

Beispiel 4.1.6. Ob die Abbildung

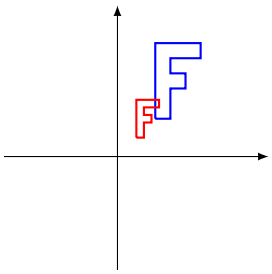
$$f \colon K \to K, \quad f(x) = x^2,$$

K-linear ist, hängt von dem Körper K ab. Es gilt $f(1 + 1) = 4$ und $f(1) + f(1) = 2$. Wenn die Charakteristik von K nicht 2 ist, dann ist $4 \neq 2$ (da $4 - 2 = 2 \neq 0$), und damit ist f auf jeden Fall keine lineare Abbildung. Aber ist $K = \mathbb{F}_2$, so ist f gleich der Identität (da $0^2 = 0$ und $1^2 = 1$), und insbesondere linear. Man kann jedoch zeigen, dass $\mathbb{F}_2$ der einzige Körper ist (bis auf Isomorphie), auf dem die Abbildung f K-linear ist. Zum Beispiel, wenn $K = \mathbb{F}_4$ (siehe Bemerkung 2.4.11), dann gilt $f(\alpha \cdot 1) = \alpha^2 = \beta \neq \alpha = \alpha \cdot f(1)$, und damit ist f nicht K-linear.

Beispiel 4.1.7 (lineare Abbildungen von $\mathbb{R}^2$ nach $\mathbb{R}^2$)**.** Beispiele von $\mathbb{R}$-linearen Abbildungen $\mathbb{R}^2 \to \mathbb{R}^2$ sind Skalierungen, Spiegelungen (an einer Ursprungsgerade), Drehungen (um den Ursprung), Scherungen, usw. In folgenden Beispielen, die rote Figur ist das Bild der blauen Figur unter der gegebenen linearen Abbildung.

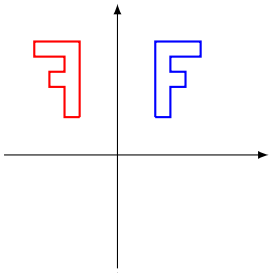
(i) Skalierung um $\frac{1}{2}$:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \frac{1}{2} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$



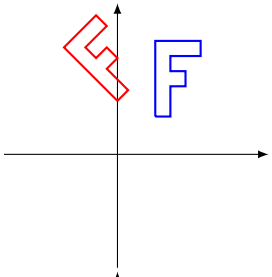
(ii) Spiegelung an $\mathbb{R}e_2$:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} -x_1 \\ x_2 \end{pmatrix}$$



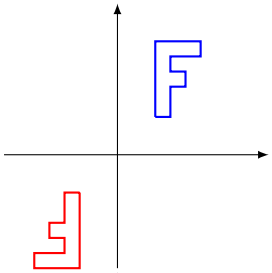
(iii) Drehung um $\frac{\pi}{4}$:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \frac{\sqrt{2}}{2} \begin{pmatrix} x_1 - x_2 \\ x_1 + x_2 \end{pmatrix}$$



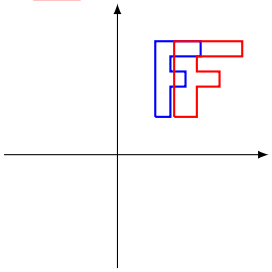
(iv) Spiegelung an 0 / Drehung um π:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} -x_1 \\ -x_2 \end{pmatrix}$$



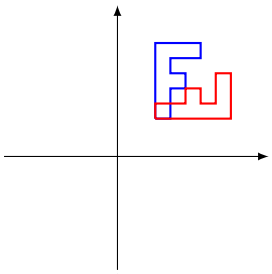
(v) Horizontale Skalierung um $\frac{3}{2}$:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} \frac{3}{2}x_1 \\ x_2 \end{pmatrix}$$



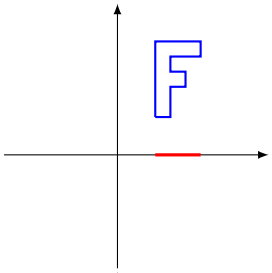
(vi) Koordinatenvertauschung / Spiegelung an $\mathbb{R}(e_1 + e_2)$:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}$$



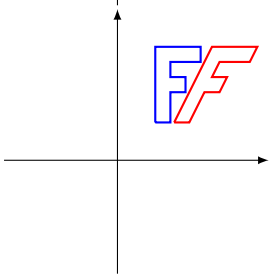
(vii) Orthogonale Projektion auf $\mathbb{R}e_1$:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} x_1 \\ 0 \end{pmatrix}$$



(viii) Horizontale Scherung um den Scherungsfaktor $\frac{1}{2}$:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} x_1 + \frac{1}{2}x_2 \\ x_2 \end{pmatrix}$$



Beispiel 4.1.8 (Auswertung)**.** Sei X eine beliebige Menge, V ein K-Vektorraum und $\mathrm{Abb}(X, V)$ der K-Vektorraum aller Abbildungen von X nach V (siehe Beispiel 3.2.5). Zu jedem $x \in X$ gibt es eine *Auswertungsabbildung*

$$\begin{aligned} \mathrm{ev}_x \colon \mathrm{Abb}(X, V) &\to V, \\ f &\mapsto f(x). \end{aligned}$$

Aus der Definition der Addition und der Skalarmultiplikation auf $\mathrm{Abb}(X, V)$ folgt unmittelbar, dass ev_x eine K-lineare Abbildung ist. Die Abbildung ev_x ist auch die kanonische Projektion π_x aus Definition 1.3.17, wenn wir die Menge $\mathrm{Abb}(X, V)$ als das Produkt $\prod_{x \in X} V$ betrachten.

Beispiel 4.1.9 (analytische Beispiele)**.** Zwei der wichtigsten Operationen in der Analysis, Differentiation und Integration, sind Beispiele von linearen Abbildungen (siehe Beispiel 3.2.14).

(i) Sei $\mathrm{Diff}(\mathbb{R}, \mathbb{R}) \subset \mathrm{Abb}(\mathbb{R}, \mathbb{R})$ der Untervektorraum aller differenzierbaren Funktionen auf $\mathbb{R}$. Dann ist die Abbildung

$$\begin{aligned} \mathrm{Diff}(\mathbb{R}, \mathbb{R}) &\to \mathrm{Abb}(\mathbb{R}, \mathbb{R}), \\ f &\mapsto f', \end{aligned}$$

$\mathbb{R}$-linear.

(ii) Seien $a < b$ reelle Zahlen und sei $\mathrm{Riem}([a, b], \mathbb{R}) \subset \mathrm{Abb}([a, b], \mathbb{R})$ der Untervektorraum aller Riemann-integrierbaren Funktionen auf $[a, b]$. Integration definiert eine $\mathbb{R}$-lineare Abbildung

$$\begin{aligned} \mathrm{Riem}([a, b], \mathbb{R}) &\to \mathbb{R}, \\ f &\mapsto \int_a^b f(x)\, dx. \end{aligned}$$

Andere analytische Beispiele stammen aus Folgen und Reihen:

(iii) Sei $\mathrm{Konv}(\mathbb{R}) \subset \mathbb{R}^{\mathbb{N}}$ der Untervektorraum aller konvergenten Folgen in $\mathbb{R}$ (siehe Beispiel 3.2.13). Die Grenzwertabbildung

$$\begin{aligned} \lim\colon \mathrm{Konv}(\mathbb{R}) &\to \mathbb{R}, \\ (a_n)_{n\in\mathbb{N}} &\mapsto \lim_{n\to\infty} a_n, \end{aligned}$$

ist dann $\mathbb{R}$-linear.

(iv) Sei $\mathrm{Konv}^{\Sigma}(\mathbb{R}) \subset \mathbb{R}^{\mathbb{N}}$ der Untervektorraum aller Folgen $(a_n)_{n\in\mathbb{N}}$, deren Reihe $\sum_{n=0}^{\infty} a_n$ konvergiert. Dann ist die Abbildung

$$\begin{aligned} \mathrm{Konv}^{\Sigma}(\mathbb{R}) &\to \mathbb{R}, \\ (a_n)_{n\in\mathbb{N}} &\mapsto \sum_{n=0}^{\infty} a_n, \end{aligned}$$

$\mathbb{R}$-linear.

Proposition 4.1.10 (Rechnen mit linearen Abbildungen)**.** *Seien V, W Vektorräume über K und sei $f\colon V \to W$ eine K-lineare Abbildung.*

(i) *Es gilt $f(0) = 0$.*

(ii) *Für alle $v \in V$ gilt $f(-v) = -f(v)$.*

(iii) *Sei I eine endliche Menge, $(v_i)_{i\in I}$ eine Familie von Vektoren aus V und $(\lambda_i)_{i\in I}$ eine Familie von Skalaren. Dann gilt*

$$f\left(\sum_{i\in I} \lambda_i \cdot v_i\right) = \sum_{i\in I} \lambda_i \cdot f(v_i).$$

Beweis. Zu (i). Da f linear ist, gilt

$$f(0) + f(0) = f(0 + 0) = f(0).$$

Wenn wir $f(0)$ von beiden Seiten subtrahieren, erhalten wir $f(0) = 0$.

Zu (ii). Nach der Linearität von f und (i) gilt

$$f(v) + f(-v) = f(v + (-v)) = f(0) = 0.$$

Da das inverse Element eindeutig ist, folgt $f(-v) = -f(v)$.

Zu (iii). Dies wird durch Induktion über die Mächtigkeit von I bewiesen. Falls $|I| = 0$ ist die Aussage dieselbe wie (i). Zum Induktionsschritt verwendet man die Gleichung $f(\lambda \cdot v + w) = \lambda \cdot f(v) + f(w)$, die direkt aus der Definition der Linearität folgt. □

Bemerkung 4.1.11. In den Beweisen von Aussagen (i) und (ii) haben wir nur Axiom (i) der Definition 4.1.1 benutzt. Insbesondere gelten diese Aussagen auch für Gruppenhomomorphismen. Es ist aber auch möglich, beide Aussagen nur durch Axiom (ii) nachzuprüfen.

Proposition 4.1.12. *Seien U, V, W Vektorräume über K.*

(i) *Sind $f\colon U \to V$ und $g\colon V \to W$ lineare Abbildungen, so ist $g \circ f\colon U \to W$ linear.*

(ii) *Sind $f, g\colon V \to W$ lineare Abbildungen, so ist ihre Summe*

$$\begin{aligned} f + g\colon V &\to W, \\ v &\mapsto f(v) + g(v), \end{aligned}$$

linear.

(iii) *Ist $f\colon V \to W$ eine lineare Abbildung und ist $\lambda \in K$, so ist*

$$\begin{aligned} \lambda \cdot f\colon V &\to W, \\ v &\mapsto \lambda \cdot f(v), \end{aligned}$$

linear.

Beweis. Jede Aussage ist eine direkte Berechnung. Wir beweisen stellvertretend (iii). Seien $v, w \in V$ und $\mu \in K$. Man berechnet:

$$\begin{aligned} (\lambda \cdot f)(v+w) &= \lambda \cdot f(v+w) \\ &= \lambda(f(v)+f(w)) \\ &= \lambda \cdot f(v) + \lambda \cdot f(w) \\ &= (\lambda \cdot f)(v) + (\lambda \cdot f)(w), \\ (\lambda \cdot f)(\mu \cdot v) &= \lambda \cdot f(\mu \cdot v) \\ &= \lambda \cdot (\mu \cdot f(v)) \\ &= \mu \cdot (\lambda \cdot f(v)) \\ &= \mu \cdot (\lambda \cdot f)(v). \end{aligned}$$

Man beachte dabei, dass zusätzlich zu den Vektorraumaxiomen auch die Kommutativität der Multiplikation auf K in der vorletzten Gleichung benutzt wurde. □

Bemerkung 4.1.13 (lineare Abbildungen und Körpererweiterungen)**.** Sei $K \subset L$ eine Körpererweiterung (Beispiel 3.2.4), seien V, W Vektorräume über L, und sei $f\colon V \to W$ eine L-lineare Abbildung. Dann ist f auch K-linear, wenn wir V und W als K-Vektorräume betrachten (siehe Bemerkung 3.2.15). Die Umkehrung gilt nicht: Beispielsweise ist die komplexe Konjugation $\mathbb{C} \to \mathbb{C}$, $a + bi \mapsto a - bi$, $\mathbb{R}$-linear aber nicht $\mathbb{C}$-linear.

Definition 4.1.14 (Isomorphismus, isomorph)**.** Seien V, W Vektorräume über K.

- Eine K-lineare Abbildung $f\colon V \to W$ heißt *Isomorphismus*, wenn eine K-lineare Abbildung $g\colon W \to V$ existiert, so dass $g \circ f = \mathrm{id}_V$ und $f \circ g = \mathrm{id}_W$. Man schreibt manchmal $f\colon V \xrightarrow{\sim} W$, wenn f ein Isomorphismus ist.
- V ist *isomorph* zu W, in Zeichen $V \cong W$, wenn ein Isomorphismus von V nach W existiert.

Die Idee hinter dieser Definition ist, dass isomorphe K-Vektorräume V, W genau dieselben Vektorraumeigenschaften haben sollen. Genauer, wenn ein Isomorphismus $f\colon V \to W$ gegeben ist, können wir jede Aussage über V, die nur die Vektorraumstruktur von V benutzt, auf eine Aussage über W durch f übertragen. Was damit gemeint ist wird nach und nach deutlich werden, aber hier sind ein paar Beispiele:

- Isomorphe Vektorräume haben dieselbe Dimension.
- Das Bild einer Basis unter einem Isomorphismus ist wieder eine Basis.

Bemerkung 4.1.15. Das Wort „Isomorphismus“ wird in vielen verschiedenen Zusammenhängen verwendet, zum Beispiel auch bei Gruppen, Körpern, partiell geordneten Mengen, usw. (siehe Bemerkung 4.1.2). Es bedeutet immer dasselbe: Ein Isomorphismus ist eine strukturerhaltende Abbildung, die ein strukturerhaltende Umkehrabbildung besitzt.

Bemerkung 4.1.16. Isomorphie ist eine Äquivalenzrelation zwischen K-Vektorräumen:

- V ist isomorph zu sich selbst, da id_V ein Isomorphismus ist.
- Ist $f\colon V \to W$ ein Isomorphismus, so ist seine Umkehrabbildung $f^{-1}\colon W \to V$ auch ein Isomorphismus.

- Die Komposition zweier Isomorphismen ist wieder ein Isomorphismus.

Folgende Proposition ist eine lineare Variante von Satz 1.3.23:

Proposition 4.1.17. *Sei $f\colon V \to W$ eine lineare Abbildung zwischen K-Vektorräumen. Die folgenden Aussagen sind äquivalent:*

(i) *f ist ein Isomorphismus.*

(ii) *f ist bijektiv.*

Beweis. Die Implikation (i) $\Rightarrow$ (ii) folgt aus dem Satz 1.3.23. Sei umgekehrt $f\colon V \to W$ bijektiv. Dann besitzt f eine (eindeutige) Umkehrabbildung $g\colon W \to V$, und es bleibt zu zeigen, dass g auch K-linear ist. Seien $w, w' \in W$ und $\lambda \in K$, und seien $v = g(w)$ und $v' = g(w')$. Da $f \circ g = \mathrm{id}_W$ gilt $f(v) = w$ und $f(v') = w'$. Wir berechnen:

$$g(w + w') = g(f(v) + f(v')) = g(f(v + v')) = v + v' = g(w) + g(w'),$$
$$g(\lambda \cdot w) = g(\lambda \cdot f(v)) = g(f(\lambda \cdot v)) = \lambda \cdot v = \lambda \cdot g(w).$$

Dabei haben wir die Linearität von f und die Gleichung $g \circ f = \mathrm{id}_V$ verwendet. □

Definition 4.1.18 (Endomorphismus, Automorphismus)**.** Sei V ein K-Vektorraum.

- Ein *Endomorphismus* von V ist eine lineare Abbildung von V nach V. Die Menge aller Endomorphismen von V wird mit $\mathrm{End}_K(V)$ bezeichnet.
- Ein Endomorphismus von V, der auch ein Isomorphismus ist, heißt *Automorphismus* von V. Die Menge aller Automorphismen von V wird mit $\mathrm{Aut}_K(V)$ bezeichnet.

Bemerkung 4.1.19. Nach Propositionen 4.1.12(i) und 4.1.17 ist die Menge $\mathrm{Aut}_K(V)$ eine Gruppe bezüglich Komposition.

4.1.1 Lineare Abbildungen und Basen

Proposition 4.1.20. *Seien V, W Vektorräume über K und $f\colon V \to W$ eine lineare Abbildung.*

(i) *Für jede Teilmenge $E \subset V$ gilt*

$$f(\mathrm{Span}_K(E)) = \mathrm{Span}_K(f(E)).$$

Insbesondere, ist f surjektiv und ist E ein Erzeugendensystem von V, so ist $f(E)$ ein Erzeugendensystem von W.

(ii) *Sei $(v_i)_{i\in I}$ eine Familie von Vektoren aus V. Ist die Familie $(f(v_i))_{i\in I}$ in W linear unabhängig, so ist $(v_i)_{i\in I}$ linear unabhängig. Die Umkehrung gilt, wenn f injektiv ist.*

Beweis. Die erste Aussage folgt aus Propositionen 3.2.18 und 4.1.10(iii). Sei $(f(v_i))_{i\in I}$ linear unabhängig, und sei $(\lambda_i)_{i\in I} \in K^{(I)}$ mit $\sum_{i\in I} \lambda_i \cdot v_i = 0$. Nach Proposition 4.1.10(i,iii) gilt dann $\sum_{i\in I} \lambda_i \cdot f(v_i) = 0$. Aus der vorausgesetzten linearen Unabhängigkeit folgt jetzt $\lambda_i = 0$ für alle $i \in I$. Also ist $(v_i)_{i\in I}$ linear unabhängig.

Sei umgekehrt $(v_i)_{i\in I}$ linear unabhängig und f injektiv, und sei $(\lambda_i)_{i\in I} \in K^{(I)}$ mit $\sum_{i\in I} \lambda_i \cdot f(v_i) = 0$. Nach Proposition 4.1.10(i,iii) gilt dann $f\left(\sum_{i\in I} \lambda_i \cdot v_i\right) = f(0)$, und damit $\sum_{i\in I} \lambda_i \cdot v_i = 0$ nach der Injektivität von f. Aus der vorausgesetzten linearen Unabhängigkeit folgt jetzt $\lambda_i = 0$ für alle $i \in I$. Also ist $(f(v_i))_{i\in I}$ linear unabhängig. □

Korollar 4.1.21. *Sei $f\colon V \to W$ ein Isomorphismus von K-Vektorräumen. Ist $(v_i)_{i\in I}$ eine Basis von V, so ist $(f(v_i))_{i\in I}$ eine Basis von W. Insbesondere gilt*

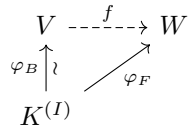
$$\dim_K(V) = \dim_K(W).$$

Beweis. Die Familie $(f(v_i))_{i\in I}$ ist erzeugend bzw. linear unabhängig nach Proposition 4.1.20 (i) bzw. (ii). □

Satz 4.1.22 (universelle Eigenschaft von Basen)**.** *Sei V ein K-Vektorraum mit Basis $(v_i)_{i\in I}$. Zu jedem K-Vektorraum W und jeder Familie $(w_i)_{i\in I}$ in W mit Indexmenge I, gibt es* genau eine *K-lineare Abbildung $f\colon V \to W$ mit $f(v_i) = w_i$ für alle $i \in I$. Außerdem:*

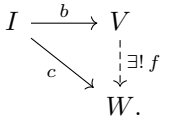
- *f ist genau dann surjektiv, wenn $(w_i)_{i\in I}$ erzeugend ist.*
- *f ist genau dann injektiv, wenn $(w_i)_{i\in I}$ linear unabhängig ist.*
- *f ist genau dann bijektiv, wenn $(w_i)_{i\in I}$ eine Basis ist.*

Beweis. Sei $B = (v_i)_{i\in I}$ and $F = (w_i)_{i\in I}$. Da B erzeugend ist, ist jedes $v \in V$ Linearkombination der Vektoren v_i. Die Eindeutigkeit von f folgt dann aus Proposition 4.1.10(iii). Da B eine Basis von V ist, ist die lineare Abbildung $\varphi_B\colon K^{(I)} \to V$ aus Konstruktion 3.3.8 bijektiv (siehe Bemerkung 3.3.18). Die lineare Abbildung $f = \varphi_F \circ \varphi_B^{-1}\colon V \to W$ hat dann die gewünschte Eigenschaft, da $\varphi_F(\varphi_B^{-1}(v_i)) = \varphi_F(e_i) = w_i$.



Die drei zusätzlichen Aussagen folgen aus den entsprechenden Aussagen für φ_F (siehe Bemerkung 3.3.18). □

Bemerkung 4.1.23. Wenn man die Basis $(v_i)_{i\in I}$ von V als eine Abbildung $b\colon I \to V$ betrachtet, kann man die universelle Eigenschaft so formulieren: Zu jeder Abbildung $c\colon I \to W$ gibt es genau eine lineare Abbildung $f\colon V \to W$ mit $f \circ b = c$:



Korollar 4.1.24 (Klassifikation von Vektorräumen bis auf Isomorphie)**.**

(i) *Jeder K-Vektorraum V ist zu $K^{(I)}$ isomorph, wobei I die Indexmenge einer Basis von V ist.*

(ii) *Zwei K-Vektorräume V und W sind genau dann isomorph, wenn sie dieselbe Dimension haben (d.h., wenn ihre jeweiligen Basen gleichmächtig sind).*

Beweis. Zu (i). Wir wenden den Satz 4.1.22 mit der Basis $(e_i)_{i\in I}$ von $K^{(I)}$ und einer beliebigen I-indizierten Basis von V an, um einen Isomorphismus $f\colon K^{(I)} \xrightarrow{\sim} V$ zu erhalten.

Zu (ii). Eine der beiden Implikationen folgt bereits aus Korollar 4.1.21. Seien umgekehrt $(v_i)_{i\in I}$ und $(w_j)_{j\in J}$ Basen von V und W, so dass I und J gleichmächtig sind. Es existiert also eine bijektive Abbildung $a\colon I \to J$. Da a bijektiv ist, ist $(w_{a(i)})_{i\in I}$ wieder eine Basis von W. Nach Satz 4.1.22 ist die eindeutige lineare Abbildung $f\colon V \to W$ mit $f(v_i) = w_{a(i)}$ ein Isomorphismus. □

Diese Klassifikation kann man auf folgende Weise zusammenfassen: Es gibt eine Bijektion

$$\{\text{Vektorräume über } K\}/\text{Isomorphie} \leftrightarrow \{\text{Mengen}\}/\text{Gleichmächtigkeit},$$
$$K^{(I)} \leftarrow\!\shortmid I.$$

Das ist aber etwas schlampig, da K-Vektorräume bzw. Mengen keine Menge bilden.

Beispiel 4.1.25. Ist V endlich-dimensional der Dimension n, so gilt $V \cong K^n$.

Bemerkung 4.1.26. Das Korollar 4.1.24 impliziert insbesondere, dass der K-Vektorraum $K^{\mathbb{N}}$ aller Folgen in K zu einem K-Vektorraum der Gestalt $K^{(I)}$ isomorph ist. Aber die Mächtigkeit von I hängt von der des Körpers K ab. Wenn $|K| \leq |\mathbb{R}|$, zum Beispiel wenn K ein Teilkörper von $\mathbb{C}$ oder ein endlicher Körper ist, dann gilt $K^{\mathbb{N}} \cong K^{(\mathbb{R})}$.

4.1.2 Kern und Bild linearer Abbildungen

Definition 4.1.27 (Kern, Bild)**.** Sei $f\colon V \to W$ eine lineare Abbildung zwischen K-Vektorräumen.

- Der *Kern* (oder der *Nullraum*) von f ist
$$\ker f := f^{-1}(\{0\}) = \{v \in V \mid f(v) = 0\}.$$
- Das *Bild* von f ist
$$\operatorname{im} f := f(V) = \{w \in W \mid \text{es existiert } v \in V \text{ mit } f(v) = w\}.$$

Proposition 4.1.28 (Kern und Bild sind Untervektorräume)**.** *Sei $f\colon V \to W$ eine lineare Abbildung zwischen K-Vektorräumen.*

(i) *Ist $U \subset W$ ein Untervektorraum, so ist $f^{-1}(U) \subset V$ ein Untervektorraum. Insbesondere ist der Kern von f ist ein Untervektorraum von V.*

(ii) *Ist $U \subset V$ ein Untervektorraum, so ist $f(U) \subset W$ ein Untervektorraum. Insbesondere ist das Bild von f ist ein Untervektorraum von W.*

Beweis. Wir verwenden das Kriterium 3.2.8.

Zu (i). $f^{-1}(U)$ enthält den Nullvektor (nach Proposition 4.1.10(i)). Sind $v, v' \in f^{-1}(U)$ und $\lambda \in K$, so gilt
$$f(v + v') = f(v) + f(v') \in U \quad \text{und} \quad f(\lambda \cdot v) = \lambda \cdot f(v) \in U,$$
und damit liegen $v + v'$ und $\lambda \cdot v$ auch in $f^{-1}(U)$.

Zu (ii). $f(U)$ enthält den Nullvektor (nach Proposition 4.1.10(i)). Seien $w, w' \in f(U)$ und $\lambda \in K$. Nach Definition von $f(U)$ existieren $v, v' \in U$ mit $f(v) = w$ und $f(v') = w'$. Dann gilt
$$f(v + v') = f(v) + f(v') = w + w' \quad \text{und} \quad f(\lambda \cdot v) = \lambda \cdot f(v) = \lambda \cdot w,$$
und damit liegen $w + w'$ und $\lambda \cdot w$ auch in $f(U)$. □

Beispiel 4.1.29. Wir berechnen den Kern und das Bild der Abbildungen aus Beispiel 4.1.4.

(i) Für die Identität $\mathrm{id}_V\colon V \to V$ gilt $\ker \mathrm{id}_V = \{0\}$ und $\operatorname{im} \mathrm{id}_V = V$.

(ii) Für die Nullabbildung $0\colon V \to W$ gilt $\ker 0 = V$ und $\operatorname{im} 0 = \{0\}$.

(iii) Für die Inklusionsabbildung $i\colon U \hookrightarrow V$ gilt $\ker i = \{0\}$ und $\operatorname{im} i = U$.

(iv) Für die Quotientenabbildung $q\colon V \twoheadrightarrow V/U$ gilt $\ker q = U$ und $\operatorname{im} q = V/U$.

(v) Für die kanonische Abbildung $\iota_1\colon V \hookrightarrow V \oplus W$ gilt $\ker \iota_1 = \{0\}$ und $\operatorname{im} \iota_1 = V \times \{0\}$.

(vi) Für die kanonische Abbildung $\pi_1\colon V \oplus W \twoheadrightarrow V$ gilt $\ker \pi_1 = \{0\} \times W$ und $\operatorname{im} \pi_1 = V$.

Beispiel 4.1.30. Sei $\lim\colon \mathrm{Konv}(\mathbb{R}) \to \mathbb{R}$ die Grenzwertabbildung aus Beispiel 4.1.9(iii). Der Kern von lim besteht aus allen Folgen, die gegen 0 konvergieren. Der $\mathbb{R}$-Vektorraum $\mathrm{Konv}^{\Sigma}(\mathbb{R})$ aus Beispiel 4.1.9(iv) ist ein Untervektorraum von $\ker(\lim)$.

Beispiel 4.1.31 (lineare Differentialgleichungen)**.** Sei $V = C^\infty(\mathbb{R}, \mathbb{R})$ der $\mathbb{R}$-Vektorraum aller glatten (d.h., beliebig oft differenzierbaren) Funktionen von $\mathbb{R}$ nach $\mathbb{R}$, und sei $D\colon V \to V$ die lineare Abbildung $f \mapsto f'$. Dann ist die Abbildung D surjektiv, und ihr Kern ist der Untervektorraum Konst$(\mathbb{R}, \mathbb{R})$ der konstanten Funktionen. Eine Abbildung $L\colon V \to V$ der Gestalt

$$L = D^n + f_{n-1} \cdot D^{n-1} + \cdots f_1 \cdot D + f_0 \cdot \mathrm{id}_V$$

mit $n \in \mathbb{N}$ und $f_i \in C^\infty(\mathbb{R}, \mathbb{R})$ heißt *linearer Differentialoperator n-ter Ordnung* (dabei werden die Potenzen von D bezüglich der Komposition genommen, d.h., $D^i(f)$ ist die i-te Ableitung von f). In der Analysis wird gezeigt, dass $\ker L$ immer ein n-dimensionaler Untervektorraum von V ist. Beispielsweise wird $\ker(D - \mathrm{id}_V)$ von der Exponentialfunktion exp erzeugt, und wird $\ker(D^2 + \mathrm{id}_V)$ von den Winkelfunktionen cos und sin erzeugt.

Proposition 4.1.32. *Sei $f\colon V \to W$ eine lineare Abbildung zwischen K-Vektorräumen.*

(i) *f ist genau dann injektiv, wenn* $\ker f = \{0\}$.

(ii) *f ist genau dann surjektiv, wenn* $\operatorname{im} f = W$.

Beweis. Die zweite Aussage ist genau die Definition der Surjektivität. Sei $f\colon V \to W$ injektiv. Dann hat $0 \in W$ höchstens ein Urbild unter f. Da $0 \in V$ ein solches Urbild ist, ist $\ker f$ genau gleich $\{0\}$. Sei umgekehrt $\ker f = \{0\}$, und seien $v, v' \in V$ mit $f(v) = f(v')$. Dann

$$f(v - v') = f(v) - f(v') = 0,$$

und damit $v - v' \in \ker f = \{0\}$, d.h., $v = v'$. Also ist f injektiv. □

Die folgende Proposition ist eine lineare Variante der universallen Eigenschaft der Quotientenmenge (Satz 1.4.9).

Proposition 4.1.33 (universelle Eigenschaft des Quotientenvektorraums)**.** *Sei V ein K-Vektorraum, $U \subset V$ ein Untervektorraum, V/U der Quotientenvektorraum (siehe Definition 3.2.24) und $q\colon V \to V/U$ die Quotientenabbildung. Zu jedem K-Vektorraum W und jeder linearen Abbildung $f\colon V \to W$ mit $U \subset \ker f$ gibt es* genau eine *lineare Abbildung $\bar{f}\colon V/U \to W$ mit $\bar{f} \circ q = f$.*

Beweis. Zur Erinnerung ist V/U die Quotientenmenge $V/\!\sim_U$, wobei $x \sim_U y \Leftrightarrow x - y \in U$. Falls $x \sim_U y$, folgt aus der Voraussetzung $f(U) = \{0\}$, dass $f(x - y) = 0$, d.h., $f(x) = f(y)$. Nach der universellen Eigenschaft der Quotientenmenge gibt es genau eine Abbildung $\bar{f}\colon V/U \to W$ mit $\bar{f} \circ q = f$. Es bleibt zu zeigen, dass $\bar{f}$ linear ist. Dies folgt aus der Linearität von f und q:

$$\begin{aligned}\bar{f}((v + U) + (v' + U)) &= \bar{f}((v + v') + U) = f(v + v') \\ &= f(v) + f(v') = \bar{f}(v + U) + \bar{f}(v' + U), \\ \bar{f}(\lambda \cdot (v + U)) &= \bar{f}(\lambda v + U) = f(\lambda v) = \lambda \cdot f(v) = \lambda \cdot \bar{f}(v + U).\end{aligned}$$

□

Bemerkung 4.1.34. Es gibt eine ähnliche universelle Eigenschaft für Untervektorräume, aber sie ist offensichtlicher: Ist $i\colon U \hookrightarrow V$ die Inklusionsabbildung eines Untervektorraums und ist $f\colon W \to V$ eine lineare Abbildung mit $\operatorname{im} f \subset U$, so gibt es genau eine lineare Abbildung $\bar{f}\colon W \to U$ mit $i \circ \bar{f} = f$.

Satz 4.1.35 (Homomorphiesatz für Vektorräume)**.** *Sei $f\colon V \to W$ eine lineare Abbildung zwischen K-Vektorräumen. Dann ist die von f induzierte Abbildung*

$$\begin{aligned}\bar{f}\colon V/\ker f &\to \operatorname{im} f, \\ v + \ker f &\mapsto f(v),\end{aligned}$$

ein Isomorphismus von K-Vektorräumen.

Beweis. Die Abbildung $\bar{f}$ ist wohldefiniert und linear nach Proposition 4.1.33, und sie ist surjektiv nach Definition von $\operatorname{im} f$. Zur Injektivität berechnen wir den Kern von $\bar{f}$. Es gilt:

$$\begin{aligned}\ker \bar{f} &= \{v + \ker f \mid f(v) = 0\} \\ &= \{v + \ker f \mid v \in \ker f\} \\ &= \{0 + \ker f\}.\end{aligned}$$

Also ist $\bar{f}$ injektiv nach Proposition 4.1.32(i). □

Bemerkung 4.1.36. Der Homomorphiesatz impliziert folgenden Zerlegungssatz für lineare Abbildungen: Jede lineare Abbildung $f\colon V \to W$ läßt sich kanonisch als Komposition einer Quotientenabbildung, eines Isomorphismus und einer Inklusionsabbildung zerlegen:

$$V \overset{q}{\twoheadrightarrow} V/\ker f \overset{\bar{f}}{\to} \operatorname{im} f \overset{i}{\hookrightarrow} W.$$

Beispiel 4.1.37. Sei $C^0(\mathbb{R},\mathbb{R})$ der $\mathbb{R}$-Vektorraum aller stetigen reellen Funktionen auf $\mathbb{R}$, und sei $C^1(\mathbb{R},\mathbb{R}) \subset C^0(\mathbb{R},\mathbb{R})$ der Untervektorraum aller stetig differenzierbaren Funktionen (d.h., Funktionen f, deren Ableitung f' existiert und stetig ist). Dann haben wir die $\mathbb{R}$-lineare Differentiationsabbildung

$$\begin{aligned}D\colon C^1(\mathbb{R},\mathbb{R}) &\to C^0(\mathbb{R},\mathbb{R}), \\ f &\mapsto f'\end{aligned}$$

(siehe Beispiel 4.1.9). In der Analysis wird gezeigt, dass jede stetige Abbildung $g \in C^0(\mathbb{R},\mathbb{R})$ eine Stammfunktion besitzt, d.h., eine Funktion f mit $f' = g$; zum Beispiel,

$$f(x) = \int_0^x g(t)\,dt.$$

Anders gesagt ist die obige Abbildung D surjektiv. In der Analysis wird auch gezeigt, dass f' genau dann gleich null ist, wenn f eine konstante Funktion ist. Der Kern von D ist also der Untervektorraum $\operatorname{Konst}(\mathbb{R},\mathbb{R}) \subset C^1(\mathbb{R},\mathbb{R})$ aller konstanten Funktionen, der offensichtlich zu $\mathbb{R}$ isomorph ist. Der Homomorphiesatz impliziert, dass D einen Isomorphismus

$$\bar{D}\colon C^1(\mathbb{R},\mathbb{R})/\operatorname{Konst}(\mathbb{R},\mathbb{R}) \xrightarrow{\sim} C^0(\mathbb{R},\mathbb{R})$$

induziert.

Korollar 4.1.38 (Dimensionsformel für lineare Abbildungen)**.** *Seien V, W Vektorräume über K und sei $f\colon V \to W$ eine K-lineare Abbildung. Ist V endlich-dimensional, so gilt*

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

Beweis. Nach dem Homomorphiesatz 4.1.35 gilt $\dim_K(\operatorname{im} f) = \dim_K(V/\ker f)$. Die gewünschte Gleichung folgt jetzt aus der Dimensionsformel für Quotientenvektorräume (Proposition 3.3.46). □

Korollar 4.1.39 (Charakterisierung von Isomorphismen)**.** *Seien V und W K-Vektorräume derselben* endlichen *Dimension und sei $f\colon V \to W$ eine lineare Abbildung. Die folgenden Aussagen sind dann äquivalent:*

(i) *f ist bijektiv, d.h., ein Isomorphismus.*

(ii) *f ist injektiv, d.h.,* $\ker f = \{0\}$.

(iii) *f ist surjektiv, d.h.,* $\operatorname{im} f = W$.

Beweis. Sei $\ker f = \{0\}$. Nach Korollar 4.1.38 gilt dann $\dim_K(\operatorname{im} f) = \dim_K(V) = \dim_K(W)$. Aus Proposition 3.3.35 folgt, dass $\operatorname{im} f = W$. Sei umgekehrt $\operatorname{im} f = W$. Aus Korollar 4.1.38 folgt dann, dass $\dim(\ker f) = 0$, d.h., dass $\ker f = \{0\}$. □

Bemerkung 4.1.40. Das Korollar 4.1.39 gilt nicht bei unendlich-dimensionalen Vektorräumen, selbst wenn wir die Dimension im Sinne der Bemerkung 3.3.31 verstehen. Die lineare Abbildung

$$K^{\mathbb{N}} \to K^{\mathbb{N}}, \quad (x_0, x_1, x_2, \dots) \mapsto (0, x_0, x_1, \dots),$$

ist injektiv aber nicht bijektiv, und die lineare Abbildung

$$K^{\mathbb{N}} \to K^{\mathbb{N}}, \quad (x_0, x_1, x_2, \dots) \mapsto (x_1, x_2, x_3, \dots),$$

ist surjektiv aber nicht bijektiv. Ein anderes Gegenbeispiel ist die Differentiationsabbildung $D\colon C^\infty(\mathbb{R}, \mathbb{R}) \to C^\infty(\mathbb{R}, \mathbb{R})$ aus Beispiel 4.1.31, die surjektiv aber nicht injektiv ist (insbesondere ist $C^\infty(\mathbb{R}, \mathbb{R})$ unendlich-dimensional).

Definition 4.1.41 (Rang)**.** Sei $f\colon V \to W$ eine lineare Abbildung zwischen K-Vektorräumen. Man nennt die Dimension des Bildes von f den *Rang* von f:

$$\operatorname{rg} f := \dim_K(\operatorname{im} f).$$

Man beachte dabei, dass $\operatorname{rg} f \leq \dim_K(V)$ (nach Satz 4.1.35) und $\operatorname{rg} f \leq \dim_K(W)$ (da $\operatorname{im} f \subset W$). Der Rang ist ein wichtiger Begriff, denn er spielt eine ähnliche Rolle bei linearen Abbildungen wie die Dimension bei Vektorräumen: Bis auf Isomorphie werden Vektorräume durch ihre Dimension bestimmt (Korollar 4.1.24), und man kann den folgenden Satz so verstehen, dass lineare Abbildungen bis auf Isomorphie durch ihren Rang bestimmt werden:

Satz 4.1.42 (Klassifikation von linearen Abbildungen bis auf Isomorphie)**.** *Seien V und W Vektorräume über K und $f\colon V \to W$ eine lineare Abbildung. Dann existieren Mengen I, J und $L \subset I \cap J$ mit $|L| = \operatorname{rg} f$, und Isomorphismen $\varphi\colon K^{(I)} \xrightarrow{\sim} V$ und $\psi\colon K^{(J)} \xrightarrow{\sim} W$, so dass*

$$\psi^{-1} \circ f \circ \varphi = \iota_L \circ \pi_L\colon K^{(I)} \to K^{(J)},$$

wobei $\pi_L\colon K^{(I)} \twoheadrightarrow K^{(L)}$ und $\iota_L\colon K^{(L)} \hookrightarrow K^{(J)}$ die kanonischen Abbildungen sind, d.h., die Einschränkung auf L und die Nullfortsetzung auf J.

Beweis. Sei $(v_i)_{i\in M}$ eine Basis von $\ker f$. Nach dem Basisergänzungssatz können wir sie zu einer Basis $B = (v_i)_{i\in I}$ von V ergänzen; sei $L = I \setminus M$ und $U = \operatorname{Span}_K\{v_i \mid i \in L\}$. Da U und $\ker f$ komplementär sind, ist die Einschränkung $f|_U$ injektiv. Nach Proposition 4.1.20(ii) ist die Familie $(f(v_i))_{i\in L}$ in W linear unabhängig, und damit eine Basis von $\operatorname{im} f$ (insbesondere gilt $|L| = \operatorname{rg} f$). Nach dem Basisergänzungssatz gibt es eine Basis $C = (w_j)_{j\in J}$ von W mit $L \subset J$ und $w_i = f(v_i)$ für alle $i \in L$. Dann gilt $\varphi_C^{-1} \circ f \circ \varphi_B = \iota_L \circ \pi_L$ nach Konstruktion. □

4.1.3 Homomorphismenräume

Seien V und W Vektorräume über K. Zur Erinnerung bezeichnen wir mit $\operatorname{Hom}_K(V, W)$ die Menge aller K-linearen Abbildungen von V nach W. Sie ist auf kanonische Weise ein K-Vektorraum:

Proposition 4.1.43. *Seien V, W Vektorräume über K. Wenn wir die Menge* $\operatorname{Abb}(V, W)$ *als K-Vektorraum bezüglich der punktweisen Addition bzw. Skalarmultiplikation betrachten (Beispiel 3.2.5), dann ist* $\operatorname{Hom}_K(V, W)$ *ein Untervektorraum von* $\operatorname{Abb}(V, W)$*.*

Beweis. Dies folgt aus Proposition 4.1.12(ii),(iii) und dem Kriterium 3.2.8 ($\operatorname{Hom}_K(V, W)$ ist nicht leer, da die Nullabbildung K-linear ist). □

Bemerkung 4.1.44. Insbesondere ist $\mathrm{End}_K(V)$ ein K-Vektorraum. Falls $V \neq \{0\}$ ist die Teilmenge $\mathrm{Aut}_K(V)$ von $\mathrm{End}_K(V)$ *kein* Untervektorraum, da die Nullabbildung kein Automorphismus ist.

Beispiel 4.1.45 (lineare Abbildungen von K nach K)**.** Sei $f\colon K \to K$ eine K-linear Abbildung. Für alle $x \in K$ gilt dann

$$f(x) = f(x \cdot 1) = x \cdot f(1) = f(1) \cdot x.$$

Das heißt, f ist gleich der Multiplikation mit $f(1) \in K$. Ist umgekehrt $\lambda \in K$, so ist $x \mapsto \lambda \cdot x$ eine K-lineare Abbildung $K \to K$ nach Beispiel 4.1.3(i), die 1 auf λ abbildet. Es gibt also zueinander inverse Bijektionen

$$\begin{aligned} K &\rightleftarrows \mathrm{Hom}_K(K, K), \\ \lambda &\mapsto (x \mapsto \lambda \cdot x), \\ f(1) &\mapsfrom f. \end{aligned}$$

Außerdem kann man leicht nachprüfen, dass beide Abbildungen linear sind. Insbesondere gilt $\mathrm{Hom}_K(K, K) \cong K$.

Proposition 4.1.46 (Funktorialität der Homomorphismenräume)**.** *Seien V, V', W, W' Vektorräume über K und $f\colon V \to V'$ und $g\colon W \to W'$ lineare Abbildungen. Dann ist die Abbildung*

$$\begin{aligned} \mathrm{Hom}_K(f, g)\colon \mathrm{Hom}_K(V', W) &\to \mathrm{Hom}_K(V, W'), \\ h &\mapsto g \circ h \circ f, \end{aligned}$$

K-linear. Insbesondere sind die Abbildungen

$$\begin{aligned} \mathrm{Hom}_K(\mathrm{id}_V, g)\colon \mathrm{Hom}_K(V, W) &\to \mathrm{Hom}_K(V, W'), \\ h &\mapsto g \circ h, \\ \mathrm{Hom}_K(f, \mathrm{id}_W)\colon \mathrm{Hom}_K(V', W) &\to \mathrm{Hom}_K(V, W), \\ h &\mapsto h \circ f \end{aligned}$$

K-linear.

Beweis. Dies folgt durch Nachrechnen aus der Linearität von g. □

Definition 4.1.47 (Linearform, Dualraum, duale Abbildung)**.**

- Sei V ein Vektorraum über K. Eine *Linearform* auf V ist eine lineare Abbildung $V \to K$. Der *Dualraum* V^* von V ist der K-Vektorraum aller Linearformen auf V:

$$V^* = \mathrm{Hom}_K(V, K).$$

- Sei $f\colon V \to W$ eine K-lineare Abbildung. Die *duale Abbildung* f^* zu f ist die K-lineare Abbildung

$$\begin{aligned} f^* = \mathrm{Hom}_K(f, \mathrm{id}_K)\colon W^* &\to V^*, \\ \alpha &\mapsto \alpha \circ f. \end{aligned}$$

Beispiel 4.1.48. Sei $n \in \mathbb{N}$. Die n kanonischen Projektionen $\pi_i\colon K^n \to K$ sind Linearformen auf K^n.

Beispiel 4.1.49. Seien $a < b$ reelle Zahlen und sei $\mathrm{Riem}([a, b], \mathbb{R})$ der $\mathbb{R}$-Vektorraum aller Riemman-integrierbaren Funktionen auf $[a, b]$. Das Riemannsche Integral $\int_a^b$ ist eine Linearform auf $\mathrm{Riem}([a, b], \mathbb{R})$.

Bemerkung 4.1.50. Sind $f\colon V \to W$ und $g\colon W \to U$ lineare Abbildungen, so gilt

$$(g \circ f)^* = f^* \circ g^*\colon U^* \to V^*.$$

Man kann auch leicht nachprüfen, dass die Abbildung

$$\begin{aligned} \mathrm{Hom}_K(V,W) &\to \mathrm{Hom}_K(W^*,V^*),\\ f &\mapsto f^*, \end{aligned}$$

linear ist.

Lemma 4.1.51. *Sei $f\colon V \to W$ eine lineare Abbildung zwischen K-Vektorräumen.*

(i) *Ist f injektiv, so ist f^* surjektiv.*

(ii) *Ist f surjektiv, so ist f^* injektiv.*

Beweis. Sei $(v_i)_{i\in I}$ eine Basis von V. Ist f injektiv, so ist die Familie $(f(v_i))_{i\in I}$ linear unabhängig (Proposition 4.1.20(ii)) und kann zu einer Basis von W ergänzt werden (Satz 3.3.20(iii)). Nach Satz 4.1.22 kann dann jede lineare Abbildung $V \to K$ zu einer linearen Abbildung $W \to K$ fortgesetzt werden, d.h., f^* ist surjektiv. Ist f surjektiv und ist $\alpha \circ f = 0$, so muss α null sein, d.h., f^* ist injektiv. □

Proposition 4.1.52 (Kern und Bild dualer Abbildungen)**.** *Sei $f\colon V \to W$ eine lineare Abbildung zwischen K-Vektorräumen.*

(i) *Die duale Abbildung der Quotientenabbildung $q\colon W \twoheadrightarrow W/\operatorname{im} f$ induziert einen Isomorphismus*

$$(W/\operatorname{im} f)^* \xrightarrow{\sim} \ker(f^*).$$

(ii) *Die duale Abbildung der Inklusionsabbildung $i\colon \ker f \hookrightarrow V$ induziert einen Isomorphismus*

$$V^*/\operatorname{im}(f^*) \xrightarrow{\sim} (\ker f)^*.$$

(iii) *Die duale Abbildung der von f induzierten Abbildung $\bar{f}\colon V \twoheadrightarrow \operatorname{im} f$ induziert einen Isomorphismus*

$$(\operatorname{im} f)^* \xrightarrow{\sim} \operatorname{im}(f^*).$$

Beweis. Zu (i). Nach Lemma 4.1.51(ii) ist die Abbildung $q^*\colon (W/\operatorname{im} f)^* \to W^*$ injektiv. Es bleibt zu zeigen, dass $\operatorname{im}(q^*) = \ker(f^*)$. Aus $q \circ f = 0$ folgt $f^* \circ q^* = (q \circ f)^* = 0$ und somit $\operatorname{im}(q^*) \subset \ker f^*$. Sei umgekehrt $\alpha \in \ker(f^*) \subset W^*$, d.h., $\alpha \circ f = 0$. Nach der universellen Eigenschaft der Quotientenvektorraum existiert $\bar{\alpha}\colon W/\operatorname{im} f \to K$ so dass $\bar{\alpha} \circ q = \alpha$, d.h., $q^*(\bar{\alpha}) = \alpha$. Insbesondere ist α im Bild von q^*.

Zu (ii). Nach Lemma 4.1.51(i) ist die Abbildung $i^*\colon V^* \to (\ker f)^*$ surjektiv. Nach Satz 4.1.35 bleibt es zu zeigen, dass $\ker(i^*) = \operatorname{im}(f^*)$. Aus $f \circ i = 0$ folgt $i^* \circ f^* = 0$ und somit $\operatorname{im}(f^*) \subset \ker(i^*)$. Sei umgekehrt $\alpha \in \ker(i^*) \subset V^*$, d.h., $\alpha \circ i = 0$. Nach der universellen Eigenschaft der Quotientenvektorraum ist α im Bild von r^*, wobei $r\colon V \to V/\ker f$ die Quotientenabbildung ist. Nach Satz 4.1.35 ist $f = j \circ r$ mit einer injektiven linearen Abbildung $j\colon V/\ker f \hookrightarrow W$. Nach Lemma 4.1.51(i) ist j^* surjektiv, und damit ist α im Bild von f^*.

Zu (iii). Nach Lemma 4.1.51(ii) ist die Abbildung $\bar{f}^*\colon (\operatorname{im} f)^* \to V^*$ injektiv. Es bleibt zu zeigen, dass $\operatorname{im}(\bar{f}^*) = \operatorname{im}(f^*)$. Es gilt $f = u \circ \bar{f}$, wobei $u\colon \operatorname{im} f \hookrightarrow W$ die Inklusionsabbildung ist. Daraus folgt $f^* = \bar{f}^* \circ u^*$. Nach Lemma 4.1.51(i) ist u^* surjektiv, und damit ist $\operatorname{im}(f^*) = \operatorname{im}(\bar{f}^*)$, wie gewünscht. □

Bemerkung 4.1.53. Sei $f\colon V \to W$ eine lineare Abbildung. Der Quotientenvektorraum $W/\operatorname{im} f$ heißt der *Kokern* von f und wird mit $\operatorname{coker} f$ bezeichnet. Aussagen (i) und (ii) der Proposition 4.1.52 sagen insbesondere, dass $\ker(f^*) \cong (\operatorname{coker} f)^*$ und $\operatorname{coker}(f^*) \cong (\ker f)^*$. Das heißt, die Dualität vertauscht den Kern und den Kokern, aber sie erhält das Bild.

Konstruktion 4.1.54 (duale Basis)**.** Sei V ein K-Vektorraum und $B = (v_i)_{i\in I}$ eine Basis von V. Nach der universellen Eigenschaft von Basen (Satz 4.1.22) gibt es zu jedem $i \in I$ genau eine Linearform $v_i^* \in V^*$, so dass

$$v_i^*(v_j) = \begin{cases} 1, & \text{falls } j = i, \\ 0, & \text{andernfalls.} \end{cases}$$

(Trotz der Notation hängt die Linearform v_i^* nicht nur von v_i ab, sondern von der ganzen Basis B!) Nach demselben Satz gibt es dann genau eine lineare Abbildung

$$\begin{aligned} \varepsilon_B \colon V &\to V^*, \\ v_i &\mapsto v_i^*. \end{aligned}$$

Die nächste Proposition zeigt, dass die Familie $B^* = (v_i^*)_{i\in I}$ in V^* immer linear unabhängig ist. Sie ist sogar eine Basis von V^*, wenn V endlich-dimensional ist. In diesem Fall heißt die Basis B^* von V^* die *duale Basis* zu B.

Beispiel 4.1.55. Sei $B = (e_1, \dots, e_n)$ die Standardbasis von K^n. Die duale Basis von $(K^n)^*$ ist $B^* = (\pi_1, \dots, \pi_n)$.

Proposition 4.1.56. *Sei V ein K-Vektorraum und $B = (v_i)_{i\in I}$ eine Basis von V.*

(i) *Die lineare Abbildung $\varepsilon_B \colon V \to V^*$ ist injektiv.*

(ii) *Ist V endlich-dimensional, so ist $\varepsilon_B \colon V \to V^*$ ein Isomorphismus.*

Beweis. Zu (i). Nach Satz 4.1.22 genügt es zu zeigen, dass die Familie $(v_i^*)_{i\in I}$ linear unabhängig ist. Sei $\sum_{i\in I} \lambda_i \cdot v_i^* = 0$ mit $(\lambda_i)_{i\in I} \in K^{(I)}$. Nach Definition von v_i^* ist der Wert von $\sum_{i\in I} \lambda_i \cdot v_i^*$ in v_j gleich λ_j, und damit sind alle λ_i gleich 0.

Zu (ii). In diesem Fall ist I endlich (Satz 3.3.27(i)). Sei $\alpha \in V^*$ beliebig. Die lineare Abbildungen α und $\sum_{i\in I} \alpha(v_i) \cdot v_i^*$ haben denselben Wert in v_j für alle $j \in I$. Nach Satz 4.1.22 stimmen sie deshalb auf ganz V überein. Dies zeigt, dass α im Bild von ε_B liegt, und damit dass ε_B surjektiv ist. □

Beispiel 4.1.57. Sei I eine Menge. Aus dem Satz 4.1.22 folgt, dass die Abbildung

$$\begin{aligned} (K^{(I)})^* &\to K^I, \\ \alpha &\mapsto (\alpha(e_i))_{i\in I}, \end{aligned}$$

ein Isomorphismus ist. Für die Basis $B = (e_i)_{i\in I}$ von $K^{(I)}$ ist dann die Komposition von $\varepsilon_B \colon K^{(I)} \hookrightarrow (K^{(I)})^*$ mit diesem Isomorphismus die Inklusionsabbildung $K^{(I)} \hookrightarrow K^I$.

Korollar 4.1.58 (Rang der dualen Abbildung)**.** *Sei $f \colon V \to W$ eine lineare Abbildung zwischen K-Vektorräumen. Ist* $\operatorname{im} f$ *endlich-dimensonal (z.B. ist V oder W endlich-dimensional), so gilt*

$$\operatorname{rg} f = \operatorname{rg} f^*.$$

Beweis. Nach Proposition 4.1.52 gilt

$$\dim_K \operatorname{im}(f^*) = \dim_K (\operatorname{im} f)^*.$$

Wenn V oder W endlich-dimensional ist, ist auch $\operatorname{im} f$ endlich-dimensional nach Korollar 4.1.38 oder Proposition 3.3.35. Nach Proposition 4.1.56(ii) gilt dann

$$\dim_K (\operatorname{im} f)^* = \dim_K \operatorname{im} f.$$

Aus diesen zwei Formeln folgt, dass $\operatorname{rg} f = \operatorname{rg} f^*$. □

Zur Erinnering gibt es zu jeder Menge X und jedem Element $x \in X$ eine Auswertungsabbildung

$$\begin{aligned} \mathrm{ev}_x \colon \mathrm{Abb}(X, K) &\to K, \\ f &\mapsto f(x), \end{aligned}$$

die K-linear ist (Beispiel 4.1.8). Wenn $X = V$ ein K-Vektorraum ist, ist insbesondere die Einschränkung von ev_x auf V^* eine Linearform auf V^*, also ein Element des Dualraums $(V^*)^*$, das wir auch mit ev_v bezeichnen.

Proposition 4.1.59 (Doppeldual)**.** *Sei V ein K-Vektorraum. Die Abbildung*

$$\begin{aligned} \mathrm{ev} \colon V &\to (V^*)^*, \\ v &\mapsto \mathrm{ev}_v \end{aligned}$$

ist K-linear und injektiv. Falls V endlich-dimensional ist, ist ev *ein Isomorphismus.*

Beweis. Zur Linearität haben wir zu zeigen:

$$\mathrm{ev}_{v+w} = \mathrm{ev}_v + \mathrm{ev}_w \quad \text{und} \quad \mathrm{ev}_{\lambda \cdot v} = \lambda \cdot \mathrm{ev}_v \,.$$

Dies folgt unmittelbar aus den Definitionen. Zum Beispiel gilt für alle $\alpha \in V^*$:

$$\mathrm{ev}_{\lambda \cdot v}(\alpha) = \alpha(\lambda \cdot v) = \lambda \cdot \alpha(v) = \lambda \cdot \mathrm{ev}_v(\alpha) = (\lambda \cdot \mathrm{ev}_v)(\alpha).$$

Zur Injektivität: Sei $v \in V \setminus \{0\}$. Nach dem Basisergänzungssatz 3.3.20 können wir v zu einer Basis B von V ergänzen. Für die Abbildung $\varepsilon_B \colon V \to V^*$ gilt dann nach Konstruktion $\varepsilon_B(v)(v) = 1 \neq 0$. Insbondere existiert $\alpha \in V^*$, so dass $\mathrm{ev}_v(\alpha) = \alpha(v) \neq 0$, und damit ist $\mathrm{ev}_v \neq 0$. Also ist ev injektiv.

Wenn V endlich-dimensional ist, dann gilt $\dim_K(V) = \dim_K(V^*) = \dim_K((V^*)^*)$ nach Proposition 4.1.56(ii). Die letzte Aussage folgt daraus, da jede injektive lineare Abbildung zwischen endlich-dimensionalen Vektorräumen derselben Dimension ein Isomorphismus ist (Korollar 4.1.39). □

Bemerkung 4.1.60. Sei V ein endlich-dimensional K-Vektorraum. Man beachte, dass der Isomorphismus $\varepsilon_B \colon V \to V^*$ von der Wahl der Basis B abhängt. Da im Allgemeinen V keine bevorzugte Basis besitzt, gibt es keinen bevorzugten Isomorphismus zwischen V und seinem Dualraum. Im Gegensatz dazu ist der Isomorphismus $\mathrm{ev} \colon V \to (V^*)^*$ unabhängig von irgendwelchen Wahlen. Deswegen wird es oft gesagt, dass ein endlich-dimensionaler Vektorraum zu seinem Doppeldual *kanonisch* isomorph ist, aber zu seinem Dual nur unkanonisch.

Bemerkung 4.1.61. Falls V unendlich-dimensional ist, dann sind weder $\varepsilon_B \colon V \to V^*$ noch $\mathrm{ev} \colon V \to (V^*)^*$ surjektiv. Man kann sogar zeigen, dass die Mächtigkeit einer Basis von V^* wirklich größer als die einer Basis von V ist. Insbesondere ist die Bijektivität von ev, oder die Existenz eines Isomorphismus zwischen V und V^*, eine *Charakterisierung* der Endlichdimensionalität eines Vektorraums V.

4.2 Matrizen

Das Ziel dieses Abschnitts ist, lineare Abbildungen zwischen Vektorräumen der Gestalt K^n ganz konkret zu verstehen. Dazu führen wir den Beriff der *Matrix* ein.

Definition 4.2.1 (Matrix)**.** Sei K ein Körper und seien $m, n \in \mathbb{N}$. Eine *$m \times n$-Matrix über K* (oder *mit Koeffizienten in K*) ist eine Familie

$$A = (a_{ij})_{(i,j) \in \{1,\dots,m\} \times \{1,\dots,n\}}$$

mit $a_{ij} \in K$. Wenn m und n festgelegt sind, schreibt man oft $(a_{ij})_{i,j}$ oder sogar (a_{ij}) für eine solche Familie. Der Skalar a_{ij} ist der (i,j)-te *Koeffizient* oder *Eintrag* der Matrix A, und wird auch mit A_{ij} bezeichnet.

Eine $m \times n$-Matrix $A = (a_{ij})_{i,j}$ wird üblicherweise als Rechteck mit m Zeilen und n Spalten dargestellt:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

Inbesondere betrachten wir $m \times 1$-Matrizen als Spaltenvektoren und $1 \times n$-Matrizen als Zeilenvektoren.

Notation 4.2.2. Die Menge aller $m \times n$-Matrizen über K wird mit $M_{m\times n}(K)$ bezeichnet. Da wir schon Elemente von K^m als Spaltenvektoren betrachten, werden wir üblicherweise $M_{m\times 1}(K)$ mit K^m identifizieren. Wir können außerdem $M_{1\times 1}(K)$ mit K identifizieren. Wenn $m = n$ schreiben wir auch $M_n(K)$ anstelle von $M_{n\times n}(K)$. Matrizen in $M_n(K)$ heißen *quadratische* Matrizen.

Notation 4.2.3. Seien $A = (a_{ij})_{i,j} \in M_{m\times n}(K)$, $i \in \{1, \dots, m\}$ und $j \in \{1, \dots, n\}$. Wir bezeichnen mit

$$A_{i*} = \begin{pmatrix} a_{i1} & a_{i2} & \dots & a_{in} \end{pmatrix}$$

die i-te Zeile von A und mit

$$A_{*j} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

die j-te Spalte von A.

Matrizen fester Größe bilden einen Vektorraum über K, indem wir Addition und Skalarmultiplikation koeffizientenweise definieren (vgl. Definition 3.1.2):

Definition 4.2.4 (Addition und Skalarmultiplikation von Matrizen)**.** Seien $m, n \in \mathbb{N}$.

- Seien $A = (a_{ij})_{i,j}$ und $B = (b_{ij})_{i,j}$ $m \times n$-Matrizen über K. Ihre Summe ist die Matrix

$$A + B := (a_{ij} + b_{ij})_{i,j}.$$

- Sei $A = (a_{ij})_{i,j}$ eine $m \times n$-Matrix über K und $\lambda \in K$ ein Skalar. Dann setzen wir

$$\lambda \cdot A := (\lambda \cdot a_{ij})_{i,j}.$$

Mit dieser Addition und Skalarmultiplikation ist es klar, dass $M_{m\times n}(K)$ ein Vektorraum über K ist (siehe Proposition 3.1.5). Der Nullvektor ist die *Nullmatrix* $0 = 0_{m,n} = (0)_{i,j}$, und das Inverse einer Matrix $A = (a_{ij})_{i,j}$ bzgl. $+$ ist $-A = (-a_{ij})_{i,j}$. Außerdem hat $M_{m\times n}(K)$ eine Basis $(E_{rs})_{(r,s)\in\{1,\dots,m\}\times\{1,\dots,n\}}$, wobei E_{rs} die Matrix ist, deren Einträge alle null sind, außer dem Eintrag in der r-ten Zeile und s-ten Spalte, welcher 1 ist. Insbesondere gilt:

$$\dim_K M_{m\times n}(K) = m \cdot n.$$

Um die Definition von E_{rs} als Formel schreiben zu können, ist folgende Notation hilfreich:

Notation 4.2.5 (das Kronecker-Delta)**.**

$$\delta_{ij} := \begin{cases} 1, & \text{falls } i = j, \\ 0, & \text{falls } i \neq j. \end{cases}$$

Mithilfe des Kronecker-Deltas kann man schreiben: $E_{rs} = (\delta_{ir}\delta_{js})_{i,j}$.

Bemerkung 4.2.6. Die Addition von Matrizen unterschiedlicher Größe ist *nicht* definiert.

Definition 4.2.7 (transponierte Matrix)**.** Seien $m, n \in \mathbb{N}$ und sei $A = (a_{ij})_{i,j}$ eine $m \times n$-Matrix über K. Die *transponierte Matrix* zu A ist die $n \times m$-Matrix

$$A^\intercal := (a_{ij})_{j,i} \in M_{n\times m}(K).$$

Beispiel 4.2.8. Für die 3×2-Matrix

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

gilt

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

Durch das Transponieren werden also Zeilen zu Spalten und umgekehrt. Die Abbildung

$$\begin{aligned} M_{m\times n}(K) &\to M_{n\times m}(K), \\ A &\mapsto A^\intercal, \end{aligned}$$

ist offensichtlich K-linear, und es gilt $(A^\intercal)^\intercal = A$.

Definition 4.2.9 (Zeilenraum, Spaltenraum)**.** Seien $m, n \in \mathbb{N}$ und sei $A \in M_{m\times n}(K)$.

- Der *Zeilenraum* $\mathrm{ZR}(A)$ von A ist der von den Zeilen von A erzeugte Untervektorraum von K^n.
- Der *Spaltenraum* $\mathrm{SR}(A)$ von A ist der von den Spalten von A erzeugte Untervektorraum von K^m.

Bemerkung 4.2.10. Es gilt $\mathrm{ZR}(A) = \mathrm{SR}(A^\intercal)$ und $\mathrm{SR}(A) = \mathrm{ZR}(A^\intercal)$.

Definition 4.2.11 (Diagonalmatrix, obere/untere Dreiecksmatrix)**.** Sei $n \in \mathbb{N}$ und sei $A = (a_{ij})_{i,j}$ eine $n \times n$-Matrix über K.

- A heißt *Diagonalmatrix*, wenn alle Koeffizienten außerhalb der Hauptdiagonale null sind, d.h., wenn $a_{ij} = 0$ für alle $i \neq j$.
- A heißt *obere Dreiecksmatrix*, wenn $a_{ij} = 0$ für alle $i > j$.
- A heißt *untere Dreicksmatrix*, wenn $a_{ij} = 0$ für alle $i < j$.
- A heißt *Dreiecksmatrix*, wenn A eine obere oder untere Dreiecksmatrix ist.

Notation 4.2.12. Man schreibt $\mathrm{diag}(d_1, \ldots, d_n)$ für die $n \times n$-Diagonalmatrix $(a_{ij})_{i,j}$ mit $a_{ii} = d_i$.

4.2.1 Multiplikation von Matrizen

Definition 4.2.13 (Matrixmultiplikation)**.** Seien $m, n, p \in \mathbb{N}$. Sei $A = (a_{ij})_{i,j}$ eine $m \times n$-Matrix über K und $B = (b_{jk})_{j,k}$ eine $n \times p$-Matrix über K. Das *Produkt* von A und B ist die $m \times p$-Matrix

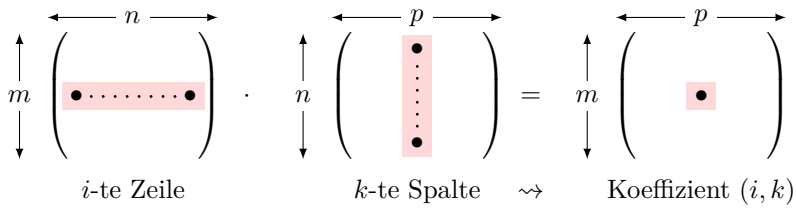
$$A \cdot B := \left(\sum_{j=1}^{n} a_{ij} \cdot b_{jk} \right)_{i,k}.$$

Um diese Definition zu verstehen ist es hilfreich, den Spezialfall $m = p = 1$ zu betrachten. In diesem Fall ist $A \in M_{1\times n}(K)$ ein Zeilenvektor und $B \in M_{n\times 1}(K)$ ein Spaltenvektor, und ihr Produkt liegt in $M_{1\times 1}(K)$, also ist ein einziger Skalar:

$$A \cdot B = \begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \sum_{j=1}^{n} a_j b_j \in K.$$

Im allgemeinen Fall mit $A \in M_{m\times n}(K)$ und $B \in M_{n\times p}(K)$ ist der (i,k)-te Eintrag in dem Produkt $A \cdot B$ das Produkt der i-te Zeile von A mit der k-te Spalte von B wie im Spezialfall:

$$A \cdot B = (A_{i*} \cdot B_{*k})_{i,k}.$$



Beispiel 4.2.14. Es gilt

$$\begin{pmatrix} 1 & -1 & 2 \\ 5 & 0 & -3 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ -2 & 4 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 4 & 1 \\ -3 & -1 \end{pmatrix},$$

$$\begin{pmatrix} 0 & 1 \\ -2 & 4 \\ 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 & 2 \\ 5 & 0 & -3 \end{pmatrix} = \begin{pmatrix} 5 & 0 & -3 \\ 18 & 2 & -16 \\ 11 & -1 & -4 \end{pmatrix}.$$

Definition 4.2.15 (Einheitsmatrix)**.** Sei $n \in \mathbb{N}$. Die $n \times n$-*Einheitsmatrix* über K ist die $n \times n$-Matrix

$$I_n := (\delta_{ij})_{i,j} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \in M_n(K).$$

Dabei ist δ_{ij} das Kronecker-Delta, siehe Notation 4.2.5.

Proposition 4.2.16 (Eigenschaften der Matrixmultiplikation)**.**

(i) *Matrixmultiplikation ist assoziativ. Das heißt, für alle $m, n, p, q \in \mathbb{N}$, $A \in M_{m\times n}(K)$, $B \in M_{n\times p}(K)$ und $C \in M_{p\times q}(K)$ gilt*

$$(A \cdot B) \cdot C = A \cdot (B \cdot C).$$

(ii) *Die Einheitsmatrix ist ein neutrales Element bzgl. Matrixmultiplikation. Das heißt, für alle $m, n \in \mathbb{N}$ und $A \in M_{m\times n}(K)$ gilt*

$$I_m \cdot A = A \quad \text{und} \quad A \cdot I_n = A.$$

(iii) *Matrixmultiplikation ist distributiv über Addition. Das heißt, für alle $m, n, p \in \mathbb{N}$, $A, B \in M_{m\times n}(K)$ und $C, D \in M_{n\times p}(K)$ gilt*

$$A \cdot (C + D) = A \cdot C + A \cdot D \quad \text{und} \quad (A + B) \cdot C = A \cdot C + B \cdot C.$$

(iv) *Matrixmultiplikation ist kompatibel mit Skalarmultiplikation im folgenden Sinne: Für alle* $m, n, p \in \mathbb{N}$, $A \in M_{m\times n}(K)$, $B \in M_{n\times p}(K)$ *und* $\lambda \in K$ *gilt*

$$(\lambda \cdot A) \cdot B = \lambda \cdot (A \cdot B) \quad \textit{und} \quad A \cdot (\lambda \cdot B) = \lambda \cdot (A \cdot B).$$

(v) *Matrixmultiplikation ist kompatibel mit Transponieren im folgenden Sinne: Für alle* $m, n, p \in \mathbb{N}$, $A \in M_{m\times n}(K)$ *und* $B \in M_{n\times p}(K)$ *gilt*

$$(A \cdot B)^\intercal = B^\intercal \cdot A^\intercal.$$

Beweis. Zu (i). Nach Definition der Matrixmultiplikaton gilt

$$(A \cdot B)_{ik} = \sum_{j=1}^{n} A_{ij}B_{jk} \quad \text{und} \quad (B \cdot C)_{jl} = \sum_{k=1}^{p} B_{jk}C_{kl},$$

und daher

$$\begin{aligned}
((A \cdot B) \cdot C)_{il} &= \sum_{k=1}^{p} (A \cdot B)_{ik} C_{kl} \\
&= \sum_{k=1}^{p} \left(\sum_{j=1}^{n} A_{ij}B_{jk} \right) C_{kl} \\
&= \sum_{k=1}^{p} \sum_{j=1}^{n} A_{ij}B_{jk}C_{kl} \\
&= \sum_{j=1}^{n} \sum_{k=1}^{p} A_{ij}B_{jk}C_{kl} \\
&= \sum_{j=1}^{n} A_{ij} \left(\sum_{k=1}^{p} B_{jk}C_{kl} \right) \\
&= \sum_{j=1}^{n} A_{ij}(B \cdot C)_{jl} \\
&= (A \cdot (B \cdot C))_{il}.
\end{aligned}$$

Dabei haben wir mehrere Eigenschaften von $+$ und $\cdot$ im Körper K verwendet: das verallgemeinerte Disbtributivgesetz (Bemerkung 2.3.3), die Assoziativität und Kommutativität von $+$ und die Assoziativität von $\cdot$. Diese Berechnung gilt für alle $i \in \{1, \ldots, m\}$ und $l \in \{1, \ldots, q\}$, und damit ist die gewünschte Matrixgleichung bewiesen.

Zu (ii). Für alle i, j gilt

$$(I_m \cdot A)_{ij} = \sum_{e=1}^{m} \delta_{ie}A_{ej} = A_{ij},$$

und somit $I_m \cdot A = A$. Der Beweis der Gleichung $A \cdot I_n = A$ ist ähnlich.

Zu (iii). Wir überprüfen nur die erste Gleichung:

$$\begin{aligned}
(A \cdot (C+D))_{ik} &= \sum_{j=1}^{n} A_{ij}(C+D)_{jk} && \text{(Definition der Matrixmultiplikation)}\\
&= \sum_{j=1}^{n} A_{ij}(C_{jk}+D_{jk}) && \text{(Definition der Matrixaddition)}\\
&= \sum_{j=1}^{n} (A_{ij}C_{jk} + A_{ij}D_{jk}) && \text{(Distributivgesetz in } K)\\
&= \sum_{j=1}^{n} A_{ij}C_{jk} + \sum_{j=1}^{n} A_{ij}D_{jk} && \text{(Assoz. \& Komm. der Addition in } K)\\
&= (A \cdot C)_{ik} + (A \cdot D)_{ik} && \text{(Definition der Matrixmultiplikation)}\\
&= ((A \cdot C) + (A \cdot D))_{ik}. && \text{(Definition der Matrixaddition)}
\end{aligned}$$

Zu (iv). Wir überprüfen nur die zweite Gleichung:

$$\begin{aligned}
(A \cdot (\lambda \cdot B))_{ik} &= \sum_{j=1}^{n} A_{ij}(\lambda \cdot B)_{jk} && \text{(Definition der Matrixmultiplikation)}\\
&= \sum_{j=1}^{n} A_{ij}(\lambda B_{jk}) && \text{(Def. der Skalarmultiplaktion von Matrizen)}\\
&= \sum_{j=1}^{n} \lambda(A_{ij}B_{jk}) && \text{(Assoz. \& Komm. der Multiplikation in } K)\\
&= \lambda \sum_{j=1}^{n} A_{ij}B_{ij} && \text{(Distributivgesetz in } K)\\
&= \lambda(A \cdot B)_{ik} && \text{(Definition der Matrixmultiplikation)}\\
&= (\lambda \cdot (A \cdot B))_{ik}. && \text{(Def. der Skalarmultiplaktion von Matrizen)}
\end{aligned}$$

Zu (v). Für alle $i \in \{1, \ldots, m\}$ und $k \in \{1, \ldots, p\}$ gilt

$$((A \cdot B)^{\intercal})_{ki} = (A \cdot B)_{ik} = \sum_{j=1}^{n} A_{ij} \cdot B_{jk} = \sum_{j=1}^{n} (B^{\intercal})_{kj} \cdot (A^{\intercal})_{ji} = (B^{\intercal} \cdot A^{\intercal})_{ki}. \qquad \square$$

Bemerkung 4.2.17. Sei $n \in \mathbb{N}$. Die Matrixmultiplikation definiert eine Verknüpfung

$$\cdot \colon M_n(K) \times M_n(K) \to M_n(K)$$

auf quadratischen $n \times n$-Matrizen. Diese Verknüpfung ist assoziativ nach Proposition 4.2.16(i) und distributiv über die Addition nach Proposition 4.2.16(iii). Wenn $n \geq 2$ ist sie aber nicht kommutativ. Zum Beispiel:

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Das Tripel $(M_n(K), +, \cdot)$ ist ein Beispiel eines *nicht-kommutativen Ringes* (siehe Bemerkung 2.3.5).

Notation 4.2.18 (Matrixpotenzen)**.** Sei $n \in \mathbb{N}$ und $A \in M_n(K)$ eine quadratische Matrix. Man definiert die Potenzen A^k mit $k \in \mathbb{N}$ rekursiv wie folgt:

$$\begin{aligned}
&A^0 = I_n,\\
&A^n = A \cdot A^{n-1} \quad \text{(für alle } n \geq 1).
\end{aligned}$$

Wegen der Assoziativität der Matrixmultiplikation (und die Neutralität von I_n) gilt dann $A^1 = A$ und

$$A^{k+l} = A^k \cdot A^l \quad \text{und} \quad (A^l)^k = A^{kl}$$

für alle $k, l \in \mathbb{N}$ (beide Formeln lassen sich leicht durch Induktion über k nachprüfen, vgl. Proposition 2.1.7).

Beispiel 4.2.19 (Fibonacci-Zahlen)**.** Die Folge $(F_n)_{n\in\mathbb{N}}$ der *Fibonacci-Zahlen* wird durch folgende Gleichungen rekursiv definiert:

$$\begin{aligned} F_0 &= 0, \\ F_1 &= 1, \\ F_{n+1} &= F_n + F_{n-1} \quad \text{(für alle } n \geq 1\text{)}. \end{aligned}$$

Die ersten Fibonacci-Zahlen sind also

$$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, \ldots$$

Diese rekursive Definition kann mithilfe der Matrixmultiplikation ausgedrückt werden:

$$\begin{pmatrix} F_{n+1} \\ F_n \end{pmatrix} = \begin{pmatrix} F_n + F_{n-1} \\ F_n \end{pmatrix} = A \cdot \begin{pmatrix} F_n \\ F_{n-1} \end{pmatrix}, \quad \text{wobei } A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \in M_{2\times 2}(\mathbb{Q}).$$

Daraus folgt die folgende nicht-rekursive Formel für die Fibonacci-Zahlen:

$$\begin{pmatrix} F_{n+1} \\ F_n \end{pmatrix} = A^n \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Damit stellt sich noch die Frage, ob man die Matrixpotenz A^n effektiv berechnen kann. Das werden wir später schaffen, und damit eine explizite Formel für F_n erhalten (siehe Beispiel 6.2.33).

Definition 4.2.20 (invertierbare Matrix, inverse Matrix)**.** Sei $n \in \mathbb{N}$. Eine quadratische Matrix $A \in M_n(K)$ heißt *invertierbar* oder *regulär*, wenn sie ein inverses Element bezüglich Matrixmultiplikation besitzt, d.h., wenn eine Matrix $B \in M_n(K)$ existiert, so dass

$$A \cdot B = I_n \quad \text{und} \quad B \cdot A = I_n.$$

Die Matrix B ist dann eindeutig bestimmt (siehe Proposition 2.1.3); sie heißt die *inverse Matrix* zu A und wird mit A^{-1} bezeichnet.

Wir werden später untersuchen, wie man die Invertierbarkeit einer Matrix bestimmen kann und wie man die inverse Matrix berechnen kann (siehe Abschnitt 5.2.1).

Beispiel 4.2.21.

(i) Eine 1×1-Matrix über K entspricht einem einzigen Skalar $a \in K$. Sie ist genau dann invertierbar, wenn $a \in K^*$.

(ii) Für die Matrix

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

gilt $A^2 = I_2$. Also ist A invertierbar mit $A^{-1} = A$.

(iii) Für $a, b \in K$ gilt

$$\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & a+b \\ 0 & 1 \end{pmatrix}.$$

Daraus folgt, dass

$$\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -a \\ 0 & 1 \end{pmatrix}.$$

Bemerkung 4.2.22. Sind $A, B \in M_n(K)$ invertierbar, so ist $A \cdot B$ invertierbar und es gilt

$$(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}.$$

Definition 4.2.23 (allgemeine lineare Gruppe)**.** Sei $n \in \mathbb{N}$. Die *allgemeine lineare Gruppe* über K ist die Menge

$$\mathrm{GL}_n(K) = \{A \in M_n(K) \mid A \text{ ist invertierbar}\}$$

versehen mit der Matrixmultiplikation.

Bemerkung 4.2.24. Wenn $n \geq 2$ ist die allgemeine lineare Gruppe $\mathrm{GL}_n(K)$ nicht abelsch (nach Bemerkung 4.2.17).

4.2.2 Lineare Abbildungen aus Matrizen

Sei A eine $m \times n$-Matrix über K. Man definiert eine Abbildung $L_A \colon K^n \to K^m$ wie folgt:

$$\begin{aligned} L_A \colon K^n &\to K^m, \\ v &\mapsto A \cdot v. \end{aligned}$$

Dabei betrachten wir v als Spaltenvektor, d.h., $v \in M_{n\times 1}(K)$, und $A \cdot v \in M_{m\times 1}$ ist das Matrixprodukt von A mit v. Konkreter, ist $A = (a_{ij})_{i,j}$, so ist

$$L_A(v) = A \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} a_{11}v_1 + a_{12}v_2 + a_{13}v_3 + \cdots + a_{1n}v_n \\ a_{21}v_1 + a_{22}v_2 + a_{23}v_3 + \cdots + a_{2n}v_n \\ \vdots \\ a_{m1}v_1 + a_{m2}v_2 + a_{m3}v_3 + \cdots + a_{mn}v_n \end{pmatrix}.$$

Bemerkung 4.2.25 (Bild der Standardeinheitsvektoren)**.** Seien $e_1, \ldots, e_n \in K^n$ die Standardeinheitsvektoren und sei $A \in M_{m\times n}(K)$. Für alle $j \in \{1, \ldots, n\}$ gilt

$$L_A(e_j) = A \cdot e_j = A_{*j} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}.$$

Das heißt: *Das Bild von e_j unter L_A ist die j-te Spalte von A.*

Proposition 4.2.26. *Seien $m, n, p \in \mathbb{N}$.*

(i) *Für alle $A \in M_{m\times n}(K)$ ist die Abbildung $L_A \colon K^n \to K^m$ K-linear.*

(ii) *Für alle $A \in M_{m\times n}(K)$ und $B \in M_{n\times p}(K)$ gilt $L_{A\cdot B} = L_A \circ L_B$.*

(iii) *Es gilt $L_{I_n} = \mathrm{id}_{K^n}$.*

(iv) *Für alle $A, B \in M_{m\times n}(K)$ gilt $L_{A+B} = L_A + L_B$.*

(v) *Für alle $A \in M_{m\times n}(K)$ und $\lambda \in K$ gilt $L_{\lambda\cdot A} = \lambda \cdot L_A$.*

Beweis. Zur ersten Aussage ist zu zeigen: Für alle $v, v' \in K^n$ und $\lambda \in K$ gelten

$$L_A(v + v') = L_A(v) + L_A(v') \quad \text{und} \quad L_A(\lambda \cdot v) = \lambda \cdot L_A(v).$$

Dies folgt aus der Definition von L_A und Proposition 4.2.16(iii,iv). Aussagen (ii) bis (v) folgen direkt aus Aussagen (i) bis (iv) der Proposition 4.2.16. □

Bemerkung 4.2.27. Nach Bemerkung 4.2.25 und Proposition 4.1.20(i) ist der Spaltenraum von $A \in M_{m\times n}(K)$ genau das Bild von L_A:

$$\mathrm{SR}(A) = \operatorname{im} L_A \subset K^m.$$

Der Zusammenhang zwischen L_A und dem Zeilenraum von A ist nicht so offensichtlich (eigentlich ist $\mathrm{ZR}(A)$ das orthogonale Komplement von $\ker L_A$ in K^n).

Bemerkung 4.2.28. Aussagen (i), (iv) und (v) der Proposition 4.2.26 können wie folgt zusammengefasst werden: Es gibt eine K-lineare Abbildung

$$\begin{aligned} L\colon M_{m\times n}(K) &\to \mathrm{Hom}_K(K^n, K^m), \\ A &\mapsto L_A. \end{aligned}$$

Dabei ist $\mathrm{Hom}_K(K^n, K^m)$ der K-Vektorraum aller K-linearen Abbildungen von K^n nach K^m (siehe Proposition 4.1.43).

Satz 4.2.29. *Seien $m, n \in \mathbb{N}$ und sei $f\colon K^n \to K^m$ eine K-lineare Abbildung. Dann existiert genau eine $m \times n$-Matrix A über K, so dass $f = L_A$. Anders gesagt ist die Abbildung*

$$L\colon M_{m\times n}(K) \to \mathrm{Hom}_K(K^n, K^m)$$

ein Isomorphismus.

Beweis. Dass höchstens eine solche Matrix A existiert folgt aus Bemerkung 4.2.25: Wenn $L_A = L_B$, dann ist inbesondere

$$A_{*j} = L_A(e_j) = L_B(e_j) = B_{*j}$$

für alle $j \in \{1, \dots, n\}$, und daher ist $A = B$. Es bleibt zu zeigen, dass A existiert. Sei A die $m \times n$-Matrix, deren j-te Spalte gleich $f(e_j) \in K^m$ ist. Wir behaupten, dass $L_A = f$. Beide Abbildungen L_A und f sind K-linear (die erste nach Proposition 4.2.26(i)). Nach Bemerkung 4.2.25 gilt $L_A(e_j) = f(e_j)$ für alle $j \in \{1, \dots, n\}$. Da $(e_1, \dots, e_n)$ eine Basis von K^n ist, folgt aus dem Satz 4.1.22, dass $L_A = f$. □

Notation 4.2.30. Die Umkehrabbildung von L bezeichnen wir mit

$$M\colon \mathrm{Hom}_K(K^n, K^m) \xrightarrow{\sim} M_{m\times n}(K).$$

Das heißt, ist $f\colon K^n \to K^m$ eine linear Abbildung, so ist $M(f)$ die $m\times n$-Matrix mit Spalten $f(e_1)$, …, $f(e_n)$.

Bemerkung 4.2.31. Nach Proposition 4.2.26(ii) ist der Isomorphismus M auch mit den multiplikativen Verknüpfungen verträglich: Für alle $m, n, p \in \mathbb{N}$ ist folgendes Diagramm kommutativ:

$$\begin{array}{ccc} \mathrm{Hom}_K(K^n, K^m) \times \mathrm{Hom}_K(K^p, K^n) & \xrightarrow{\circ} & \mathrm{Hom}_K(K^p, K^m) \\ \downarrow{\scriptstyle M\times M} & & \downarrow{\scriptstyle M} \\ M_{m\times n}(K) \times M_{n\times p}(K) & \xrightarrow{\cdot} & M_{m\times p}(K). \end{array}$$

Beispiel 4.2.32. Wir listen die Matrizen $A \in M_2(\mathbb{R})$ auf, deren zugeordnete $\mathbb{R}$-lineare Abbildungen $L_A\colon \mathbb{R}^2 \to \mathbb{R}^2$ die im Beispiel 4.1.7 sind:

(i) Skalierung um $\frac{1}{2}$: $\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$.

(ii) Spiegelung an $\mathbb{R}e_2$: $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$.

(iii) Drehung um $\frac{\pi}{4}$: $\begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}$.

(iv) Spiegelung an 0: $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$.

(v) Horizontale Skalierung um $\frac{3}{2}$: $\begin{pmatrix} \frac{3}{2} & 0 \\ 0 & 1 \end{pmatrix}$.

(vi) Koordinatenvertauschung: $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

(vii) Orthogonale Projektion auf $\mathbb{R}e_1$: $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.

(viii) Horizontale Scherung um $\frac{1}{2}$: $\begin{pmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{pmatrix}$.

Beispiel 4.2.33 (Drehmatrizen)**.** Sei $\alpha \in \mathbb{R}$ eine reelle Zahl und sei

$$D(\alpha) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \in M_2(\mathbb{R}).$$

Die entsprechende lineare Abbildung $L_{D(\alpha)} \colon \mathbb{R}^2 \to \mathbb{R}^2$ ist eine Drehung um den Winkel α um den Nullpunkt. Insbesondere gilt $L_{D(\alpha+\beta)} = L_{D(\alpha)} \circ L_{D(\beta)}$. Aus Proposition 4.2.26(ii) und der Injektivität von L folgt:

$$D(\alpha + \beta) = D(\alpha) \cdot D(\beta).$$

Diese Gleichung liefert die gewöhnlichen Formeln für $\cos(\alpha + \beta)$ und $\sin(\alpha + \beta)$.

Bemerkung 4.2.34. Man kann den Satz 4.2.29 auf Vektorräume der Gestalt $K^{(I)}$ verallgemeinern, aber die Aussage wird komplizierter. Man nennt eine $I \times J$-indizierte Familie in K eine $I \times J$-*Matrix* über K. Aus dem Satz 4.1.22 erhalten wir einen Isomorphismus

$$L \colon M_{I\times J}(K) \xrightarrow{\sim} \mathrm{Hom}_K(K^{(J)}, K^I).$$

Ist $A \in M_{I\times J}(K)$, so landet die lineare Abbildung $L_A \colon K^{(J)} \to K^I$ in $K^{(I)}$ genau dann, wenn jede Spalte von A in $K^{(I)}$ liegt, d.h., wenn jede Spalte von A nur endlich viele Koeffizienten enthält, die nicht null sind. Eine solche Matrix heißt *spaltenendlich*. Dementsprechend schränkt sich der obige Isomorphismus zu einem Isomorphismus zwischen $\mathrm{Hom}_K(K^{(J)}, K^{(I)})$ und dem Vektorraum der spaltenendlichen $I \times J$-Matrizen ein.

Definition 4.2.35 (Rang einer Matrix)**.** Seien $m, n \in \mathbb{N}$ und sei $A \in M_{m\times n}(K)$. Der *Rang* von A ist der Rang von L_A (siehe Definition 4.1.41):

$$\mathrm{rg}\, A := \mathrm{rg}\, L_A = \dim_K(\mathrm{im}\, L_A).$$

Bemerkung 4.2.36. Da das Bild von L_A der Spaltenraum von A ist (Bemerkung 4.2.27), gilt

$$\mathrm{rg}\, A = \dim_K \mathrm{SR}(A).$$

Wir werden später beweisen, dass auch $\mathrm{rg}\, A = \dim_K \mathrm{ZR}(A)$ (siehe Korollar 4.2.49).

Proposition 4.2.37 (Charakterisierung der Invertierbarkeit)**.** *Sei $n \in \mathbb{N}$ und $A \in M_n(K)$. Die folgenden Aussagen sind äquivalent:*

(i) *A ist invertierbar.*

(ii) *A ist von links invertierbar, d.h., es existiert $B \in M_n(K)$ mit $B \cdot A = I_n$.*

(iii) *A ist von rechts invertierbar, d.h., es existiert $B \in M_n(K)$ mit $A \cdot B = I_n$.*

(iv) *Die lineare Abbildung $L_A \colon K^n \to K^n$ ist bijektiv.*

(v) *Die lineare Abbildung $L_A \colon K^n \to K^n$ ist injektiv.*

(vi) *Die lineare Abbildung $L_A \colon K^n \to K^n$ ist surjektiv.*

(vii) *Es gilt* $\operatorname{rg} A = n$.

Beweis. Die Implikationen (i) $\Rightarrow$ (ii) und (i) $\Rightarrow$ (iii) sind klar. Die Äquivalenz von (iv), (v) und (vi) folgt aus Korollar 4.1.39, und die Äquivalenz von (vi) und (vii) folgt aus Proposition 3.3.35: $L_A \colon K^n \to K^n$ ist genau dann surjektiv, wenn $\dim_K(\operatorname{im} L_A) = n$. Die Implikation (ii) $\Rightarrow$ (v) folgt aus Proposition 4.2.26(ii,iii), denn $B \cdot A = I_n$ impliziert $L_B \circ L_A = \mathrm{id}_{K^n}$, und damit ist L_A injektiv (nach Proposition 1.3.25(iii)). Die Implikation (iii) $\Rightarrow$ (vi) folgt auf ähnliche Weise.

Zum Schluss beweisen wir (iv) $\Rightarrow$ (i). Sei g die Umkehrabbildung von L_A und sei $B = M(g)$ die zugehörige Matrix. Nach Bemerkung 4.2.31 gilt

$$B \cdot A = M(g) \cdot M(L_A) = M(g \circ L_A) = M(\mathrm{id}_{K^n}) = I_n,$$

und ebenso $A \cdot B = I_n$. □

4.2.3 Darstellung von linearen Abbildungen

Nach Satz 4.2.29 können wir jede K-lineare Abbildung $f \colon K^n \to K^m$ als eine $m \times n$-Matrix A auffassen. Diese Darstellung von linearen Abbildungen als Matrizen ist sehr nützlich für Berechnungen. Zum Beispiel entspricht die Komposition von linearen Abbildungen der Matrixmultiplikation.

In diesem Abschnitt wollen wir diese Matrixdarstellung auf lineare Abbildungen $f \colon V \to W$ zwischen beliebigen endlich-dimensionalen K-Vektorräumen verallgemeinern. Wenn $B = (v_1, \dots, v_n)$ eine Basis von V ist, gibt es bekanntlich einen Isomorphismus

$$\varphi_B \colon K^n \xrightarrow{\sim} V$$

mit $\varphi_B(e_i) = v_i$ (siehe Bemerkung 3.3.18). Ist $v \in V$, so heißt der Spaltenvektor

$$[v]_B := \varphi_B^{-1}(v) \in K^n$$

der *Koordinatenvektor* von v bzgl. B (Definition 3.3.17). Sind V und W Vektorräume über K der Dimension n und m mit Basen B und C, so erhalten wir einen Isomorphismus

$$\begin{aligned} \mathrm{Hom}_K(V, W) &\xrightarrow{\sim} M_{m\times n}(K), \\ f &\mapsto [f]_C^B, \end{aligned}$$

als die Komposition der Isomorphismen

$$\mathrm{Hom}_K(V, W) \xrightarrow{\mathrm{Hom}_K(\varphi_B, \varphi_C^{-1})} \mathrm{Hom}_K(K^n, K^m) \xrightarrow{M} M_{m\times n}(K).$$

Seine Umkehrabbildung ist die Komposition

$$M_{m\times n}(K) \xrightarrow{L} \mathrm{Hom}_K(K^n, K^m) \xrightarrow{\mathrm{Hom}_K(\varphi_B^{-1}, \varphi_C)} \mathrm{Hom}_K(V, W).$$

Definition 4.2.38 (Darstellungsmatrix)**.** Seien V und W endlich-dimensionale K-Vektorräume mit Basen B und C und sei $f \colon V \to W$ eine lineare Abbildung. Die Matrix $[f]_C^B$ heißt die *Darstellungsmatrix* oder *Abbildungsmatrix* von f bzgl. der Basen B und C.

Seien $B = (v_1, \ldots, v_n)$ und $C = (w_1, \ldots, w_m)$. Nach Definition ist $[f]_C^B$ die $m \times n$-Matrix entsprechend der linearen Abbildung

$$\varphi_C^{-1} \circ f \circ \varphi_B \colon K^n \to K^m,$$

d.h., $[f]_C^B = M(\varphi_C^{-1} \circ f \circ \varphi_B)$. Die j-te Spalte von $[f]_C^B$ ist also

$$(\varphi_C^{-1} \circ f \circ \varphi_B)(e_j) = \varphi_C^{-1}(f(v_j)) = [f(v_j)]_C.$$

Anders gesagt, ist $[f]_C^B = (a_{ij})_{i,j}$, so gilt

$$f(v_j) = \sum_{i=1}^{m} a_{ij} \cdot w_i$$

für alle $j \in \{1, \ldots, n\}$, und die Matrix $[f]_C^B$ ist dadurch eindeutig bestimmt.

Proposition 4.2.39. *Seien V, W und U endlich-dimensionale K-Vektorräume mit Basen B, C und D.*

(i) *Für jede lineare Abbildung $f \colon V \to W$ und jeden Vektor $v \in V$ gilt*

$$[f(v)]_C = [f]_C^B \cdot [v]_B.$$

(ii) *Für lineare Abbildungen $f \colon V \to W$ und $g \colon W \to U$ gilt*

$$[g \circ f]_D^B = [g]_D^C \cdot [f]_C^B.$$

(iii) *Es gilt*

$$[\mathrm{id}_V]_B^B = I_n,$$

wobei $n = \dim_K(V)$.

(iv) *Für jeden Isomorphismus $f \colon V \xrightarrow{\sim} W$ ist die Matrix $[f]_C^B$ invertierbar, und es gilt*

$$[f^{-1}]_B^C = ([f]_C^B)^{-1}.$$

Beweis. Zu (i). Es gilt

$$[f]_C^B \cdot [v]_B = (\varphi_C^{-1} \circ f \circ \varphi_B)(\varphi_B^{-1}(v)) = \varphi_C^{-1}(f(v)) = [f(v)]_C.$$

Zu (ii). Dies folgt aus der Gleichung

$$(\varphi_D^{-1} \circ g \circ \varphi_C) \circ (\varphi_C^{-1} \circ f \circ \varphi_B) = \varphi_D^{-1} \circ (g \circ f) \circ \varphi_B,$$

indem wir M auf beide Seiten anwenden.

Zu (iii). $[\mathrm{id}_V]_B^B = M(\varphi_B^{-1} \circ \mathrm{id}_V \circ \varphi_B) = M(\mathrm{id}_{K^n}) = I_n$.

Zu (iv). Dies folgt aus (ii) and (iii). □

Bemerkung 4.2.40 (Rang einer Darstellungsmatrix)**.** Der Rang der Darstellungsmatrix $[f]_C^B$ ist gleich dem Rang von f. Denn es gilt

$$\mathrm{rg}\, [f]_C^B = \mathrm{rg}(\varphi_C^{-1} \circ f \circ \varphi_B) = \dim_K \mathrm{im}(\varphi_C^{-1} \circ f \circ \varphi_B)$$

und

$$\mathrm{im}(\varphi_C^{-1} \circ f \circ \varphi_B) = \mathrm{im}(\varphi_C^{-1} \circ f) = \varphi_C^{-1}(\mathrm{im}\, f) \cong \mathrm{im}\, f,$$

da φ_B surjektiv ist und φ_C^{-1} injektiv ist.

Definition 4.2.41 (Basiswechselmatrix)**.** Seien B und B' zwei Basen eines n-dimensionalen K-Vektorraums V. Die *Basiswechselmatrix* $T_{B'}^{B}$ von B nach B' ist die Matrix

$$T_{B'}^{B} := [\mathrm{id}_V]_{B'}^{B} = M(\varphi_{B'}^{-1} \circ \varphi_B) \in M_n(K).$$

Nach Proposition 4.2.39(i) gilt also

$$T_{B'}^{B} \cdot [v]_B = [v]_{B'}$$

für alle Vektoren $v \in V$. Nach Proposition 4.2.39(iv) ist die Basiswechselmatrix $T_{B'}^{B}$ invertierbar, mit

$$(T_{B'}^{B})^{-1} = T_{B}^{B'}.$$

Beispiel 4.2.42. Sei $B = (v_1, \ldots, v_n)$ eine Basis von K^n. Die Basiswechselmatrix von B nach der Standardbasis ist die Matrix $\begin{pmatrix} v_1 & \ldots & v_n \end{pmatrix}$.

Proposition 4.2.43 (Basiswechselformel)**.** *Seien V und W endlich-dimensionale Vektorräume über K. Seien B, B' Basen von V und C, C' Basen von W. Für jede lineare Abbildung $f\colon V \to W$ gilt*

$$[f]_{C'}^{B'} = T_{C'}^{C} \cdot [f]_{C}^{B} \cdot T_{B}^{B'}.$$

Insbesondere: Für jeden Endomorphismus $f\colon V \to V$ gilt

$$[f]_{B'}^{B'} = T_{B'}^{B} \cdot [f]_{B}^{B} \cdot T_{B}^{B'}.$$

Beweis. Dies folgt unmittelbar aus Proposition 4.2.39(ii):

$$T_{C'}^{C} \cdot [f]_{C}^{B} \cdot T_{B}^{B'} = [\mathrm{id}_W]_{C'}^{C} \cdot [f]_{C}^{B} \cdot [\mathrm{id}_V]_{B}^{B'} = [\mathrm{id}_W \circ f \circ \mathrm{id}_V]_{C'}^{B'} = [f]_{C'}^{B'}. \qquad \square$$

Beispiel 4.2.44. Sei $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ die lineare Abbildung $f = L_A$ mit

$$A = \begin{pmatrix} 3 & -1 \\ 2 & 0 \end{pmatrix}.$$

Wir betrachten die Basis $B = (v_1, v_2)$ von $\mathbb{R}^2$ mit

$$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{und} \quad v_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Die Basiswechselmatrix von B nach der Standardbasis E ist also

$$T_{E}^{B} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix},$$

und man kenn leicht durch Multiplizieren nachprüfen, dass

$$T_{B}^{E} = (T_{E}^{B})^{-1} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}.$$

Nach Proposition 4.2.43 gilt

$$[f]_{B}^{B} = T_{B}^{E} \cdot A \cdot T_{E}^{B} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}.$$

Diese Darstellungsmatrix zeigt, dass die lineare Abbildung f folgende einfache geometrische Beschreibung besitzt: Sie skaliert um den Faktor 2 in Richtung von v_1, und sie lässt die Vektoren in Richtung von v_2 fest. Die lineare Abbildung f ist ein Beispiel einer *diagonalisierbaren* Abbildung (siehe Abschnitt 6.2.1).

Dieses Beispiel zeigt folgendes: Selbst wenn wir nur an Vektorräumen der Gestalt K^n Interesse haben, ist es oft hilfreich, die Darstellungsmatrix einer linearen Abbildung bezüglich beliebiger Basen zu betrachten. Denn mit einer geeigneten Wahl von Basen kann die Darstellungsmatrix besonders einfach sein, was uns hilft, die Abbildung besser zu verstehen. Man kann insbesondere die folgenden Fragen stellen:

Frage 4.2.45. Seien V und W endlich-dimensionale Vektorräume über K.

(i) Für eine lineare Abbildung $f\colon V \to W$, wie kann man Basen B von V und C von W finden, so dass die Matrix $[f]_C^B$ so einfach wie möglich ist?

(ii) Für einen Endomorphismus $g\colon V \to V$, wie kann man eine Basis B von V finden, so dass die Matrix $[g]_B^B$ so einfach wie möglich ist?

Die möglichst einfache Darstellungsmatrix $[f]_C^B$ bzw. $[g]_B^B$ heißt die *Smithsche Normalform* der linearen Abbildung f bzw. die *Jordansche Normalform* des Endomorphismus g. Die Existenz der Smithschen Normalform können wir bereits beweisen:

Satz 4.2.46 (Smithsche Normalform linearer Abbildungen)**.** *Seien V und W endlich-dimensionale K-Vektorräume der Dimension n und m, und sei $f\colon V \to W$ eine lineare Abbildung. Dann existieren Basen B von V und C von W, so dass*

$$[f]_C^B = \begin{pmatrix} I_r & 0_{r,n-r} \\ 0_{m-r,r} & 0_{m-r,n-r} \end{pmatrix},$$

wobei $r = \operatorname{rg} f$.

Beweis. Das ist der endlich-dimensionale Fall des Satzes 4.1.42, aber wir geben wieder einen vollständigen Beweis. Nach der Dimensionsformel für lineare Abbildungen hat der Kern von f die Dimension $n - r$. Nach dem Basisergänzungssatz existiert eine Basis $B = (v_1, \dots, v_n)$ von V, so dass $(v_{r+1}, \dots, v_n)$ eine Basis von $\ker f$ ist. Sei $U = \operatorname{Span}_K\{v_1, \dots, v_r\}$. Da U und $\ker f$ komplementär sind, ist die Einschränkung $f|_U$ injektiv, und damit ist die Familie $(f(v_1), \dots, f(v_r))$ in W linear unabhängig (Proposition 4.1.20(ii)). Nach dem Basisergänzungssatz gibt es weitere Vektoren $w_{r+1}, \dots, w_m$, so dass die Familie

$$C = (f(v_1), \dots, f(v_r), w_{r+1}, \dots, w_m)$$

eine Basis von W ist. Für $i \le r$ gilt dann $[f(v_i)]_C = e_i$ und für $i > r$ gilt $f(v_i) = 0$. Also hat die Matrix $[f]_C^B$ die gewünschte Form. □

Es stellt sich noch die Frage, wie man solche Basen B und C im konkreten Fall $V = K^n$ und $W = K^m$ finden kann. Eine effektive Methode zu solchen Berechnungen werden wir im Kapitel 5 erklären, und damit wird die Frage 4.2.45(i) vollständig beantwortet werden. Die Frage 4.2.45(ii) ist wesentlich schwieriger. Wir werden sie im Kapitel 6 nur teilweise beantworten; eine vollständige Antwort ist ein Ziel der Vorlesung *Lineare Algebra II*.

Zum Abschluss dieses Kapitels erklären wir den Zusammenhang zwischen der dualen Abbildung (Definition 4.1.47) und der transponierten Matrix (Definition 4.2.7):

Proposition 4.2.47 (Darstellungsmatrix der dualen Abbildung)**.** *Seien V und W endlich-dimensionale K-Vektorräume mit Basen B und C, und seien B^* und C^* die dualen Basen von V^* und W^*. Für jede lineare Abbildung $f\colon V \to W$ gilt*

$$[f^*]_{B^*}^{C^*} = ([f]_C^B)^\intercal.$$

Beweis. Seien

$$B = (v_1, \dots, v_n) \quad \text{und} \quad C = (w_1, \dots, w_m)$$

die gegebenen Basen und seien

$$B^* = (v_1^*, \dots, v_n^*) \quad \text{und} \quad C^* = (w_1^*, \dots, w_m^*),$$

die zugehörigen dualen Basen. Sei $[f]_C^B = (a_{ij})_{i,j}$. Nach Definition der Matrix $[f]_C^B$ ist ihre j-te Spalte der Koordinatenvektor $[f(v_j)]_C$, das heißt:

$$f(v_j) = \sum_{i=1}^{m} a_{ij} \cdot w_i.$$

Unser Ziel ist, die Koordinatenvektoren $[f^*(w_i^*)]_{B^*}$ zu berechnen. Es gilt

$$f^*(w_i^*)(v_j) = w_i^*(f(v_j)) = \sum_{k=1}^{m} a_{kj} \cdot w_i^*(w_k) = a_{ij},$$

da $w_i^*(w_k) = \delta_{ik}$. Daraus folgt, dass

$$f^*(w_i^*) = \sum_{j=1}^{n} a_{ij} \cdot v_j^*,$$

da beide linearen Abbildungen $V \to K$ dieselbe Werte auf der Basis B von V nehmen. Diese Gleichung bedeutet, dass der Koordinatenvektor $[f^*(w_i^*)]_{B^*}$ gleich der i-ten Zeile von $[f]_C^B$ ist, d.h., dass $[f^*]_{B^*}^{C^*}$ die transponierte Matrix zu $[f]_C^B$ ist. □

Bemerkung 4.2.48. Sei $A \in M_{m\times n}(K)$. Die Standardbasis von K^n induziert einen Isomorphismus $\varepsilon\colon K^n \xrightarrow{\sim} (K^n)^*$ mit $\varepsilon(e_i) = \pi_i$ (siehe Konstruktion 4.1.54). Proposition 4.2.47, angewendet auf die lineare Abbildung L_A und die Standardbasen, liefert ein kommutatives Quadrat

$$\begin{array}{ccc} K^m & \xrightarrow{L_{A^\intercal}} & K^n \\ {\scriptstyle\varepsilon}\downarrow\wr & & {\scriptstyle\varepsilon}\downarrow\wr \\ (K^m)^* & \xrightarrow{L_A^*} & (K^n)^*. \end{array}$$

Korollar 4.2.49 (Rang der transponierten Matrix)**.** *Seien $m, n \in \mathbb{N}$ und sei $A \in M_{m\times n}(K)$. Dann gilt*

$$\operatorname{rg} A^\intercal = \operatorname{rg} A,$$

das heißt, nach Bemerkung 4.2.36,

$$\dim_K \operatorname{ZR}(A) = \dim_K \operatorname{SR}(A).$$

Beweis. Wir wenden Proposition 4.2.47 mit der Abbildung $L_A\colon K^n \to K^m$ an. Die Matrix $A^\intercal$ ist also die Darstellungsmatrix der dualen Abbildung L_A^* bzgl. geeigneter Basen. Aus der Bemerkung 4.2.40 und dem Korollar 4.1.58 folgt

$$\operatorname{rg} A^\intercal = \operatorname{rg} L_A^* = \operatorname{rg} L_A = \operatorname{rg} A.$$ □

Kapitel 5

Lineare Gleichungen

Es gibt viele Arten von Gleichungen in der Mathematik: Polynomgleichungen, Diophantische Gleichungen, gewöhnliche und partielle Differentialgleichungen, usw. Ganz allgemein gesagt ist eine Gleichung ein Ausdruck der Gestalt $f(x) = g(x)$, wobei $f, g\colon X \to Y$ zwei Abbildungen mit derselben Definitions- und Zielmenge sind. Eine solche Gleichung zu *lösen* bedeutet, alle Elemente $x \in X$ zu finden, für die $f(x) = g(x)$ gilt. In den meisten Fällen ist g eine konstante Abbildung, d.h., die Gleichung hat die Form $f(x) = b$ mit einem $b \in Y$, und dann suchen wir alle Urbilder von b unter f (wenn Y eine Gruppe ist, ist das keine Beschränkung der Allgemeinheit, da jede Gleichung $f(x) = g(x)$ als $f(x)g(x)^{-1} = e$ umgeschrieben werden kann). Eine *lineare* Gleichung ist eine Gleichung $f(x) = b$, wobei f eine lineare Abbildung ist. Beispiele davon sind lineare Differentialgleichungen (siehe Beispiel 4.1.31).

In diesem Kapitel betrachten wir den konkreten Spezialfall, in dem f eine lineare Abbildung von K^n nach K^m ist, mit K einem beliebigen Körper. Das heißt, wir betrachten Gleichungen der Gestalt $A \cdot x = b$, wobei A eine $m \times n$-Matrix über K ist. Man kann solche Gleichungen als Systeme von m Gleichungen mit n Unbekannten $x_1, \ldots, x_n \in K$ auffassen. Wir werden insbesondere einen Algorithmus lernen, das *Gaußsche Eliminationsverfahren*, mit dem man solche Gleichungssysteme systematisch und vollständig lösen kann. Dieser Algorithmus hat viele weitere Anwendungen in der linearen Algebra und auch außerhalb der Mathematik, da viele praktische Probleme durch lineare Gleichungssysteme modelliert werden können. Mit ihm kann man auch effektiv bestimmen, ob eine quadratische Matrix invertierbar ist, und falls ja, ihre inverse Matrix berechnen.

5.1 Lineare Gleichungssysteme

Definition 5.1.1 (lineares Gleichungssystem, Lösungsmenge)**.** Seien $m, n \in \mathbb{N}$. Ein *lineares Gleichungssystem* über K mit m Gleichungen und n Unbekannten (oder Variablen) ist ein Paar (A, b) bestehend aus einer Matrix $A \in M_{m\times n}(K)$ und einem Spaltenvektor $b \in K^m$. Das Gleichungssystem (A, b) heißt *homogen*, wenn $b = 0$.

Eine *Lösung* von (A, b) ist ein Spaltenvektor $x \in K^n$, so dass $A \cdot x = b$. Die *Lösungsmenge* von (A, b) ist die Menge aller Lösungen von (A, b); sie wird mit $\mathcal{L}(A, b)$ bezeichnet:

$$\mathcal{L}(A, b) = \{x \in K^n \mid A \cdot x = b\} \subset K^n.$$

Eine Lösung von (A, b) besteht konkreter aus n Skalaren $x_1, \ldots, x_n \in K$, die gleichzeitig

die folgenden m Gleichungen erfüllen:

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

Bemerkung 5.1.2. Es gilt $\mathcal{L}(A,b) = L_A^{-1}(\{b\})$, wobei $L_A\colon K^n \to K^m$ die der Matrix A zugeordnete lineare Abbildung ist. Insbesondere ist die homogene Lösungsmenge

$$\mathcal{L}(A,0) = \ker L_A$$

ein Untervektorraum von K^n (nach Proposition 4.1.28), und es gilt

$$\mathcal{L}(A,b) \neq \varnothing \iff b \in \operatorname{im} L_A.$$

Proposition 5.1.3 (Struktur der Lösungsmenge eines linearen Gleichungssystems)**.** *Seien $m,n \in \mathbb{N}$, $A \in M_{m\times n}(K)$ und $b \in K^m$. Dann ist $\mathcal{L}(A,b)$ entweder leer oder eine Verschiebung des Untervektorraums $\mathcal{L}(A,0)$. Genauer: Für jede Lösung $x \in \mathcal{L}(A,b)$ gilt*

$$\mathcal{L}(A,b) = x + \mathcal{L}(A,0).$$

Außerdem gilt $\dim_K \mathcal{L}(A,0) = n - r$, wobei $r = \operatorname{rg} A$.

Beweis. Zu $\supset$. Sei $v \in \mathcal{L}(A,0)$. Dann gilt

$$A \cdot (x+v) = A \cdot x + A \cdot v = b + 0 = b,$$

und damit ist $x + v \in \mathcal{L}(A,b)$.

Zu $\subset$. Sei $v \in \mathcal{L}(A,b)$. Dann ist $v = x + (v - x)$, und es bleibt zu zeigen, dass $v - x \in \mathcal{L}(A,0)$:

$$A \cdot (x - v) = A \cdot x - A \cdot v = b - b = 0. \qquad \square$$

Nach Definition ist $\operatorname{rg} A = \operatorname{rg} L_A = \dim_K(\operatorname{im} L_A)$. Da $\mathcal{L}(A,0) = \ker L_A$ folgt die letzte Aussage aus der Dimensionsformel für die lineare Abbildung L_A (Korollar 4.1.38).

Um ein lineares Gleichungssystem (A,b) explizit zu lösen, soll man zuerst bestimmen, ob eine Lösung existiert. Wenn ja, genügt es nach Proposition 5.1.3 eine besondere Lösung $v_0 \in \mathcal{L}(A,b)$ und eine Basis $(v_1,\ldots,v_{n-r})$ von $\mathcal{L}(A,0)$ zu finden. Die allgemeine Lösung von $A \cdot x = b$ ist dann

$$x = v_0 + \sum_{i=1}^{n-r} \lambda_i v_i$$

mit beliebigen Skalaren $\lambda_i \in K$.

5.1.1 Zeilenstufenform

Es gibt besondere Matrizen A, die in sogenannter *Zeilenstufenform*, für die man die Lösungsmengen aller Gleichungssysteme (A,b) einfach bestimmen kann.

Definition 5.1.4 (Zeilenstufenform, Pivotelemente, Pivotspalten, reduzierte Zeilenstufenform)**.** Seien $m,n \in \mathbb{N}$ und sei $A = (a_{ij})_{i,j}$ eine $m \times n$-Matrix über K. Man sagt, dass A in *Zeilenstufenform* ist, wenn es Indizes $r \in \{0,\ldots,m\}$ und $1 \leq k_1 < \ldots < k_r \leq n$ gibt, so dass:

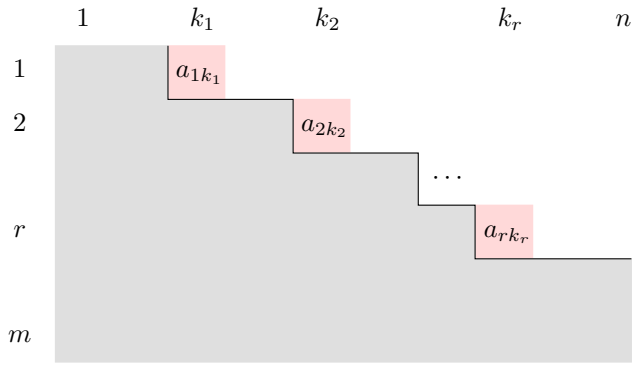
- $a_{ij} = 0$ für alle $i \leq r$ und alle $j < k_i$.

- $a_{ik_i} \neq 0$ für alle $i \leq r$.
- $a_{ij} = 0$ für alle $i > r$ und alle j.

Die r Koeffizienten $a_{1k_1}, \ldots, a_{rk_r}$ heißen die *Pivotelemente* von A, und ihre Spalten A_{*k_1}, $\ldots$, A_{*k_r} sind die *Pivotspalten* von A.

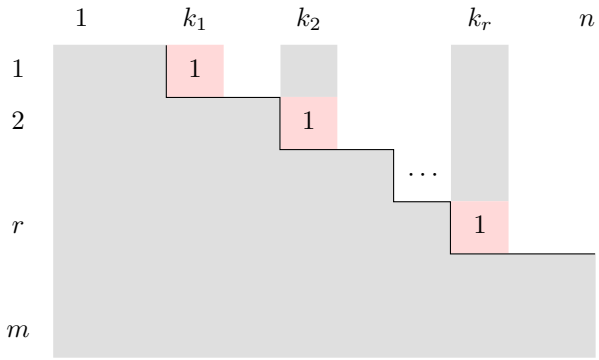
Man sagt, dass A in *reduzierter Zeilenstufenform* ist, wenn außerdem $A_{*k_i} = e_i$ für alle $i \leq r$. In einer Matrix in reduzierter Zeilenstufenform sind insbesondere alle Pivotelemente gleich 1.

Eine $m \times n$-Matrix A in Zeilenstufenform sieht wie folgt aus:



Das graue Gebiet besteht nur aus Nullen, die rote Einträge sind die Pivotelemente und das weisse Gebiet enthält beliebige Elemente aus K. Alle Spalten A_{*j} mit $j < k_1$, sowie alle Zeilen A_{i*} mit $i > r$, sind null (es kann aber sein, dass $k_1 = 1$ oder $r = m$).

Wenn A in *reduzierter* Zeilenstufenform ist, dann ist die i-te Pivotspalte A_{*k_i} der Standardeinheitsvektor e_i, d.h., alle Pivotelemente sind gleich 1 und alle darüberliegenden Koeffizienten sind null:



Proposition 5.1.5 (Zeilen- und Spaltenraum einer Matrix in Zeilenstufenform)**.** *Sei A eine $m \times n$-Matrix über K in Zeilenstufenform, und seien $k_1 < \cdots < k_r$ die Indizes der Pivotspalten mit $r \in \{0, \ldots, m\}$. Dann:*

(i) *Die ersten r Zeilen von A bilden eine Basis von* $\mathrm{ZR}(A)$.

(ii) *Es gilt* $\mathrm{SR}(A) = K^r \times \{0\} \subset K^m$ *und die Pivotspalten von A bilden eine Basis von* $\mathrm{SR}(A)$. *Insbesondere gilt* $\operatorname{rg} A = r$.

Beweis. Zu (i). Der Zeilenraum ist von den ersten r Zeilen erzeugt, weil alle anderen Zeilen null sind. Wegen der Zeilenstufenform ist es klar, dass die ersten r Zeilen linear unabhängig sind. Sie bilden also eine Basis von $\mathrm{ZR}(A)$.

Zu (ii). Da jede Spalte von A im Untervektorraum $K^r \times \{0\}$ liegt, gilt $\mathrm{SR}(A) \subset K^r \times \{0\}$. Wegen der Zeilenstufenform ist es aber klar, dass die r Pivotspalten linear unabhängig sind, so dass $\dim \mathrm{SR}(A) \geq r$. Aus Proposition 3.3.35 folgt, dass $\mathrm{SR}(A) = K^r \times \{0\}$. □

Rezept 5.1.6 (Lösung eines linearen Gleichungssystems in Zeilenstufenform)**.** Sei (A, b) ein lineares Gleichungssystem über K mit m Gleichungen und n Unbekannten, wobei A in Zeilenstufenform ist. Seien $k_1 < \cdots < k_r$ die Indizes der Pivotspalten von A. Dann kann die Lösungsmenge $\mathcal{L}(A, b)$ wie folgt bestimmt werden:

- Falls es einen Index $i > r$ mit $b_i \neq 0$ gibt, dann ist $\mathcal{L}(A, b) = \varnothing$. Denn die i-te Gleichung des Systems ist $0 \cdot x = b_i$ und ist nicht mit $x \in K^n$ lösbar.
- Andernfalls ist $\mathcal{L}(A, b)$ nicht leer, und man kann die allgemeine Lösung $x \in K^n$ wie folgt bestimmen:
  - Für die Unbekannten x_{k_i} mit $i \in \{1, \ldots, r\}$ gilt
$$x_{k_i} = \frac{1}{a_{ik_i}} \left(b_{k_i} - \sum_{j=k_i+1}^{n} a_{ij} x_j \right).$$
  - Es gibt keine weitere Bedingungen, d.h., die Unbekannten x_j mit $j \notin \{k_1, \ldots, k_r\}$ sind beliebig. Diese Unbekannten heißen die *freien Variablen*.
  - Durch *Rückwärtssubstitution* kann man die x_{k_i} nur durch die freien Variablen ausdrücken (dieser Schritt ist nicht nötig, wenn die Matrix A in *reduzierter* Zeilenstufenform ist).

  Dann erhalten wir die allgemeine Lösung in der Form
$$x = v_0 + \sum_{j \notin \{k_1, \ldots, k_r\}} x_j v_j, \quad x_j \text{ beliebig},$$
  mit geeigneten Vektoren $v_0, v_j \in K^n$. Dabei ist also v_0 eine besondere Lösung von (A, b) und ist $(v_j)_{j \notin \{k_1, \ldots, k_r\}}$ eine Basis von $\mathcal{L}(A, 0)$.

Beispiel 5.1.7.

(i) Seien
$$A = \begin{pmatrix} 2 & -1 & 4 & 0 \\ 0 & 3 & 1 & -3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{und} \quad b = \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix}.$$
Wenn $\mathrm{char}(K) \notin \{2, 3\}$ ist die Matrix A in Zeilenstufenform mit drei Pivotspalten. Für das System $A \cdot x = b$ erhalten wir
$$\begin{aligned} x_1 &= \frac{1}{2}(x_2 - 4x_3) \\ x_2 &= \frac{1}{3}(3 - x_3 + 3x_4) \\ x_4 &= 1, \end{aligned}$$
und die Variable x_3 ist frei. Durch Rückwärtssubstitution erhalten wir
$$\begin{aligned} x_4 &= 1, \\ x_2 &= \frac{1}{3}(3 - x_3 + 3) = 2 - \frac{1}{3}x_3, \\ x_1 &= \frac{1}{2}\left(2 - \frac{1}{3}x_3 - 4x_3\right) = 1 - \frac{13}{6}x_3, \end{aligned}$$
was die allgemeine Lösung des Gleichungssystems ergibt:
$$x = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 1 \end{pmatrix} + x_3 \begin{pmatrix} -\frac{13}{6} \\ -\frac{1}{3} \\ 1 \\ 0 \end{pmatrix}, \quad x_3 \text{ beliebig}.$$

(ii) Seien

$$A = \begin{pmatrix} 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad c = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}.$$

Die Matrix A ist in reduzierter Zeilenstufenform mit zwei Pivotspalten. Das System $A \cdot x = b$ hat keine Lösung, weil $b_3 \neq 0$. Für das System $A \cdot x = c$ erhalten wir

$$\begin{aligned} x_2 &= 2 - 3x_3, \\ x_4 &= 1, \end{aligned}$$

und die Variablen x_1 und x_3 sind frei. Da A in reduzierter Zeilenstufenform war, braucht man hier keine Rüchwärtssubstitution durchzuführen. Die allgemeine Lösung ist also

$$x = \begin{pmatrix} 0 \\ 2 \\ 0 \\ 1 \end{pmatrix} + x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ -3 \\ 1 \\ 0 \end{pmatrix}, \quad x_1, x_3 \text{ beliebig.}$$

5.2 Das Gaußsche Eliminationsverfahren

Das Gaußsche Eliminationsverfahren ist ein Algorithmus, der beliebige lineare Gleichungssysteme löst. Dieser Algorithmus ist sowohl von theoretischer als auch von praktischer Bedeutung: Er ist sehr effizient und wird tatsächlich von Computern zur Lösung großer linearer Gleichungssysteme verwendet. Andererseits werden wir diesen Algorithmus verwenden, um weitere Sätze zu beweisen.

Die Strategie zur Lösung eines beliebigen linearen Gleichungsystems ist die folgende:

- Wie bereits im Abschnitt 5.1.1 erklärt, gibt es besondere Matrizen A, die in sogenannter *Zeilenstufenform*, für die man die Lösungsmengen aller Gleichungssysteme (A, b) einfach bestimmen kann.
- Zu jedem linearen Gleichungssystem (A, b) kann man ein lineares Gleichungssystem (A', b') mit derselben Lösungsmenge effektiv finden, durch sogenannte *elementare Zeilenumformungen*, wobei A' in Zeilenstufenform ist.

Definition 5.2.1 (Zeilenumformung)**.** Seien $m, n \in \mathbb{N}$. Eine *Zeilenumformung* oder *Zeilenoperation* auf $m \times n$-Matrizen ist eine Abbildung der Gestalt

$$\begin{aligned} M_{m\times n}(K) &\to M_{m\times n}(K), \\ A &\mapsto Z \cdot A, \end{aligned}$$

wobei Z eine invertierbare $m \times m$-Matrix ist.

Nach Definition der Matrixmultiplikation ergibt sich die umgeformte Matrix $Z \cdot A$ aus A, indem man jede Zeile durch eine Linearkombination der Zeilen ersetzt: Für alle $k \in \{1, \ldots, m\}$ gilt

$$(Z \cdot A)_{k*} = \sum_{i=1}^{m} Z_{ki} \cdot A_{i*}.$$

Da Z invertierbar ist, ist außerdem die entsprechende Zeilenumformung $A \mapsto Z \cdot A$ bijektiv, mit Umkehrabbildung $A \mapsto Z^{-1} \cdot A$.

Proposition 5.2.2 (Zeilen- und Spaltenraum bei Zeilenumformungen)**.** *Sei A eine $m \times n$-Matrix über K und sei $Z \in \mathrm{GL}_m(K)$.*

(i) *Es gilt* $\mathrm{ZR}(Z \cdot A) = \mathrm{ZR}(A)$.

(ii) *Es gilt* $\mathrm{SR}(Z \cdot A) = L_Z(\mathrm{SR}(A))$ *und insbesondere* $\mathrm{SR}(Z \cdot A) \cong \mathrm{SR}(A)$.

Beweis. Zu (i). Jede Zeile von $Z \cdot A$ ist eine Linearkombination der Zeilen von A und damit liegt im Zeilenraum von A. Dies zeigt, dass $\mathrm{ZR}(Z \cdot A) \subset \mathrm{ZR}(A)$. Umgekehrt gilt $\mathrm{ZR}(A) = \mathrm{ZR}(Z^{-1} \cdot Z \cdot A) \subset \mathrm{ZR}(Z \cdot A)$.

Zu (ii). Mit der Bemerkung 4.2.27 und der Proposition 4.2.26(ii) berechnen wir

$$\mathrm{SR}(Z \cdot A) = \operatorname{im} L_{Z \cdot A} = \operatorname{im}(L_Z \circ L_A) = L_Z(\operatorname{im} L_A) = L_Z(\mathrm{SR}(A)).$$

Da Z invertierbar ist, ist L_Z nach Proposition 4.2.37 ein Isomorphismus und damit gilt $L_Z(\mathrm{SR}(A)) \cong \mathrm{SR}(A)$. □

Proposition 5.2.3 (Invarianz der Lösungsmenge bei Zeilenumformungen)**.** *Sei* (A, b) *ein lineares Gleichungssystem über* K *mit* m *Gleichungen und* n *Unbekannten, und sei* $Z \in \mathrm{GL}_m(K)$*. Dann gilt*

$$\mathcal{L}(A, b) = \mathcal{L}(Z \cdot A, Z \cdot b).$$

Beweis. Sei $x \in \mathcal{L}(A, b)$, d.h., $A \cdot x = b$. Daraus folgt, dass $Z \cdot A \cdot x = Z \cdot b$, d.h., $x \in \mathcal{L}(Z \cdot A, Z \cdot b)$. Sei umgekehrt $x \in \mathcal{L}(Z \cdot A, Z \cdot b)$, d.h., $Z \cdot A \cdot x = Z \cdot b$. Dann gilt

$$A \cdot x = Z^{-1} \cdot Z \cdot A \cdot x = Z^{-1} \cdot Z \cdot b = b,$$

d.h., $x \in \mathcal{L}(A, b)$. □

Zur Erinnerung ist E_{rs} die Matrix mit $(E_{rs})_{ij} = \delta_{ir}\delta_{js}$.

Definition 5.2.4 (Elementarmatrizen, elementare Zeilenumformung)**.** Sei $m \in \mathbb{N}$. Die folgenden $m \times m$-Matrizen über K heißen *Elementarmatrizen*:

- Seien $i, j \in \{1, \ldots, m\}$ mit $i < j$. Die Matrix $V_{ij} \in M_m(K)$ ist
$$V_{ij} := I_m - E_{ii} - E_{jj} + E_{ij} + E_{ji}.$$
- Seien $i \in \{1, \ldots, m\}$ und $\lambda \in K^*$. Die Matrix $M_i(\lambda) \in M_m(K)$ ist
$$M_i(\lambda) := I_m - E_{ii} + \lambda \cdot E_{ii}.$$
- Seien $i, j \in \{1, \ldots, m\}$ mit $i \neq j$ und sei $\alpha \in K$. Die Matrix $A_{ij}(\alpha) \in M_m(K)$ ist
$$A_{ij}(\alpha) := I_m + \alpha \cdot E_{ij}.$$

Eine Zeilenumformung $A \mapsto Z \cdot A$ mit Z einer Elementarmatrix heißt *elementare Zeilenumformung.*

Bei dieser Definition soll man nachprüfen, dass Elementarmatrizen invertierbar sind. Ist Z eine Elementarmatrix, so gilt folgendes für die Zeilen von $Z \cdot A$:

$$(V_{ij} \cdot A)_{k*} = \begin{cases} A_{j*}, & \text{falls } k = i, \\ A_{i*}, & \text{falls } k = j, \\ A_{k*}, & \text{andernfalls,} \end{cases}$$

$$(M_i(\lambda) \cdot A)_{k*} = \begin{cases} \lambda \cdot A_{i*}, & \text{falls } k = i, \\ A_{k*}, & \text{andernfalls,} \end{cases}$$

$$(A_{ij}(\alpha) \cdot A)_{k*} = \begin{cases} A_{i*} + \alpha \cdot A_{j*}, & \text{falls } k = i, \\ A_{k*}, & \text{andernfalls.} \end{cases}$$

Es folgt daraus (und aus der Proposition 4.2.37 (i) $\Leftrightarrow$ (iii)), dass Elementarmatrizen invertierbar sind, mit Inversen

$$V_{ij}^{-1} = V_{ij}, \quad M_i(\lambda)^{-1} = M_i(\lambda^{-1}), \quad A_{ij}(\alpha)^{-1} = A_{ij}(-\alpha).$$

Diese Beobachtungen sind in folgender Tabelle zusammengefasst:

Matrix Z	Zeilenumformung $A \mapsto Z \cdot A$	inverse Matrix Z^{-1}
V_{ij}	Vertauschung der i-ten und j-ten Zeilen	V_{ij}
$M_i(\lambda)$	Multiplikation der i-ten Zeile mit λ	$M_i(\lambda^{-1})$
$A_{ij}(\alpha)$	Addition des α-fachen der j-ten Zeile zur i-ten Zeile	$A_{ij}(-\alpha)$

Bemerkung 5.2.5. Multiplikation mit einer invertierbaren Matrix von *rechts* ergibt eine *Spaltenumformung*, indem jede Spalte durch eine Linearkombination der Spalten ersetzt wird. Elementarmatrizen liefern dabei *elementare Spaltenumformungen*. Da $A \cdot Z = (Z^\intercal \cdot A^\intercal)^\intercal$ und die transponierte Matrix zu einer Elementarmatrix wieder eine Elementarmatrix ist, sind elementare Spaltenumformungen die offensichtlichen Entsprechungen von elementaren Zeilenumformungen.

Satz 5.2.6 (Gaußsches Eliminationsverfahren)**.** *Seien $m, n \in \mathbb{N}$. Jede $m \times n$-Matrix A über K kann durch elementare Zeilenumformungen auf reduzierte Zeilenstufenform gebracht werden. Das heißt, es existiert Elementarmatrizen $Z_1, \ldots, Z_k$, so dass die Matrix $Z_k \cdot \ldots \cdot Z_1 \cdot A$ in reduzierter Zeilenstufenform ist.*

Das Gaußsche Eliminationsverfahren ist sogar ein Algorithmus, das jede Matrix auf reduzierte Zeilenstufenform bringt. Um diesen Algorithmus zu verstehen, ist es hilfreich, ein Bild einer Matrix in Zeilenstufenform parat zu haben. Wir erklären zunächst, wie man eine Matrix durch elementare Zeilenumformungen auf Zeilenstufenform bringen kann:

Algorithmus 5.2.7 (Gaußsches Eliminationsverfahren I: Eine Matrix auf Zeilenstufenform bringen)**.** Sei $A = (a_{ij})_{i,j}$ eine $m \times n$-Matrix über K. Man zeigt durch Induktion über $j \in \{0, \ldots, n\}$, wie man die ersten j Spalten von A auf Zeilenstufenform bringen kann. Wenn $j = 0$ gibt es nichts zu tun. Sei also $j \in \{1, \ldots, n\}$. Angenommen sind die ersten $j-1$ Spalten von A schon in Zeilenstufenform. Sei $s - 1$ die Anzahl der bisherigen Pivotspalten ($s \geq 1$). Falls $a_{ij} = 0$ für alle $i \geq s$, dann sind die ersten j Spalten von A schon in Zeilenstufenform. Andernfalls entspricht j einer neuen Pivotspalte.

(i) Man wähle $i \geq s$ mit $a_{ij} \neq 0$ und man vertausche die i-te une s-te Zeilen, so dass $a_{sj} \neq 0$. Das Element a_{sj} ist das s-te Pivotelement.

(ii) Für alle $i > s$ verwende man die elementare Zeilenumformung $A_{is}(-a_{ij}/a_{sj})$, um a_{ij} zu null zu machen. Die ersten j Spalten von A sind jetzt in Zeilenstufenform.

Wenn eine Matrix in Zeilenstufenform ist, kann man sie weiter folgendermaßen auf reduzierte Zeilenstufenform bringen:

Algorithmus 5.2.8 (Gaußsches Eliminationsverfahren II: Eine Matrix in Zeilenstufenform auf reduzierte Zeilenstufenform bringen)**.** Sei $A = (a_{ij})_{i,j}$ eine $m \times n$-Matrix über K in Zeilenstufenform. Seien $1 \leq k_1 < \cdots < k_r \leq n$ die Indizes der Pivotspalten von A. Für $i = r, r-1, \ldots, 1$ tue man nacheinander folgendes:

(i) Man multipliziere die i-te Zeile mit $1/a_{ik_i}$, um das Pivotelement a_{ik_i} zu 1 zu machen.

(ii) Für alle $e < i$ verwende man die elementare Zeilenumformung $A_{ei}(-a_{ek_i})$, um a_{ek_i} zu null zu machen.

Rezept 5.2.9 (Lösung eines beliebigen linearen Gleichungssystems)**.** Sei (A, b) ein lineares Gleichungssystem über K. Um seine Lösungsmenge zu bestimmen, betrachten wir die *erweiterte Matrix* $(A \,|\, b)$, indem wir den Spaltenvektor b am Ende der Matrix A hinzufügen. Mit dem Gaußschen Eliminationsverfahren können wir die Matrix $(A \,|\, b)$ durch elementare Zeilenumformungen in eine Matrix $(A' \,|\, b')$ bringen, wobei A' in (reduzierter) Zeilenstufenform ist. Das heißt, es gilt $A' = Z_k \cdot \ldots \cdot Z_1 \cdot A$ und $b' = Z_k \cdot \ldots \cdot Z_1 \cdot b$ mit geeigneten Elementarmatrizen $Z_1, \ldots, Z_k$. Nach Proposition 5.2.3 gilt dann $\mathcal{L}(A, b) = \mathcal{L}(A', b')$, und mit dem Rezept 5.1.6 können wir diese Lösungsmenge völlig bestimmen.

Beispiel 5.2.10. Seien

$$A = \begin{pmatrix} 1 & 6 & 2 & -5 & -2 \\ -2 & -12 & -2 & 2 & 3 \\ 3 & 18 & 4 & -7 & -3 \end{pmatrix}, \quad b = \begin{pmatrix} -4 \\ 11 \\ -9 \end{pmatrix}, \quad c = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}.$$

Wir lösen die Gleichungssysteme $A \cdot x = b$ und $A \cdot x = c$ gleichzeitig durch das Gaußsche Eliminationsverfahren:

$$\left(\begin{array}{ccccc|c|c} 1 & 6 & 2 & -5 & -2 & -4 & 1 \\ -2 & -12 & -2 & 2 & 3 & 11 & 2 \\ 3 & 18 & 4 & -7 & -3 & -9 & 0 \end{array}\right) \xrightarrow[A_{31}(-3)]{A_{21}(2)} \left(\begin{array}{ccccc|c|c} 1 & 6 & 2 & -5 & -2 & -4 & 1 \\ 0 & 0 & 2 & -8 & -1 & 3 & 4 \\ 0 & 0 & -2 & 8 & 3 & 3 & -3 \end{array}\right)$$

$$\xrightarrow{A_{32}(1)} \left(\begin{array}{ccccc|c|c} 1 & 6 & 2 & -5 & -2 & -4 & 1 \\ 0 & 0 & 2 & -8 & -1 & 3 & 4 \\ 0 & 0 & 0 & 0 & 2 & 6 & 1 \end{array}\right).$$

Wenn $\mathrm{char}(K) \neq 2$ ist das System jetzt in Zeilenstufenform mit drei Pivotspalten, und die freien Variablen sind x_2 und x_4. Für das System $A \cdot x = b$ erhalten wir durch Rückwärtssubstitution

$$\begin{aligned} x_5 &= 3, \\ x_3 &= \frac{1}{2}(3 + 8x_4 + x_5) = 3 + 4x_4, \\ x_1 &= -4 - 6x_2 - 2x_3 + 5x_4 + 2x_5 = -4 - 6x_2 - 3x_4. \end{aligned}$$

Die allgemeine Lösung von (A, b) ist also

$$x = \begin{pmatrix} -4 \\ 0 \\ 3 \\ 0 \\ 3 \end{pmatrix} + x_2 \begin{pmatrix} -6 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -3 \\ 0 \\ 4 \\ 1 \\ 0 \end{pmatrix},$$

mit $x_2, x_4 \in K$ beliebig. Die zwei letzten Vektoren bilden also eine Basis von $\mathcal{L}(A, 0)$. Für das System $A \cdot x = c$ erhalten wir durch Rückwärtssubstitution

$$\begin{aligned} x_5 &= \frac{1}{2}, \\ x_3 &= \frac{1}{2}(4 + 8x_4 + x_5) = \frac{9}{4} + 4x_4, \\ x_1 &= 1 - 6x_2 - 2x_3 + 5x_4 + 2x_5 = -\frac{5}{2} - 6x_2 - 3x_4. \end{aligned}$$

Die allgemeine Lösung von (A, c) ist also

$$x = \frac{1}{4} \begin{pmatrix} -10 \\ 0 \\ 9 \\ 0 \\ 2 \end{pmatrix} + x_2 \begin{pmatrix} -6 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -3 \\ 0 \\ 4 \\ 1 \\ 0 \end{pmatrix},$$

mit $x_2, x_4 \in K$ beliebig.

Wenn $\mathrm{char}(K) = 2$ haben wir das System

$$\left(\begin{array}{ccccc|c|c} 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right),$$

das bereits in reduzierter Zeilenstufenform ist. Jetzt gibt es zwei Pivotspalten und die freien Variablen sind x_2, x_3, x_4. Für das System (A, b) gilt $x_5 = 1$ und $x_1 = -x_4 = x_4$. Aber für das System (A, c) gilt $\mathcal{L}(A, c) = \varnothing$ wegen der dritten Gleichung $0 = 1$.

Das Gaußsche Eliminationsverfahren liefert einen alternativen, konkreteren Beweis der Tatsache, dass der Zeilenraum und der Spaltenraum einer beliebigen $m \times n$-Matrix immer die gleiche Dimension haben (Korollar 4.2.49):

Alternativer Beweis des Korollars 4.2.49. Sei A' eine Zeilenstufenform von A, d.h., A' ist in Zeilenstufenform und wird aus A durch elementare Zeilenumformungen erhalten. Nach Proposition 5.2.2 gelten

$$\dim_K \mathrm{ZR}(A) = \dim_K \mathrm{ZR}(A'),$$
$$\dim_K \mathrm{SR}(A) = \dim_K \mathrm{SR}(A').$$

Nach Proposition 5.1.5 gilt

$$\dim_K \mathrm{ZR}(A') = \dim_K \mathrm{SR}(A').$$

Aus diesen drei Gleichungen folgt, dass $\dim_K \mathrm{ZR}(A) = \dim_K \mathrm{SR}(A)$. □

Korollar 5.2.11. *Seien $n \in \mathbb{N}$ und $A \in M_n(K)$. Die folgenden Aussagen sind äquivalent:*

(i) *A ist invertierbar.*

(ii) *A ist ein Produkt von Elementarmatrizen.*

Insbesondere ist jede Zeilenumformung eine Komposition elementarer Zeilenumformungen.

Beweis. Die Implikation (ii) ⇒ (i) folgt daraus, dass Elementarmatrizen invertierbar sind. Sei umgekehrt A invertierbar, d.h., $\mathrm{rg}\, A = n$ (Proposition 4.2.37). Nach Satz 5.2.6 existieren Elementarmatrizen $Z_1, \ldots, Z_k$, so dass $A' = Z_k \cdot \ldots \cdot Z_1 \cdot A$ in reduzierter Zeilenstudenform ist. Nach Proposition 5.2.2(ii) ist $\mathrm{rg}\, A' = \mathrm{rg}\, A = n$, und damit hat A' n Pivotspalten (Proposition 5.1.5). Aber die einzige $n \times n$-Matrix in reduzierter Zeilenstufenform mit n Pivotspalten ist die Einheitsmatrix I_n. Daraus folgt, dass $A = Z_1^{-1} \cdot \ldots \cdot Z_k^{-1}$. □

Ein theoretischer Vorteil der *reduzierten* Zeilenstufenform gegenüber der Zeilenstufenform ist die folgende Eindeutigkeitsaussage:

Proposition 5.2.12 (Eindeutigkeit der reduzierten Zeilenstufenform)**.** *Seien $m, n \in \mathbb{N}$ und $A \in M_{m\times n}(K)$. Seien $Z, Z' \in \mathrm{GL}_m(K)$, so dass die Matrizen $B = Z \cdot A$ und $B' = Z' \cdot A$ in reduzierter Zeilenstufenform sind. Dann ist $B = B'$.*

**Beweis.* Nach Proposition 5.2.2(ii) haben B und B' denselben Rang. Es genügt also die folgende Aussage zu beweisen: Sind B und B' $m\times n$-Matrizen in reduzierter Zeilenstufenform mit demselben Rang r, so dass jede Zeile von B' eine Linearkombination der Zeilen von B ist, so gilt $B = B'$. Seien $k_1 < \cdots < k_r$ bzw. $k'_1 < \cdots < k'_r$ die Indizes der Pivotspalten von B bzw. B'. Wir beweisen die Aussage durch Induktion über r. Wenn $r = 0$ sind beide B und B' null. Sei $r \geq 1$ und seien C und C' die $(m-1) \times n$-Matrizen, die aus B und B' entstehen, wenn man die ersten Zeilen streicht. Die Matrizen C und C' sind wieder in reduzierter Zeilenstufenform und haben Rang $r-1$. Da die Zeile B'_{1*} eine Linearkombination der Zeilen B_{i*} ist und die Spalten B_{*j} für alle $j < k_1$ null sind, muss $k'_1 \geq k_1$ sein. Für jedes $i \in \{2, \ldots, m\}$ gilt insbesondere $B'_{ik_1} = 0$. In einer Linearkombination der Zeilen B_{1*}, …, B_{m*}, die B'_{i*} ergibt, muss deswegen B_{1*} mit dem Koeffizient 0 vorkommen. Das heißt, jede Zeile von C' ist eine Linearkombination der Zeilen von C. Nach der Induktionsvoraussetzung ist $C = C'$; insbesondere ist $k_e = k'_e$ für alle $e \in \{2, \ldots, r\}$. Es bleibt zu zeigen, dass $B_{1*} = B'_{1*}$. Sei $B'_{1*} = \sum_{e=1}^{r} \lambda_e B_{e*}$. Für alle $e \geq 2$ gilt $B'_{1k_e} = 0$ aber $B_{ek_e} = 1$, und deswegen muss $\lambda_e = 0$ sein. Das heißt, es gilt $B'_{1*} = \lambda_1 B_{1*}$, und die einzige Möglichkeit ist dann $\lambda_1 = 1$. □

Bemerkung 5.2.13. Seien $A, B \in M_{m\times n}(K)$. Man sagt, dass A *zeilenäquivalent* zu B ist, wenn eine Matrix $Z \in \mathrm{GL}_m(K)$ mit $Z \cdot A = B$ existiert. Es ist klar, dass Zeilenäquivalenz eine Äquivalenzrelation auf $M_{m\times n}(K)$ ist. Das Gaußsche Eliminationsverfahren und die Proposition 5.2.12 implizieren, dass jede Äquivalenzklasse *genau eine* Matrix in reduzierter Zeilenstufenform enthält. Insbesondere können wir durch das Gaußsche Eliminationsverfahren effektiv bestimmen, ob zwei Matrizen zeilenäquivalent sind, indem wir beide Matrizen in reduzierte Zeilenstufenform bringen.

5.2.1 Rezepte

Mit dem Gaußschen Eliminationsverfahren können wir jetzt beliebige lineare Gleichungssysteme effektiv lösen (siehe Rezept 5.2.9). Aber dieser Algorithmus hilft auch bei vielen anderen Problemen in der linearen Algebra. In diesem Abschnitt erklären wir einige solcher weiteren Anwendungen.

Alle nachfolgenden Rezepte werden mit Vektorräumen der Gestalt K^n und Matrizen formuliert. Aber diese Rezepte können auch auf beliebige endlich-dimensionale Vektorräume und lineare Abbildungen zwischen denen angewendet werden, sofern Basen dieser Vektorräume bekannt sind: Dann kann das Problem auf Vektorräume der Gestalt K^n übertragen werden.

Rezept 5.2.14 (Berechnung des Rangs)**.** Gegeben sei eine Matrix $A \in M_{m\times n}(K)$. Gesucht ist der Rang von A. Man überführt A mit dem Gaußschen Eliminationsverfahren in Zeilenstufenform. Nach Propositionen 5.2.2(ii) und 5.1.5(ii) ist die Anzahl der Pivotspalten gleich dem Rang von A

Rezept 5.2.15 (Berechnung des Kerns)**.** Gegeben sei eine Matrix $A \in M_{m\times n}(K)$. Gesucht ist eine Basis von $\ker L_A$. Da $\ker L_A = \mathcal{L}(A, 0)$, ist das ein Sonderfall vom Rezept 5.2.9.

Rezept 5.2.16 (Berechnung des Bildes)**.** Gegeben sei eine Matrix $A \in M_{m\times n}(K)$. Gesucht ist eine Basis von $\operatorname{im} L_A$. Man bringt A auf Zeilenstufenform A' mit dem Gaußschen Eliminationsverfahren. Sind $k_1 < \cdots < k_r$ die Indizes der Pivotspalten von A', so ist $(A_{*k_1}, \ldots, A_{*k_r})$ eine Basis von $\operatorname{im} L_A$. Denn $A' = Z \cdot A$ mit einem $Z \in \mathrm{GL}_n(K)$, und die Pivotspalten $Z \cdot A_{*k_i}$ bilden eine Basis von $\mathrm{SR}(Z \cdot A)$ (Proposition 5.1.5(ii)). Nach Proposition 5.2.2(ii) bilden dann die Spalten A_{*k_i} eine Basis von $\mathrm{SR}(A) = \operatorname{im} L_A$.

Alternativ dazu führt man das Gaußsche Eliminationsverfahren mit der transponierten Matrix $A^\intercal$ durch, bis sie in Zeilenstufenform ist. Nach Propositionen 5.2.2(i) und 5.1.5(i) bilden jetzt die Nicht-Null-Zeilen eine Basis von $\mathrm{ZR}(A^\intercal) = \mathrm{SR}(A) = \operatorname{im} L_A$.

Rezept 5.2.17 (Test auf Invertierbarkeit)**.** Gegeben sei eine Matrix $A \in M_n(K)$. Zu bestimmen ist, ob A invertierbar ist. Man berechnet den Rang von A mit dem Rezept 5.2.14. Nach Proposition 4.2.37 ist A genau dann invertierbar, wenn $\operatorname{rg} A = n$.

Rezept 5.2.18 (Berechnung der inversen Matrix)**.** Gegeben sei eine Matrix $A \in M_n(K)$. Gesucht ist die inverse Matrix A^{-1}, wenn sie existiert. Man führt das Gaußsche Eliminationsverfahren mit der erweiterten Matrix $(A \mid I_n)$ durch, bis A in reduzierter Zeilenstufenform ist (sofern die Zeilenstufenform n Pivotspalten besitzt, sonst ist A nicht invertierbar). Die erweiterte Matrix hat jetzt die Form $(I_n \mid B)$, und es gilt $A^{-1} = B$. Denn es gibt ein $Z \in \mathrm{GL}_n(K)$ mit $Z \cdot A = I_n$ und $Z \cdot I_n = B$, und damit ist $B = Z = A^{-1}$.

Beispiel 5.2.19. Sei

$$A = \begin{pmatrix} 0 & 1 & 6 \\ 0 & 1 & 7 \\ 1 & 6 & 0 \end{pmatrix} \in M_3(K).$$

Wir führen das Gaußsche Eliminationsverfahren durch:

$$\left(\begin{array}{ccc|ccc} 0 & 1 & 6 & 1 & 0 & 0 \\ 0 & 1 & 7 & 0 & 1 & 0 \\ 1 & 6 & 0 & 0 & 0 & 1 \end{array}\right) \xrightarrow{V_{13}} \left(\begin{array}{ccc|ccc} 1 & 6 & 0 & 0 & 0 & 1 \\ 0 & 1 & 7 & 0 & 1 & 0 \\ 0 & 1 & 6 & 1 & 0 & 0 \end{array}\right)$$

$$\xrightarrow{A_{32}(-1)} \left(\begin{array}{ccc|ccc} 1 & 6 & 0 & 0 & 0 & 1 \\ 0 & 1 & 7 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 & -1 & 0 \end{array}\right).$$

Die linke Seite ist jetzt in Zeilenstufenform und hat Rang 3, so dass A invertierbar ist. Wir führen das Eliminationsverfahren weiter, bis die linke Seite die Einheitsmatrix wird:

$$\xrightarrow{M_3(-1)} \left(\begin{array}{ccc|ccc} 1 & 6 & 0 & 0 & 0 & 1 \\ 0 & 1 & 7 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 1 & 0 \end{array}\right) \xrightarrow[A_{12}(-6)]{A_{23}(-7)} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -42 & 36 & 1 \\ 0 & 1 & 0 & 7 & -6 & 0 \\ 0 & 0 & 1 & -1 & 1 & 0 \end{array}\right)$$

Nach Rezept 5.2.18 gilt also

$$A^{-1} = \begin{pmatrix} -42 & 36 & 1 \\ 7 & -6 & 0 \\ -1 & 1 & 0 \end{pmatrix}.$$

Man kann natürlich diese Antwort durch Multiplizieren mit A bestätigen.

Rezept 5.2.20 (Test auf Enthaltensein)**.** Gegeben seien eine Familie $(v_1, \dots, v_k)$ und ein Vektor v in K^n. Zu bestimmen ist, ob v in $\operatorname{Span}_K\{v_1, \dots, v_k\}$ enthalten ist. Das gilt genau dann, wenn das System (A, v) eine Lösung besitzt, wobei A die Matrix mit Spalten v_1, …, v_k ist. Man verwendet also das Rezept 5.2.9.

Rezept 5.2.21 (Test auf Erzeugendensystem)**.** Gegeben sei eine Familie $F = (v_1, \dots, v_k)$ in K^n. Zu bestimmen ist, ob F erzeugend ist. Sei A die $n \times k$-Matrix mit Spalten v_1, …, v_k. Man berechnet den Rang von A mit dem Rezept 5.2.14. Dann ist F genau dann erzeugend, wenn $\operatorname{rg} A = n$. Denn $\operatorname{Span}_K\{v_1, \dots, v_k\} = \operatorname{SR}(A)$ und $\dim_K \operatorname{SR}(A) = \operatorname{rg} A$.

Rezept 5.2.22 (Test auf lineare Unabhängigkeit)**.** Gegeben sei eine Familie $F = (v_1, \dots, v_k)$ in K^n. Zu bestimmen ist, ob F linear unabhängig ist. Sei A die $n \times k$-Matrix mit Spalten v_1, …, v_k. Man berechnet den Rang von A mit dem Rezept 5.2.14. Dann ist F genau dann linear unabhängig, wenn $\operatorname{rg} A = k$. Denn $\operatorname{rg} A = k - \dim_K \ker L_A$ nach der Dimensionsformel für die lineare Abbildung L_A, und $\ker L_A$ besteht aus allen k-Tupeln $(\lambda_1, \dots, \lambda_k)$ mit $\sum_{i=1}^k \lambda_i v_i = 0$.

Rezept 5.2.23 (Einschränkung zu einer Basis)**.** Gegeben sei eine Familie $(v_1, \dots, v_k)$ in K^n. Gesucht ist eine Teilfamilie, die eine Basis von $\operatorname{Span}_K\{v_1, \dots, v_k\}$ ist. Sei A die $n \times k$-Matrix mit Spalten v_1, …, v_k. Seien $k_1 < \cdots < k_r$ die Indizes der Pivotspalten einer Zeilenstufenform von A. Dann ist $(v_{k_1}, \dots, v_{k_r})$ eine Basis von $\operatorname{Span}_K\{v_1, \dots, v_k\}$. Denn eine Zeilenstufenform von A hat die Form $Z \cdot A$ mit einem $Z \in \operatorname{GL}_n(K)$, und ihre Pivotspalten $Z \cdot v_{k_i}$ bilden eine Basis von $\operatorname{SR}(Z \cdot A)$ (Proposition 5.1.5(ii)). Nach Proposition 5.2.2(ii) bilden dann die Spalten v_{k_i} eine Basis von $\operatorname{SR}(A)$.

Dieses Rezept hat die folgende zusätzliche Eigenschaft: War die Teilfamilie $(v_1, \dots, v_l)$ bereits linear unabhängig, so gilt $k_i = i$ für alle $i \le l$. Denn die Matrix $Z \cdot \begin{pmatrix} v_1 & \dots & v_l \end{pmatrix}$ ist in Zeilenstufenform und hat Rang l, und damit sind die ersten l Spalten von $Z \cdot A$ Pivotspalten.

Beispiel 5.2.24. Wir betrachten die Vektoren

$$v_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad v_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix}, \quad v_4 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 2 \end{pmatrix}, \quad v_5 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 2 \end{pmatrix}$$

in $\mathbb{F}_3^4$, und wir verwenden das Rezept 5.2.23, um eine Teilfamilie von $(v_1, \dots, v_5)$ zu finden, die eine Basis von $U = \mathrm{Span}_K\{v_1, \dots, v_5\}$ ist:

$$\begin{pmatrix} 2 & 0 & 1 & 0 & 1 \\ 1 & 0 & 2 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 2 & 2 \end{pmatrix} \xrightarrow{A_{21}(1)} \begin{pmatrix} 2 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 2 & 2 \end{pmatrix}$$
$$\xrightarrow[V_{23}]{A_{43}(-1)} \begin{pmatrix} 2 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix} \xrightarrow{A_{43}(-1)} \begin{pmatrix} 2 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Die Indizes der Pivotspalten sind 1, 2 und 4. Deshalb ist (v_1, v_2, v_4) eine Basis von U.

Rezept 5.2.25 (Ergänzung zu einer Basis)**.** Gegeben sei eine linear unabhängige Familie $(v_1, \dots, v_k)$ in K^n. Gesucht sind Vektoren v_{k+1}, …, v_n, so dass $(v_1, \dots, v_n)$ eine Basis von K^n ist. Dazu verwendet man das Rezept 5.2.23 mit der Familie $(v_1, \dots, v_k, e_1, \dots, e_n)$.

Rezept 5.2.26 (Durchschnitt von Untervektorräumen)**.** Gegeben seien Untervektorräume $U = \mathrm{Span}_K\{u_1, \dots, u_k\}$ und $W = \mathrm{Span}_K\{w_1, \dots, w_l\}$ von K^n. Gesucht ist eine Basis von $U \cap W$. Sei A die $n \times k$-Matrix mit Spalten u_1, …, u_k und sei B die $n \times l$-Matrix mit Spalten w_1, …, w_l. Man führt das Gaußsche Eliminationsverfahren mit der erweiterten Matrix $(A \,|\, B)$ durch, bis sie in Zeilenstufenform $(A' \,|\, B')$ ist. Sei s die Anzahl der Pivotspalten von A' und sei B'' die Matrix bestehend aus den Zeilen $B'_{(s+1)*}$, …, B'_{n*} von B'. Die Matrix B'' ist dann in Zeilenstufenform (und man kann sie weiter in reduzierte Zeilenstufenform bringen). Mithilfe von Rezept 5.1.6 erhält man eine Basis $(v_1, \dots, v_t)$ von $\mathcal{L}(B'', 0) \subset K^l$. Die Menge $\{B \cdot v_1, \dots, B \cdot v_t\}$ ist jetzt ein Erzeugendensystem von $U \cap W$ und man verwendet das Rezept 5.2.23, um eine Basis daraus auszuwählen. Denn $\mathcal{L}(B'', 0) \subset K^l$ ist das Bild von $\mathcal{L}(A' \,|\, B', 0) = \mathcal{L}(A \,|\, B, 0) \subset K^{k+l}$ unter der Projektion $K^{k+l} \to K^l$ auf die letzten l Koordinaten, und $\mathcal{L}(A \,|\, B, 0)$ besteht aus allen Vektoren $(\lambda_1, \dots, \lambda_k, \mu_1, \dots, \mu_l)$, so dass

$$\sum_{i=1}^{k} \lambda_i v_i + \sum_{j=1}^{l} \mu_j w_j = 0.$$

Damit besteht $U \cap W$ aus aller Linearkombinationen $\sum_{j=1}^{l} \mu_j w_j$, deren Koeffizienten die letzten l Koordinaten eines Vektors aus $\mathcal{L}(A \,|\, B, 0)$ sind, d.h., $U \cap W = \{B \cdot v \,|\, v \in \mathcal{L}(B'', 0)\}$.

Beispiel 5.2.27. Wir betrachten die folgenden Untervektorräume von $\mathbb{R}^3$:

$$U = \mathrm{Span}_{\mathbb{R}} \left\{ \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix} \right\}, \quad W = \mathrm{Span}_{\mathbb{R}} \left\{ \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right\}.$$

Wir berechnen den Durchschnitt $U \cap W$ mit Rezept 5.2.26 :

$$(A \,|\, B) = \left(\begin{array}{cc|cc} 0 & 3 & 2 & 1 \\ 1 & 4 & 0 & -1 \\ 2 & 5 & 1 & 0 \end{array}\right) \xrightarrow[V_{12}]{A_{32}(-2)} \left(\begin{array}{cc|cc} 1 & 4 & 0 & -1 \\ 0 & 3 & 2 & 1 \\ 0 & -3 & 1 & 2 \end{array}\right)$$
$$\xrightarrow{A_{32}(1)} \left(\begin{array}{cc|cc} 1 & 4 & 0 & -1 \\ 0 & 3 & 2 & 1 \\ 0 & 0 & 3 & 3 \end{array}\right) = (A' \,|\, B').$$

Die Matrix A' hat zwei Pivotspalten, so dass $B'' = \begin{pmatrix} 3 & 3 \end{pmatrix}$. Der Nullraum $\mathcal{L}(B'', 0) \subset \mathbb{R}^2$ ist von $e_1 - e_2$ erzeugt. Daher bildet der Vektor

$$B \cdot (e_1 - e_2) = \begin{pmatrix} 2 & 1 \\ 0 & -1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

ein Erzeugendensystem von $U \cap W$, und damit auch eine Basis.

Rezept 5.2.28 (Summe von Untervektorräumen)**.** Gegeben seien Untervektorräume $U = \mathrm{Span}_K\{u_1, \ldots, u_k\}$ und $W = \mathrm{Span}_K\{w_1, \ldots, w_l\}$ von K^n. Gesucht ist eine Basis von $U+W$. Dazu wendet man das Rezept 5.2.23 mit der Familie $(u_1, \ldots, u_k, w_1, \ldots, w_l)$ an.

Rezept 5.2.29 (komplementäre Untervektorräume)**.** Gegeben sei ein Untervektorraum $U = \mathrm{Span}_K\{u_1, \ldots, u_k\}$ von K^n. Gesucht ist eine Basis eines zu U komplementären Untervektorraums. Man verwendet das Rezept 5.2.25, um eine Basis von K^n der Gestalt $(u_{i_1}, \ldots, u_{i_r}, e_{j_1}, \ldots, e_{j_s})$ zu finden. Dann ist $(e_{j_1}, \ldots, e_{j_s})$ eine Basis eines zu U komplementären Untervektorraums.

Rezept 5.2.30 (Quotientenvektorräume)**.** Gegeben sei ein Untervektorraum $U = \mathrm{Span}_K\{u_1, \ldots, u_k\}$ von K^n. Gesucht ist eine Basis des Quotientenvektorraums K^n/U. Man verwendet das Rezept 5.2.29, um eine Basis $(w_1, \ldots, w_l)$ eines zu U komplementären Untervektorraums zu finden. Dann ist $(w_1 + U, \ldots, w_l + U)$ eine Basis von K^n/U.

Beispiel 5.2.31. Sei $K = \mathbb{R}$ und

$$U = \mathrm{Span}_{\mathbb{R}} \left\{ \begin{pmatrix} 0 \\ 1 \\ -1 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} \right\} \subset \mathbb{R}^4.$$

Wir verwenden die Rezepte 5.2.29 und 5.2.30, um einen komplementären Untervektorraum zu U und eine Basis von $\mathbb{R}^4/U$ zu bestimmen:

$$A = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 1 & 0 \\ 3 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow[A_{42}(-3)]{\substack{A_{32}(1) \\ V_{12}}} \begin{pmatrix} 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 1 & 0 \\ 0 & 3 & 0 & -3 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow[A_{42}(-3)]{A_{32}(1)} \begin{pmatrix} 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & -3 & -3 & 0 & 1 \end{pmatrix} \xrightarrow{A_{43}(3)} \begin{pmatrix} 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 3 & 1 \end{pmatrix} = A'.$$

Die Spalten von A entsprechend den Pivotspalten von A' bilden eine Basis von $\mathbb{R}^4$. Damit ist (e_1, e_3) eine Basis eines zu U komplementären Untervektorraums, und ist $(e_1 + U, e_3 + U)$ eine Basis von $\mathbb{R}^4/U$.

Rezept 5.2.32 (Smithsche Normalform)**.** Gegeben sei eine Matrix $A \in M_{m\times n}(K)$. Gesucht sind Basen B von K^n und C von K^m, so dass $[L_A]_C^B$ in Smithscher Normalform ist (siehe Satz 4.2.46). Man verwendet das Rezept 5.2.15, um eine Basis $(v_1, \ldots, v_k)$ von $\ker L_A$ zu finden, und das Rezept 5.2.25, um die zu einer Basis $B = (v_1, \ldots, v_n)$ von K^n zu ergänzen. Dann verwendet man wieder das Rezept 5.2.25, um $(A \cdot v_{k+1}, \ldots, A \cdot v_n)$ zu einer Basis C von K^m zu ergänzen.

Alternativ dazu kann man beobachten, dass eine Matrix A genau dann in Schmithscher Normalform ist, wenn beide A und $A^\intercal$ in reduzierter Zeilenstufenform sind. Man kann erstens A auf reduzierte Zeilenstufenform A' durch elementare Zeilenumformungen bringen, und zweitens A' auf Smithsche Normalform A'' durch elementare Spaltenumformungen bringen. Dabei findet man Elementarmatrizen $Z_1, \ldots, Z_k$, $W_1, \ldots, W_l$, so dass

$$A'' = Z_k \cdot \ldots \cdot Z_1 \cdot A \cdot W_1 \cdot \ldots \cdot W_l.$$

Nach der Basiswechselformel (Proposition 4.2.43) bilden die Spalten von $W_1 \cdot \ldots \cdot W_l$ und $Z_1^{-1} \cdot \ldots \cdot Z_k^{-1}$ Basen B und C mit der gewünschten Eigenschaft.

5.3 Die Determinante

Sei $n \in \mathbb{N}$. Die Determinante ist eine kanonische Abbildung $\det\colon M_n(K) \to K$ mit folgender wichtigen Eigenschaft: Eine $n \times n$-Matrix A ist genau dann invertierbar, wenn $\det(A)$ nicht null ist. Außerdem erlaubt uns die Determinante, direkte Formeln für die Koeffizienten von A^{-1} und für die Lösung eines linearen Gleichungssystems $A \cdot x = b$ zu schreiben (obwohl solche Formeln nur von theoretischer Bedeutung sind; praktisch ist das Gaußsche Eliminationsverfahren viel schneller).

5.3.1 Das Vorzeichen einer Permutation

Sei $n \in \mathbb{N}$. Zur Erinnerung ist die symmetrische Gruppe S_n die Menge aller Permutationen von $\{1, \dots, n\}$ mit der Verknüpfung $\circ$ (siehe Abschnitt 2.2.2). Falls $n \geq 3$ ist diese Gruppe nicht abelsch.

Definition 5.3.1 (Zyklus, Transposition)**.** Sei $n \in \mathbb{N}$ und sei $k \in \{1, \dots, n\}$. Eine Permutation $\sigma \in S_n$ heißt *Zyklus* der Länge k, wenn es paarweise verschiedene Elemente $a_1, \dots, a_k$ gibt, so dass $\sigma(a_i) = a_{i+1}$ für alle $i < k$, $\sigma(a_k) = a_1$, und $\sigma(b) = b$ für alle anderen Elemente b. Man schreibt dann

$$\sigma = (a_1\ a_2\ \dots\ a_k).$$

Ein Zyklus der Länge 2 heißt *Transposition.*

Beispiel 5.3.2. Mit der Zyklenschreibweise gilt:

$$S_2 = \{\mathrm{id}, (1\ 2)\},$$
$$S_3 = \{\mathrm{id}, (1\ 2), (1\ 3), (2\ 3), (1\ 2\ 3), (1\ 3\ 2)\}.$$

Bemerkung 5.3.3. Für einen Zyklus $(a_1\ a_2\ \dots\ a_k)$ gilt

$$(a_1\ a_2\ \dots\ a_k) = (a_1\ a_2) \circ \dots \circ (a_{k-1}\ a_k).$$

Insbesondere ist jeder Zyklus der Länge k ein Produkt von $k - 1$ Transpositionen.

Lemma 5.3.4. *Sei $n \in \mathbb{N}$. In der Gruppe S_n ist jedes Element ein Produkt von Transpositionen.*

Beweis. Wir verwenden Induktion über n. Falls $n = 0$ besteht S_n nur aus dem neutralen Element, das ein leeres Produkt ist. Angenommen gilt die Aussage für S_n, und sei $\sigma \in S_{n+1}$. Sei $\tau \in S_{n+1}$ die wie folgt definierte Permutation:

$$\tau = \begin{cases} \mathrm{id}, & \text{falls } \sigma(n+1) = n+1, \\ (n+1\ \sigma(n+1)), & \text{andernfalls.} \end{cases}$$

Dann gilt $(\tau \circ \sigma)(n+1) = n+1$. Das heißt, die Permutation $\tau \circ \sigma$ shränkt sich zu einer Permutation von $\{1, \dots, n\}$ ein. Nach der Induktionsvoraussetzung gibt es also Transpositionen $\tau_1, \dots, \tau_k$ in S_{n+1}, die $n+1$ festhalten, so dass $\tau \circ \sigma = \tau_k \circ \dots \circ \tau_1$. Dann gilt $\sigma = \tau \circ \tau_k \circ \dots \circ \tau_1$. □

Satz 5.3.5. *Sei $n \in \mathbb{N}$. Es gibt genau eine Abbildung*

$$\mathrm{sgn}\colon S_n \to \{1, -1\}$$

mit folgenden zwei Eigenschaften:

(i) sgn *ist ein* Gruppenhomomorphismus*, d.h.: Für alle $\sigma, \tau \in S_n$ gilt*

$$\mathrm{sgn}(\tau \circ \sigma) = \mathrm{sgn}(\tau) \cdot \mathrm{sgn}(\sigma).$$

(ii) *Für alle Transpositionen* $\sigma \in S_n$ *gilt* $\operatorname{sgn}(\sigma) = -1$.

Außerdem:

(iii) *Für alle Zyklen* $\sigma \in S_n$ *der Länge* k *gilt* $\operatorname{sgn}(\sigma) = (-1)^{k-1}$.

Beweis. Die Aussage (iii) folgt aus (i) und (ii), da jeder Zyklus der Länge k die Komposition von $k-1$ Transpositionen ist (siehe Bemerkung 5.3.3).

Zur Eindeutigkeit. Seien $s, t\colon S_n \to \{1, -1\}$ zwei Abbildungen, die (i) und (ii) erfüllen, und sei $\sigma \in S_n$. Nach Lemma 5.3.4 ist σ eine Komposition von Transpositionen $\tau_n \circ \cdots \circ \tau_1$. Nach (i) und (ii) gilt:

$$s(\sigma) = s(\tau_n) \cdot \ldots \cdot s(\tau_1) = t(\tau_n) \cdot \ldots \cdot t(\tau_1) = t(\sigma).$$

Also ist $s = t$.

Zur Existenz. Sei $\mathcal{Z}$ die Menge aller zweielementigen Teilmengen von $\{1, \ldots, n\}$:

$$\mathcal{Z} = \{I \subset \{1, \ldots, n\} \mid |I| = 2\}.$$

Sei $\sigma \in S_n$ und $I \in \mathcal{Z}$. Falls $I = \{i, j\}$ ist die rationale Zahl

$$e_I(\sigma) := \frac{\sigma(i) - \sigma(j)}{i - j} \in \mathbb{Q}^*$$

unabhängig von der Reihenfolge von i und j. Für $\sigma, \tau \in S_n$ gilt:

$$e_I(\tau \circ \sigma) = \frac{\tau(\sigma(i)) - \tau(\sigma(j))}{i - j} = \frac{\tau(\sigma(i)) - \tau(\sigma(j))}{\sigma(i) - \sigma(j)} \cdot \frac{\sigma(i) - \sigma(j)}{i - j} = e_{\sigma(I)}(\tau) \cdot e_I(\sigma). \quad (5.3.6)$$

Man definiert:

$$\begin{aligned} \operatorname{sgn}\colon S_n &\to \mathbb{Q}^*, \\ \sigma &\mapsto \prod_{I \in \mathcal{Z}} e_I(\sigma). \end{aligned}$$

Die Abbildung $I \mapsto \sigma(I)$ ist eine Permutation der endlichen Menge $\mathcal{Z}$, und damit gilt

$$\prod_{I \in \mathcal{Z}} e_{\sigma(I)}(\tau) = \prod_{I \in \mathcal{Z}} e_I(\tau). \quad (5.3.7)$$

Aus (5.3.6) und (5.3.7) folgt:

$$\operatorname{sgn}(\tau \circ \sigma) = \prod_{I \in \mathcal{Z}} e_{\sigma(I)}(\tau) \cdot \prod_{I \in \mathcal{Z}} e_I(\sigma) = \prod_{I \in \mathcal{Z}} e_I(\tau) \cdot \prod_{I \in \mathcal{Z}} e_I(\sigma) = \operatorname{sgn}(\tau) \cdot \operatorname{sgn}(\sigma).$$

Das heißt, $\operatorname{sgn}\colon S_n \to \mathbb{Q}^*$ ist ein Gruppenhomomorphismus.

Sei σ die Transposition $(k\ l)$. Wir zeigen jetzt, dass $\operatorname{sgn}(\sigma) = -1$. Sei $I = \{i, j\} \in \mathcal{Z}$. Wenn $I \cap \{k, l\} = \varnothing$, dann ist $e_I(\sigma) = \frac{i-j}{i-j} = 1$, und damit können wir diese Faktoren in dem Produkt $\operatorname{sgn}(\sigma)$ ignorieren. Für jedes $m \notin \{k, l\}$ gilt $e_{\{k,m\}}(\sigma) = \frac{l-m}{k-m} = e_{\{l,m\}}(\sigma)^{-1}$, so dass diese Faktoren in dem Produkt $\operatorname{sgn}(\sigma)$ paarweise verschwinden. Es bleibt übrig die Teilmenge $I = \{k, l\}$ selbst, für die gilt $e_{\{k,l\}}(\sigma) = \frac{l-k}{k-l} = -1$. Also ist $\operatorname{sgn}(\sigma) = -1$, wie behauptet. Aus Lemma 5.3.4 folgt schließlich, dass $\operatorname{sgn}(S_n) \subset \{1, -1\}$, und damit erhalten wir einen Gruppenhomomorphismus $\operatorname{sgn}\colon S_n \to \{1, -1\}$ mit den gewünschten Eigenschaften. □

Definition 5.3.8 (Vorzeichen, gerade/ungerade Permutationen)**.** Sei $n \in \mathbb{N}$ und $\sigma \in S_n$. Die Zahl $\operatorname{sgn}(\sigma) \in \{1, -1\}$ heißt das *Vorzeichen* oder das *Signum* von σ. Permutationen σ mit $\operatorname{sgn}(\sigma) = 1$ heißen *gerade* und die mit $\operatorname{sgn}(\sigma) = -1$ heißen *ungerade*.

Bemerkung 5.3.9. Nach dem Lemma 5.3.4 und dem Satz 5.3.5 ist eine Permutation $\sigma \in S_n$ genau dann gerade bzw. ungerade, wenn sie die Komposition einer geraden bzw. ungeraden Anzahl von Transpositionen ist.

Bemerkung 5.3.10 (alternierende Gruppe)**.** Da $\operatorname{sgn}\colon S_n \to \{1, -1\}$ ein Gruppenhomomorphismus ist, ist die Komposition zweier geraden Permutationen sowie das Inverse einer geraden Permutation wieder gerade. Insbesondere ist die Teilmenge $A_n \subset S_n$ aller geraden Permutationen eine Gruppe bzgl. $\circ$. Sie heißt die *alternierende Gruppe* vom Grad n.

5.3.2 Determinantenfunktionen

Um die Determinante $\det\colon M_n(K) \to K$ zu verstehen, ist es hilfreich, der Vektorraum $M_n(K)$ mit $(K^n)^n$ zu identifizieren, indem wir eine Matrix A als das n-Tupel ihrer Spalten $(A_{*1}, \ldots, A_{*n})$ auffassen. Die Determinante $\det\colon (K^n)^n \to K$ ist dann *keine* lineare Abbildung, aber sie ist linear bezüglich jedes ihrer n Argumente im folgenden Sinne:

Definition 5.3.11 (multilineare Abbildung)**.** Seien $n \in \mathbb{N}$ und $V_1, \ldots, V_n, W$ Vektorräume über K. Eine Abbildung

$$f\colon V_1 \times \cdots \times V_n \to W$$

heißt *multilinear* oder *n-linear* (*bilinear* wenn $n = 2$), wenn sie bezüglich jedes ihrer n Argumente linear ist, d.h.: Für jedes $i \in \{1, \ldots, n\}$ und jedes

$$(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n) \in V_1 \times \cdots \times V_{i-1} \times V_{i+1} \times \cdots \times V_n$$

ist die folgende Abbildung linear:

$$\begin{aligned} V_i &\to W, \\ v &\mapsto f(v_1, \ldots, v_{i-1}, v, v_{i+1}, \ldots, v_n). \end{aligned}$$

Definition 5.3.12 (symmetrisch, antisymmetrisch, alternierend)**.** Sei $n \in \mathbb{N}$, seien V, W Vektorräume über K und sei $f\colon V^n \to W$ eine n-lineare Abbildung.

- f heißt *symmetrisch*, wenn für alle $i, j \in \{1, \ldots, n\}$ mit $i < j$ und alle $(v_1, \ldots, v_n) \in V^n$ gilt
$$f(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) = f(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n).$$
- f heißt *antisymmetrisch*, wenn für alle $i, j \in \{1, \ldots, n\}$ mit $i < j$ und alle $(v_1, \ldots, v_n) \in V^n$ gilt
$$f(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) = -f(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n).$$
- f heißt *alternierend*, wenn für alle $i, j \in \{1, \ldots, n\}$ mit $i < j$ und alle $(v_1, \ldots, v_n) \in V^n$ mit $v_i = v_j$ gilt
$$f(v_1, \ldots, v_n) = 0.$$

Beispiel 5.3.13.

(i) Die Abbildung
$$K^n \times K^n \to K, \quad (x, y) \mapsto \sum_{i=1}^{n} x_i y_i,$$
ist eine symmetrische bilineare Abbildung.

(ii) Sei $K = \mathbb{R}$. Die Multiplikation von komplexen Zahlen $\mathbb{C} \times \mathbb{C} \to \mathbb{C}$ ist eine symmetrische bilineare Abbildung. Die Multiplikation von Quaternionen $\mathbb{H} \times \mathbb{H} \to \mathbb{H}$ ist eine bilineare Abbildung, die weder symmetrisch noch antisymmetrisch ist.

Bemerkung 5.3.14 (alternierend vs. antisymmetrisch)**.** Jede alternierende n-lineare Abbildung $f\colon V^n \to W$ ist antisymmetrisch. Ohne Beschränkung der Allgemeinheit (indem man alle Argumente von f bis auf zwei festhält), können wir $n = 2$ annehmen. Für alle $(v, v') \in V^2$ gilt dann:

$$0 = f(v + v', v + v') = f(v, v) + f(v, v') + f(v', v) + f(v', v') = f(v, v') + f(v', v),$$

und daher $f(v, v') = -f(v', v)$. Die Umkehrung gilt wenn die Charakteristik von K nicht 2 ist: In diesem Fall impliziert die Gleichung $f(v, v) = -f(v, v)$, dass $f(v, v) = 0$. Aber wenn die Charakteristik von K gleich 2 ist, dann sind „symmetrisch“ und „antisymmetrisch“ äquivalent, und „alternierend“ ist eine stärkere Bedingung.

Bemerkung 5.3.15. Multilineare Abbildungen $V_1 \times \cdots \times V_n \to W$ bilden einen Untervektorraum des Vektorraums $\mathrm{Abb}(V_1 \times \cdots \times V_n, W)$. Wenn $V_1 = \cdots = V_n = V$, dann bilden symmetrische, antisymmetrische und alternierende Abbildungen $V^n \to W$ weitere Untervektorräume davon.

Definition 5.3.16 (Determinantenfunktion)**.** Sei V ein K-Vektorraum der endlichen Dimension n. Eine *Determinantenfunktion* auf V ist eine alternierende n-lineare Abbildung $V^n \to K$. Der K-Vektorraum aller Determinantenfunktionen auf V wird mit $\mathrm{Det}(V)$ bezeichnet.

Proposition 5.3.17. *Sei $n \in \mathbb{N}$. Die Abbildung*

$$\begin{aligned} \Delta\colon (K^n)^n &\to K, \\ (v_1, \dots, v_n) &\mapsto \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot \prod_{i=1}^{n} v_{i\sigma(i)}, \end{aligned}$$

ist eine Determinantenfunktion auf K^n mit $\Delta(e_1, \dots, e_n) = 1$.

Beweis. Es gilt

$$\Delta(e_1, \dots, e_n) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot \prod_{i=1}^{n} \delta_{i\sigma(i)} = 1,$$

da das Produkt $\prod_{i=1}^{n} \delta_{i\sigma(i)}$ immer null ist, außer wenn $\sigma = \mathrm{id}$, in welchem Fall ist es gleich 1.

- Δ *ist multilinear.* Seien $i \in \{1, \dots, n\}$, $v_1, \dots, v_i, v_i', \dots, v_n \in K^n$ und $\lambda, \lambda' \in K$. Dann gilt

$$\begin{aligned} \Delta(v_1, \dots, \lambda v_i + \lambda' v_i', \dots, v_n) &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot v_{1\sigma(1)} \cdot \ldots \cdot (\lambda v_{i\sigma(i)} + \lambda' v_{i\sigma(i)}') \cdot \ldots \cdot v_{n\sigma(n)} \\ &= \lambda \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot v_{1\sigma(1)} \cdot \ldots \cdot v_{i\sigma(i)} \cdot \ldots \cdot v_{n\sigma(n)} \\ &\quad + \lambda' \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot v_{1\sigma(1)} \cdot \ldots \cdot v_{i\sigma(i)}' \cdot \ldots \cdot v_{n\sigma(n)} \\ &= \lambda \Delta(v_1, \dots, v_i, \dots, v_n) + \lambda' \Delta(v_1, \dots, v_i', \dots, v_n). \end{aligned}$$

- Δ *ist alternierend.* Sei $v_i = v_j$ mit $i < j$. Man beachte, dass die Abbildung

$$A_n \to S_n \setminus A_n, \quad \sigma \mapsto \sigma \circ (i\ j),$$

bijektiv ist (mit Umkehrabbildung $\tau \mapsto \tau \circ (i\ j)$). Es gilt

$$\begin{aligned}\Delta(v_1, \dots, v_n) &= \sum_{\sigma \in A_n} v_{1\sigma(1)} \cdots v_{i\sigma(i)} \cdots v_{j\sigma(j)} \cdots v_{n\sigma(n)} \\ &\quad - \sum_{\tau \in S_n \setminus A_n} v_{1\tau(1)} \cdots v_{i\tau(i)} \cdots v_{j\tau(j)} \cdots v_{n\tau(n)} \\ &= \sum_{\sigma \in A_n} v_{1\sigma(1)} \cdots v_{i\sigma(i)} \cdots v_{j\sigma(j)} \cdots v_{n\sigma(n)} \\ &\quad - \sum_{\sigma \in A_n} v_{1\sigma(1)} \cdots v_{i\sigma(j)} \cdots v_{j\sigma(i)} \cdots v_{n\sigma(n)} \\ &= 0,\end{aligned}$$

da $v_{ik} = v_{jk}$ für alle $k \in \{1, \dots, n\}$. □

Die Formel für die Determinantenfunktion Δ sieht vielleicht eigenartig aus. Aber wir zeigen jetzt, dass $\{\Delta\}$ ein Erzeugendensystem von $\mathrm{Det}(K^n)$ ist, das heißt: Jede Determinantenfunktion auf K^n ist gleich $\lambda \cdot \Delta$ mit einem Skalar $\lambda \in K$.

Lemma 5.3.18. *Seien V, W Vektorräume über K, $n \in \mathbb{N}$ und $f\colon V^n \to W$ eine antisymmetrische n-lineare Abbildung. Für alle Vektoren $v_1, \dots, v_n \in V$ und alle Permutationen $\sigma \in S_n$ gilt*

$$f(v_{\sigma(1)}, \dots, v_{\sigma(n)}) = \mathrm{sgn}(\sigma) \cdot f(v_1, \dots, v_n).$$

Beweis. Wenn σ eine Transposition ist, folgt die Aussage aus der Antisymmetrie von f. Gilt die Aussage für σ und τ, so gilt sie auch für $\sigma \circ \tau$, da $\mathrm{sgn}(\sigma \circ \tau) = \mathrm{sgn}(\sigma) \cdot \mathrm{sgn}(\tau)$. Da jede Permutation eine Komposition von Transpositionen ist (Lemma 5.3.4), gilt die Aussage im Allgemeinen. □

Bemerkung 5.3.19. Ist $f\colon V^n \to W$ symmetrisch, so folgt unmittelbar aus Lemma 5.3.4, dass

$$f(v_{\sigma(1)}, \dots, v_{\sigma(n)}) = f(v_1, \dots, v_n)$$

für alle $v_1, \dots, v_n \in V$ und alle $\sigma \in S_n$. Das ist tatsächlich der geschichtliche Grund dafür, dass die Gruppe S_n als symmetrische Gruppe bezeichnet wird. Ist andererseits $f\colon V^n \to W$ antisymmetrisch (z.B. alternierend), so ist f invariant gegenüber *geraden* Permutationen seiner Argumente, und deswegen wird die Gruppe A_n als alternierende Gruppe bezeichnet.

Lemma 5.3.20 (Funktorialität von Determinantenfunktionen)**.** *Sei $f\colon V \to W$ eine lineare Abbildung zwischen K-Vektorräumen derselben endlichen Dimension n und sei $\delta \in \mathrm{Det}(W)$. Dann ist die Abbildung*

$$\begin{aligned} f^*(\delta)\colon V^n &\to K, \\ (v_1, \dots, v_n) &\mapsto \delta(f(v_1), \dots, f(v_n)),\end{aligned}$$

eine Determinantenfunktion auf V. Außerdem ist die so definierte Abbildung $f^*\colon \mathrm{Det}(W) \to \mathrm{Det}(V)$ linear.

Beweis. Man hat zu zeigen, dass $f^*(\delta)$ n-linear und alternierend ist. Die n-Linearität folgt aus der n-Linearität von δ und der Linearität von f, und es ist klar, dass $f^*(\delta)$ auch alternierend ist. Die Linearität von f^* ist auch klar. □

Satz 5.3.21 (Klassifikation von Determinantenfunktionen)**.** *Sei V ein K-Vektorraum der endlichen Dimension n und sei $B = (b_1, \dots, b_n)$ eine Basis von V. Dann ist die lineare Abbildung*

$$\begin{aligned} \mathrm{Det}(V) &\to K, \\ (\delta\colon V^n \to K) &\mapsto \delta(b_1, \dots, b_n),\end{aligned}$$

ein Isomorphismus. Die Umkehrabbildung schickt $1 \in K$ *auf die Determinantenfunktion*

$$\begin{aligned} \Delta_B \colon V^n &\to K, \\ (v_1, \dots, v_n) &\mapsto \Delta([v_1]_B, \dots, [v_n]_B). \end{aligned}$$

Beweis. Zur Injektivität. Sei $\delta \in \mathrm{Det}(V)$ eine Determinantenfunktion mit $\delta(b_1, \dots, b_n) = 0$. Es gilt dann $\delta(b_{i_1}, \dots, b_{i_n}) = 0$ für alle $i_1, \dots, i_n \in \{1, \dots, n\}$: Falls zwei Indizes i_k gleich sind, folgt das aus der Definition einer alternierenden Abbildung, sonst ist

$$\delta(b_{i_1}, \dots, b_{i_n}) = \pm\delta(b_1, \dots, b_n) = 0$$

nach Lemma 5.3.18. Seien nun $v_1, \dots, v_n \in V$ beliebige Vektoren, und sei $v_i = \sum_{j=1}^n \lambda_{ij} b_j$. Da δ multilinear ist, gilt

$$\delta(v_1, \dots, v_n) = \sum_{j_1=1}^n \cdots \sum_{j_n=1}^n \lambda_{1j_1} \dots \lambda_{nj_n} \delta(b_{j_1}, \dots, b_{j_n}) = 0.$$

Also ist $\delta = 0$.

Zur Surjektvität. Es genügt zu zeigen, dass Δ_B eine Determinantenfunktion auf V mit $\Delta_B(b_1, \dots, b_n) = 1$ ist. Nach Definition ist $\Delta_B = (\varphi_B^{-1})^*(\Delta)$, wobei $\varphi_B \colon K^n \xrightarrow{\sim} V$ der der Basis B zugehörige Isomorphismus ist. Die gewünschten Eigenschaften von Δ_B folgen also aus dem Lemma 5.3.20 und der Proposition 5.3.17. □

Bemerkung 5.3.22. Sei $K = \mathbb{R}$. Für die Determinantenfunktion $\Delta \colon (\mathbb{R}^2)^2 \to \mathbb{R}$ gilt

$$|\Delta(v_1, v_2)| = \text{der Flächeninhalt des Parallelograms mit Eckpunkten } 0, v_1, v_2, v_1 + v_2.$$

Wenn $\Delta(v_1, v_2)$ nicht null ist, ist es genau dann positiv, wenn der kürzeste Winkel von v_1 nach v_2 gegen den Uhrzeigersinn läuft. Allgemeiner ist $|\Delta(v_1, \dots, v_n)|$ das Volumen des von $v_1, \dots, v_n$ aufgespannten Parallelotops in $\mathbb{R}^n$, im Sinne der Maßtheorie. Man kann deshalb Determinantenfunktionen als Varianten mit Vorzeichen des Volumens auffassen, die auch über beliebigen Körpern K sinnvoll sind.

5.3.3 Die Determinante einer Matrix

Definition 5.3.23 (Determinante einer Matrix)**.** Seien $n \in \mathbb{N}$ und $A \in M_n(K)$. Die *Determinante* von A ist der Skalar

$$\det(A) = \Delta(A_{*1}, \dots, A_{*n}),$$

wobei $\Delta \colon (K^n)^n \to K$ die Determinantenfunktion aus Proposition 5.3.17 ist.

Nach Definition von Δ gilt also

$$\det(A) = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \cdot \prod_{i=1}^n A_{\sigma(i)i}. \tag{5.3.24}$$

Diese Formel für die Determinante heißt die *Leibniz-Formel*.

Beispiel 5.3.25. Es gilt

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$$

Proposition 5.3.26 (Eigenschaften der Determinante)**.** *Sei* $n \in \mathbb{N}$ *und seien* $A, B \in M_n(K)$*. Dann gilt:*

(i) $\det(I_n) = 1$.

(ii) $\det(A \cdot B) = \det(A) \cdot \det(B)$.

(iii) *Ist A invertierbar, so ist* $\det(A^{-1}) = \det(A)^{-1}$.

(iv) $\det(A^\intercal) = \det(A)$.

Beweis. Zu (i). $\det(I_n) = \Delta(e_1, \ldots, e_n) = 1$.

Zu (ii). Nach Lemma 5.3.20 ist $L_A^*(\Delta)$ eine Determinantenfunktion auf K^n mit

$$L_A^*(\Delta)(e_1, \ldots, e_n) = \Delta(A_{*1}, \ldots, A_{*n}) = \det(A).$$

Nach der Klassifikation von Determinantenfunktionen (Satz 5.3.21) gilt $L_A^*(\Delta) = \det(A) \cdot \Delta$. Daher gilt

$$\det(A \cdot B) = L_A^*(\Delta)(B_{*1}, \ldots, B_{*n}) = \det(A) \cdot \Delta(B_{*1}, \ldots, B_{*n}) = \det(A) \cdot \det(B).$$

Zu (iii). Dies folgt unmittelbar aus (i) und (ii).

Zu (iv). Wir verwenden die Leibniz-Formel (5.3.24). Da S_n eine Gruppe ist, ist die Abbildung $S_n \to S_n$, $\tau \mapsto \tau^{-1}$, eine Permutation von S_n. Außerdem ist nach Definition jedes $\tau \in S_n$ eine Permutation der Indexmenge $\{1, \ldots, n\}$. Es gilt also

$$\begin{aligned}
\det(A) &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot \prod_{i=1}^{n} A_{\sigma(i)i} && \text{(Leibniz-Formel)} \\
&= \sum_{\tau \in S_n} \operatorname{sgn}(\tau^{-1}) \cdot \prod_{i=1}^{n} A_{\tau^{-1}(i)i} && \text{(Permutation } \tau \mapsto \tau^{-1}\text{)} \\
&= \sum_{\tau \in S_n} \operatorname{sgn}(\tau^{-1}) \cdot \prod_{j=1}^{n} A_{\tau^{-1}(\tau(j))\tau(j)} && \text{(Permutation } j \mapsto \tau(j)\text{)} \\
&= \sum_{\tau \in S_n} \operatorname{sgn}(\tau) \cdot \prod_{j=1}^{n} A_{j\tau(j)} && \text{(sgn ist ein Gruppenhomomorphismus)} \\
&= \det(A^\intercal). && \text{(Leibniz-Formel)}
\end{aligned}$$

□

Bemerkung 5.3.27. Die Gleichung $\det(A) = \det(A^\intercal)$ bedeutet, dass

$$\Delta(A_{*1}, \ldots, A_{*n}) = \Delta(A_{1*}, \ldots, A_{n*}).$$

Das heißt, es macht keinen Unterschied, ob wir die Spalten oder die Zeilen von A in der Definition der Determinante verwenden.

Bemerkung 5.3.28. Nach Proposition 5.3.26(ii,iii) schränkt sich die Determinante zu einem Gruppenhomomorphismus $\det\colon \mathrm{GL}_n(K) \to K^*$ ein.

Proposition 5.3.29 (Determinante und elementare Zeilenumformungen)**.** *Sei $n \in \mathbb{N}$ und $A \in M_n(K)$.*

(i) *Seien $i, j \in \{1, \ldots, n\}$ mit $i < j$. Dann ist*

$$\det(V_{ij} \cdot A) = -\det(A).$$

(ii) *Seien $i \in \{1, \ldots, n\}$ und $\lambda \in K^*$. Dann ist*

$$\det(M_i(\lambda) \cdot A) = \lambda \cdot \det(A).$$

(iii) *Seien $i, j \in \{1, \ldots, n\}$ mit $i \neq j$ und sei $\alpha \in K$. Dann ist*

$$\det(A_{ij}(\alpha) \cdot A) = \det(A).$$

Beweis. Nach Proposition 5.3.26(iv) ist $\det(A) = \Delta(A_{1*}, \dots, A_{n*})$. Die elementare Zeilenumformung $A \mapsto V_{ij} \cdot A$ vertauscht zwei Zeilen. Die erste Aussage folgt daraus, dass Δ antisymmetrisch ist. Die zweite und dritte Aussagen folgen ähnlich daraus, dass Δ n-linear und alternierend ist. □

Notation 5.3.30. Sei A eine $m \times n$-Matrix, $i \in \{1, \dots, m\}$ und $j \in \{1, \dots, n\}$. Wir bezeichnen mit $A[i, j]$ die $(m-1) \times (n-1)$-Matrix, die aus A entsteht, wenn man die i-te Zeile und j-te Spalte streicht.

Satz 5.3.31 (Laplacescher Entwicklungssatz)**.** *Sei $n \in \mathbb{N}$ und sei $A = (a_{ij})_{i,j}$ eine $n \times n$-Matrix über K.*

(i) (Entwicklung nach der j-ten Spalte) *Für jedes $j \in \{1, \dots, n\}$ gilt*

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A[i, j]).$$

(ii) (Entwicklung nach der i-ten Zeile) *Für jedes $i \in \{1, \dots, n\}$ gilt*

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A[i, j]).$$

Beweis. Die zweite Aussage folgt aus der ersten und Proposition 5.3.26(iv). Wir beweisen (i) mithilfe von der Leibniz-Formel. Sei $S_n^{j \mapsto i} \subset S_n$ die Teilmenge aller Permutationen, die j auf i abbilden. Für festes j ist dann $\{S_n^{j \mapsto 1}, \dots, S_n^{j \mapsto n}\}$ eine Partition von S_n. Es gibt außerdem eine Bijektion

$$S_n^{j \mapsto i} \to S_{n-1}, \quad \sigma \mapsto \sigma_{ij},$$

so dass folgendes Quadrat kommmutiert:

$$\begin{array}{ccc} \{1, \dots, n\} \setminus \{j\} & \xrightarrow{\sigma} & \{1, \dots, n\} \setminus \{i\} \\ \uparrow & & \uparrow \\ \{1, \dots, n-1\} & \xrightarrow{\sigma_{ij}} & \{1, \dots, n-1\}. \end{array}$$

Dabei sind die vertikalen Pfeile die ordnungserhaltenden Bijektionen.

Behauptung. Es gilt $\operatorname{sgn}(\sigma) = (-1)^{i+j} \operatorname{sgn}(\sigma_{ij})$.

Mit dieser Behauptung können wir berechnen:

$$\begin{aligned} \det(A) &= \sum_{i=1}^{n} \sum_{\sigma \in S_n^{j \mapsto i}} \operatorname{sgn}(\sigma) \cdot \prod_{k=1}^{n} a_{\sigma(k)k} \\ &= \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \sum_{\sigma \in S_n^{j \mapsto i}} \operatorname{sgn}(\sigma_{ij}) \cdot \prod_{l=1}^{n-1} A[i, j]_{\sigma_{ij}(l)l} \\ &= \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \sum_{\tau \in S_{n-1}} \operatorname{sgn}(\tau) \cdot \prod_{l=1}^{n-1} A[i, j]_{\tau(l)l} \\ &= \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A[i, j]). \end{aligned}$$

Es bleibt die Behauptung nachzuprüfen. Sei $\tilde{\sigma}_{ij} \in S_n$ die Fortsetzung von σ_{ij} auf $\{1, \dots, n\}$, die n auf n abbildet, so dass $\operatorname{sgn}(\sigma_{ij}) = \operatorname{sgn}(\tilde{\sigma}_{ij})$. Dann gilt

$$\sigma = (i \; i+1 \; \dots \; n) \circ \tilde{\sigma}_{ij} \circ (n \; \dots \; j+1 \; j).$$

Nach dem Satz 5.3.5(iii) ist $\operatorname{sgn}(i\ i+1\ \dots\ n) = (-1)^{n-i}$ und $\operatorname{sgn}(n\ \dots\ j+1\ j) = (-1)^{n-j}$, und damit

$$\operatorname{sgn}(\sigma) = (-1)^{2n-i-j}\operatorname{sgn}(\tilde{\sigma}_{ij}) = (-1)^{i+j}\operatorname{sgn}(\sigma_{ij}),$$

wie behauptet. □

Beispiel 5.3.32. Sei

$$A = \begin{pmatrix} 1 & 0 & -2 \\ 3 & 2 & -1 \\ -3 & 0 & 5 \end{pmatrix}.$$

Der Laplacesche Entwicklungssatz liefert sechs verschiedene Formeln für $\det(A)$, eine für jede Spalte und Zeile.

(i) Die einfachste Berechnung von $\det(A)$ ist voraussichtlich durch Entwicklung nach der zweiten Spalte, die nur einen Nicht-Null-Koeffizient enthält. Dies ergibt:

$$\det(A) = (-1)^{2+2} \cdot 2 \cdot \det \begin{pmatrix} 1 & -2 \\ -3 & 5 \end{pmatrix} = 2 \cdot (1 \cdot 5 - (-2) \cdot (-3)) = -2.$$

(ii) Entwicklung nach der ersten Zeile ergibt:

$$\det(A) = 1 \cdot \det \begin{pmatrix} 2 & -2 \\ 0 & 5 \end{pmatrix} + (-2) \cdot \det \begin{pmatrix} 3 & 2 \\ -3 & 0 \end{pmatrix} = 10 - 2 \cdot 6 = -2.$$

(iii) Entwicklung nach der dritten Spalte ergibt:

$$\begin{aligned} \det(A) &= (-2) \cdot \det \begin{pmatrix} 3 & 2 \\ -3 & 0 \end{pmatrix} - (-1) \cdot \det \begin{pmatrix} 1 & 0 \\ -3 & 0 \end{pmatrix} + 5 \cdot \det \begin{pmatrix} 1 & 0 \\ 3 & 2 \end{pmatrix} \\ &= -2 \cdot 6 + 1 \cdot 0 + 5 \cdot 2 = -2. \end{aligned}$$

Beispiel 5.3.33. Sei

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

Um die Determinante von A effektiv zu berechnen, entwickeln wir sie nach der dritten Spalte und danach nach der ersten Spalte:

$$\det(A) \overset{3.S}{=} -\det \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix} \overset{1.S}{=} -\det \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = -(-1) = 1.$$

Korollar 5.3.34 (Determinante einer Dreiecksmatrix)**.** *Sei $n \in \mathbb{N}$ und sei $A = (a_{ij})_{i,j}$ eine $n \times n$-Dreiecksmatrix, d.h., so dass $a_{ij} = 0$ für alle $i > j$ bzw. für alle $i < j$. Dann gilt*

$$\det(A) = \prod_{i=1}^{n} a_{ii}.$$

Beweis. Wir verwenden Induktion über n. Wenn $n = 0$ ist $A = I_n$ und damit $\det(A) = 1 = \prod_{i=1}^{0} a_{ii}$. Sei also $n \geq 1$. Entwicklung nach der ersten Spalte (bzw. nach der ersten Zeile) ergibt $\det(A) = a_{11} \cdot \det(A[1,1])$. Nach der Induktionsvoraussetzung ist $\det(A[1,1]) = \prod_{i=2}^{n} a_{ii}$, was die gewünschte Formel ergibt. □

Beispiel 5.3.35. Nach dem Korollar 5.3.34 und der Proposition 5.3.29 kann man auch die Determinante durch das Gaußsche Eliminationsverfahren berechnen. Denn eine quadratische Matrix in Zeilenstufenform ist insbesondere eine Dreiecksmatrix. Als Beispiel berechnen wir die Determinante der folgenden Matrix A:

$$A = \begin{pmatrix} 3 & 1 & 6 \\ 1 & 1 & 7 \\ 2 & 6 & -3 \end{pmatrix} \xrightarrow[A_{32}(-2)]{A_{12}(-3)} \xrightarrow{V_{12}} \begin{pmatrix} 1 & 1 & 7 \\ 0 & -2 & -15 \\ 0 & 4 & -17 \end{pmatrix}$$
$$\xrightarrow{A_{32}(2)} \begin{pmatrix} 1 & 1 & 7 \\ 0 & -2 & -15 \\ 0 & 0 & -47 \end{pmatrix} = A'.$$

Damit ist $\det(A) = -\det(A') = -94$.

Die folgende Verallgemeinerung des Korollars 5.3.34 ist auch nützlich:

Korollar 5.3.36 (Determinante einer Blockdreiecksmatrix)**.** *Seien $k \in \mathbb{N}$, $n_1, \ldots, n_k \in \mathbb{N} \setminus \{0\}$, $n = \sum_{i=1}^{k} n_i$ und sei A eine $n \times n$-Matrix der Gestalt*

$$A = \begin{pmatrix} A_1 & & & * \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_k \end{pmatrix} \quad \textit{bzw.} \quad A = \begin{pmatrix} A_1 & & & 0 \\ & A_2 & & \\ & & \ddots & \\ * & & & A_k \end{pmatrix},$$

wobei $A_i \in M_{n_i}(K)$. Dann gilt

$$\det(A) = \prod_{i=1}^{k} \det(A_i).$$

Beweis. Es genügt den ersten Fall zu behandeln, indem man die transponierte Matrix betrachtet. Wir verwenden Induktion über n. Wenn $n = 0$ ist die Aussage trivial. Sei also $n \geq 1$. Entwicklung von $\det(A)$ nach der ersten Spalte liefert

$$\det(A) = \sum_{e=1}^{n_1} (-1)^{e+1} (A_1)_{e1} \det(A[e,1]).$$

Nach der Induktionsvoraussetzung ist $\det(A[e,1]) = \det(A_1[e,1]) \prod_{i=2}^{k} \det(A_i)$, und damit

$$\det(A) = \left(\sum_{e=1}^{n_1} (-1)^{e+1} (A_1)_{e1} \det(A_1[e,1]) \right) \cdot \prod_{i=2}^{k} \det(A_i) = \det(A_1) \cdot \prod_{i=2}^{k} \det(A_i),$$

wie gewünscht. □

Definition 5.3.37 (Kofaktor, Kofaktormatrix, adjunkte Matrix)**.** Seien $n \in \mathbb{N}$ und A eine $n \times n$-Matrix über K.

- Der (i,j)-*Kofaktor* von A ist $(-1)^{i+j} \det(A[i,j]) \in K$.
- Die $n \times n$-Matrix
$$\operatorname{cof}(A) = \left((-1)^{i+j} \det(A[i,j]) \right)_{i,j}$$
heißt die *Kofaktormatrix* von A.
- Die *adjunkte Matrix* zu A ist die $n \times n$-Matrix $\operatorname{adj}(A) = \operatorname{cof}(A)^{\intercal}$.

Beispiel 5.3.38. Es gilt

$$\operatorname{adj}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Korollar 5.3.39. *Sei $n \in \mathbb{N}$ und sei $A \in M_n(K)$. Dann gilt*

$$A \cdot \operatorname{adj}(A) = \operatorname{adj}(A) \cdot A = \det(A) \cdot I_n.$$

Beweis. Wir berechnen zunächst $A \cdot \operatorname{adj}(A)$. Es gilt

$$(A \cdot \operatorname{adj}(A))_{ij} = \sum_{k=1}^{n} A_{ik} \operatorname{adj}(A)_{kj} = \sum_{k=1}^{n} (-1)^{k+j} A_{ik} \det(A[j,k]).$$

Wenn $i = j$ ist diese Summe gleich $\det(A)$ nach dem Laplaceschen Entwicklungssatz (Entwicklung nach der i-ten Zeile). Es bleibt also zu zeigen, dass diese Summe null ist, wenn $i \neq j$. Sei A_j die Matrix, die aus A entsteht, wenn man die j-te Zeile durch A_{i*} ersetzt. Entwicklung nach der j-ten Zeile zeigt, dass die obige Summe gleich $\det(A_j)$ ist. Aber $\det(A_j) = \Delta((A_j)_{1*}, \ldots, (A_j)_{n*}) = 0$, da die i-te und j-te Zeilen von A_j gleich sind und Δ alternierend ist. Die Berechnung von $\operatorname{adj}(A) \cdot A$ erfolgt auf ähnliche Weise durch Spaltenentwicklung. □

Korollar 5.3.40 (Determinante und Invertierbarkeit)**.** *Sei $n \in \mathbb{N}$. Eine Matrix $A \in M_n(K)$ ist genau dann invertierbar, wenn $\det(A) \in K^*$. In diesem Fall gilt*

$$A^{-1} = \det(A)^{-1} \cdot \operatorname{adj}(A).$$

Beweis. Wenn A invertierbar ist, dann ist $\det(A) \in K^*$ nach Proposition 5.3.26(iii). Sei umgekehrt $\det(A) \in K^*$. Dann ist $\det(A)^{-1} \cdot \operatorname{adj}(A)$ die inverse Matrix zu A nach Korollar 5.3.39. □

Beispiel 5.3.41. Die 2×2-Matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

ist genau dann invertierbar, wenn $ad - bc \neq 0$, in welchem Fall gilt

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

(siehe Beispiele 5.3.25 und 5.3.38).

Beispiel 5.3.42. Die Matrix $A \in M_3(K)$ aus Beispiel 5.3.35 ist genau dann invertierbar, wenn $\operatorname{char}(K) \notin \{2, 47\}$. Denn 2 und 47 sind die Primzahlen, die $\det(A) = -94$ teilen.

Korollar 5.3.43 (Cramersche Regel)**.** *Sei $n \in \mathbb{N}$ und sei (A, b) ein lineares Gleichungssystem über K mit n Gleichungen und n Unbekannten. Sei $A_j[b]$ die Matrix, die aus A entsteht, wenn man die j-te Spalte durch b ersetzt. Ist $\det(A) \in K^*$, so hat (A, b) genau eine Lösung $x \in K^n$, für die gilt*

$$x_j = \frac{\det(A_j[b])}{\det(A)}.$$

Beweis. Nach Korollar 5.3.40 ist A invertierbar, und die einzige Lösung von (A, b) ist

$$x = A^{-1} \cdot b = \det(A)^{-1} \cdot \operatorname{adj}(A) \cdot b.$$

Die j-te Koordinate ist also

$$x_j = \det(A)^{-1} \cdot \sum_{i=1}^{n} \operatorname{adj}(A)_{ji} b_i = \det(A)^{-1} \cdot \sum_{i=1}^{n} (-1)^{i+j} b_i \det(A[i,j]),$$

und die Summe ist genau die Entwicklung von $\det(A_j[b])$ nach der j-ten Spalte. □

Bemerkung 5.3.44. Wie schon gesagt, die Formeln aus Korollaren 5.3.40 und 5.3.43 sind nur von theoretischer Bedeutung. Zum Beispiel impliziert Korollar 5.3.40, dass die Koeffizienten von A^{-1} als Polynome in den Koeffizienten von A geschrieben werden können, was in der algebraischen Geometrie wichtig ist. Bei großen Matrizen ist aber das Gaußsche Eliminationsverfahren viel schneller, um A^{-1} zu berechnen oder das lineare Gleichungssystem (A, b) zu lösen, und es ist nicht sinnvoll, diese Korollare zu verwenden.

5.3.4 Die Determinante eines Endomorphismus

In diesem Abschnitt definieren wir die Determinante $\det(f)$ eines Endomorphismus $f\colon V \to V$ eines endlich-dimensionalen K-Vektorraums V. Mithilfe der Determinante von Matrizen kann man einfach $\det(f)$ als $\det([f]_B^B)$ definieren, wobei B eine beliebige Basis von V ist. Dabei muss man nachprüfen, dass $\det([f]_B^B)$ unabhängig von B ist, was aus der Basiswechselformel (Proposition 4.2.43) und der Multiplikativität der Determinante (Proposition 5.3.26) folgt. Im Folgenden geben wir aber eine begrifflichere Definition von $\det(f)$, die keine Wahl einer Basis von V erfordert, und danach beweisen wir, dass $\det(f) = \det([f]_B^B)$ (siehe Proposition 5.3.49).

Lemma 5.3.45. *Sei L ein K-Vektorraum der Dimension 1. Dann ist die Abbildung*

$$\begin{aligned} K &\to \operatorname{End}_K(L), \\ \lambda &\mapsto (v \mapsto \lambda \cdot v), \end{aligned}$$

ein Isomorphismus.

Beweis. Die Injektivität folgt aus Proposition 3.2.6(iv). Aus $L \cong K$ folgt $\operatorname{End}_K(L) \cong \operatorname{End}_K(K) \cong K$ (siehe Beispiel 4.1.45), und damit $\dim_K \operatorname{End}_K(L) = 1$. Die Sujektivität folgt dann aus Korollar 4.1.39. □

Zur Erinnerung: Ist V ein endlich-dimensionaler K-Vektorraum, so ist die Menge $\operatorname{Det}(V)$ aller Determinantenfunktionen auf V ein K-Vektorraum der Dimension 1 (Satz 5.3.21). Haben außerdem V und W dieselbe endliche Dimension, so induziert jede lineare Abbildung $f\colon V \to W$ eine lineare Abbildung $f^*\colon \operatorname{Det}(W) \to \operatorname{Det}(V)$ (Lemma 5.3.20).

Definition 5.3.46 (Determinante eines Endomorphismus)**.** Sei V ein endlich-dimensionaler K-Vektorraum und $f\colon V \to V$ ein Endomorphismus. Die *Determinante* von f ist der Skalar $\det(f) \in K$, so dass die von f induzierte Abbildung $f^*\colon \operatorname{Det}(V) \to \operatorname{Det}(V)$ gleich der Skalarmultiplikation mit $\det(f)$ ist (siehe Lemma 5.3.45), d.h.: Es gilt $f^*(\delta) = \det(f) \cdot \delta$ für alle $\delta \in \operatorname{Det}(V)$.

Bemerkung 5.3.47. Sei $K = \mathbb{R}$ und $f \in \operatorname{End}_{\mathbb{R}}(\mathbb{R}^n)$. Nach der Definition von $\det(f)$ und der Bemerkung 5.3.22 skaliert f das Volumen eines Parallelotops in $\mathbb{R}^n$ um den Faktor $|\det(f)|$. Die Determinante von f misst also die durch f bewirkte Volumenveränderung mit einem geeigneten Vorzeichen.

Proposition 5.3.48 (Eigenschaften der Determinante)**.** *Sei V ein endlich-dimensionaler K-Vektorraum und seien $f, g\colon V \to V$ zwei Endomorphismen. Dann gilt:*

(i) $\det(\mathrm{id}_V) = 1$.

(ii) $\det(g \circ f) = \det(g) \cdot \det(f)$.

(iii) *Ist f ein Automorphismus, so gilt* $\det(f^{-1}) = \det(f)^{-1}$.

(iv) *Für die duale Abbildung $f^*\colon V^* \to V^*$ gilt* $\det(f^*) = \det(f)$.

Beweis. Jede Aussage folgt aus Proposition 5.3.49(ii) unten und der entsprechenden Aussage über die Determinante von Matrizen (Proposition 5.3.26). Wir geben aber zusätzlich matrixfreie Beweise.

Zu (i). Dies folgt daraus, dass $\mathrm{id}_V^* \colon \mathrm{Det}(V) \to \mathrm{Det}(V)$ die Identität ist.

Zu (ii). Sei $\delta \in \mathrm{Det}(V)$. Dann gilt

$$(g \circ f)^*(\delta) = f^*(g^*(\delta)) = f^*(\det(g) \cdot \delta) = \det(g) \cdot f^*(\delta) = \det(g) \cdot \det(f) \cdot \delta.$$

Zu (iii). Dies folgt unmittelbar aus (i) und (ii).

Zu (iv). In diesem Beweis schreiben wir $\mathrm{Det}(f)$ für die von f induzierte lineare Abbildung $\mathrm{Det}(V) \to \mathrm{Det}(V)$, um sie nicht mit der dualen Abbildung f^* zu verwechseln. Sei n die Dimension von V. Wir definieren die Abbildung $\Theta \colon (V^*)^n \to \mathrm{Det}(V)$ durch

$$\Theta(\alpha_1, \ldots, \alpha_n)(v_1, \ldots, v_n) = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \alpha_{\sigma(1)}(v_1) \ldots \alpha_{\sigma(n)}(v_n).$$

Man kann leicht nachprüfen, dass $\Theta(\alpha_1, \ldots, \alpha_n)$ eine Determinantenfunktion auf V ist, und dass Θ selbst n-linear und alternierend ist (vgl. den Beweis der Proposition 5.3.17). Die Abbildung Θ ist außerdem surjektiv: Ist $B = (b_1, \ldots, b_n)$ eine Basis von V, so ist der Vektorraum $\mathrm{Det}(V)$ von Δ_B erzeugt (Satz 5.3.21), und es gilt $\Delta_B = \Theta(b_1^*, \ldots, b_n^*)$. Man erhält daraus eine injektive lineare Abbildung

$$\begin{aligned} \vartheta \colon \mathrm{Det}(V)^* &\to \mathrm{Det}(V^*), \\ \varepsilon &\mapsto \varepsilon \circ \Theta. \end{aligned}$$

Da beide $\mathrm{Det}(V)^*$ und $\mathrm{Det}(V^*)$ die Dimension 1 haben, ist ϑ sogar ein Isomorphismus. Es gilt zudem $\mathrm{Det}(f^*) \circ \vartheta = \vartheta \circ \mathrm{Det}(f)^*$. Da die Abbildung $\mathrm{Det}(f)$ durch Skalarmultiplikation mit $\det(f)$ gegeben ist, gilt das Gleiche für $\mathrm{Det}(f)^*$ und damit für $\mathrm{Det}(f^*)$, d.h., $\det(f^*) = \det(f)$. □

Proposition 5.3.49.

(i) *Seien $n \in \mathbb{N}$ und $A \in M_n(K)$. Dann gilt* $\det(L_A) = \det(A)$.

(ii) *Sei f ein Endomorphismus eines endlich-dimensionalen K-Vektorraum V. Für alle Basen B von V gilt* $\det(f) = \det([f]_B^B)$.

Beweis. Die erste Aussage ist der Sonderfall der zweiten mit $V = K^n$ und $B = (e_1, \ldots, e_n)$. Sei $\delta \in \mathrm{Det}(V)$. Nach Satz 5.3.21 ist $\delta = \lambda \cdot \Delta_B$ mit einem $\lambda \in K$. Sei $B = (b_1, \ldots, b_n)$. Es gilt

$$\begin{aligned} f^*(\delta)(b_1, \ldots, b_n) &= \lambda \cdot \Delta_B(f(b_1), \ldots, f(b_n)) = \lambda \cdot \Delta([f(b_1)]_B, \ldots, [f(b_n)]_B) \\ &= \lambda \cdot \Delta([f]_B^B \cdot [b_1]_B, \ldots, [f]_B^B \cdot [b_n]_B) = \lambda \cdot \Delta([f]_B^B \cdot e_1, \ldots, [f]_B^B \cdot e_n) \\ &= \lambda \cdot \det([f]_B^B) = \delta(b_1, \ldots, b_n) \cdot \det([f]_B^B). \end{aligned}$$

Die Determinantenfunktionen $f^*(\delta)$ und $\det([f]_B^B) \cdot \delta$ stimmen also auf $(b_1, \ldots, b_n)$ überein. Aus dem Satz 5.3.21 folgt, dass $f^*(\delta) = \det([f]_B^B) \cdot \delta$, und damit dass $\det(f) = \det([f]_B^B)$. □

Proposition 5.3.50. *Sei V ein endlich-dimensionaler K-Vektorraum. Ein Endomorphismus $f \colon V \to V$ ist genau dann ein Automorphismus, wenn* $\det(f) \in K^*$.

Beweis. Sei B eine Basis von V. Dann ist f genau dann ein Automorphismus, wenn die Darstellungsmatrix $[f]_B^B$ invertierbar ist. Die Aussage folgt jetzt aus Proposition 5.3.49(ii) und Korollar 5.3.40. □

Kapitel 6

Eigenwerte und Diagonalisierbarkeit

In diesem Kapitel beschäftigen wir uns mit *Endomorphismen* von Vektorräumen. Das Endziel ist es, die Frage 4.2.45(ii) vollständig zu beantworten, d.h., Endomorphismen von endlich-dimensionalen Vektorräumen bis auf Isomorphie zu klassifizieren. Das werden wir aber nur in der Vorlesung *Lineare Algebra II* schaffen. In diesem Kapitel führen wir die wichtigen Begriffe von *Eigenwert* und *Eigenvektoren* ein, und erhalten wir eine Klassifikation von sogenannten *diagonalisierbaren* Endomorphismen, die bis auf Isomorphie durch ihre Eigenwerte bestimmt sind. Außerdem entwickeln wir mehrere Hilfsmittel zum besseren Verständnis von Endomorphismen, wie zum Beispiel das *charakteristische Polynom*. Die meisten Begriffe in diesem Kapitel sind tatsächlich auch sinnvoll bei unendlich-dimensionalen Vektorräumen, und es gibt wie immer in diesem Fall interessante Beispiele aus der Analysis. Deshalb berücksichtigen wir in unserer Untersuchung so weit wie möglich beliebige Vektorräume.

Ein beliebiger Grundkörper K wird immer noch festgelegt.

6.1 Präliminarien zu Endomorphismen

6.1.1 Direkte Summen von Vektorräumen

Die direkte Summe zweier K-Vektorräume haben wir bereits definiert (Definition 3.3.45). Als Nächstes wollen wir diese Konstruktion auf mehr als zwei Vektorräume verallgemeinern. Bei unendlich vielen Vektorräumen muss man aber zwischen dem Produkt und der direkten Summe unterscheiden, die verschiedene universelle Eigenschaften haben.

Definition 6.1.1 (Produkt und direkte Summe einer Familie von Vektorräumen)**.** Sei I eine beliebige Menge und $(V_i)_{i\in I}$ eine Familie von K-Vektorräumen.

- Das *Produkt* von $(V_i)_{i\in I}$ ist der K-Vektorraum

$$\prod_{i\in I} V_i = \{(v_i)_{i\in I} \mid v_i \in V_i \text{ für alle } i \in I\}$$

  mit der punktweisen Addition bzw. Skalarmultiplikation. Ist $e \in I$, so bezeichnen wir mit $\pi_e \colon \prod_{i\in I} V_i \to V_e$ die lineare Abbildung mit $\pi_e((v_i)_{i\in I}) = v_e$.

- Die *direkte Summe* von $(V_i)_{i\in I}$ ist der Untervektorraum

$$\bigoplus_{i\in I} V_i \subset \prod_{i\in I} V_i$$

bestehend aus allen Familien $(v_i)_{i\in I}$, die null sind außerhalb einer endlichen Teilmenge von I. Ist $e \in I$, so bezeichnen wir mit $\iota_e \colon V_e \to \bigoplus_{i\in I} V_i$ die lineare Abbildung mit

$$\iota_e(v)_i = \begin{cases} v, & \text{falls } i = e, \\ 0, & \text{andernfalls.} \end{cases}$$

Bemerkung 6.1.2. Es ist $K^I = \prod_{i\in I} K$ und $K^{(I)} = \bigoplus_{i\in I} K$.

Proposition 6.1.3 (universelle Eigenschaften des Produkts und der direkten Summe)**.** *Sei $(V_i)_{i\in I}$ eine Familie von K-Vektorräumen.*

(i) *Zu jedem K-Vektorraum W und jeder Familie $(f_i \colon W \to V_i)_{i\in I}$ von linearen Abbildungen gibt es* genau eine *lineare Abbildung $f \colon W \to \prod_{i\in I} V_i$, so dass $\pi_i \circ f = f_i$ für alle $i \in I$.*

(ii) *Zu jedem K-Vektorraum W und jeder Familie $(f_i \colon V_i \to W)_{i\in I}$ von linearen Abbildungen gibt es* genau eine *lineare Abbildung $f \colon \bigoplus_{i\in I} V_i \to W$, so dass $f \circ \iota_i = f_i$ für alle $i \in I$.*

Beweis. Zu (i). Man definiert f durch $f(w) = (f_i(w))_{i\in I}$, so dass nach Definition gilt $\pi_i \circ f = f_i$. Die Abbildung f ist linear, da der Vektorraumstruktur auf $\prod_{i\in I} V_i$ punktweise definiert ist. Die Eindeutigkeit ist klar, da die i-te Koordinate von $f(w)$ gleich $\pi_i(f(w)) = f_i(w)$ sein muss.

Zu (ii). Man definiert f durch $f((v_i)_{i\in I}) = \sum_{i\in I} f_i(v_i)$; dies ist sinnvoll (und offensichtlich linear), da nur endlich viele Summanden nicht null sind. Nach Definition der direkten Summe ist jedes Element von $\bigoplus_{i\in I} V_i$ eine Linearkombination von Elemente der Gestalt $\iota_i(v_i)$ mit $i \in I$ und $v_i \in V_i$. Deswegen gibt es höchstens ein f mit den geforderten Eigenschaften. □

Bemerkung 6.1.4. Das Produkt und die Summe einer Mengenfamilie (Definition 1.2.15) haben analoge universelle Eigenschaften bzgl. beliebiger statt linearer Abbildungen.

Bemerkung 6.1.5 (Dimension einer direkten Summe)**.** Ist $(V_i)_{i\in I}$ eine endliche Familie von endlich-dimensionalen K-Vektorräumen, so gilt

$$\dim_K \left(\bigoplus_{i\in I} V_i \right) = \sum_{i\in I} \dim_K V_i.$$

Denn sind B_i Basen von V_i, so ist die Zusammensetzung der Familien $\iota_i(B_i)$ eine Basis von $\bigoplus_{i\in I} V_i$.

Die folgende Definition ist eine Verallgemeinerung von Definition 3.3.36:

Definition 6.1.6 (Summe einer Familie von Untervektorräumen)**.** Sei V ein K-Vektorraum und $(U_i)_{i\in I}$ eine Familie von Untervektorräumen. Die *Summe* der Familie $(U_i)_{i\in I}$ ist der Untervektorraum

$$\sum_{i\in I} U_i := \operatorname{Span}_K \left(\bigcup_{i\in I} U_i \right) \subset V.$$

Ist $(U_i)_{i\in I}$ eine Familie von Untervektorräumen von V, so gibt es nach der universellen Eigenschaft der direkten Summe eine kanonische surjektive lineare Abbildung

$$\bigoplus_{i\in I} U_i \to \sum_{i\in I} U_i.$$

Wenn diese Abbildung bijektiv ist, sagt man auch, dass die Summe $\sum_{i\in I} U_i$ *direkt* ist.

Proposition 6.1.7. *Sei V ein K-Vektorraum und $(U_i)_{i\in I}$ eine Familie von Untervektorräumen. Dann sind die folgenden Aussagen äquivalent:*

(i) *Die kanonische Abbildung $\bigoplus_{i\in I} U_i \to \sum_{i\in I} U_i$ ist ein Isomorphismus.*

(ii) *Für jede endliche Teilmenge $J \subset I$ ist jede Familie $(u_i)_{i\in J}$ mit $u_i \in U_i \setminus \{0\}$ linear unabhängig.*

Beweis. Zu (i) $\Rightarrow$ (ii). Sei $\sum_{i\in J} \lambda_i u_i = 0$. Die Linearkombination $\sum_{i\in J} \lambda_i u_i$ ist das Bild von $(\lambda_i u_i)_{i\in J}$ unter der kanonischen Abbildung

$$\bigoplus_{i\in J} U_i \to \bigoplus_{i\in I} U_i \to \sum_{i\in I} U_i,$$

die injektiv ist. Daraus folgt, dass $\lambda_i u_i = 0$ und damit $\lambda_i = 0$ für alle $i \in J$.

Zu (ii) $\Rightarrow$ (i). Die kanonische Abbildung ist surjektiv nach Definition der Summe. Sei $(u_i)_{i\in I}$ ein Element ihres Kerns, d.h., $\sum_{i\in I} u_i = 0$. Sei $J \subset I$ die endliche Teilmenge aller Indizes i mit $u_i \neq 0$. Nach (ii) ist die Familie $(u_i)_{i\in J}$ linear unabhängig. Daraus folgt, dass $J = \varnothing$, sonst würde die Gleichung $\sum_{i\in J} u_i = 0$ im Widerspruch zu der linearen Unabhängigkeit stehen. □

6.1.2 Invariante Untervektorräume

Definition 6.1.8 (invarianter Untervektorraum)**.** Sei $f\colon V \to V$ ein Endomorphismus eines K-Vektorraums V. Ein Untervektorraum $U \subset V$ heißt *f-invariant*, wenn $f(U) \subset U$.

Wenn $U \subset V$ f-invariant ist, dann schränkt sich f zu einem Endomorphismus $f_U \in \mathrm{End}_K(U)$. Nach der universellen Eigenschaft des Quotientenvektorraums (Proposition 4.1.33) induziert auch f einen Endomorphismus $\bar{f} \in \mathrm{End}_K(V/U)$ mit $\bar{f}(v + U) = f(v) + U$.

Beispiel 6.1.9. Sei $f\colon V \to V$ ein Endomorphismus. Dann sind $\{0\}$, V, $\ker f$ und $\operatorname{im} f$ f-invariante Untervektoräume von V.

Proposition 6.1.10. *Sei $f\colon V \to V$ ein Endomorphismus eines K-Vektorraums V und sei $(U_i)_{i\in I}$ eine Familie von f-invarianten Untervektorräumen.*

(i) *Der Durchschnitt $\bigcap_{i\in I} U_i$ ist f-invariant.*

(ii) *Die Summe $\sum_{i\in I} U_i$ ist f-invariant.*

Beweis. Dies folgt aus den mengentheoretischen Formeln

$$f\left(\bigcap_{i\in I} U_i\right) \subset \bigcap_{i\in I} f(U_i) \quad \text{und} \quad f\left(\bigcup_{i\in I} U_i\right) = \bigcup_{i\in I} f(U_i)$$

und der Proposition 4.1.20(i). □

Weitere Beispiele erhalten wir durch die Potenzen von f:

Notation 6.1.11 (Potenzen eines Endomorphismus)**.** Sei $f\colon V \to V$ ein Endomorphismus eines K-Vektorraums V und sei $k \in \mathbb{N}$. Man definiert die k-te Potenz f^k rekursiv wie folgt:

$$f^0 = \mathrm{id}_V, \quad f^{k+1} = f \circ f^k.$$

Proposition 6.1.12. *Sei V ein K-Vektorraum, $f \in \mathrm{End}_K(V)$ ein Endomorphismus und $g \in \mathrm{End}_K(V)$ eine Linearkombination der Familie $(f^k)_{k\in\mathbb{N}}$. Dann:*

(i) $\ker g$ *und* $\operatorname{im} g$ *sind f-invariant.*

(ii) *Ist $U \subset V$ ein f-invarianter Untervektorraum, so ist U auch g-invariant.*

Beweis. *Zu* (i). Sei $g = \sum_{k=0}^{n} \lambda_k f^k$. Die Linearität von f impliziert, dass $g \circ f = f \circ g$, denn:

$$g(f(v)) = \sum_{k=0}^{n} \lambda_k f^{k+1}(v) = f\left(\sum_{k=0}^{n} \lambda_k f^k(v)\right) = f(g(v)).$$

Sei $v \in \ker g$. Dann ist $g(f(v)) = f(g(v)) = f(0) = 0$, und damit ist $f(v) \in \ker g$. Sei nun $v \in \operatorname{im} g$, d.h., $v = g(w)$ mit einem $w \in V$. Dann ist $f(v) = f(g(w)) = g(f(w))$, und damit ist $f(w) \in \operatorname{im} g$.

Zu (ii). Durch Induktion über k schließen wir unmittelbar, dass $f^k(U) \subset U$ für alle $k \in \mathbb{N}$. Da U ein Untervektorraum ist, folgern wir, dass $g(U) \subset U$. □

6.1.3 Isomorphie von Endomorphismen

Definition 6.1.13 (Isomorphie von Endomorphismen)**.** Seien V, W Vektorräume über K und seien $f \in \operatorname{End}_K(V)$ und $g \in \operatorname{End}_K(W)$ Endomorphismen. Man sagt, dass die Paare (V, f) und (W, g) *isomorph* sind, und man schreibt $(V, f) \cong (W, g)$, wenn ein Isomorphismus $\varphi\colon V \xrightarrow{\sim} W$ existiert, so dass $\varphi \circ f = g \circ \varphi$:

$$\begin{array}{ccc} V & \xrightarrow[\sim]{\varphi} & W \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle g} \\ V & \xrightarrow[\sim]{\varphi} & W. \end{array}$$

Bemerkung 6.1.14. Man kann leicht nachprüfen, dass Isomorphie eine Äquivalenzrelation zwischen Paaren (V, f) ist (vgl. Bemerkung 4.1.16).

Beispiel 6.1.15.

(i) Jedes (V, f) ist isomorph zu einem Paar der Gestalt $(K^{(I)}, g)$: Man wählt einen Isomorphismus $\varphi\colon V \xrightarrow{\sim} K^{(I)}$ und setzt $g = \varphi \circ f \circ \varphi^{-1}$.

(ii) Sei V ein n-dimensionaler K-Vektorraum mit einer Basis B. Durch den Isomorphismus $\varphi_B^{-1}\colon V \to K^n$ ist jedes Paar (V, f) zu $(K^n, L_{[f]_B^B})$ isomorph.

Bemerkung 6.1.16. Die Determinante von Endomorphismen (Definition 5.3.46) ist eine *Isomorphie-Invariante* im folgenden Sinne: Ist V endlich-dimensional und sind (V, f) und (W, g) isomorph, so ist $\det(f) = \det(g)$. Denn sei $\varphi\colon V \xrightarrow{\sim} W$ ein Isomorphismus mit $\varphi \circ f = g \circ \varphi$ und sei $\delta \in \operatorname{Det}(W)$ eine Determinantenfunktion. Dann gilt

$$\begin{aligned} g^*(\delta) &= (\varphi \circ f \circ \varphi^{-1})^*(\delta) \\ &= (\varphi^{-1})^*(f^*(\varphi^*(\delta))) \\ &= (\varphi^{-1})^*(\det(f) \cdot \varphi^*(\delta)) \\ &= \det(f) \cdot (\varphi^{-1})^*(\varphi^*(\delta)) \\ &= \det(f) \cdot (\varphi \circ \varphi^{-1})^*(\delta) \\ &= \det(f) \cdot \delta, \end{aligned}$$

und damit $\det(g) = \det(f)$.

Definition 6.1.17 (Ähnlichkeit von Matrizen)**.** Seien $n \in \mathbb{N}$ und $A, B \in M_n(K)$. Man sagt, dass A *ähnlich* oder *konjugiert* zu B ist, wenn eine invertierbare Matrix $S \in \mathrm{GL}_n(K)$ existiert, so dass $S^{-1} \cdot A \cdot S = B$.

Proposition 6.1.18. *Ähnlichkeit ist eine Äquivalenzrelation auf* $M_n(K)$.

Beweis. Sie ist reflexiv, da $A = I_n^{-1} \cdot A \cdot I_n$. Sie ist symmetrisch, denn:

$$B = S^{-1} \cdot A \cdot S \quad \Longrightarrow \quad A = (S^{-1})^{-1} \cdot B \cdot S^{-1}.$$

Zur Transitivät, seien $B = S^{-1} \cdot A \cdot S$ und $C = T^{-1} \cdot B \cdot T$. Dann gilt:

$$(S \cdot T)^{-1} \cdot A \cdot (S \cdot T) = T^{-1} \cdot (S^{-1} \cdot A \cdot S) \cdot T = T^{-1} \cdot B \cdot T = C. \qquad \square$$

Proposition 6.1.19. *Sei $n \in \mathbb{N}$.*

(i) *Seien $A, B \in M_n(K)$. Dann sind A und B genau dann ähnlich, wenn (K^n, L_A) und (K^n, L_B) isomorph sind.*

(ii) *Seien V und W n-dimensionale K-Vektorräume mit Endomorphismen $f \in \mathrm{End}_K(V)$ und $g \in \mathrm{End}_K(W)$ und mit Basen B und C. Dann sind (V, f) und (W, g) genau dann isomorph, wenn $[f]_B^B$ und $[g]_C^C$ ähnlich sind.*

Beweis. Die erste Aussage ist der Sonderfall der zweiten, mit $V = W = K^n$ und $B = C$ der Standardbasis. Sei $\varphi\colon V \to W$ ein Isomorphismus mit $\varphi \circ f \circ \varphi^{-1} = g$. Für $S = [\varphi^{-1}]_B^C$ gilt dann $S^{-1}[f]_B^B S = [g]_C^C$ nach Proposition 4.2.39. Sei umgekehrt $S \in \mathrm{GL}_n(K)$ mit $S^{-1}[f]_B^B S = [g]_C^C$. Sei $\varphi\colon V \to W$ der Isomorphismus mit $[\varphi]_C^B = S^{-1}$. Nach Proposition 4.2.39 gilt dann $[\varphi \circ f \circ \varphi^{-1}]_C^C = S^{-1}[f]_B^B S = [g]_C^C$, und daher $\varphi \circ f \circ \varphi^{-1} = g$. $\square$

Man kann die letzte Proposition wie folgt zusammenfassen: Es gibt eine bijektive Abbildung

$$\begin{aligned} M_n(K)/\text{Ähnlichkeit} &\to \{(V, f) \mid \dim_K V = n \text{ und } f \in \mathrm{End}_K(V)\}/\text{Isomorphie}, \\ [A] &\mapsto [(K^n, L_A)], \end{aligned}$$

deren Umkehrabbildung die Äquivalenzklasse von (V, f) auf die von $[f]_B^B$ abbildet, wobei B eine beliebige Basis von V ist. In der Praxis bedeutet das folgendes: Wenn wir eine Abbildung $\Phi\colon M_n(K) \to X$ definiert haben, so dass $\Phi(A) = \Phi(B)$ wann immer A und B ähnlich sind, dann erhalten wir aus jedem Paar (V, f) ein wohldefiniertes Element $\Phi(f) \in X$, so dass $\Phi(f) = \Phi([f]_B^B)$ für jede Basis B von V. Außerdem gilt dann $\Phi(f) = \Phi(g)$, wenn die Paare (V, f) und (W, g) isomorph sind, d.h., Φ ist automatisch eine *Isomorphie-Invariante*. Als Beispiel dieser Methode definieren wir jetzt die *Spur* eines Endomorphismus.

Definition 6.1.20 (Spur einer Matrix)**.** Seien $n \in \mathbb{N}$ und $A = (a_{ij})_{i,j}$ eine $n \times n$-Matrix über K. Die *Spur* von A ist

$$\mathrm{tr}(A) := \sum_{i=1}^{n} a_{ii} \in K.$$

Proposition 6.1.21 (Spureigenschaft)**.** *Seien $m, n \in \mathbb{N}$, $A \in M_{m \times n}(K)$ und $B \in M_{n \times m}(K)$. Dann gilt*

$$\mathrm{tr}(A \cdot B) = \mathrm{tr}(B \cdot A).$$

Beweis. $\mathrm{tr}(A \cdot B) = \displaystyle\sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij} B_{ji} = \sum_{j=1}^{n} \sum_{i=1}^{m} B_{ji} A_{ij} = \mathrm{tr}(B \cdot A).$ $\square$

Bemerkung 6.1.22. Seien $A_1, \ldots, A_k \in M_n(K)$ und $\sigma \in S_k$. Die Spureigenschaft impliziert, dass

$$\mathrm{tr}(A_1 \cdot \ldots \cdot A_k) = \mathrm{tr}(A_{\sigma(1)} \cdot \ldots \cdot A_{\sigma(k)}),$$

sofern σ ein Zyklus der Länge k ist. Im Gegensatz zu der Determinante, gilt das aber im Allgemeinen nicht für eine beliebige Permutation σ.

Korollar 6.1.23. *Sei $n \in \mathbb{N}$ und seien $A, B \in M_n(K)$. Sind A und B ähnlich, so gilt* $\mathrm{tr}(A) = \mathrm{tr}(B)$.

Beweis. Sei $S^{-1} \cdot A \cdot S = B$. Nach Proposition 6.1.21 gilt nun

$$\operatorname{tr}(B) = \operatorname{tr}(S^{-1} \cdot A \cdot S) = \operatorname{tr}(A \cdot S \cdot S^{-1}) = \operatorname{tr}(A). \qquad \square$$

Definition 6.1.24 (Spur eines Endomorphismus)**.** Sei V ein endlich-dimensionaler K-Vektorraum und $f \in \operatorname{End}_K(V)$. Die *Spur* von f ist

$$\operatorname{tr}(f) := \operatorname{tr}([f]_B^B),$$

wobei B eine Basis von V ist. Nach Korollar 6.1.23 ist $\operatorname{tr}(f)$ unabhängig von der Wahl der Basis B.

Bemerkung 6.1.25. Es ist auch möglich, eine begrifflichere matrixfreie Definition der Spur $\operatorname{tr}(f)$ zu formulieren, wie bei der Determinante (siehe Abschnitt 5.3.4). Dazu braucht man aber den Begriff des *Tensorprodukts* von Vektorräumen, den wir noch nicht besprochen haben.

6.2 Eigenvektoren und Eigenwerte

Definition 6.2.1 (Eigenraum, Eigenvektor, Eigenwert)**.** Sei $f\colon V \to V$ ein Endomorphismus eines K-Vektorraums V und sei $\lambda \in K$.

- Der *Eigenraum* zu λ von f ist der Untervektorraum

$$\operatorname{Eig}_\lambda(f) := \{v \in V \mid f(v) = \lambda \cdot v\} = \ker(\lambda \cdot \mathrm{id}_V - f) \subset V.$$

- Ein *Eigenvektor* zu λ von f ist ein Element von $\operatorname{Eig}_\lambda(f) \setminus \{0\}$, d.h., ein Vektor $v \neq 0$, so dass $f(v) = \lambda \cdot v$.

- λ heißt *Eigenwert* von f, wenn $\operatorname{Eig}_\lambda(f) \neq \{0\}$, d.h., wenn ein Eigenvektor zu λ existiert.

Ist $n \in \mathbb{N}$ und ist A eine $n \times n$-Matrix über K, so bezeichnen wir als *Eigenräume*, *Eigenvektoren* und *Eigenwerte* von A die Eigenräume, Eigenvektoren und Eigenwerte von $L_A\colon K^n \to K^n$.

Bemerkung 6.2.2. Es gilt $\operatorname{Eig}_0(f) = \ker f$.

Bemerkung 6.2.3. Seien $f \in \operatorname{End}_K(V)$ und $g \in \operatorname{End}_K(W)$ Endomorphismen, und sei $\varphi\colon V \xrightarrow{\sim} W$ ein Isomorphismus mit $\varphi \circ f = g \circ \varphi$. Für alle $\lambda \in K$ ist dann $\varphi(\operatorname{Eig}_\lambda(f)) = \operatorname{Eig}_\lambda(g)$, denn:

$$f(v) = \lambda \cdot v \iff \varphi(f(v)) = \varphi(\lambda \cdot v) \iff g(\varphi(v)) = \lambda \cdot \varphi(v).$$

Insbesondere haben f und g dieselben Eigenwerte.

Beispiel 6.2.4. Sei V ein K-Vektorraum.

(i) Für die Identität id_V gilt

$$\operatorname{Eig}_\lambda(\mathrm{id}_V) = \begin{cases} V, & \text{falls } \lambda = 1, \\ \{0\}, & \text{falls } \lambda \neq 1. \end{cases}$$

Falls $V \neq \{0\}$ ist also $1 \in K$ der einzige Eigenwert von id_V, und alle Vektoren $v \neq 0$ sind Eigenvektoren dazu.

(ii) Allgemeiner, ist $\mu \in K$ und $V \neq \{0\}$, so ist μ der einzige Eigenwert von $\mu \cdot \mathrm{id}_V$, und alle Vektoren $v \neq 0$ sind Eigenvektoren dazu.

(iii) Sei $t\colon V \oplus V \to V \oplus V$ die lineare Abbildung $t(v,w) = (w,v)$. Es gilt

$$\begin{aligned}\mathrm{Eig}_\lambda(t) &= \{(v,w) \in V \oplus V \mid w = \lambda \cdot v \text{ und } v = \lambda \cdot w\} \\ &= \{(v, \lambda \cdot v) \mid v \in V \text{ und } (\lambda^2 - 1) \cdot v = 0\}.\end{aligned}$$

Nach Proposition 3.2.6(iv) ist $(\lambda^2 - 1) \cdot v = 0$ nur möglich, wenn $v = 0$ oder $\lambda^2 - 1 = 0$. Daher gilt

$$\mathrm{Eig}_\lambda(t) = \{(0,0)\} \cup \{(v, \lambda \cdot v) \mid v \in V \setminus \{0\} \text{ und } \lambda^2 - 1 = 0\}.$$

Falls $V \neq \{0\}$ sind also die Eigenwerte von t alle Skalare λ mit $\lambda^2 - 1 = 0$. Aus der Gleichung $\lambda^2 - 1 = (\lambda - 1)(\lambda + 1)$ und der Nullteilerfreiheit von K folgt, dass 1 und -1 die einzigen Eigenwerte von t sind. Die Eigenvektoren zu 1 sind (v,v) mit $v \neq 0$ und die Eigenvektoren zu -1 sind $(v,-v)$ mit $v \neq 0$.

(iv) Sei $D = \mathrm{diag}(d_1, \ldots, d_n)$ eine $n \times n$-Diagonalmatrix über K. Dann ist der Eigenraum zu λ von D der Untervektorraum $\mathrm{Span}_K\{e_i \mid d_i = \lambda\} \subset K^n$. Insbesondere ist jeder Standardeinheitsvektor $e_i \in K^n$ ein Eigenvektor zum Eigenwert d_i von D.

Beispiel 6.2.5 (Drehungen)**.** Sei $\alpha \in \mathbb{R} \setminus \pi\mathbb{Z}$ eine reelle Zahl, die kein ganzzahliges Vielfaches von π ist, und sei $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ die Drehung um den Winkel α um den Nullpunkt. Dann ist $\mathrm{Eig}_\lambda(f) = \{0\}$ für alle $\lambda \in K$. Denn für alle $v \neq 0$ liegen die Vektoren v und $f(v)$ auf keiner gemeinsamen Ursprungsgerade. Der Endomorphismus f hat also keine Eigenwerte und somit keine Eigenvektoren.

Beispiel 6.2.6 (Abhängigkeit vom Grundkörper)**.** Sei

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in M_2(\mathbb{R}).$$

Die Abbildung $L_A\colon \mathbb{R}^2 \to \mathbb{R}^2$ ist eine Drehung um den Winkel $\pi/2$, die keine Eigenwerte besitzt (Beispiel 6.2.5). Betrachtet man jedoch A als komplexe Matrix, so hat $L_A\colon \mathbb{C}^2 \to \mathbb{C}^2$ die Eigenwerte i und $-i$, mit zugehörigen Eigenvektoren $\left(\begin{smallmatrix} i \\ 1 \end{smallmatrix}\right)$ und $\left(\begin{smallmatrix} 1 \\ i \end{smallmatrix}\right)$:

$$A \begin{pmatrix} i \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ i \end{pmatrix} = i \begin{pmatrix} i \\ 1 \end{pmatrix}, \quad A \begin{pmatrix} 1 \\ i \end{pmatrix} = \begin{pmatrix} -i \\ 1 \end{pmatrix} = -i \begin{pmatrix} 1 \\ i \end{pmatrix}.$$

Beispiel 6.2.7 (Eigenwerte und Eigenvektoren der Differentiation)**.** Sei $C^\infty(\mathbb{R},\mathbb{R})$ der $\mathbb{R}$-Vektorraum aller beliebig oft differenzierbaren Funktionen auf $\mathbb{R}$, sei $D\colon C^\infty(\mathbb{R},\mathbb{R}) \to C^\infty(\mathbb{R},\mathbb{R})$ die Differentiationsabbildung (siehe Beispiel 4.1.31) und sei $\lambda \in \mathbb{R}$. Dann ist

$$\mathrm{Eig}_\lambda(D) = \{f \in C^\infty(\mathbb{R},\mathbb{R}) \mid f' = \lambda \cdot f\}.$$

In der Analysis wird gezeigt, dass der Eigenraum $\mathrm{Eig}_\lambda(D)$ 1-dimensional ist und von der Exponentialfunktion $x \mapsto \exp(\lambda x)$ erzeugt wird. Insbesondere sind alle reellen Zahlen Eigenwerte von D, und die Eigenvektoren zu einem λ sind die Funktionen $x \mapsto c\exp(\lambda x)$ mit $c \in \mathbb{R}^*$.

Beispiel 6.2.8. Sei $D^2\colon C^\infty(\mathbb{R},\mathbb{R}) \to C^\infty(\mathbb{R},\mathbb{R})$ der Endomorphismus $f \mapsto f''$ und sei $\lambda \in \mathbb{R}$. Dann gilt $\dim_{\mathbb{R}} \mathrm{Eig}_\lambda(D^2) = 2$ und zwar

$$\mathrm{Eig}_\lambda(D^2) = \begin{cases} \mathrm{Span}_{\mathbb{R}}\{x \mapsto \cosh(\sqrt{\lambda}\,x), x \mapsto \sinh(\sqrt{\lambda}\,x)\}, & \text{falls } \lambda > 0, \\ \mathrm{Span}_{\mathbb{R}}\{x \mapsto 1, x \mapsto x\}, & \text{falls } \lambda = 0, \\ \mathrm{Span}_{\mathbb{R}}\{x \mapsto \cos(\sqrt{-\lambda}\,x), x \mapsto \sin(\sqrt{-\lambda}\,x)\}, & \text{falls } \lambda < 0. \end{cases}$$

Lemma 6.2.9. *Sei $f\colon V \to V$ ein Endomorphismus eines K-Vektorraums V.*

(i) *Für alle $\lambda \in K$ ist der Eigenraum* $\mathrm{Eig}_\lambda(f) \subset V$ *f-invariant.*

(ii) *Ist $\lambda \neq \mu$, so gilt* $\mathrm{Eig}_\lambda(f) \cap \mathrm{Eig}_\mu(f) = \{0\}$.

Beweis. Die erste Aussage ist der Sonderfall der Proposition 6.1.12(i) mit $g = \lambda \cdot \mathrm{id}_V - f$. Sei $v \in \mathrm{Eig}_\lambda(f) \cap \mathrm{Eig}_\mu(f)$. Dann gilt $f(v) = \lambda \cdot v = \mu \cdot v$, und damit $(\lambda - \mu) \cdot v = 0$. Da $\lambda \neq \mu$ folgt aus Proposition 3.2.6(iv), dass $v = 0$. □

Proposition 6.2.10 (lineare Unabhängigkeit von Eigenvektoren)**.** *Sei $f\colon V \to V$ ein Endomorphismus eines K-Vektorraums V. Sei $\Lambda \subset K$ eine Teilmenge bestehend aus Eigenwerten von f, und zu jedem $\lambda \in E$ sei $v_\lambda \in V$ ein Eigenvektor zum Eigenwert λ. Dann ist die Familie $(v_\lambda)_{\lambda\in\Lambda}$ linear unabhängig.*

Beweis. Nach Definition der linearen Unabhängigkeit dürfen wir voraussetzen, dass Λ endlich ist. Dann verwenden wir Induktion über die Mächtigkeit von Λ. Wenn $\Lambda = \varnothing$ ist die Aussage trivial. Sei also $\lambda_0 \in \Lambda$ und sei $\sum_{\lambda\in\Lambda} \mu_\lambda \cdot v_\lambda = 0$ mit $\mu_\lambda \in K$. Dann gilt:

$$0 = (\lambda_0 \,\mathrm{id}_V - f)\left(\sum_{\lambda\in\Lambda} \mu_\lambda \cdot v_\lambda\right) = \sum_{\lambda\in\Lambda\setminus\{\lambda_0\}} \mu_\lambda \cdot (\lambda_0 \,\mathrm{id}_V - f)(v_\lambda).$$

Nach Lemma 6.2.9(i) liegt jeder Vektor $(\lambda_0 \,\mathrm{id}_V - f)(v_\lambda)$ in $\mathrm{Eig}_\lambda(f)$, und nach Lemma 6.2.9(ii) ist er nicht null, sonst wäre $v_\lambda \in \mathrm{Eig}_\lambda(f) \cap \mathrm{Eig}_{\lambda_0}(f) = \{0\}$, aber $v_\lambda \neq 0$ nach Definition von Eigenvektor. Also ist jedes $(\lambda_0 \,\mathrm{id}_V - f)(v_\lambda)$ wieder ein Eigenvektor zu λ. Aus der Induktionsvoraussetzung folgt, dass $\mu_\lambda = 0$ für alle $\lambda \in \Lambda \setminus \{\lambda_0\}$. Dann ist auch $\mu_{\lambda_0} \cdot v_{\lambda_0} = 0$ und damit $\mu_{\lambda_0} = 0$. □

Korollar 6.2.11. *Sei $f\colon V \to V$ ein Endomorphismus eines K-Vektorraums V und sei $\Lambda \subset K$ eine Teilmenge. Dann ist die kanonische Abbildung*

$$\bigoplus_{\lambda\in\Lambda} \mathrm{Eig}_\lambda(f) \to \sum_{\lambda\in\Lambda} \mathrm{Eig}_\lambda(f)$$

ein Isomorphismus.

Beweis. Dies folgt aus Propositionen 6.2.10 und 6.1.7. □

Definition 6.2.12 (geometrische Vielfachheit)**.** Sei $f\colon V \to V$ ein Endomorphismus eines K-Vektorraums V und sei $\lambda \in K$. Die Dimension von $\mathrm{Eig}_\lambda(f)$ heißt die *geometrische Vielfachheit* von λ bzgl. f und wird mit $\mu_f^{\mathrm{geom}}(\lambda)$ bezeichnet.

Beispiel 6.2.13. Sei $A = \mathrm{diag}(d_1, \ldots, d_n)$ eine $n \times n$-Diagonalmatrix über K. Aus Beispiel 6.2.4(iv) folgt, dass $\mu_A^{\mathrm{geom}}(\lambda) = |\{i \in \{1, \ldots, n\} \mid d_i = \lambda\}|$.

Proposition 6.2.14 (Charakterisierung von Eigenwerten)**.** *Sei V ein endlich-dimensionaler K-Vektorraum, $f \in \mathrm{End}_K(V)$ ein Endomorphismus und $\lambda \in K$. Dann sind folgende Aussagen äquivalent:*

(i) *λ ist ein Eigenwert von f.*

(ii) *Es gilt* $\mathrm{rg}(\lambda \cdot \mathrm{id}_V - f) < \dim_K V$.

(iii) *Es gilt* $\det(\lambda \cdot \mathrm{id}_V - f) = 0$.

Beweis. Nach Definition ist λ genau dann ein Eigenwert von f, wenn die Abbildung $\lambda \,\mathrm{id}_V - f$ nicht injektiv ist. Aber wenn V endlich-dimensional ist, sind die Injektivität, Surjektivität und Bijketivität von $\lambda \,\mathrm{id}_V - f$ äquivalent (Korollar 4.1.39). Aussage (ii) bedeutet, dass $\lambda \,\mathrm{id}_V - f$ nicht surjektiv ist (nach Proposition 3.3.35), und Aussage (iii) bedeutet, dass $\lambda \,\mathrm{id}_V - f$ nicht bijektiv ist (Proposition 5.3.50). Deswegen sind alle Aussagen äquivalent. □

Nach Proposition 6.2.14 kann man die Eigenwerte von f finden, indem man die Gleichung $\det(\lambda \cdot \mathrm{id}_V - f) = 0$ löst. Nach der Leibniz-Formel ist $\det(\lambda \cdot \mathrm{id}_V - f)$ eine Polynomfunktion von λ, und es gibt leider keine allgemeine Methode zur Lösung solcher Gleichungen. Wir kommen auf diesen Punkt im Abschnitt 6.3 zurück.

In der Funktionalanalysis spielt die folgende Verallgemeinerung von Eigenwerten eine wichtige Rolle:

Definition 6.2.15 (Spektrum, Spektralwert)**.** Sei $f\colon V \to V$ ein Endomorphismus eines K-Vektorraums V. Das *Spektrum* $\sigma(f)$ von f ist die Menge aller Skalare λ, so dass die Abbildung $\lambda \cdot \mathrm{id}_V - f$ nicht bijektiv ist. Elemente von $\sigma(f)$ heißen *Spektralwerte* von f.

Bemerkung 6.2.16 (Eigenwerte vs. Spektralwerte)**.** Eigenwerte von f sind auch Spektralwerte von f. Die Umkehrung gilt, wenn V endlich-dimensional ist (nach Korollar 4.1.39), aber nicht im Allgemeinen. Zum Beispiel hat der Endomorphismus

$$f\colon K^{\mathbb{N}} \to K^{\mathbb{N}}, \quad (x_0, x_1, \dots) \mapsto (0, x_0, x_1, \dots),$$

keine Eigenwerte, aber 0 ist ein Spektralwert von f, da f nicht surjektiv ist. Es ist eigentlich $\sigma(f) = \{0\}$.

Korollar 6.2.17 (Mächtigkeit des Spektrums mit geometrischer Vielfachheit gerechnet)**.** *Sei $f\colon V \to V$ ein Endomorphismus eines K-Vektorraums V der endlichen Dimension n. Dann gelten $|\sigma(f)| \leq n$ und $\sum_{\lambda \in \sigma(f)} \mu_f^{\mathrm{geom}}(\lambda) \leq n$.*

Beweis. Dies folgt aus dem Korollar 6.2.11, da $\sum_{\lambda \in \sigma(f)} \mu_f^{\mathrm{geom}}(\lambda) = \dim_K\big(\bigoplus_{\lambda \in \sigma(f)} \mathrm{Eig}_\lambda(f)\big)$ und $\mu_f^{\mathrm{geom}}(\lambda) \geq 1$ für alle $\lambda \in \sigma(f)$. □

Mit dem Gaußschen Eliminationsverfahren können wir die geometrischen Vielfachheiten bzgl. einer Matrix sowie Basen der zugehörigen Eigenräume leicht bestimmen:

Rezept 6.2.18 (Berechnung der geometrischen Vielfachheit)**.** Gegeben seien eine Matrix $A \in M_n(K)$ und ein Skalar $\lambda \in K$. Gesucht ist $\mu_A^{\mathrm{geom}}(\lambda)$. Man berechnet den Rang von $\lambda I_n - A$ mit dem Rezept 5.2.14. Nach der Dimensionsformel für lineare Abbildungen ist dann $\mu_A^{\mathrm{geom}}(\lambda) = n - \mathrm{rg}(\lambda I_n - A)$.

Rezept 6.2.19 (Berechnung des Eigenraums)**.** Gegeben seien eine Matrix $A \in M_n(K)$ und ein Skalar $\lambda \in K$. Gesucht ist eine Basis von $\mathrm{Eig}_\lambda(A)$. Dazu verwendet man das Rezept 5.2.15 mit der Matrix $\lambda I_n - A$.

6.2.1 Diagonalisierbarkeit

Definition 6.2.20 (diagonalisierbarer Endomorphismus)**.** Sei V ein K-Vektorraum. Ein Endomorphismus $f \in \mathrm{End}_K(V)$ heißt *diagonalisierbar*, wenn eine Basis von V bestehend aus Eigenvektoren von f existiert.

Eine quadratische Matrix A heißt *diagonalisierbar*, wenn L_A diagonalisierbar ist.

Beispiel 6.2.21.

(i) Sei V ein K-Vektorraum und sei $t\colon V \oplus V \to V \oplus V$, $(v, w) \mapsto (w, v)$ (siehe Beispiel 6.2.4). Ist $(v_i)_{i \in I}$ eine Basis von V und ist $\mathrm{char}(K) \neq 2$, so bilden die Eigenvektoren (v_i, v_i) und $(v_i, -v_i)$ von t eine Basis von $V \oplus V$, und damit ist t diagonalisierbar. Aber wenn $\mathrm{char}(K) = 2$ und $V \neq \{0\}$, dann erzeugen die Eigenvektoren von t den Untervektorraum $\{(v, v) \mid v \in V\}$ von $V \oplus V$, und damit ist t nicht diagonalisierbar.

(ii) Sei $n \in \mathbb{N}$ und sei D eine $n \times n$-Diagonalmatrix über K. Dann besteht der Standardbasis von K^n aus Eigenvektoren von D, und insbesondere ist D diagonalisierbar.

Beispiel 6.2.22. Der Endomorphismus D von $C^\infty(\mathbb{R},\mathbb{R})$ ist nicht diagonalisierbar, da nicht alle Funktionen aus $C^\infty(\mathbb{R},\mathbb{R})$ Linearkombinationen der Funktionen $x \mapsto \exp(\lambda x)$ sind (siehe Beispiel 6.2.7). Denn eine solche Linearkombination (außer der Nullfunktion) muss unbeschränkt sein, aber es existiert beschränkte beliebig oft differenzierbare Funktionen (z.B. konstante Funktionen).

Proposition 6.2.23. *Sei $f\colon V \to V$ ein Endomorphismus eines K-Vektorraums V. Die folgenden Aussagen sind äquivalent:*

(i) *f ist diagonalisierbar.*

(ii) *Jeder Vektor $v \in V$ ist eine Linearkombination von Eigenvektoren von f.*

(iii) *Die kanonische lineare Abbildung $\bigoplus_{\lambda\in K} \mathrm{Eig}_\lambda(f) \to V$ ist ein Isomorphismus.*

Beweis. Die Implikation (i) $\Rightarrow$ (ii) ist klar, und die Implikation (ii) $\Rightarrow$ (iii) folgt aus Korollar 6.2.11. Ist die Abbildung $\bigoplus_{\lambda\in K} \mathrm{Eig}_\lambda(f) \to V$ ein Isomorphismus, so erhält man eine Basis von V bestehend aus Eigenvektoren, indem man Basen aller Eigenräume $\mathrm{Eig}_\lambda(f)$ zusammensetzt. □

Satz 6.2.24 (Charakterisierung der Diagonalisierbarkeit)**.** *Sei V ein endlich-dimensionaler K-Vektorraum und sei $f \in \mathrm{End}_K(V)$. Dann sind die folgenden Aussagen äquivalent:*

(i) *f ist diagonalisierbar.*

(ii) *Es existiert eine Basis B von V, so dass $[f]_B^B$ eine Diagonalmatrix ist.*

(iii) *Es gilt*
$$\sum_{\lambda\in\sigma(f)} \mu_f^{\mathrm{geom}}(\lambda) = \dim_K V.$$

Beweis. Zu (i) $\Leftrightarrow$ (ii). Sei B eine Basis von V. Dann ist $[f]_B^B$ genau dann eine Diagonalmatrix, wenn B aus Eigenvektoren von f besteht.

Zu (i) $\Leftrightarrow$ (iii). Nach Korollar 6.2.11 ist $\sum_{\lambda\in\sigma(f)} \mu_f^{\mathrm{geom}}(\lambda)$ genau dann gleich $\dim_K V$, wenn die kanonische Abbildung $\bigoplus_{\lambda\in\sigma(f)} \mathrm{Eig}_\lambda(f) \to V$ ein Isomorphismus ist. Nach Proposition 6.2.23 ist das Letztere zur Diagonalisierbarkeit von f äquivalent. □

Korollar 6.2.25. *Sei V ein K-Vektorraum der endlichen Dimension n und sei $f \in \mathrm{End}_K(V)$. Hat f n paarweise verschiedene Eigenwerte, so ist f diagonalisierbar.*

Beweis. Nach Definition ist die geometrische Vielfachheit eines Eigenwerts mindestens 1. Wenn f n paarweise verschiedene Eigenwerte besitzt, dann ist $\sum_{\lambda\in\sigma(f)} \mu_f^{\mathrm{geom}}(\lambda) \geq n$. Andererseits ist $\sum_{\lambda\in\sigma(f)} \mu_f^{\mathrm{geom}}(\lambda) \leq n$ nach Korollar 6.2.17. Nach Satz 6.2.24 (iii) $\Rightarrow$ (i) ist also f diagonalisierbar. □

Korollar 6.2.26 (Diagonalisierbarkeit von Matrizen)**.** *Seien $n \in \mathbb{N}$ und $A \in M_n(K)$. Die folgenden Aussagen sind äquivalent:*

(i) *A ist diagonalisierbar.*

(ii) *A ist ähnlich zu einer Diagonalmatrix.*

Beweis. Dies folgt aus Satz 6.2.24 (i) $\Leftrightarrow$ (ii) und Proposition 6.1.19. □

Rezept 6.2.27 (Test auf Diagonalisierbarkeit)**.** Gegeben seien eine Matrix $A \in M_n(K)$ und ihre Eigenwerte $\lambda_1, \ldots, \lambda_k$. Zu bestimmen ist, ob A diagonalisierbar ist. Wenn $k = n$ ist, ist A diagonalisierbar nach Korollar 6.2.25. Sonst berechnet man die geometrischen Vielfachheiten $\mu_A^{\mathrm{geom}}(\lambda_i)$ mit Rezept 6.2.18. Nach Satz 6.2.24 ist die Matrix A genau dann diagonalisierbar, wenn $\sum_{i=1}^k \mu_A^{\mathrm{geom}}(\lambda_i) = n$.

Rezept 6.2.28 (Diagonalisierung einer Matrix)**.** Gegeben seien eine Matrix $A \in M_n(K)$ und ihre Eigenwerte $\lambda_1, \dots, \lambda_k$. Gesucht ist eine Matrix $S \in \mathrm{GL}_n(K)$, wenn sie existiert, so dass $S^{-1}AS$ eine Diagonalmatrix ist. Mit Rezept 6.2.19 findet man Basen der Eigenräume $\mathrm{Eig}_{\lambda_i}(A)$, und dadurch wird auch bestimmt, ob A diagonalisierbar ist (siehe Rezept 6.2.27). Wenn ja, erhält man eine Basis $B = (v_1, \dots, v_n)$ von K^n, indem man die gefundenen Basen der Eigenräume zusammensetzt. Sei S die Matrix mit Spalten $v_1, \dots, v_n$, d.h., die Basiswechselmatrix T_E^B von B nach der Standardbasis E. Nach der Basiswechselformel (Proposition 4.2.43) hat dann die Matrix S die gewünschte Eigenschaft. Man kann außerdem die inverse Matrix S^{-1} mit Rezept 5.2.18 berechnen.

Beispiel 6.2.29. Wir können jetzt das Beispiel 4.2.44 erklären. Sei

$$A = \begin{pmatrix} 3 & -1 \\ 2 & 0 \end{pmatrix} \in M_2(\mathbb{R}).$$

Um die Eigenwerte von A zu finden, müssen wir die Gleichung $\det(\lambda I_2 - A) = 0$ lösen:

$$\det(\lambda I_2 - A) = (\lambda - 3)\lambda + 2 = \lambda^2 - 3\lambda + 2 = (\lambda - 1)(\lambda - 2),$$

und daher sind 1 und 2 die Eigenwerte von A. Insbesondere ist A diagonalisierbar. Mit Rezept 6.2.19 berechnen wir die zugehörigen Eigenräume:

$$I_2 - A = \begin{pmatrix} -2 & 1 \\ -2 & 1 \end{pmatrix} \xrightarrow{A_{21}(-1)} \begin{pmatrix} -2 & 1 \\ 0 & 0 \end{pmatrix} \implies \mathrm{Eig}_1(A) = \mathbb{R} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix},$$
$$2I_2 - A = \begin{pmatrix} -1 & 1 \\ -2 & 2 \end{pmatrix} \xrightarrow{A_{21}(-2)} \begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix} \implies \mathrm{Eig}_2(A) = \mathbb{R} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Also ist $B = \left(\binom{1}{1}, \binom{1}{2}\right)$ eine Basis von $\mathbb{R}^2$ bestehend aus Eigenvektoren von A, und damit ist

$$S^{-1}AS = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{wobei} \quad S = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

Wir berechnen noch die inverse Matrix S^{-1} mit Rezept 5.2.18:

$$(S|I_2) = \left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{array}\right) \xrightarrow{A_{21}(-1)} \left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 0 & 1 & -1 & 1 \end{array}\right) \xrightarrow{A_{12}(-1)} \left(\begin{array}{cc|cc} 1 & 0 & 2 & -1 \\ 0 & 1 & -1 & 1 \end{array}\right) = (I_2|S^{-1}).$$

Beispiel 6.2.30. Sei

$$A = \begin{pmatrix} -5 & 2 & -4 \\ -2 & 0 & -2 \\ 4 & -2 & 3 \end{pmatrix} \in M_3(\mathbb{Q}).$$

Wir versuchen A mit Rezept 6.2.28 zu diagonalisieren. Es ist

$$\begin{aligned} \det(\lambda I_3 - A) &= \det \begin{pmatrix} \lambda + 5 & -2 & 4 \\ 2 & \lambda & 2 \\ -4 & 2 & \lambda - 3 \end{pmatrix} \\ &\overset{\text{1.S}}{=} (\lambda + 5)(\lambda(\lambda - 3) - 4) - 2(-2(\lambda - 3) - 8) - 4(-4 - 4\lambda) \\ &= \lambda^3 + 2\lambda^2 + \lambda = \lambda(\lambda + 1)^2. \end{aligned}$$

Die Eigenwerte von A sind also 0 und -1. Als Nächstes berechnen wir die Eigenräume:

- *Eigenraum zu* -1.

$$-I_3 - A = \begin{pmatrix} 4 & -2 & 4 \\ 2 & -1 & 2 \\ -4 & 2 & -4 \end{pmatrix} \xrightarrow[V_{12}]{\substack{A_{12}(-2) \\ A_{32}(2)}} \begin{pmatrix} 2 & -1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Der Rang von $-I_2 - A$ ist also gleich 1, so dass $\mu_A^{\text{geom}}(-1) = 3-1 = 2$. Daraus können wir bereits schließen, dass A diagonalisierbar ist. Die Vektoren

$$v_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \quad \text{und} \quad v_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

bilden eine Basis von $\text{Eig}_{-1}(A) = \mathcal{L}(-I_3 - A, 0)$.

- *Eigenraum zu* 0.

$$0I_3 - A = \begin{pmatrix} 5 & -2 & 4 \\ 2 & 0 & 2 \\ -4 & 2 & -3 \end{pmatrix} \xrightarrow[V_{12}]{M_2(\frac{1}{2})} \begin{pmatrix} 1 & 0 & 1 \\ 5 & -2 & 4 \\ -4 & 2 & -3 \end{pmatrix}$$
$$\xrightarrow[A_{31}(4)]{A_{21}(-5)} \begin{pmatrix} 1 & 0 & 1 \\ 0 & -2 & -1 \\ 0 & 2 & 1 \end{pmatrix} \xrightarrow{A_{32}(1)} \begin{pmatrix} 1 & 0 & 1 \\ 0 & -2 & -1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Damit wird $\text{Eig}_0(A) = \mathcal{L}(A, 0)$ von folgendem Vektor erzeugt:

$$v_3 = \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}.$$

Daraus erhalten wir $S^{-1}AS = \text{diag}(-1, -1, 0)$, wobei

$$S = \begin{pmatrix} v_1 & v_2 & v_3 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 2 \\ 2 & 0 & 1 \\ 0 & 1 & -2 \end{pmatrix}.$$

Schließlich berechnen wir die inverse Matrix S^{-1}:

$$(S|I_3) = \left(\begin{array}{ccc|ccc} 1 & -1 & 2 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & -2 & 0 & 0 & 1 \end{array}\right) \xrightarrow{A_{21}(-2)} \left(\begin{array}{ccc|ccc} 1 & -1 & 2 & 1 & 0 & 0 \\ 0 & 2 & -3 & -2 & 1 & 0 \\ 0 & 1 & -2 & 0 & 0 & 1 \end{array}\right)$$
$$\xrightarrow[V_{23}]{A_{23}(-2)} \left(\begin{array}{ccc|ccc} 1 & -1 & 2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 0 & 0 & 1 \\ 0 & 0 & 1 & -2 & 1 & -2 \end{array}\right) \xrightarrow[A_{13}(-2)]{A_{23}(2)} \left(\begin{array}{ccc|ccc} 1 & -1 & 0 & 5 & -2 & 4 \\ 0 & 1 & 0 & -4 & 2 & -3 \\ 0 & 0 & 1 & -2 & 1 & -2 \end{array}\right)$$
$$\xrightarrow{A_{12}(1)} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -4 & 2 & -3 \\ 0 & 0 & 1 & -2 & 1 & -2 \end{array}\right) = (I_3|S^{-1}).$$

Beispiel 6.2.31. Seien $\lambda, \alpha \in K$. Falls $\alpha \neq 0$ ist die Dreiecksmatrix

$$A = \begin{pmatrix} \lambda & \alpha \\ 0 & \lambda \end{pmatrix}$$

nicht diagonalisierbar. Denn A hat den einzigen Eigenwert λ, und seine geometrische Vielfachheit ist nur 1:

$$\lambda I_2 - A = \begin{pmatrix} 0 & -\alpha \\ 0 & 0 \end{pmatrix} \quad \Longrightarrow \quad \text{Eig}_\lambda(A) = K \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Bemerkung 6.2.32 (Potenzen einer diagonalisierbaren Matrix)**.** Sei $A \in M_n(K)$ eine quadratische Matrix. Ist A diagonalisierbar, so können wir die Potenzen A^k von A wie folgt berechnen. Sei $S \in \text{GL}_n(K)$, so dass $S^{-1}AS = \text{diag}(\lambda_1, \ldots, \lambda_n)$. Dann ist

$$A^k = (S\,\text{diag}(\lambda_1, \ldots, \lambda_n)S^{-1})^k = S\,\text{diag}(\lambda_1, \ldots, \lambda_n)^k S^{-1} = S\,\text{diag}(\lambda_1^k, \ldots, \lambda_n^k)S^{-1}.$$

Beispiel 6.2.33 (Fibonacci-Zahlen)**.** Sei $(F_n)_{n\in\mathbb{N}}$ die Folge der Fibonacci-Zahlen. Im Beispiel 4.2.19 haben wir die folgende Formel erreicht:

$$\begin{pmatrix} F_{n+1} \\ F_n \end{pmatrix} = A^n \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \text{wobei} \quad A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \in M_2(\mathbb{R}).$$

Es ist $\det(\lambda I_2 - A) = \lambda(\lambda - 1) - 1 = \lambda^2 - \lambda - 1$, und damit hat A die zwei Eigenwerte

$$\frac{1 \pm \sqrt{5}}{2}.$$

Insbesondere ist A diagonalisierbar. Durch die Methode aus Bemerkung 6.2.32 erhalten wir die explizite Formel

$$F_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right).$$

Proposition 6.2.34 (Ähnlichkeit von Diagonalmatrizen)**.** *Sei $n \in \mathbb{N}$ und seien $D = \operatorname{diag}(d_1, \ldots, d_n)$ und $D' = \operatorname{diag}(d'_1, \ldots, d'_n)$ zwei Diagonalmatrizen über K. Dann sind die folgenden Aussagen äquivalent:*

(i) *D und D' sind ähnlich.*

(ii) *Es gibt eine Permutation $\sigma \in S_n$, so dass $d'_i = d_{\sigma(i)}$ für alle $i \in \{1, \ldots, n\}$.*

Beweis. Für die Elementarmatrix V_{ij} gilt $V_{ij}^{-1} D V_{ij} = \operatorname{diag}(d_{\tau(1)}, \ldots, d_{\tau(n)})$, wobei τ die Transposition $(i\ j)$ ist. Da jede Permutation σ eine Komposition von Transpositionen ist (Lemma 5.3.4), folgt daraus die Implikation (ii) $\Rightarrow$ (i). Sind umgekehrt D und D' ähnlich, so haben D und D' die gleichen Eigenwerte mit den gleichen geometrischen Vielfachheiten (siehe Bemerkung 6.2.3). Aber die Diagonalkoeffizienten einer Diagonalmatrix sind genau ihre Eigenwerte, und ein Eigenwert kommt so oft vor wie seine geometrische Vielfachheit (siehe Beispiel 6.2.13). Deshalb gilt die Implikation (i) $\Rightarrow$ (ii). □

Bemerkung 6.2.35 (Klassifikation von diagonalisierbaren Endomorphismen bis auf Isomorphie)**.** Sei $\mathfrak{Diag}\mathfrak{End}_n$ die Menge aller Isomorphieklassen von Paaren (V, f), wobei V ein n-dimensionaler K-Vektorraum ist und f ein diagonalisierbarer Endomorphismus von V ist. Sei $\sim$ die folgende Äquivalenzrelation auf K^n: $x \sim y$ genau dann, wenn eine Permutation $\sigma \in S_n$ existiert, so dass $y_i = x_{\sigma(i)}$ für alle $i \in \{1, \ldots, n\}$. Nach Proposition 6.2.34 gibt es dann eine bijektive Abbildung

$$\begin{aligned} K^n/\!\sim \;&\to \mathfrak{Diag}\mathfrak{End}_n, \\ [(d_1, \ldots, d_n)] &\mapsto [(K^n, L_{\operatorname{diag}(d_1,\ldots,d_n)})]. \end{aligned}$$

Die Umkehrabbildung schickt einen diagonalisierbaren Endomorphimus $f \in \operatorname{End}_K(V)$ auf das n-Tupel seiner Eigenwerte mit geometrischer Vielfachheit gezählt.

Proposition 6.2.36 (Determinante und Spur diagonalisierbarer Endomorphismen)**.** *Sei V ein endlich-dimensionaler K-Vektorraum und sei $f \in \operatorname{End}_K(V)$ ein diagonalisierbarer Endomorphismus. Dann gilt:*

(i) $\det(f) = \prod_{\lambda \in \sigma(f)} \lambda^{\mu_f^{\text{geom}}(\lambda)}$.

(ii) $\operatorname{tr}(f) = \sum_{\lambda \in \sigma(f)} \mu_f^{\text{geom}}(\lambda) \cdot \lambda$.

Beweis. Sei B eine Basis von V, so dass $[f]_B^B$ eine Diagonalmatrix ist. Dann ist $\det(f)$ bzw. $\operatorname{tr}(f)$ das Produkt bzw. die Summe der Diagonalkoeffizienten von $[f]_B^B$, die genau die Eigenwerte von f sind, mit geometrischer Vielfachheit gezählt. □

6.3 Das charakteristische Polynom

6.3.1 Polynome

Zur Erinnerung ist $K^{(\mathbb{N})}$ der K-Vektorraum aller Folgen $(a_n)_{n\in\mathbb{N}}$ in K, die null außerhalb einer endlichen Teilmenge von $\mathbb{N}$ ist. Die Folgen $(\delta_{in})_{n\in\mathbb{N}}$ mit $i \in \mathbb{N}$ bilden eine Basis von $K^{(\mathbb{N})}$ (siehe Beispiel 3.3.13).

Sei T ein Symbol, das wir als *Variable* oder *Unbestimmte* benennen. Dann bezeichnen wir mit $K[T]$ der K-Vektorraum $K^{(\mathbb{N})}$, in dem wir die Folge $(\delta_{in})_{n\in\mathbb{N}}$ als T^i schreiben. Also ist $(T^i)_{i\in\mathbb{N}}$ eine Basis des K-Vektorraums $K[T]$, so dass jedes Element $p \in K[T]$ als Linearkombination

$$p = \sum_{i\in\mathbb{N}} a_i T^i$$

geschrieben werden kann, wobei $a_i \in K$ und nur endlich viele der a_i nicht null sind. Außerdem schreiben wir T anstelle von T^1 und identifizieren wir die von T^0 aufgespannte Gerade in $K[T]$ mit K durch die injektive lineare Abbildung

$$K \hookrightarrow K[T], \quad a \mapsto aT^0.$$

Definition 6.3.1 (Polynom, Monom, Glied, Absolutglied)**.** Mit der obigen Schreibweise bezeichnen wir Elemente von $K[T]$ als *Polynome über K* (oder *mit Koeffizienten in K*) in der Variablen T. Ist $p = \sum_{i\in\mathbb{N}} a_i T^i$ ein Polynom, so heißen die Skalare a_i die *Koeffizienten* von p. Der Nullvektor $0 \in K[T]$ heißt das *Nullpolynom*.

Ein Polynom mit höchstens einem Nicht-Null-Koeffizient heißt *Monom*. Die Monome a_iT^i heißen die *Glieder* des Polynoms $\sum_{i\in\mathbb{N}} a_iT^i$, und der Koeffizient a_0 heißt das *Absolutglied*.

Definition 6.3.2 (Multiplikation von Polynomen)**.** Seien $p = \sum_{i\in\mathbb{N}} a_iT^i$ und $q = \sum_{i\in\mathbb{N}} b_iT^i$ zwei Polynome über K. Man definiert das Produkt $p \cdot q$ durch

$$p \cdot q = \sum_{i\in\mathbb{N}} \sum_{j\in\mathbb{N}} a_i b_j T^{i+j} = \sum_{k\in\mathbb{N}} \left(\sum_{i+j=k} a_i b_j \right) T^k.$$

Man beachte dabei, dass diese Summen nur endlich viele Summanden enthalten, die nicht null sind, so dass $p \cdot q$ ein wohldefiniertes Polynom ist.

Beispiel 6.3.3.

(i) Sind $\lambda, \mu \in K$, so gilt $(T + \lambda) \cdot (T + \mu) = T^2 + (\lambda + \mu)T + \lambda\mu$. Insbesondere ist $(T + \lambda) \cdot (T - \lambda) = T^2 - \lambda^2$.

(ii) Für jedes $n \in \mathbb{N} \setminus \{0\}$ gilt $T^n - 1 = (T - 1) \cdot (T^{n-1} + \ldots + T + 1)$ (die rechte Seite ist eine „Teleskopsumme“).

Proposition 6.3.4. *Die Multiplikation auf $K[T]$ ist assoziativ und kommutativ, sie hat das neutrale Element $1 = T^0$ und sie ist distributiv über die Addition.*

Beweis. Seien $p = \sum_{i\in\mathbb{N}} a_iT^i$, $q = \sum_{i\in\mathbb{N}} b_iT^i$ und $r = \sum_{i\in\mathbb{N}} c_iT^i$ Polynome über K. Die Assoziativität folgt daraus, dass beide $p \cdot (q \cdot r)$ und $(p \cdot q) \cdot r$ gleich

$$\sum_{i\in\mathbb{N}} \sum_{j\in\mathbb{N}} \sum_{k\in\mathbb{N}} a_i b_j c_k T^{i+j+k}$$

sind. Die Kommutativität und die Neutralität von T^0 sind klar. Zur Distributivität berechnen wir:

$$\begin{aligned} p \cdot (q + r) &= \sum_{i\in\mathbb{N}} \sum_{j\in\mathbb{N}} a_i(b_j + c_j)T^{i+j} \\ &= \sum_{i\in\mathbb{N}} \sum_{j\in\mathbb{N}} a_i b_j T^{i+j} + \sum_{i\in\mathbb{N}} \sum_{j\in\mathbb{N}} a_i c_j T^{i+j} \\ &= p \cdot q + p \cdot r. \end{aligned}$$

□

Bemerkung 6.3.5. Insbesondere ist $(K[T], +, \cdot)$ ein kommutativer Ring (siehe Bemerkung 2.3.5), den wir als *Polynomring* über K bezeichnen. Außerdem ist die Multiplikation auf $K[T]$ mit der Skalarmultiplikation kompatibel: Sind $p, q \in K[T]$ und ist $\lambda \in K$, so gilt

$$(\lambda \cdot p) \cdot q = \lambda \cdot (p \cdot q) = p \cdot (\lambda \cdot q).$$

Ein kommutativer Ring mit einer kompatiblen Struktur von K-Vektorraum in diesem Sinne heißt eine *kommutative K-Algebra*.

Definition 6.3.6 (Grad, Leitkoeffizient, monisches Polynom)**.** Der *Grad* eines Polynoms $p = \sum_{i \in \mathbb{N}} a_i T^i$ über K ist

$$\deg(p) := \sup\{i \in \mathbb{N} \mid a_i \neq 0\} \in \{-\infty\} \cup \mathbb{N},$$

wobei $\sup \varnothing = -\infty$. Das Nullpolynom ist also das einzige Polynom vom Grad $-\infty$, und ein Polynom $p \in K[T]$ vom Grad $d \geq 0$ kann wie folgt geschrieben werden, mit $a_d \neq 0$:

$$p = a_d T^d + a_{d-1} T^{d-1} + \cdots + a_1 T + a_0.$$

Der Koeffizient $a_d \in K \setminus \{0\}$ heißt der *Leitkoeffizient* von p. Ein Polynom $p \in K[T]$ heißt *monisch*, wenn $\deg(p) \geq 0$ und der Leitkoeffizient von p gleich 1 ist.

Proposition 6.3.7 (Eigenschaften des Grades)**.** *Seien $p, q \in K[T]$ Polynome und $\lambda \in K$.*

(i) $\deg(p) = -\infty \iff p = 0$.

(ii) *Es gilt* $\deg(p \cdot q) = \deg(p) + \deg(q)$.

(iii) *Es gilt* $\deg(p + q) \leq \max\{\deg(p), \deg(q)\}$. *Die Gleichheit gilt, wenn* $\deg(p) \neq \deg(q)$.

(iv) *Es gilt* $\deg(\lambda \cdot p) \leq \deg(p)$. *Die Gleichheit gilt, wenn $\lambda \in K^*$.*

Beweis. Dies folgt unmittelbar aus den Definitionen. □

Definition 6.3.8 (Einsetzung, Polynomfunktion)**.** Sei $p = \sum_{i \in \mathbb{N}} a_i T^i \in K[T]$ und $\lambda \in K$. Die *Einsetzung* von λ in p ist $p(\lambda) := \sum_{i \in \mathbb{N}} a_i \lambda^i \in K$. Die Abbildung

$$\begin{aligned} K &\to K, \\ \lambda &\mapsto p(\lambda), \end{aligned}$$

heißt die dem Polynom p zugehörige *Polynomfunktion* auf K. Man bezeichnet mit $\mathrm{Poly}(K, K)$ die Menge aller Polynomfunktionen auf K, d.h., das Bild der Abbildung $K[T] \to \mathrm{Abb}(K, K)$, die p auf $\lambda \mapsto p(\lambda)$ abbildet. Man kann leicht nachprüfen, dass die letztere Abbildung linear ist, so dass $\mathrm{Poly}(K, K)$ ein Untervektorraum von $\mathrm{Abb}(K, K)$ ist.

Beispiel 6.3.9. Die Polynomfunktionen vom Grad $\leq n$ auf $\mathbb{R}$ bilden den Kern der $(n+1)$-ten Differentiationsabbildung $D^{n+1} \colon C^\infty(\mathbb{R}, \mathbb{R}) \to C^\infty(\mathbb{R}, \mathbb{R})$.

Bemerkung 6.3.10 (Polynome vs. Polynomfunktionen)**.** Man sollte Polynome über K nicht mit Polynomfunktionen auf K verwechseln: Ist K ein endlicher Körper, so gibt es nur endlich viele Polynomfunktionen $K \to K$, aber trotzdem unendlich viele Polynome über K. Insbesondere gibt es in diesem Fall viele verschiedene Polynome über K, die dieselbe Polynomfunktion induzieren. Zum Beispiel induzieren alle Polynome T^{2i+1} über $\mathbb{F}_3$ die Identität auf $\mathbb{F}_3$.

Definition 6.3.11 (Teilbarkeit)**.** Seien $f, g \in K[T]$. Man sagt, dass g *teilt* f oder dass f durch g *teilbar* ist, und man schreibt $g | f$, wenn ein Polynom $q \in K[T]$ existiert, so dass $f = qg$.

Satz 6.3.12 (Polynomdivision mit Rest)**.** *Seien $f, g \in K[T]$ zwei Polynome mit $g \neq 0$. Dann existieren eindeutige Polynome $q, r \in K[T]$ mit $f = qg + r$ und $\deg(r) < \deg(g)$.*

Beweis. Zur Eindeutigkeit. Seien $f = q_1 g + r_1$ und $f = q_2 g + r_2$ mit $\deg(r_1) < \deg(g)$ und $\deg(r_2) < \deg(g)$. Dann gilt:

$$0 = f - f = (q_1 - q_2)g + (r_1 - r_2)$$

und damit $(q_1 - q_2)g = r_2 - r_1$. Nach Proposition 6.3.7 gelten $\deg((q_1 - q_2)g) = \deg(q_1 - q_2) + \deg(g)$ und $\deg(r_2 - r_1) \leq \max\{\deg(r_1), \deg(r_2)\} < \deg(g)$. Folglich ist $\deg(q_1 - q_2) < 0$, d.h., $q_1 = q_2$. Dann ist auch $r_1 = f - q_1 g = f - q_2 g = r_2$.

Zur Existenz. Seien $n = \deg(f)$ und $m = \deg(g) \geq 0$. Wir beweisen die Existenz von q und r durch vollständige Induktion über n. Falls $n < m$ kann man $q = 0$ und $r = f$ nehmen. Falls $n \geq m$ schreibt man $f = aT^n + \bar{f}$ und $g = bT^m + \bar{g}$ mit $a, b \in K \setminus \{0\}$, $\deg(\bar{f}) < n$ und $\deg(\bar{g}) < m$. Man setzt $q_1 = b^{-1}aT^{n-m}$ und $f_1 = f - q_1 g$, so dass $f = q_1 g + f_1$. Dann ist

$$\deg(f_1) = \deg(\bar{f} - b^{-1}aT^{n-m}\bar{g}) \leq \max\{\deg(\bar{f}), \deg(\bar{g}) + n - m\} < n.$$

Nach der Induktionsvoraussetzung existieren $q', r \in K[T]$ mit $f_1 = q'g + r$ und $\deg(r) < \deg(g)$, und damit ist

$$f = (q_1 + q')g + r,$$

wie gewünscht. □

Beispiel 6.3.13. Der Beweis der Existenz von q und r im Satz 6.3.12 ist völlig konstruktiv und liefert einen Algorithmus, der ganz ähnlich wie die gewöhnliche schriftliche Division mit Rest von ganzen Zahlen ist. Als Beispiel sei $f = T^3 + T + 1$ und $g = T - 2$. Dann erhalten wir $q = T^2 + 2T + 5$ und $r = 11$:

$$\begin{array}{rrr}
\multicolumn{3}{l}{\big(T^3 \quad + T \quad + 1\big) : \big(T - 2\big) = T^2 + 2T + 5 \text{ mit Rest } 11.} \\
-T^3 & +2T^2 & \\
\hline
 & 2T^2 & +T \quad +1 \\
 & -2T^2 & +4T \\
\hline
 & & 5T \quad +1 \\
 & & -5T \quad +10 \\
\hline
 & & 11
\end{array}$$

Insbesondere ist $T^3 + T + 1$ genau dann durch $T - 2$ teilbar, wenn $\mathrm{char}(K) = 11$.

Definition 6.3.14 (Nullstelle)**.** Sei $p \in K[T]$ ein Polynom. Eine *Nullstelle* von p ist ein Skalar $a \in K$, so dass $p(a) = 0$.

Proposition 6.3.15. *Sei $p \in K[T]$ und $a \in K$. Die folgenden Aussagen sind äquivalent:*

(i) *a ist eine Nullstelle von p.*

(ii) *p ist durch $T - a$ teilbar.*

Beweis. Ist $p = (T - a)q$, so ist $p(a) = (a - a)q(a) = 0$. Sei umgekehrt a eine Nullstelle von p. Nach Satz 6.3.12 gibt es Polynome $q, r \in K[T]$ mit $p = (T - a)q + r$ und $\deg(r) < \deg(T - a) = 1$, d.h., $r \in K$. Dann gilt $0 = p(a) = (a - a)q(a) + r = r$, und damit ist $p = (T - a)q$. □

Beispiel 6.3.16. Sei $p \in \mathbb{N}$ eine Primzahl. Nach dem *kleinen Fermatschen Satz*, der in der *Algebra* Vorlesung bewiesen wird, gilt $n^p \equiv n \mod p$ für jede ganze Zahl n. Anders

gesagt ist jedes $a \in \mathbb{F}_p$ eine Nullstelle des Polynoms $T^p - T$. Durch p Anwendungen der Proposition 6.3.15 erhalten wir die Gleichung

$$T^p - T = \prod_{a \in \mathbb{F}_p} (T - a)$$

in $\mathbb{F}_p[T]$.

Korollar 6.3.17. *Sei $p \in K[T]$ ein Polynom vom Grad $d \geq 0$. Dann hat p höchstens d verschiedene Nullstellen.*

Beweis. Wir verwenden Induktion über d. Ein Polynom vom Grad 0 hat keine Nullstellen. Falls $d \geq 1$ und λ eine Nullstelle von p ist, dann existiert $q \in K[T]$ mit $p = (T - \lambda)q$ nach Proposition 6.3.15. Nach der Nullteilerfreiheit von K müssen alle anderen Nullstellen von p auch Nullstellen von q sein. Nach der Induktionsvoraussetzung hat q höchstens $d - 1$ Nullstellen, und damit hat p höchstens d Nullstellen. □

Korollar 6.3.18. *Ist K unendlich, so ist die lineare Abbildung $K[T] \to \mathrm{Poly}(K, K)$ aus Definition 6.3.8 ein Isomorphismus.*

Beweis. Sei p ein Polynom im Kern dieser Abbildung. Dann ist jedes $a \in K$ eine Nullstelle von p. Da K unendlich ist, folgt aus Korollar 6.3.17, dass $\deg(p) = -\infty$, d.h., dass $p = 0$. □

Definition 6.3.19 (Vielfachheit)**.** Sei $p \in K[T]$ und $a \in K$. Die *Vielfachheit* von a in p ist

$$v_a(p) := \sup\{n \in \mathbb{N} \mid (T - a)^n \text{ teilt } p\} \in \mathbb{N} \cup \{+\infty\}.$$

Bemerkung 6.3.20. Nach Proposition 6.3.15 ist a ist genau dann eine Nullstelle von p, wenn $v_a(p) \geq 1$. Für das Nullpolynom 0 gilt $v_a(0) = +\infty$.

Sind $p, q \in K[T]$, so gilt

$$v_a(pq) = v_a(p) + v_a(q).$$

Die Ungleichung $v_a(p)+v_a(q) \leq v_a(pq)$ ist klar. Angenommen, es wäre $v_a(p)+v_a(q) < v_a(pq)$. Dann gibt es Polynome $r, \bar{p}, \bar{q} \in K[T]$, so dass $pq = (T - a)^{v_a(p)+v_a(q)} r$, $p = (T - a)^{v_a(p)}\bar{p}$ und $q = (T - a)^{v_a(q)}\bar{q}$ mit $r(a) = 0$, $\bar{p}(a) \neq 0$ und $\bar{q}(a) \neq 0$. Aus der Eindeutigkeitsaussage im Satz 6.3.12 folgt $r = \bar{p}\bar{q}$. Insbesondere ist $r(a) = \bar{p}(a)\bar{q}(a) \neq 0$, was ein Widerspruch ist.

Definition 6.3.21 (Zerfall in Linearfaktoren)**.** Man sagt, dass ein Polynom $p \in K[T]$ *in seine Linearfaktoren zerfällt*, wenn p ein Produkt von Polynomen vom Grad ≤ 1 ist.

Bemerkung 6.3.22. Sei $p \in K[T] \setminus \{0\}$. Nach dem Satz 6.3.12 und der Bemerkung 6.3.20 kann man schreiben

$$p = q \cdot \prod_{a \in K} (T - a)^{v_a(p)}$$

mit einem eindeutig bestimmten Polynom q, das keine Nullstellen hat. Insbesondere ist

$$\sum_{a \in K} v_a(p) \leq n,$$

und die Gleichheit gilt genau dann, wenn p in seine Linearfaktoren zerfällt.

Korollar 6.3.23. *Ist K algebraisch abgeschlossen (Definition 2.4.5), so zerfällt jedes Polynom $p \in K[T]$ in seine Linearfaktoren.*

Beweis. Wir verwenden Induktion über $\deg(p)$. Wenn $\deg(p) \leq 1$ gibt es nichts zu zeigen. Falls $\deg(p) \geq 2$, dann existiert eine Nullstelle a von p nach Definition eines algebraisch abgeschlossenen Körpers. Nach Korollar 6.3.17 ist $p = (T - a)q$ mit $\deg(q) = \deg(p) - 1$. Nach der Induktionsvoraussetzung zerfällt q in seine Linearfaktoren, und somit auch p. □

Bemerkung 6.3.24. Sei $p = aT^2 + bT + c \in K[T]$ ein Polynom vom Grad 2 (d.h., $a \neq 0$). Falls $\mathrm{char}(K) \neq 2$ kann man die Nullstellen von p durch die gewöhnliche Mitternachtsformel ausdrücken:

$$\frac{-b \pm \sqrt{\Delta}}{2a}, \quad \Delta = b^2 - 4ac,$$

wobei $\pm\sqrt{\Delta}$ die Quadratwurzeln von Δ sind (wenn sie existieren; sonst hat p keine Nullstellen in K). Wenn $\mathrm{char}(K) \notin \{2, 3\}$ gibt es auch kompliziertere Wurzelausdrücke für die Nullstellen eines allgemeinen Polynoms vom Grad 3 oder 4. Bei Polynomen des fünften Grades und höher existiert aber keine allgemeine Wurzelausdruck für die Nullstellen (das wird in der Vorlesung *Algebra* genauer formuliert und auch bewiesen).

6.3.2 Das charakteristische Polynom

Zu jedem Endomorphismus $f\colon V \to V$ eines endlich-dimensionalen K-Vektorraums V gibt es ein kanonisches monisches Polynom $\chi_f \in K[T]$ vom Grad $\dim_K V$, das *charakteristische Polynom* von f, mit folgender Eigenschaft: Die Nullstellen von χ_f sind genau die Eigenwerte von f. Außerdem sind die Koeffizienten von χ_f eine gemeinsame Verallgemeinerung der Determinante $\det(f)$ und der Spur $\mathrm{tr}(f)$.

Wir definieren zunächst das charakteristische Polynom χ_A einer quadratischen Matrix A, und danach zeigen wir, dass $\chi_{[f]_B^B}$ unabhängig von der Wahl der Basis B ist. Dazu braucht man *Matrizen von Polynomen* zu betrachten: Eine $m \times n$-Matrix mit Koeffizienten in $K[T]$ ist einfach eine $\{1, \dots, m\} \times \{1, \dots, n\}$-indizierte Familie in $K[T]$.

Definition 6.3.25 (Determinante einer Matrix von Polynomen)**.** Sei $n \in \mathbb{N}$ und sei A eine $n \times n$-Matrix mit Koeffizienten in $K[T]$. Die *Determinante* von A ist das Polynom

$$\det(A) := \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \cdot \prod_{i=1}^{n} A_{\sigma(i)i} \in K[T].$$

Obwohl $(K[T], +, \cdot)$ kein Körper ist, viele (aber nicht alle) der Resultate über die Determinante von Matrizen, die wir im Abschnitt 5.3.3 bewiesen haben, bleiben gültig für Matrizen von Polynomen, mit denselben Beweisen. Zum Beispiel gelten der Laplacesche Entwicklungssatz und seine Korollare, d.h., man kann die Determinante einer Matrix von Polynomen durch Spalten- oder Zeilenentwicklung berechnen. Für $A, B \in M_n(K[T])$ gilt auch $\det(A \cdot B) = \det(A) \cdot \det(B)$ in $K[T]$, denn man diese Formel direkt aus der Leibniz-Formel nachrechnen kann.

Definition 6.3.26 (charakteristisches Polynom einer Matrix)**.** Sei A eine $n \times n$-Matrix über K. Das *charakteristische Polynom* von A ist das Polynom

$$\chi_A := \det(T \cdot I_n - A) \in K[T].$$

Bemerkung 6.3.27. Das charakteristische Polynom von A wird manchmal als die Determinante von $A - T \cdot I_n$ definiert. Nach Proposition 5.3.29(ii) gilt

$$\det(A - T \cdot I_n) = (-1)^n \chi_A.$$

Für viele Zwecke (z.B. um die Nullstellen zu bestimmen) macht das Vorzeichen $(-1)^n$ keinen Unterschied. Der Vorteil unserer Definition ist, dass χ_A immer ein monisches Polynom ist (siehe Proposition 6.3.32(i)).

Beispiel 6.3.28.

(i) Ist $A = (a_{ij})_{i,j}$ eine $n \times n$-Dreiecksmatrix (Definition 4.2.11), so ist

$$\chi_A = \prod_{i=1}^{n} (T - a_{ii})$$

nach Korollar 5.3.34.

(ii) Sei

$$A = \begin{pmatrix} 2 & 0 & -1 \\ 0 & 1 & -2 \\ 3 & -1 & 5 \end{pmatrix}.$$

Entwicklung nach der zweiten Spalte ergibt:

$$\begin{aligned} \chi_A &= \det \begin{pmatrix} T-2 & 0 & 1 \\ 0 & T-1 & 2 \\ -3 & 1 & T-5 \end{pmatrix} \\ &= (T-1)\det \begin{pmatrix} T-2 & 1 \\ -3 & T-5 \end{pmatrix} - \det \begin{pmatrix} T-2 & 1 \\ 0 & 2 \end{pmatrix} \\ &= (T-1)((T-2)(T-5)+3) - 2(T-2) = T^3 - 8T^2 + 18T - 9. \end{aligned}$$

Proposition 6.3.29. *Sei $n \in \mathbb{N}$ und seien $A, B \in M_n(K)$. Sind A und B ähnlich, so gilt $\chi_A = \chi_B$.*

Beweis. Sei $S \in \mathrm{GL}_n(K)$ mit $B = S^{-1} \cdot A \cdot S$. Dann ist

$$TI_n - B = TI_n - S^{-1}AS = S^{-1}(TI_n - A)S$$

in $M_n(K[T])$, und damit $\chi_B = \det(S)^{-1} \cdot \chi_A \cdot \det(S) = \chi_A$. □

Definition 6.3.30 (charakteristisches Polynom eines Endomorphismus)**.** Sei V ein K-Vektorraum der endlichen Dimension n und sei $f \in \mathrm{End}_K(f)$ ein Endomorphismus von V. Das *charakteristische Polynom* von f ist das Polynom

$$\chi_f := \chi_{[f]_B^B},$$

wobei B eine Basis von V ist. Nach Proposition 6.3.29 ist χ_f unabhängig von der Wahl der Basis B.

Bemerkung 6.3.31. Es ist auch möglich, eine matrixfreie Definition von χ_f zu geben. Dazu braucht man aber mehrere Erweiterungen der bisherigen betrachteten Begriffe. Man betrachtet nämlich die Menge $V[T]$ von Polynomen mit Koeffizienten in V, die ein „Vektorraum über $K[T]$“ ist. Man kann dann die Menge $\mathrm{Det}(V[T])$ von $K[T]$-multilinearen Determinantenfunktionen auf $V[T]$ definieren, und man kann zeigen, dass $\mathrm{End}_{K[T]}(\mathrm{Det}(V[T]))$ zu $K[T]$ kanonisch isomorph ist (vgl. Lemma 5.3.45). Die Determinante einer $K[T]$-linearen Abbildung $F\colon V[T] \to V[T]$ ist dann das eindeutige Polynom $\det(F)$, so dass $F^*(\Delta) = \det(F) \cdot \Delta$ für alle Determinantenfunktionen $\Delta \in \mathrm{Det}(V[T])$. Dann ist $\chi_f = \det(T \cdot \mathrm{id}_{V[T]} - f_{V[T]})$, wobei die Abbildung $f_{V[T]}\colon V[T] \to V[T]$ ein Polynom $\sum_{i\in\mathbb{N}} v_i T^i$ auf $\sum_{i\in\mathbb{N}} f(v_i)T^i$ abbildet.

Proposition 6.3.32 (Koeffizienten und Nullstellen des charakteristischen Polynoms)**.** *Sei V ein K-Vektorraum der endlichen Dimension n und sei $f \in \mathrm{End}_K(V)$.*

(i) *χ_f ist ein monisches Polynom vom Grad n.*

(ii) *Der Koeffizient von T^{n-1} in χ_f ist $-\mathrm{tr}(f)$.*

(iii) *Das Absolutglied von χ_f ist $(-1)^n \det(f)$.*

(iv) *Die Nullstellen von χ_f sind genau die Eigenwerte von f.*

Beweis. Es genügt die entsprechenden Aussagen über das charakteristische Polynom einer $n \times n$-Matrix A zu beweisen. Nach Definition ist

$$\chi_A = \sum_{\sigma \in S_n} \chi_A^\sigma, \quad \text{wobei} \quad \chi_A^\sigma := \mathrm{sgn}(\sigma) \cdot \prod_{i=1}^{n} (TI_n - A)_{\sigma(i)i}.$$

Wenn $\chi_A^\sigma \neq 0$ ist der Grad von χ_A^σ gleich der Anzahl der Fixpunkte von σ, d.h., der Indizes i mit $\sigma(i) = i$. Also ist $\deg(\chi_A^\sigma) \leq n$ für alle $\sigma \in S_n$, und es kann nur $\geq n-1$ sein, wenn $\sigma = \mathrm{id}$. In diesem Fall ist

$$\chi_A^{\mathrm{id}} = \prod_{i=1}^{n}(T - A_{ii}) = T^n - \left(\sum_{i=1}^{n} A_{ii}\right) T^{n-1} + \cdots,$$

und damit ist χ_A monisch und ist der Koeffizient von T^{n-1} gleich $-\operatorname{tr}(A)$. Das Absolutglied erhält man, indem man 0 für T in χ_A einsetzt:

$$\chi_A(0) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot \prod_{i=1}^{n}(-A)_{\sigma(i)i} = \det(-A) = (-1)^n \det(A).$$

Allgemeiner ist $\chi_A(\lambda) = \det(\lambda I_n - A)$. Nach Proposition 6.2.14 ist deshalb ein Skalar genau dann eine Nullstelle von χ_A, wenn er ein Eigenwert von A ist. □

Beispiel 6.3.33. Für einen Endomorphismus f eines 2-dimensionalen K-Vektorraums gilt $\chi_f = T^2 - \operatorname{tr}(f)T + \det(f)$.

Definition 6.3.34 (algebraische Vielfachheit)**.** Sei V ein endlich-dimensionaler K-Vektorraum, sei $f \in \mathrm{End}_K(V)$ und sei $\lambda \in K$. Die *algebraische Vielfachheit* von λ bzgl. f ist die Vielfachheit von λ in χ_f im Sinne der Definition 6.3.19. Sie wird mit $\mu_f^{\mathrm{alg}}(\lambda) \in \mathbb{N}$ bezeichnet.

Proposition 6.3.35 (Zerlegung des charakteristischen Polynoms)**.** *Sei V ein endlich-dimensionaler K-Vektorraum und $f \in \mathrm{End}_K(V)$ ein Endomorphismus.*

(i) *Sei $U \subset V$ ein f-invarianter Untervektorraum und seien $f_U \in \mathrm{End}_K(U)$ und $\bar{f} \in \mathrm{End}_K(V/U)$ die von f induzierten Endomorphismen. Dann gilt:*

$$\chi_f = \chi_{f_U} \cdot \chi_{\bar{f}}.$$

(ii) *Seien $U, W \subset V$ komplementäre f-invariante Untervektorräume und seien $f_U \in \mathrm{End}_K(U)$ und $f_W \in \mathrm{End}_K(W)$ die von f induzierten Endomorphismen. Dann gilt:*

$$\chi_f = \chi_{f_U} \cdot \chi_{f_W}.$$

Beweis. Man wählt eine Basis $C = (v_1, \ldots, v_m)$ von U und eine Familie $D = (v_{m+1}, \ldots, v_n)$, so dass $B = (v_1, \ldots, v_n)$ eine Basis von V ist. Dann ist $\bar{D} = (v_{m+1} + U, \ldots, v_n + U)$ eine Basis von V/U, und die Darstellungsmatrix von f bzgl. B hat die Form

$$[f]_B^B = \begin{pmatrix} [f_U]_C^C & * \\ 0 & [\bar{f}]_{\bar{D}}^{\bar{D}} \end{pmatrix}.$$

Aus Korollar 5.3.36 folgt, dass $\chi_f = \chi_{f_U} \cdot \chi_{\bar{f}}$. Falls W ein f-invariantes direktes Komplement von U ist, kann man für D eine Basis von W wählen. Dann ist

$$[f]_B^B = \begin{pmatrix} [f_U]_C^C & 0 \\ 0 & [f_W]_D^D \end{pmatrix},$$

und damit $\chi_f = \chi_{f_U} \cdot \chi_{f_W}$. □

Korollar 6.3.36 (geometrische vs. algebraische Vielfachheit)**.** *Sei $f\colon V \to V$ ein Endomorphismus eines endlich-dimensionalen K-Vektorraums V.*

(i) *Für alle $\lambda \in K$ gilt $\mu_f^{\mathrm{geom}}(\lambda) \leq \mu_f^{\mathrm{alg}}(\lambda)$.*

(ii) *Ist f diagonalisierbar, so gilt $\mu_f^{\mathrm{alg}}(\lambda) = \mu_f^{\mathrm{geom}}(\lambda)$ für alle $\lambda \in K$.*

Beweis. Zu (i). Der Eigenraum $\mathrm{Eig}_\lambda(f) \subset V$ ist f-invariant nach Lemma 6.2.9(i). Sei $g \in \mathrm{End}_K(\mathrm{Eig}_\lambda(f))$ die Einschränkung von f. Nach Proposition 6.3.35(i) ist χ_f durch χ_g teilbar. Aber $g = \lambda \cdot \mathrm{id}_{\mathrm{Eig}_\lambda(f)}$ und damit ist $\chi_g = (T-\lambda)^{\mu_f^{\mathrm{geom}}(\lambda)}$. Deswegen ist die Vielfachheit von λ in χ_f mindestens $\mu_f^{\mathrm{geom}}(\lambda)$.

Zu (ii). Sei f diagonalisierbar. Nach (i) gilt:

$$\dim_K V = \sum_{\lambda \in \sigma(f)} \mu_f^{\mathrm{geom}}(\lambda) \leq \sum_{\lambda \in \sigma(f)} \mu_f^{\mathrm{alg}}(\lambda) \leq \deg(\chi_f) = \dim_K V.$$

Dabei haben wir auch Satz 6.2.24, Bemerkung 6.3.22 und Proposition 6.3.32(i) verwendet. Daraus folgt

$$\sum_{\lambda \in \sigma(f)} \mu_f^{\mathrm{geom}}(\lambda) = \sum_{\lambda \in \sigma(f)} \mu_f^{\mathrm{alg}}(\lambda),$$

und nach (i) ist dies nur möglich, wenn $\mu_f^{\mathrm{alg}}(\lambda) = \mu_f^{\mathrm{geom}}(\lambda)$ für alle $\lambda \in K$. □

Definition 6.3.37 (Einsetzung von Endomorphismen)**.** Sei $f\colon V \to V$ ein Endomorphismus eines K-Vektorraums V. Ist $p = \sum_{i=0}^n a_i T^i$ ein Polynom über K, so definieren wir die *Einsetzung* von f in p durch

$$p(f) := \sum_{i=0}^n a_i f^i \in \mathrm{End}_K(V).$$

Man definiert auf ähnliche Weise die Einseztung $p(A) \in M_n(K)$ einer Matrix $A \in M_n(K)$ in p.

Dabei muss man sich daran erinnern, dass f^0 die Identität auf V ist. Ist zum Beispiel $p = \lambda - T$ mit $\lambda \in K$, so ist $p(f) = \lambda \cdot \mathrm{id}_V - f$.

Lemma 6.3.38. *Seien $p, q \in K[T]$ Polynome über K, $f\colon V \to V$ ein Endomorphismus eines K-Vektorraums V und A eine quadratische Matrix über K.*

(i) *Es ist $(p \cdot q)(f) = p(f) \circ q(f)$.*

(ii) *Es ist $(p \cdot q)(A) = p(A) \cdot q(A)$.*

(iii) *Ist V endlich-dimensional mit einer Basis B, so ist $[p(f)]_B^B = p([f]_B^B)$.*

Beweis. Ist p eine Linearkombination von Polynomen p_i und gilt (i) für jedes p_i, so gilt (i) für p. Deswegen können wir annehmen, dass $p = T^d$ mit einem $d \in \mathbb{N}$. Ist $q = \sum_{i=0}^n b_i T^i$, so gilt

$$(T^d \cdot q)(f) = \sum_{i=0}^n b_i f^{d+i} = f^d \circ \left(\sum_{i=0}^n b_i f^i \right) = f^d \circ q(f),$$

da f^d linear ist. Die dritte Aussage folgt aus Proposition 4.2.39(ii,iii), und die zweite Aussage folgt aus (i) und (iii) (und kann auch leicht direkt überprüft werden). □

Satz 6.3.39 (Satz von Cayley–Hamilton)**.** *Sei V ein endlich-dimensionaler K-Vektorraum und $f \in \mathrm{End}_K(V)$ ein Endomorphismus. Dann gilt $\chi_f(f) = 0$ in $\mathrm{End}_K(V)$.*

Beweis. Sei B eine Basis von V und sei $A = [f]_B^B$. Nach Lemma 6.3.38(iii) ist dann $[\chi_f(f)]_B^B = \chi_A(A)$. Es genügt also zu zeigen, dass $\chi_A(A) = 0$ für jede $n \times n$-Matrix A. Nach dem Korollar 5.3.39 existiert eine Matrix $B \in M_n(K[T])$ mit

$$(TI_n - A) \cdot B = \chi_A I_n,$$

nämlich $B = \operatorname{adj}(TI_n - A)$. Nach Definition der adjunkten Matrix sind die Koeffizienten von B Polynome vom Grad $\leq n-1$. Man kann also schreiben $B = \sum_{i=0}^{n-1} T^i B_i$ mit $B_i \in M_n(K)$. Man setzt auch $B_n = 0$ und $B_{-1} = 0$. Dann erhalten wir

$$\chi_A I_n = \sum_{i=0}^{n-1} (T^{i+1} B_i - T^i A B_i) = \sum_{i=0}^{n} T^i (B_{i-1} - A B_i).$$

Ist $\chi_A = \sum_{i=0}^{n} c_i T^i$, so folgt

$$c_i I_n = B_{i-1} - A B_i \quad \text{und daher} \quad c_i A^i = A^i B_{i-1} - A^{i+1} B_i$$

für alle $i \in \{0, \ldots, n\}$. Nimmt man die Summe über i, so erhält man

$$\chi_A(A) = \sum_{i=0}^{n} c_i A^i = \sum_{i=0}^{n} (A^i B_{i-1} - A^{i+1} B_i).$$

Die rechte Seite ist jetzt eine Teleskopsumme, die gleich null ist, wie gewünscht. □

Bemerkung 6.3.40. Es mag scheinen, dass der Satz von Cayley–Hamilton trivial sein soll, da

$$\chi_A(A) \overset{!?}{=} \det(AI_n - A) = \det(0) = 0.$$

Das Problem mit diesem „Beweis“ ist, dass $\chi_A(A)$ eine $n \times n$-Matrix ist während $\det(AI_n - A)$ ein Skalar ist. Es macht also gar keinen Sinn, sie gleichzusetzen. Nach dem Satz von Cayley–Hamilton gilt eigentlich

$$\chi_A(A) = \det(AI_n - A) \cdot I_n,$$

da beide Seiten die Nullmatrix sind, aber diese Gleichung folgt *nicht* aus allgemeinen Gründen: Ist $p = \sum_{i=0}^{d} T^i C_i \in M_n(K[T])$, so ist im Allgemeinen $\det(p)(A) \neq \det(p(A)) \cdot I_n$.

6.4 Hauptvektoren

Hauptvektoren sind eine Verallgemeinerung von Eigenvektoren, die relevant bei nicht-diagonalisierbaren Endomorphismen sind. Das einfachste Beispiel einer nicht-diagonalisierbaren Matrix ist die Scherungsmatrix

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

(siehe Beispiel 6.2.31). Die Matrix A hat den einzigen Eigenwert 1 (da $\chi_A = (T-1)^2$), aber der zugehörige Eigenraum ist die Gerade $K \cdot e_1$. Man kann aber bemerken, dass für den Vektor e_2 gilt $(I_2 - A)e_2 = -e_1$, und daher $(I_2 - A)^2 e_2 = 0$.

In diesem Beispiel gibt es einen Eigenwert λ von einem $f \in \operatorname{End}_K(V)$ und einen Vektor $v \in V$, der nicht im Kern von $\lambda \cdot \mathrm{id}_V - f$ liegt, aber der im Kern einer *Potenz* von $\lambda \cdot \mathrm{id}_V - f$ liegt. Es stellt sich heraus, dass ein solcher Vektor v kein Eigenvektor zu einem anderen $\mu \neq \lambda$ sein kann (siehe Lemma 6.4.5(ii)), und deswegen ist die Existenz von v eine Obstruktion zur Diagonalisierbarkeit von f.

Diese Beobachtung führt zum Begriff von *Hauptvektor* zu einem Eigenwert und dem zusammenhängenden Begriff der *Trigonalisierbarkeit*, die wir in diesem Abschnitt untersuchen. Als Konsequenz werden wir auch eine geometrische Interpretation der algebraischen Vielfachheit erhalten (Proposition 6.4.12).

Lemma 6.4.1. *Sei V ein K-Vektorraum und $(U_n)_{n\in\mathbb{N}}$ eine Folge von Untervektorräumen von V, so dass $U_n \subset U_{n+1}$ für alle $n \in \mathbb{N}$. Dann ist die Vereinigung $\bigcup_{n\in\mathbb{N}} U_n$ ein Untervektorraum von V.*

Beweis. Wir verwenden das Kriterium 3.2.8. Es ist klar, dass $\bigcup_{n\in\mathbb{N}} U_n$ nicht leer ist. Seien $u, u' \in \bigcup_{n\in\mathbb{N}} U_n$ und $\lambda \in K$. Nach Definition der Vereinigung existieren $n, n' \in \mathbb{N}$, so dass $u \in U_n$ und $u' \in U_{n'}$. Dann liegen $\lambda \cdot u$ in U_n und $u + u'$ in $U_{\max\{n,n'\}}$. □

Definition 6.4.2 (Hauptraum, Hauptvektor, Stufe)**.** Sei $f\colon V \to V$ ein Endomorphismus eines K-Vektorraums V und sei $\lambda \in K$.

- Der *Hauptraum* (oder *verallgemeinerte Eigenraum*) zu λ von f ist der Untervektorraum

$$\begin{aligned}\mathrm{Hau}_\lambda(f) &:= \{v \in V \mid \text{es existiert } n \in \mathbb{N} \text{ mit } (\lambda \cdot \mathrm{id}_V - f)^n(v) = 0\}\\ &= \bigcup_{n\in\mathbb{N}} \ker\big((\lambda \cdot \mathrm{id}_V - f)^n\big) \subset V\end{aligned}$$

  (siehe Lemma 6.4.1).

- Ein *Hauptvektor* (oder *verallgemeinerter Eigenvektor*) zu λ von f ist ein Element von $\mathrm{Hau}_\lambda(f) \setminus \{0\}$. Die *Stufe* eines Hauptvektors v zu λ ist das kleinste $n \in \mathbb{N} \setminus \{0\}$ mit $(\lambda \cdot \mathrm{id}_V - f)^n(v) = 0$.

Ist $n \in \mathbb{N}$ und ist A eine $n \times n$-Matrix über K, so bezeichnen wir als *Haupträume* und *Hauptvektoren* von A die Haupträume und Hauptvektoren von $L_A\colon K^n \to K^n$.

Bemerkung 6.4.3. Nach Definition gilt $\mathrm{Eig}_\lambda(f) \subset \mathrm{Hau}_\lambda(f)$, und die Eigenvektoren von f sind genau die Hauptvektoren von f der Stufe 1. Außerdem ist ein Skalar $\lambda \in K$ genau dann Eigenwert von f, wenn $\mathrm{Hau}_\lambda(f) \neq \{0\}$, denn: Ist v ein Hauptvektor zu λ der Stufe n, so ist $(\lambda \cdot \mathrm{id}_V - f)^{n-1}(v)$ ein Eigenvektor zu λ.

Beispiel 6.4.4.

(i) Sei $\lambda \in K$ und sei $A \in M_n(K)$ eine Dreiecksmatrix der Gestalt

$$A = \begin{pmatrix} \lambda & & * \\ & \ddots & \\ 0 & & \lambda \end{pmatrix} \quad \text{bzw.} \quad A = \begin{pmatrix} \lambda & & 0 \\ & \ddots & \\ * & & \lambda \end{pmatrix}.$$

Für die Matrix $\lambda I_n - A$ gilt dann $(\lambda I_n - A)^n = 0$. Deswegen ist $\mathrm{Hau}_\lambda(A) = K^n$, d.h., jeder Vektor $v \in K^n \setminus \{0\}$ ist ein Hauptvektor zu λ von A.

(ii) Sei $D\colon C^\infty(\mathbb{R},\mathbb{R}) \to C^\infty(\mathbb{R},\mathbb{R})$ die Differentiationsabbildung. Dann ist

$$\mathrm{Hau}_0(D) = \{f \in C^\infty(\mathbb{R},\mathbb{R}) \mid \text{es gibt } n \in \mathbb{N} \text{ mit } D^n(f) = 0\} = \mathrm{Poly}(\mathbb{R},\mathbb{R}).$$

Lemma 6.4.5. *Sei $f\colon V \to V$ ein Endomorphismus eines K-Vektorraums V.*

(i) *Für alle $\lambda \in K$ ist der Hauptraum $\mathrm{Hau}_\lambda(f) \subset V$ f-invariant.*

(ii) *Ist $\lambda \neq \mu$, so gilt $\mathrm{Hau}_\lambda(f) \cap \mathrm{Hau}_\mu(f) = \{0\}$.*

Beweis. Zu (i). Aus Proposition 6.1.12(i) folgt, dass alle Untervektorräume $\ker((\lambda\cdot\mathrm{id}_V - f)^n)$ f-invariant sind, und daher auch ihre Vereinigung.

Zu (ii). Sei $v \in \mathrm{Hau}_\lambda(f) \cap \mathrm{Hau}_\mu(f)$. Angenommen ist $v \neq 0$. Sei n die Stufe von v als Hauptvektor zu λ. Dann ist $w = (\lambda \cdot \mathrm{id}_V - f)^{n-1}(v)$ ein Eigenvektor zu λ. Nach (i) gilt auch $w \in \mathrm{Hau}_\mu(f)$. Sei m die Stufe von w als Hauptvektor zu μ. Dann ist $u = (\mu \cdot \mathrm{id}_V - f)^{m-1}(w)$ ein Eigenvektor zu μ, und nach Lemma 6.2.9(i) ist u auch ein Eigenvektor zu λ. Das steht aber im Widerspruch zum Lemma 6.2.9(ii). □

Proposition 6.4.6 (lineare Unabhängigkeit von Hauptvektoren)**.** *Sei $f\colon V \to V$ ein Endomorphismus eines K-Vektorraums V. Sei $\Lambda \subset K$ eine Teilmenge bestehend aus Eigenwerten von f, und zu jedem $\lambda \in E$ sei $v_\lambda \in V$ ein Hauptvektor zum Eigenwert λ. Dann ist die Familie $(v_\lambda)_{\lambda\in\Lambda}$ linear unabhängig.*

Beweis. Nach Definition der linearen Unabhängigkeit dürfen wir voraussetzen, dass Λ endlich ist. Dann verwenden wir Induktion über die Mächtigkeit von Λ. Wenn $\Lambda = \varnothing$ ist die Aussage trivial. Sei also $\lambda_0 \in \Lambda$ und sei $\sum_{\lambda\in\Lambda} \mu_\lambda \cdot v_\lambda = 0$ mit $\mu_\lambda \in K$. Es existiert $n \in \mathbb{N}$, so dass $(\lambda_0 \operatorname{id}_V - f)^n(v_{\lambda_0}) = 0$. Dann gilt:

$$0 = (\lambda_0 \operatorname{id}_V - f)^n \left(\sum_{\lambda\in\Lambda} \mu_\lambda \cdot v_\lambda \right) = \sum_{\lambda\in\Lambda\setminus\{\lambda_0\}} \mu_\lambda \cdot (\lambda_0 \operatorname{id}_V - f)^n(v_\lambda).$$

Nach Lemma 6.4.5(i) liegt jeder Vektor $(\lambda_0 \operatorname{id}_V - f)^n(v_\lambda)$ in $\operatorname{Hau}_\lambda(f)$, und nach Lemma 6.4.5(ii) ist er nicht null, sonst wäre $v_\lambda \in \operatorname{Hau}_\lambda(f) \cap \operatorname{Hau}_{\lambda_0}(f) = \{0\}$, aber $v_\lambda \neq 0$ nach Definition von Hauptvektor. Also ist jedes $(\lambda_0 \operatorname{id}_V - f)^n(v_\lambda)$ wieder ein Hauptvektor zu λ. Aus der Induktionsvoraussetzung folgt, dass $\mu_\lambda = 0$ für alle $\lambda \in \Lambda \setminus \{\lambda_0\}$. Dann ist auch $\mu_{\lambda_0} \cdot v_{\lambda_0} = 0$ und damit $\mu_{\lambda_0} = 0$. □

Korollar 6.4.7. *Sei $f\colon V \to V$ ein Endomorphismus eines K-Vektorraums V und sei $\Lambda \subset K$ eine Teilmenge. Dann ist die kanonische Abbildung*

$$\bigoplus_{\lambda\in\Lambda} \operatorname{Hau}_\lambda(f) \to \sum_{\lambda\in\Lambda} \operatorname{Hau}_\lambda(f)$$

ein Isomorphismus.

Beweis. Dies folgt aus Propositionen 6.4.6 und 6.1.7. □

6.4.1 Trigonalisierbarkeit

Der Begriff der Trigonalisierbarkeit erhalten wir, indem wir Eigenvektoren durch Hauptvektoren in der Definition der Diagonalisierbarkeit ersetzt:

Definition 6.4.8 (trigonalisierbarer Endomorphismus)**.** Sei V ein K-Vektorraum. Ein Endomorphismus $f \in \operatorname{End}_K(V)$ heißt *trigonalisierbar*, wenn eine Basis von V bestehend aus Hauptvektoren von f existiert.

Eine quadratische Matrix A heißt *trigonalisierbar*, wenn L_A trigonalisierbar ist.

Beispiel 6.4.9.

(i) Da jeder Eigenvektor ein Hauptvektor ist, ist jeder diagonalisierbare Endomorphismus auch trigonalisierbar.

(ii) Sei $V \neq \{0\}$ ein K-Vektorraum und sei $t\colon V \oplus V \to V \oplus V$, $(v, w) \mapsto (w, v)$. Nach Beispiel 6.2.21(i) ist der Endomorphismus t genau dann diagonalisierbar, wenn $\operatorname{char}(K) \neq 2$. Aber t ist immer trigonalisierbar: Ist $\operatorname{char}(K) = 2$, so sind alle Vektoren $(v, w) \neq (0, 0)$ Hauptvektoren zum Eigenwert 1 von t, denn es gilt $(\operatorname{id} - t)(v, w) = (v - w, w - v)$ und daher $(\operatorname{id} - t)^2(v, w) = (2v - 2w, 2w - 2v) = (0, 0)$.

(iii) Sei $\lambda \in K$ und sei A eine $n \times n$-Dreiecksmatrix über K mit allen Diagonalkoeffizienten gleich λ. Nach Beispiel 6.4.4(i) ist dann $\operatorname{Hau}_\lambda(A) = K^n$ und insbesondere ist A trigonalisierbar. Wir werden später beweisen, dass *alle* Dreiecksmatrizen trigonalisierbar sind (Satz 6.4.14).

Proposition 6.4.10. *Sei $f\colon V \to V$ ein Endomorphismus eines K-Vektorraums V. Die folgenden Aussagen sind äquivalent:*

(i) *f ist trigonalisierbar.*

(ii) *Jeder Vektor $v \in V$ ist eine Linearkombination von Hauptvektoren von f.*

(iii) *Die kanonische lineare Abbildung $\bigoplus_{\lambda\in K} \mathrm{Hau}_\lambda(f) \to V$ ist ein Isomorphismus.*

Beweis. Die Implikation (i) $\Rightarrow$ (ii) ist klar, und die Implikation (ii) $\Rightarrow$ (iii) folgt aus Korollar 6.4.7. Ist die Abbildung $\bigoplus_{\lambda\in K} \mathrm{Hau}_\lambda(f) \to V$ ein Isomorphismus, so erhält man eine Basis von V bestehend aus Hauptvektoren, indem man Basen aller Haupträume $\mathrm{Hau}_\lambda(f)$ zusammensetzt. □

Lemma 6.4.11. *Seien $p \in K[T]$, $a \in K$ und $n \in \mathbb{N}$. Ist $p(a) \neq 0$, so existieren Polynome $u, v \in K[T]$, so dass $u(T-a)^n + vp = 1$.*

Beweis. Wenn $n = 0$ leisten die Polynome $u = 1$ und $v = 0$ das Gewünschte. Nach Satz 6.3.12 existieren $q \in K[T]$ und $r \in K$ mit $p = (T-a)q + r$. Aus $p(a) \neq 0$ folgt $r \neq 0$. Setzt man $u = -r^{-1}q$ und $v = r^{-1}$, so erhält man $u(T-a) + vp = 1$. Ist $n \geq 1$, so hat die n-te Potenz von $u(T-a) + vp$ die Form $u^n(T-a)^n + wp$, und sie ist immer noch gleich 1. □

Proposition 6.4.12 (Haupträume und algebraische Vielfachheit)**.** *Sei V ein endlich-dimensionaler K-Vektorraum und $f \in \mathrm{End}_K(V)$ ein Endomorphismus. Für alle $\lambda \in K$ gilt:*

(i) $\dim_K \mathrm{Hau}_\lambda(f) = \mu_f^{\mathrm{alg}}(\lambda)$.

(ii) $\mathrm{Hau}_\lambda(f) = \ker\left((\lambda \cdot \mathrm{id}_V - f)^{\mu_f^{\mathrm{alg}}(\lambda)}\right)$.

Beweis. Sei $m = \mu_f^{\mathrm{alg}}(\lambda)$. Nach Definition der algebraischen Vielfachheit ist $\chi_f = (T-\lambda)^m p$ mit einem $p \in K[T]$, so dass $p(\lambda) \neq 0$. Nach Lemma 6.4.11 gibt es zu jedem $n \in \mathbb{N}$ Polynome $u_n, v_n \in K[T]$, so dass

$$u_n(T-\lambda)^n + v_n p = 1. \tag{6.4.13}$$

Seien $U_n = \ker((f - \lambda\,\mathrm{id})^n)$, $U = U_m$ und $W = \ker p(f)$. Nach Definition ist $\mathrm{Hau}_\lambda(f) = \bigcup_{n\in\mathbb{N}} U_n$. Wir behaupten, dass U und W komplementär in V sind:

- $\mathrm{Hau}_\lambda(f) \cap W = \{0\}$. Sei $x \in U_n \setminus \{0\}$. Nach (6.4.13) und Lemma 6.3.38(i) ist $(v_n(f) \circ p(f))(x) = x$, und insbesondere ist $p(f)(x) \neq 0$, d.h., $x \notin W$.
- $U + W = V$. Sei $x \in V$. Nach (6.4.13) und Lemma 6.3.38(i) ist

  $$x = (v_m(f) \circ p(f))(x) + (u_m(f) \circ (f - \lambda\,\mathrm{id}_V)^m)(x).$$

  Der erste Summand liegt in U, da $(T-\lambda)^m v_m p = v_m \chi_f$ und $\chi_f(f) = 0$ nach dem Satz von Cayley–Hamilton. Genauso liegt der zweite Summand in W.

Daraus folgt die zweite Aussage, denn: Jedes $x \in \mathrm{Hau}_\lambda(f)$ kann als $x = y + z$ mit $y \in U$ und $z \in W$ geschrieben werden. Da $U \subset \mathrm{Hau}_\lambda(f)$ liegt auch z in $\mathrm{Hau}_\lambda(f)$, so dass $z \in \mathrm{Hau}_\lambda(f) \cap W = \{0\}$.

Nach Proposition 6.1.12(i) sind U und W f-invariant. Mit der Proposition 6.3.35 erhalten wir die Zerlegung

$$\chi_f = \chi_{f_U} \cdot \chi_{f_W}.$$

Der Endomorphismus f_U ist trigonalisierbar mit dem einzigen Eigenwert λ. Nach der Implikation (i) $\Rightarrow$ (iv) im Satz 6.4.14 (deren Beweis unabhängig von der aktuellen Proposition ist), gilt $\chi_{f_U} = (T-\lambda)^{\dim_K U}$, und somit $m \geq \dim_K U$. Auf der anderen Seite ist $\chi_{f_W}(\lambda) \neq 0$, weil W keinen Eigenvektor zu λ enthält, und deshalb muss $\chi_{f_U}(f)$ durch $(T-\lambda)^m$ teilbar sein. Daraus folgt $m \leq \dim_K U$, und damit $m = \dim_K U$. □

Satz 6.4.14 (Charakterisierung der Trigonalisierbarkeit)**.** *Sei V ein endlich-dimensionaler K-Vektorraum und sei $f \in \mathrm{End}_K(V)$. Dann sind die folgenden Aussagen äquivalent:*

(i) *f ist trigonalisierbar.*

(ii) *Es existiert eine Basis B von V, so dass $[f]_B^B$ eine obere Dreiecksmatrix ist.*

(iii) *Es existiert eine Basis B von V, so dass $[f]_B^B$ eine untere Dreiecksmatrix ist.*

(iv) *Das charakteristische Polynom χ_f zerfällt in seine Linearfaktoren.*

(v) *Es gilt*

$$\sum_{\lambda\in\sigma(f)} \mu_f^{\mathrm{alg}}(\lambda) = \dim_K V.$$

Beweis. Zu (i) $\Rightarrow$ (ii). Wir beweisen die Aussage durch Induktion über $n = \dim_K V$. Sie ist trivial, wenn $n = 0$; sonst existiert nach Bemerkung 6.4.3 ein Eigenvektor $v \in V$ zu einem Eigenwert λ. Da $f(Kv) \subset Kv$ induziert f nach der universellen Eigenschaft des Quotientenvektorraums einen Endomorphismus $\bar{f}$ von V/Kv. Im Quotientenvektorraum V/Kv ist wieder jeder Vektor eine Linearkombination von Hauptvektoren, d.h., $\bar{f}$ ist wieder trigonalisierbar (Proposition 6.4.10). Da $\dim_K(V/Kv) = n-1$ gibt es nach Induktionsvoraussetzung eine Basis $C = (\bar{v}_2, \dots, \bar{v}_n)$ von V/Kv, so dass die Matrix $[\bar{f}]_C^C$ eine obere Dreiecksmatrix ist. Seien $v_2, \dots, v_n \in V$ Urbilder der Vektoren $\bar{v}_2, \dots, \bar{v}_n$. Dann ist $B = (v, v_2, \dots, v_n)$ eine Basis von V, so dass

$$[f]_B^B = \begin{pmatrix} \lambda & * \\ 0 & [\bar{f}]_C^C \end{pmatrix}.$$

Insbesondere ist $[f]_B^B$ eine obere Dreiecksmatrix.

Zu (ii) $\Rightarrow$ (iii). Sei $B = (b_1, \dots, b_n)$ und sei $B' = (b_n, \dots, b_1)$. Ist $[f]_B^B$ eine obere Dreiecksmatrix, so ist $[f]_{B'}^{B'}$ eine untere Dreiecksmatrix.

Zu (iii) $\Rightarrow$ (iv). Dies folgt aus dem Korollar 5.3.34.

Zu (iv) $\Rightarrow$ (v). Dies folgt aus der Definition von μ_f^{alg}.

Zu (v) $\Rightarrow$ (i). Nach Korollar 6.4.7 und Proposition 6.4.12(i) gilt

$$\dim_K \left(\sum_{\lambda\in\sigma(f)} \mathrm{Hau}_\lambda(f) \right) = \dim_K \left(\bigoplus_{\lambda\in\sigma(f)} \mathrm{Hau}_\lambda(f) \right) = \sum_{\lambda\in\sigma(f)} \mu_f^{\mathrm{alg}}(\lambda) = \dim_K V,$$

und damit ist $V = \sum_{\lambda\in\sigma(f)} \mathrm{Hau}_\lambda(f)$. Aus Proposition 6.4.10 folgt nun, dass f trigonalisierbar ist. □

Korollar 6.4.15. *Sei K ein algebraisch abgeschlossener Körper und V ein endlich-dimensionaler K-Vektorraum. Dann ist jeder Endomorphismus $f \in \mathrm{End}_K(V)$ trigonalisierbar.*

Beweis. Nach Korollar 6.3.23 zerfällt χ_f in seine Linearfaktoren. Nach Satz 6.4.14 (v) $\Rightarrow$ (i) ist f trigonalisierbar. □

Korollar 6.4.16 (Trigonalisierbarkeit von Matrizen)**.** *Seien $n \in \mathbb{N}$ und $A \in M_n(K)$. Dann sind die folgenden Aussagen äquivalent:*

(i) *A ist trigonalisierbar.*

(ii) *A ist ähnlich zu einer oberen Dreiecksmatrix.*

(iii) *A ist ähnlich zu einer unteren Dreiecksmatrix.*

Beweis. Dies folgt aus Satz 6.4.14 (i) $\Leftrightarrow$ (ii) $\Leftrightarrow$ (iii) und Proposition 6.1.19. □

Bemerkung 6.4.17. Um eine Klassifikation von trigonalisierbaren Endomorphismen bis auf Isomorphie zu erhalten, braucht man noch Dreiecksmatrizen bis auf Ähnlichkeit zu klassifizieren. Das werden wir in der Vorlesung *Lineare Algebra II* weiter untersuchen.

Rezept 6.4.18 (Test auf Trigonalisierbarkeit)**.** Gegeben seien eine Matrix $A \in M_n(K)$ und ihre Eigenwerte $\lambda_1, \ldots, \lambda_k$. Zu bestimmen ist, ob A trigonalisierbar ist. Wenn $k = n$ ist, ist A sogar diagonalisierbar nach Korollar 6.2.25. Sonst berechnet man die algebraischen Vielfachheiten $\mu_A^{\text{alg}}(\lambda_i)$, indem man χ_A durch $T - \lambda_i$ so oft wie möglich dividiert. Nach Satz 6.4.14 ist die Matrix A genau dann trigonalisierbar, wenn $\sum_{i=1}^{k} \mu_A^{\text{alg}}(\lambda_i) = n$.

Proposition 6.4.19 (Determinante und Spur trigonalisierbarer Endomorphismen)**.** *Sei V ein endlich-dimensionaler K-Vektorraum und $f \in \operatorname{End}_K(V)$ ein trigonalisierbarer Endomorphismus. Dann gilt:*

(i) $\det(f) = \prod_{\lambda \in \sigma(f)} \lambda^{\mu_f^{\text{alg}}(\lambda)}$.

(ii) $\operatorname{tr}(f) = \sum_{\lambda \in \sigma(f)} \mu_f^{\text{alg}}(\lambda) \cdot \lambda$.

Beweis. Sei B eine Basis von V, so dass $[f]_B^B$ eine Dreiecksmatrix ist. Dann ist $\det(f)$ bzw. $\operatorname{tr}(f)$ das Produkt bzw. die Summe der Diagonalkoeffizienten von $[f]_B^B$, die genau die Eigenwerte von f sind, mit algebraischer Vielfachheit gezählt. □

Index